

Erhöhung der Löslichkeit und Stabilität verschiedener Proteine durch rationales Design und gelenkte Evolution

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der naturwissenschaftlichen Fakultät III – Biologie und vorklinische Medizin –
der Universität Regensburg



vorgelegt von

Tobias Seitz

aus Schrobenhausen

November 2009

Promotionsgesuch eingereicht am: 11. November 2009

Die Arbeit wurde angeleitet von: Prof. Dr. Reinhard Sterner

Prüfungsausschuss:

Vorsitzender: Prof. Dr. Reinhard Wirth

1. Prüfer: Prof. Dr. Reinhard Sterner

2. Prüfer: PD Dr. Winfried Hausner

3. Prüfer: Prof. Dr. Herbert Tschochner

Die vorliegende Arbeit wurde in der Zeit von April 2006 bis November 2009 am Lehrstuhl Biochemie II des Institutes für Biophysik und physikalische Biochemie der Naturwissenschaftlichen Fakultät III der Universität zu Regensburg unter Leitung von Herrn Prof. Dr. Reinhard Sterner in Kooperation mit F. Hoffmann – LaRoche, *Pharma Research Discovery*, Basel (Schweiz) durchgeführt. Von März 2008 bis April 2008 erfolgten Arbeiten bei F. Hoffmann – LaRoche in Basel.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis	VI
Tabellenverzeichnis	X
Formelverzeichnis	XI
Abkürzungsverzeichnis	XII
Zusammenfassung	1
1 Einleitung	4
1.1 Faltung, Stabilität und Löslichkeit von Proteinen <i>in vitro</i>	4
1.1.1 Proteinfaltung unter Gleichgewichtsbedingungen	4
1.1.2 Strukturelle Grundlagen der Proteinstabilität.....	5
1.1.3 Löslichkeit von Proteinen	7
1.2 Proteinfaltung und Aggregation <i>in vivo</i>	8
1.3 Erhöhung der Proteinstabilität	9
1.3.1 Extrinsische Optimierung der heterologen Expression und der Kristallisation von Proteinen..	9
1.3.2 Intrinsische Optimierung der Stabilität und Löslichkeit von Proteinen durch Mutagenese.....	10
1.4 Das grün fluoreszierende Protein (GFP) und seine Anwendungen	13
1.5 Durchflusszytometrie	16
1.6 Die humane Proprotein Convertase Subtilisin/Kexin Typ 9 (hPCSK9)	16
1.7 Der humane Glucocorticoid-Rezeptor (hGR)	19
1.8 Zielsetzung der Arbeit	23
2 Material	25
2.1 Geräte	25
2.2 Verbrauchsmaterialien	27
2.3 Chemikalien	28
2.4 Kits	29
2.5 Enzyme	29
2.6 Antikörper	29
2.7 Bakterienstämme	30
2.8 Vektoren	31
2.8.1 pET-Vektoren.....	31
2.8.1.1 pET21a(+)	32
2.8.1.2 pER Vektoren.....	32

2.8.1.2.1 pER13a und pER13b.....	33
2.8.1.2.2 pER14.....	33
2.8.1.2.3 pER13cat.....	34
2.8.1.2.4 pERN7.....	34
2.8.2 Konstitutive Expressionsvektoren.....	35
2.8.2.1 pTNA	35
2.8.2.2 pTNA13a.....	35
2.8.2.3 pHCE-IIB	35
2.8.2.4. pHCE13a.....	36
2.8.3 pCFN1	36
2.9 Oligodesoxyribonukleotide.....	37
2.9.1 Vektorspezifische Amplifikations- und Sequenzierungsprimer.....	37
2.9.2 Primer zur Herstellung neuer Fusionsvektoren.....	38
2.9.3 Primer zur Amplifikation von <i>hkas</i> und <i>tmhisF</i>	38
2.9.4 Primer zur Amplifikation und Mutagenisierung von <i>hpcsk9</i> ³¹⁻⁴⁵⁴	38
2.9.5 Primer zur Amplifikation und Mutagenisierung von <i>hgr-lbd</i>	39
2.10 DNA- und Protein-Längenstandards	41
2.11 Puffer und Lösungen.....	42
2.11.1 Arbeiten mit <i>E. coli</i>	42
2.11.2 Arbeiten mit DNA.....	43
2.11.3 Agarosegelelektrophorese	43
2.11.4 Arbeiten mit Proteinen	44
2.11.4.1 Allgemeine Arbeiten	44
2.11.4.2 Arbeiten mit <i>hPCSK9</i> ³¹⁻⁴⁵⁴	44
2.11.4.3 Arbeiten mit <i>hGR-LBD</i>	45
2.11.5 SDS-Polyacrylamidgelelektrophorese (12,5 % SDS-Gele)	46
2.11.6 Lösungen für <i>Western-Blotting</i>	47
2.12 Nährmedien.....	48
2.13 Software.....	49
3 Methoden.....	50
3.1 Vorbereitung von Geräten und Lösungen	50
3.2 Mikrobiologische Methoden.....	50
3.2.1 Anzucht und Lagerung von <i>E. coli</i> -Stämmen	50
3.2.2 Herstellung chemisch kompetenter <i>E. coli</i> Zellen (Inoue <i>et al.</i> , 1990).....	50
3.2.3 Chemische Transformation (Inoue <i>et al.</i> , 1990)	51
3.2.4 Herstellung elektrokompenter <i>E. coli</i> Zellen (Dower <i>et al.</i> , 1988).....	51
3.2.5 Vorbereitung von DNA für die Elektroporation	52
3.2.6 Elektroporation (Dower <i>et al.</i> , 1988)	52
3.2.7 Bestimmung der Transformationseffizienz.....	52
3.2.8 <i>In vivo</i> Analyse der Proteinlöslichkeit mit Hilfe von Reporter-Proteinen	53

3.2.8.1 <i>In vivo</i> Analyse der Proteinlöslichkeit mit Hilfe von eGFP als Reporter	53
3.2.8.2 <i>In vivo</i> Analyse der Proteinlöslichkeit mit Hilfe der Chloramphenicol-Acetyltransferase (CAT) als Reporter.....	54
3.2.9 Analyse und Anreicherung fluoreszierender <i>E. coli</i> Zellen durch Durchflusszytometrie.....	55
3.2.10 Fluoreszenzmikroskopie von <i>E. coli</i> Zellen.....	58
3.3 Molekularbiologische Methoden.....	59
3.3.1 Amplifikation von DNA-Fragmenten (Mullis & Faloona, 1987; Saiki <i>et al.</i> , 1988).....	59
3.3.2 Kolonie-PCR.....	60
3.3.3 PCR-Methoden zur gerichteten und zufälligen Mutagenese.....	60
3.3.3.1 Megaprimer-PCR (Sarkar & Sommer, 1990)	60
3.3.3.2 <i>Overlap Extension</i> – PCR (OE-PCR)	61
3.3.3.3 Zufallsmutagenese durch <i>error prone</i> PCR (<i>epPCR</i>)	62
3.3.4 Agarosegelelektrophorese.....	63
3.3.5 Isolierung von DNA-Fragmenten aus Agarosegelen	64
3.3.6 Enzymatische Manipulation von dsDNA.....	64
3.3.6.1 Spaltung von dsDNA mit Restriktionsendonukleasen	64
3.3.6.2 Ligation von DNA-Fragmenten	65
3.3.7 Isolierung und Reinigung von Plasmid-DNA aus <i>E. coli</i>	65
3.3.7.1 Isolierung von Plasmid-DNA in analytischen Mengen (Mini-Präparation)	65
3.3.7.2 Isolierung von Plasmid-DNA in präparativen Mengen (Midi/Maxi-Präparation).....	65
3.3.8 Bestimmung der DNA-Konzentration	66
3.3.9 Herstellung von Genbanken.....	66
3.3.10 DNA-Sequenzierung.....	68
3.4 Proteinchemische Methoden	69
3.4.1 Genexpression und Proteinanreicherung.....	69
3.4.1.1 Expression und Anreicherung von Proteinen im analytischen Maßstab.....	69
3.4.1.2 Expression und Anreicherung von Proteinen im präparativen Maßstab.....	70
3.4.2 Proteinreinigung aus der löslichen Zellfraktion.....	71
3.4.2.1 Metallchelate-Affinitätschromatographie	71
3.4.2.2 Enzymatische Entfernung des His ₆ -Glutathion-S-Transferase Fusionsproteins	73
3.4.2.3 Gekoppelte Glutathion- und Metallchelate-Affinitätschromatographie	73
3.4.2.4 Präparative Gelfiltration.....	74
3.4.2.5 Dialyse von Proteinlösungen.....	75
3.4.2.6 Einkonzentrieren von Proteinlösungen	75
3.4.2.7 Lagerung von gereinigten Proteinen	75
3.5 Analytische Methoden.....	76
3.5.1 Konzentrationsbestimmung von Proteinen	76
3.5.1.1 Bestimmung der Proteinkonzentration nach Bradford (1976)	76
3.5.1.2 Bestimmung der Proteinkonzentration durch Absorptionsspektroskopie	76
3.5.2 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE).....	77
3.5.3 Analyse der Proteinlöslichkeit <i>in vivo</i> durch Densitometrie	78
3.5.4 <i>Western-Blotting</i>	79

3.5.5 Analytische Gelfiltration.....	80
3.5.6 Analytische <i>reversed-phase</i> HPLC.....	80
3.5.7 Circular-Dichroismus-Spektroskopie.....	81
3.5.8 Fluoreszenz-Spektroskopie.....	83
3.5.9. Refraktometrie.....	83
3.5.10 Bestimmung der Proteinstabilität durch schrittweise Denaturierung.....	84
3.5.11 Analyse der Proteinlöslichkeit <i>in vitro</i> durch fraktionierte Präzipitation.....	88
4 Ergebnisse und Diskussion.....	89
4.1 Rationales Design der Oberfläche von hPCSK9³¹⁻⁴⁵⁴.....	89
4.1.1 Analyse der Oberfläche von hPCSK9 ³¹⁻⁴⁵⁴	89
4.1.1.1 Homologie-Modellierung der hPCSK9 ³¹⁻⁴⁵⁴ - Struktur.....	89
4.1.1.2 Lösungsmittel-Zugänglichkeit der Aminosäuren von hPCSK9 ³¹⁻⁴⁵⁴	91
4.1.2 Klonierung von <i>hpcsk9</i> ³¹⁻⁴⁵⁴ (S386A) ^{opt}	93
4.1.3 Heterologe Herstellung und Reinigung der hPCSK9 ³¹⁻⁴⁵⁴ Varianten.....	94
4.1.4 Biophysikalische Charakterisierung der hPCSK9 ³¹⁻⁴⁵⁴ Varianten.....	95
4.1.4.1 Analyse der Löslichkeit <i>in vitro</i> durch fraktionierte Präzipitation.....	95
4.1.4.2 Analyse der konformationellen Stabilität durch Auffaltung in Harnstoff.....	96
4.2 Benutzung von eGFP als Löslichkeits- und Stabilitäts-Reporter <i>in vivo</i>.....	98
4.2.1 Auswahl eines geeigneten Expressionssystems und Zielproteins.....	98
4.2.1.1 Klonierung verschiedener <i>egfp</i> -Fusionsvektoren.....	99
4.2.1.2 Analyse verschiedener Expressionssysteme.....	101
4.2.1.2.1 Konstitutive Expression von <i>egfp</i> -fusionierten Genen.....	101
4.2.1.2.2 Induzierbare Expression von <i>egfp</i> -fusionierten Genen.....	104
4.2.1.2.2.1 eGFP-Fusionsproteine mit langem Linker (13 AS).....	104
4.2.1.2.2.2 eGFP-Fusionsproteine mit kurzem Linker (3 AS).....	106
4.2.2 Exemplarische Anreicherung fluoreszierender Zellen durch Durchflusszytometrie.....	106
4.2.3 Nähere Charakterisierung von hGR-LBD-eGFP.....	108
4.3 Erhöhung der Löslichkeit und der Stabilität von hGR-LBD durch gelenkte Evolution mit Hilfe von eGFP als Reporter.....	110
4.3.1 Herstellung von <i>hgr-lbd</i> Genbanken in pER13b.....	111
4.3.2 <i>Screening</i> der Genbanken.....	113
4.3.2.1 <i>Screening</i> bei 37 °C in Abwesenheit eines Liganden.....	113
4.3.2.2 <i>Screening</i> bei 30 °C in Anwesenheit von Dexamethason.....	117
4.3.3 Charakterisierung der angereicherten hGR-LBD Varianten.....	124
4.3.3.1 Analyse der Löslichkeit der angereicherten hGR-LBD Varianten <i>in vivo</i> mit Hilfe der Chloramphenicol-Acetyltransferase als Reporter.....	124
4.3.3.2 Expression der angereicherten <i>hgr-lbd</i> Varianten in Abwesenheit von Reporterprotein....	126
4.3.4 Einzel-Analyse von Mutationen aus hGR-LBD A5-1 - A5-4 und hGR-LBD A8.....	128
4.3.5 Einfluss von M560K, F602S, A605V, V702A, E705G und M752T auf hGR-LBD und hGR-LBD-eGFP.....	131
4.3.5.1 Klonierung von <i>hgr-lbd</i> Einzel- und Kombinationsvarianten in pER13a und pERN7.....	132

4.3.5.2 Analyse der Fluoreszenz von <i>E. coli</i> durch die eGFP-fusionierten hGR-LBD Einzel- und Kombinationsvarianten.....	133
4.3.5.3 Quantitative Analyse der löslichen Herstellung der hGR-LBD Einzel- und Kombinationsvarianten in <i>E. coli</i>	136
4.3.5.4 Reinigung der hGR-LBD Einzel- und Kombinationsvarianten	138
4.3.5.5 Konformationelle Stabilität der hGR-LBD Einzel- und Kombinationsvarianten	141
4.3.5.6 Weitere biophysikalische Charakterisierung von hGR-LBD.....	145
4.3.5.6.1 Analyse der Löslichkeit <i>in vitro</i>	145
4.3.5.6.2 Analyse des Oligomerisierungsgrades durch analytische Gelfiltration	146
4.3.5.6.3 Spektroskopische Analyse	148
4.3.5.7 Reinigung von Ligand-freiem hGR-LBD	150
4.3.5.8 Strukturelle Analyse der vorteilhaften Mutationen in hGR-LBD	152
4.3.5.8.1 A605V in hGR-LBD	152
4.3.5.8.2 Röntgenkristallstrukturanalyse von mGR-LBD(F608S), mGR-LBD(F608S+A611V) und mGR-LBD(A611V+V708A+E711G).....	156
4.3.5.8.3 V702A (V708A) und E705G (E711G) in hGR-LBD (mGR-LBD).....	160
4.3.5.8.4 M752T.....	161
5 Abschließende Diskussion.....	164
5.1 Rationales Design einer hydrophilen Proteinoberfläche	164
5.2 Erhöhung der Löslichkeit und Stabilität von Proteinen durch gelenkte Evolution mit Hilfe von eGFP als Faltungsreporter und Durchflusszytometrie als Hochdurchsatz <i>Screening</i>-Verfahren	166
5.3 Löslichere und stabilere Mutanten der Ligand-Bindungsdomäne des Glucocorticoid-Rezeptors.....	173
6 Literaturverzeichnis.....	177
7 Anhang	194
7.1 Protokoll zur Zell-Sortierung mittels Durchflusszytometrie	194
7.2 Bestimmung der Proteinkonzentration durch Fluoreszenz-Spektroskopie	195
7.3 Multiples Sequenzalignment der mittels FACS isolierten hGR-LBD Varianten.....	196
7.4 Molekulargewichte, spezifische Extinktionskoeffizienten bei 280 nm und Ausbeuten der aufgereinigten hGR-LBD Einzel- und Kombinationsvarianten	199
7.5 Kalibrierung der Superdex-75 (16/60) Gelfiltrationssäule.....	200
7.6 Multiples Sequenzalignment hGR-LBD - homologer Proteine.....	201
7.7 Superpositionierung von mGR-LBD mit hRXRα – hPPARγ Nuklearhormonrezeptor-Komplex	202
7.8 Teilpublikationen zur Stabilisierung und Strukturaufklärung des künstlich hergestellten ($\beta\alpha$)$_8$-Barrel Proteins HisF-C*C	203
8 Danksagung.....	231
9 Eidesstattliche Erklärung	233

Abbildungsverzeichnis

Abbildung 1: Faltung eines Proteins mit einem auftretenden Intermediat.....	4
Abbildung 2: Elektronenmikroskopische Aufnahme von <i>inclusion bodies</i> in <i>Escherichia coli</i> Zellen.	9
Abbildung 3: Die Biolumineszenz der Kristallqualle <i>Aequorea victoria</i> durch Aequorin und GFP...	13
Abbildung 4: Chromophor des GFP bestehend aus Ser65, Tyr66 und Gly67.	14
Abbildung 5: Zellulärer Signalweg von LDL, LDL-Rezeptoren und PCSK9 nach Lambert <i>et al.</i> (2008).	17
Abbildung 6: Aufbau von hPCSK9.	18
Abbildung 7: Mechanismen der Transkriptionsregulation durch den Glucocorticoid-Rezeptor nach Lu & Cidlowski (2006) und Heitzer <i>et al.</i> (2007).	20
Abbildung 8: Aufbau der häufigsten Isoform des humanen Glucocorticoid-Rezeptors (hGR α -A). ...	21
Abbildung 9: Schematische Darstellung des pET-Expressionssystems.	31
Abbildung 10: pET-21a(+) mit multipler Klonierungsstelle (MCS).	32
Abbildung 11: pER-Vektoren.	33
Abbildung 12: pTNA-Vektor.....	35
Abbildung 13: pHCE-IIB Vektor mit multipler Klonierungsstelle.....	35
Abbildung 14: Multiple Klonierungsstelle von pCFN1.....	36
Abbildung 15: DNA- und Protein-Längenstandards.....	41
Abbildung 16: Schematische Darstellung der <i>in vivo</i> Analyse von Proteinlöslichkeit mit Hilfe von eGFP als Reporter.	53
Abbildung 17: Schematische Darstellung der <i>in vivo</i> Analyse von Proteinlöslichkeit mit Hilfe von CAT als Reporter.	55
Abbildung 18: Schematischer Aufbau des MoFlo [®] Zytometers und exemplarisches Histogramm einer FACS-Analyse.....	56
Abbildung 19: Prinzip der Megaprimer-PCR.	61
Abbildung 20: Prinzip der <i>Overlap Extension</i> - PCR.	62
Abbildung 21: Schematische Darstellung des Vorgehens bei einer Probeexpression.	70
Abbildung 22: Anordnung in der Tankblot-Apparatur.	79
Abbildung 23: Refraktometrische Bestimmung der Ammoniumsulfat-Konzentration.	84
Abbildung 24: Vergleich der CD- (a) und Fluoreszenz-Spektren (b) von nativem und denaturiertem Protein.	85
Abbildung 25: Schematische Darstellung der Auswertung eines thermischen (a) bzw. chemischen (b) Auffaltungsexperiments.	87
Abbildung 26: Homologie-Modell I von hPCSK9 ³¹⁻⁴⁵⁴	91
Abbildung 27: Lösungsmittel-zugängliche Threonine, Leucine, Valine, Tyrosine, Phenylalanine und Tryptophane des hPCSK9 ³¹⁻⁴⁵⁴ -Homologie-Modells I.	92
Abbildung 28: Heterologe Expression in <i>E. coli</i> BL21(DE3) und Reinigung der hPCSK9 ³¹⁻⁴⁵⁴ Varianten.....	94
Abbildung 29: Löslichkeit von hPCSK9 ³¹⁻⁴⁵⁴ (S386A) und hPCSK9 ³¹⁻⁴⁵⁴ (S386A) ^{opt} <i>in vitro</i>	96
Abbildung 30: Auffaltung von hPCSK9 ³¹⁻⁴⁵⁴ (S386A) und hPCSK9 ³¹⁻⁴⁵⁴ (S386A) ^{opt} mit Harnstoff. ...	97

Abbildung 31: Analyse der schwach konstitutiven Expression von <i>egfp</i> -fusionierten Genen durch pTNA13a.	102
Abbildung 32: Analyse der stark konstitutiven Expression von <i>egfp</i> -fusionierten Genen durch pHCE13a.....	103
Abbildung 33: Analyse der induzierbaren Expression von <i>egfp</i> -fusionierten Genen durch pER13a.....	105
Abbildung 34: Analytisches FACS-Histogramm von tmHisF-eGFP(GSG).....	106
Abbildung 35: Exemplarische Anreicherung stark fluoreszierender Zellen durch Durchflusszytometrie.	107
Abbildung 36: Zweite Runde der exemplarischen Anreicherung.....	108
Abbildung 37: Analytische FACS-Histogramme von Zellen mit eGFP-fusioniertem hGR-LBD (schwarz) und hGR-LBD(F602S) (grün).	109
Abbildung 38: Strategie zur Erhöhung der Löslichkeit und Stabilität von hGR-LBD durch gelenkte Evolution.....	111
Abbildung 39: Analytische FACS-Histogramme von Zellen mit hGR-LBDwt (schwarz) und hGR-LBD Bank 1+2 (rot), nach Expression bei 37 °C ohne Zugabe eines Liganden.	113
Abbildung 40: Sortierung bei 37 °C ohne Zugabe eines Liganden. Analytisches FACS-Histogramm der Anreicherungsrunde 3 der Versuchsreihe I.....	115
Abbildung 41: Sortierung bei 37 °C ohne Zugabe eines Liganden. Analytisches FACS-Histogramm der Anreicherungsrounden 1, 5 und 6 der Versuchsreihe II.....	116
Abbildung 42: Quelle falsch positiver Klone beim <i>Screening</i> bei 37 °C ohne Zugabe eines Liganden. Sequenzausschnitt des <i>hgr-lbd</i> -Gens von Base 192 bis 226.	117
Abbildung 43: Analytische FACS-Histogramme von Zellen mit hGR-LBDwt (schwarz) und hGR-LBD Bank 1+2 (rot), nach Expression bei 30 °C in Anwesenheit von Dexamethason.	118
Abbildung 44: Sortierung bei 30 °C in Anwesenheit von Dexamethason. Analytisches FACS-Histogramm der Anreicherungsrounden 1, 2, 6 und 7.....	120
Abbildung 45: Quelle falsch positiver Klone beim <i>Screening</i> bei 30 °C in Anwesenheit von Dexamethason. Ergebnis analytischer Kolonie-PCR von je 10 Klonen aus Anreicherung 5 bzw. 8 (1 %iges Agarosegel).....	121
Abbildung 46: Analytische FACS-Histogramme von Zellen mit hGR-LBDwt (schwarz) und hGR-LBD A5-1 - A5-4 bzw. A8 (grün), bei 30 °C in Anwesenheit von Dexamethason.	123
Abbildung 47: Wachstum von transformierten <i>E. coli</i> T7 Express Zellen auf LB-Platten in Anwesenheit unterschiedlicher Konzentrationen an Chloramphenicol.....	125
Abbildung 48: Analytische Herstellung His ₆ -GST-fusionierter hGR-LBD Varianten in <i>E. coli</i> BL21(DE3), dokumentiert durch SDS-PAGE (12,5 %).	127
Abbildung 49: Lokalisation der 14 ausgewählten Einzel-Mutationen aus den sortierten hGR-LBD Varianten A5-1 - A5-4 und A8 (rot) in der Kristallstruktur von hGR-LBD(F602S) (pdb 1m2z) (grau).....	129
Abbildung 50: Analytische Herstellung His ₆ -GST-fusionierter hGR-LBD Einzelvarianten in <i>E. coli</i> BL21(DE3), dokumentiert durch SDS-PAGE (12,5 %).	130
Abbildung 51: FACS-Analyse und Fluoreszenz-Mikroskopie von Zellen mit hGR-LBDwt (schwarz; mittlere Fluoreszenz (mf): 83,8) im Vergleich zu Zellen mit hGR-LBD Einzel- und Kombinationsvarianten, nach Expression bei 30 °C in Anwesenheit von Dexamethason.....	134

Abbildung 52: Mittlere Fluoreszenz von <i>E. coli</i> in der Durchflusszytometrie durch (a) gezielt hergestellte Einzel- und Kombinationsvarianten von hGR-LBD-eGFP und (b) durch FACS angereicherte Populationen (links) bzw. sortierte Varianten (rechts).	135
Abbildung 53: Quantitative Analyse der löslichen Herstellung His ₆ -GST-fusionierter hGR-LBD Einzel- und Kombinationsvarianten in <i>E. coli</i> BL21(DE3), dokumentiert durch SDS-PAGE (12,5 %).	137
Abbildung 54: Korrelation zwischen der hGR-LBD-eGFP vermittelten Fluoreszenz und der Menge an löslich hergestelltem His ₆ -GST-hGR-LBD in <i>E. coli</i>	138
Abbildung 55: Reinigung des rekombinant hergestellten hGR-LBD(M560K+F602S+A605V+M752T) Proteins aus dem löslichen Zellextrakt von <i>E. coli</i> BL21(DE3).	139
Abbildung 56: <i>Reversed-phase</i> HPLC Analyse der einzelnen Reinigungsschritte des rekombinant hergestellten hGR-LBD(A605V+M752T) Proteins.	140
Abbildung 57: Thermische Auffaltungen der hGR-LBD Einzel- und Kombinationsvarianten.	142
Abbildung 58: Korrelation zwischen der hGR-LBD-eGFP vermittelten Fluoreszenz in <i>E. coli</i> und der konformationellen Stabilität von gereinigtem hGR-LBD (Agonist-Konformation).	144
Abbildung 59: Löslichkeit von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) <i>in vitro</i>	146
Abbildung 60: Analytische Gelfiltration von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+ E705G+M752T).	147
Abbildung 61: Fern-UV CD-Spektren von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+ E705G+M752T).	148
Abbildung 62: Fluoreszenzemissionsspektren von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+ V702A+E705G+M752T) nach Anregung bei 280 nm (a) bzw. 295 nm (b).	149
Abbildung 63: Charakterisierung von Ligand-freiem hGR-LBD(F602S+A605V+V702A+E705G+M752T).	151
Abbildung 64: Superpositionierung der Ligand-Bindungsdomänen des humanen Glucocorticoid-Rezeptors (hGR-LBD, pdb 1m2z) und des Progesteron-Rezeptors (huPR-LBD, pdb 1a28) mit gebundenem Liganden.	153
Abbildung 65: Analyse von hGR-LBD(F602S) mit unterschiedlichen Mutationen von A605.	155
Abbildung 66: Röntgenkristallstruktur von mGR-LBD(A611V+V708A+E711G) bzw. mGR-LBD(F608S) mit gebundenem Dexamethason und TIF2 Coaktivator-Peptid bei einer Auflösung von 1,5 bzw. 1,9 Å.	157
Abbildung 67: Etablierung eines Wasserstoffbrückennetzwerkes in mGR-LBD durch F608S.	159
Abbildung 68: Umgebung von V702(V708) und E705(E711) in hGR-LBD (mGR-LBD).	160
Abbildung 69: Bindung von Coaktivator-Peptid TIF2 an α -Helix 12 von hGR-LBD (Agonist-Konformation).	162
Abbildung 70: Schematische Darstellung des Vorgehens zur Verbesserung der Löslichkeit und Stabilität eines Proteins mit Hilfe von eGFP als Faltungsreporter und Durchflusszytometrie als Hochdurchsatz <i>Screening</i> -Methode.	172
Abbildung A1: Schematische Darstellung des Vorgehens bei einer FACS-Sortierung	194
Abbildung A2: Eichkurven zur Bestimmung der Proteinkonzentration durch Fluoreszenz-Spektroskopie.	195

Abbildung A3: Kalibrierung der Superdex S-75 Gelfiltrationssäule	200
Abbildung A4: Multiples Sequenzalignment hGR-LBD – homologer Ligand-Bindungs Domänen verschiedener Nuklearhormon-Rezeptoren, erstellt durch MCoffee.	201
Abbildung A5: Superpositionierung von mGR-LBD mit hRXR α – hPPAR γ Nuklearhormonrezeptor-Komplex (pdb 3dzy; Chandra <i>et al.</i> , 2008), erstellt durch PyMol (http://www.pymol.org/).....	202

Tabellenverzeichnis

Tabelle 1: Verwendete <i>error prone</i> PCR (<i>epPCR</i>) Reaktionsbedingungen.	63
Tabelle 2: Protokolle für die Reinigung verschiedener Proteine durch Metallchelate-Affinitäts-chromatographie.	72
Tabelle 3: Zusammensetzung 12,5 %iger SDS-PAGE Gele.	77
Tabelle 4: Referenz-Strukturen für Homologie-Modellierungen von hPCSK9 ³¹⁻⁴⁵⁴	90
Tabelle 5: Homologie-Modelle von hPCSK9 ³¹⁻⁴⁵⁴	90
Tabelle 6: Lösungsmittel-zugängliche Threonine, Leucine, Valine, Tyrosine, Phenylalanine und Tryptophane von hPCSK9 ³¹⁻⁴⁵⁴	92
Tabelle 7: Biophysikalische Parameter der hPCSK9 ³¹⁻⁴⁵⁴ Varianten.	95
Tabelle 8: Klonierte Konstrukte zur Analyse der Expressionsbedingungen von <i>egfp</i> -Fusionsgenen.	101
Tabelle 9: Anzahl und Verteilung der Nukleotid-Austausche bei zufällig ausgewählten Klonen der <i>hgr-lbd</i> Genbank 1 und 2.	112
Tabelle 10: FACS-Anreicherungen von <i>hgr-lbd</i> Genbank 1+2 exprimierenden <i>E. coli</i> BL21(DE3) Zellen bei 37 °C ohne Zugabe eines Liganden.	114
Tabelle 11: FACS-Anreicherungen von <i>hgr-lbd</i> Genbank 1+2 exprimierenden <i>E. coli</i> BL21(DE3) Zellen bei 30 °C in Anwesenheit von Dexamethason.	119
Tabelle 12: <i>Screening</i> bei 30 °C in Anwesenheit von Dexamethason. Isolierte hGR-LBD Varianten aus Anreicherung 5, 7 und 8.	122
Tabelle 13: Wachstum von transformierten <i>E. coli</i> T7 Express Zellen auf LB-Platten in Anwesenheit unterschiedlicher Konzentrationen an Chloramphenicol.	125
Tabelle 14: hGR-LBD Einzel- und Kombinationsvarianten zur Analyse von M560K, F602S, A605V, V702A, E705G und M752T.	132
Tabelle 15: Apparente T_M -Werte (T_M^{app}) der hGR-LBD Einzel- und Kombinationsvarianten.	143
Tabelle 16: Bindungsaffinität von TIF2 Coaktivator- bzw. NCoR Corepressor-Peptid an hGR-LBD Varianten, ermittelt durch Oberflächenplasmonresonanz (BIAcore TM).	163
Tabelle 17: Falsch positive Artefakte beim <i>Screening</i> nach stark fluoreszierenden eGFP-Fusionsproteinen <i>in vivo</i>	170
Tabelle A1: Molekulargewichte, molare Extinktionskoeffizienten und Ausbeuten der in dieser Arbeit aufgereinigten hGR-LBD Einzel- und Kombinationsvarianten.	199
Tabelle A2: Zur Kalibrierung der Gelfiltrationssäule Superdex S-75 verwendete Proteine.	200

Formelverzeichnis

Formel 1: Thermodynamische Stabilität eines Proteins.	4
Formel 2: Gibbs-Helmholtz Gleichung. Enthalpie-und Entropie-Abhängigkeit der thermodynamischen Stabilität eines Proteins.	5
Formel 3: Berechnung der Transformationseffizienz.....	52
Formel 4: Berechnung der Schmelztemperatur eines Oligonukleotids.	59
Formel 5: Berechnung der optimalen Anlagerungstemperatur der Primer in einer PCR-Reaktion.	60
Formel 6: Bestimmung der DNA-Konzentration.	66
Formel 7: Bestimmung der Ligationseffizienz in Prozent.	67
Formel 8: Bestimmung der Größe einer Genbank aus Formel 7 und Formel 8.	67
Formel 9: Berechnung der durchschnittlichen Basenaustausche pro Gen durch <i>ep</i> PCR.	68
Formel 10: Berechnung des molaren Extinktionskoeffizienten ϵ_{280}	76
Formel 11: Berechnung des spezifischen Extinktionskoeffizienten $^{0,1\%}A_{280}$	76
Formel 12: Berechnung der Proteinkonzentration mit Hilfe des spezifischen Extinktionskoeffizienten $^{0,1\%}A_{280}$	77
Formel 13: Berechnung der molaren Elliptizität pro Aminosäure (θ_{MRW}).	82
Formel 14: Refraktometrische Bestimmung der Guanidiniumchlorid-Konzentration (Pace, 1986)....	83
Formel 15: Refraktometrische Bestimmung der Harnstoff-Konzentration (Pace, 1986).....	84
Formel 16: Anteil an nativem (F_N) und denaturiertem (F_U) Protein unter Annahme des Zweizustands-Modells der Auffaltung.....	86
Formel 17: Zusammensetzung des Signalwertes y unter Annahme des Zweizustands-Modells der Auffaltung.....	86
Formel 18: Bestimmung des Anteils an ungefaltetem Protein aus den bei einem Auffaltungsexperiment gemessenen Signalintensitäten, unter Annahme des Zweizustands-Modells der Auffaltung.....	86

Abkürzungsverzeichnis

Aminosäuren wurden gemäß der Einbuchstaben- oder Dreibuchstaben-Schreibweise abgekürzt.

A	Adenosin, Aggregate, FACS-Anreicherung
A_x	Absorption bei x nm
Å	Ångström (10^{-10} m)
$0,1\% A_x$	spezifischer Extinktionskoeffizient bei x nm
A_{Basen}	Durchschnittliche Anzahl der Basenaustausche pro Gen
a_{Basen}	Anzahl der Basenaustausche pro analysiertem Gen
Amp	Ampicillin
APOB	Apoprotein B100
APS	Ammoniumperoxosulfid
ARH	Autosomal rezessive Hypercholesterinämie
AS	Aminosäure
ATP	Adenosintriphosphat
<i>bla</i>	β -Lactamase-Gen
bp	Basenpaare
bzw.	beziehungsweise
°C	Grad Celsius
C	Cytosin
c	Konzentration
ca.	circa
C α	zentrales C-2 Kohlenstoffatom von Aminosäuren
CD	Circular Dichroismus
C	<i>Crude extract</i> (lösliche Fraktion eines Zellextrakts)
C*	gereinigter <i>Crude extract</i>
CAT	Chloramphenicolacetyltransferase
d	Schichtdicke
D	denaturierter, ungefalteter Zustand eines Proteins
$[D_{1/2}]^{\text{app}}$	Konzentration an Denaturierungsmittel bei welcher 50 % des vorliegenden Proteins in nicht nativem Zustand vorliegt
ΔG_{xy}	Differenz der freien Energie zwischen Zustand x und y
ΔH_{xy}	Enthalpie-Differenz zwischen Zustand x und y
Δn	Differentieller Brechungsindex
ΔS_{xy}	Entropie-Differenz zwischen Zustand x und y
Da	Dalton
DBD	DNA-bindende Domäne
d.h.	das heißt
DNA	Desoxyribonukleinsäure
dsDNA	doppelsträngige DNA
dNTP	Desoxyribonukleotidtriphosphat
DTT	Dithiothreitol
eGFP	<i>enhanced GFP</i>
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraessigsäure
<i>et al.</i>	und andere
<i>epPCR</i>	<i>error prone PCR</i>
eps	analyisierte Ereignisse pro Sekunde bei FACS-Messungen

ER	Endoplasmatisches Reticulum
EtBr	Ethidiumbromid
f	Verdünnungsfaktor
FACS	<i>Fluorescence Activated Cell Sorting</i>
FL1	aufgenommenes Fluoreszenz Signal des MoFlo [®] Zytometers
F _N	Anteil an gefaltetem Protein
F _U	Anteil an ungefaltetem Protein
FPLC	<i>fast protein liquid chromatography</i>
FSC	aufgenommenes <i>Forward Scatter</i> Signal des MoFlo [®] Zytometers
G	Größe einer Genbank, Guanodin, freie Energie
g	Gramm
GC%	GC-Gehalt von Nucleotidsequenzen
G _D	freie Energie des entfaltenen Zustands eines Proteins
G _N	freie Energie des gefalteten Zustands eines Proteins
GdmCl	Guanidiniumchlorid
GFP	grün fluoreszierendes Protein aus <i>Aequorea victoria</i>
GST	Glutathion-S-Transferase
h	Stunde
H	Hinge-Region
hAR	humaner Androgen-Rezeptor α
hER	humaner Estrogen-Rezeptor α
His ₆	Anhang von 6 Histidinen an ein Protein
His ₆ -GST	Glutathion-S-Transferase N-terminal fusioniert mit His ₆
hGR	humaner Glucocorticoid-Rezeptor
<i>hgr-lbd</i>	Gen für die Ligand-Bindungsdomäne (Aminosäuren 521-777) des humanen Glucocorticoid-Rezeptors
hGR-LBD	Ligand-Bindungsdomäne (Aminosäuren 521-777) des humanen Glucocorticoid-Rezeptors
<i>hkas</i>	Gen für die β -Ketoacyl-Synthase Domäne der humanen Fettsäure-Synthase
hKAS	humane β -Ketoacyl-Synthase Domäne der Fettsäure-Synthase
hMR	humaner Mineralocorticoid-Rezeptor
hPCSK9	humane Proprotein Convertase Subtilisin/Kexin Typ 9
hPPAR	humaner Peroxisom-Proliferator-aktivierter Rezeptor γ
hPR	humaner Progesteron-Rezeptor
hRAR	humaner Retinsäure-Rezeptor γ
hRXR	humaner Retinoid-X-Rezeptor α
hTR	humaner Thyroid Hormon Rezeptor α
HPLC	<i>high performance liquid chromatography</i>
I	Intermediat
IPTG	Isopropyl- β -D-thiogalactopyranosid
K	Gleichgewichtskonstante; Kolonien
Kan	Kanamycin
kb	Kilobase
kDa	Kilodalton
KP	Kaliumphosphat-Puffer
kU	<i>kiloUnits</i> = 1000 U
L	Ligationseffizienz
l	Liter
<i>lacZ</i>	Gen für die β -Galactosidase

$[L_{1/2}]^{app}$	Salzkonzentration bei der 50% von zugegebenem Protein präzipitiert
LB	Luria-Bertani (-Medium)
LBD	Ligand-bindende Domäne
LDL	<i>Low Density</i> Lipoprotein
LDL-R	LDL-Rezeptoren
LMW	<i>LowMolecularWeight</i> Protein-Längenstandard
ln	natürlicher Logarithmus
log	dekadischer Logarithmus
M	Molar [mol/l]; Mutationen pro Gen bei Randomisierung einer Sequenz
mA	Milliampère
mg	Milligramm
mGR-LBD	Ligand-Bindungsdomäne (Aminosäuren 528-783) des Glucocorticoid-Rezeptors aus der Maus (<i>Mus Musculus</i>)
min	Minute
ml	Milliliter
mM	Millimolar
MPa	Megapascal
M_R	molekulare Masse
MW	Molekulargewicht eines Proteins
mwco	<i>molecular weight cut off</i>
N	nativer, gefalteter Zustand eines Proteins
$n_{Kolonien}$	Anzahl der Kolonien
n_{seq}	Anzahl sequenzierter Gene
n_T	Anzahl an transformierten Zellen
nm	Nanometer
NTD	N-terminale Domäne
OD_x	optische Dichte bei x nm
OE-PCR	<i>Overlap Extension</i> PCR
P	Pellet (unlösliche Fraktion eines Zellextraktes)
PA	Polyacrylamid
PAGE	Polyacrylamidgelelektrophorese
PCR	Polymerasekettenreaktion
<i>pcsk9</i>	Gen für die humane Proprotein Convertase Subtilisin/Kexin Typ 9
PDB	<i>Protein Data Bank</i>
<i>Pwo</i>	<i>Pyrococcus woeseii</i>
R	Gaskonstante ($8,314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$)
RMSD	<i>root mean square deviation</i>
rpm	Umdrehungen pro Minute
RT	Raumtemperatur
SB	Sortier-Bereich bei einer FACS-Anreicherung
SDS	<i>sodium dodecyl sulfate</i> ; Natriumdodecylsulfat
sek	Sekunde
SSC	aufgenommenes <i>Side Scatter</i> Signal des MoFlo [®] Zytometers
ssDNA	einzelsträngige DNA
SV	Säulenvolumen
T	Temperatur; Thymidin
T_A	<i>Annealing</i> -Temperatur eines <i>Primers</i>
T_E	Transformationeffizienz
T_M	Schmelztemperatur eines Oligonukleotids
T_M^{app}	Temperatur bei welcher 50 % des vorliegenden Proteins in nicht nativem Zustand vorliegt

<i>Tag</i>	fusionierter Anhang an ein Protein
<i>Taq</i>	<i>Thermus aquaticus</i>
TCEP	Tris(2-carboxyethyl)phosphin
TEMED	N,N,N',N'-Tetramethylethyldiamin
TFA	Trifluoressigsäure
<i>tmhisF</i>	Gen für die Synthase-Untereinheit der Imidazolglycerinphosphat-Synthase aus <i>Thermotoga maritima</i>
<i>tmHisF</i>	Synthase-Untereinheit der Imidazolglycerinphosphat-Synthase aus <i>Thermotoga maritima</i>
TPR	Tetratricopeptid Repeat Proteine
Ts	Transitionen
Tv	Transversionen
U	<i>Unit</i>
u.a.	unter anderem
üN	über Nacht
UV	Ultraviolett
V	Volt
Vol	Volumen
(v/v)	Volumen pro Volumen
wt	Wildtyp
(w/v)	Masse pro Volumen
y	Messwert
y_N	Signalwert von vollständig gefaltetem Protein
y_U	Signalwert von vollständig entfaltetem Protein
z.B.	zum Beispiel
ϵ_x	molarer Extinktionskoeffizient bei x nm
λ	Wellenlänge
θ_{MRW}	durchschnittliche molare Elliptizität pro Aminosäure
θ_{obs}	gemessene Elliptizität
μg	Mikrogramm
μl	Mikroliter
μM	Mikromolar

Zusammenfassung

Zur biophysikalischen Charakterisierung oder zur Aufklärung der Struktur eines Proteins muss dieses rekombinant in löslicher Form und in großen Mengen produziert werden. Die intrinsische Instabilität vieler eukaryotischer Proteine macht dies jedoch oft sehr schwierig bzw. gänzlich unmöglich. Obwohl dieses Problem in manchen Fällen durch Variation der Expressions- und/oder Reinigungsbedingungen umgangen werden kann, ist in vielen Fällen eine Optimierung der Aminosäuresequenz mittels Mutagenese erforderlich.

Bei Kenntnis der Struktur eines Proteins oder bei Vorliegen eines zuverlässigen Modells kann die Optimierung durch *rationales Design* erfolgen. Im ersten Projekt dieser Arbeit wurde auf der Basis mehrerer Homologie-Modelle die Löslichkeit der N-terminalen Domäne der humanen Proprotein Convertase Subtilisin/Kexin Typ 9 durch den Austausch 15 hydrophober Aminosäuren der Proteinoberfläche gegen hydrophilere Reste signifikant erhöht: die Ausbeute an gereinigtem Protein stieg auf das Fünffache des Ausgangswertes an und die zur Ausfällung von 50 % des Proteins benötigte Konzentration an Ammoniumsulfat erhöhte sich von 0,23 M auf 0,52 M. Allerdings war die optimierte Proteinvariante etwas weniger stabil gegenüber chemischer Denaturierung: die zur Entfaltung von 50 % des Proteins benötigte Konzentration an Harnstoff erniedrigte sich von 3,75 M auf 3,02 M. Diese Ergebnisse zeigen, dass Löslichkeit und konformationelle Stabilität eines Proteins nicht in jedem Fall korreliert sein müssen.

In Abwesenheit struktureller Information kann die Aminosäuresequenz eines Proteins mittels *gelenkter Evolution* optimiert werden. Dabei wird zunächst durch Zufallsmutagenese ein großes Repertoire an Proteinvarianten erzeugt, aus dem anschließend durch effiziente Selektions- oder *Screening*-Verfahren diejenigen mit den besten Eigenschaften isoliert werden können.

Im zentralen zweiten Teil dieser Arbeit wurde gelenkte Evolution zur Erhöhung der Löslichkeit und Stabilität der Ligand-Bindungsdomäne des humanen Glucocorticoid-Rezeptors (hGR-LBD) durchgeführt, wobei fusioniertes grün fluoreszierendes Protein (GFP) als Reporter diente. GFP weist ein leicht messbares Fluoreszenzsignal auf, dessen Intensität von der Löslichkeit bzw. Stabilität des fusionierten hGR-LBD abhängt und im Hochdurchsatz *Screening*-Verfahren mittels Fluoreszenz-aktivierter Zellsortierung (FACS) ausgelesen werden kann. Zunächst wurden durch Zufallsmutagenese mittels *error prone* PCR zwei plasmidkodierte *hgr-lbd-egfp* Genbanken mit hoher bzw. niedriger Mutationsfrequenz

hergestellt. Anschließend wurden die mit den beiden Banken transformierten und gemischten *E. coli* Zellen insgesamt acht FACS-Anreicherungsrounds unterzogen, wobei schrittweise nach höherer GFP-Fluoreszenz im Hochdurchsatz (> 10.000 Zellen/Sekunde) sortiert wurde. Hierdurch gelang die Isolierung von insgesamt fünf hGR-LBD Varianten (vier aus Anreicherungsrunde 5 und eine aus Anreicherungsrunde 8), die alle aus der hochmutagenen Genbank stammten und insgesamt 30 unterschiedliche Aminosäureaustausche enthielten.

Von diesen 30 Austauschen wurden 14 einzeln getestet, wobei sich herausstellte, dass die Mutationen A605V, V702A, E705G und M752T, ähnlich wie die in der Literatur beschriebene Mutation F602S (Bledsoe *et al.*, 2002), zu einer Erhöhung der Fluoreszenz des fusionierten GFP und zu einem Anstieg der Löslichkeit von hGR-LBD in Abwesenheit von GFP führen. Die Kombination der vorteilhaften Mutationen zeigte, dass die Effekte der einzelnen Austausche auf die Fluoreszenz und die Löslichkeit additiv sind und ein linearer Zusammenhang zwischen den beiden Parametern besteht. Die erhöhte Löslichkeit der optimierten hGR-LBD-Variante ermöglichte ihre Reinigung mit hoher Ausbeute und die Untersuchung ihrer konformationellen Stabilität mittels thermischer Auffaltung. Es zeigte sich, dass alle Austausche den apparenten Schmelzpunkt sowohl mit gebundenem Agonisten Dexamethason als auch (mit einer Ausnahme) mit gebundenem Antagonisten Mifepriston erhöhen und ein linearer Zusammenhang zwischen der Fluoreszenz von fusioniertem GFP und der thermischen Stabilität von hGR-LBD in Abwesenheit von GFP besteht. Die Einführung der in dieser Arbeit identifizierten vorteilhaften Mutationen (A605V, V702A, E705G und M752T) führt zu einer 26-fach verbesserten Reinigungsausbeute des Proteins aus der löslichen Zellfraktion von *E. coli*. Darüber hinaus wird durch die Austausche der apparente Schmelzpunkt sowohl der Agonist- als auch der Antagonist-Konformation um ca. 8 °C erhöht.

Im Laufe des FACS-basierten Hochdurchsatz-Screenings wurden verschiedene Artefakte beobachtet, welche die Isolierung optimierter hGR-LBD Varianten erschwerten. Die Ursachen dieser Artefakte, die sowohl auf Gen- als auch auf Proteinebene auftreten können, wurden aufgeklärt und Möglichkeiten zu ihrer Identifizierung und Umgehung erarbeitet.

Die stabilisierenden Mutationen A605V, V702A und E705G ermöglichten es, in Kooperation mit F.Hoffmann – LaRoche (Basel) die Röntgenkristallstruktur der homologen GR-LBD aus der Maus (mGR-LBD; 95,3 % Sequenz-Identität mit hGR-LBD) mit einer verbesserten Auflösung von 1,5 Å anstelle der bisher erzielten 1,9 Å aufzuklären. Durch die Analyse der Struktur konnten bisher nicht bekannte atomare Details sichtbar gemacht und der Effekt der stabilisierenden Mutationen erklärt werden. In weiterführenden Arbeiten können nun die

optimierten GR-LBD-Varianten zur Strukturaufklärung des Proteins mit neuen Ligand-Leitstrukturen verwendet werden, was bei der Entwicklung pharmazeutischer Wirkstoffe von Nutzen sein kann.

Im dritten Teil der Arbeit wurde ein in meiner Diplomarbeit begonnenes Projekt zur Stabilisierung des artifiziellen ($\beta\alpha$)₈-*Barrel* Proteins HisF-C**C*, welches aus zwei identischen, fusionierten C-terminalen Hälften von tmHisF (Synthase-Untereinheit der Imidazolglycerinphosphat-Synthase aus *Thermotoga maritima*) besteht, mit der Aufklärung einer hoch aufgelösten Röntgenkristallstruktur zum vorläufigen Abschluss gebracht. Dieses Projekt resultierte in zwei Publikationen (Seitz *et al.*, 2007; Höcker *et al.*, 2009), die sich im Anhang zu dieser Arbeit finden.

1 Einleitung

1.1 Faltung, Stabilität und Löslichkeit von Proteinen *in vitro*

1.1.1 Proteinfaltung unter Gleichgewichtsbedingungen

Im einfachsten Fall liegen Proteine in einem Gleichgewicht zwischen dem nativen, gefalteten Zustand (N) und dem denaturierten, ungefalteten Zustand (D) vor ($N \leftrightarrow D$).

Unter der thermodynamischen Stabilität eines Proteins (Gibbsche freie Ennergie der Proteinfaltung) versteht man die Differenz zwischen der freien Energie von D (G_D) und N (G_N) unter physiologischen Bedingungen (ΔG_{DN}) (Pace *et al.*, 2004) (Formel 1).

$$\Delta G_{DN} = G_D - G_N = -RT \ln K = -RT \ln [D]/[N]$$

Formel 1: Thermodynamische Stabilität eines Proteins.

[D] und [N]: Konzentrationen an D bzw. N im Gleichgewicht; G_D und G_N : freie Energien von D bzw. N; K: Gleichgewichtskonstante, ΔG_{DN} : Gibbsche freie Energie der Proteinfaltung, R: ideale Gaskonstante, T: Temperatur

Dieses Zwei-Zustands-Modell der Proteinfaltung trifft allerdings in den meisten Fällen nicht zu, da während der Faltung ein oder mehrere Faltungsintermediate (I) auftreten können (Jaenicke, 1999). Der Energieunterschied des nativen und denaturierten Zustands weist demnach mehrere Stufen auf, welche die Energieunterschiede zwischen den einzelnen, miteinander im Gleichgewicht stehenden Zuständen, beschreiben (Sancho *et al.*, 2002). Partiiell gefaltetes (I) bzw. ungefaltetes Protein (D) kann dem Gleichgewicht durch Aggregation entzogen werden, da bei unvollständiger Faltung hydrophobe Stellen des Proteins exponiert sein können (Abbildung 1).

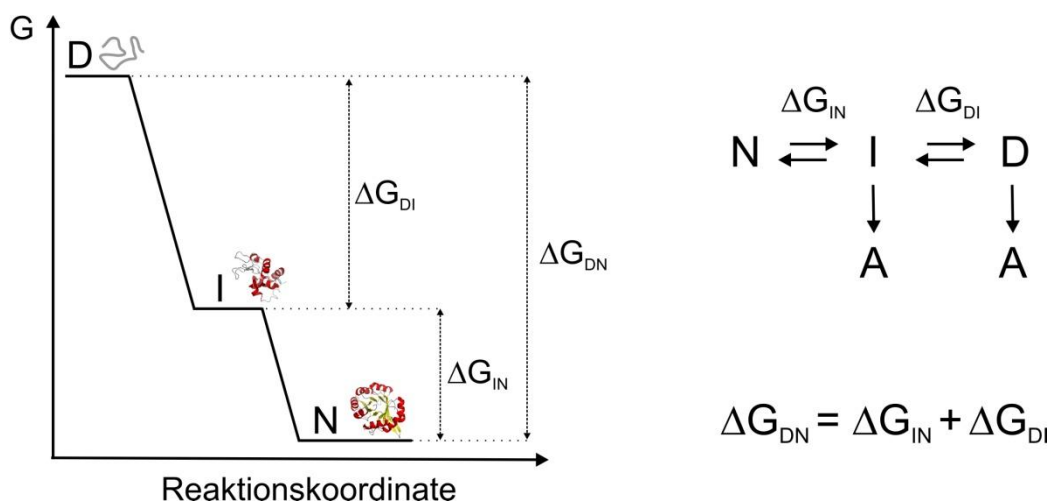


Abbildung 1: Faltung eines Proteins mit einem auftretenden Intermediat.

Übergangszustände sind nicht dargestellt.

G: freie Energie, D: denaturierter, entfalteter Zustand, N: nativer, gefalteter Zustand, I: Faltungsintermediat, A: Aggregate, ΔG_{xy} : Gibbsche freie Energie zwischen Zustand x und y, ΔG_{DN} : Gibbsche freie Energie der Proteinfaltung.

Zum Verständnis der Proteinfaltung ist es essentiell, Intermediate näher zu charakterisieren, was in einigen Fällen gelungen ist. Im Fall des α -Lactalbumins reichert sich beispielsweise bei niedrigen pH-Werten ein Faltungs-Intermediat an (Nozaka *et al.*, 1978; Kuwajima, 1996), welches über Circular Dichroismus Spektroskopie (Kuwajima *et al.*, 1976) und Wasserstoff-Deuterium Austausch-Versuche (Alexandrescu *et al.*; 1993, Chyan *et al.*, 1993) identifiziert und charakterisiert werden konnte. Dieses Faltungsintermediat besitzt native Sekundärstruktur und ist relativ kompakt. Allerdings wird keine konstante, sondern eine fluktuierende Tertiärstruktur beobachtet (Dolgikh *et al.*, 1981). Proteine mit solchen Eigenschaften werden auch als *Molten Globules* bezeichnet (Dolgikh *et al.*, 1983; Ptitsyn *et al.*, 1990) und stellen Zwischenschritte des Faltungswegs dar, die dabei helfen können das globale Energieminimum des nativen Zustands zu erreichen (Dobson, 1994).

1.1.2 Strukturelle Grundlagen der Proteinstabilität

Da kovalente Bindungen im gefalteten und entfalteten Zustand eines Proteins gleich stark sind, haben sie mit Ausnahme von Disulfidbrücken (siehe unten) keine Bedeutung für die konformationelle Proteinstabilität ΔG_{DN} (Whitford, 2005). Die Proteinstabilität wird demnach durch nicht-kovalente Interaktionen festgelegt (Dill, 1990, Pace *et al.*, 1996), wobei sich ΔG_{DN} nach der Gibbs-Helmholtz-Gleichung aus der Enthalpie- und der Entropie-Differenz des gefalteten und entfalteten Zustands zusammensetzt (Santoro & Bolen, 1988) (Formel 2).

$$\Delta G_{DN} = \Delta H_{DN} - T\Delta S_{DN}$$

Formel 2: Gibbs-Helmholtz Gleichung.

ΔG_{DN} : Gibbsche freie Energie der Proteinfaltung, ΔH_{DN} : Enthalpie-Differenz zwischen entfaltetem (D) und gefaltetem (N) Zustand des Proteins, ΔS_{DN} : Entropie-Differenz zwischen entfaltetem (D) und gefaltetem (N) Zustand des Proteins, T: Temperatur

Die treibende Kraft der Entfaltung ist die konformationelle Entropie, welche den denaturierten Zustand stabilisiert. Die Rotation um die vielen Einfachbindungen innerhalb einer Polypeptidkette ist im ungefalteten Zustand wesentlich einfacher möglich als im gefalteten Protein (Pace *et al.*, 2004). So begünstigt beispielsweise die konformationelle Entropie den entfalteten Zustand der RNAseA bei 25 °C um ca. 165 kcal · mol⁻¹ (Pace *et al.*, 1998). Um eine intakte Faltung zu erreichen muss das Protein enthalpisch und entropisch dieser konformationellen Entropie entgegenwirken, indem verschiedene Interaktionen nur im nativen Zustand ausgebildet werden (Pace *et al.*, 1996). Zu den stabilisierenden Faktoren zählen dabei der hydrophobe Effekt, Wasserstoffbrücken, ionische Bindungen und van-der-Waals Kräfte. Lange wurde angenommen, dass der hydrophobe Effekt die dominierende

Triebkraft für die Proteinfaltung darstellt (Kauzmann, 1959; Pace *et al.*, 2004). Mittlerweile ist jedoch bekannt, dass Wasserstoffbrücken und auch van-der-Waals Kräfte zwischen polaren Gruppen einen vergleichbar wichtigen Beitrag zur Proteinstabilität leisten (Pace *et al.*, 2004). So werden bei der Proteinfaltung zusätzlich ca. 1,1 H-Brücken pro Seitenkette ausgebildet (Stickle *et al.*, 1992), die jeweils eine Stabilisierung von etwa $1 \text{ kcal} \cdot \text{mol}^{-1}$ (Myers & Pace, 1996) bewirken. Ionische Wechselwirkungen zwischen ungleich geladenen Seitenketten tragen in der Regel weniger zur Stabilität bei, da sie durch die unvorteilhafte Dehydratisierung der geladenen Gruppen und durch die Erniedrigung der Seitenkettenentropie teilweise kompensiert werden (Hendsch & Tidor, 1994; Pace *et al.*, 2000; Bosshard *et al.*, 2004). Allerdings können Salzbrücken unter bestimmten Umständen die Stabilität um mehr als $4 \text{ kcal} \cdot \text{mol}^{-1}$ erhöhen (Anderson *et al.*, 1990; Pace *et al.*, 2004). Dies ist vor allem der Fall wenn sie, wie häufig bei Proteinen aus Hyperthermophilen, in ausgedehnten ionischen Netzwerken vorliegen (Kelly *et al.*, 1993; Britton *et al.*, 1995; Chan *et al.*, 1995; Hennig *et al.*, 1997; Korndörfer *et al.*, 1997; Zhang *et al.*, 2001), da dort die elektrostatischen Interaktionen kooperativ wirken können (Karshikoff & Ladenstein, 2001). Die erniedrigte Dielektrizitätskonstante von Wasser bei höheren Temperaturen sorgt zudem dafür, dass die Dehydratisierung von geladenen Aminosäuren weniger ungünstig wird, was das gehäufte Vorkommen von Salzbrücken bei Proteinen aus Hyperthermophilen im Vergleich zu Mesophilen erklärt (Elcock, 1998; Sterner & Liebl, 2001; Karshikoff & Ladenstein, 2001). Disulfidbindungen sind die einzigen kovalenten Verknüpfungen welche die thermodynamische Stabilität erhöhen können. Sie senken die konformationelle Entropie des entfalteten Zustands und erhöhen somit seine freie Energie (Pace *et al.*, 1990).

Insgesamt heben sich die Entropie und Enthalpie des gefalteten und entfalteten Zustands eines Proteins nahezu gegenseitig auf und der Netto-Energiebetrag der konformationellen Proteinstabilität ist relativ klein (Pace *et al.*, 1996). So liegen die ΔG_{DN} -Werte von in der Natur vorkommenden Proteinen zwischen 5 und $15 \text{ kcal} \cdot \text{mol}^{-1}$ (Pace, 1990). Der genaue Wert für ein Protein hängt neben den beschriebenen intrinsischen Eigenschaften der Polypeptidkette auch von äußeren Bedingungen wie Temperatur, pH-Wert und Osmolarität ab (Pace, 1990).

1.1.3 Löslichkeit von Proteinen

Aufgrund der Ausbildung von Wasserstoffbrücken zwischen geladenen bzw. polaren Aminosäuren-Seitenketten der Proteinoberfläche und Wassermolekülen der Umgebung sind Proteine von einer Hydrat-Hülle umgeben (Schein, 1990). Die Stärke der ausgebildeten Interaktionen zwischen Protein und Hydrat-Hülle hängt von der Aminosäuren-Zusammensetzung der Proteinoberfläche ab, wobei Aspartat bzw. Tryptophan die hydrophilste bzw. hydrophobeste Aminosäure darstellt (Trevino *et al.*, 2007). Die Löslichkeit von Proteinen kann durch Variation von Ionen-Stärke, pH und Temperatur beeinflusst werden (Trevino *et al.*, 2008). Bei niedrigen Salz-Konzentrationen in der Lösung kann die Proteinlöslichkeit verbessert werden („*salting-in*“), da die Bindung von Salz-Ionen an geladene Reste der Proteinoberfläche zu einer erhöhten Hydratisierung des Proteins führt (Trevino *et al.*, 2008). Bei hohen Salz-Konzentrationen erfolgt dagegen die Präzipitation der Proteine („*salting-out*“) (Arakawa & Timasheff, 1984), da die starke Hydratisierung der freien Salz-Ionen in der Lösung den Verlust der Protein-Hydrathülle zur Folge hat; durch hydrophobe Wechselwirkungen kommt es in der Folge zur Aggregation und Präzipitation der Proteine. Die Stärke dieses Aussalzungseffektes hängt nach den sogenannten „Hofmeister-Serien“ von der Art des Salzes ab (Kationen: $\text{NH}_4^+ > \text{K}^+ > \text{Na}^+ > \text{Li}^+ > \text{Mg}^{2+} > \text{Ca}^{2+}$; Anionen: $\text{SO}_4^{2-} \approx \text{HPO}_4^{2-} > \text{Acetat} > \text{Cl}^- > \text{Br}^- > \text{I}^- \approx \text{ClO}_4^- > \text{SCN}^-$), wobei dem Anion generell mehr Bedeutung zugesprochen wird (Cacace *et al.*, 1997; Trevino *et al.*, 2008).

Daneben können osmolytisch aktive Substanzen Proteine destabilisieren bzw. stabilisieren (Timasheff, 1992; Timasheff, 1993; Timasheff, 1995). Die entscheidende Rolle bei diesen Effekten spielt Wasser, da das Protein durch den Einfluss der Verbindungen auf die Hydrat-Hülle stabilisiert oder destabilisiert wird (Timasheff, 1993). Chaotrope Salze wie Guanidiniumchlorid oder Harnstoff stören generell die Ausbildung von Wasserstoffbrücken innerhalb des Solvents und verhindern somit die Ausbildung einer intakten Hydrathülle des gefalteten Zustands des Proteins (Schein, 1990). Zudem bindet Harnstoff sowohl an die polaren als auch die unpolaren Bestandteile eines Proteins. Dadurch wird der hydrophobe Effekt, welcher für die Proteinfaltung essentiell ist, abgeschwächt und der entfaltete Zustand des Proteins wird besser solvatisiert (Rossky, 2008). Stabilisierende Verbindungen, wie bestimmte Zucker (z.B. Sucrose, Trehalose, Sorbitol, Glycerol), Aminosäuren (z.B. Prolin, Glycin), Methylamine (z.B. Sarcosin, Betain) (Knapp *et al.*, 1999) und kompatible Solute in Hyperthermophilen (z.B. Cyclic-2,3-Diphosphoglycerat, Diglycerol-Phosphat) (Sternner & Liebl, 2001) besitzen eine niedrigere Affinität zum Protein als Wasser. Dadurch werden sie bevorzugt vom Protein separiert und es kommt durch eine verbesserte

Hydratisierung des Proteins zur Stabilisierung des gefalteten Zustands (Timasheff & Arakawa, 1997).

1.2 Proteinfaltung und Aggregation *in vivo*

Die Faltung von Proteinen *in vivo* wird durch spezielle Helfer-Proteine, den sogenannten Chaperonen unterstützt. Diese binden, stabilisieren und remodellieren nicht-native Proteinspezies (Baneyx & Mujacic, 2004). Die Expression vieler dieser Chaperone wird bei einem Hitze-Schock hochreguliert, weshalb einige von ihnen auch als Hitze-Schock-Proteine (Hsp) bezeichnet werden (Becker & Craig, 1994). Neben Prolyl-Peptidyl Cis/Trans-Isomerasen (PPIasen), können molekulare Chaperone in 3 Klassen eingeteilt werden (Baneyx & Mujacic, 2004): „Faltungs“-Chaperone (z.B. Hsp70, Hsp60) (Bukau & Horwich, 1998), die ATP-abhängig die Faltung fehlerhafter Proteinspezies katalysieren, „Halte“-Chaperone (z.B. IbpB, Hsp31) (Lund, 2001; Sastry *et al.*, 2002), die an partiell gefaltete Proteine binden und sie den anderen Chaperonen bereitstellen und „disaggregierende“-Chaperone (z.B. Hsp100 = ClpB) (Parsell *et al.*, 1994; Horwich, 2004), welche die Auflösung aggregierter Proteine katalysieren. Bei rekombinanter Expression von Proteinen in *Escherichia coli* ist die Herstellung gefalteten, löslichen Proteins dennoch eine große Herausforderung, da intrazelluläre Konzentrationen von 300 – 400 mg/ml erreicht werden (Ellis *et al.*, 2003). Die fehlende Fähigkeit von Bakterien zur Durchführung jedweder post-translationaler Modifikation erhöht zudem die Wahrscheinlichkeit der Fehlfaltung heterolog exprimierter, eukaryontischer Proteine (Villaverde & Carrió, 2003). So werden z.B. im reduzierenden Milieu des Bakterien-Cytoplasmas keine Disulfidbrücken ausgebildet (Fahey *et al.*, 1977; Makrides, 1996), wodurch disulfidreiche Proteine (z.B. Fab Antikörper-Fragmente) mit hoher Sicherheit aggregieren. Bei Anhäufung von fehlgefalteten Proteinspezies im Cytoplasma kommt es in *E. coli* zum Abbau des Proteins oder zur Bildung sogenannter „inclusion bodies“ (Baneyx & Mujacic, 2004). Diese wachsen durch Aggregation von partiell strukturierten Faltungs-Intermediaten und bilden poröse Kugeln oder Zylinder mit einer maximalen Länge von ca. 1 μm und einem maximalen Volumen von ca. 0,6 μm^3 (Abbildung 2) (Bowden *et al.*, 1991; Taylor *et al.*, 1986; Carrió *et al.*, 1998). Die aggregierten Proteine können zum Teil nach Solubilisierung und Reinigung der *inclusion bodies in vitro* rückgefaltet werden (Weir & Sparks, 1987; Tsuji *et al.*, 1987; Tsumoto *et al.*, 2003; Gustafsson *et al.*, 2009), was allerdings aufwendig ist und für jedes Protein optimiert werden muss (Vallejo & Rinas, 2004).



Abbildung 2: Elektronenmikroskopische Aufnahme von *inclusion bodies* in *Escherichia coli* Zellen.

(<http://web.mit.edu/king-lab/www/research/Scott/Scott-Research.html>)

Die zu *inclusion bodies* aggregierten Proteine sind an den Polen der Zelle als elektronendichte Partikel erkennbar.

1.3 Erhöhung der Proteinstabilität

Es wurden zahlreiche Strategien entwickelt, um Proteine für die heterologe Expression in *E. coli* bzw. für biophysikalische Untersuchungen wie z.B. die Röntgenkristallographie zu optimieren. Neben der Modifikation der Versuchsbedingungen oder des Aufbaus des zu untersuchenden Proteinkonstrukts („extrinsische“ Optimierung, 1.2.1) kann dies auch durch Mutagenese bestimmter Aminosäuren erfolgen („intrinsische“ Optimierung, 1.2.2).

1.3.1 Extrinsische Optimierung der heterologen Expression und der Kristallisation von Proteinen

Um die Ausbeute an löslichem Protein bei heterologer Expression in *E. coli* extrinsisch zu erhöhen kann man Chaperone koexprimieren (Goloubinoff *et al.*, 1989; Roman *et al.*, 1995), die Expressionsrate erniedrigen (Hockney, 1994), stabilisierende Verbindungen wie Sorbitol oder Betain zugeben (Blackwell & Horgan, 1991), osmotischen Schock durch NaCl induzieren um die intrazelluläre Menge an Chaperonen und Osmoprotektanten zu erhöhen (Chopra *et al.*, 1994), die Expression bei niedrigeren Temperaturen (Georgiou & Valax, 1996) und in Anwesenheit von stabilisierenden Liganden (Bledsoe *et al.*, 2002) durchführen, die Proteine in das Periplasma der Zellen sekretieren lassen um Disulfidbindungen zu ermöglichen (Becker & Hsiung, 1986; Humphreys *et al.*, 2004) oder die Zielproteine an extrem lösliche Fusionsproteine wie z.B. das Maltose-Bindeprotein (Sachdev & Chirgwin; 1998), Thioredoxin (LaVallie *et al.*, 2000) oder die Glutathion-S-Transferase (Smith, 2000) koppeln.

Mit Hilfe solcher unterstützender Proteine konnte auch die Röntgenkristallstrukturanalyse einiger an sie fusionierter Proteine erleichtert werden (Kobe *et al.*, 1999; Liu *et al.*, 2001; Ke & Wolfberger, 2003), wobei die Heterogenität von Fusionsproteinen die Kristallisation im allgemeinen eher hemmt (Smyth *et al.*, 2003). Eine andere Strategie zur Erleichterung der Kristallisation schwieriger Zielproteine mit hoher Flexibilität oder stark hydrophoben Oberflächen (z.B. bei Membranproteinen) stellt die Stabilisierung durch spezifische Bindeproteine dar. So führte die Bindung von Antikörper-Fragmenten (Ostermeier *et al.*, 1995; Hunte *et al.*, 2000; Zhou *et al.*, 2001; Jiang *et al.*, 2003; Day *et al.*, 2007; Rasmussen *et al.*, 2007), „Affibodies“ (Wahlberg *et al.*, 2003), „Monobodies“ (Koide *et al.*, 2007) oder designer Ankyrin-Repeat-Proteine (DARPs) (Binz *et al.*, 2004; Kohl *et al.*, 2005; Schweizer *et al.*, 2007; Sennhauser *et al.*, 2007; Bandejas *et al.*, 2008) zur erfolgreichen Strukturaufklärung von Proteinen. Die Erhöhung lokaler Rigidität zur Erhöhung der Kristallisationswahrscheinlichkeit kann auch noch durch andere Vorgehensweisen erreicht werden. So gelang die Kristallisation des K⁺-Kanals Kir3.1 der Maus nach Einführung einiger Stellen aus dem homologen Protein von *B. xenovorans* (Nishida *et al.*, 2007), während die Struktur des humanen β_1 -adrenergen G-Protein-gekoppelten Rezeptors nach Einführung stabilisierender Punktmutationen gelöst werden konnte (Serrano-Vega *et al.*, 2008; Warne *et al.*, 2008). Die Kristallisation des humanen β_2 -adrenergen G-Protein-gekoppelten Rezeptors erforderte die Insertion von T4-Lysozym in einen flexiblen Bereich (Cherezov *et al.*, 2007; Rosenbaum *et al.*, 2007). Letztere Strategien fallen in den Bereich der intrinsischen Optimierung, die im nächsten Abschnitt genauer besprochen wird.

1.3.2 Intrinsische Optimierung der Stabilität und Löslichkeit von Proteinen durch Mutagenese

Zur Erhöhung der löslichen Expression in *E. coli* bzw. der Kristallisationswahrscheinlichkeit bietet sich zum einen rationales Proteindesign an (van den Burg & Eijsink, 2002; Eijsink *et al.*, 2004). Dadurch wurde zum Beispiel die Rigidität von Proteinen durch Einführung von Gly→Ala oder Xxx→Pro Austausch (Matthews *et al.*, 1987) bzw. von Disulfidbrücken (Matsumura *et al.*, 1989; Mansfeld *et al.*, 1997; Tigerström *et al.*, 2004; Hagihara *et al.*, 2007; Searens *et al.*, 2008) erhöht, wodurch die Entropiebegünstigung des entfalteten Zustands geringer wird. Reste, die eingeführt wurden, um mit einem α -Helix-Dipol zu interagieren (Nicholson *et al.*, 1988; Nicholson *et al.*, 1991), auf andere Weise α -Helices (Serrano *et al.*, 1992; Blaber *et al.*, 1993) oder andere Sekundärstrukturbestandteile wie β -turns (Simpson *et al.*, 2005) optimieren, können ebenfalls zur Stabilisierung von Proteinen führen. Ähnlich

wirksam kann die Einführung von Interaktionen zwischen aromatischen Seitenketten (Serrano *et al.*, 1991; Puchkaev *et al.*, 2003), die Erhöhung des Anteils an Konsensus-Resten (Lehmann *et al.*, 2002; Bledsoe *et al.*, 2002) oder die Übertragung von Resten aus homologen Proteinen hyperthermophiler Organismen sein (Littlechild *et al.*, 2007). Durch verbesserte Protomer-Interaktionen kann die Quartärstruktur bei oligomeren Proteinen stabilisiert werden (Kaneko *et al.*, 2005). Die Einführung bzw. Optimierung von Salzbrücken-Netzwerken kann ebenfalls zu einer deutlichen Erhöhung der konformationellen Stabilität führen (Sun *et al.*, 1991; Waldburger *et al.*, 1995; Pace *et al.*, 2000; Makhatadze *et al.*, 2003; Schwehm *et al.*, 2003; Höcker *et al.*, 2004; Gribenko *et al.*, 2009). In vielen Fällen führten bereits wenige Austausche zu einer deutlichen Stabilisierung eines Proteins (Mansfeld *et al.*, 1997; Perl *et al.*, 2000; Pedone *et al.*, 2001). Das eindrucksvollste Beispiel stellt bis dato die Stabilisierung der Triosephosphat-Isomerase aus *Leishmania mexicana* um 26 °C durch eine einzige Punktmutation dar (Williams *et al.*, 1999).

Neben rationalem Design, für das im idealen Fall eine 3D-Struktur vorhanden sein sollte, besteht auch die Möglichkeit, Proteine durch „gelenkte Evolution“ im Labor zu stabilisieren (Arnold, 1998; Eijnsink *et al.*, 2005; Roodveldt *et al.*, 2005). Dabei wird das für das Protein kodierende Gen durch eine geeignete Technik zunächst global oder lokal randomisiert mittels *error prone* PCR (Leung *et al.*, 1989), DNA-Shuffling (Stemmer *et al.*, 1994a,b) oder einer PCR mit degenerierten Oligonukleotiden als Primer (Neylon, 2004; Reetz *et al.*, 2008). Die dadurch erzeugten Genbanken können anschließend mittels geeigneter Selektions- oder *Screening*-Verfahren auf stabilisierte Varianten durchsucht werden.

Eine Methode zum *Screening* löslich exprimierter Proteine in *E. coli* ist der „*Colony Filtration Blot*“ (Cornvik *et al.*, 2005; Dalroth *et al.*, 2006). Hierbei kann nach Übertragung der auf Nährplatten gewachsenen *E. coli* - Kolonien auf eine Nitrocellulosemembran die lösliche Expression des Zielproteins durch einen spezifischen Antikörper nachgewiesen werden. Ein anderes System beruht darauf, die zelluläre Stressantwort auf falsch gefaltete Proteine zu analysieren. Durch Kopplung der Stressantwort an eine Farbreaktion können instabile Proteinvarianten erkannt und aussortiert werden (Lesley *et al.*, 2002). Weiterhin besteht die Möglichkeit, die erhöhte Protease-Resistenz eines stabilisierten Proteins an die Infektiosität von Phagen zu koppeln („*Proside*“; Sieber *et al.*, 1998; Martin *et al.*, 2001, Martin *et al.*, 2003; Kather *et al.*, 2008). Eine andere Methode stabilisierte Proteinvarianten zu identifizieren stellt die Selektion in thermophilen Wirten dar. So wurden stabilisierte Varianten der 3-Isopropylmalat-Dehydrogenase aus der Leucin-Biosynthese von Hefe über ihre Fähigkeit zur Komplementation eines auxotrophen *Thermus thermophilus* Stammes bei

schrittweiser Erhöhung der Wachstumstemperatur von 50 – 70 °C isoliert (Tamakoshi *et al.*, 2001).

Neben diesen Systemen gibt es noch andere Methoden der gelenkten Evolution, bei denen das Signal eines fusionierten Reporterproteins analysiert wird. Diese Verfahren beruhen darauf, dass die Stabilität des Zielproteins die Löslichkeit des Reporterproteins beeinflusst. So beeinflusst zum Beispiel ein an das α -LacZ gekoppeltes Zielprotein dessen Fähigkeit zur strukturellen Komplementation von ω -LacZ (Wigley *et al.*, 2001). Ein anderes Reporter-System basiert darauf, dass unlösliches Zielprotein zur Aggregation der fusionierten Chloramphenicol-Acetyltransferase (CAT) führt, wodurch die vermittelte Antibiotikum-Resistenz erniedrigt wird (Maxwell *et al.*, 1999; Sieber *et al.*, 2001; Sieber, 2003).

Das grün fluoreszierende Protein (GFP) aus *Aequorea victoria* (siehe 1.4) kann ebenfalls als Reporterprotein benutzt werden. Hierbei führt unlösliches Zielprotein zur Fehlfaltung des fusionierten GFP, wodurch dessen Fluoreszenz erniedrigt wird (Waldo *et al.*, 1999; Drew *et al.*, 2001; Waldo, 2003). Transformierte *E. coli* Zellen „leuchten“ damit umso intensiver, je löslicher das Zielprotein ist. Mit Hilfe von GFP als Reporterprotein konnte die Nucleosid-Diphosphat-Kinase aus *Pyrobaculum aerophilum* (Pedelacq *et al.*, 2002) und das Genprodukt Rv2002 aus *Mycobacterium tuberculosis* (Yang *et al.*, 2002; Yang *et al.*, 2003) stabilisiert werden, was die Röntgenstrukturanalyse dieser Proteine ermöglichte. Bei der Beurteilung der Löslichkeit des Zielproteins durch die GFP-Fluoreszenz kann es zu falsch positiven Artefakten kommen. So kann beispielsweise das Zielprotein durch intermolekulare Wechselwirkungen das GFP-Chromophor beeinflussen (Cabantous *et al.* 2005) oder es kann aufgrund von Proteolyse bzw. neu eingeführten ribosomalen Bindestellen zu einem stark verkürzten Zielprotein kommen, dessen Löslichkeit keinen Einfluss mehr auf die Faltung des fusionierten GFP hat (Cabantous *et al.*, 2008). Um solche Artefakte während des Screenings zu vermeiden, wurde das GFP-Reportersystem in einigen Arbeiten abgewandelt. Hierbei wurde das Zielprotein nur an einen Teil eines modifizierten *Superfolder*-GFPs (Pedelacq *et al.*, 2006) (AS 216-228) fusioniert und der Rest (AS 1-215) koexprimiert („Split-GFP“, Cabantous *et al.* 2005; Cabantous & Waldo, 2006) oder das Zielprotein in das *Superfolder*-GFP inseriert (Cabantous *et al.*, 2008). Dadurch konnten intermolekulare Wechselwirkungen bzw. artifizielle Verkürzungen, welche zu einer falsch positiven Erhöhung des Fluoreszenz-Signals führen, ausgeschlossen werden. Mit Hilfe von Pipettierrobotern wurde das Split-GFP System soweit automatisiert, dass mit Hilfe von 96er-Well Platten bis zu 2304 Klone pro Woche analysiert werden konnten (Listwan *et al.*, 2009). Da in dieser Arbeit vorwiegend GFP als Reporterprotein zur Anwendung kam, wird es im nächsten Kapitel näher vorgestellt.

1.4 Das grün fluoreszierende Protein (GFP) und seine Anwendungen

Das grün fluoreszierende Protein (GFP) wurde 1962 zum ersten Mal aus der Kristallqualle *Aequorea victoria* (Abbildung 3a) isoliert (Shimomura *et al.*, 1962). Im Jahre 1994 gelang die Klonierung der cDNA, wodurch GFP molekularbiologischen und biochemischen Anwendungen zugänglich gemacht werden konnte (Chalfie *et al.*, 1994). GFP besteht aus 238 Aminosäuren, besitzt ein Molekulargewicht von 28 kDa, und bildet als Tertiärstruktur ein β -Fass aus elf antiparallelen β -Strängen (Abbildung 3b) (Yang *et al.*, 1996).

Es absorbiert Licht maximal bei Wellenlängen von 395 nm (UV) oder 475 nm (blau), woraufhin Licht mit einer Wellenlänge von $\lambda_{\max} = 508$ nm bzw. $\lambda_{\max} = 503$ nm (grün) emittiert wird (Heim *et al.*, 1994; Tsien, 1998). Die in der Tiefsee in absoluter Dunkelheit lebende Qualle stellt das zur Anregung benötigte blaue Licht mit Hilfe des Photoproteins Aequorin selbst her (Abbildung 3b) (Morise *et al.*, 1974). Aequorin bindet kovalent das Luciferin-Molekül Coelenterazin (Shimomura, 2005). Bei Bindung von Calciumionen erfolgt eine Konformationsänderung, woraufhin die kovalente Bindung zwischen Aequorin und Coelenterazin gespalten wird, Coelenterazin unter CO_2 -Abspaltung zu Coelentaramid abreagiert und Licht mit einer Wellenlänge von $\lambda_{\max} = 470$ nm (blau) freigesetzt wird (Shimomura, 2005).

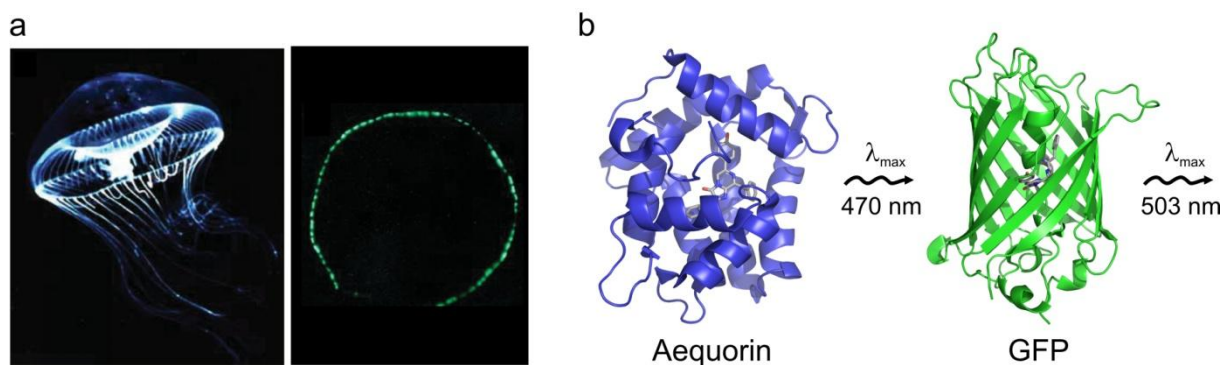


Abbildung 3: Die Biolumineszenz der Kristallqualle *Aequorea victoria* durch Aequorin und GFP.

a Die Biolumineszenz der Kristallqualle *Aequorea victoria* findet nur in einem schmalen, grün leuchtenden Ring am Rand des Schirms statt, wo sich das GFP in sogenannten photophoren Zellen befindet (aus Veith & Veith, 2005). **b** GFP (pdb 1gfl) emittiert grünes Licht ($\lambda_{\max} = 503$ nm) nach Anregung mit blauem Licht. Dieses wird durch das Photoprotein Aequorin (pdb 1ej3) mit Hilfe von Ca^{2+} und eines kovalent gebundenen Coelenterazin-Moleküls erzeugt.

Bei GFP ist der Chromophor nicht wie bei Aequorin ein zusätzlich gebundenes Molekül, sondern wird von einem Teil der Polypeptidkette selbst gebildet. So liegt im Zentrum des β -Fasses eine kurze α -Helix, in deren Mitte die Aminosäuren Serin 65, Tyrosin 66 und Glycin 67 durch eine autokatalytische Zyklisierungs-Reaktion ein konjugiertes Elektronenresonanzsystem ausbilden (Abbildung 4) (Cubitt *et al.*, 1995; Tsien, 1998).

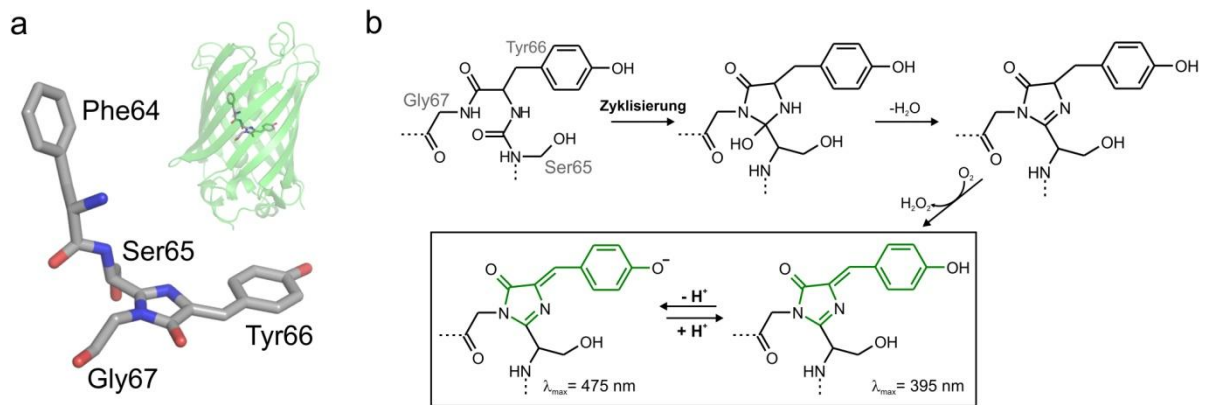


Abbildung 4: Chromophor des GFP bestehend aus Ser65, Tyr66 und Gly67.

a Der Chromophor des GFP (pdb 1gfl) wird aus 3 Aminosäuren gebildet, die im Inneren des elfsträngigen β -Fasses sitzen. In *enhanced* GFP (eGFP) ist Phe64 zu Leucin und Ser65 zu Threonin mutiert. **b** Mechanismus der Chromophor-Bildung in GFP nach Cubitt *et al.* (1995). Der erste Ring des Chromophors (p-Hydroxybenzylidenimidazolinon) wird durch eine autokatalytische Zyklisierung gebildet, wobei die Amidgruppe von Gly67 die Carbonylgruppe von Ser65 nukleophil angreift und anschließend H_2O abgespalten wird (= Imidazolinon). Durch O_2 wird die α - β -Bindung von Tyr66 unter Bildung von H_2O_2 dehydriert, wodurch die phenolische Seitenkette in Konjugation mit dem Imidazolinonring gebracht wird. Das Elektronenresonanzsystem (grün) kann in einer protonierten (neutrales Phenol) und deprotonierten (Phenolat-Anion) Form vorliegen, was die beiden unterschiedlichen Anregungsmaxima von GFP erklärt.

Zur molekularbiologischen und biotechnologischen Nutzung von GFP wurden zahlreiche GFP-Derivate entwickelt, bei denen die Fluoreszenzeigenschaften und die Stabilität des Proteins variiert wurden (Tsien, 1998). So gibt es mittlerweile viele in ihrer Leuchtkraft modifizierte Versionen des GFP, die eine erhöhte Fluoreszenz aufweisen („*enhanced* GFP“ = eGFP: F64L+S65T) (Heim *et al.*, 1995; Yang *et al.*, 1996; Cormack *et al.*, 1996; Patterson *et al.*, 1997) bzw. blau („BFP“: Y66H), gelb („YFP“: F64L+S65G+T203Y) oder cyan („CFP“: Y66W) leuchten (Shaner *et al.*, 2007). Diese neuen GFP-Varianten können für FRET (Fluoreszenz-Resonanz-Energie-Transfer) -Experimente benutzt werden, bei denen die Überlagerung des Emissionsspektrums einer GFP-Variante (z.B. CFP) mit dem Anregungsspektrum einer anderen GFP-Variante (z.B. YFP) ausgenutzt wird (Tsien *et al.*, 1993). Ein Beispiel für solche FRET-Experimente ist die Verwendung eines CFP-Calmodulin-YFP Fusionsproteins als Ca^{2+} -Indikator. Hierbei kommt es nach Ca^{2+} -Bindung zur Annäherung von CFP und YFP, weshalb die Emission des Fusionsproteins nach Anregung des CFP von cyan nach gelb umschlägt (Miyawaki *et al.*, 1999).

Split-GFP- bzw. BiFC- („*Bimolecular fluorescence complementation*“) Systeme (siehe 1.3.2) werden bevorzugt zur Analyse von Protein-Protein-Interaktionen benutzt. Hierbei werden die potentiellen Interaktionspartner mit sich ergänzenden Hälften von GFP fusioniert und nur bei intakter Interaktion kann sich das fluoreszierende Protein rekonstituieren (Hu *et al.*, 2002; Hu & Kerppola, 2003; Hu *et al.*, 2006; Barnard *et al.*, 2008; Horvath *et al.*, 2009).

Modifizierte GFP-Derivate werden daneben auch vielfach in der Zellbiologie angewendet zur Analyse von Genexpression (Chalfie *et al.*, 1994) bzw. Gentransfer (Cheng *et al.*, 1996; Muldoon *et al.*, 1997), zur Visualisierung von Proteinen, Zellkompartimenten bzw. Zell-Transportvorgängen (Yokoe & Meyer, 1996; Hanakam *et al.*, 1996; DeGiorgi *et al.*, 1996; Straight *et al.*, 1997; Presley *et al.*, 1997; Grebenok *et al.*, 1997) und zur Messung des intrazellulären pH-Wertes (Llopis *et al.*, 1998) bzw. des Membranpotentials (Siegel & Isacoff, 1997). Zur selektiven bzw. induzierbaren Markierung stehen auch photoaktivierbare GFP-Derivate zur Verfügung (Patterson *et al.*, 2004; Haigh *et al.*, 2007; Baltrusch *et al.*, 2008). Aufgrund der beschriebenen enormen Bedeutung des grün fluoreszierenden Proteins für die Molekularbiologie, Biochemie und medizinische Forschung erhielten Osamu Shimomura, Marty Chalfie und Roger Tsien für die Entdeckung und Weiterentwicklung von GFP den Nobelpreis für Chemie 2008.

Im Rahmen dieser Arbeit wurde eine Variante von GFP verwendet, die u.a. die Austausch von eGFP (F64L und S65T) (Heim *et al.*, 1995; Yang *et al.*, 1996; Cormack *et al.*, 1996; Patterson *et al.*, 1997) trägt. Durch diese beiden Austausch erfolgt aufgrund der Bevorzugung des Phenolat-Anions im Chromophor (Abbildung 4) eine Unterdrückung des Anregungsmaximums bei 395 nm, während das höherwellige Anregungsmaximum eine Rotverschiebung von 475 nach 488 nm erfährt und die Fluoreszenz-Emission stark erhöht wird (Heim *et al.*, 1995; Cormack *et al.*, 1996). So zeigt eGFP nach Anregung bei 488 nm eine 30-fach stärkere Fluoreszenz als das wildtypische GFP, wobei das Emissionsmaximum bei 507 nm liegt (Cormack *et al.*, 1996). Die Fluoreszenz des eGFP kann somit optimal mit einem Argonlaser angeregt werden, weshalb es sich sehr gut für die Durchflusszytometrie eignet (Cormack *et al.*, 1996). Neben diesen Mutationen trägt die in dieser Arbeit verwendete eGFP-Variante zusätzlich die sogenannten „Cycle-3“-Mutationen (F99S, M153T, V163A), welche die Faltungseigenschaften von GFP bei 37 °C verbessern und die Aggregationstendenz bei hohen Konzentrationen reduzieren (Cramer *et al.*, 1996). Zusätzlich ist der weit verbreitete, neutrale Aminosäureaustausch Q80R enthalten, der bei der ersten Klonierung durch einen PCR-Fehler eingeführt wurde (Chalfie *et al.*, 1994). Die beschriebene eGFP-Variante wurde in der vorliegenden Arbeit, wie in Kapitel 1.2.2 erläutert, als Reporter für die Stabilität eines fusionierten Proteins *in vivo* benutzt, wobei die mit den verschiedenen Konstrukten transformierten *E. coli*-Zellen mit Hilfe eines MoFlo[®] Zytometers analysiert und sortiert wurden.

1.5 Durchflusszytometrie

Die Durchflusszytometrie bzw. FACS (*Fluorescence activated cell sorting*) – Analyse dient zur quantitativen Hochdurchsatz-Analyse und -Sortierung von Zell-Populationen. Bei dieser Technik kann je nach Gerät die Zellform und/oder die Fluoreszenz (bis zu 4 verschiedene Laser zur Anregung) von mehreren Tausend Zellen pro Sekunde analysiert werden (Ibrahim & van den Engh, 2007). Die Durchflusszytometrie hat breite Anwendung in der klinischen Forschung und Diagnostik gefunden, wie z.B. bei der CD4 T-Helferzell-Bestimmung zur HIV-Diagnostik (Hengel & Nicholson, 2001; Mandy, 2004), der Anreicherung von Stammzellen (Sasaki *et al.*, 1995) oder der *in vitro* Fertilisation (Johnson *et al.*, 1993). Auch in der Zellbiologie gibt es zahlreiche FACS-Anwendungen, z.B. zur Analyse des Zellzykluses (Schorl & Sedivy, 2007) oder der Apoptose (Hauptmann & Lehle, 2008). In der angewandten Biochemie wird die Durchflusszytometrie zum Hochdurchsatz-Screening bei der gelenkten Evolution von Enzymen (Aharoni *et al.*, 2005; Mastrobattista *et al.*, 2005; Griswold *et al.*, 2005; Griswold *et al.*, 2006; Aharoni *et al.*, 2006; Miller *et al.*, 2006; Becker *et al.*, 2007; Lipovsek *et al.*, 2007) oder Membranproteinen (Sarkar *et al.*, 2008) mit fluorogenen Substraten bzw. Liganden benutzt. Die Funktionsweise des in dieser Arbeit benutzten MoFlo[®] Zytometers wird in 3.2.9 näher erläutert.

1.6 Die humane Proprotein Convertase Subtilisin/Kexin Typ 9 (hPCSK9)

Die familiäre Hypercholesterinämie (FH) ist gekennzeichnet durch einen erhöhten Spiegel an *Low-Density* Lipoprotein (LDL) im Blutplasma. Dies führt zu Ablagerungen von LDL-Cholesterin in Geweben, Haut und Arterien, wodurch Arteriosklerose und frühzeitige kardiovaskuläre Erkrankungen gefördert werden (Goldstein & Brown, 2001). Die FH kann vererbt werden, wobei mehrere molekulare Ursachen für die Störung des Lipid-Stoffwechsels existieren. In sehr seltenen Fällen (< 1 von 10⁷ Menschen) tritt eine autosomal rezessive FH auf. Diese wird durch mutierte Formen des Adaptor-Proteins ARH (autosomal rezessive Hypercholesterinämie) hervorgerufen, welches an den cytoplasmatischen Schwanz der LDL-Rezeptoren (LDL-R) bindet und die Anhäufung von LDL-R in Clathrin-umhüllte Vesikel organisiert (Abbildung 5) (Garuti *et al.*, 2005). Im Gegensatz dazu ist die autosomal dominante Form der FH (ADH) eine der häufigsten genetisch determinierten Krankheiten überhaupt (1 von 500 Menschen) (Goldstein & Brown, 2001), wobei mittlerweile drei verschiedene Proteine identifiziert wurden, die für diese Form der FH verantwortlich sein können. Bis 2003 konnten ca. 65 % bzw. 10 % aller ADH-Fälle auf Mutationen im LDL-Rezeptor (LDL-R) (Goldstein & Brown, 1973) bzw. in der Ligand-Bindungsdomäne des

Apolipoproteins B100 (APOB) (Vega & Grundy, 1986) der LDL-Partikel zurückgeführt werden (Goldstein & Brown, 2001). Durch die von Abifadel *et al.* (2003) durchgeführte genetische Analyse von 23 ADH-anfälligen Familien, bei denen der Einfluss von LDL-R oder APOB ausgeschlossen wurde, konnte die Proprotein Convertase Subtilisin/Kexin Typ9 (PCSK9) als drittes ursächliche Protein für ADH identifiziert werden. PCSK9 ist eine Proprotein Convertase, die aus drei Domänen besteht (Pro-Domäne, katalytische Domäne, C-terminale V-Domäne), wobei die katalytische Domäne verwandt ist zur Proteinase K Subfamilie der Serin-Proteasen (Naureckiene *et al.*, 2003). PCSK9 wird hauptsächlich in der Leber, dem Darm, und in geringerem Maße in der Niere, der Haut und im Gehirn exprimiert (Seidah *et al.*, 2003). Hierbei erfolgt die Expression im endoplasmatischen Reticulum (ER) als 72 kDa großes Zymogen (Benjannet *et al.*, 2004). Erst nach autokatalytischer, intramolekularer Spaltung zwischen Pro- und katalytischer Domäne kann die Sekretion über den Golgi-Apparat erfolgen (Lagace *et al.*, 2006), wobei die 14 kDa Pro-Domäne und die restliche 63 kDa-Einheit weiterhin nicht-kovalent verbunden bleiben (Seidah & Prat, 2007; Cunningham *et al.*, 2007). Die Pro-Domäne wirkt als Chaperon und ist für die richtige Faltung der katalytischen Domäne notwendig (Ikemura *et al.*, 1987).

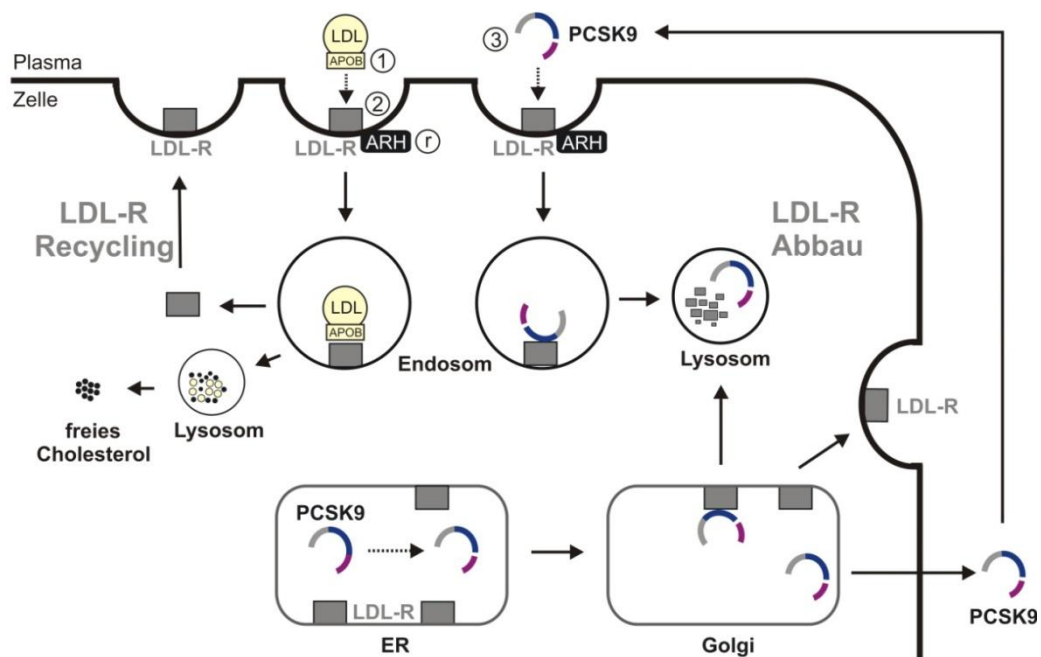


Abbildung 5: Zellulärer Signalweg von LDL, LDL-Rezeptoren und PCSK9 nach Lambert *et al.* (2008).

LDL-Partikel im Plasma binden über das Apoprotein B100 (APOB) an LDL-Rezeptoren (LDL-R) auf der Cytoplasmamembran, welche über ARH in Clathrin-umhüllten Vesikeln organisiert sind. Nach erfolgter Endozytose erfolgt der Abbau der LDL-Partikel zu freiem Cholesterol. Die LDL-Rezeptoren werden wieder an der Zelloberfläche exponiert (LDL-R Recycling). PCSK9 wird als Zymogen im ER exprimiert. Nach autokatalytischer Spaltung zu einem selbst-inhibiertem Enzym, bei dem die Pro-Domäne nicht-kovalent an die katalytische Domäne gebunden bleibt, erfolgt die Sekretion über den Golgi-Apparat. PCSK9 bindet sowohl intrazellulär im Golgi als auch als sekretierter Faktor an LDL-Rezeptoren und dirigiert sie zum lysosomalen Abbau. Familiäre Hypercholesterinämie kann durch Mutationen in ARH (r) (autosomal rezessive Form) bzw. in APOB (1), LDL-R (2) oder PCSK9 (3) (autosomal dominante Formen) hervorgerufen werden.

Das prozessierte PCSK9 bindet an LDL-Rezeptoren im Golgi (Park *et al.*, 2004) bzw. auf der Plasmamembran (Lagace *et al.*, 2006) und dirigiert sie zum lysosomalen Abbau (Abbildung 5). Der molekulare Mechanismus ist jedoch noch nicht im Detail verstanden.

PCSK9 wirkt also sowohl intrazellulär als auch als sekretierter Faktor (Lambert *et al.*, 2009), wobei die katalytische Aktivität von PCSK9 außer bei der Auto-Prozessierung keine weitere Rolle spielt (McNutt *et al.*, 2007). So hat eine inaktive 63 kDa-Einheit, koexprimiert mit der Pro-Domäne, den gleichen Einfluss auf den LDL-R - Abbau wie der Wildtyp (Li *et al.*, 2007). Die Bindung von PCSK9 an LDL-R erfolgt Calcium-abhängig über die katalytische Domäne an die EGF-A Domäne von LDL-R (Kwon *et al.*, 2008).

Bisher wurden drei Röntgenkristallstrukturen des humanen PCSK9 (hPCSK9) veröffentlicht, die alle die gespaltene und aus zwei nicht-kovalent verbundenen Ketten bestehende Form zeigen (Abbildung 6a) (Cunningham *et al.*, 2007; Hampton *et al.*, 2007; Piper *et al.*, 2007).

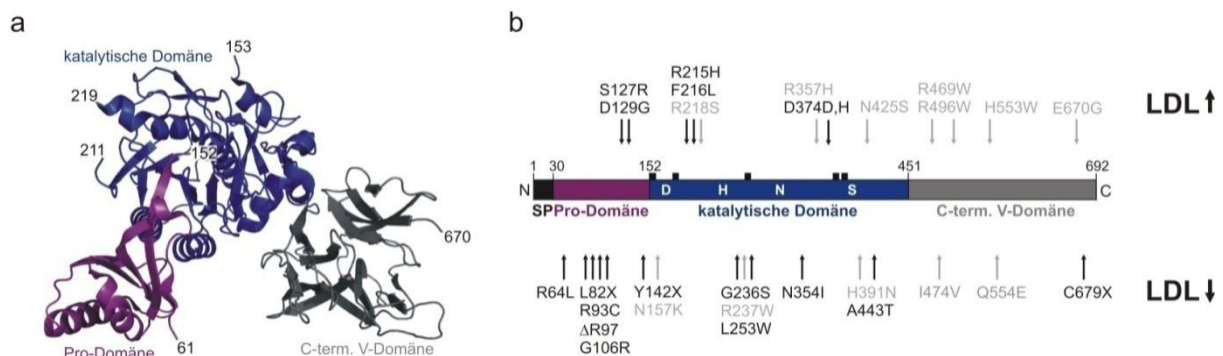


Abbildung 6: Aufbau von hPCSK9.

a hPCSK9 besteht aus einer Pro- (magenta), einer katalytischen (blau) und einer C-terminalen V-Domäne (grau). Autokatalytische Spaltung erfolgt zwischen Q152 und S153, wonach die Pro- und die katalytische Domäne nicht-kovalent verbunden bleiben. Einige Bereiche des Proteins konnten in der Röntgenkristallstruktur (pdb 2pmw) nicht aufgelöst werden (Δ 168-175, Δ 212-218, Δ 448-451, Δ 572-584, Δ 660-669). **b** Schematischer Aufbau von hPCSK9; SP: Signalpeptid; Die gekennzeichneten Mutationen in PCSK9 können zu einer erhöhten (\rightarrow LDL im Plasma steigt) oder einer erniedrigten (\rightarrow LDL im Plasma sinkt) Aktivität führen (Lambert *et al.*, 2008). Mutationen bei denen unklar ist ob sie allein für den beobachteten Phenotyp verantwortlich sein können sind grau dargestellt. Die katalytische Triade (D186, H226, S386) bzw. N317 (Ausbildung des Oxyanion-Lochs) sind weiß markiert. Die Positionen der Reste, welche H-Brücken zur EFG-A Domäne der LDL-Rezeptoren ausbilden (S153, R194, D238, T377, F379) sind als ■ gekennzeichnet.

Die vier C-terminalen Reste (149-152) der Pro-Domäne ragen hierbei in die katalytische Domäne und blockieren den Zugang zum katalytischen Zentrum, wodurch es zur Selbst-Inhibition der Protease-Aktivität kommt. Die C-terminale V-Domäne bildet einen neuen Faltungstyp, der aus drei symmetrischen, homologen Subdomänen (β -Sandwich aus 6 antiparallelen β -Strängen) besteht und durch insgesamt neun Disulfidbrücken stabilisiert wird.

Es wurden mittlerweile zahlreiche Mutationen in hPCSK9 identifiziert, die sowohl vermehrten (*missense* Mutationen → „*gain-of-function*“) als auch verringertem (*nonsense* Mutationen → „*loss-of-function*“) Abbau von LDL-R zur Folge haben können (Abbildung 6b) (Lambert *et al.*, 2008). Aufgrund der dadurch erniedrigten bzw. erhöhten Anzahl an LDL-Rezeptoren kommt es zu einem Anstieg bzw. Abfall des LDL-Spiegels im Plasma. Es konnte in beiden Fällen die Verknüpfung zwischen dem hPCSK9-Genotyp und ADH (Abifadel *et al.*, 2003; Allard *et al.*, 2005; Cameron *et al.*, 2008) bzw. einer erniedrigten Anfälligkeit für kardiovaskuläre Erkrankungen (Cohen *et al.*, 2005; Zhao *et al.*, 2006) festgestellt werden, wodurch die hohe klinische Relevanz des Proteins deutlich wird (Alborn *et al.*, 2007). Weitere *in vivo* Untersuchungen zeigten, dass die therapeutische Inhibition von PCSK9 eine neue Strategie zur Bekämpfung arteriosklerotischer Herz-Kreislauf-Erkrankungen darstellen kann (Park *et al.*, 2004; Graham *et al.*, 2007; Chan *et al.*, 2009).

1.7 Der humane Glucocorticoid-Rezeptor (hGR)

In der Medizin spielen Glucocorticoide eine sehr wichtige Rolle bei der gezielten Suppression der Immunantwort und bei der Behandlung von akuten und chronischen Entzündungen (Rhen & Cidlowski, 2005). Steroid-Hormone dieser Klasse und deren Derivate werden bereits seit dem 19. Jahrhundert untersucht und gehören heute zu den wichtigsten Medikamenten. So erhielten 1950 Edward Kendall, Tadeusz Reichstein und Philip Hench für ihre Studien über die Struktur und die physiologischen Effekte von Glucocorticoiden den Nobelpreis für Medizin. Die Wirkung dieser Steroid-Hormone begründet sich auf die Bindung und Aktivierung des Glucocorticoid-Rezeptors (GR). Dieser Hormon-Rezeptor kann direkt oder indirekt zahlreiche Zielgene regulieren, die am Glucose-Stoffwechsel, der Zell-Differenzierung, an inflammatorischen Prozessen und vielen anderen Signalwegen beteiligt sind (Reichardt *et al.*, 2000). So werden ca. 20 % aller exprimierten Gene in humanen Leukozyten positiv oder negativ durch Glucocorticoide reguliert (Galon *et al.*, 2002). Mutationen im GR werden mit dem *Cushing*-Syndrom und verschiedenen Krebs-Erkrankungen in Verbindung gebracht (Werner & Brönnegard, 1996). Daneben kann es zu gewebeabhängiger Glucocorticoid-Resistenz kommen, was chronische Entzündungs- oder Autoimmunerkrankungen wie z.B. rheumatoide Arthritis, Osteoarthritis, *Morbus Crohn*, *Colitis ulcerosa* oder Asthma zur Folge haben kann (Kino *et al.*, 2003). Die Transkriptions-Regulation durch den Glucocorticoid-Rezeptor kann über mehrere Mechanismen erfolgen, die in Abbildung 7 dargestellt sind.

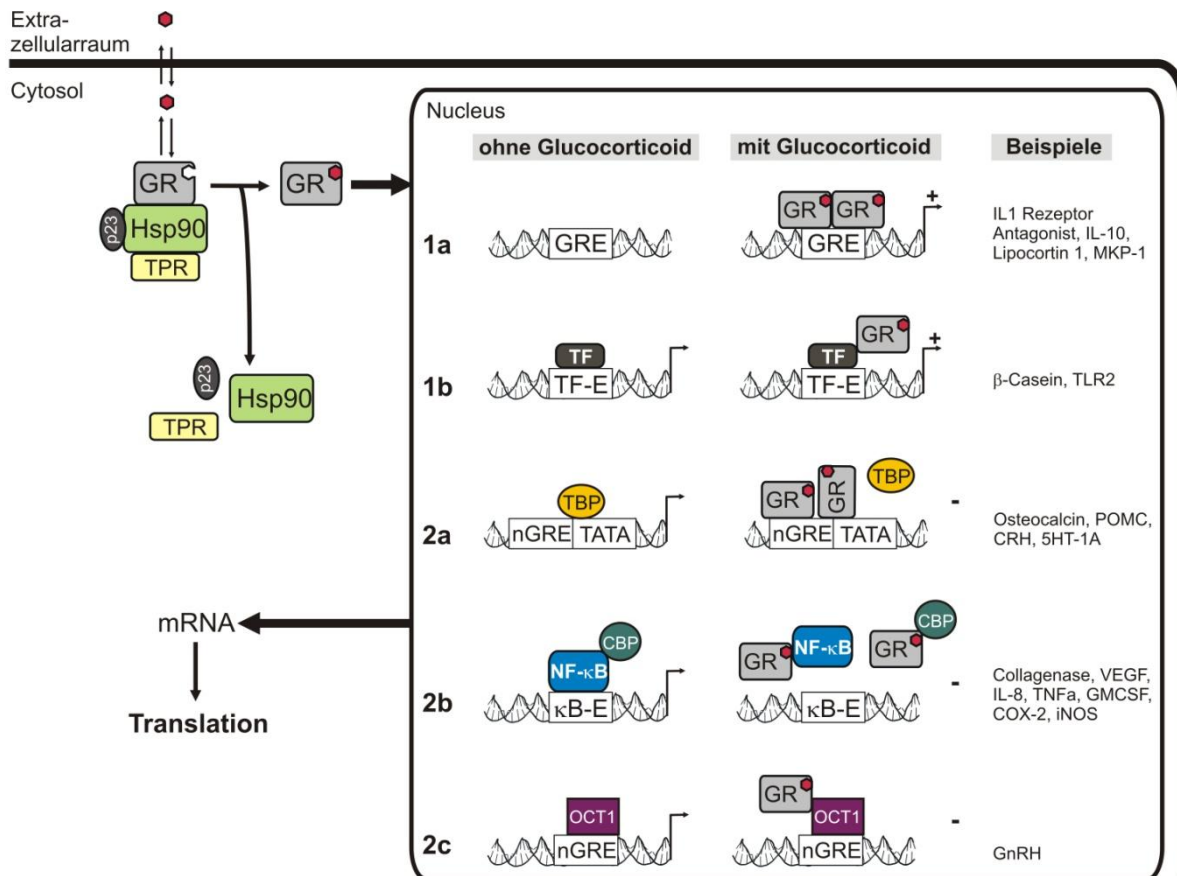


Abbildung 7: Mechanismen der Transkriptionsregulation durch den Glucocorticoid-Rezeptor nach Lu & Cidlowski (2006) und Heitzer *et al.* (2007).

GR liegt im Cytosol in einem Heterokomplex mit Hsp90, p23 und einem Tetratricopeptid Repeat Protein (TPR) vor. Nach Bindung eines Glucocorticoids (rot) dissoziiert GR aus diesem Komplex ab und kann in den Nucleus transloziert werden, wo es in unterschiedlicher Weise die Expression zahlreicher Gene positiv (1) oder negativ (2) beeinflusst. **1a** GR bindet als Homodimer an spezifische GRE (Glucocorticoid *Response* Elemente) - Sequenzen in der Promotor-Region von Zielgenen und kann deren Transkription nach Rekrutierung zahlreicher anderer Faktoren (Coaktivatoren, Chromatin *Remodelling* Komplexe, Teile der basalen Transkriptionsmaschinerie und Transkriptionsfaktoren) anschalten. **1b** GR kann auch GRE-unabhängig direkt an Transkriptionsfaktoren (TF) binden und mit diesen synergistisch die Transkription aktivieren. **2a** GR erkennt auch negative GREs (nGRE) und verhindert die Transkription durch Verdrängung anderer Faktoren, wie z.B. des TATA-Box bindenden Proteins (TBP). **2b** GR kann direkt an manche Transkriptionsfaktoren (z.B. NF- κ B), oder Coregulatoren (z.B. CREB-bindendes Protein CBP) binden und sie vom Promotor-Bereich absondern. **2c** Die Bindung von GR kann auch DNA-gebundene Transkriptionsfaktoren (z.B. OCT-1) inaktivieren. Abkürzungen: COX-2: Cyclooxygenase 2; CRH: Corticotropin-*Releasing* Hormon; GMCSF: Granulozyt Makrophagen *Colony Stimulating Factor*; GnRH: Gonadotropin-*Releasing* Hormon; iNOS: induzierbare Stickstoffmonoxid-Synthase; IL: Interleukine; κ B-E: κ B-Erkennungssequenz; MKP-1: MAP Kinase Phosphatase-1; OCT-1: Octamer Bindungs-Protein 1; POMC: Proopiomelanocortin; TF-E: Transkriptionsfaktor-Erkennungssequenz; TNF α : Tumor Nekrose Faktor α ; VEGF: Vaskulärer Endothel Wachstumsfaktor.

Einerseits kann der aktivierte Glucocorticoid-Rezeptor die Transkription beeinflussen, nachdem er über spezifische DNA-Bindestellen (Glucocorticoid *Response* Elemente = GRE) direkt an die DNA gebunden hat. Hierbei können positive (GRE) und negative (nGRE) *Response* Elemente unterschieden werden, welche die Transkriptions-Aktivierung bzw. -Hemmung durch GR induzieren. Andererseits kann der Glucocorticoid-Rezeptor über die Interaktion mit anderen Transkriptionsfaktoren ebenfalls Einfluss auf die Genregulation

nehmen, indem er deren Funktion unterstützt oder stört (Kumar & Thompson, 2005). So geht man davon aus, dass die entzündungshemmende Wirkung von Glucocorticoiden neben der GR-induzierten Expression von anti-inflammatorischen Proteinen (Phen & Cidlowski, 2005) vor allem durch die GR-vermittelte Repression von Nuklear-Faktor κB (NF- κB) und Aktivator-Protein 1 (AP-1) an Promotoren pro-inflammatorischer Gene hervorgerufen wird (Scheinman *et al.*, 1995; Smoak & Cidlowski, 2004; Phen & Cidlowski, 2005).

Der Glucocorticoid-Rezeptor gehört zur Superfamilie der Ligand-aktivierten Nuklearhormonrezeptoren (NHR). Wie alle Mitglieder dieser Superfamilie besteht GR aus drei Domänen: N-terminale-Domäne (NTD), DNA-bindende Domäne (DBD) und Ligand-bindende Domäne (LBD) (Heitzer *et al.*, 2007). In Abbildung 8 ist der Aufbau der am häufigsten vorkommenden Isoform des humanen GR (hGR) gezeigt (hGR α -A: 92 kDa). Diese kommt in nahezu allen Gewebetypen und Zellen des Menschen vor. Daneben werden allerdings auch noch zahlreiche andere Isoformen des Proteins durch alternatives *Splicing* des für hGR kodierenden Gens oder durch alternative Initiation der Translation gebildet. Diese unterscheiden sich in ihren posttranslationalen Modifikationen, der Art ihrer Transkriptions-Regulation und ihrer Häufigkeit in verschiedenen Gewebetypen (Lu & Cidlowski, 2006).

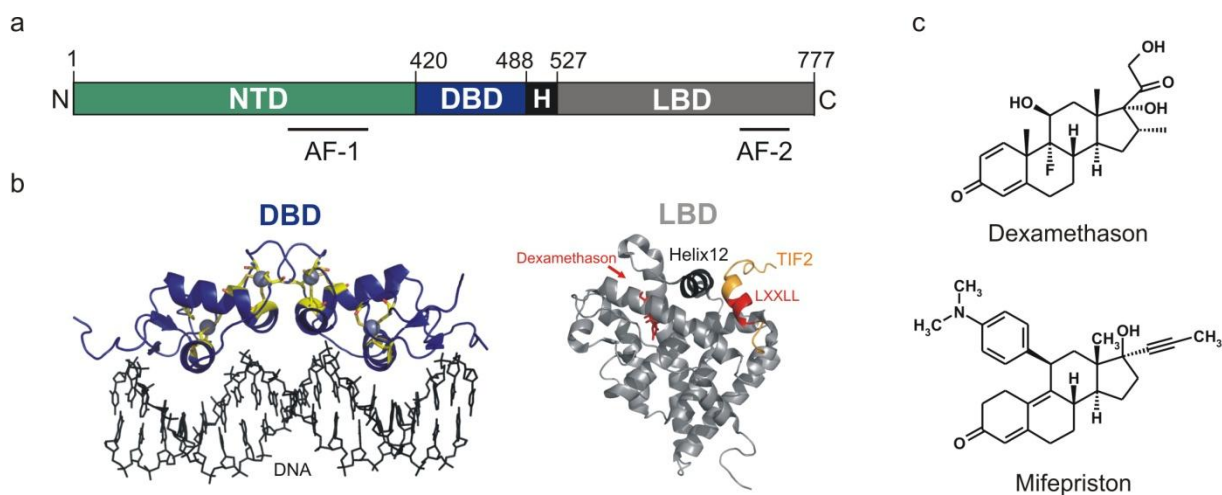


Abbildung 8: Aufbau der häufigsten Isoform des humanen Glucocorticoid-Rezeptors (hGR α -A).

a Schematische Darstellung der Domänen-Organisation von hGR α -A. NTD: N-terminale Domäne, DBD: DNA-bindende Domäne, H: Hinge-Region, LBD: Ligand-bindende Domäne. In NTD und LBD sind Regionen vorhanden (AF-1, AF-2), die für die Interaktion mit anderen Faktoren der Transkriptionsregulation verantwortlich sind. **b** Röntgenkristallstrukturen des DBD-Dimers (blau), ko-kristallisiert mit DNA (Cysteine und gebundene Zn^{2+} -Ionen der Zinkfinger-Motive sind gelb bzw. als graue Kugeln dargestellt) (pdb 1glu) und der LBD, ko-kristallisiert mit Dexamethason und dem Coaktivator-Peptid TIF2, welches über ein LXXLL-Motiv an Helix 12 (= AF-2) bindet (pdb 1m2z). **c** Strukturformeln des hGR-Agonisten Dexamethason bzw. des hGR-Antagonisten Mifepriston.

Die NTD von hGR α -A enthält einen stark negativ geladenen Bereich („AF-1“ = *activation function-1* oder „*tau-1*“), der für die Interaktion mit der basalen Transkriptionsmaschinerie und vielen anderen Transkriptionsfaktoren verantwortlich ist (McEwan *et al.*, 1993) und die wichtigste Aktivierungsregion von hGR darstellt (Giguere *et al.*, 1986; Dieken & Miesfeld, 1992). Die 3D-Struktur der NTD konnte bis jetzt noch nicht gelöst werden, da AF-1 intrinsisch unstrukturiert ist (Dahlman-Wright *et al.*, 1995). Man geht davon aus, dass AF-1 eine potentiell helikale Struktur besitzt (Dahlman-Wright *et al.*, 1995), die erst dann ausgebildet wird, sobald die Bindung an andere Proteine der Transkriptionsmaschinerie erfolgt, was zur spezifischen Transkriptionsregulation ausgenutzt werden kann (Almlöf *et al.*, 1998; Wärnmark *et al.*, 2000; Kumar *et al.*, 2004). Darüber hinaus scheint die korrekte Faltung von AF-1 ebenfalls nach Bindung von hGR an DNA induziert zu werden, was auch zur Regulation der hGR-Aktivität benutzt werden könnte (Kumar *et al.*, 1999). Die DNA-bindende Domäne von hGR enthält acht Cysteine, die zwei Zink-Fingermotive ausbilden, wobei der aminoterminaler Zinkfinger verschiedene DNA-*Response* Elemente unterscheiden kann und der carboxyl-terminale Zinkfinger eine wichtige Rolle bei der Dimerisierung von hGR spielt (Hard *et al.*, 1990; Luisi *et al.*, 1991). Die DBD ist mit der Ligand-bindenden Domäne über eine kurze Hinge-Domäne verbunden, die dem Rezeptor erlaubt unterschiedliche Konformationen bezüglich der Domänen-Positionierung anzunehmen (Tsai & O'Malley, 1994).

Die hochkonservierte C-terminale, Ligand-bindende Domäne (LBD) erfüllt neben der spezifischen Glucocorticoid-Bindung auch noch andere Aufgaben. So binden Hsp90 und verschiedene Co-Chaperone an die Apo-Form der LBD im Cytosol, wodurch einerseits die korrekte Faltung von hGR sichergestellt und andererseits eine Kernimport-Signalsequenz der LBD maskiert wird. Erst nach Hormon-Bindung, dem die Abdissoziation des Hsp90-Komplexes folgt, kann es somit zum Transport von hGR in den Zellkern kommen (Abbildung 7) (Savory *et al.*, 1999). Desweiteren kann die LBD über α -Helix 12 („AF-2“ = *activation function 2*) hormonabhängig mit zusätzlichen Faktoren der Transkriptionsregulation interagieren, wobei AF-1 synergistisch wirken kann (Beato & Klug, 2000). Nach Bindung unterschiedlicher Glucocorticoide, wie z.B. Dexamethason (Agonist) oder Mifepriston (RU-486) (Antagonist) (Abbildung 8c) wird die Konformation von AF-2 so verändert, dass sich spezifische Faktoren, welche die Transkription aktivieren (Coaktivatoren) oder hemmen (Corepressoren), anlagern können (Bledsoe *et al.*, 2002; Kauppi *et al.*, 2003). Die meisten Coaktivatoren können mit Hilfe sogenannter NR-Boxen, die ein konserviertes LXXLL-Motiv enthalten, mit mehreren nuklearen Hormonrezeptoren interagieren (Heery *et al.*, 1997). Jeder

Coaktivator enthält mehrere NR-Boxen und die unterschiedliche Affinität zu den verschiedenen NHRs wird zur Regulation der spezifischen Hormonantwort ausgenutzt (Heitzer *et al.*, 2007). Mittlerweile sind ca. 200 NHR-Coaktivatoren bekannt, welche die Steroidhormonrezeptor-abhängige Expression von Zielgenen durch eine Vielzahl an Mechanismen unterstützen (Lonard & O'Malley, 2006). So können sie zum Beispiel die Transkriptionsaktivierung verstärken, indem sie als Histon-Acetyltransferasen oder Methyltransferasen wirken oder indem sie den Rezeptor mit dem Transkriptionsinitiationskomplex, anderen Transkriptionsfaktoren oder weiteren Histon-modifizierenden Enzymen verbinden (Heitzer *et al.*, 2007). Corepressoren können ebenfalls über LXXLL-Motive an die LBD (Loinder & Söderström, 2004) binden und sind in der Lage die Transkriptionsaktivierung zu hemmen, indem sie Histon-Deacetylasen rekrutieren, deren Aktivität zu einer geschlossenen Chromatin-Struktur führt (Jones & Shi, 2003). Bisher sind einige Röntgenkristallstrukturen der hGR-LBD gelöst worden, welche die isolierte Domäne mit gebundenem Agonist (Dexamethason, Fluticason- Furoat, Amino-pyrazole-2,6-dichloro-N-ethylbenzamide1 oder Deacylcortivazol) und gebundenem Coaktivator-Peptid (TIF-2 oder SRC-1) (Bledsoe *et al.*, 2002; Biggadike *et al.*, 2008; Madauss *et al.*, 2008; Suino-Powell *et al.*, 2008) zeigen. Bis heute wird nur in einer Veröffentlichung die 3D-Struktur der hGR-LBD mit gebundenem Antagonist (Mifepriston) beschrieben, wobei allerdings verschiedene Kristallformen gezeigt werden, bei denen die wichtige Helix 12 unterschiedlich positioniert ist (Kauppi *et al.*, 2003).

1.8 Zielsetzung der Arbeit

Im Rahmen der vorliegenden Arbeit sollte die Stabilität und Löslichkeit mehrerer Proteine verbessert werden, um sie nach Expression in *E. coli* mit höherer Ausbeute reinigen zu können und ihre Kristallisationstendenz zu steigern.

In diesem Zusammenhang sollte zunächst versucht werden, die Hydrophilie der Oberfläche von hPCSK9 durch rationales Design gezielt zu erhöhen.

Im Mittelpunkt der Arbeit stand jedoch gelenkte Evolution zur Optimierung von hGR-LBD, wobei eGFP als Faltungsreporter und Hochdurchsatz-FACS als *Screening*-System in der Arbeitsgruppe etabliert werden sollten. Die Ergebnisse sollten zeigen, inwieweit sich dieses in der Literatur beschriebene System für schwierig zu handhabende menschliche Proteine eignet und durch welche Maßnahmen auftretende Probleme erkannt und umgangen werden können.

Daneben sollte die mit Hilfe der Chloramphenicol-Acetyltransferase als Reporterprotein erfolgte Stabilisierung eines künstlich hergestellten $(\beta\alpha)_8$ -Barrel Proteins weitergeführt

werden (Seitz, 2006). Dazu sollten vorteilhafte Mutationen kombiniert werden, um die Kristallisation und die Röntgenstrukturanalyse des Proteins zu ermöglichen. Die durchgeführten Experimente und Ergebnisse dieses Teilprojekts der Dissertation sind in zwei Publikationen dokumentiert (Seitz *et al.*, 2007; Höcker *et al.*, 2009), welche im Anhang (7.8) angefügt sind.

2 Material

2.1 Geräte

Abbé Refraktometer	Carl Zeiss, Jena
Autoklaven	
München 25	MMM, München
VE-40	SYSTEC GmbH, Wettenberg
Brutschränke	BINDER GmbH, Tuttlingen
Chromatographieanlagen:	
ÄKTA™ basic better	GE HEALTHCARE, München
ÄKTA™ purifier 10	GE HEALTHCARE, München
FPLC LCC-501 plus	GE HEALTHCARE, München
HPLC Agilent 1200	Agilent Technologies
verwendete Säulen:	
HisTrap HP 5 ml	GE HEALTHCARE, München
HisTrap FF Crude 1 ml	GE HEALTHCARE, München
GSTrap FF 5 ml	GE HEALTHCARE, München
Superdex 75 HR 10/30	GE HEALTHCARE, München
Superdex 200 pg HiLoad 16/60	GE HEALTHCARE, München
Superdex 200 pg HiLoad 26/60	GE HEALTHCARE, München
Poroshell C8 (2.1x50 mm)	Agilent Technologies
Cell Density Meter Ultrospec 10	GE HEALTHCARE, München
Circular Dichroismus Spektropolarimeter J-815	JASCO GmbH, Groß-Umstadt
Computer	Dell Optiplex Systems
Elektroporator 2510	EPPENDORF, HAMBURG
<i>Fluorescence Activated Cell Sorter (FACS)</i>	DAKO CYTOMATION, Fort Collins, USA
MoFlo® Zytometer mit Cyclone Sortiereinheit und Coherent EnterpriseII Laser	
Fluoreszenzmikroskop Axiovert 200M	CARL ZEISS AG, Jena
Fluoreszenz-Spektrometer Cary Eclipse	VARIAN, Darmstadt
Gasbrenner Gasprofi 2SCS	WLD Tec GmbH, Göttingen
Gefrierschrank -80 °C	MDF-U72V, SANYO, Tokyo, Japan
Gefrierschrank -20 °C	LIEBHERR, Nussbaumen
Gelelektrophoresekammern:	
Agarose-Elektrophorese-Apparatur	HOEFER Pharmacia Biotech, USA

SDS-Elektrophoresekammer	
Mighty Small II	HOEFER Pharmacia Biotech, USA
Multi Gel Caster Gelgießstand	GE HEALTHCARE, München
Glaspipetten & Glaswaren	SCHOTT, Mainz
	FISCHER SCIENTIFIC, Schwerte
	NOVOGLAS, Bern, Schweiz
Hamilton Mikroliterspritze, 25 µl	HAMILTON, Banskuduz, Schweiz
Heizblock-Thermostat HBT-2 131	HLC, Bovenden
Inkubationsschüttler:	
GFL 3033	GFL, Burgwedel
Certomat H	BRAUN Biotech, Melsungen
Certomat BS-1	BRAUN Biotech, Melsungen
Kühlschrank Privileg de luxe (8 °C)	PRIVILEG, Fürth
Magnetrührer:	
MR0, MR2000	HEIDOLPH, Kehlheim
MR1, MR3001 (beheizbar)	HEIDOLPH, Kehlheim
Microliterpipetten: Research 10, 100, 1000	EPPENDORF, Hamburg
Mikrowelle HMT 842C	BOSCH, Nürnberg
Multi-Doc-It Digital Imaging System	UVP Inc., USA
Netzgeräte:	
Power Supply EPS 301	GE HEALTHCARE, München
Power Pack P25	BIOMETRA, Göttingen
PCR-Geräte:	
Mastercycler personal	EPPENDORF, Hamburg
Mastercycler gradients	EPPENDORF, Hamburg
Peristaltikpumpe Miniplus 2	GILSON Medical Electronics, Frankreich
pH-Meter inolab Level1	INOLAB, Weilheim
Pipettus-accu	HIRSCHMANN, Laborgeräte, Eberstadt
Potter-Homogenisator	B. Braun, Melsungen AG
Quarzküvetten:	
101-QS (Schichtdicke 10 mm)	HELLMA GmbH & Co. KG, Müllheim
105-QS (Schichtdicke 10 mm)	HELLMA GmbH & Co. KG, Müllheim
21/Q/1/CD (Schichtdicke 1 mm)	Starna GmbH, Pfungstadt
Schütteltisch Rocking Platform	BIOMETRA, Göttingen

Überkopffrotierer Eurostar digital	IKA Labortechnik AG
Ultrafree-20 Nanopore-Wasser-Anlage	MILLIPORE, Eschborn
Ultraschallgerät Branson Sonifier 250D	HEINEMANN, Schwäbisch Gmünd
UV-Leuchttisch Reprostar	CAMAG Chemie Erzeugnisse, Berlin
UV-Vis Spektralphotometer Cary 100 Bio	VARIAN, Darmstadt
UV-Vis Spektralphotometer V650	JASCO GmbH, Groß-Umstadt
UV-Vis Biophotometer	EPPENDORF, Hamburg
Vakuumpumpe ME 2C	VACUUMBRAND, Wertheim
Vortex Genie 2	SCIENTIFIC IND., Bohemia, USA
Waagen:	
MC1	SARTORIUS, Göttingen
SI-114	DENVER INSTRUMENT, Göttingen
Wasserbad 1003	GFL, Burgwedel
WetBlot <i>Western-Blot</i> Apparatur	Peqlab, Erlangen
PerfectBlue Tank-Elektroblotter	
Zentrifugen:	
Centrifuge 5810R	EPPENDORF, Hamburg
Centrifuge 5415D	EPPENDORF, Hamburg
Centrifuge 5415R	EPPENDORF, Hamburg
Sorvall RC 2B, 5C plus	DU PONT Instruments, Bad Homburg
mit SS34- bzw. GS3-Rotoren	

2.2 Verbrauchsmaterialien

AlignFlow, Plus flow cytometry alignment beads for 488 nm excitation, 6 µm	INVITROGEN Molecular Probes, Carlsbad, USA
Dialyseschlauch: Visking, Typ 27/32, 14 kDa	ROTH GmbH +Co, Karlsruhe
Einwegspritzen Omnifix® 2 ml, 5 ml und 60 ml	BRAUN Biotech, Melsungen
Elektroporationsküvetten 2 mm	PEQLAB, Erlangen
Faltenfilter	MACHEREY-NAGEL, Düren
Filterpapier	Whatman, Maidstone, England
Konzentratoren:	
Amicon Ultra 4 bzw. 15 (mwco: 10 kDa)	Millipore, Bedford, USA
Kunststoffküvetten: ½ Mikroküvetten	Sarstedt, Nümbrecht
Kunststoffröhrchen:	
15 ml, 50 ml	SARSTEDT, Nümbrecht

Falcon 15 ml bzw. 50 ml	Becton Dickinson, Franklin Lakes, USA
Latex-Handschuhe Sempermed	SEMPERIT Technische Produkte GmbH, Wien, Österreich
Membranfilter ME24 Ø47 mm; 0,2 µm	SCHLEICHER&SCHUELL, Dassel
Nitrocellulosefilter (Ø13 mm; Cat#: VSWP01300)	MILLIPORE, Eschborn
NAP5, NAP10 und NAP25 Säulen	GE HEALTHCARE, München
Nickel-Sepharose	GE HEALTHCARE, München
Nitril Untersuchungshandschuhe	VWR, Leuven, Belgien
Parafilm „M“ Laboratory Film	PECHINEY, Plastic Packaging, Menasha, USA
Pasteurpipetten	HIRSCHMANN, Ebermannstadt
PCR-Gefäße 0,2 ml	PEQLAB, Erlangen
Petrischalen 150/25 und 94/16	GREINER bio-one, Nürtingen
Pipettenspitzen 10 µl, 200 µl, 1000 µl	SARSTEDT, Nümbrecht
PVDF Membran Immobilon –P	Millipore, Bedford, USA
Reaktionsgefäße 1,5 ml und 2 ml	ROTH, Karlsruhe
Reaktionsgefäße mit Schraubdeckel 2 ml	SARSTEDT, Nümbrecht
Röntgenfilme Fuji Super RX	FujiFilm, Düsseldorf
Spritzenfilter Porengröße 0,22 µm, 0,45 µm	RENNER GmbH, Daunstadt

2.3 Chemikalien

Die verwendeten Chemikalien besaßen den höchstmöglichen Reinheitsgrad und wurden von folgenden Firmen bezogen:

AMERSHAM PHARMACIA Biotech	(Freiburg)
BIO101 Inc.	(Carlsbad, USA)
BIORAD LABORATORIES	(München)
BIOZYM	(Hess. Oldendorf)
BOEHRINGER MANNHEIM	(Mannheim)
DIFCO	(Dreieich)
FLUKA	(Neu-Ulm)
GERBU Biotechnik GmbH	(Gailberg)
GIBCO/BRL	(Eggestein)
MERCK	(Darmstadt)

LUCERNA-CHEM AG	(Luzern, Schweiz)
NATIONAL DIAGNOSTICS	(Simerville, NJ, USA)
OXOID	(Wesel)
RIEDEL-DE HAEN	(Seelze)
ROCHE DIAGNOSTICS	(Mannheim)
ROTH	(Karlsruhe)
SERVA	(Heidelberg)
SIGMA-ALDRICH	(Deisenhofen)

2.4 Kits

Biorad-Lösung für Bradford Test	Biorad Laboratories GmbH, München
Gel Filtration Calibration Kit	GE Healthcare, München
GeneJET™ Plasmid Miniprep Kit	MBI FERMENTAS, St.-Leon-Rot
NucleoBond®PC100/500 <i>DNA-Extraction</i> Kit	MACHEREY-NAGEL, Düren
NucleoSpin® Extract II	MACHEREY-NAGEL, Düren
SuperSignal West Pico Kit	Pierce, USA
X-Cell <i>SureLock</i> ™ Mini Cell	Invitrogen, Carlsbad USA
NuPAGE® Novex Bis-Tris Gele (4-12 %)	
Instant Blue SDS-Gel Färbelösung	Expedeon, Lucerna-Chem AG

2.5 Enzyme

DNase I	ROCHE DIAGNOSTICS, Mannheim
Restriktionsendonukleasen	NEW ENGLAND BIOLABS, Bad Schwalbach
T4-DNA-Ligase	ROCHE DIAGNOSTICS, Mannheim
	MBI FERMENTAS, St. Leon-Rot
DNA-Polymerasen:	
<i>Pwo</i> -DNA-Polymerase	ROCHE DIAGNOSTICS, Mannheim
GoTaq® DNA-Polymerase	PROMEGA, Mannheim
Thrombin (from bovine plasma)	SIGMA-ALDRICH, Deisenhofen

2.6 Antikörper

Anti-GFP Antikörper <i>mouse</i> monoklonal 11E5 (A11121)	MolecularProbes, Invitrogen
Peroxidase-konjugierter anti- <i>mouse</i> IgG NA931VS	GE Healthcare, München

2.7 Bakterienstämme

E. coli DH5 α (Hanahan, 1983)

F', *endA1*, *hsdR17* ($r_k^- m_k^-$), *supE44*, *thi1*, *recA1*, *gyrA* (Nal^r), *relA1*, $\Delta(lacZYA-argF)$ U169, $\Phi 80$, *lacZ* Δ M15

E. coli 5 α (NEW ENGLAND BIOLABS, Frankfurt am Main)

fhuA2, $\Delta(argF-lacZ)$ U169, *phoA*, *glnV44*, $\Phi 80$, $\Delta(lacZ)$ M15, *gyrA96*, *recA1*, *relA1*, *endA1*, *thi-1*, *hsdR17*

Bei diesen Zellen handelt es sich um ein *E. coli* DH5 α Derivat. Durch die Deletion des Gens *fhuA2* fehlt der Eisen-Hydroxamat-Aufnahme-Rezeptor auf der Zelloberfläche, was verhindert, dass der Bacteriophage T1, der an diesen Rezeptor bindet, die Zellen befällt. Somit sind diese Zellen resistent gegen T1-Phagen.

E. coli XL1 Blue MRF' (Stratagene)

$\Delta(mcrA)$ 183 $\Delta(mcrCB-hsdSMR-mrr)$ 173 *endA1* *supE44* *thi-1* *recA1* *gyrA96* *relA1* *lac* [F' *proAB lacIqZ* Δ M15 Tn10 (Tet^r)]

In XL1 Blue MRF' Zellen sind alle in *E. coli* K12 bekannten Restriktionsendonucleasen auf Genomebene inaktiviert.

E. coli Turbo (NEW ENGLAND BIOLABS, Frankfurt am Main)

F', *proA*⁺*B*⁺, *lacI*^q, $\Delta(lacZ)$ M15/*fhuA2*, $\Delta(lac-proAB)$, *glnV*, *gal* R(*zgb-210::Tn10*), Tet^s, *endA1*, *thi-1*, $\Delta(hsdS-mcrB)$ 5

E. coli Turbo Zellen sind wie *E. coli* 5 α -Zellen T1-phagenresistent. Diese Resistenz ist mit einem langsameren Zellwachstum verbunden. Dieser Nachteil wird bei *E. coli* Turbo Zellen kompensiert, wodurch sie innerhalb von 8 h bei Inkubation bei 37 °C sichtbare Kolonien bilden können.

E. coli BL21 (DE3) (Studier & Moffat, 1986)

hsdS, *gal*, [λ cI, *ts857*, *cmd1*, *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi1*, *relA1*]

Diese Zellen tragen auf ihrem Chromosom ein Gen für die T7 RNA-Polymerase, die zur Genexpression im pET-System verwendet wird (Abbildung 9).

***E. coli* T7 Express** (NEW ENGLAND BIOLABS, Frankfurt am Main)

fhuA2 lacZ::T7gene1 lon ompT gal sulA11 dcm R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr)114::IS10 R(mcr-73::miniTn10--Tet^S)2

Bei diesen Zellen handelt es sich um ein *E. coli* BL21(DE3) Derivat. Durch die Deletion des Gens *fhuA2* fehlt der Eisen-Hydroxamat-Aufnahme-Rezeptor, wodurch diese Zellen resistent gegen den Bacteriophagen T1 sind. Zusätzlich enthalten die Zellen eine Tetrazyklin-Resistenz.

2.8 Vektoren**2.8.1 pET-Vektoren**

Gene, die in die multiple Klonierungsstelle (MCS) von pET-Vektoren (NOVAGEN, *plasmid for expression by T7 RNA Polymerase*) eingefügt wurden, werden durch die RNA-Polymerase des Phagen T7 (Studier *et al.*, 1990) transkribiert. Die Expression der Gene erfolgt in speziellen *E. coli*-Stämmen, die eine chromosomale Kopie der T7 RNA-Polymerase besitzen (Abbildung 9). Die Expression des T7 RNA-Polymerasegens erfolgt unter der Kontrolle eines *lacUV5*-Promotor-Operators und wird durch die Zugabe von Isopropyl-β-D-Thiogalaktopyranosid (IPTG) induziert. Das zur Unterdrückung der Expression erforderliche Gen für den *lac*-Repressor (*lacI*) liegt auf dem Plasmid und wird konstitutiv exprimiert.

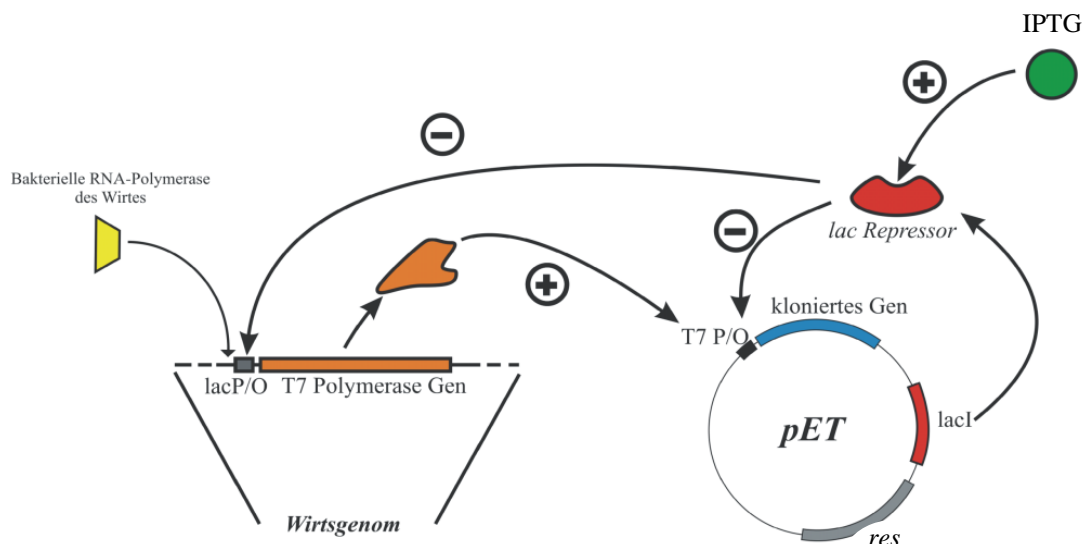


Abbildung 9: Schematische Darstellung des pET-Expressionssystems.

P/O: Promotor und Operator

+/- : Aktivierung /Repression von Expression

res: Antibiotikum-Resistenz vermittelndes Gen

2.8.1.1 pET21a(+)

Der pET21a(+)-Vektor ist so konstruiert, dass unmittelbar stromabwärts der multiplen Klonierungsstelle sechs Histidin-Codone liegen (Abbildung 10). Das hergestellte Protein trägt dadurch C-terminal einen Hexahistidin (His₆) - Tag, der eine Aufreinigung durch Metallchelat-Affinitätschromatographie ermöglicht. Nach der Transformation mit dem pET21a(+)-Vektor wird mittels der enthaltenen Ampicillin-Resistenz (*bla*) auf plasmidenthaltende *E.coli*-Zellen selektiert.

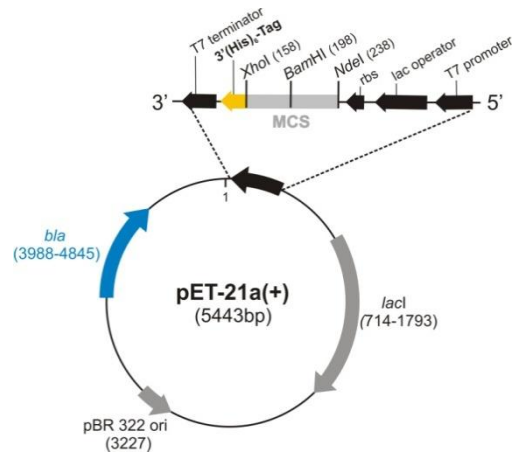


Abbildung 10: pET-21a(+) mit multipler Klonierungsstelle (MCS).

2.8.1.2 pER Vektoren

pER-Vektoren basieren auf dem pET24a(+)-Vektor, der bis auf ein unterschiedliches Antibiotika-Resistenzgen (*kan* statt *bla*) identisch zum pET21a(+)-Vektor ist (Abbildung 11). Die pER-Vektoren unterscheiden sich allerdings von pET24a(+) und untereinander durch ihre multiple Klonierungsstelle, wobei verschiedene Fusionskonstrukte kodiert werden, sobald ein Gen zwischen die Restriktionsschnittstellen *NdeI* und *NotI* eingefügt wird. Die pER-Vektoren pER13a, pER13b, pERN7 und pER14 wurden von Dr. Ralf Thoma (Hoffmann-La Roche, Pharma Research Discovery, Basel) konstruiert und zur Verfügung gestellt. pER13cat wurde im Rahmen dieser Arbeit erstellt. Bei pER13a, pER13b, pER14 und pER13cat wurde die multiple Klonierungsstelle auf zwei Schnittstellen für die Restriktionsenzyme *NdeI* und *NotI* reduziert. Zwischen den beiden Schnittstellen wurden Stoppcodone eingebracht, die jedes der drei möglichen Leseraster abdecken. Zusätzlich kommt es aufgrund einer Deletion zwischen den beiden Restriktionsschnittstellen zu einer Leserasterverschiebung. Diese Deletion wird beim Einfügen des jeweiligen Zielgens durch eine zusätzliche Base zwischen dem Ende des Gens und der *NotI*-Schnittstelle kompensiert (im Rahmen dieser Arbeit wurde am 3'-Ende der verwendeten Gene ein Guanin eingefügt). Dadurch erfolgt eine Synchronisation des Leserasters des Zielgens mit dem Reporter-gen, das auf die *NotI*-Schnittstelle folgt. Durch die Stoppcodone und die eingefügte Deletion soll bei pER13a, pER13b und pER13cat die Expression des stromabwärts fusionierten Reporter-Gens ohne eingefügtes Zielgen verhindert werden. Bei pERN7 liegt eine andere multiple Klonierungsstelle ohne diese Eigenschaften vor (Abbildung 11c).

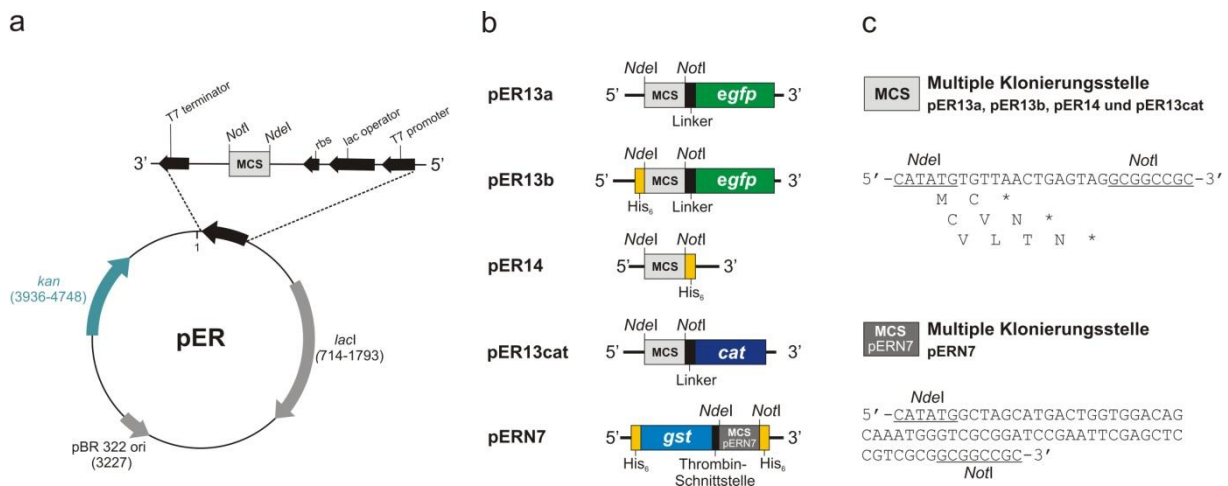


Abbildung 11: pER-Vektoren.

a Allgemeiner Aufbau der pER-Vektoren (pET24a(+)-Derivate).

b 5'- und 3'-Umgebung der multiplen Klonierungsstelle bei den verschiedenen pER-Vektoren.

c Die multiple Klonierungsstelle von pER13a, pER13b, pER14 und pER13cat beinhaltet neben einer Verschiebung des Leserasters zusätzlich in jedem Leseraster ein Stoppcodon. Die kodierte Aminosäuresequenz jedes Leserasters beginnend beim ATG der *NdeI*-Schnittstelle ist bis zum jeweiligen Stoppcodon unterhalb der DNA-Sequenz gezeigt. Die multiple Klonierungsstelle von pERN7 weist diese Eigenschaften nicht auf.

2.8.1.2.1 pER13a und pER13b

Im pER13a-Vektor (6003 bp) folgt auf die *NotI*-Schnittstelle eine für dreizehn Aminosäuren kodierende Linkerregion (GGRAGSAAGSGEF, nach Waldo *et al.*, 1999) und anschließend das Gen für eine eGFP (*enhanced green fluorescent protein*) – Variante (Abbildung 11b). Das verwendete eGFP unterscheidet sich von GFP aus der Qualle *Aequoria victoria* durch einen neutralen Aminosäureaustausch (Q80R, Chalfie *et al.*, 1994), die „Cycle-3“- (F99S, M153T, V163A; Cramer *et al.*, 1996) und die „eGFP“-Mutationen (F64L, S65T; Cormack *et al.*, 1996), welche die Stabilität und die Fluoreszenz-Intensität erhöhen (siehe 1.4). Zudem wird die Fluoreszenz von eGFP nicht wie die von GFP bei 396 nm maximal angeregt, sondern bei 488 nm, was der Wellenlänge des emittierten Lichts des Argonlasers des verwendeten MoFlo[®] Zytometers entspricht. Der pER13b-Vektor (6062 bp) entspricht im Aufbau und in der Funktion dem pER13a-Vektor mit dem Unterschied, dass vor der *NdeI*-Schnittstelle ein *His₆*-Tag eingefügt wurde.

2.8.1.2.2 pER14

Der pER14-Vektor ist in gleicher Weise konstruiert wie der pER13a-Vektor. An die *NotI*-Schnittstelle schließt sich jedoch ohne Linker direkt die kodierende Sequenz für einen *His₆*-Tag an (Abbildung 11b). Dieser Vektor dient dazu, Gene aus dem pER13a-, pER13b- oder

pER13cat-Vektor über einen *NdeI/NotI*-Restriktionsverdau direkt umklonieren zu können und damit vom Fusionspartner zu trennen. Anschließend kann die Herstellung des Genprodukts ohne Fusionsprotein und die Reinigung über einen C-terminalen *His₆-Tag* erfolgen.

2.8.1.2.3 pER13cat

Der pER13cat-Vektor ist ebenfalls in gleicher Weise konstruiert wie der pER13a-Vektor. Nach der *NotI*-Schnittstelle folgt jedoch eine für acht Aminosäuren kodierende Linkerregion (GGRLEAAR, nach Maxwell *et al.*, 1999) und anschließend das Gen für die Chloramphenicol-Acetyltransferase (*cat*) (Abbildung 11b). Dieser Vektor dient dazu, Gene aus pER13a über einen *NdeI/NotI*-Restriktionsverdau direkt so umzuklonieren, dass sie an *cat* fusioniert vorliegen.

2.8.1.2.4 pERN7

Der pERN7-Vektor unterscheidet sich in seiner multiplen Klonierungstelle von den anderen pER-Vektoren (Abbildung 11b). Diese beinhaltet keine Verschiebung des Leserasters, da dieser Vektor nicht für Selektions- oder *Screening*-Methoden verwendet wird. Das über die Restriktionsschnittstellen *NdeI* und *NotI* eingefügte Gen wird 3'-terminal an die kodierende Sequenz für einen *His₆-Tag* und die Glutathion-S-Transferase (*His₆-GST*) fusioniert. Bei allen in dieser Arbeit klonierten pERN7-Konstrukten wurde am 3'-Ende des Zielgens ein Stopp-Codon eingeführt, wodurch das durch pERN7 hergestellte Fusionsprotein keinen zusätzlichen C-terminalen *His₆-Tag* aufweist. Das *His₆-GST*-Fusionsprotein bindet sowohl an Glutathion-Sepharose- als auch Metallchelate-Säulen mit hoher Affinität. Zwischen *His₆-GST* und dem Zielprotein befindet sich die Erkennungssequenz für die spezifische Protease Thrombin, wodurch *His₆-GST* nachträglich vom Zielprotein abgespalten werden kann.

2.8.2 Konstitutive Expressionsvektoren

2.8.2.1 pTNA

Der pTNA-Vektor (Abbildung 12) ist eine Variation des pDS56/RBSII-Vektors (Bujard *et al.*, 1987; Stüber *et al.*, 1990). Die T5 Promotor und *lac*-operator Sequenzen wurden gegen die (-30)-(-10)-Region des Promotors des Tryptophanase-Operons (P_{TNA}) von *E. coli* ausgetauscht, welcher zu einer schwachen konstitutiven Expression führt (Merz *et al.*, 2000). Die Klonierung erfolgt in der Regel über die *Sph*I- und *Hind*III-Schnittstellen. Der pTNA-Vektor vermittelt Ampicillin-Resistenz aufgrund eines funktionellen *bla*-Gens. Der Vektor wurde von der Arbeitsgruppe Yanofsky (Stanford University, USA) zur Verfügung gestellt.

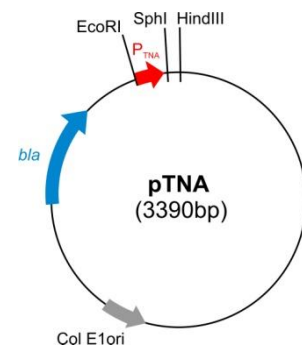


Abbildung 12: pTNA-Vektor.

2.8.2.2 pTNA13a

Bei pTNA13a handelt es sich um einen pTNA-Vektor, dessen multiple Klonierungsstelle (plus Linker und *egfp*-Gen) identisch zu pER13a (Abbildung 11b) ist, wobei allerdings die *Nde*I-Schnittstelle durch eine *Sph*I-Restriktionsschnittstelle ersetzt ist. Einfügen eines Gens über die Schnittstellen *Sph*I und *Not*I führt demnach zur schwach konstitutiven Expression des mit *egfp* fusionierten Gens.

2.8.2.3 pHCE-IIB

Der pHCE-IIB Vektor ist ein Derivat des kommerziellen pUC19-Vektors (FERMENTAS), wobei der induzierbare *lacZ*-Promotor durch den stark konstitutiv exprimierenden HCE Promotor des D-AAT Gens aus *Geobacillus toebii* ersetzt wurde (Poo *et al.*, 2002). Dieser ermöglicht im Gegensatz zum pET-System eine Überexpression rekombinanter Proteine ohne Induktion. Die Klonierung des Zielgens erfolgte in der Regel über die *Nco*I- und *Hind*III-Schnittstellen

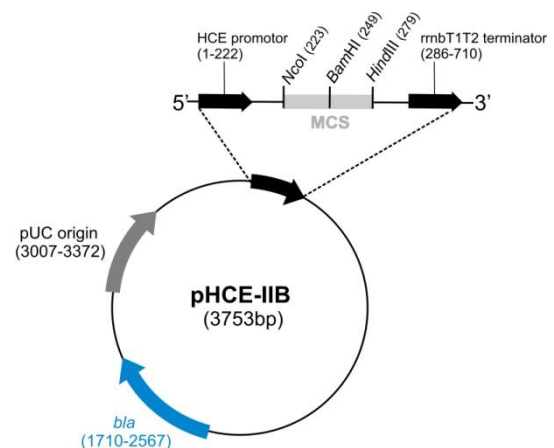


Abbildung 13: pHCE-IIB Vektor mit multipler Klonierungsstelle.

(Abbildung 13). Die Selektion plasmidtragender *E. coli* Zellen erfolgt über eine Ampicillin-Resistenz, die durch Expression des *bla*-Gens ausgeprägt wird. Der Vektor wurde freundlicherweise von Dr. Moon Hee Sung (BIOLEADERS CORPORATION, Daejeon, Korea) zur Verfügung gestellt.

2.8.2.4. pHCE13a

Bei pHCE13a handelt es sich um einen pHCE-IIB-Vektor, dessen multiple Klonierungsstelle (plus Linker und *egfp*-Gen) identisch zu pER13a (Abbildung 11b) ist, wobei allerdings die *Nde*I-Schnittstelle durch eine *Nco*I-Restriktionsschnittstelle ersetzt ist. Einfügen eines Gens über die Schnittstellen *Nco*I und *Not*I führt demnach zur stark konstitutiven Expression des mit *egfp* fusionierten Gens.

2.8.3 pCFN1

Der Vektor pCFN1 (Maxwell *et al.*, 1999) wurde so konstruiert, dass ein in die Klonierungsstellen *Bgl*II und *Xba*I eingefügtes Gen mit der stromabwärts kodierten Chloramphenicolacetyltransferase fusioniert wird. Das eingesetzte Gen wird durch eine für acht Aminosäuren kodierende Linkerregion

(GGRLEAAR) mit *cat* verbunden. Zwischen dem Insert und dem *cat*-Gen liegt zusätzlich ein *amber*-Stopcodon (Abbildung 14). Dadurch ist es möglich, in einem *amber*-Suppressorstamm das Fusionsprotein herzustellen, während in einem anderen Stamm die Expression nach dem inklonierten Gen abbricht. N-terminal wird das Fusionsprotein mit einem His₆-Tag und einem FLAG-Epitop zum Antikörpernachweis versehen. Der pCFN1-Vektor enthält einen starken, durch IPTG induzierbaren Promotor P_{trc}, pBR322 und Phagen f1 Replikationsursprünge, ein Ampicillinresistenzgen (*bla*) und ein *lac*I-Gen.

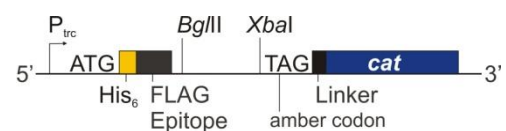


Abbildung 14: Multiple Klonierungsstelle von pCFN1.

2.9 Oligodesoxyribonukleotide

Erkennungssequenzen von Restriktionsenzymen sind unterstrichen. Bei Mutationsprimern sind diejenigen Basen, die von der Originalsequenz abweichen, blau markiert. Wurde ein Primer von Dritten zur Verfügung gestellt, erfolgt die Angabe der Person in Klammern.

2.9.1 Vektorspezifische Amplifikations- und Sequenzierungsprimer

T7 Promotor	88-69 bp <i>upstream</i> des Startcodons von pET21a(+) 5'- TAA TAC GAC TCA CTA TAG GG -3'
T7 Terminator	59-77 bp <i>downstream</i> des Stoppcodons von pET21a(+) bzw. 112-130 bp <i>downstream</i> der <i>NotI</i> -Schnittstelle von pERN7 5'- GCT AGT TAT TGC TCA GCG G -3'
CyRI	137-117 bp <i>upstream</i> des Startcodons des pTNA- und pTNA13a-Vektors 5'- TCA CGA GGC CCT TTC GTC TT -3'
CyPstI	110-132 bp <i>downstream</i> des Stoppcodons des pTNA-Vektors 5'- TCG CCA AGC TAG CTT GGA TTC T -3'
pERN7up	141-171 bp <i>upstream</i> des Startcodons des pERN7-Vektors 5'- TTC CCA AAA TTA GTT TGT TTT AAA AAA CGT -3'
5'pER13	114-134 bp bzw. 173-193 bp <i>upstream</i> des Startcodons des pER13a-, pER13cat-, pER14- bzw. pER13b-Vektors 5'- CCA CGA TGC GTC CGG CGT AG -3'
3'pER13	103-127 bp <i>downstream</i> der <i>NotI</i> -Schnittstelle der pER13a-, pER13b-, pTNA13a- und pHCE13a-Vektoren 5'- CAC CTT CAC CCT CTC CAC TGA CAG -3'
pCFNup	112-92 bp <i>upstream</i> des Startcodons des pCFN1-Vektors 5'- AAC GGT TCT GGC AAA TAT TC -3'
pCFNdown	28-47 bp <i>downstream</i> des <i>amber</i> -Stoppcodons des pCFN1- Vektors bzw. 159-178 bp <i>downstream</i> der <i>NotI</i> -Schnittstelle des pER13cat-Vektors 5'- TGG TTA TAG GTA CAT TGA G -3'
5'pUC	260-284 bp <i>upstream</i> des Startcodons des pHCE-IIB- bzw. pHCE13a -Vektors 5'- AGC GGA TAA CAA TTT CAC ACA GGA -3' (A. Lochner)
3'pUC-Ersatz	234-262 bp <i>downstream</i> des Stoppcodons des pHCE-IIB - Vektors 5'- TTG ATG CCT GGC AGT TCC CTA CTC TCG C -3' (A. Lochner)

2.9.2 Primer zur Herstellung neuer Fusionsvektoren

5' <i>cat</i> (<i>NotI</i>)	5'- GAC AAG <u>GGC GGC CGC</u> CTA GAG GCC GCT CGC GAG AAA AAA ATC ACT GGA TAT ACC ACC GTT G -3'
3' <i>cat</i> (<i>XhoI</i>)	5'- CAG TCA <u>CTC GAG</u> TCA TTA CGC CCC GCC CTG CCA CTC ATC -3'
5'N-eGFP(<i>NcoI</i>)	5'- ATA <u>CCA TGG</u> GTT AAC TGA GTA GGC GGC CGC G -3' (J. Claren)
5'N-eGFP(<i>SphI</i>)	5'- AGC <u>TGC ATG CGT</u> TAA CTG AGT AGG CGG CCG CGC T -3' (J. Claren)
3'C-eGFP(<i>HindIII</i>)	5'- ATG <u>CGA AGC TTA</u> TTT GTA GAG CTC ATC CAT GCC ATG TGT -3' (J. Claren)
5'pER13a_GSG	5'- AAA AGG GAT CCG GTA AAG GAG AAG AAC TTT TCA CTG GAG TTG TCC CAA TTC TTG -3'
3'pER13a_GSG	5'- CTT CTC CTT TAC CGG ATC CCT TTT GAT GAA ACA GAA GTT TTT TGA TAT TTC CAT TTG -3'

2.9.3 Primer zur Amplifikation von *hkas* und *tmhisF*

5' <i>hkas</i> (<i>NdeI</i>)	5'- ATA ATA <u>CAT ATG</u> GAG GAG GTG GTG ATT GCC GGC ATG TCC GGG -3'
5' <i>hkas</i> (<i>NcoI</i>)	5'- GAA GGA GAT ATA <u>CCC ATG GAG</u> GAG GTG GTG ATT GCC -3'
5' <i>hkas</i> (<i>SphI</i>)	5'- GAA GGA GAT ATA <u>CGC ATG CAG</u> GAG GTG GTG ATT GCC -3'
3' <i>hkas</i> (<i>NotI</i>)	5'- CAG TAT <u>GCG GCC GCC</u> CTG CGT GTT GGG CCT CAG GAT GAT GTG -3'
5' <i>tmhisF</i> -N(<i>NdeI</i>)	5'- AGC <u>CAT ATG</u> CTC GCT AAA AGA ATA ATC GCG -3' (J. Claren)
5' <i>tmhisF</i> (<i>NcoI</i>)	5'- GAA GGA GAT ATA <u>CCC ATG GTG</u> GCT AAA AGA ATA ATC -3'
5' <i>tmhisF</i> (<i>SphI</i>)	5'- GAA GGA GAT ATA <u>CGC ATG CTC</u> GCT AAA AGA ATA ATC -3'
3' <i>tmhisF</i> (<i>NotI</i>)	5'- CTT TCG GGC TTT GTT AGC <u>AGC GGC CGC</u> CTC ACA ACC CCT C -3' (J. Claren)

2.9.4 Primer zur Amplifikation und Mutagenisierung von *hpsk9*³¹⁻⁴⁵⁴

5' <i>hpsk9</i> (<i>NdeI</i>)	5'- GAA GGA GAT ATA <u>CAT ATG</u> CAG GAA GAT -3'
5' <i>hpsk9</i> (<i>SphI</i>)	5'- TTT AAG AAG GAG ATA TAC <u>GCA TGC</u> AGG AAG ATG AAG ATG G -3'

5' <i>hpcsk9</i> (<i>NcoI</i>)	5'- TTT AAG AAG GAG ATA TAC <u>CCA TGG</u> AAG ATG AAG ATG GTG ATT ATG -3'
3' <i>hpcsk9</i> (<i>NotI</i>)	5'- CAT AGC <u>GCG GCC GCC</u> TTC GGA TCC CTG CCA ACC GGC ACC ATG G -3'
3' <i>hpcsk9</i> (<i>XhoI</i>)	5'- GTG GTG GTG <u>CTC GAG</u> TTC GGA TCC CTG -3'
3' <i>hpcsk9_delNotI_1</i>	5'- ACA CCC GCA TCG CGT CCG CTA ACA ACG CCC -3'
5' <i>hpcsk9_delNotI_2</i>	5'- CAG AGC GGT ACC GCC CAG GCA GCC GCG CAT -3' GTG GCC G -3'
3' <i>hpcsk9_delSphI</i>	5'- CAG CAC ACG CAG GCT ACG CAT <u>CGA</u> GGC GCC TTT CGC AAC ACC -3'
3' <i>hpcsk9_delNcoI</i> (<i>NotI</i>)	5'- CAT AGC <u>GCG GCC GCC</u> TTC GGA TCC CTG CCA ACC GGC ACC <u>GTG CGT</u> GCT CGG CGG CAG CGC CGC -3'
3' <i>hpcsk9</i> (T60S+T61S+T63S)	5'- GCA GCG ATG AAA <u>GCT</u> CGC <u>GCT</u> <u>GCT</u> ACC GTG TTC -3'
3' <i>hpcsk9</i> (L179E)	5'- GAT AAA CTT CCA CCT <u>CGC</u> TGC CAC CAT CCG G -3'
5' <i>hpcsk9</i> (L276S+V277S)	5'- CGC AAA AGC CAG <u>GAG AGC</u> CAG CCG GTG GGC CCG C -3'
5' <i>hpcsk9</i> (T447S+W453S)	5'- CCG CCG AGC <u>AGC</u> CAT GGT GCC GGT <u>AGC</u> CAG GGA TCC GAA -3'
5' <i>hpcsk9</i> (Y293H)	5'- GCG GGT GGC <u>CAT</u> AGC CGC GTG CTG -3'
3' <i>hpcsk9</i> (T214S+F216H)	5'- CTG ACG ATG <u>GTG</u> ACG <u>GCT</u> ACC ATC TTC TTC -3'
5' <i>hpcsk9</i> (L324S)	5'- CTT TCG TGA TGA TGC GTG <u>TTC</u> GTA TAG CCC GGC -3'
3' <i>hpcsk9</i> (T377S)	5'- CTG GCT CAC AAA ACA <u>GCT</u> GCT GCA ATC GCT GC -3'
3' <i>hpcsk9</i> (T350S+L351S)	5'- CCA AAA TTG GTG CCC <u>GAG GAG</u> CCC AGG GTA ACC GGC -3'

2.9.5 Primer zur Amplifikation und Mutagenisierung von *hgr-lbd*

5' <i>hgr-lbd</i> (<i>NcoI</i>)	5'- GAA GGA GAT ATA <u>CCC ATG GTT</u> CCT GCA ACG TTA CCA CAA CTC ACC -3'
5' <i>hgr-lbd</i> (<i>SphI</i>)	5'- GAA GGA GAT ATA <u>CGC ATG CTG</u> CCT GCA ACG TTA CCA CAA CTC ACC -3'
3' <i>hgr-lbd</i> (<i>NotI/Stop</i>)	5'- TCG CCA GAA CCA GCA GCG GAG CCA <u>GCG CGG</u> <u>CCG CCT</u> ATT ACT TTT GAT G -3'
5' <i>hgr-lbd</i> _OE1	5'- TGT AAA CAC ATG CTG TAT GTT TCC TCT GAG -3'

3'hgr-lbd _OE1	5'- CTC AGA GGA AAC ATA CAG CAT GTG TTT ACA -3'
5'hgr-lbd _OE2	5'- TAT CTC TGT ATG AAA ACC TTA CTG CTT CTC T -3'
3'hgr-lbd _OE2	5'- AGA GAA GCA GTA AGG TTT TCA TAC AGA GAT A -3'
3'hgr-lbd (L544F)	5'- CAT ATC CTG CAT AAA ACA CTT CAG GTT CAA TAA CC -3'
3'hgr-lbd (I559N)	5'- GTT GAG CGT AGT CAT GTT CCT CCA AGT TGA GTC -3'
3'hgr-lbd (M560K)	5'- CAT GTT GAG CGT AGT CTT GAT CCT CCA AGT TG -3'
3'hgr-lbd (F602S)	5'- CAC CCC AGA GCA AAT GCC ATA AGA GAC ATC CAG GAG TA -3'
3'hgr-lbd (A605V)	5'- CAC CCC AGA GCA AAT ACC ATA AGA AAC ATC CAG GAG TA -3'
3'hgr-lbd (F602S+A605V)	5'- CAC CCC AGA GCA AAT ACC ATA AGA GAC ATC CAG GAG TA -3'
3'hgr-lbd (N619D)	5'- GCA AAA CAC AGC AGG TCT GCA CTT GAT TGT C -3'
3'hgr-lbd (H645R)	5'- GGA AAC ATA CAG CAT GCG TTT ACA TTG GTC -3'
5'hgr-lbd (K695E)	5'- GAA TGA CCT ACA TCG AAG AGC TAG GAA AAG CC -3'
5'hgr-lbd (V702A)	5'- GGA AAA GCC ATT GCC AAG AGG GAA GGA AAC -3'
5'hgr-lbd (E705G)	5'- CAT TGT CAA GAG GGG CGG AAA CTC CAG CCA GAA C -3'
5'hgr-lbd (V702A+E705G)	5'- GGA AAA GCC ATT GCC AAG AGG GGC GGA AAC TCC AGC CAG -3'
5'hgr-lbd (Q710R)	5'- GGA AGG AAA CTC CAG CCG CAA CTG GCA GCG G -3'
5'hgr-lbd (Q738H)	5'- CCT TAA CTA TTG CTT CCA CAC ATT TTT GGA TAA G -3'
5'hgr-lbd (M752T)	5'- GAA TTC CCC GAG ACG TTA GCT GAA ATC ATC -3'
5'hgr-lbd (N759D)	5'- CTG AAA TCA TCA CCG ATC AGA TAC CAA AAT ATT C 3'
3'hgr-lbd (F774S/NotI)	5'- CAG CGG AGC CAG CGC GGC CGC CTA TTA CTT TTG ATG AGA CAG AAG TTT TTT G -3'
3'hgr-lbd (F602S+A605F)	5'- GAT CTC CAC CCC AGA GCA AAA AAC ATA AGA GAC ATC CAG -3'

3' *hgr-lbd* (F602S+A605I) 5'- GAT CTC CAC CCC AGA GCA AAG ATC ATA AGA
GAC ATC CAG -3'

3' *hgr-lbd* (F602S+A605L) 5'- GAT CTC CAC CCC AGA GCA AAC AGC ATA AGA
GAC ATC CAG -3'

3' *hgr-lbd* (F602S+A605M) 5'- GAT CTC CAC CCC AGA GCA AAC ATC ATA AGA
GAC ATC CAG -3'

2.10 DNA- und Protein-Längenstandards

Zur Größenbestimmung von DNA-Fragmenten in Agarosegelen wurde der GeneRuler™ 1 kb Plus DNA Ladder (MBI FERMENTAS, St. Leon-Rot) verwendet (Abbildung 15a).

Zur Größenbestimmung von Proteinen bei SDS-PA Gelen bzw. bei *Western-Blots* wurden der Unstained Protein Molecular Weight Marker (MBI FERMENTAS) (Abbildung 15b), der Precision Plus Protein All Blue Prestained Marker (Bio-Rad) (Abbildung 15c) bzw. der Prestained PageRuler™ Protein Ladder Marker (MBI FERMENTAS) (Abbildung 15d) verwendet.

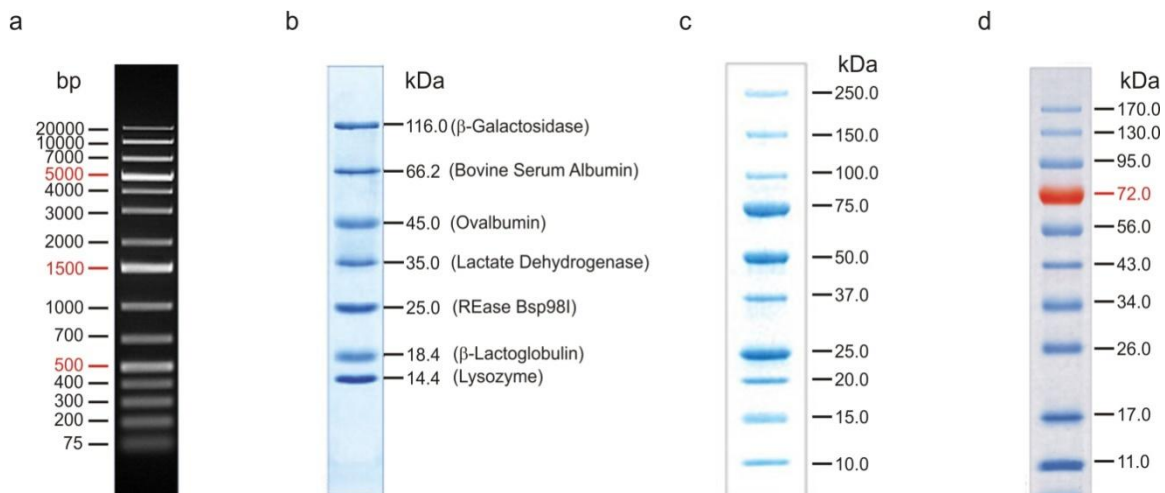


Abbildung 15: DNA- und Protein-Längenstandards.

a GeneRuler™ 1kb Plus DNA Ladder (MBI FERMENTAS) für Agarosegele.

b Unstained Protein Molecular Weight Marker (MBI FERMENTAS) für SDS-PAGE (12,5 %).

c Precision Plus Protein All Blue Prestained Marker (Bio-Rad) für SDS-PAGE (Bis-Tris 4-12 % Gele)

d Prestained PageRuler™ Protein Ladder (MBI FERMENTAS) für *Western-Blotting*.

2.11 Puffer und Lösungen

Soweit nicht anders angegeben wurden die verwendeten Puffer sterilfiltriert und bei Raumtemperatur gelagert. Es wurde MILLIPORE bzw. bidestilliertes Wasser verwendet.

2.11.1 Arbeiten mit *E. coli*

Antibiotika

Die langfristige Lagerung der gelösten Antibiotika erfolgte nach Sterilfiltrierung bei 20 °C, die kurzfristige Lagerung erfolgte bei 4 °C.

Ampicillin (1000 x):	150 mg/ml Ampicillin (Na-Salz) in Wasser gelöst und sterilfiltriert.
Chloramphenicol (1000 x):	30 mg/ml Chloramphenicol in 100 % EtOH gelöst und sterilfiltriert.
Kanamycin (1000 x):	75 mg/ml Kanamycin in Wasser gelöst und sterilfiltriert.
Dulbecco's PBS Puffer:	Sigma-Aldrich
Glucose-Stammlösung:	20 % (w/v) Glucose in Wasser gelöst und autoklaviert.
Glycerin (87 %)	
IPTG-Stammlösung	1 M IPTG in Wasser gelöst, sterilfiltriert, bei -20 °C gelagert.
KCl-Stammlösung:	1 M KCl in Wasser gelöst, sterilfiltriert und bei -20 °C gelagert.
MgCl ₂ -Stammlösung:	1 M MgCl ₂ in Wasser gelöst, sterilfiltriert und bei -20 °C gelagert.
MgSO ₄ -Stammlösung:	1 M MgSO ₄ in Wasser gelöst, sterilfiltriert und bei -20 °C gelagert
TFB I-Puffer:	100 mM KCl, 50 mM MnCl ₂ , 30 mM KOAc, 10 mM CaCl, 15 % Glycerol. Die Stammlösungen der Einzelkomponenten wurden bei 4 °C gelagert. Der Puffer wurde direkt vor der Verwendung frisch angesetzt.

TFB II-Puffer: 100 mM Tris/HCl pH 7,0, 10 mM KCl, 75 mM CaCl₂. Die Stammlösungen der Einzelkomponenten wurden bei 4 °C gelagert. Der Puffer wurde direkt vor der Verwendung frisch angesetzt.

2.11.2 Arbeiten mit DNA

dNTP-Lösung für PCR: je 2 mM dNTP (N = A, C, G oder T) in Wasser gelöst und bei -20 °C gelagert.

MgCl₂ für *ep*PCR: 50 mM MgCl₂ Stammlösung von Phusion High-Fidelity DNA-Polymerase (NEW ENGLAND BIOLABS); bei -20 °C gelagert.

MnCl₂ für *ep*PCR: 1 M MnCl₂ in Wasser gelöst, autoklaviert und bei -20 °C gelagert.

2.11.3 Agarosegelelektrophorese

Agarose (1 %): Eine entsprechende Menge an Agarose in 0,5 x TBE gelöst, aufgekocht und bei 60 °C aufbewahrt.

Ethidiumbromid-Stammlösung (10 mg/ml)

Sukrose-Farbmarker: 60 % (w/v) Sukrose, 0,1 % (w/v) Bromphenolblau, 0,1 % (w/v) Xylencyanol FF in 0,5 x TBE gelöst.

TBE (5 x): 445 mM Borsäure, 12,5 mM EDTA, 445 mM Tris (pH-Wert automatisch bei 8,15).

2.11.4 Arbeiten mit Proteinen

2.11.4.1 Allgemeine Arbeiten

EDTA-Stammlösung:	500 mM EDTA unter Zugabe von ca. 20 Plätzchen KOH in 1 l Wasser gelöst und anschließend mit 1 M KOH den pH-Wert auf 7,5 eingestellt. Die Lösung wurde bei RT gelagert.
Labelpuffer für Metallchelat-Affinitätschromatographie:	50 mM EDTA, 1 M NaCl zum Herunterwaschen der alten Ni ²⁺ -Ionen; 250 mM NiCl ₂ zum Beladen der Säule mit frischen Ni ²⁺ -Ionen; 500 mM NaCl zum Herunterwaschen unspezifisch gebundener Metallionen.

2.11.4.2 Arbeiten mit hPCSK9³¹⁻⁴⁵⁴

Ammoniumsulfat-Lösung zur fraktionierten Präzipitation:	3 M (NH ₄) ₂ SO ₄ in 50 mM Tris/Cl, 5 mM DTT, pH 8,0. Die exakte Konzentration wurde refraktometrisch bestimmt (3.5.9).
Aufschlusspuffer	
bei Protein-Herstellung im analytischen Maßstab:	50 mM Tris/Cl, 2 mM EDTA, pH 8,0.
bei präparativer Proteinreinigung:	50 mM Tris/Cl, 1 mM DTT, pH 8,0.
Dialyse-Puffer:	50 mM Tris/Cl, 5 mM DTT, pH 8,0.
DTT-Stammlösung:	500 mM DTT in Wasser gelöst, sterilfiltriert, bei -20 °C gelagert.
Puffer für Metallchelat-Affinitätschromatographie:	
Laufpuffer:	50 mM Tris/Cl, 300 mM NaCl, 5 mM Imidazol, pH 8,0.
Elutionspuffer:	50 mM Tris/Cl, 300 mM NaCl, 1 M Imidazol, pH 8,0.

Harnstoff-Stammlösung

für Auffaltungsexperimente: 10 M Urea unter Erhitzen in 50 mM Tris/Cl, pH 8,0 gelöst. Danach wurde zusätzlich 5 mM DTT zugegeben. Die exakte Konzentration des Harnstoffs wurde refraktometrisch bestimmt (3.5.9).

Tris/Cl-Puffer Stammlösung: 1 M Tris mit HCl auf pH 8,0 eingestellt und autoklaviert.

2.11.4.3 Arbeiten mit hGR-LBD

Ammoniumsulfat-Lösung zur

fraktionierten Präzipitation: 2,5 M $(\text{NH}_4)_2\text{SO}_4$ in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 50 μM Dexamethason, pH 7,9. Die exakte Konzentration wurde refraktometrisch bestimmt (3.5.9).

Aufschlusspuffer

bei Protein-Herstellung im analytischen Maßstab:

50 mM Tris/Cl, 400 mM NaCl, 0,5 % CHAPS, 10 % Glycerin, pH 7,9 ($\pm 50 \mu\text{M}$ Dexamethason).

bei präparativer Proteinreinigung:

50 mM Tris/Cl, 2 M Urea, 100 mM NaCl, 5 mM MgCl_2 , 40 mM Imidazol, 2 mM TCEP, pH 7,9 ($\pm 50 \mu\text{M}$ Dexamethason); direkt vor Benutzung wurde DNase I (Roche) (1 mg/ 50 ml) und EDTA-freier Protease Inhibitor Cocktail *complete* (Roche) (1 Tablette / 50 ml) zugesetzt.

Dexamethason-Stammlösung: 20 mM Dexamethason in EtOH gelöst, sterilfiltriert und bei 4 °C gelagert.

Dialyse-Puffer für Thrombin-Verdau: 50 mM Tris/Cl, 500 mM NaCl, 2 mM CaCl_2 , 2 mM TCEP, 0.5 % CHAPS, 10 % Glycerin, pH 7,9 ($\pm 10 \mu\text{M}$ Dexamethason).

Mifepriston-Stammlösung: 20 mM Mifepriston in EtOH gelöst, sterilfiltriert und bei 4 °C gelagert.

Puffer für Metallchelate-Affinitätschromatographie:

Laufpuffer: 50 mM Tris/Cl, 500 mM NaCl, 40 mM Imidazol, 2 mM TCEP, 10 % Glycerin, pH 7,9 (\pm 50 μ M Dexamethason).

Elutionspuffer: 50 mM Tris/Cl, 500 mM NaCl, 340 mM Imidazol, 2 mM TCEP, 10 % Glycerin, pH 7,9 (\pm 50 μ M Dexamethason).

Puffer für Glutathion-Affinitätschromatographie:

Laufpuffer: 50 mM Tris/Cl, 500 mM NaCl, 2 mM TCEP, 0.25 % CHAPS, 10 % Glycerin, pH 7,9 (\pm 10 μ M Dexamethason).

Elutionspuffer: 50 mM Tris/Cl, 500 mM NaCl, 2 mM TCEP, 0.25 % CHAPS, 10 % Glycerin, 10 mM L-Glutathion, pH 7,9 (\pm 10 μ M Dexamethason).

Puffer für präparative Gelfiltration: 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9.

Puffer für analytische Gelfiltration: 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 20 μ M Dexamethason, pH 7,9.

Puffer für HPLC

polares Laufmittel: H₂O, 0,1 % TFA

unpolares Laufmittel: Acetonitril, 0,08 % TFA

Tris/Cl-Puffer Stammlösung: 1 M Tris mit HCl auf pH 7,9 eingestellt und autoklaviert.

2.11.5 SDS-Polyacrylamidgelelektrophorese (12,5 % SDS-Gele)**Ammoniumperoxodisulfid (APS)-**

Stammlösung (25 % und 40 %): 25 % bzw. 40 % (w/v) APS in Wasser gelöst, sterilfiltriert, bei -20 °C gelagert.

Coomassie Färbelösung: 0,2 % (w/v) Coomassie Brilliant Blue G250 und R250, 50 % (v/v) Ethanol, 10 % (v/v) Eisessig, filtriert, bei RT lichtgeschützt gelagert.

Protogel™: 30 % (v/v) Acrylamid, 0,8 % (v/v) Bisacrylamid, in Wasser gelöst und bei 4 °C gelagert.

SDS-PAGE Trenngelpuffer: 0,4 % (w/v) SDS, 1,5 M Tris HCl, pH 8,8

SDS-PAGE Sammelgelpuffer:	0,4 % (w/v) SDS, 0,5 M Tris HCl, pH 6,8
SDS-PAGE Laufpuffer:	0,1 % (w/v) SDS, 0,025 M Tris, 0,2 M Glycin (pH-Wert automatisch bei 8,5).
SDS-PAGE Probenpuffer (2x):	2 % (w/v) SDS, 10 % (w/v) Glycerin, 5 % (v/v) β -Mercaptoethanol, 2 % (w/v) SDS, 0,01 % (w/v) Bromphenolblau, 1,25 M Tris/HCl, pH 6,8.
SDS-PAGE Probenpuffer (5x):	5 % (w/v) SDS, 25 % (w/v) Glycerin, 12,5 % (v/v) β -Mercaptoethanol, 5 % (w/v) SDS, 0,025 % (w/v) Bromphenolblau, 1,25 M Tris/HCl, pH 6,8.

2.11.6 Lösungen für *Western-Blotting*

Towbin Transfer Puffer:	10 mM Tris, 75 mM Glycin, 10 % Methanol, 0,01 % SDS.
PBS (10x):	91 mM Na_2HPO_4 , 17 mM NaH_2PO_4 , 0,15 M NaCl, auf pH 7,4 mit NaOH eingestellt.
Blockierungspuffer PBS/T+M:	1x PBS, 0,1 % Tween-20, 5 % Milchpulver.
Antikörperlösungen:	

Die Lösungen wurden stets frisch hergestellt und verwendet.

Anti-GFP Antikörper <i>mouse</i> monoklonal:	1:1000 in 1x PBS, 0,1 % Tween-20.
Peroxidase-konjugierter anti- <i>mouse</i> IgG:	1:5000 in 1x PBS, 0,1 % Tween-20, 1 % Milchpulver.
Waschpuffer PBS/T:	1x PBS, 0,1 % Tween-20.

2.12 Nährmedien

Zur Sterilisation wurden die Medien für 20 min bei 121 °C und 2 bar autoklaviert. Bei Selektivmedien wurden die entsprechenden Antibiotika nachträglich, nach weitgehendem Abkühlen der Medien, in Form von sterilfiltrierten 1000-fach konzentrierten Stammlösungen zugegeben.

Luria-Bertani-(LB)-Medium:	0,5 % (w/v) Hefeextrakt 1 % (w/v) Trypton 1 % (w/v) NaCl
LB-Agar:	LB-Medium plus 1,5 % (w/v) Bacto-Agar.
LB-Agar für CAT-Löslichkeitstest (3.2.8.2):	LB-Medium plus 1,5 % (w/v) Bacto-Agar; nach weitestgehendem Abkühlen werden verschiedene Konzentrationen an Chloramphenicol und 200 µg/ml (0,84 mM) IPTG zugegeben.
SOB-Medium:	0,5 % Hefeextrakt 2 % (w/v) Trypton 0,05 % NaCl Nach dem Autoklavieren wurden 10 mM MgSO ₄ , 10 mM MgCl ₂ und 2,5 mM KCl (alles steril) hinzugefügt.
SOC-Medium:	SOB-Medium plus Zugabe von zusätzlich 20 mM Glucose (sterilfiltriert) nach dem Autoklavieren.
SOC-Agar:	SOC-Medium plus 1,5 % (w/v) Bacto- Agar.

2.13 Software

ÄKTA Unicorn Version 5.01 (318)	© GE HEALTHCARE
AxioVision 5.7	© Carl Zeiss AG, Jena
BioEdit V.7.0.9	© IBIS BIOSCIENCES, Tom Hall
Cary Eclipse Version 1.1 (175)	© VARIAN Australia Pty. Ltd. 2002
Cary Win UV Version 3.0 (182)	© VARIAN Australia Pty. Ltd. 2002
Chromas Version 1.45	© 1998 Conor McCarthy http://www.technelysium.com.au/chromas.html
Corel Draw Version 12	© 2003 COREL Corp.
Deep View Swiss PDB Version 3.7	© Guex & Peitsch 1997 http://www.expasy.org/spdbv
Endnote Version 9	© 2000 WINTERTREE Software Inc. http://www.endnote.com
Spectra Manager 1	JASCO GmbH, Groß-Umstadt
OptiQuant Version 02.50	© Packard Instrument Co.
PyMOL™ 0.99rc6	© 2006 DELANO SCIENTIFIC LCC. http://www.pymol.org/
Seqman™II	© 1989-1999 DNASTAR Inc.
SigmaPlot 8.0/10.0	© 2004/2006 SPSS INC.
Summit Version 3.1	© CYTOMATION Inc., 2001

Frei zugängliche, servergestützte Anwendungen

3djigsaw	http://bmm.cancerresearchuk.org/~3djigsaw
CPHmodels 3.0	http://www.cbs.dtu.dk/services/CPHmodels
DaliLite	http://www.ebi.ac.uk/DaliLite
K2D	http://www.embl-heidelberg.de/~andrade/k2d/
MCoffee	http://www.igs.cnrsMrs.fr/Tcoffee/tcoffee_cgi/index.cgi?stage1=1&daction=MCOFFEE::Regular
Polyview	http://polyview.cchmc.org/index_old.html
MPI Tübingen Bioinformatics Toolkit	http://toolkit.tuebingen.mpg.de/
HHpred	http://toolkit.tuebingen.mpg.de/hhpred
Modeller	http://toolkit.tuebingen.mpg.de/modeller

3 Methoden

3.1 Vorbereitung von Geräten und Lösungen

Alle hitzestabilen Lösungen und Medien wurden vor Gebrauch 20 min bei 121 °C und 2 bar autoklaviert. Glaswaren und hitzestabile Verbrauchsmaterialien, wie z.B. Pipettenspitzen, wurden ebenso autoklaviert und anschließend bei 50 °C im Trockenschrank getrocknet. Hitzeelabile Lösungen wurden in Stammlösungen angesetzt und je nach Volumen durch einen Membranfilter der Porengröße 0,2 µm mittels einer Vakuumpumpe oder durch einen Spritzenfilter der Porengröße 0,22 µm filtriert. Puffer für die analytische oder präparative Gelfiltration wurden mindestens 30 min im Exsikkator entgast.

3.2 Mikrobiologische Methoden

3.2.1 Anzucht und Lagerung von *E. coli*-Stämmen

Alle verwendeten *E. coli* Stämme wurden bei 37 °C unter Schütteln bei 150 rpm (1 Liter Kultur) bzw. 220 rpm (5, 10 bzw. 50 ml Kultur) angezogen. Die Anzucht erfolgte in der Regel in LB-Medium. Bei plasmidtragenden Stämmen wurde dem Medium das entsprechende Antibiotikum (150 µg/ml Ampicillin, 30 µg/ml Chloramphenicol, 75 µg/ml Kanamycin) aus einer tausendfach konzentrierten, sterilfiltrierten Stammlösung zugegeben.

Um Einzelkolonien zu erhalten, wurde die Zellsuspension auf Agarplatten mit den entsprechenden Antibiotika ausplattiert und üN bei 37 °C im Brutschrank inkubiert.

Zur kurzzeitigen Lagerung wurden die Platten oder Suspensionen gut verschlossen und im Kühlschrank bei 8 °C aufbewahrt. Zur dauerhaften Lagerung wurden Glycerinkulturen hergestellt, welche bei -80 °C aufbewahrt wurden. Hierzu wurde ein Aliquot einer üN-Kultur im Verhältnis 1:1 in 87 % Glycerin in einem sterilen Schraubdeckelgefäß aufgenommen.

3.2.2 Herstellung chemisch kompetenter *E. coli* Zellen (Inoue *et al.*, 1990)

Zur Präparation chemisch kompetenter *E. coli* Zellen wurde 1 Liter LB-Medium mit dem entsprechenden Stamm auf eine OD₆₀₀ von 0,1 angeimpft und bei 37 °C und 150 rpm bis zu einer OD₆₀₀ von 0,5 - 0,6 angezogen. Die Kultur wurde 15 min auf Eis gekühlt, in 50 ml-Röhrchen überführt und durch Zentrifugation (EPPENDORF Centrifuge 5810R, 4.000 rpm, 10 min, 4 °C) geerntet. Die sedimentierten Zellen wurden in 100 ml eiskaltem TFB I-Puffer resuspendiert, erneut unter gleichen Bedingungen abzentrifugiert und das erhaltene Pellet in 10 ml eiskaltem TFB II-Puffer aufgenommen. Sofort nach der Resuspendierung wurden

100 µl Aliquots der Zellsuspension in Eppendorf-Reaktionsgefäße auf Eis überführt und bei -80 °C gelagert.

3.2.3 Chemische Transformation (Inoue *et al.*, 1990)

Zur Transformation chemisch kompetenter Zellen wurde ein 100 µl Aliquot der Zellen langsam auf Eis aufgetaut und mit etwa 100 ng Plasmid-DNA (maximales Volumen: 20 µl) versetzt. Nach fünf-minütiger Inkubation auf Eis wurde der Transformationsansatz für 45 s einem Hitzeschock bei 42 °C ausgesetzt. Danach wurde der Ansatz nochmals für fünf Minuten auf Eis gestellt. Nach Zugabe von 800 µl LB-Medium folgte eine Kur von einer Stunde bei 37 °C im Brutschrank, um die Antibiotikaresistenz auszubilden. Bei Transformation eines gereinigten Plasmids wurden 100 µl der Zellsuspension auf Selektivagar ausplattiert. Bei Transformation eines Ligationsansatzes wurde die komplette Zellsuspension kurz abzentrifugiert (EPPENDORF Centrifuge 5415D, 4.000 rpm, 2 min, RT), in ca. 100 µl LB-Medium resuspendiert und anschließend ausplattiert.

3.2.4 Herstellung elektrokompetenter *E. coli* Zellen (Dower *et al.*, 1988)

Zur Präparation elektrokompetenter *E. coli* Zellen wurden 50 oder 200 ml SOB-Medium mit den entsprechenden Zellen auf eine OD₆₀₀ von 0,1 angeimpft und bei 37 °C und 150 rpm bis zu einer OD₆₀₀ von 0,5 - 0,6 angezogen. Die Kultur wurde 30 min auf Eis gekühlt und die Zellen anschließend in 50 ml-Röhrchen abzentrifugiert (EPPENDORF Centrifuge 5810R, 4.000 rpm, 10 min, 4 °C). Das Zellpellet wurde in 50 ml eiskaltem, sterilem Wasser resuspendiert und erneut 15 min auf Eis gehalten. Nach nochmaliger Zentrifugation wurden die Zellen in 20 ml eiskaltem, sterilem Wasser resuspendiert, daraufhin erneut 15 min auf Eis gehalten, zentrifugiert und schließlich in 10 ml eiskaltem, sterilem Wasser resuspendiert. Nach einem letzten Inkubations- und Zentrifugationsschritt wurde der Überstand vollständig abgossen und die Zellen auf Eis gestellt. Das erhaltene Zellpellet wurde in 2 – 3 ml sterilem Wasser bei sofortiger Verwendung bzw. in 10 % Glycerin bei Lagerung der Zellen aufgenommen und zu je 100 µl Aliquots in Eppendorf-Reaktionsgefäße überführt. Die so erhaltenen elektrisch kompetenten Zellen wurden entweder sofort zur Transformation frisch verwendet oder bei -80 ° gelagert.

3.2.5 Vorbereitung von DNA für die Elektroporation

Bei einer Elektroporation muss die verwendete DNA-Lösung eine möglichst geringe Konzentration an Ionen enthalten, da diese beim elektrischen Puls in der Küvette einen Kurzschluss auslösen können. Deshalb muss jede DNA-Lösung, die in einem salzhaltigen Puffer vorliegt (z.B. Ligationsansatz), gegen Wasser dialysiert werden. Hierfür wurden 10 - 20 µl der DNA-Lösung für mindestens 1 h auf einen Nitrocellulosefilter (Millipore) aufgebracht, der in einer kleinen Petrischale auf sterilem Wasser aufgelegt wurde.

3.2.6 Elektroporation (Dower *et al.*, 1988)

Für die Transformation wurden 100 µl der frischen elektrokompetenten Zellen mit maximal 20 µl salzfreier DNA-Lösung versetzt. Nach fünfminütiger Inkubation auf Eis wurde der Transformationsansatz in eine vorgekühlte Elektroporationsküvette (Spaltbreite 2 mm) überführt. Nach dem elektrischen Puls (2500 V, 25 µF, 200 Ω) im Elektroporator (bei erfolgreicher Elektroporation lag die Zeitkonstante zwischen 4,0 und 6,0 ms) wurden die Zellen sofort in 1,5 ml SOC-Medium aufgenommen und für 1 h bei 37 °C im Schüttler (220 rpm) inkubiert. Danach wurden Aliquots auf Selektivagar ausplattiert bzw. zur Beimpfung von LB-Selektivmedium benutzt. Die Inkubation erfolgte jeweils üN bei 37 °C.

3.2.7 Bestimmung der Transformationseffizienz

Zur Bestimmung der Transformationseffizienz (T_E) wurde ein Aliquot der kompetenten Zellen wie oben beschrieben (3.2.3 / 3.2.6) mit 100 ng Plasmid-DNA transformiert. Nach dem Kuren wurden je 100 µl einer Verdünnungsreihe (unverdünnt, 1:10, 1:100, 1:1000) auf Selektivagar ausplattiert. Die Transformationseffizienz wurde folgendermaßen berechnet:

$$T_E = \frac{n_{\text{Kolonien}} \cdot f}{m_{\text{DNA}}}$$

Formel 3: Berechnung der Transformationseffizienz.

T_E : Transformationseffizienz (Kolonien pro µg DNA)

n_{Kolonien} : Anzahl der Kolonien

f : Verdünnungsfaktor

m_{DNA} : eingesetzte DNA-Menge [µg]

3.2.8 *In vivo* Analyse der Proteinlöslichkeit mit Hilfe von Reporter-Proteinen

3.2.8.1 *In vivo* Analyse der Proteinlöslichkeit mit Hilfe von eGFP als Reporter

Zur Abschätzung der Löslichkeit von Proteinen *in vivo* wird bei diesem System das grün fluoreszierende Protein (GFP) C-terminal an die zu untersuchenden Proteine fusioniert. Ist GFP an ein unlösliches, aggregierendes Protein gekoppelt, ist die Bildung des GFP-Chromophors erschwert. Zellen mit einem derartigen Konstrukt fluoreszieren deshalb schwächer als solche, in denen GFP mit einem löslichen Protein fusioniert ist (Waldo *et al.*, 1999; Waldo, 2003) (Abbildung 16). Zur Expression von mit *egfp* fusionierten Genen in *E. coli* wurden in der vorliegenden Arbeit die Vektoren pER13a, pER13b (2.8.1.2.1), pTNA13a (2.8.2.2) und pHCE13a (2.8.2.4) benutzt.

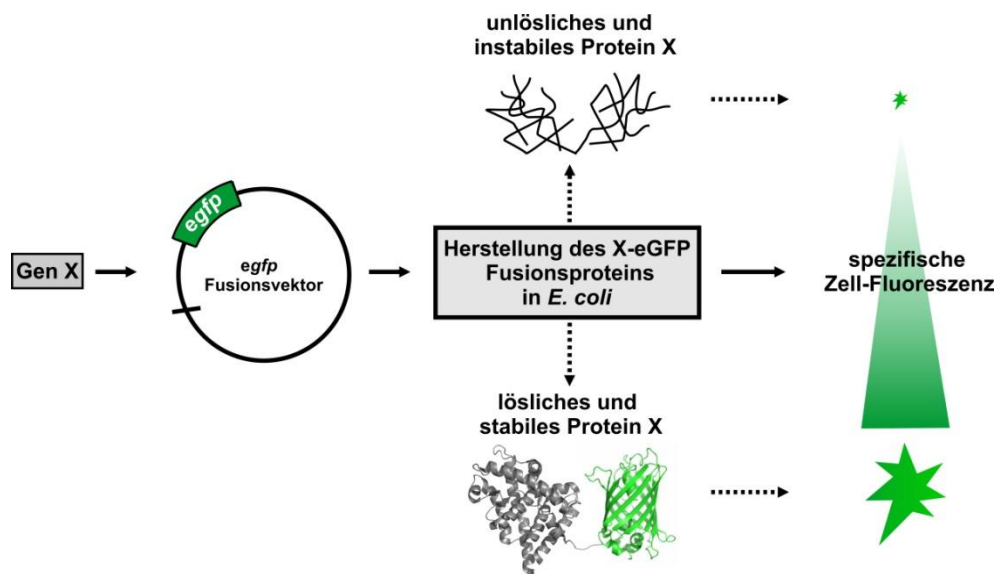


Abbildung 16: Schematische Darstellung der *in vivo* Analyse von Proteinlöslichkeit mit Hilfe von eGFP als Reporter.

Der Grad an Zell-Fluoreszenz korreliert mit der korrekten Faltung des eGFP in der Zelle. Diese hängt von den Eigenschaften des fusionierten Proteins X ab.

Einerseits wurde mit Hilfe dieser Methode die *in vivo* Löslichkeit von bestimmten Proteinen und deren Mutanten bewertet. Andererseits konnte dieses System benutzt werden, um Genbanken eines Proteins (3.3.9), welche in einen *egfp*-Fusionsvektor kloniert wurden, nach löslichen Varianten zu durchsuchen.

Nach Transformation (3.2.6) frischer elektrokompenter *E. coli* BL21(DE3) Zellen (3.2.4) mit einem *egfp*-Fusionsvektor, einer *egfp*-fusionierten Genbank oder einem bereits durch FACS-Sortierung erhaltenen Plasmid-Gemisch (3.2.9) wurden 50 ml LB-Selektivmedium mit 1 ml des Transformationsansatzes (Gesamt-Volumen: 1,5 ml; siehe 3.2.6) als Vorkultur angeimpft und bei 37 °C mit 220 rpm üN geschüttelt. Die Anzahl an angeimpften Transformanten wurde parallel bestimmt, indem verschiedene Verdünnungen desselben

Transformationsansatzes auf LB-Selektivagar ausplattiert wurden. Nach Inkubation in bei 37 °C wurden die gewachsenen Kolonien ausgezählt und darüber die Anzahl an unabhängigen Transformanten in dem zur Vorkultur-Inokulation benutzten 1 ml-Aliquot berechnet. Bei der Transformation von Genbanken bzw. FACS-Anreicherungen sollte die Anzahl der Transformanten die Genbank-Größe (Anzahl unabhängiger Klone) bzw. die Anzahl an zuvor sortierten Zellen mindestens um den Faktor zehn übersteigen. Dadurch soll sichergestellt werden, dass möglichst alle Mutanten, welche in der Genbank bzw. der Anreicherung enthalten sind, in der Vorkultur vertreten sind.

Es wurden 10 ml LB-Selektivmedium mit der hergestellten Vorkultur auf eine OD₆₀₀ von 0,1 angeimpft und bei 37 °C mit 220 rpm geschüttelt. Bei mit den konstitutiven Expressionsvektoren pTNA13a und pHCE13a transformierten Zellen erfolgte ein weiteres Wachstum von 6 h bei 37 °C mit 220 rpm. Bei mit den induzierbaren Expressionsvektoren pER13a oder pER13b transformierten Zellen wurde die Expression bei einer OD₆₀₀ von 0,6 mit 0,5 mM IPTG für 4 h bei 37 °C und 220 rpm induziert. Zum Teil wurde die Expression auch bei niedrigeren Temperaturen durchgeführt. Hierbei wurden die Zellen ab einer OD₆₀₀ von 0,6 für 30 min bei 30 °C und 220 rpm inkubiert, woraufhin die Expression mit 0,5 mM IPTG für 6 h induziert wurde. Bei der Analyse und der Sortierung von Zellen, die verschiedene Mutanten von *hgr-lbd* fusioniert an *egfp* exprimierten, wurde teilweise parallel zu IPTG zusätzlich 500 µM des hGR-LBD Agonisten Dexamethason zugegeben.

Von den so angezogenen Expressionskulturen wurde jeweils 1 ml abgenommen und zentrifugiert (EPPENDORF CENTRIFUGE 5415D, 4000 rpm, RT, 2 min). Die erhaltenen Zellpellets wurden in je 200 µl 1x PBS aufgenommen. Die spezifische Fluoreszenz der Zellen wurde anschließend mittels Fluoreszenzmikroskopie (3.2.10) oder Durchflusszytometrie (3.2.9) analysiert. Die Herstellung der eGFP-Fusionsproteine in *E. coli* konnte zudem durch *Western-Blotting* mit Anti-GFP Antikörpern (3.5.4) überprüft werden.

3.2.8.2 *In vivo* Analyse der Proteinlöslichkeit mit Hilfe der Chloramphenicol-Acetyltransferase (CAT) als Reporter

Zur Abschätzung der Löslichkeit von Proteinen *in vivo* wird bei diesem System die Chloramphenicol-Acetyltransferase (CAT) C-terminal an die zu untersuchenden Proteine fusioniert. Zellen, die ein relativ unlösliches Protein fusioniert an die CAT herstellen, tolerieren geringere Chloramphenicol-Konzentrationen als Zellen mit einem sehr löslichen CAT-Fusionsprotein (Abbildung 17).

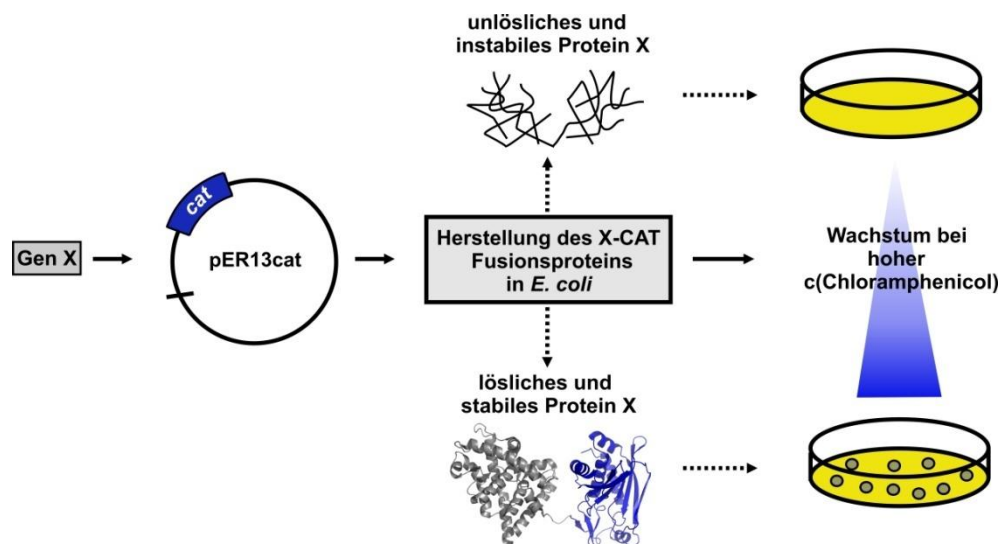


Abbildung 17: Schematische Darstellung der *in vivo* Analyse von Proteinlöslichkeit mit Hilfe von CAT als Reporter.

Der Grad an Antibiotikum-Resistenz der Zellen korreliert mit der korrekten Faltung der Chloramphenicol-Acetyltransferase (CAT) in der Zelle. Diese hängt von den Eigenschaften des fusionierten Proteins X ab.

Im Rahmen dieser Arbeit wurde der zum *egfp*-Fusionsvektor pER13a identisch konstruierte *cat*-Fusionsvektor pER13cat (2.8.1.2.3) benutzt. Die verbesserte Löslichkeit *in vivo* von Mutanten die nach *Screening* mit eGFP als Reporterprotein erhalten wurden, kann so mit einem alternativen Reporter überprüft werden. Nach erfolgreicher Umklonierung des Gens des zu untersuchenden Proteins in pER13cat wurden frisch hergestellte (3.2.4) elektrokompente *E. coli* T7 Express Zellen (2.7) mit dem Vektor transformiert (3.2.6). Nach je 30 min Inkubation bei 220 rpm und zuerst 37 °C und dann 30 °C erfolgte die Zugabe von 200 µg/ml IPTG und 500 µM des hGR-LBD Agonisten Dexamethason. Nach einer weiteren Stunde Inkubation bei 30 °C im Schüttler wurden die Zellen auf LB-Agarplatten mit IPTG (200 µg/ml) und verschiedenen Chloramphenicol-Konzentrationen (40 – 800 µg/ml) bzw. 75 µg/ml Kanamycin (Nullkontrolle) mit sterilen Glasstäben ausgestrichen und bei 30 °C im Brutschrank inkubiert. Das Wachstum sichtbarer Kolonien wurde in einem Zeitraum von 85 h verfolgt.

3.2.9 Analyse und Anreicherung fluoreszierender *E. coli* Zellen durch Durchflusszytometrie

Bei der Durchflusszytometrie mittels *Fluorescence Activated Cell Sorting* (FACS) können Zellen auf der Basis unterschiedlicher physikalischer und molekularer Eigenschaften analysiert und sortiert werden. Mit dem verwendeten Gerät (MoFlo[®] Zytometer mit Cyclone-Sortiereinheit und Coherent Enterprise II Laser) können bei der Analyse von bis zu 70.000 Zellen pro Sekunde drei verschiedene Parameter detektiert werden: das Vorwärts-Streulicht

(*Forward Scatter*, FSC), das Seitwärts-Streulicht (*Side Scatter*, SSC) und die Fluoreszenzintensität nach Anregung mit einem Argonlaser (488 nm) (FL1). Der FSC liefert ein Maß für die Größe der analysierten Zellen, die Detektion des SSC erfolgt zur Analyse der Granularität der Zellen. Im Rahmen dieser Arbeit wurde jedoch ausschließlich die Fluoreszenz der Zellen nach Anregung bei 488 nm analysiert, wobei die Emission über einen 530/40 nm (FITC) Bandpassfilter detektiert wurde. Durch hydrodynamische Fokussierung (Watson, 1999) in einer Düse wird die in das MoFlo[®] Zytometer eingeleitete Zellsuspension (in 1x PBS) in einem Strahl mit definiertem Durchmesser (verwendete Keramik-Düse: 70 µm) am gebündelten Laserstrahl vorbeigeleitet, wobei stets nur eine Zelle im Querschnitt des Meßbereiches vorhanden ist. Um eine Sortierung von einzelnen Zellen vornehmen zu können wird die Düse so in Schwingungen versetzt, dass der Strahl in einzelne Tropfen bricht, wobei jeder Tropfen eine Zelle enthält. Die Detektoren für das Seitwärts (SSC)- und Vorwärtsstreulicht (FSC) bzw. die Fluoreszenz (FL1) nehmen dabei das Signal auf, das jede einzelne Zelle liefert. Wird dabei eine Zelle mit den zuvor definierten Eigenschaften erkannt, kann diese heraus sortiert werden. Dies erfolgt durch ein elektrisches Feld, in welchem der entsprechende Tropfen elektrostatisch aufgeladen wird und anschließend seitlich in ein Auffanggefäß oder auf eine LB-Selektivagarplatte über geladene Platten abgelenkt wird (Abbildung 18).

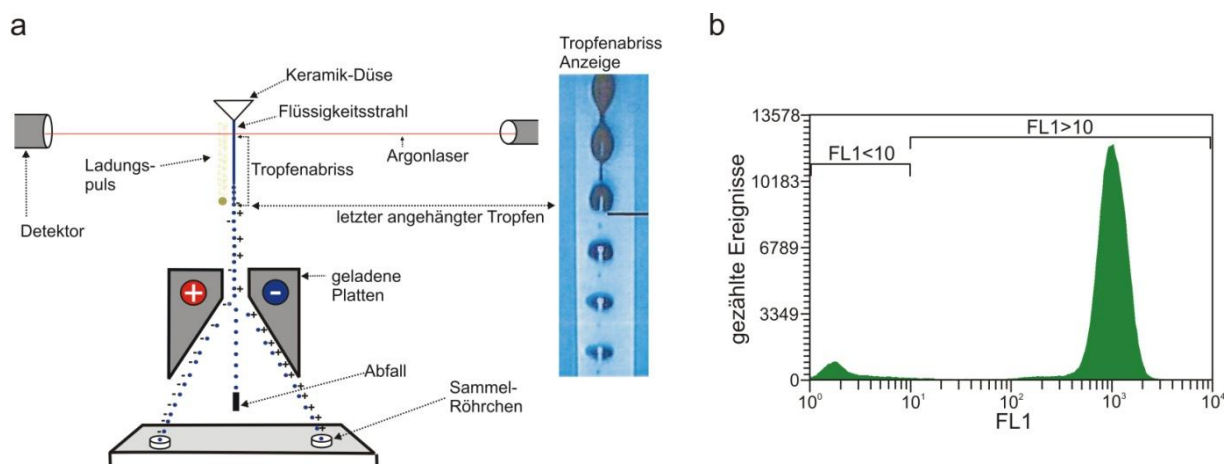


Abbildung 18: Schematischer Aufbau des MoFlo[®] Zytometers und exemplarisches Histogramm einer FACS-Analyse.

a Graphische Darstellung des Sortier-Prinzips des MoFlo[®] Zytometers. Nach erfolgreicher Kalibrierung des Tropfenabrisse und des Argon-Lasers auf den Flüssigkeitsstrahl können einzelne Zellen durch Ablenkung geladener Tropfen sortiert werden.

b Analyse von 300.000 *E. coli* BL21(DE3) Zellen, 4 h nach Induktion der Expression bei 37 °C von tmHisF-eGFP. Das FACS-Histogramm zeigt einen homogenen Peak, wobei 277.453 gezählte Ereignisse mit FL1>10 eine mittlere Fluoreszenz (mf) von mf = 1035,1 zeigen. Die 22.357 gezählten Ereignisse mit FL1<10 gehen auf Signalrauschen und Zellen mit fehlerhafter Expression zurück. Sie werden bei der Bestimmung des mittleren Fluoreszenzwerts der Zellen vernachlässigt.

Die Kalibrierung des Geräts mit Hilfe von fluoreszierenden *Beads* (Durchmesser: 6 μM) (AlignFlow[®], Invitrogen) erfolgte bei einer Photomultiplier-Spannung von 400 Volt für alle drei Detektoren (FSC, SSC und FL1). Der Strahlengang und der Tropfenabrisspunkt des MoFlo[®] Zytometers wurden vor jedem Lauf justiert und fokussiert. Die Datenverwaltung erfolgte durch das Softwarepaket SUMMIT[™] 3.1. Zur Reinigung nach jedem Lauf wurde das System für 15 min mit 1 \times PBS und anschließend für 15 min mit 70 % Ethanol über die Probenaufnahme gespült.

Analyse und Sortierung fluoreszierender Zellen erfolgte mit folgenden Einstellungen am MoFlo[®] Zytometer:

Frequenz:	100.000 Hz (für 70 μm Düse)
Amplitude der Düse:	10–15 Volt (legt die Form der Tropfen fest)
Spannung der elektrischen Platten bei Sortierung:	2500 Volt
Photomultiplier-Spannung:	FSC (<i>Forward Scatter</i>): 400 Volt SSC (<i>Side Scatter</i>): 750 Volt FL1 (Fluoreszenz): 600 Volt
Probendruck:	57–59 psi

Die Empfindlichkeit des Gerätes wird über den *Threshold* eingestellt, wobei das Hintergrundrauschen der Signal-Detektion niedrig gehalten werden muss. Dazu wird das Gerät so eingestellt, dass ca. 100 - 200 Ereignisse pro Sekunde gezählt werden, wenn über die Probenaufnahme nur 1 \times PBS (ohne Zellen) zugeführt wird.

Die zu vermessenden Zellen der *E. coli*-Kulturen mit eGFP-Fusionsproteinen (3.2.8.1) wurden in einem 5 ml Probenröhrchen so verdünnt, dass das Gerät üblicherweise 10.000 bis 15.000 Ereignisse (Zellen) pro Sekunde (eps) analysierte.

Zur Analyse der Fluoreszenzintensität von Expressionskulturen wurden stets jeweils $3 \cdot 10^5$ Zellen vermessen. Bei Quantifizierung der mittleren Fluoreszenz (mf) der Zellen wurden Fluoreszenz-Ereignisse mit einem Signal kleiner als zehn vernachlässigt, da es sich hierbei um Hintergrundrauschen des Geräts oder *E. coli* - Zellen mit fehlerhafter Expression handelt (Abbildung 18). Die täglich neu durchgeführte Kalibrierung des Gerätes kann zur Verschiebung der Fluoreszenz identischer Kulturen an verschiedenen Tagen führen, weswegen alle in dieser Arbeit gezeigten Analysen innerhalb eines FACS-Histogramms stets parallel aufgenommen wurden.

Bei Sortierungen von Zellen wurden in der ersten Sortierrunde 10^7 - 10^8 Ereignisse analysiert. Diejenigen Zellen, deren Fluoreszenzintensität in einem zuvor definierten

Screening-Bereich lag, wurden in einem leeren Probenröhrchen gesammelt. Um die Qualität der Auslese zu erhöhen wurden die gesammelten Zellen ohne erneute Anzucht direkt nochmals im gleichen *Screening*-Bereich einer Sortierung unterzogen („Resort“). Die hierbei sortierten Zellen wurden in einem mit LB-Medium gefüllten Probenröhrchen aufgefangen, womit anschließend 50 ml LB-Selektivmedium angeimpft werden konnten. Nach Inkubation üN bei 37 °C im Schüttler mit 220 rpm konnte von dieser Mischkultur die Plasmid-DNA präpariert werden (3.3.7.1). Nach Transformation von frischen elektrokompenten *E. coli* BL21(DE3) Zellen (3.2.4/3.2.6) mit dem so gewonnenen Plasmid-Gemisch und anschließender Expression der sortierten *egfp*-Fusionsvarianten (3.2.8.1) konnte die nächste Anreicherungsrunde durchgeführt werden. Dieser Zyklus wurde so lange wiederholt, bis eine ausreichende Anreicherung von Zellen im eingestellten *Screening*-Bereich erfolgt war. Diese wurden dann mit Hilfe der Cyclone Sortiereinheit auf LB-Selektivagarplatten vereinzelt, wobei nur mit einer Rate von 500 Ereignissen pro Sekunde sortiert wurde. Die LB-Selektivagarplatten wurden üN bei 37 °C inkubiert und anschließend bei 4 °C gelagert. Um sicherzustellen, dass keine Varianten während der Anreicherungszyklen verloren gehen, wurden in jeder neuen Runde mindestens 10x mehr Varianten analysiert als in der vorherigen Runde sortiert worden waren. Darüber hinaus wurde bei jedem Transformationsschritt die Anzahl der unabhängigen Transformanten über parallel ausplattierte Verdünnungsreihen bestimmt, um sicherzustellen dass mindestens 10x mehr Zellen transformiert wurden als in der vorherigen Runde sortiert worden waren (siehe 3.2.8.1). Im Anhang dieser Arbeit ist ein detailliertes Protokoll zur näheren Erläuterung der Vorgehensweise bei FACS-Sortierungen gezeigt (siehe 7.1; Abbildung A1). Darüber hinaus wird in 4.2.2 eine exemplarische Anreicherung stark fluoreszierender Zellen aus einem Zell-Gemisch dargestellt.

3.2.10 Fluoreszenzmikroskopie von *E. coli* Zellen

Die in 1x PBS resuspendierten *E. coli* BL21(DE3) Zellen, welche *egfp*-fusionierte Gene exprimieren (3.2.8.1), wurden auf Objektträgern mit 0,8 % Agarose (in 1x PBS) fixiert. Bilder der fluoreszierenden Zellen wurden mit Hilfe eines Zeiss Axiovert 200M Mikroskops aufgenommen, ausgestattet mit einem 63x/1,4 Objektiv, einer Zeiss AxioCamMR Digital Kamera und der AxioVision 5.7 Software. Die Detektion der eGFP-Fluoreszenz, erfolgte nach Anregung bei 488 nm mit Hilfe eines Bandpassfilters (Emission: BP 510/50, Zeiss FS37 Filter). Die Belichtungsdauer betrug je nach Fluoreszenz-Intensität zwischen 50 und 10000 ms.

3.3 Molekularbiologische Methoden

3.3.1 Amplifikation von DNA-Fragmenten (Mullis & Faloona, 1987; Saiki *et al.*, 1988)

Die Polymerasekettenreaktion (PCR; Mullis & Faloona, 1987; Saiki *et al.*, 1988) ermöglicht die Vervielfältigung definierter DNA-Fragmente *in vitro*. Dies geschieht über zyklisches Wiederholen von Denaturierung der doppelsträngigen DNA, gefolgt von Anlagerung (*Annealing*) zweier Primer (synthetische Oligonukleotide, die das zu amplifizierende DNA-Fragment flankieren) und enzymatischer DNA-Synthese (*Extension*). Die Amplifikation von DNA-Fragmenten erfolgt exponentiell. Die Reaktion wurde in einem Volumen von 50-200 µl in einem Thermocycler (Deckeltemperatur 110 °C) durchgeführt. Die Reaktionen erfolgten mit je 5–100 ng zu amplifizierender DNA-Matrize, 0,025 U/µl GoTaq[®] DNA-Polymerase, 5× Green GoTaq[®] Reaktionspuffer (enthält 7,5 mM MgCl₂ (Endkonzentration: 1,5 mM MgCl₂) und Auftragspuffer), je 0,2 mM dNTPs und jeweils 1 µM der beiden Primer. Zur fehlerfreien Amplifikation von DNA-Fragmenten für Klonierungen wurde zu diesem Ansatz noch 2,5 U *Pwo*-Polymerase zugegeben, da diese im Gegensatz zur *Taq*-Polymerase eine 3' → 5'-*Proofreading*-Aktivität besitzt. Ein Standard PCR-Programm setzte sich wie folgt zusammen.

Schritt	Temperatur [°C]	Dauer
1. Denaturierung	95	5 min
2. Denaturierung	95	45 s
3. Primer-Anlagerung	T _A	45 s
4. Extension	72	1 min/kb
5. Finale Extension	72	10 min
7. Lagerung	4	∞

Die Schritte 2 bis 4 wurden 35-mal wiederholt.

Die optimale Anlagerungstemperatur T_A der Primer wurde nach folgenden Gleichungen (Chester & Marshak, 1993) ermittelt, wobei der GC-Gehalt und die Länge der Primer berücksichtigt werden:

$$T_M = 69,3 + 0,41 \cdot (\%GC) - \frac{650}{n}$$

Formel 4: Berechnung der Schmelztemperatur eines Oligonukleotids.

T_M: Schmelztemperatur des Primers [°C]

%GC: GC-Gehalt des Primers [%]

n: Anzahl der Nukleotide des Primers

$$T_A = \left(\frac{T_{M1} + T_{M2}}{2} \right) - 3^\circ\text{C}$$

Formel 5: Berechnung der optimalen Anlagerungstemperatur der Primer in einer PCR-Reaktion.

T_A : *Annealing*-Temperatur [$^\circ\text{C}$]
 T_{M1} & T_{M2} : Schmelzpunkte der eingesetzten Primer

Die optimale *Annealing*-Temperatur wurde zum Teil auch experimentell bestimmt. Hierfür wurden mehrere PCR-Ansätze im Gradienten-Cycler (EPPENDORF Mastercycler gradients) parallel angesetzt. Mittels Agarosegelelektrophorese (3.3.4) konnte anschließend ermittelt werden, welche *Annealing*-Temperatur die höchste Ausbeute an spezifischem Amplifikationsprodukt lieferte. Die Spanne des Gradienten lag zumeist im Bereich einer Anlagerungstemperatur zwischen 50°C und 70°C .

3.3.2 Kolonie-PCR

Zur Überprüfung der korrekten Insertion eines Gens in die multiple Klonierungsstelle eines Vektors wurde ein *Insert-Screening* mittels Kolonie-PCR durchgeführt. Dazu wurden auf Selektivplatten gewachsene Einzelkolonien dem PCR-Ansatz hinzugefügt und bei dem Denaturierungsschritt von 95°C aufgeschlossen. Hierzu wurden die Zellen mit einer Pipettenspitze von der Platte gepickt und kurz auf dem Boden des PCR-Gefäßes verstrichen. Die freigesetzte DNA diente in den folgenden Amplifikationsrunden als Matrize. Als Primer wurden in der Regel genflankierende Plasmidprimer eingesetzt. Das Volumen des Ansatzes betrug jeweils $20\ \mu\text{l}$. Im Ansatz waren je $0,5\ \text{U}$ GoTaq[®] DNA-Polymerase, Green GoTaq[®] Reaktionspuffer, $0,1\ \text{mM}$ dNTPs und je $0,5\ \mu\text{M}$ der beiden Primer enthalten. Die Amplifikation erfolgte mittels Standard-Protokoll wie in 3.3.1 beschrieben.

3.3.3 PCR-Methoden zur gerichteten und zufälligen Mutagenese

3.3.3.1 Megaprimer-PCR (Sarkar & Sommer, 1990)

Die Megaprimer-PCR (MP-PCR) ist eine Form der DNA-Manipulation, mit der gerichtet Punktmutationen eingeführt werden können. In einer ersten PCR wird der Megaprimer hergestellt, indem mit Hilfe eines genflankierenden Primers und eines Primers, der die Mutation enthält, ein Fragment (der Megaprimer) amplifiziert wird. Je nach Lage der Mutation wird die Mutation über einen fehlerhaften 3'-Primer in Kombination mit einem genflankierenden 5'-Primer oder über einen fehlerhaften 5'-Primer in Kombination mit einem genflankierenden 3'-Primer eingeführt. Der Megaprimer wurde mittels präparativer

Gelelektrophorese (3.3.5) gereinigt und diente zusammen mit dem das Gen auf der gegenüberliegenden Seite flankierenden Primer zur Amplifikation des gesamten Gens in einer zweiten PCR (Abbildung 19).

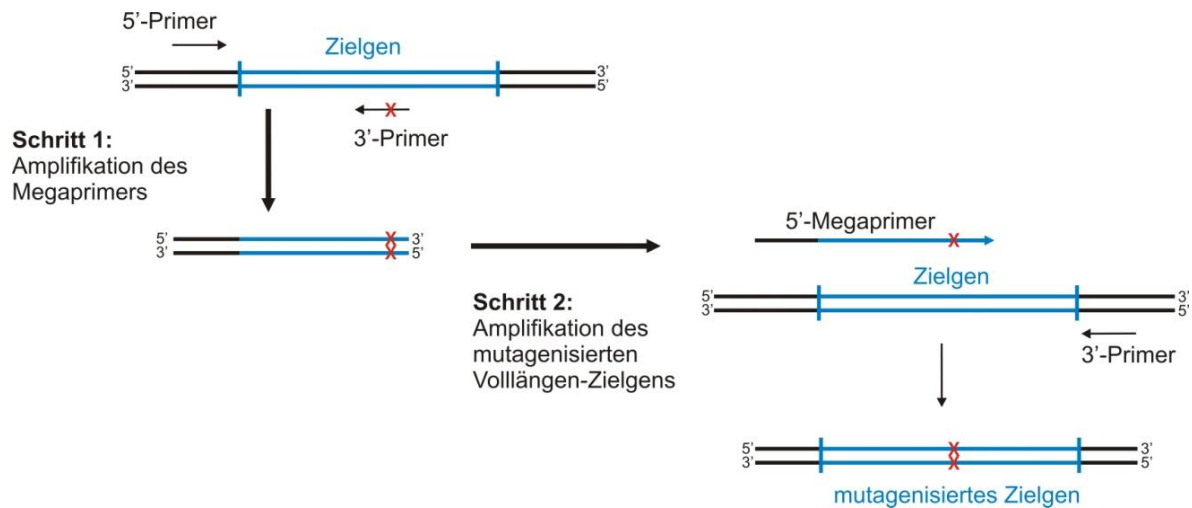


Abbildung 19: Prinzip der Megaprimer-PCR.

Gezielte Mutationen werden im 1. Schritt durch zum Zielgen komplementäre fehlerhafte Primer eingeführt.

Der entstehende Megaprimer sollte möglichst kurz sein, um eine effektive Anlagerung an die Matrize im 2. Schritt gewährleisten zu können. Die erste und zweite PCR wurden nach beschriebener Methode (3.3.1) durchgeführt, wobei jedoch bei der Amplifikation des Vollängen-Zielgens die Dauer der Primer-Anlagerung auf 1 min erhöht wurde. Die optimale Anlagerungs-Temperatur des Megaprimers wurde wenn nötig experimentell ermittelt.

3.3.3.2 *Overlap Extension* – PCR (OE-PCR)

Bei der *Overlap Extension* – PCR (OE-PCR) zur gerichteten Mutagenese werden zwei zueinander komplementäre Mutationsprimer von 25 bis 35 bp Länge definiert, die komplementär zum Zielgen sind und im zentralen Bereich den gewünschten Nukleotidaustausch tragen. Im ersten Schritt werden in getrennten PCR-Ansätzen mit diesen Primern in Kombination mit je einem genflankierenden Primer von der gleichen Matrize zwei überlappende Genfragmente amplifiziert und mittels präparativer Gelelektrophorese gereinigt (3.3.5). Diese Fragmente überlappen im Bereich der ausgetauschten Nukleotide und dienen in einer zweiten PCR als Matrize, um in Kombination mit genflankierenden Primern das vollständige Gen mit den ausgetauschten Nukleotiden zu amplifizieren (Abbildung 20).

Die Amplifikationsschritte erfolgten mittels Standard-Protokoll wie in 3.3.1 beschrieben.

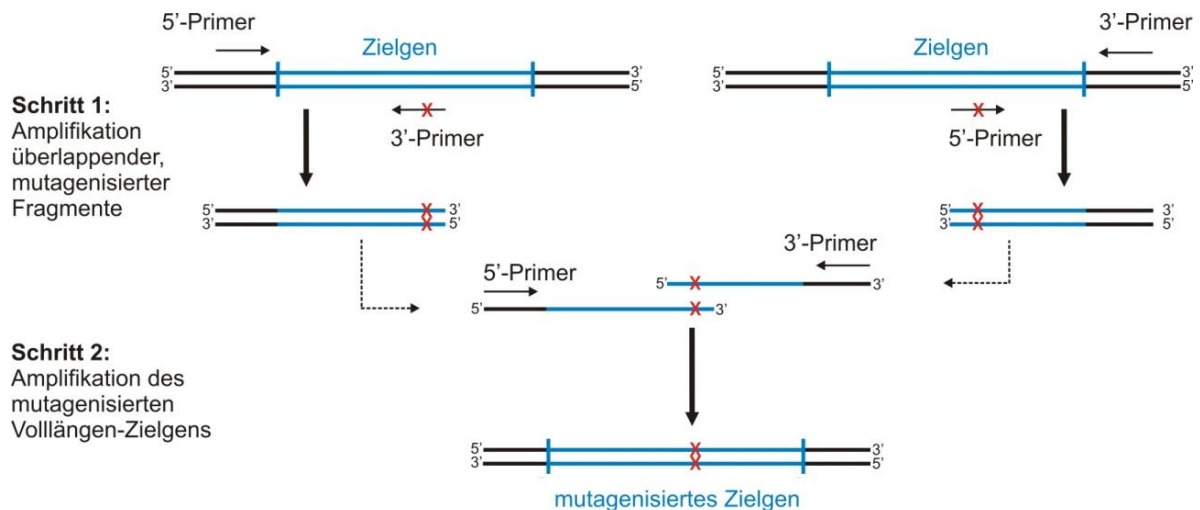


Abbildung 20: Prinzip der *Overlap Extension* - PCR.

In zwei getrennten Ansätzen, in denen das Zielgen jeweils als Matrize dient, werden zwei Fragmente amplifiziert, die in einem überlappenden Bereich den gewünschten Austausch tragen. Im 2. Schritt dienen diese Fragmente als Matrize und mit Hilfe genflankierender Primer kann das mutagenisierte Vollängen-Zielgen amplifiziert werden.

3.3.3.3 Zufallsmutagenese durch *error prone* PCR (*epPCR*)

Zur Einführung zufälliger Punktmutationen in ein Gen wurde die Methode der *error-prone* PCR (*epPCR*) benutzt (Leung *et al.*, 1989). Hierbei wurde die GoTaq[®] DNA-Polymerase verwendet. Diese besitzt keine 3' → 5'-*Proofreading*-Aktivität, wodurch es im Schnitt zu $8 \cdot 10^{-6}$ bis $2 \cdot 10^{-4}$ Fehlern pro Nukleotid kommt (Eckert & Kunkel, 1990; Cline *et al.*, 1996). Die Fehlerrate kann durch Verwendung ungleicher Konzentrationen der verschiedenen Desoxynukleotide (Fromant *et al.*, 1995; Vanhercke *et al.*, 2005) sowie durch den Zusatz von Mn^{2+} weiter gesteigert werden (Beckmann *et al.*, 1985; Leung *et al.*, 1989). In der vorliegenden Arbeit wurden zwei verschiedene *epPCR*-Protokolle verwendet (Tabelle 1). Die einzelnen Amplifikationsschritte wurden wie in 3.3.1 beschrieben durchgeführt, wobei die Primer-Anlagerung bei 55 °C erfolgte und insgesamt nur 30 Amplifikations-Zyklen durchgeführt wurden.

Tabelle 1: Verwendete *error prone* PCR (*epPCR*) Reaktionsbedingungen.
Das Gesamt-Volumen aller gezeigter Reaktionsansätze beträgt 200 μ l.

	konventionelle PCR	<i>ep</i> PCR -1	<i>ep</i> PCR -2
MgCl ₂ [mM]	1,5	1,5	2,3
MnCl ₂ [mM]	-	0,375	0,64
dATP [mM]	0,2	0,2	0,35
dCTP [mM]	0,2	0,2	0,4
dGTP [mM]	0,2	0,2	0,2
dTTP [mM]	0,2	0,2	1,35
Primer [μ M]	1	1	1
DNA-Matrize [ng]	20-400	200	100
<i>GoTaq</i> -DNA-Polymerase [U]	5	20	20

Bei *epPCR*-1 wurde, wie bei Vanhercke *et al.* (2005) beschrieben, lediglich MnCl₂ zum PCR-Ansatz hinzugefügt (Verhältnis Mg²⁺ : Mn²⁺ = 4 : 1). Bei *epPCR*-2 wurden zum Einen die zugegebene Mengen an DNA-Matrize und MgCl₂ bzw. MnCl₂ (Verhältnis Mg²⁺ : Mn²⁺ = 3,6 : 1) verändert und zum Anderen, wie bei Fromant *et al.* (1995) bzw. Drummond *et al.* (2005) beschrieben, zusätzlich die Konzentrationen der dNTPs variiert.

Bei beiden *epPCR* Ansätzen wurden 20 U *Taq*-Polymerase zugegeben, um die Produktmenge zu erhöhen. Findet ein Fehlereinbau in den ersten Zyklen der PCR statt, kann diese Mutation durch die exponentielle Vermehrung der Fragmente im finalen PCR-Produkt überrepräsentiert sein und somit die Vielfältigkeit der Genbank verringern. Um dies zu vermeiden, wurden die Ansätze zuerst in einem Gesamtvolumen von 200 μ l pipettiert, dann aber in acht Aliquots aufgeteilt. Nach erfolgter PCR wurden die Ansätze wieder vereinigt und die Menge an amplifiziertem Fragment mittels Agarosegelelektrophorese (3.3.4) überprüft.

3.3.4 Agarosegelelektrophorese

Im Agarosegel werden DNA-Fragmente ihrer Länge entsprechend aufgetrennt. Durch Zugabe des DNA-interkalierenden Fluoreszenz-Farbstoffs Ethidiumbromid werden die Banden im Gel unter UV-Licht sichtbar (Sharp *et al.*, 1973). Zur Herstellung eines Agarosegels wurde 1 % (w/v) Agarose in 0,5 % TBE-Puffer durch Aufkochen in einer Mikrowelle gelöst. Nach Abkühlen auf 50 – 60 °C wurden 0,2 μ l der Ethidiumbromidstammllösung (10 mg/ml) pro ml Agarose hinzugegeben, die Lösung in eine Gelkammer gegossen und ein Taschenkamm eingesetzt, der nach dem Erstarren wieder entfernt wurde. Das fertige Gel wurde mit 0,5 % TBE-Puffer überschichtet. Die DNA-Proben wurden nun bei Bedarf mit 10 x DNA-

Probenpuffer versetzt und in die Geltaschen pipettiert. Die Elektrophorese wurde bei 190 V für etwa 20 min durchgeführt. Die negativ geladene DNA wandert durch die angelegte Spannung zur Anode, wobei verschieden lange DNA-Fragmente durch das Agarosenetz unterschiedlich stark gebremst werden.

Mit UV-Licht der Wellenlänge 302 nm wurden die DNA-Fragmente detektiert und mit Hilfe des Imager Multi-Doc-It Digital Imaging System die Gele dokumentiert. Als Größenstandard wurden 5 µl des GeneRuler™ 1 kb Plus DNA Ladder von Fermentas verwendet.

3.3.5 Isolierung von DNA-Fragmenten aus Agarosegelen

Unter UV-Licht der Wellenlänge 302 nm wurden gewünschte Fragmentbanden mit einem sauberen Skalpell aus dem Gel ausgeschnitten und in ein Reaktionsgefäß überführt. Die Extraktion der DNA erfolgte mit Hilfe des Nucleospin®II-Kits von Macherey-Nagel laut Anleitung des Herstellers. Die isolierte DNA wurde in 30 µl Wasser aufgenommen und bei -20 °C gelagert.

3.3.6 Enzymatische Manipulation von dsDNA

3.3.6.1 Spaltung von dsDNA mit Restriktionsendonukleasen

Für die spezifische Spaltung von dsDNA wurden Restriktionsendonukleasen vom TypII verwendet (Sambrook *et al.*, 1989, Wilson & Murray, 1991), die eine palindromische Erkennungssequenz besitzen. Diese erzeugen einzelsträngige Überhänge, sogenannte *sticky-ends* mit 3'-Hydroxyl- und 5'-Phosphatenden. Für analytische Spaltungen wurde etwa 1 µg DNA mit 20 U von jedem Restriktionsenzym für 2 – 3h in einem Volumen von ca. 20 µl in entsprechendem Puffer bei 37 °C inkubiert. Für präparative Spaltungen zur späteren Ligation von Fragmenten wurde die gesamte Menge an erhaltenem PCR-Produkt bzw. 2 µg Vektor-DNA mit 10 – 30 U von jedem Restriktionsenzym in einem Gesamtvolumen von je 50 µl verdaut. Die Inkubation erfolgte bei 37 °C für 2 h bzw. üN. Die Reaktionen wurden in dem jeweils vom Hersteller empfohlenen Puffer durchgeführt. Der Anteil der Restriktionsenzyme im Reaktionsansatz sollte nicht mehr als 10 % des Volumens betragen, da die glycerinhaltige Lösung, in der die Enzyme gelagert werden, deren Aktivität beeinflussen kann. Die Fragmente wurden mittels Agarosegelelektrophorese analysiert (3.3.4) und für die Ligation aus dem Gel aufgereinigt (3.3.5).

3.3.6.2 Ligation von DNA-Fragmenten

Zur Ligation wurden geschnittener Vektor und Fragment in einem abgeschätzten molaren Verhältnis von 1 zu 4 gemischt. Die Reaktion erfolgte in einem Gesamtvolumen von 20 µl mit 1 U T4-DNA-Ligase (Fermentas) im entsprechenden Puffer üN bei 16 °C im Thermocycler oder 1 h bei RT. Danach wurden kompetente *E. coli* Zellen mit dem Ansatz chemisch (3.2.3) transformiert.

Zur Ligation von Genbanken wurde geschnittener Vektor und randomisiertes Fragment ebenfalls in einem abgeschätzten molaren Verhältnis von 1 zu 4 gemischt. Die Reaktion erfolgte stets in einem Gesamtvolumen von 20 µl mit 1 U T4-DNA-Ligase (Roche) im entsprechenden Puffer üN bei 16 °C im Thermocycler. Nach der Entfernung von Salz aus dem Ligationsansatz (3.2.5) wurden kompetente *E. coli* Zellen mit dem Ansatz elektrisch transformiert (3.2.6).

3.3.7 Isolierung und Reinigung von Plasmid-DNA aus *E. coli*

Die Präparation von Plasmid-DNA aus Zellen erfolgte nach dem Prinzip der alkalischen Lyse (Le Gouill *et al.*, 1994). Durch Zugabe von SDS und Natriumhydroxid werden zunächst die Zellen aufgeschlossen. Anschließend werden die Proteine und die genomische DNA durch Neutralisation der Suspension unter Zugabe von Ammoniumacetat gefällt, während die zirkuläre Plasmid-DNA in Lösung bleibt.

3.3.7.1 Isolierung von Plasmid-DNA in analytischen Mengen (Mini-Präparation)

Die analytische Isolierung und Reinigung von Plasmid-DNA aus *E. coli* erfolgte entweder mit Hilfe des QIAprep Spin Miniprep Kit der Firma Qiagen oder mit Hilfe des GeneJET™ Plasmid Miniprep Kit der Firma Fermentas. Dazu wurden aus einer frischen üN Kultur 5 ml Zellsuspension abzentrifugiert (EPPENDORF 5810R, 4000 rpm, 5 min, RT). Die Isolierung der Plasmid-DNA erfolgte anschließend nach den Angaben des Herstellers. Zur Elution wurden stets 50 µl steriles Wasser verwendet. Die gewonnene DNA wurde bei -20 °C gelagert.

3.3.7.2 Isolierung von Plasmid-DNA in präparativen Mengen (Midi/Maxi-Präparation)

Die Isolierung von Plasmid-DNA zu präparativen Zwecken erfolgte durch das NucleoBond® PC 100 (Midi-Präparationen) oder NucleoBond® PC 500 Kit (Maxi-Präparationen) von Macherey-Nagel.

Für Midi-Präparationen wurden 50 ml einer *E. coli* üN-Kultur abzentrifugiert (EPPENDORF 5810R, 4.000 rpm, 10 min, 4 °C). Die anschließende Isolierung der Plasmid-DNA erfolgte nach Angaben des Herstellers unter Verwendung der AX100 Säulen, wobei die final erhaltene Plasmid-DNA in 50 µl sterilem Wasser aufgenommen wurde. Die Plasmidpräparation wurde bei -20 °C gelagert.

Maxi-Präparationen wurden zur Gewinnung von Plasmid-Genbanken durchgeführt. Dazu wurden die auf großen Platten (Ø 14,5 cm) gewachsenen Zellen mit Hilfe von flüssigem LB-Medium (5 ml für je 3 Platten) und einem sterilen Spatel abgeschabt. Die so gewonnene Zellmasse wurde abzentrifugiert (EPPENDORF 5810R 4000 rpm, 10 min, 4 °C) und das Plasmidgemisch im präparativen Maßstab nach Angaben des Herstellers isoliert, wobei die erhaltene Plasmid-DNA in 300 µl sterilem Wasser aufgenommen und bei -20 °C gelagert wurde. Pro Maxi-Präparation wurden maximal 5 g Zellmasse verwendet.

3.3.8 Bestimmung der DNA-Konzentration

Die Konzentration von DNA-Lösungen wurde photometrisch bei einer Wellenlänge von 260 nm bestimmt. Nach dem Lambert-Beerschen-Gesetz entspricht eine OD₂₆₀ von 1 (mit $0,1\% A_{260} = 20 \text{ cm}^2 \text{ mg}^{-1}$ und einer Küvetten-Schichtdicke von 1 cm) einer DNA-Konzentration von 50 µg/ml dsDNA (bzw. 35 µg/ml RNA oder 33 µg/ml ssDNA).

Dadurch kann die DNA-Konzentration wie folgt berechnet werden:

$$c_{dsDNA} = \frac{A_{260} \cdot 50 \cdot f}{1000}$$

Formel 6: Bestimmung der DNA-Konzentration.

c_{dsDNA} : Konzentration an doppelsträngiger DNA [µg/µl]

A_{260} : Absorption bei 260 nm

f: Verdünnungsfaktor

Eine saubere DNA-Lösung sollte oberhalb von 300 nm keine messbare Absorption zeigen und einen OD₂₆₀/OD₂₈₀ Quotienten von mindestens 1,8 aufweisen.

3.3.9 Herstellung von Genbanken

Zunächst wurde das Gen, von dem eine Bank angelegt werden sollte, mit Hilfe von *ep*PCR (3.3.3.3) randomisiert und die Reaktionsprodukte sowie 2 µg des Zielvektors, mit den entsprechenden Restriktionsendonukleasen geschnitten (3.3.6.1). Die Produkte der *ep*PCR und des Restriktionsverdau wurden jeweils über Agarosegelelektrophorese (3.3.4)

aufgetrennt und aus dem Gel aufgereinigt (3.3.5), wobei die DNA jeweils in 50 °C warmen sterilen Wasser eluiert wurde, um eine optimale DNA-Ausbeute zu gewährleisten. Die Ligation (3.3.6.2) erfolgte in 8 - 10 Ansätzen à 20 µl bei 16 °C üN. Frisch hergestellte (3.2.4) elektrokompente *E. coli* XL1 Blue MRF^c Zellen (2.7) wurden mit jeweils 5 µl des dialysierten Ligationsansatzes (3.2.5) transformiert (3.2.6), d. h. pro Genbank wurden 32 - 40 Transformationen durchgeführt. Nach Inkubation der in je 1,5 ml SOC-Medium aufgenommenen Transformationsansätze (siehe 3.2.6) in 15 ml Röhrchen für 1 h bei 37 °C und 220 rpm wurden die Zellen vereinigt, abzentrifugiert (EPPENDORF Centrifuge 5810R, 4000 rpm, 10 min, 25 °C) und danach in 8 ml SOC-Medium resuspendiert. Daraufhin wurde die Zellsuspension in Aliquots à 200 µl auf großen SOC- Selektivagarplatten (Ø 14,5 cm, 40 Stück) ausgestrichen.

Nach dem Anwachsen der Zellen bei 37 °C üN im Brutschrank wurden die SOC-Platten mit einem sterilen Spatel und LB-Medium abgeschabt und die dabei entstehende Bakteriensuspension vereinigt und abzentrifugiert (EPPENDORF Centrifuge 5810R, 4000 rpm, 10 min, 4 °C). Das Zellpellet wurde gewogen und die Plasmidgenbank über Maxi-Präparationen (3.3.7.2) isoliert.

Die Anzahl der transformierten Zellen (n_T) wurde bestimmt indem die angewachsenen Kolonien auf einer der vierzig großen SOC-Selektivagarplatten (Ø 14,5 cm) ausgezählt wurden, wodurch auf die Gesamtanzahl der gewachsenen Kolonien hochgerechnet werden konnte. Anschließend wurden 20 Einzelkolonien zufällig ausgewählt und über eine Kolonie-PCR (3.3.2) die Ligationseffizienz (L) nach Formel 7 bestimmt.

$$L = \frac{n_V}{n_G} \cdot 100$$

Formel 7: Bestimmung der Ligationseffizienz in Prozent.

n_V : Anzahl der Kolonien mit Vollängenkonstrukt

n_G : Anzahl der analysierten Kolonien

Die Größe der Genbank (G) berechnet sich demnach nach Formel 8 aus dem Produkt der Anzahl an transformierten Zellen (n_T) und der Ligationseffizienz (L).

$$G = n_T \cdot \frac{L}{100}$$

Formel 8: Bestimmung der Größe einer Genbank aus Formel 7 und Formel 8.

Um zusätzlich die Effizienz der für die Genbank verwendeten *ep*PCR zu überprüfen, wurden jeweils zehn Klone vollständig sequenziert und hinsichtlich der Zahl der Mutationen, der auftretenden Insertionen bzw. Deletionen und des Verhältnisses von Transitionen zu Transversionen (Ts/Tv) analysiert. Die Anzahl der Basenaustausche pro Gen wurde wie folgt berechnet:

$$A_{\text{Basen}} = \frac{\sum a_{\text{Basen}}}{n_{\text{Seq}}}$$

Formel 9: Berechnung der durchschnittlichen Basenaustausche pro Gen durch *ep*PCR.

A_{Basen} : durchschnittliche Anzahl der Basenaustausche pro Gen.

a_{Basen} : Anzahl der Basenaustausche pro analysiertem Gen.

n_{Seq} : Anzahl der sequenzierten Gene.

3.3.10 DNA-Sequenzierung

Alle in dieser Arbeit klonierten Gene wurden vollständig mit vektorspezifischen bzw. genflankierenden Primern sequenziert. Die Bestimmung der Nukleotidsequenzen erfolgte durch die Firma Geneart (Regensburg, www.geneart.com). Es wurde das Kettenabbruch-Verfahren von Sanger (Sanger *et al.*, 1977) auf einem ABI PRISM® 3700 DNA Analyzer von Applied Biosystems verwendet. Bei diesem Verfahren werden verschiedene fluoreszenzmarkierte 2',3'-Didesoxynukleotide zusätzlich zu den 2'-Desoxynukleotiden in der *Thermal Cycle Sequencing* Reaktion (Sears *et al.*, 1992; Slatko, 1994) eingesetzt. Dadurch entstehen 3'-terminal fluoreszenzmarkierte Einzelstrang-Abbruchfragmente, die gelektrophoretisch aufgetrennt werden können. Die Nukleotidsequenz kann anschließend einfach abgelesen werden. Die abgegebenen Proben mit einem Gesamtvolumen von 8 µl enthielten zwischen 100 und 300 ng DNA und jeweils 10 pmol eines Sequenzierungsprimers, wodurch die DNA linear amplifiziert wurde. Die im ABI Format erhaltenen Sequenzen wurden mit den Programmen BioEdit, Chromas und Seqman II analysiert.

3.4 Proteinchemische Methoden

3.4.1 Genexpression und Proteinanreicherung

3.4.1.1 Expression und Anreicherung von Proteinen im analytischen Maßstab

Um zu überprüfen, ob sich die klonierten Gene überexprimieren lassen und in welcher Fraktion des Zellextraktes sich das Protein anschließend befindet, wurde eine Expression im analytischen Maßstab durchgeführt („Probeexpression“).

Nach Transformation chemisch kompetenter *E. coli* Zellen mit den entsprechenden Vektoren wurden 5 ml LB-Selektivmedium mit 50 µl des Transformationsansatzes angeimpft und bei 37 °C mit 220 rpm üN geschüttelt. Anschließend wurden 10 ml LB-Selektivmedium mit dieser Vorkultur auf eine OD₆₀₀ von 0,1 angeimpft und bei 37 °C mit 220 rpm geschüttelt. Nach Absenkung der Inkubationstemperatur für 30 min auf 20 °C (hGR-LBD Varianten) oder 18 °C (hPCSK9³¹⁻⁴⁵⁴-Varianten) ab einer OD₆₀₀ von 0,6 wurde die Expression mit 0,5 mM (hGR-LBD Varianten) oder 0,1 mM IPTG (hPCSK9³¹⁻⁴⁵⁴-Varianten) üN induziert.

Bei Expression verschiedener *hgr-lbd* Varianten wurde nach Absenkung der Inkubationstemperatur auf 20 °C teilweise parallel zu IPTG zusätzlich 500 µM des hGR-LBD Agonisten Dexamethason zugegeben. Bei der Expression von *hpcsk9*³¹⁻⁴⁵⁴-Varianten wurde parallel zur Absenkung der Inkubationstemperatur auf 18 °C ein osmotischer Schock durch Zugabe von 0,5 M NaCl induziert, welcher zur erhöhten Expression zelleigener Chaperone führt.

Standardmäßig wurden nach der Expression von jeder Kultur 2 ml Zellsuspension entnommen, welche zur Präparation der unlöslichen (Pellet P) und löslichen Zellfraktion (*crude extract* C) verwendet wurde (Abbildung 21).

Bei quantitativer Analyse der löslichen Herstellung von His₆-GST fusionierten hGR-LBD Varianten (3.5.3) wurde die entnommene Zellsuspensions-Menge jedoch auf die Zellzahl normiert. Hierfür wurde die Zelldichte über die optische Dichte der Zellsuspension bestimmt, um anschließend jeweils exakt 2×10^9 Zellen (entspricht 2 ml Zellen einer OD₆₀₀ von 3,2) für die Probenpräparation zu entnehmen. Die lösliche Zellfraktion wurde bei diesen analytischen Probeexpressionen zudem über Nickel-Sepharose gereinigt (gereinigter *crude extract* C = C*), um exakte Aussagen über die Menge an gewünschtem, löslichem Ziel-Protein treffen zu können (Abbildung 21, blau markiert). Die Nickel-Sepharose wurde vor Verwendung dreimal mit Aufschlusspuffer gewaschen.

Die mit SDS-Probenpuffer versetzten Proteinfractionen wurden je für 5 min bei 95 °C inkubiert und auf ein SDS-Polyacrylamidgel aufgetragen (3.5.2). Hierbei wurden von den C-

und P-Fractionen stets gleiche Mengen aufgetragen (je nach Geltaschen-Größe 10 µl oder 25 µl). Von der C*-Fraktion wurde etwas weniger für die SDS-PAGE verwendet (je nach Geltaschen-Größe 5 µl oder 12 µl), wobei jedoch bei Analyse verschiedener Proteinvarianten die jeweils aufgetragene Menge gleich war. Dadurch konnte die Ausbeute an gereinigtem, löslichem Protein direkt verglichen werden.

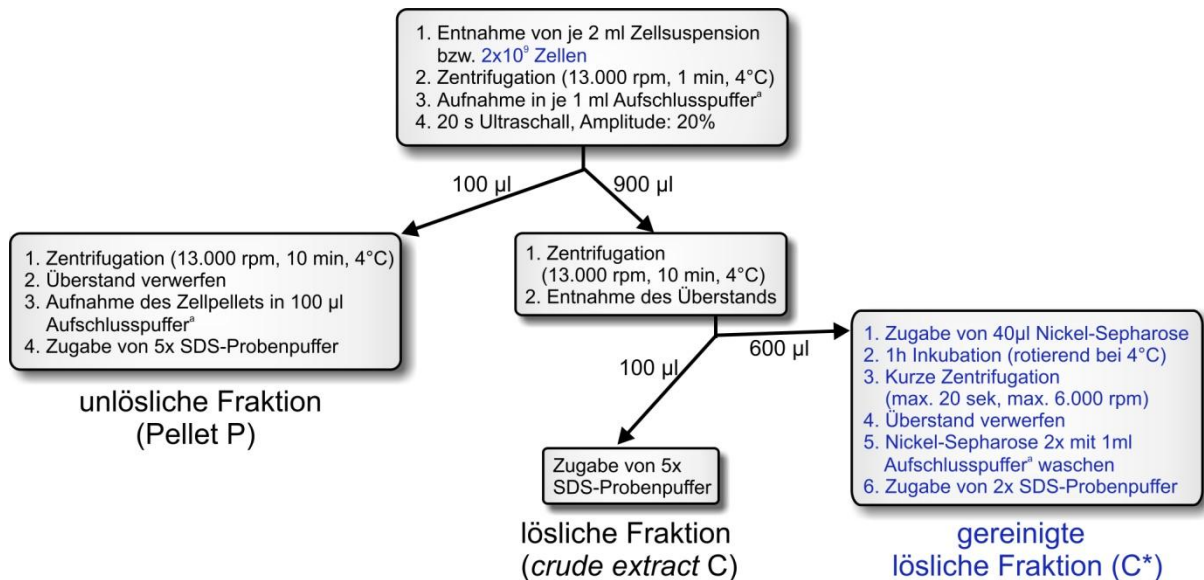


Abbildung 21: Schematische Darstellung des Vorgehens bei einer Probeexpression.

^a: Aufschlusspuffer: bei His₆-GST-hGR-LBD Varianten: 50 mM Tris/Cl, 400 mM NaCl, 10 % Glycerin, 0,5 % CHAPS, 2 mM TCEP, pH 7,9 (± 50 µM Dexamethason); bei hPCSK9³¹⁻⁴⁵⁴-Varianten: 50 mM Tris/Cl, 2 mM EDTA, pH 8,0. Zusätzliche Schritte bei analytischen Probeexpressionen zur Herstellung von Präparationen für die Quantifizierung des löslichen Zielproteins (3.5.3) sind in blau dargestellt.

3.4.1.2 Expression und Anreicherung von Proteinen im präparativen Maßstab

Zur Aufreinigung der Proteine wurden 2-10 l LB-Medium (1 l in einem 3 l Glaskolben) mit den entsprechenden Antibiotika mit einer frischen Vorkultur einer Einzelkolonie auf eine OD₆₀₀ von 0,1 angeimpft und bei 37 °C und 150 rpm im Schüttler inkubiert. Nach Absenkung der Inkubationstemperatur für 30 min auf 20 °C (His₆-GST-hGR-LBD Varianten) oder 18 °C (hPCSK9³¹⁻⁴⁵⁴-Varianten) bei einer OD₆₀₀ von 0,6 wurde die Expression mit 0,5 mM (His₆-GST-hGR-LBD Varianten) oder 0,1 mM IPTG (hPCSK9³¹⁻⁴⁵⁴-Varianten) üN induziert.

Bei Expression verschiedener *hgr-lbd* Varianten fusioniert an *his₆-gst*, wurde nach Absenkung der Inkubationstemperatur auf 20 °C teilweise parallel zu IPTG zusätzlich 500 µM des hGR-LBD Agonisten Dexamethason zugegeben. Bei Expression verschiedener *hpcsk9*³¹⁻⁴⁵⁴-Varianten wurde parallel zur Absenkung der Inkubationstemperatur auf 18 °C ein osmotischer Schock durch Zugabe von 0,5 M NaCl induziert, welcher zur erhöhten Expression zelleigener Chaperone führt.

Die Zellen wurden nach Inkubation ün geerntet (Sorvall RC5B plus, GS3 Rotor, 4000 rpm, 30 min, 4 °C). Das Zellpellet wurde mit 50 ml 1x PBS pro Liter Zellkultur gewaschen, woraufhin das Nassgewicht der Zellen bestimmt wurde. Falls nötig wurden die Zellen bei -20 °C gelagert. Zum Zellaufschluss wurden die Zellen in 25 ml Aufschlusspuffer (His₆-GST-hGR-LBD Varianten: 50 mM Tris/Cl, 2 M Urea, 100 mM NaCl, 5 mM MgCl₂, 40 mM Imidazol, 2 mM TCEP, DNase I (Roche) (1 mg/ 50 ml), EDTA-freier Protease Inhibitor Cocktail *complete* (Roche) (1 Tablette / 50 ml), pH 7,9 (± 50 µM Dexamethason). hPCK9³¹⁻⁴⁵⁴-Varianten: 50 mM Tris/Cl, 1 mM DTT, pH 8,0.) pro Gramm Zellmasse aufgenommen. Die Zellsuspension wurde anschließend mit einem Glas-Potter homogenisiert und auf Eis sonifiziert (Amplitude: 50 %, 2 x 2 min mit 15 s Ultraschall/15 s Pause). Nach Zentrifugation (Sorvall RC5B plus, SS34 Rotor, 15000 rpm, 30 min, 4 °C) wurde der lösliche Überstand zur weiteren Entfernung von Schwebeteilchen mit Hilfe von Spritzenfiltern (Porengröße: 0,45 µm) filtriert, um anschließend weiterverarbeitet werden zu können.

3.4.2 Proteinreinigung aus der löslichen Zellfraktion

3.4.2.1 Metallchelate-Affinitätschromatographie

Die Metallchelate-Affinitätschromatographie wurde an der Chromatographie-Anlage ÄKTA™ Basic 10 better oder ÄKTA™ Purifier 10 (GE HEALTHCARE) mit der Säule HisTrap HP 5 ml (GE HEALTHCARE; Säulenvolumen (SV): 5 ml; max. Druck: 0,3 MPa) bzw. HisTrap™ FF crude 1 ml (GE HEALTHCARE; Säulenvolumen (SV): 1 ml; max. Druck: 0,3 MPa) durchgeführt. Diese Säulen bestehen aus Sepharose 6 (FastFlow) Medium und dem daran gebundenem Chelator Iminodiaktische Säure (IDA). IDA besitzt sechs Koordinationsstellen, wobei an drei Koordinationsstellen Metallionen (Ni²⁺, Cu²⁺, Co²⁺ oder Fe³⁺) immobilisiert werden können, während die anderen drei für die Komplexbildung von Stickstoff- und Schwefelatomen von Histidinen, Tryptophanen oder Cysteinen zur Verfügung stehen. Proteine, die gentechnisch mit einem N- oder C-terminalem His₆-Tag fusioniert wurden, binden nach Beladung der Säule mit Metallionen (in dieser Arbeit wurde stets Ni²⁺ verwendet) sehr spezifisch an die Säule, da die Metallionen zusätzliche Koordinationsstellen für die Histidine bereitstellen. Die Elution von gebundenem Protein erfolgt durch Anlegung eines Imidazolgradienten, wobei die Imidazolmoleküle mit den chemisch äquivalenten Histidinseitenketten des Proteins um die Bindestellen an der Säule konkurrieren. In den Auftragspuffern dürfen sich keine größeren Mengen an EDTA oder DTT

befinden. EDTA würde die gebundenen Nickelionen komplexieren und damit von der Säule entfernen, während DTT die Ionen zu elementarem Nickel reduzieren würde.

Die Beladung der Säule mit Nickelionen erfolgte nach Angaben des Herstellers (siehe 2.11.4.1) und wurde bei Bedarf durchgeführt.

Die Reinigung von His₆-GST-hGR-LBD und hPCSK9³¹⁻⁴⁵⁴ Varianten wurde nach folgenden Methoden durchgeführt:

Tabelle 2: Protokolle für die Reinigung verschiedener Proteine durch Metallchelate-Affinitätschromatographie.

Reinigung von	His ₆ -GST-hGR-LBD Varianten	hPCSK9 ³¹⁻⁴⁵⁴ Varianten
verwendete Säule:	HisTrap HP 5 ml	HisTrap FF Crude 1 ml
Flussrate:	4 ml/min	1 ml/min
Äquilibrierung:	2 SV Laufpuffer ^a	2 SV Laufpuffer ^d
Beladung:	50–250 ml Proteinlösung in Aufschlusspuffer ^b mit 2 ml/min	50–100 ml Proteinlösung in Aufschlusspuffer ^e
Waschschritt:	20 SV Laufpuffer ^a	5 SV Laufpuffer ^d
Elution:	12 SV Gradient von 40–340 mM Imidazol mit Elutionspuffer ^c ; es wurden Fraktionen à 5 ml gesammelt	30 SV Gradient von 5–700 mM Imidazol mit Elutionspuffer ^f ; es wurden Fraktionen à 0,5 ml gesammelt
Reinigung der Säule:	5 SV Elutionspuffer ^c	5 SV Elutionspuffer ^f
Äquilibrierung:	5 SV Laufpuffer ^a	5 SV Laufpuffer ^d
Spülen und Lagerung der Säule:	5 SV H ₂ O; 3 SV 20 % Ethanol	5 SV H ₂ O; 3 SV 20 % Ethanol

^a: Laufpuffer für His₆-GST-hGR-LBD Reinigungen: 50 mM Tris/Cl, 500 mM NaCl, 40 mM Imidazol, 2 mM TCEP, 10 % Glycerin, pH 7,9 (\pm 50 μ M Dexamethason).

^b: Aufschlusspuffer für His₆-GST-hGR-LBD Reinigungen: 50 mM Tris/Cl, 2 M Urea, 100 mM NaCl, 5 mM MgCl₂, 40 mM Imidazol, 2 mM TCEP, DNase I (Roche) (1 mg/ 50 ml), EDTA-freier Protease Inhibitor Cocktail *complete* (Roche) (1 Tablette / 50 ml) pH 7,9 (\pm 50 μ M Dexamethason).

^c: Elutionspuffer für His₆-GST-hGR-LBD Reinigungen: 50 mM Tris/Cl, 500 mM NaCl, 340 mM Imidazol, 2 mM TCEP, 10 % Glycerin, pH 7,9 (\pm 50 μ M Dexamethason).

^d: Laufpuffer für hPCSK9³¹⁻⁴⁵⁴-Reinigungen: 50 mM Tris/Cl, 300 mM NaCl, 5 mM Imidazol, pH 8,0.

^e: Aufschlusspuffer für hPCSK9³¹⁻⁴⁵⁴-Reinigungen: 50 mM Tris/Cl, 1 mM DTT, pH 8,0.

^f: Elutionspuffer für hPCSK9³¹⁻⁴⁵⁴-Reinigungen: 50 mM Tris/Cl, 300 mM NaCl, 1 M Imidazol, pH 8,0.

Die Elution der Proteine wurde über die Messung der Absorption bei 260 und 280 nm verfolgt. Die weitere Analyse der Elutionsfraktionen erfolgte über SDS-PAGE (3.5.2) und zum Teil über analytische *reversed-phase* HPLC (3.5.6). Saubere Fraktionen wurden vereinigt. Anschließend konnte mittels Dialyse (3.4.2.5) das enthaltene Imidazol entfernt

werden. Bei gereinigten His₆-GST-hGR-LBD Varianten wurde parallel dazu der fusionierte His₆-GST-Tag enzymatisch abgespalten (3.4.2.2).

3.4.2.2 Enzymatische Entfernung des His₆-Glutathion-S-Transferase Fusionsproteins

Die His₆-Glutathion-S-Transferase (His₆-GST), die bei der Expression von Genen in pERN7 (2.8.1.2.4) N-terminal an die entsprechenden Proteine fusioniert wird, kann mit Hilfe der Protease Thrombin spezifisch abgespalten werden. Dazu wurde die nach Metallchelataffinitätschromatographie (3.4.2.1) erhaltene Proteinlösung mit 1 – 3 kU Thrombin (Sigma-Aldrich) versetzt. Der Ansatz wurde üN bei 4 °C bei gleichzeitiger Dialyse (3.4.2.5) gegen 2 l 50 mM Tris/Cl, 500 mM NaCl, 2 mM CaCl₂, 2 mM TCEP, 0.5 % CHAPS, 10 % Glycerin, pH 7,9 (±10 µM Dexamethason) inkubiert. Nach weiterer Zugabe von 1 kU Thrombin wurde der Ansatz nochmals 15 min bei RT und 30 min auf Eis inkubiert. Die Vollständigkeit der Spaltung wurde mittels SDS-PAGE (3.5.2) und zum Teil über analytische *reversed-phase* HPLC (3.5.6) überprüft.

3.4.2.3 Gekoppelte Glutathion- und Metallchelataffinitätschromatographie

Zur Entfernung von abgespaltenem His₆-GST und ungespaltenem His₆-GST-Fusionsprotein nach Metallchelataffinitätschromatographie (3.4.2.1) und enzymatischem Verdau (3.4.2.2) wurde eine gekoppelte Glutathion- und Metallchelataffinitätschromatographie durchgeführt. Hierfür wurde an der Chromatographie-Anlage ÄKTA™ Basic 10 better oder ÄKTA™ Purifier 10 (GE HEALTHCARE) eine GSTrap FF 5 ml Säule vor eine HisTrap FF Crude 1 ml (3.4.2.1) Säule geschaltet.

GSTrap Säulen bestehen aus Glutathion Sepharose 4 (FastFlow) Medium und binden die Glutathion-S-Transferase (GST) mit hoher Affinität. GST bzw. GST-Fusionsproteine binden somit sehr spezifisch an die Säule und können durch L-Glutathion, welches mit der Säule um die Bindung an GST konkurriert, wieder eluiert werden.

Die Beladung der gekoppelten GSTrap FF 5 ml - und HisTrap FF Crude 1 ml – Säulen mit der verdauten Proteinpräparation (3.4.2.2) führt dazu, dass sowohl abgespaltenes His₆-GST als auch ungespaltenes His₆-GST-Fusionsprotein effektiv an beide Säulen bindet. Das saubere abgespaltene Zielprotein befindet sich somit im Durchfluss, welcher in Fraktionen à 5 ml gesammelt wurde. Der Chromatographie-Lauf wurde über die Messung der Absorption bei 260 und 280 nm verfolgt. Nach der Beladung wurden beide Säulen gereinigt, indem His₆-GST

und ungespaltenes His₆-GST-Fusionsprotein durch L-Glutathion bzw. durch Imidazol eluiert wurde. Die Reinigung wurde nach folgender Methode durchgeführt:

Flussrate:	1 ml/min
Äquilibrierung:	2 SV Laufpuffer ^a
Beladung und Sammeln des Durchflusses:	40–120 ml verdaute Proteinlösung (3.4.2.2)
Waschschritt:	2 SV Laufpuffer ^a
Reinigung der Säule:	5 SV Elutionspuffer 1 ^b 5 SV Elutionspuffer 2 ^c
Spülen und Lagerung der Säule:	5 SV H ₂ O; 3 SV 20 % Ethanol

^a: Laufpuffer: 50 mM Tris/Cl, 500 mM NaCl, 2 mM TCEP, 0.25 % CHAPS, 10 % Glycerin, pH 7,9 (± 10 µM Dexamethason).

^b: Elutionspuffer 1: 50 mM Tris/Cl, 500 mM NaCl, 2 mM TCEP, 0.25 % CHAPS, 10 % Glycerin, 10 mM L-Glutathion, pH 7,9 (± 10 µM Dexamethason).

^c: Elutionspuffer 2: 50 mM Tris/Cl, 500 mM NaCl, 340 mM Imidazol, 2 mM TCEP, 10 % Glycerin, pH 7,9 (± 50 µM Dexamethason).

Die Analyse der Durchfluss-Fractionen erfolgte mittels SDS-PAGE (3.5.2) und zum Teil über analytische *reversed-phase* HPLC (3.5.6). Saubere Fractionen wurden vereinigt und anschließend mit Hilfe präparativer Gelfiltration (3.4.2.4) weiter verarbeitet.

3.4.2.4 Präparative Gelfiltration

Diese Methode beruht auf dem Prinzip eines umgekehrten Molekularsiebes (Ausschluss-Chromatographie). Die auf die Säule aufgetragenen Moleküle wandern bei konstanter Flussrate durch das poröse Säulenmaterial. Kleine Moleküle wandern aufgrund des für sie größer vorhandenen Diffusionsraumes langsamer als große Moleküle. Aufgetragene Proteine werden somit im präparativen Maßstab ihrer Größe nach aufgetrennt.

Die präparative Gelfiltration wurde bei RT an einer ÄKTA™ Purifier - Anlage (GE HEALTHCARE) oder im Kühlraum bei 4 °C an einer FPLC LCC-501 - Anlage (GE HEALTHCARE) durchgeführt. Die verwendete Säulen Superdex 200 pg HiLoad 16/60 bzw. Superdex 200 pg HiLoad 26/60 (GE HEALTHCARE; Säulenvolumen (SV): 120 bzw. 320 ml) bestehen aus quervernetzter Agarose und Dextran mit einer Partikelgröße von 24 - 44 µM. Die aufgetragene Probe durfte nicht mehr als 5 ml betragen, so dass sie gegebenenfalls einkonzentriert werden musste (3.4.2.6), bevor sie mit Hilfe eines Superloops auf die mit entgastem Laufpuffer äquilibrierte Säule aufgetragen wurde.

Die präparative Gelfiltration wurde als finaler Reinigungsschritt für hGR-LBD Varianten genutzt, wobei noch vorhandenes Dexamethason ebenfalls größtenteils entfernt wurde. Hierbei wurde die Proteinpräparation aus der gekoppelten Glutathion- und Metallchelataffinitätschromatographie (3.4.2.3) auf 5 ml einkonzentriert und auf die mit 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9 äquilibrierte Säule aufgetragen. Anschließend wurde mit 150 ml (Superdex 200 pg HiLoad 16/60) bzw. 360 ml (Superdex 200 pg HiLoad 26/60) desselben Puffers eluiert, wobei die Laufgeschwindigkeit 1 ml/min betrug (Rückdruck: max. 0,3 mPa). Der Gelfiltrationslauf wurde über die Absorption bei 280 nm und 260 nm (ÄKTA™ Purifier) bzw. bei 280 nm (FPLC LCC-501) verfolgt. Es wurden Fraktionen mit einem Volumen von je 4 ml gesammelt. Die Analyse der Fraktionen erfolgte über SDS-PAGE (3.5.2) und zum Teil über analytische *reversed-phase* HPLC (3.5.6).

3.4.2.5 Dialyse von Proteinlösungen

Falls bei einer Proteinlösung der Puffer gewechselt oder Salze entfernt werden mussten, wurde gegen einen ausreichenden Volumenüberschuss des Puffers bei 4 °C im Kühlraum dialysiert. Hierzu wurde ein Dialyseschlauch (Visking) mit einer molekularen Ausschlussgrenze von 14 kDa verwendet, welcher das Protein zurückhält, während niedermolekulare Substanzen die Membran passieren können.

3.4.2.6 Einkonzentrieren von Proteinlösungen

Die Proteinlösungen wurden mit Hilfe der Amicon® Zentrifugenröhrchen (Amicon Ultra 4 bzw. 15; molekulare Ausschlussgrenze: je 10 kDa) über eine semipermeable Membran nach Herstellerangaben (Eppendorf Centrifuge 5810 R, 4000 rpm, 4 °C) einkonzentriert. Die Dauer der Zentrifugation richtete sich nach dem Volumen des einzukonzentrierenden Proteins.

3.4.2.7 Lagerung von gereinigten Proteinen

Die gereinigten und konzentrierten Proteine wurden in flüssigen Stickstoff getropft und die gefrorenen Kügelchen bei -80 °C gelagert.

3.5 Analytische Methoden

3.5.1 Konzentrationsbestimmung von Proteinen

3.5.1.1 Bestimmung der Proteinkonzentration nach Bradford (1976)

Der Bradfordtest ist ein Assay zur Bestimmung des Gesamtproteingehalts von wässrigen Lösungen. Er beruht darauf, dass sich das Absorptionsmaximum einer sauren *Coomassie Brilliant Blue* G250-Lösung bei Bindung des Farbstoffes an Proteine von 465 auf 595 nm verschiebt (Bradford, 1976). Es handelt sich um eine unspezifische Bindung des Farbstoffes vorwiegend an kationische, nichtpolare und hydrophobe Seitenketten. Die Intensitätszunahme der Farbreaktion ist über einen Bereich von 0,2 - 1,4 mg Protein pro ml Lösung annähernd linear. Es wurde der Protein-Assay Kit von BIO-RAD verwendet. Dabei wurden 200 µl des *Coomassie*-Reagenz mit 800 µl der Proteinlösung vermischt und nach ca. 5 min die Absorption der Lösung bei 595 nm gemessen. Als Referenz wurde eine Kalibrierungsgerade mit bovinem Serumalbumin (BSA) erstellt. Es ist zu beachten, dass der Assay durch SDS (> 0,1 %) und starke Basen beeinflusst werden kann. Das Ergebnis ist jedoch auch abhängig von der Aminosäurezusammensetzung des Proteins, so dass die aus der BSA-Kalibrierungskurve abgeleiteten Proteinkonzentrationen immer nur einen Anhaltspunkt darstellen.

3.5.1.2 Bestimmung der Proteinkonzentration durch Absorptionsspektroskopie

Die aromatischen Aminosäuren Tryptophan, Tyrosin und Phenylalanin sowie Disulfidbrücken (Cystine) absorbieren unterschiedlich stark UV-Licht in einem Wellenlängenbereich von 250 bis 300 nm. Aus der Aminosäuresequenz kann nach Pace *et al.* (1995) der molare Extinktionskoeffizient ϵ_{280} (Formel 10) und daraus, unter Einbeziehung des Molekulargewichtes, der spezifische Extinktionskoeffizient ($^{0,1\%}A_{280}$) (Formel 11) errechnet werden.

$$\epsilon_{280} = \sum \text{Trp} \cdot 5500 + \sum \text{Tyr} \cdot 1490 + \sum \text{Cystine} \cdot 125$$

Formel 10: Berechnung des molaren Extinktionskoeffizienten ϵ_{280} .

ϵ_{280} : molarer Extinktionskoeffizient bei 280 nm [$M^{-1}cm^{-1}$]

$$^{0,1\%}A_{280} = \epsilon_{280} / MW$$

Formel 11: Berechnung des spezifischen Extinktionskoeffizienten $^{0,1\%}A_{280}$.

$^{0,1\%}A_{280}$: spezifischer Extinktionskoeffizient bei 280 nm [cm^2/mg]

MW: Molekulargewicht des Proteins [g/mol]

Damit kann mit Hilfe des Lambert-Beer'schen Gesetzes die Proteinkonzentration durch Absorptionsmessung bei 280 nm bestimmt werden (Formel 12).

$$A_{280} = \epsilon_{280}^{0,1\%} \cdot c \cdot d$$

$$\rightarrow c = A_{280} / (\epsilon_{280}^{0,1\%} \cdot d)$$

Formel 12: Berechnung der Proteinkonzentration mit Hilfe des spezifischen Extinktionskoeffizienten $\epsilon_{280}^{0,1\%}$.

A_{280} : gemessene Absorption bei 280 nm

c: Konzentration [mg/ml]

d: Schichtdicke der Küvette [cm]

$\epsilon_{280}^{0,1\%}$: spezifischer Extinktionskoeffizient bei 280 nm [cm²/mg]

Es wurden Spektren von 220 – 350 nm aufgenommen, wobei sich das Absorptionsmaximum des Proteins bei 278 nm befinden sollte. Der Quotient A_{280}/A_{250} sollte mindestens 1,8 betragen und es sollte keine Absorption oberhalb von 300 nm auftreten, um eine Verfälschung des Ergebnisses durch auf Aggregatbildung zurückgehende Lichtstreuung ausschließen zu können. Bei deutlicher Absorption oberhalb von 300 nm wurde als erste Näherung eine lineare Zunahme der Lichtstreuung mit abnehmender Wellenlänge angenommen und die ermittelte Proteinkonzentration entsprechend korrigiert.

3.5.2 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE)

Proteine werden durch das Detergenz Natriumdodecylsulfat (SDS) denaturiert und negativ geladen. Dabei bindet SDS an das Protein gewöhnlich in einem Verhältnis von einem Molekül SDS pro 1,4 Aminosäureresten. Die Eigenladung des Proteins kann neben der Negativladung des mit SDS komplexierten Proteins vernachlässigt werden, so dass SDS-beladene Proteine identische Ladungs-Masse-Verhältnisse zeigen. Die Auftrennung beruht daher allein auf dem Siebeffekt, wobei die Laufgeschwindigkeit dem Logarithmus der Masse umgekehrt proportional ist (Laemmli, 1970). Tabelle 3 zeigt die Zusammensetzung der verwendeten 12,5 %igen SDS-Gele.

Tabelle 3: Zusammensetzung 12,5 %iger SDS-PAGE Gele.

Die Mengenangaben gelten für 13 Gele.

	Trenngel (12,5 %)	Sammelgel (12,5 %)
Trenn-/Sammelgel-Puffer	19,5 ml	7,38 ml
Acryamid-SL (30 %)	26,2 ml	5,9 ml
H ₂ O	31,58 ml	15,95 ml
TEMED	0,089 ml	0,029 ml
APS (10 %)	0,195 ml	0,089 ml

Die Proben wurden mit 2 x oder 5 x SDS- Probenpuffer versetzt und 5 min bei 95 °C inkubiert. Es wurden je 5–25 µl der Proben mit *Gelloader*-Spitzen in die Geltaschen gefüllt. Die Elektrophorese wurde bei 40 mA und 300 V durchgeführt und dauerte ca. 30 min. Anschließend wurden die Gele mit SDS-PAGE-Färbelösung gefärbt, wobei die Nachweisgrenze des Farbstoffs Coomassie Brilliant Blue G250 200–500 ng Protein/mm² beträgt. Die Gele wurden 30 min in der Färbelösung geschwenkt und im Anschluss wurde durch wiederholtes Aufkochen in Wasser (Mikrowelle 900 W) die Hintergrundfärbung entfernt.

Zum Teil wurden auch bereits hergestellte NuPAGE[®] Novex Bis-Tris Gele (4-12 %) des X-Cell *SureLock*[™] Mini Cell Systems von Invitrogen verwendet, welche mit Instant Blue (Expedeon, Lucerna-Chem AG) - Lösung gefärbt wurden.

3.5.3 Analyse der Proteinlöslichkeit *in vivo* durch Densiometrie

Zur Analyse der Proteinlöslichkeit *in vivo* wurden Expressionen im analytischen Maßstab durchgeführt. Die Präparation der Proben erfolgte wie in 3.4.1.1 beschrieben, wobei von den Expressionskulturen stets jeweils 2×10^9 Zellen entnommen und analysiert wurden. Zur Quantifizierung wurde nur die gereinigte lösliche Zellfraktion (C*) verwendet, da die lösliche Zellfraktion (C) oftmals zu viele Fremdproteine für eine genaue Bestimmung der Zielproteinmenge enthielt.

Die gereinigten löslichen Zellfraktionen (C*) der verschiedenen Proteinvarianten und des Wildtyps wurden jeweils in unterschiedlichen Verdünnungen auf ein SDS-Polyacrylamidgel aufgetragen (3.5.2) und das Gel mit einer frischen Coomassie Blue-Färbelösung gefärbt. Die aufgetragenen Proteinmengen konnten anschließend durch Quantifizierung der Proteinbanden mit der Software OptiQuant (Version 02.50) bestimmt werden. Um Quantifizierungsfehler durch nichtlineare Färbung ausschließen zu können, wurden nur Banden mit ähnlicher Intensität aber unterschiedlicher Verdünnung verglichen, woraufhin auf die unverdünnten Lösungen zurückgerechnet werden konnte.

Pro Proteinvariante wurden stets drei Präparationen mit je mindestens zwei verschiedenen SDS-PA Gelen analysiert, wodurch die Bestimmung von Mittelwerten und Standardabweichungen möglich war.

Die quantifizierte Menge der löslichen Proteinvariante in der C* - Zellfraktion wurde auf die Wildtyp-Menge normiert und diente als Maß für die Proteinlöslichkeit *in vivo*.

3.5.4 Western-Blotting

Die durch die SDS-PAGE (3.5.2) aufgetrennten Proteinproben wurden mit Hilfe eines Tankblotters (Peqlab) vom SDS-Gel auf eine PVDF-Membran übertragen. Die Polyvinyliden-Fluorid-Membran (PVDF) besteht aus einem Polymer aus $-(CH_2-CF_2)-$ Untereinheiten und ermöglicht die Bindung von Protein über hydrophobe Wechselwirkungen. Gegenüber der traditionell verwendeten Nitrocellulose hat sie einige Vorteile, unter anderem ist sie robuster. Die verwendete Membran Immobilon-P (Millipore) hat eine Porengröße von $0,45\ \mu\text{m}$ und wurde gemäß Herstellerangaben vor Gebrauch mit 100 % Methanol für 1 min aktiviert. Nach dem Zuschneiden wurde die Membran zusammen mit dem SDS-PA Gel und Filterpapieren im Transfer-Puffer für 10 min inkubiert und wie in Abbildung 22 dargestellt in der Tankblot-Apparatur angeordnet.

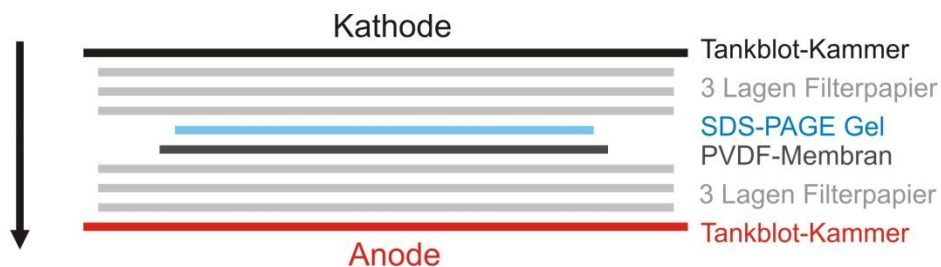


Abbildung 22: Anordnung in der Tankblot-Apparatur.

Der Pfeil deutet die Transferrichtung der Proteine an.

Der Transfer erfolgte bei 220 mA für 1 h. Der Erfolg des Transfers war aufgrund der Verwendung des vorgefärbten Standards Prestained PageRuler™ Protein Ladder Marker (MBI FERMENTAS, St. Leon-Rot) (2.10) ersichtlich.

Die auf der Membran frei zugänglichen Proteine wurden mit Hilfe von spezifischen Antikörpern sichtbar gemacht. Dazu wurden zuerst unspezifische Bindungsstellen durch Inkubation der Membran üN bei $4\ ^\circ\text{C}$ in Blockierungspuffer (1x PBS, 0,1 % Tween-20 = PBS/T + 5 % Milchpulver) geblockt. Anschließend wurde die Membran mit dem in 20 ml PBS/T 1:1000 verdünnten primären Antikörper (Anti-GFP-Antikörper *mouse* monoklonal; siehe 2.6) 90 min bei RT inkubiert. Nach 3-maligen Waschen der Membran für 10 min mit PBS/T, wurde sie mit dem in PBS/T + 1 % Milchpulver 1:5000 verdünnten sekundären Antikörper (Peroxidase-konjugierter anti-*mouse* IgG; siehe 2.6) 60 min bei RT inkubiert. Nachdem die Membran nochmals 3 x für 10 min mit PBS/T gewaschen wurde, erfolgte die spezifische Detektion mit Hilfe des SuperSignal West Pico Kits (Pierce) nach Anleitung. Der sekundäre Antikörper ist an die Meerrettich-Peroxidase

konjugiert und diese kann Luminol als Substrat (im SuperSignal West Pico-Kit enthalten) umsetzen. Dabei wird Luminol oxidiert zu 3-Aminophthalat, das bei 425 nm luminesziert. Diese Lichtsignale wurden durch Auflegen und Entwickeln von Röntgenfilmen (Fuji) sichtbar gemacht.

3.5.5 Analytische Gelfiltration

Mit Hilfe der analytischen Gelfiltration lässt sich das apparente Molekulargewicht und somit auch der Oligomerisierungsgrad eines Proteins bestimmen. Das Trennprinzip entspricht dem der präparativen Gelfiltration (3.4.2.4), wobei ein linearer Zusammenhang zwischen der Elutionsdauer und dem Logarithmus des Molekulargewichtes besteht. Es wurde die Säule Superdex 75 HR 10/30 (GE HEALTHCARE) mit einem Säulenvolumen von 24 ml an der Chromatographie-Anlage ÄKTA Basic 10 better (GE HEALTHCARE) verwendet. Auf die mit Laufpuffer (25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 20 µM Dexamethason, pH 7,9) äquilibrierte Säule wurden 100 – 500 µl Proteinlösung einer Konzentration von 1 – 14 µM über einen Superloop aufgetragen. Die Elution erfolgte mit einer Flussrate von 0,5 ml/min und wurde durch Absorptionsmessung bei 280 nm und 260 nm verfolgt. Mit Hilfe einer Kalibrierungsgerade, die mit Proteinen bekannter Größe erstellt wurde, konnte das apparente Molekulargewicht (MW_{app}) der analysierten Proteine ermittelt werden.

3.5.6 Analytische *reversed-phase* HPLC

Mit Hilfe von HPLC (*High Performance Liquid Chromatography*) können die in einer Lösung enthaltenen Komponenten mit einer sehr hohen Empfindlichkeit und Trennschärfe analysiert werden. Hierfür wird die Probe chromatographisch auf einer Säule getrennt und durch Absorptions- oder Fluoreszenzmessung ein Elutionsprofil des Laufs erstellt. Im Gegensatz zur FPLC (*Fast Protein Liquid Chromatography*) (z.B. ÄKTA Basic 10 better) erfolgt die Chromatographie bei höherem Druck (bis 400 bar) und mit Säulen niedrigeren Durchmessers (2 – 4 mm) und kleinerer Partikelgröße (3 – 10 µm). Dadurch wird im Vergleich zur herkömmlichen Säulenchromatographie das Trennvermögen etwa 100 x besser. Unterschiedliche Substanzen in einer Probe zeichnen sich durch verschiedene Retentionszeiten aus, über die sie dann im Vergleich zu Standardsubstanzen identifiziert werden können.

Im Rahmen dieser Arbeit wurde das Ergebnis verschiedener Schritte der Proteinreinigung mit Hilfe einer *reversed phase* – Säule analysiert, um den erreichten Reinheitsgrad der

Proteinpräparationen exakt bestimmen zu können. Bei *reversed phase* –Säulen handelt es sich um Kieselgele, die mit langen Kohlenwasserstoffketten modifiziert wurden. Dadurch erhält man eine sehr unpolare Oberfläche, an die Substanzen durch hydrophobe Wechselwirkungen binden können. Zur Elution wurde eine Gradiententrennung mit einem wässrigen Puffer und einem organischen Lösungsmittel verwendet. Durch Zumischen des organischen Lösungsmittels zur wässrigen Phase sinkt die Polarität des Laufmittels, wodurch seine Elutionskraft für unpolare Substanzen steigt. Die an die Säule gebundenen Substanzen eluieren also gemäß ihrer Polarität, wobei sich polare Stoffe vor den Unpolaren von der Säule lösen. Der *reversed-phase* HPLC Lauf wurde an einer Agilent 1200-Anlage mit einer C8 (Poroshell 2.1x50 mm) – Säule durchgeführt, wobei die Absorption bei 215, 260 und 280 nm verfolgt wurde. Bezüglich der Peaks stehen ganze Spektren zur Verfügung, da ein Dioden-Array verwendet wurde. Nach Auftrag von 2 µl Probe erfolgte die Auftrennung nach folgendem Protokoll:

Flussrate:	3 ml/min
Äquilibrierung:	90 % polares Laufmittel ^a , 10 % unpolares Laufmittel ^b
Elution:	Gradient bis 40 % polares Laufmittel ^a , 60 % unpolares Laufmittel ^b
Spülen:	25 % polares Laufmittel ^a , 75 % unpolares Laufmittel ^b
Temperatur:	70 °C

^a: polares Laufmittel: H₂O, 0,1 % TFA

^b: unpolares Laufmittel: Acetonitril, 0,08 % TFA

3.5.7 Circular-Dichroismus-Spektroskopie

Circular-Dichroismus-Spektroskopie (CD-Spektroskopie) ist eine spezielle Form der UV/Vis-Spektroskopie, die auf Wechselwirkungen von zirkular polarisiertem Licht mit optisch aktiven Substanzen, wie Proteinen, beruht. Optisch aktive Substanzen besitzen unterschiedliche Brechungsindizes für rechts- und links-polarisiertes Licht, was zu unterschiedlichen Verzögerungen beim Durchgang der Lösung und somit zu einer Phasenverschiebung ϕ führt. Weiterhin unterscheiden sich bei optisch aktiven Substanzen die Extinktionskoeffizienten für rechts- und links-polarisiertes Licht, was beim Durchgang der Lösung zu elliptisch polarisiertem Licht führt. Die Elliptizität (θ_{obs}) wird bei CD-Messungen als Signal verwendet.

Die CD-Spektroskopie eignet sich besonders zu Untersuchungen von Proteinkonformationen. Die Absorption im Wellenlängenbereich von 170 bis 250 nm (Fern-UV CD) wird hauptsächlich von den Peptidbindungen bestimmt, so dass dieser Teil des Spektrums Aussagen über die Sekundärstruktur ermöglicht. Die verschiedenen Sekundärstrukturelemente wie α -Helix, β -Faltblatt oder *Random-coil* besitzen sehr distinkte Fern-UV CD-Spektren (Johnson *et al.*, 1990). Aus diesen verschiedenen Signalen setzt sich ein Proteinspektrum zusammen (Stark, 1965), wobei α -Helices die anderen Strukturelemente in der Signalintensität dominieren. Als charakteristische Merkmale eines Fern-UV CD-Spektrums gelten Maxima, Nulldurchgang und Minima. Die Absorption im Bereich von 250 bis 320 nm (Nah-UV CD) ist auf die aromatischen Aminosäuren zurückzuführen. Kleine Änderungen in der Umgebung eines Chromophors führen hierbei zu einer deutlichen Änderung des Signals. Das Nah-UV CD-Spektrum ist somit eine empfindliche Methode zur Untersuchung der Tertiärstruktur von Proteinen. In dieser Arbeit wurden ausschließlich Fern-UV CD-Spektren aufgenommen, um die Integrität der Sekundärstrukturen von Proteinvarianten zu kontrollieren.

Für das CD-Signal von Proteinen in der Amidregion (Fern-UV) findet eine Normierung auf die durchschnittliche molare Elliptizität pro Aminosäure (θ_{MRW}) nach Formel 13 statt (Schmid, 1997).

$$\theta_{\text{MRW}} = \frac{\theta_{\text{obs}} \cdot 100 \cdot \text{MRW}}{c \cdot d} = \frac{\theta_{\text{obs}} \cdot 100 \cdot \text{MW}}{c \cdot d \cdot N_{\text{A}}}$$

Formel 13: Berechnung der molaren Elliptizität pro Aminosäure (θ_{MRW}).

- θ_{MRW} : durchschnittliche Elliptizität pro Aminosäure [$\text{deg cm}^2 \text{dmol}^{-1}$]
- θ_{obs} : gemessene Elliptizität [mGrad]
- MRW : mittlere Molmasse der Aminosäurenreste (*mean residue weight*) [kDa]
- c : Konzentration der Proteinlösung [mg/ml]
- d : Schichtdicke der Küvette [cm]
- MW : Molare Masse des Proteins [kDa]
- N_{A} : Anzahl der Aminosäuren

Alle CD-Messungen wurden in Quarzküvetten der Schichtdicke 1 mm durchgeführt. Zur Aufnahme von Fern-UV CD Spektren wurde das CD-Signal, gemessen als Elliptizität θ_{obs} , in Abständen von 1,0 nm aufgenommen (Scan-Geschwindigkeit: 100 nm/min). Jede Messung wurde mindestens 10-mal wiederholt und die Spektren anschließend gemittelt. Zusätzlich wurde stets ein reines Pufferspektrum aufgenommen, welches vom Proteinspektrum abgezogen wurde. Die Werte wurden anschließend wie oben beschrieben in molare Elliptizität pro Aminosäure θ_{MRW} umgewandelt.

3.5.8 Fluoreszenz-Spektroskopie

Absorbiert ein Molekül Licht bestimmter Wellenlänge, so wird es von einem energetischen Grundzustand in einen angeregten Zustand transferiert. Die Rückkehr in den Grundzustand kann unter anderem durch die Emission eines Photons geschehen. Dieser Vorgang ist als Fluoreszenz messbar. Die Energie eines emittierten Photons ist geringer als die eines absorbierten, was einer Verschiebung zu größerer Wellenlänge, also einer Rotverschiebung, entspricht. Die Fluoreszenz von Proteinen beruht im Wesentlichen auf den aromatischen Aminosäuren Tryptophan und Tyrosin, wobei das Tryptophan eine größere Rolle spielt. Aromatische Seitenketten absorbieren maximal bei einer Wellenlänge von etwa 280 nm. Nach Anregung bei 295 nm ist die Fluoreszenzemission ausschließlich auf Tryptophane zurückzuführen. Die Fluoreszenz von Tryptophan zeigt eine Abhängigkeit von der Polarität der Umgebung. Eine stark polare Umgebung bewirkt eine Rotverschiebung der Emission zu einem Maximum bei ca. 350 nm gegenüber einem Maximum bei 320 – 330 nm in unpolarer Umgebung. Über die Messung von Fluoreszenzspektren kann also eine Aussage über den Grad der Lösungsmittelexposition eines Tryptophans in einem Protein gemacht werden. Daraus wiederum kann auf den Faltungszustand eines Proteins geschlossen werden (Schmid, 1997). Die Messungen wurden am Fluoreszenz-Spektrophotometer Cary Eclipse (Varian, Darmstadt) in einer vierseitig geschliffenen Quarzküvette durchgeführt. Die Fluoreszenz wurde bei einer Wellenlänge von 280 nm (Tryptophan- und Tyrosinfluoreszenz) oder 295 nm (nur Tryptophanfluoreszenz) angeregt und das emittierte Signal in einem Bereich von 300 bis 400 nm aufgenommen. Die Spaltbreite betrug bei der Excitation und der Emission jeweils 5 nm.

3.5.9. Refraktometrie

Die genaue Konzentration von Detergenz- bzw. Salz-Stammlösungen wurde vor deren Verwendung bei Auffaltungs- (3.5.10) bzw. fraktionierten Präzipitations- (3.5.11) Experimenten bestimmt. Hierfür wurde mit Hilfe eines Abbé Refraktometers (Carl Zeiss, Jena) die Differenz der Brechungsindizes (Δn) der Lösungen mit und ohne Detergenz bzw. Salz bestimmt, wodurch die Konzentration mit Hilfe der folgenden Formeln berechnet werden konnte.

$$c(\text{GdmCl}) [\text{M}] = 57,147 \cdot \Delta n + 38,68 \cdot (\Delta n)^2 - 91,60 \cdot (\Delta n)^3$$

Formel 14: Refraktometrische Bestimmung der Guanidiniumchlorid-Konzentration (Pace, 1986).

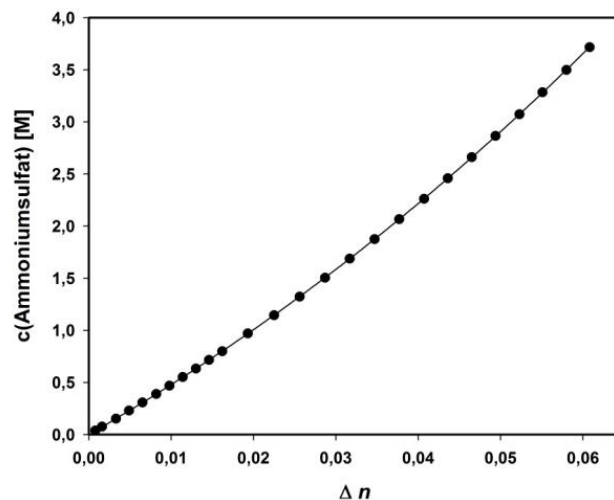
Δn : Differenz der Brechungsindizes der Lösung mit und ohne GdmCl.

$$c(\text{Harnstoff}) [\text{M}] = 117,66 \cdot \Delta n + 29,753 \cdot (\Delta n)^2 + 185,56 \cdot (\Delta n)^3$$

Formel 15: Refraktometrische Bestimmung der Harnstoff-Konzentration (Pace, 1986).

Δn : Differenz der Brechungsindizes der Lösung mit und ohne Harnstoff.

Zur Bestimmung der Ammoniumsulfat-Konzentration konnte aus der tabellarischen Auflistung von konzentrationsabhängigen Δn -Werten aus dem CRC Handbook of Chemistry and Physics (Weast, 1982-1983) über einen polynomischen Fit ($y = a \cdot x + b \cdot x^2 + c \cdot x^3$) (Abbildung 23) ebenfalls eine Formel abgeleitet werden.



$$c(\text{Ammoniumsulfat}) [\text{M}] = 45,624 \cdot \Delta n + 226,389 \cdot (\Delta n)^2 + 476,336 \cdot (\Delta n)^3$$

Abbildung 23: Refraktometrische Bestimmung der Ammoniumsulfat-Konzentration.

Einzelne Datenpunkte wurden aus Weast (1982-1983) entnommen.

Δn : Differenz der Brechungsindizes der Lösung mit und ohne Ammoniumsulfat.

3.5.10 Bestimmung der Proteinstabilität durch schrittweise Denaturierung

Zur Untersuchung der Stabilität von Proteinen wurden diese durch Erhöhung der Temperatur bzw. durch Zugabe von steigenden Konzentrationen an chaotropen Reagenzien schrittweise denaturiert. Zur chemischen Denaturierung wurde in dieser Arbeit stets Harnstoff verwendet, dessen denaturierende Wirkung seit über 100 Jahren bekannt ist und genutzt wird (Spiro, 1900; Ramsden, 1902). Der Auffaltungsvorgang der Proteine wurde durch CD-Spektroskopie (3.5.7) oder Fluoreszenz-Spektroskopie (3.5.8) verfolgt.

Durch CD-Spektroskopie wurde die thermische Auffaltung der Sekundärstruktur von Proteinen verfolgt. Hierfür wurde mit Hilfe des im JASCO J-815 CD-Spektrometers vorhandenen PTC-423S/15-Peltiers die Proteinlösung in einer Konzentration von 2,5 – 10 μM mit einer Rate von 1 $^{\circ}\text{C} / \text{min}$ aufgeheizt. Die Messungen wurden in Quarzküvetten mit einer

Schichtdicke von 1 mm durchgeführt, wobei die Temperatur im Küvettenhalter detektiert wurde. Die Auffaltung wurde stets bei der für α -Helices typischen Wellenlänge von 220 nm verfolgt, da dort der Signalunterschied zwischen nativem und aufgefaltetem Protein sehr groß ist (Abbildung 24a). Falls die Auffaltungskurven aufgrund niedriger Proteinkonzentration (2,5 μ M) starkes Signalrauschen im nativen und denaturierten Bereich zeigten, wurden sie mindestens dreimal wiederholt und gemittelt.

Mit Hilfe von Fluoreszenz-Spektroskopie wurde die chemische Auffaltung der Tertiärstruktur von Proteinen verfolgt. Hierfür wurde aus einer Proteinlösung, einer Pufferstammlösung und einer hochkonzentrierten Harnstoff-Lösung in gleichem Puffer eine Messreihe mit gleich bleibender Proteinkonzentration (3 μ M) und ansteigenden Harnstoffkonzentrationen von 0 – 8 M erstellt. Die Auffaltungsexperimente fanden bei Raumtemperatur statt. Die Einstellung des Gleichgewichts zwischen nativem und aufgefaltetem Zustand des Proteins wurde nach verschiedenen Zeitpunkten überprüft. Sobald die Messpunkte sich nicht mehr veränderten, galt das Gleichgewicht als erreicht. Nach Anregung bei 280 nm wurde von den einzelnen Proben der Messreihe die relative Intensität der Emission bei derjenigen Wellenlänge gemessen, bei der die Differenz zwischen nativem und vollständig denaturiertem Protein am größten war (Abbildung 24b). Jede Probe wurde 1 min lang im Fluoreszenz-Spektrophotometer vermessen und die erhaltene Intensität gemittelt.

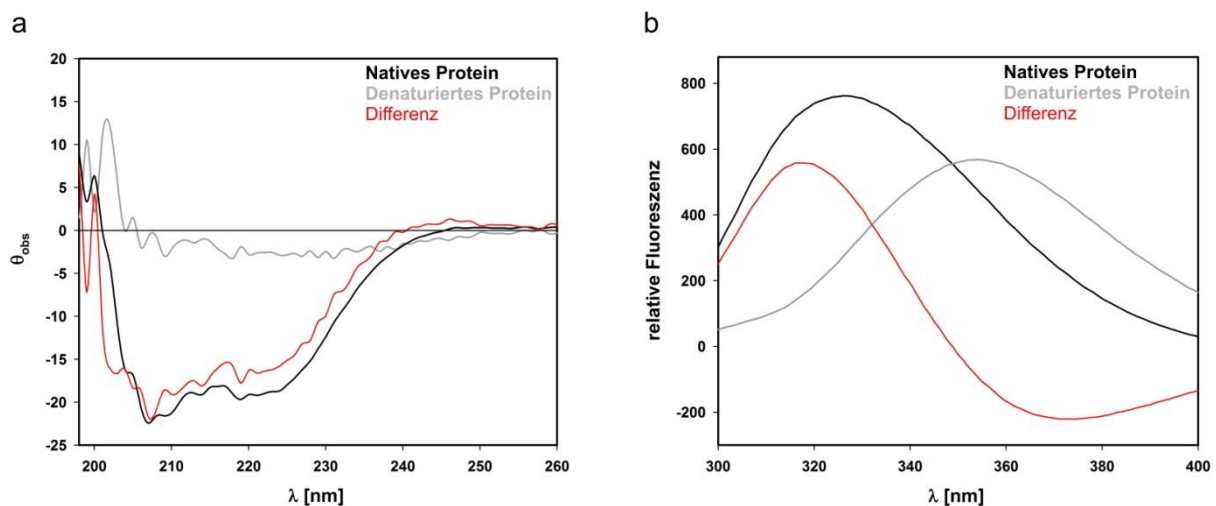


Abbildung 24: Vergleich der CD- (a) und Fluoreszenz-Spektren (b) von nativem und denaturiertem Protein.

a: Die CD-Spektren von nativem (25 °C) und denaturiertem (75 °C) Protein (hier gezeigt: hGR-LBD (F602S+A605V+V702A+E705G+M752T)) wurden wie in 3.5.7 beschrieben aufgenommen. Das Differenzspektrum ist in rot dargestellt. Die Auffaltung der Sekundärstruktur wurde bei 220 nm verfolgt.

b: Emissions-Fluoreszenzspektren von nativem (0 M Harnstoff) und denaturiertem (8 M Harnstoff) Protein (hier gezeigt: HisF-C**C, Seitz *et al.* (2007)). Die Spektren wurden nach Anregung mit einer Wellenlänge von 280 nm aufgenommen. Das Differenzspektrum ist in rot dargestellt und weist hier ein Maximum bei 317 nm auf, weshalb diese Wellenlänge für die Verfolgung der Tertiärstruktur-Auffaltung benutzt wurde.

Sowohl bei thermischen als auch bei chemischen Auffaltungsexperimenten wurden die erhaltenen Messwerte anschließend normiert und als Funktion der Temperatur bzw. der Konzentration an Denaturierungsmittel aufgetragen. Hierfür wurde der Zweizustands-Faltungsmechanismus angenommen, bei dem das Protein nur in der gefalteten (N) oder ungefalteten (U) Konformation vorliegen kann, Faltungsintermediate werden hierbei nicht berücksichtigt (1.1.1). Die Anteile an gefaltetem Protein (F_N) und ungefaltetem Protein (F_U) summiert sich damit zu:

$$F_N + F_U = 1$$

Formel 16: Anteil an nativem (F_N) und denaturiertem (F_U) Protein unter Annahme des Zweizustands-Modells der Auffaltung.

Somit werden die Signalwerte y_N von vollständig nativem bzw. y_U von vollständig denaturiertem Protein als 0 %-ige bzw. 100 %-ige Auffaltung des Proteins definiert und jeder Signalwert y lässt sich durch Formel 17 beschreiben.

$$y = y_N \cdot F_N + y_U \cdot F_U$$

Formel 17: Zusammensetzung des Signalwertes y unter Annahme des Zweizustands-Modells der Auffaltung.

y_N und y_U stellen den für N bzw. U (bei der jeweiligen Konzentration an Denaturierungsmittel) typischen Signalwert dar.

Durch Kombination von Formel 16 und Formel 17 ergibt sich für F_U :

$$F_U = \frac{(y_N - y)}{(y_N - y_U)}$$

Formel 18: Bestimmung des Anteils an ungefaltetem Protein aus den bei einem Auffaltungsexperiment gemessenen Signalintensitäten, unter Annahme des Zweizustands-Modells der Auffaltung.

F_U : Fraktion an ungefaltetem Protein
 y_N : Signalwert des nativen Proteins
 y_U : Signalwert des denaturierten Proteins
 y : Messwert

Die Änderung der y -Werte im Prä- bzw. Postübergangsbereich zeigt die Abhängigkeit von y_N bzw. y_U von der Temperatur bzw. der Denaturierungsmittelkonzentration an. Diese Abhängigkeit war annähernd linear, weshalb sich y_N und y_U bei der jeweiligen Temperatur bzw. Harnstoff-Konzentration durch lineare Regression bestimmen ließen. Dazu wurde eine

Ausgleichsgerade durch die Datenpunkte im Prä- sowie im Postübergangsbereich gelegt (Abbildung 25).

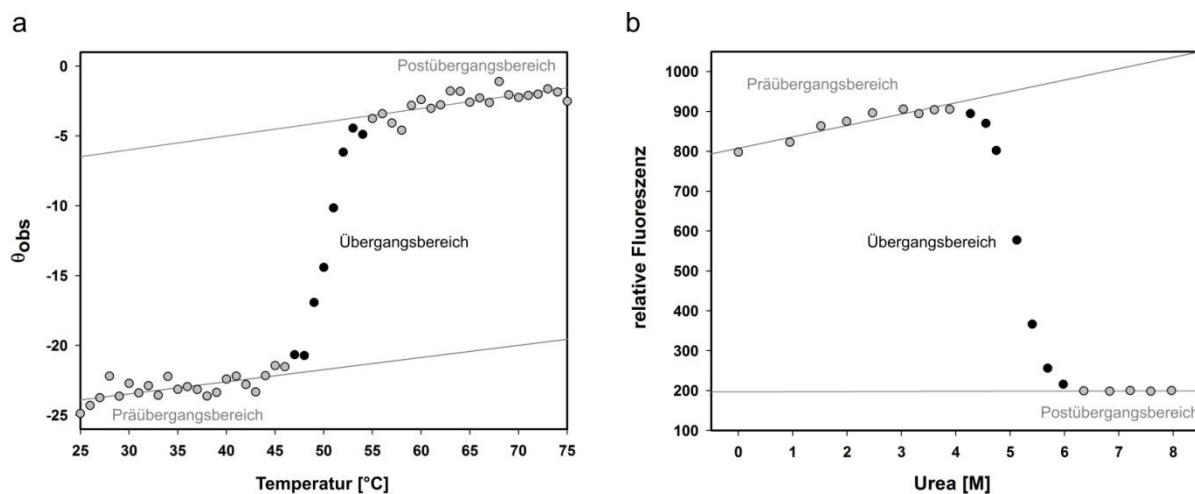


Abbildung 25: Schematische Darstellung der Auswertung eines thermischen (a) bzw. chemischen (b) Auffaltungsexperiments.

Die Werte stammen aus CD-spektroskopischen (a) bzw. fluoreszenzspektroskopischen (b) Messungen. Die grau markierten Punkte zeigen den Prä- sowie den Postübergangsbereich. Durch lineare Regression dieser Bereiche kann die Temperatur- bzw. die Denaturierungsmittel-Abhängigkeit der ermittelten Signalwerte von N bzw. U bestimmt werden.

Mit Hilfe der ermittelten Geradengleichungen im Prä- und Postübergangsbereich wurde für jede Temperatur bzw. Denaturierungsmittelkonzentration der theoretische Signalwert von nativem (y_N) und ungefaltetem Protein (y_U) bestimmt. Anschließend wurde für jeden Messwert gemäß Formel 18 der Anteil an ungefaltetem Protein (F_U) berechnet.

Falls die Denaturierung eines Proteins in erster Näherung tatsächlich mit dem Zwei-Zustands-Modell beschrieben werden kann, ist es möglich auf die freie Energie (ΔG_{DN}) der Auffaltung unter nicht denaturierenden Bedingungen zu schließen (Formel 1). Hierbei wird ausgehend von der Übergangsregion der Auffaltung zurück auf native Bedingungen extrapoliert (Pace & Scholtz, 1997). Parallel stattfindende Auffaltung der Sekundär- und Tertiärstruktur ist ein guter Hinweis dafür, dass ein Protein nur zwei Zustände (N bzw. U) aufweist und keine Faltungsintermediate vorhanden sind.

Bei chemischen Auffaltungsexperimenten wird die Konzentration an Denaturierungsmittel, bei welcher das Protein zur Hälfte in nicht-nativer Form vorliegt als $[D_{1/2}]^{app}$ -Wert definiert. Bei thermischen Auffaltungsexperimenten wird entsprechend die apparente Schmelztemperatur (T_M^{app}) ermittelt, welche der Temperatur entspricht bei der die Hälfte des Proteins in nicht-nativem Zustand vorliegt. Sowohl $[D_{1/2}]^{app}$ als auch T_M^{app} können unabhängig von der Gültigkeit des Zweizustandsmodells ermittelt werden und wurden im Rahmen dieser Arbeit als operationelles Maß für die Proteinestabilität verwendet.

3.5.11 Analyse der Proteinlöslichkeit *in vitro* durch fraktionierte Präzipitation

Die Löslichkeit von Proteinen wird durch Protein-Protein-, Protein-Wasser-, Protein-Ion- und Ion-Wasser-Interaktionen bestimmt (Trevino *et al.*, 2008). Bei Zugabe von Salzen in hoher Konzentration hydratisieren die geladenen Salzionen, wodurch den Proteinen ihre Hydrat-Hülle entzogen wird. Durch hydrophobe Wechselwirkungen kommt es zur Aggregation und schließlich zur Präzipitation der Proteine (1.1.3). Im Gegensatz zu chaotropen Reagenzien (z.B. Harnstoff oder Guanidiniumchlorid), welche den entfalteten Zustand eines Proteins stabilisieren, werden Proteine durch anti-chaotrope (kosmotrope) Salze wie Ammoniumsulfat nicht denaturiert. Die native Struktur und Funktion der Proteine bleibt demnach im aggregierten Zustand erhalten (Trevino *et al.*, 2008).

In dieser Arbeit sollte die Löslichkeit verschiedener Proteinvarianten *in vitro* bestimmt werden. Hierfür wurde aus einer Proteinlösung, einer Pufferstammlösung und einer hochkonzentrierten Ammoniumsulfat-Lösung in gleichem Puffer eine Messreihe mit gleich bleibender Proteinkonzentration und ansteigenden Ammoniumsulfat-Konzentrationen von 0 – 1,2 M erstellt. Die Präzipitationsexperimente fanden bei Raumtemperatur statt. Nach 5 min Inkubation wurden alle Proteinproben 60 min zentrifugiert (EPPENDORF 5415D, 13000 rpm, RT) um das präzipitierte Protein zu pelletieren. Die Konzentration an löslichem Protein im Überstand der verschiedenen Proben wurde mit Hilfe einer zuvor erstellten Eichgerade durch Fluoreszenz-Spektroskopie ermittelt. Hierfür wurde nach Anregung bei 280 nm die relative Intensität der Emission des Überstands (1:8 verdünnt in gleicher Pufferlösung) bei derjenigen Wellenlänge gemessen, bei dem das Protein maximale Emission zeigte. Jede Probe wurde 1 min lang im Fluoreszenz-Spektrophotometer vermessen und die erhaltenen Intensität gemittelt. Diese Methode zur Konzentrationsbestimmung stellte sich als sehr genau heraus, wobei stets sichergestellt wurde, dass im untersuchten Konzentrationsbereich ein linearer Zusammenhang zwischen Fluoreszenz-Signal und Proteinkonzentration besteht und somit ein innerer Filtereffekt ausgeschlossen werden kann (siehe Anhang 7.2).

Die Proteinkonzentration im löslichen Überstand der unterschiedlichen Proben wurde jeweils prozentual auf die nach Inkubation mit 0 M Ammoniumsulfat erhaltene Konzentration an löslichem Protein normiert und als Funktion der Salz-Konzentration aufgetragen. Als operationelles Maß für die Proteinlöslichkeit wurde der $[L_{1/2}]^{\text{app}}$ -Wert eingeführt, welcher der Ammoniumsulfat-Konzentration entspricht bei der die Hälfte des eingesetzten Proteins präzipitiert.

4 Ergebnisse und Diskussion

4.1 Rationales Design der Oberfläche von hPCSK9³¹⁻⁴⁵⁴

Die Löslichkeit von Proteinen kann laut Trevino *et al.* (2007) durch Optimierung der Proteinoberfläche verbessert werden. Hierzu muss die Anzahl an exponierten hydrophilen Aminosäure-Resten durch gezielt eingeführte Mutationen erhöht werden. Relativ konservative Austausche die zu einer höheren Hydrophilizität führen (z.B. Threonin → Serin) können hierbei bereits die Löslichkeit des Proteins verbessern. Bei vielen Proteinen ist keine durch Röntgenkristallographie oder NMR-Spektroskopie bestimmte 3D-Struktur verfügbar. Um in solchen Fällen die Proteinoberfläche zu analysieren, müssen über bioinformatische Methoden erstellte Struktur-Modelle (z.B. Homologie-Modelle) verwendet werden. Dieses Vorgehen sollte am Beispiel der humanen Proprotein Convertase/Subtilisin Kexin Typ 9 (hPCSK9) getestet werden. Um die heterologe Herstellung von hPCSK9 in *Escherichia coli* überhaupt zu ermöglichen wurde im Rahmen dieser Arbeit mit dem modifizierten Konstrukt hPCSK9³¹⁻⁴⁵⁴(S386A) gearbeitet. Es handelt sich hierbei um ein verkürztes Fragment von hPCSK9, welches nur aus der Pro- und der katalytischen Domäne des Proteins besteht (Abbildung 6). Die disulfidreiche C-terminale V-Domäne wurde entfernt, da diese sich im Cytoplasma von *E. coli* nicht richtig falten könnte. Darüber hinaus wurde die katalytische Aktivität der Protease entfernt, indem das katalytisch essentielle Serin 386 gegen ein Alanin ausgetauscht wurde. Dies sorgt dafür, dass keine autokatalytische Abspaltung der Pro- von der katalytischen Domäne erfolgt. Während der Laufzeit dieser Dissertation wurden drei hochaufgelöste Röntgenkristallstrukturen von hPCSK9 veröffentlicht (Cunningham *et al.*, 2007; Hampton *et al.*, 2007; Piper *et al.*, 2007), die im Folgenden zur Evaluierung der erstellten Homologie-Modelle dienen.

4.1.1 Analyse der Oberfläche von hPCSK9³¹⁻⁴⁵⁴

4.1.1.1 Homologie-Modellierung der hPCSK9³¹⁻⁴⁵⁴ - Struktur

Um zuverlässige Aussagen über die Oberfläche des Proteins zu erhalten wurden zu einem Zeitpunkt an dem noch keine experimentell ermittelten 3D-Strukturen von hPCSK9 verfügbar waren, mehrere Homologie-Modelle mit Hilfe unterschiedlicher, öffentlich zugänglicher Server erstellt. Homologe Proteine zu hPCSK9³¹⁻⁴⁵⁴, deren 3D-Struktur in der PDB-Datenbank (<http://www.rcsb.org/pdb/home/home.do>) hinterlegt sind, wurden mit Hilfe von HHpred (<http://toolkit.tuebingen.mpg.de/hhpred>) identifiziert (Tabelle 4).

Tabelle 4: Referenz-Strukturen für Homologie-Modellierungen von hPCSK9³¹⁻⁴⁵⁴.

Proteine die eine ähnliche Sequenz zu hPCSK9³¹⁻⁴⁵⁴ aufweisen und deren 3D-Struktur bekannt ist wurden über HHpred (<http://toolkit.tuebingen.mpg.de/hhpred>) identifiziert.

Struktur	Protein	Organismus	Sequenzidentität mit hPCSK9 ³¹⁻⁴⁵⁴
pdb2b6n	Proteinase K-Like Enzym	<i>Serratia sp.</i>	35%
pdb1sh7	Subtilisin-Like Serin Protease	<i>Vibrio sp.</i>	35%
pdb1ic6	Proteinase K	<i>T. album limber</i>	33%
pdb1gci	Subtilisin	<i>B. lentus</i>	31%
pdb1to2	Subtilisin BPN ^c	<i>B. amyloliquefaciens</i>	31%
pdb1y1k	Subtilisin BPN ^c	<i>B. amyloliquefaciens</i>	31%
pdb1r0r	Subtilisin Carlsberg	<i>B. licheniformis</i>	30%
pdb1thm	Thermitase	<i>T. vulgaris</i>	29%
pdb2e1p	Subtilisin	<i>T. kodakarensis</i>	25%
pdb1y9z	Subtilisin-Like Serin Protease	<i>Pseudoalteromonas sp.</i>	25%
pdb1r6v	Fervidolysin	<i>F. pennivorans</i>	23%
pdb1ea7	Sphaericase	<i>B. sphaericus</i>	23%
pdb1r64	Kexin	<i>S. cerevisae</i>	20%
pdb1p8j	Furin	<i>M. musculus</i>	19%

Mit Hilfe des Programmes Modeller (<http://toolkit.tuebingen.mpg.de/modeller>) wurden anschließend sieben Homologie-Modelle (Modelle I-VII) für hPCSK9³¹⁻⁴⁵⁴ auf der Grundlage unterschiedlicher Referenz-Strukturen generiert. Zusätzlich wurde mit Hilfe der Homologie-Modellierungs – Server 3djigsaw (<http://bmm.cancerresearchuk.org/~3djigsaw/>) (Modell VIII) und CPH models 3.0 (<http://www.cbs.dtu.dk/services/CPHmodels/>) (Modell IX) jeweils ein Homologie-Modell berechnet, wobei diese Programme jeweils selbstständig eine Referenz-Struktur wählten. Mit Hilfe von DaliLite (<http://www.ebi.ac.uk/DaliLite>) wurden paarweise Superpositionierungen aller Modelle mit Modell I durchgeführt. Der Vergleich ergab, dass sich die Modelle in ihrer globalen Struktur stark ähneln, was für die Qualität der berechneten Strukturen spricht (Tabelle 5).

Tabelle 5: Homologie-Modelle von hPCSK9³¹⁻⁴⁵⁴.

RMSD-Werte (C α Atome) wurden über DaliLite bestimmt.

Modell	Programm	Referenzstruktur(en)	RMSD zu Modell I
I	Modeller	pdb2b6n, pdb1r6v, pdb1ic6, pdb1sh7, pdb1r0r, pdb1gci, pdb1y9z, pdb1to2, pdb1thm, pdb1ea7	-
II	Modeller	pdb2b6n	1,6 Å
III	Modeller	pdb1p8j	2,0 Å
IV	Modeller	pdb1r64	2,1 Å
V	Modeller	pdb2e1p	2,2 Å
VI	Modeller	pdb1r6v	2,5 Å
VII	Modeller	pdb1r6v, pdb2e1p	2,7 Å
VIII	3djigsaw	pdb1y1k	2,0 Å
IX	CPH models 3.0	pdb2b6n	1,7 Å

Die korrekt modellierte Anordnung des katalytischen S386 an die Autoprozessierungs-Stelle Q152-S153 spricht ebenfalls für die Qualität der Homologie-Modelle (Abbildung 26a). Superpositionierungen von Modell I mit den drei später veröffentlichten, experimentell ermittelten Strukturen pdb 2pmw, pdb 2qtw und pdb 2p4e (Piper *et al.*, 2007; Hampton *et al.*, 2007; Cunningham *et al.*, 2007) zeigten, dass das Modell global eine hohe Ähnlichkeit zur tatsächlichen Struktur aufweist (RMSD: 2,5 Å) (Abbildung 26b). Hierbei muss zusätzlich beachtet werden, dass es sich bei den Kristallstrukturen jeweils um die prozessierte Form des Proteins handelt bei der die Pro- und die katalytische Domäne im Gegensatz zu Modell I nicht-kovalent verbunden sind, weswegen es im Bereich der Pro-Domäne zu größeren Abweichungen kommt.

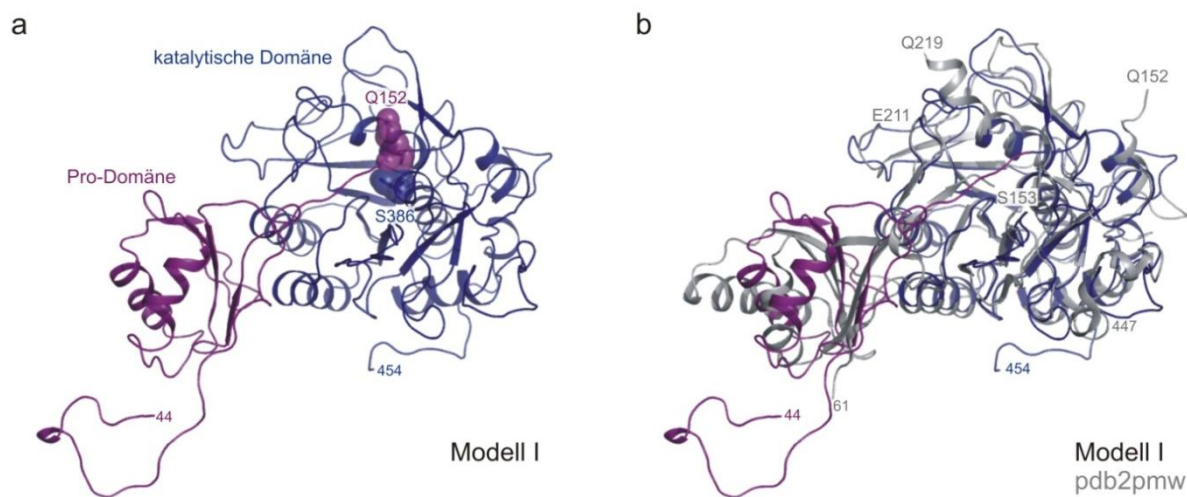


Abbildung 26: Homologie-Modell I von hPCSK9³¹⁻⁴⁵⁴.

a Das Modell zeigt deutlich separierte Pro (magenta)- und katalytische Domäne (blau), wobei die autokatalytische Spaltstelle bei Q152 korrekt am katalytisch essentiellen S386 positioniert ist. **b** Die Superpositionierung von Modell I mit der Kristallstruktur pdb 2pmw (Piper *et al.*, 2007) (grau) wurde mit DaliLite (<http://www.ebi.ac.uk/DaliLite>) durchgeführt. 268 verglichene Aminosäuren-Reste überlagerten mit einem RMSD von 2,5 Å (Z-Score: 35,5). Superpositionierungen von Modell I mit pdb 2p4e (Cunningham *et al.*, 2007) und pdb 2qtw (Hampton *et al.*, 2007) ergaben ebenfalls einen RMSD von 2,5 Å. Einige Stellen der Struktur sind bei pdb 2pmw nicht aufgelöst ($\Delta 168-175$, $\Delta 212-218$, $\Delta 448-451$, $\Delta 572-584$, $\Delta 660-669$). Modell I zeigt bis auf eine N-terminale Verkürzung die vollständige Struktur von hPCSK9³¹⁻⁴⁵⁴ (Aminosäure 44 bis 454).

4.1.1.2 Lösungsmittel-Zugänglichkeit der Aminosäuren von hPCSK9³¹⁻⁴⁵⁴

Die Oberflächen der generierten Homologie-Modelle wurden mit Hilfe von Polyview (http://polyview.cchmc.org/index_old.html) analysiert. Das Programm gibt die Lösungsmittel-Zugänglichkeit der einzelnen Aminosäuren auf einer Skala von 0 (Seitenkette komplett im Protein vergraben) bis 9 (Seitenkette zu 100 % Lösungsmittel-zugänglich) an. Es konnten so 15 hydrophobe Aminosäuren (Valin, Leucin, Phenylalanin, Tyrosin, Tryptophan) bzw. Threonine in hPCSK9³¹⁻⁴⁵⁴ identifiziert werden, die bei der Mehrzahl der Homologie-Modelle an der Oberfläche des Proteins exponiert sind (Tabelle 6, Abbildung 27). Die

nachträgliche Überprüfung der Proteinoberfläche von hPCSK9 mit Hilfe der Röntgenkristallstrukturen pdb 2pmw, pdb 2qtw und pdb 2p4e (Cunningham *et al.*, 2007; Hampton *et al.*, 2007; Piper *et al.*, 2007) zeigte, dass die getroffenen Vorhersagen qualitativ sehr gut mit der tatsächlichen Situation übereinstimmen (Tabelle 6).

Tabelle 6: Lösungsmittel-zugängliche Threonine, Leucine, Valine, Tyrosine, Phenylalanine und Tryptophane von hPCSK9³¹⁻⁴⁵⁴.

Die Analyse der Oberflächen der Homologie-Modelle und der Röntgenkristallstrukturen erfolgte mit Hilfe von Polyview. Die Lösungsmittelzugänglichkeit der einzelnen Aminosäuren wurde auf einer Skala von 0 (vollständig abgeschirmt) bis 9 (vollständig exponiert) bewertet. Bei einigen Strukturen waren manche Aminosäuren nicht aufgelöst (n.a.).

	Homologie-Modelle									Röntgenkristallstrukturen				
	I	II	III	IV	V	VI	VII	VIII	IX	∅	pdb2pmw	pdb2qtw	pdb2p4e	∅
Thr60	9	n.a.	n.a.	n.a.	n.a.	5	9	n.a.	n.a.	7,7	n.a.	n.a.	n.a.	n.a.
Thr61	6	n.a.	n.a.	n.a.	n.a.	5	7	n.a.	n.a.	6,0	9	9	9	9,0
Thr63	6	n.a.	n.a.	n.a.	n.a.	1	7	n.a.	n.a.	4,7	4	4	4	4,0
Leu179	8	7	9	6	7	9	8	7	5	7,3	5	4	5	4,7
Thr214	7	4	4	3	1	0	8	0	9	4,0	n.a.	n.a.	n.a.	n.a.
Phe216	5	7	4	3	6	4	3	6	8	5,1	n.a.	n.a.	n.a.	n.a.
Leu276	8	9	0	7	6	8	0	7	5	5,6	5	6	6	5,7
Val277	6	3	0	8	4	7	0	6	2	4,0	8	7	6	7,0
Tyr293	4	5	6	5	5	6	0	4	5	4,4	5	5	6	5,3
Leu324	3	5	3	8	0	0	1	5	5	3,3	5	4	5	4,7
Thr350	0	0	0	0	6	6	1	0	6	2,1	9	9	9	9,0
Leu351	3	0	0	2	2	6	0	7	3	2,6	4	4	4	4,0
Thr377	4	5	4	3	7	5	5	2	5	4,4	6	6	6	6,0
Thr448	5	n.a.	n.a.	n.a.	n.a.	n.a.	5	n.a.	n.a.	5,0	n.a.	6	7	6,5
Trp453	4	n.a.	n.a.	n.a.	n.a.	n.a.	8	n.a.	n.a.	6,0	4	4	5	4,3

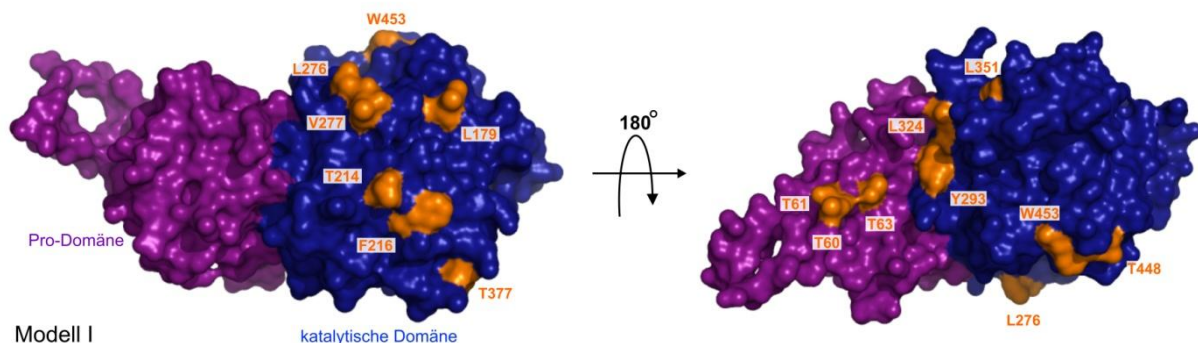


Abbildung 27: Lösungsmittel-zugängliche Threonine, Leucine, Valine, Tyrosine, Phenylalanine und Tryptophane des hPCSK9³¹⁻⁴⁵⁴-Homologie-Modells I.

Die über Polyview identifizierte Threonine, Leucine, Valine, Tyrosine, Phenylalanine und Tryptophane auf der Oberfläche von Modell I sind orange dargestellt. Dieselben Aminosäuren wurden durch Analyse der anderen Homologie-Modelle identifiziert (Tabelle 6), wobei noch eine zusätzliche Aminosäure als Lösungsmittel-zugänglich definiert wurde, die in diesem Modell vergraben vorliegt (T350).

Nach der Identifikation dieser 15 exponierten Aminosäuren, wurde eine Variante von hPCSK9(S386A)³¹⁻⁴⁵⁴ hergestellt, bei der diese Reste durch gezielte Mutagenese gegen hydrophilere ausgetauscht wurden. Im Allgemeinen wurde ein Serin anstelle des wildtypischen Restes eingeführt (Threonin 60, 61, 63, 214, 350, 377, 448; Leucin 324, 351; Valin 277; Tryptophan 453). Serin ist die hydrophilste Aminosäure ohne Ladung und stellt somit den konservativsten hydrophilen Austausch dar. Die sehr stark exponierten Leucin-Reste 179 und 276 wurden gegen das noch hydrophilere Glutamat (Trevino *et al.*, 2007) ausgetauscht, da bei diesen Positionen davon ausgegangen werden kann, dass die Ladung ins Solvent zeigt und es zu keinen ungünstigen elektrostatischen Wechselwirkungen mit anderen Aminosäure-Resten kommt. Die aromatischen Aminosäuren Phenylalanin 216 und Tyrosin 293 wurden gegen ein Histidin ausgetauscht, um einerseits eine hydrophilere Aminosäure einzuführen, andererseits aber mögliche *Stacking*-Wechselwirkungen oder Wasserstoffbrücken zu anderen Aminosäuren beizubehalten.

4.1.2 Klonierung von *hpcsk9*³¹⁻⁴⁵⁴(S386A)^{opt}

Zu Beginn dieser Arbeit lag die kodierende Sequenz für hPCSK9³¹⁻⁴⁵⁴(S386A) im Expressionsvektor pET21a(+) vor. Dieses Konstrukt wurde freundlicherweise von Dr. Daniel Schlatter (F. Hoffmann – LaRoche, *Pharma Research Discovery*, Basel) zur Verfügung gestellt. *hpcsk9*³¹⁻⁴⁵⁴(S386A) war hierbei so über *NdeI* und *XhoI* in den Vektor kloniert, dass dem translatierten hPCSK9³¹⁻⁴⁵⁴(S386A) Protein ein kurzes Peptid (GSELEHHHHHH) C-terminal angehängt wird. Der darin enthaltene His₆-Tag ermöglicht die Aufreinigung des Proteins über Metallchelate-Affinitätschromatographie (3.4.2.1).

Die Nukleotid-Austausche, welche für die 15 Aminosäuren-Mutationen T60S, T61S, T63S, L179E, T214S, F216H, L276E, V277S, Y293H, L324S, T350S, L351S, T377S, T447S und W453S kodieren, wurden schrittweise mit Hilfe von Megaprimer-PCR (3.3.3.1) in *hpcsk9*³¹⁻⁴⁵⁴(S386A) eingeführt. Der Megaprimer wurde hierbei stets mit einem Mutationsprimer, welcher einen oder mehrere Austausche beinhaltet (3'*hpcsk9*(T60S+T61S+T63S), 3'*hpcsk9*(L179E), 3'*hpcsk9*(T214S+F216H), 3'*hpcsk9*(T377S), 3'*hpcsk9*(T350S+L351S) bzw. 5'*hpcsk9*(L276S+V277S), 5'*hpcsk9*(T447S+W453S), 5'*hpcsk9*(Y293H), 5'*hpcsk9*(L324S)), und einem genflankierenden Primer (5'*hpcsk9*(*NdeI*) bzw. 3'*hpcsk9*(*XhoI*)) mittels PCR (3.3.1) amplifiziert, wobei zuerst pET21a(+)_{*hpcsk9*³¹⁻⁴⁵⁴}(S386A) und dann sukzessive das pET21a(+)_{*hpcsk9*³¹⁻⁴⁵⁴}(S386A) Konstrukt mit den zuvor eingefügten Austauschen als Templat diente. Die Vollängen-Gene wurden anschließend mit Hilfe der aus einem Agarosegel aufgereinigten (3.3.5) Megaprimer

und des entsprechenden genflankierenden Primers (3' *hpcsk9*(*XhoI*) bzw. 5' *hpcsk9*(*NdeI*)) mit Hilfe desselben Templat wie zuvor durch PCR amplifiziert. Mittels der Restriktionsenzyme *NdeI* und *XhoI* wurde das erhaltene mutagenisierte Gen in einen leeren pET21a(+)-Vektor kloniert und diente als Templat für die nächste Megaprimer-PCR, bei welcher ein neuer Austausch hinzugefügt werden konnte. Nach jeder Klonierung wurde die Nukleotidsequenz durch Sequenzierung mit den vektorspezifischen Primern T7 Promotor und T7 Terminator bestätigt. Nach neun Klonierungsrunden wurde das Konstrukt pET21a(+)_{*hpcsk9*³¹⁻⁴⁵⁴(S386A)^{opt}} erhalten, in welchem alle oben erwähnten Mutationen in *hpcsk9*³¹⁻⁴⁵⁴(S386A) enthalten sind.

4.1.3 Heterologe Herstellung und Reinigung der hPCSK9³¹⁻⁴⁵⁴ Varianten

Die Expression von *hpcsk9*³¹⁻⁴⁵⁴(S386A) bzw. *hpcsk9*³¹⁻⁴⁵⁴(S386A)^{opt} erfolgte in mit den entsprechenden pET21a(+)-Vektoren transformierten *E. coli* BL21(DE3) Zellen, wobei die Expression nach einem osmotischen Schock mit NaCl durch Zugabe von IPTG bei 18 °C üN induziert wurde (3.4.1.2). Die Produktion der Proteine wurde zuvor mittels Probeexpression (3.4.1.1) überprüft, wobei bei hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} ein deutlich größerer Anteil des Proteins in der löslichen Zellfraktion zu finden war als bei hPCSK9³¹⁻⁴⁵⁴(S386A) (Abbildung 28a). Die Reinigung beider Proteine erfolgte mit Hilfe des C-terminalen His₆-Tags über Metallchelate-Affinitätschromatographie (3.4.2.1) (Abbildung 28b). Die Proteine eluierten jeweils im Bereich von 80 – 300 mM Imidazol.

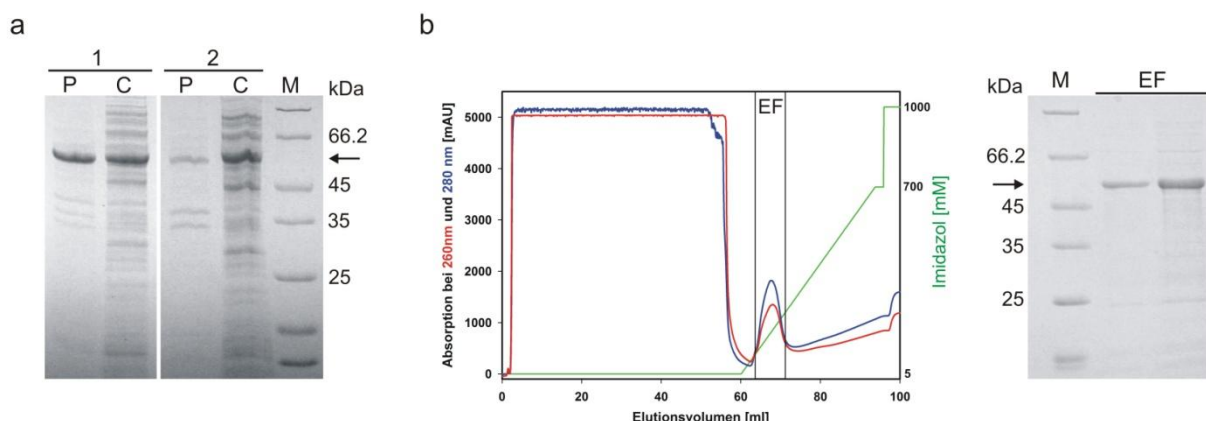


Abbildung 28: Heterologe Expression in *E. coli* BL21(DE3) und Reinigung der hPCSK9³¹⁻⁴⁵⁴ Varianten.

P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*), M: Markerproteine (Unstained Protein Molecular Weight Marker, MBI Fermentas). Bei den SDS-PAGE Gelen sind die Banden der hPCSK9³¹⁻⁴⁵⁴(S386A) Varianten mit einem Pfeil markiert.

a Probeexpressionen, analysiert über SDS-PAGE (12,5 %). Bei der Oberflächen-optimierten Variante hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} (2) befindet sich mehr Protein in der löslichen Zellfraktion als bei hPCSK9³¹⁻⁴⁵⁴(S386A) (1).

b Reinigung von hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} über Metallchelate-Affinitätschromatographie. Das Elutionsprofil wurde aufgenommen über die Messung der Absorption bei 260 nm (rot) und 280 nm (blau). Die Elution erfolgte über einen Imidazolgradienten von 0 – 700 mM (grün). Die Elutionsfraktionen (EF) von 80 – 300 mM Imidazol enthielten sauberes Protein (analysiert über SDS-PAGE (12,5 %)).

Die Fraktionen mit sauberem Protein wurden jeweils vereinigt und zur Entfernung des Imidazols zweimal gegen 1 l 50 mM Tris/Cl, 5 mM DTT, pH 8,0 dialysiert (3.4.2.5).

Die Ausbeute betrug bei hPCSK9³¹⁻⁴⁵⁴(S386A) ca. 1 mg und bei hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} ca. 5 mg Protein pro Liter Zellkultur. Die Proteine wurden in flüssigen Stickstoff eingetropfelt und bei -80 °C gelagert (3.4.2.7). Die verbesserte Reinigungs-Ausbeute von hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} im Vergleich zu hPCSK9³¹⁻⁴⁵⁴(S386A) zeigt, dass die 15 eingeführten Mutationen auf der Oberfläche des Proteins tatsächlich zu einer Erhöhung der Löslichkeit *in vivo* führen.

4.1.4 Biophysikalische Charakterisierung der hPCSK9³¹⁻⁴⁵⁴ Varianten

Die in Tabelle 7 gezeigten molaren und spezifischen Extinktionskoeffizienten der Varianten wurden über deren Aminosäure-Sequenz und deren Molekulargewicht ermittelt (siehe 3.5.1.2, Formel 10, Formel 11).

Tabelle 7: Biophysikalische Parameter der hPCSK9³¹⁻⁴⁵⁴ Varianten.

Molekulargewichte (MW) unter Berücksichtigung des C-terminalen His₆-Tags, Anzahl der enthaltenen Tryptophan - (Trp) bzw. Tyrosin (Tyr) - Reste, sowie molare (ϵ_{280}) und spezifische ($^{0,1\%}A_{280}$) Extinktionskoeffizienten.

Protein	MW [Da]	Σ Trp + Tyr	$\epsilon_{280}[\text{M}^{-1}\text{cm}^{-1}]$	$^{0,1\%}A_{280}[\text{cm}^2/\text{mg}]$
hPCSK9 ³¹⁻⁴⁵⁴ (S386A)	47312	4 + 9	35410	0,75
hPCSK9 ³¹⁻⁴⁵⁴ (S386A) ^{opt}	47046	3 + 8	28420	0,60

4.1.4.1 Analyse der Löslichkeit *in vitro* durch fraktionierte Präzipitation

Um die Löslichkeit der hPCSK9³¹⁻⁴⁵⁴ Varianten näher zu charakterisieren wurden die Proteine in 50 mM Tris/Cl, 5 mM DTT, pH 8,0 mit Ammoniumsulfat fraktionell präzipitiert (3.5.1.1) (Abbildung 29). Hierbei zeigte sich, dass hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} auch *in vitro* eine höhere Löslichkeit aufweist als hPCSK9³¹⁻⁴⁵⁴(S386A). So fällt hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} erst bei höheren Ammoniumsulfat-Konzentrationen aus als hPCSK9³¹⁻⁴⁵⁴(S386A). Um 50 % des Proteins zu präzipitieren ist bei hPCSK9³¹⁻⁴⁵⁴(S386A) eine Ammoniumsulfat-Konzentration von 0,23 M nötig, im Vergleich zu 0,52 M bei hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt}.

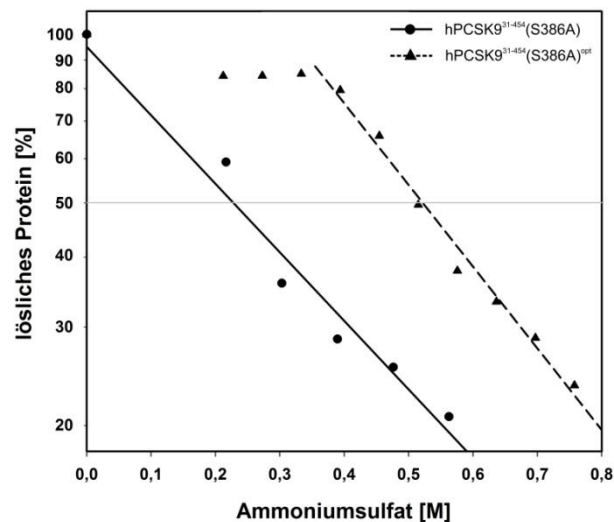


Abbildung 29: Löslichkeit von hPCSK9³¹⁻⁴⁵⁴(S386A) und hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} in vitro.

Die Proteine wurden jeweils mit einer Konzentration von 4 μM in 50 mM Tris/Cl, 5 mM DTT, pH 8,0 bei RT fraktionell mit steigenden Konzentrationen an Ammoniumsulfat präzipitiert. Die Konzentration des im Überstand verbleibenden, löslichen Proteins wurde über Fluoreszenz-Spektroskopie anhand einer zuvor bestimmten Eichgerade (siehe Anhang 7.2) bestimmt. Bei jeder Probe wurde von einer 1:8 Verdünnung des löslichen Überstands (in 50 mM Tris/Cl, 5 mM DTT, pH 8,0) nach Anregung bei 280 nm die Emission bei 344 nm für 1 min gemessen und das erhaltene Signal gemittelt (Schichtdicke der Küvette: 1 cm; Photomultiplier-Spannung: 910 V). Die für hPCSK9³¹⁻⁴⁵⁴(S386A) bzw. hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} ermittelten $[L_{1/2}]^{\text{app}}$ -Werte betragen 0,23 M bzw. 0,52 M.

4.1.4.2 Analyse der konformationellen Stabilität durch Auffaltung in Harnstoff

Die hPCSK9³¹⁻⁴⁵⁴ Varianten wurden durch schrittweise Denaturierung mit Harnstoff auf ihre konformationelle Stabilität hin untersucht (3.5.10). Der Anteil an gefaltetem bzw. entfaltetem Protein jeder Probe wurde durch Fluoreszenz-Spektroskopie ermittelt, wobei nach Anregung mit 280 nm die unterschiedliche Emission des nativen und denaturierten Zustands bei 318 nm bzw. 319 nm als Signal ausgenutzt werden konnte (Abbildung 30a). Die Auffaltungsexperimente in 50 mM Tris/Cl, 5 mM DTT, pH 8,0 zeigen, dass hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} bereits bei niedrigeren Konzentrationen an Harnstoff auffaltet als hPCSK9³¹⁻⁴⁵⁴(S386A) (Abbildung 30b). Der entsprechende $[D_{1/2}]^{\text{app}}$ -Wert beträgt bei hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} 3,02 M Harnstoff, im Vergleich zu 3,75 M bei hPCSK9³¹⁻⁴⁵⁴(S386A). Die eingeführten 15 Oberflächen-Mutationen führen demnach zu einer Destabilisierung des Proteins.

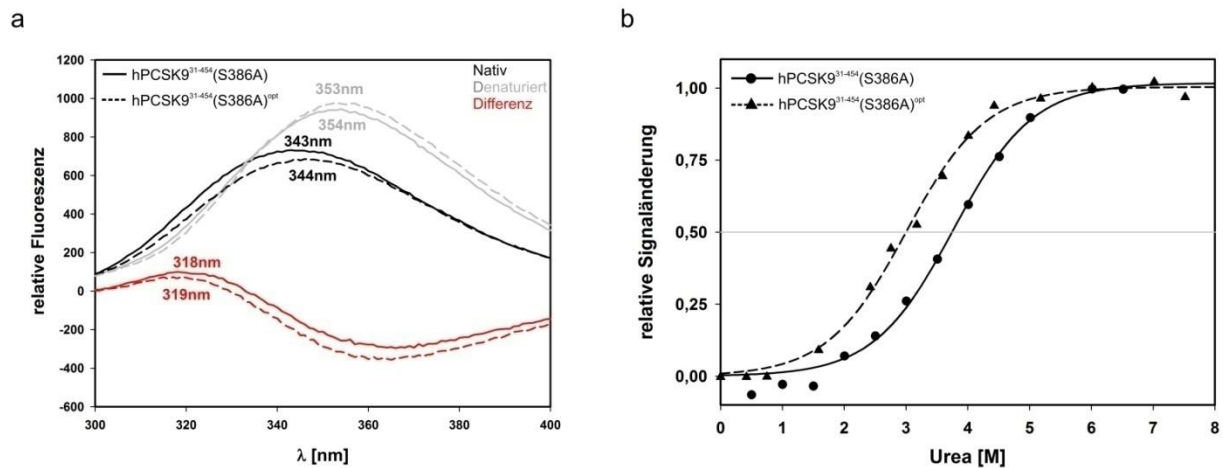


Abbildung 30: Auffaltung von hPCSK9³¹⁻⁴⁵⁴(S386A) und hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} mit Harnstoff.

a Fluoreszenzemissionsspektren von hPCSK9³¹⁻⁴⁵⁴(S386A) bzw. hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} in 50 mM Tris/Cl, 5 mM DTT, pH 8,0 (nativ) und 50 mM Tris/Cl, 5 mM DTT, 8 M Urea, pH 8,0 (denaturiert). Die Anregung erfolgte bei 280 nm. Es wurden jeweils 3 μ M Protein eingesetzt (Schichtdicke der Küvette: 1 cm; Photomultiplier-Spannung: 750 V). Die Maxima der Emission sind jeweils angegeben. Die Differenz der nativen und denaturierten Spektren ist in rot dargestellt und zeigte ein Maximum bei 318 nm bzw. 319 nm.

b Die Auffaltungskurven von hPCSK9³¹⁻⁴⁵⁴(S386A) bzw. hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} wurden durch Fluoreszenz-Spektroskopie verfolgt (Excitation: 280 nm, Emission: 318 nm bzw. 319 nm; Schichtdicke: 1 cm, Photomultiplier-Spannung: 810 V). Die Messungen wurden in 50 mM Tris/Cl, 5 mM DTT, pH 8,0 bei RT durchgeführt. Das fluorimetrische Signal jeder Probe wurde 1 min lang aufgenommen und gemittelt. Die für hPCSK9³¹⁻⁴⁵⁴(S386A) bzw. hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} ermittelten $[D_{1/2}]^{app}$ -Werte betragen 3,75 M bzw. 3,02 M.

Die gezielte Oberflächen-Mutagenese von hPCSK9³¹⁻⁴⁵⁴(S386A) zur Optimierung der Löslichkeit und Stabilität hat somit sein Ziel nicht vollständig erreicht. Einerseits wurde durch die aufgrund von Homologie-Modellen ausgewählten Mutationen (4.1.1) zwar die Proteinlöslichkeit tatsächlich verbessert (4.1.3, 4.1.4.1). Andererseits wurde aber die Oberfläche durch die hohe Anzahl an Mutationen so stark verändert, dass die konformationelle Stabilität und somit die strukturelle Integrität des Proteins erniedrigt wurde. Die gelenkte Evolution von Proteinen durch Zufallsmutagenese mit anschließendem *Screening* nach optimierten Varianten (1.2.2) stellt eine alternative Methode zur Erhöhung der Löslichkeit und Stabilität von Proteinen dar. Diese Vorgehensweise benötigt keinerlei strukturelle Information über das zu verbessernde Zielprotein. Es werden somit ungünstige Mutationen, welche durch fehlerbehaftete Struktur-Modelle vorgeschlagen werden, ausgeschlossen. Darüber hinaus können überall im Protein Mutationen identifiziert werden, welche die Löslichkeit bzw. die konformationelle Stabilität des Proteins auf die unterschiedlichste Art und Weise verbessern können. Im nächsten Teil dieser Arbeit wird eine Methode zur gelenkten Evolution hoher Löslichkeit und Stabilität von Proteinen vorgestellt, bei der eGFP als Reporterprotein dient (siehe 1.3.2, 3.2.8.1).

4.2 Benutzung von eGFP als Löslichkeits- und Stabilitäts-Reporter *in vivo*

Wie in 1.2.2 erläutert, wurde die Verwendung des grün fluoreszierenden Proteins (GFP) aus *Aequorea victoria* als Faltungsreporter *in vivo* bereits in mehreren Publikationen beschrieben. Im Rahmen dieser Arbeit sollte nun eine Methode der gelenkten Evolution von Proteinen hin zu hoher Löslichkeit und Stabilität etabliert werden, bei der mit Hilfe eines MoFlo[®] Zytometers Zellen mit GFP-fusionierten Proteinen analysiert und sortiert werden können. Durch die Etablierung von Durchflusszytometrie (3.2.9) als Hochdurchsatz *Screening* – Methode könnten auch große GFP-fusionierte Banken ($> 10^5$ unabhängige Varianten) eines Zielproteins zügig und effizient nach verbesserten Mutanten abgesucht werden. Hierfür wurde eine für FACS-Anwendungen optimierte eGFP-Variante (F64L+S65T+Q80R+F99S+M153T+V163A) benutzt (siehe 1.4).

Im Folgenden wird zunächst beschrieben wie verschiedene Expressionssysteme mit den Proteinen hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP getestet wurden, um das für das *Screening* am besten geeignete Plasmidsystem und Zielprotein zu identifizieren (4.2.1). Das weitere Vorgehen bei der Analyse und der Anreicherung fluoreszierender Zellen mittels Durchflusszytometrie wird anschließend an einem einfachen Fall exemplarisch erläutert (4.2.2)

4.2.1 Auswahl eines geeigneten Expressionssystems und Zielproteins

Bei Waldo *et al.* (1999) wurde zur Expression der *egfp*-fusionierten Konstrukte ein induzierbarer pET-Vektor verwendet, bei dem nach Genexpression das Zielprotein mit eGFP über einen 12 Aminosäuren langen Linker (GSAGSAAGSGEF) verbunden wird. Diese grundsätzliche Konstruktion wurde beibehalten, wobei zur Ermittlung der besten Bedingungen für die gelenkte Evolution hoher Proteinlöslichkeit *in vivo* folgende Expressionssysteme für die *egfp*-fusionierten Gene getestet wurden: starke induzierbare Genexpression durch den Vektor pER13a/b (2.8.1.2.1), stark konstitutive Genexpression durch den Vektor pHCE13a (2.8.2.4) und schwach konstitutive Genexpression durch den Vektor pTNA13a (2.8.2.2). Vorversuche zeigten, dass bei Zellen mit pER13a/b die Genexpression nur einmal induziert werden kann. Überimpfung induzierter Zellen, gefolgt von neuer Induktion führte zu nicht reproduzierbaren Fluoreszenz-Intensitäten (Daten nicht gezeigt). Um bei einer FACS-Sortierung eine neue Anreicherungsrunde zu starten müssen deshalb die Plasmide angereicherter Zellen präpariert und neue Zellen damit transformiert werden. Bei konstitutiven Expressionssystemen werden die eGFP-Fusionsproteine permanent

in der Zelle hergestellt, wodurch die benötigte Zeit für den *Screening*-Prozess deutlich verkürzt werden würde. Die angereicherten Zellen könnten nach Anwachsen in LB-Selektivmedium sofort neu sortiert werden, womit die FACS-Anreicherungsrounden täglich fortgesetzt werden könnten. Andererseits entspricht starke induzierbare Expression durch die T7-Polymerase in *E. coli* den Bedingungen bei der heterologen Herstellung der Proteine im präparativen Maßstab. *Screening* unter diesen Bedingungen würde die Wahrscheinlichkeit erhöhen, dass nur solche Varianten gefunden werden, welche auch bei der späteren präparativen Reinigung der isolierten Zielproteine ohne eGFP verbesserte Eigenschaften aufweisen.

Als potentielle Zielproteine wurden hPCSK9³¹⁻⁴⁵⁴(S386A) (Pro- und inaktive katalytische Domäne der humanen Proprotein Convertase/Subtilisin/Kexin Typ9) und hGR-LBD (Ligand-Bindungs Domäne des humanen Glucocorticoid-Rezeptors) analysiert. Als Kontrollen wurden tmHisF (Synthase-Untereinheit der Imidazolglycerinphosphat-Synthase aus *Thermotoga maritima*) bzw. hKAS (β -Ketoacyl-Synthase Domäne der humanen Fettsäure-Synthase) verwendet, welche eine sehr hohe (Seitz *et al.*, 2007) bzw. niedrige Löslichkeit (Dr. Ralf Thoma, mündliche Mitteilung) in *E. coli* zeigen. Die verschiedenen Vektoren sind so konstruiert, dass das kodierte Zielprotein C-terminal mit eGFP über einen 13 AS langen Linker (GGRAGSAAGSGEF) verbunden wird. Als Alternative wurde zudem das Konstrukt pER13a_GSG-tmhisF hergestellt. Dies führt zur Herstellung von tmHisF-eGFP, wobei die beiden Fusionspartner nur durch drei Aminosäuren (Gly-Ser-Gly) verbunden werden. Dadurch sollte die potentielle Anfälligkeit der flexiblen Linker-Region für proteolytischen Abbau herabgesetzt werden.

4.2.1.1 Klonierung verschiedener *egfp*-Fusionsvektoren

Die pET24a(+)-basierten *egfp*-Fusionsvektoren pER13a und pER13b (2.8.1.2.1) wurden freundlicherweise von Dr. Ralf Thoma (F. Hoffmann – LaRoche, *Pharma Research Discovery*, Basel) zur Verfügung gestellt. Die darin enthaltene kodierende Sequenz für eGFP (inklusive der Linkerregion) wurde in die Vektoren pHCE-IIB (2.8.2.3) und pTNA (2.8.2.1) eingefügt. Hierzu wurde die betreffende Region mit den Primern 5'N-eGFP(*Nco*I) bzw. 5'N-eGFP(*Sph*I) und 3'C-eGFP(*Hind*III) aus dem Templat pER13a mittels PCR (3.3.1) amplifiziert. Die erhaltenen Produkte konnten anschließend über die Restriktionsschnittstellen *Nco*I bzw. *Sph*I und *Hind*III in pHCE-IIB bzw. pTNA eingefügt werden, wodurch die stark bzw. schwach konstitutiven *egfp*-Expressionsvektoren pHCE13a (2.8.2.4) bzw. pTNA13a (2.8.2.2) hergestellt wurden.

Die *hpcsk9*³¹⁻⁴⁵⁴(S386A)-Sequenz enthält zwei *NotI*- und jeweils eine *NcoI*- und *SphI*-Erkennungssequenz, welche parallel zu den Klonierungen in die verschiedenen *egfp*-Fusionsvektoren entfernt wurden. Zur Klonierung in pER13a wurde *hpcsk9*³¹⁻⁴⁵⁴(S386A) durch zwei zuvor mit Hilfe der Primerpaare 5'*hpcsk9*(*NdeI*) / 3'*hpcsk9_delNotI_1* und 5'*hpcsk9_delNotI_2* / 3'*hpcsk9*(*NotI*) durch Megaprimer-PCR (3.3.3.1) hergestellten Megaprimern aus dem Templat pET21a(+)*_hpcsk9*³¹⁻⁴⁵⁴(S386A) (von Dr. Daniel Schlatter zur Verfügung gestellt) mittels PCR amplifiziert, wobei pET21a(+)*_hpcsk9*³¹⁻⁴⁵⁴(S386A) wieder als Templat diente. Das erhaltene Produkt wurde über die Restriktionsschnittstellen *NdeI* und *NotI* in pER13a eingefügt. Zur Klonierung in pHCE13a wurde *hpcsk9*³¹⁻⁴⁵⁴(S386A) mit den Primern 5'*hpcsk9*³¹⁻⁴⁵⁴(*NcoI*) und 3'*hpcsk9*³¹⁻⁴⁵⁴*_delNcoI*(*NotI*) aus dem Templat pER13a*_hpcsk9*³¹⁻⁴⁵⁴(S386A) mittels PCR amplifiziert und das erhaltene Produkt über die Restriktionsschnittstellen *NcoI* und *NotI* in pHCE13a eingefügt. Zur Klonierung in pTNA13a wurde *hpcsk9*³¹⁻⁴⁵⁴(S386A) durch einen zuvor mit den Primern 5'*hpcsk9*(*SphI*) und 3'*hpcsk9_delSphI* aus dem Templat pER13a*_hpcsk9*³¹⁻⁴⁵⁴(S386A) durch Megaprimer-PCR (3.3.3.1) hergestellten Megaprimer und 3'*hpcsk9*³¹⁻⁴⁵⁴(*NotI*) mittels PCR amplifiziert, wobei pER13a*_hpcsk9*³¹⁻⁴⁵⁴(S386A) wieder als Templat diente. Das erhaltene Produkt wurde über die Restriktionsschnittstellen *SphI* und *NotI* in pTNA13a eingefügt.

Ausgehend von dem Templat pER13a*_hgr-lbd* (von Dr. Ralf Thoma zur Verfügung gestellt) wurde *hgr-lbd* mit den Primern 5'*hgr-lbd*(*NcoI*) bzw. 5'*hgr-lbd*(*SphI*) und 3'pER13 mittels PCR amplifiziert. Die erhaltenen Produkte konnten anschließend jeweils über die Restriktionsschnittstellen *NcoI* bzw. *SphI* und *NotI* in pHCE13a und pTNA13a eingefügt werden.

tmhisF wurde mit den Primern 5'*tmhisF-N*(*NdeI*), 5'*tmhisF*(*NcoI*) bzw. 5'*tmhisF*(*SphI*) und 3'*tmhisF*(*NotI*) aus dem Templat pET24a(+)*_tmhisF* (von Dr. Jörg Claren zur Verfügung gestellt) mittels PCR (3.3.1) amplifiziert. Die erhaltenen Produkte konnten anschließend über die Restriktionsschnittstellen *NdeI*, *NcoI* bzw. *SphI* und *NotI* in pER13a, pHCE13a und pTNA13a eingefügt werden. Zur Verkürzung des Linkers zwischen *tmhisF* und *egfp* in pER13a*_tmhisF* wurde die gesamte kodierende Sequenz für tmHisF-Linker-eGFP über OE-PCR (3.3.3.2) mit Hilfe der Primerpaare 5'pER13 / 3'pER13a_GSG und 5'pER13a_GSG / T7 Terminator aus dem Templat pER13a*_tmhisF* amplifiziert, wobei die beiden inneren Primer (3'pER13a_GSG und 5'pER13a_GSG) die verkürzte Linkersequenz einfügen. Das erhaltene Amplifikationsprodukt wurde anschließend über die Restriktionsschnittstellen *NdeI* und *NotI* in pER13a kloniert, wodurch pER13a_GSG-*tmhisF* erhalten wurde.

Ausgehend von dem Templat pER13a_hkas (von Dr. Ralf Thoma zur Verfügung gestellt) wurde hkas mit den Primern 5'hkas (*NcoI*) bzw. 5'hkas (*SphI*) und 3'hkas (*NotI*) mittels PCR amplifiziert. Die erhaltenen Produkte konnten anschließend jeweils über die Restriktionsschnittstellen *NcoI* bzw. *SphI* und *NotI* in pHCE13a und pTNA13a eingefügt werden. Die Nukleotidsequenzen aller erstellten Konstrukte (Tabelle 8) wurden jeweils durch Sequenzierung mit vektorspezifischen Primern bestätigt.

Tabelle 8: Klonierte Konstrukte zur Analyse der Expressionsbedingungen von *egfp*-Fusionsgenen.

<i>egfp</i> - Fusionsplasmid	eingefügte Inserts	kodierter Linker zu eGFP	Expressions-system
pER13a	<i>hpcsk9</i> ³¹⁻⁴⁵⁴ (S386A), <i>hgr-lbd</i> , <i>tmhisF</i> , <i>hkas</i>	GGRAGSAAGSGEF	stark induzierbar
pER13a_GSG	<i>tmhisF</i>	GSG	stark induzierbar
pHCE13a	<i>hpcsk9</i> ³¹⁻⁴⁵⁴ (S386A), <i>hgr-lbd</i> , <i>tmhisF</i> , <i>hkas</i>	GGRAGSAAGSGEF	stark konstitutiv
pTNA13a	<i>hpcsk9</i> ³¹⁻⁴⁵⁴ (S386A), <i>hgr-lbd</i> , <i>tmhisF</i> , <i>hkas</i>	GGRAGSAAGSGEF	schwach konstitutiv

4.2.1.2 Analyse verschiedener Expressionssysteme

Die Expression der verschiedenen *egfp*-Fusionsgene zur *in vivo* Analyse der Proteinlöslichkeit wurde bei 37 °C in *E. coli* BL21(DE3) durchgeführt (3.2.8.1) und mit Hilfe von Durchflusszytometrie (3.2.9), Fluoreszenzmikroskopie (3.2.10) und *Western-Blotting* (3.5.4) überprüft. Darüber hinaus wurden die Expressionskulturen, welche die *egfp*-Fusionskonstrukte konstitutiv exprimieren, überimpft und die Zell-Fluoreszenz nochmals durch Durchflusszytometrie analysiert. Dies sollte zeigen, ob die verwendeten konstitutiven Expressionssysteme reproduzierbare Ergebnisse liefern und somit für zeitlich schnell aufeinanderfolgende *Screening*-Runden geeignet wären.

4.2.1.2.1 Konstitutive Expression von *egfp*-fusionierten Genen

Bei schwach konstitutiver Expression durch den Vektor pTNA13a blieb die durch die verschiedenen eGFP-Fusionsproteine verursachte Zell-Fluoreszenz bei überimpften Kulturen konstant (Abbildung 31a). Laut *Western-Blot* Analyse werden allerdings die eGFP-Fusionsproteine bis auf tmHisF-eGFP nicht korrekt hergestellt (Abbildung 31b). So wird die Negativkontrolle hKAS-eGFP überhaupt nicht hergestellt und bei den potentiellen Zielproteinen hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP sind lediglich geringe Mengen eines eGFP-beihaltenden Abbauprodukts vorhanden (~30 kDa), was zu einer sehr schwachen Fluoreszenz der Zellen führt (Abbildung 31a,c). So musste in diesen Fällen bei der

Fluoreszenz-Mikroskopie auch eine sehr hohe Belichtungsdauer angelegt werden (10000 ms), um die Zellen überhaupt sichtbar machen zu können (Abbildung 31c). Aufgrund dieser Artefakt-Anfälligkeit eignet sich pTNA13a nicht als *egfp*-Fusionsvektor zur Analyse von Proteinlöslichkeit *in vivo*.

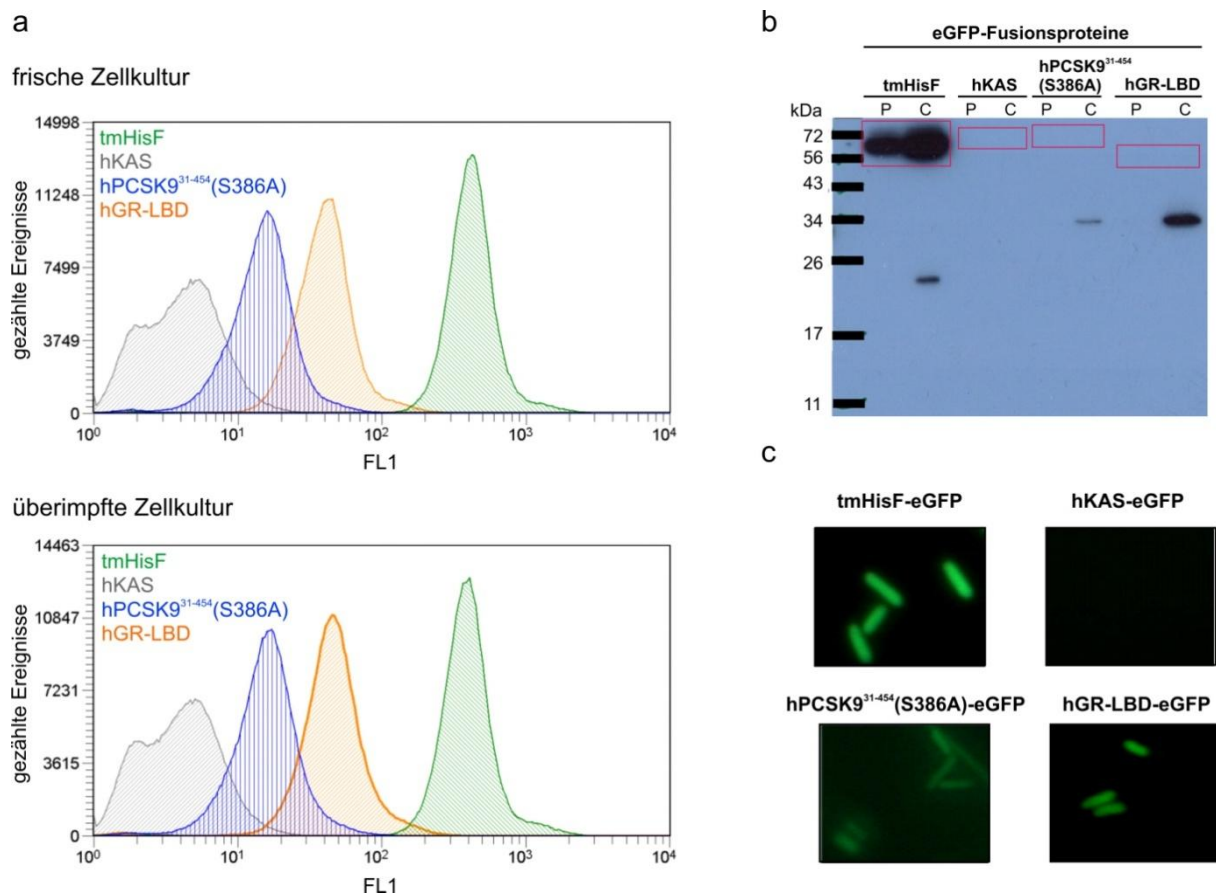


Abbildung 31: Analyse der schwach konstitutiven Expression von *egfp*-fusionierten Genen durch pTNA13a.

Die Herstellung der verschiedenen eGFP-Fusionsproteine erfolgte in *E. coli* BL21(DE3), 6 h bei 37 °C. **a** Analytische FACS-Histogramme der verschiedenen eGFP-Fusionsproteine. Es wurden jeweils 300.000 Ereignisse einer frischen und einer überimpften Zellkultur analysiert. **b** Western-Blot Detektion der erzeugten eGFP-Fusionsproteine durch Anti-GFP Antikörper. Die Bereiche mit den beobachteten bzw. erwarteten Banden der eGFP-Fusionsproteine sind durch rot umrandete Rechtecke markiert (tmHisF-eGFP: 57 kDa, hKAS-eGFP: 72 kDa, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP: 74 kDa, hGR-LBD-eGFP: 56,5 kDa); P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*). Die Abbau-Banden bei ~30 kDa stellen vermutlich wildtypisches eGFP dar. **c** Fluoreszenz-Mikroskopie der eGFP-Fusionsprotein - herstellenden Zellen. Die Belichtungsdauer betrug 500 ms (tmHisF-eGFP) bzw. 10000 ms (hKAS-eGFP, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP).

Bei stark konstitutiver Expression durch den Vektor pHCE13a werden alle Proteine zum Teil korrekt gebildet, wobei sich wie erwartet die Positivkontrolle tmHisF-eGFP zu einem höheren Anteil in der löslichen Zellfraktion befindet als die Negativkontrolle hKAS-eGFP bzw. die potentiellen Zielproteine hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP (Abbildung 32b). Dies führt bei Zellen die tmHisF-eGFP beinhalten auch zu einer starken Fluoreszenz im Cytoplasma (Abbildung 32c). Zellen mit hKAS-eGFP, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP oder

hGR-LBD-eGFP zeigen eine schwächere Fluoreszenz (Abbildung 32a), wobei sich diese neben dem Cytoplasma (Lokalisation des löslichen eGFP-Fusionsproteins) vor allem an den Zell-Polen aufkonzentriert (Abbildung 32c), wo bekanntermaßen unlösliche Proteine als *inclusion bodies* abgelagert werden (siehe 1.1.3, Abbildung 2). Die höhere Löslichkeit der Positivkontrolle tmHisF-eGFP im Vergleich zu hKAS-eGFP, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP wird demnach in diesem System wie erwünscht durch die Fluoreszenz der Zellen dargestellt.

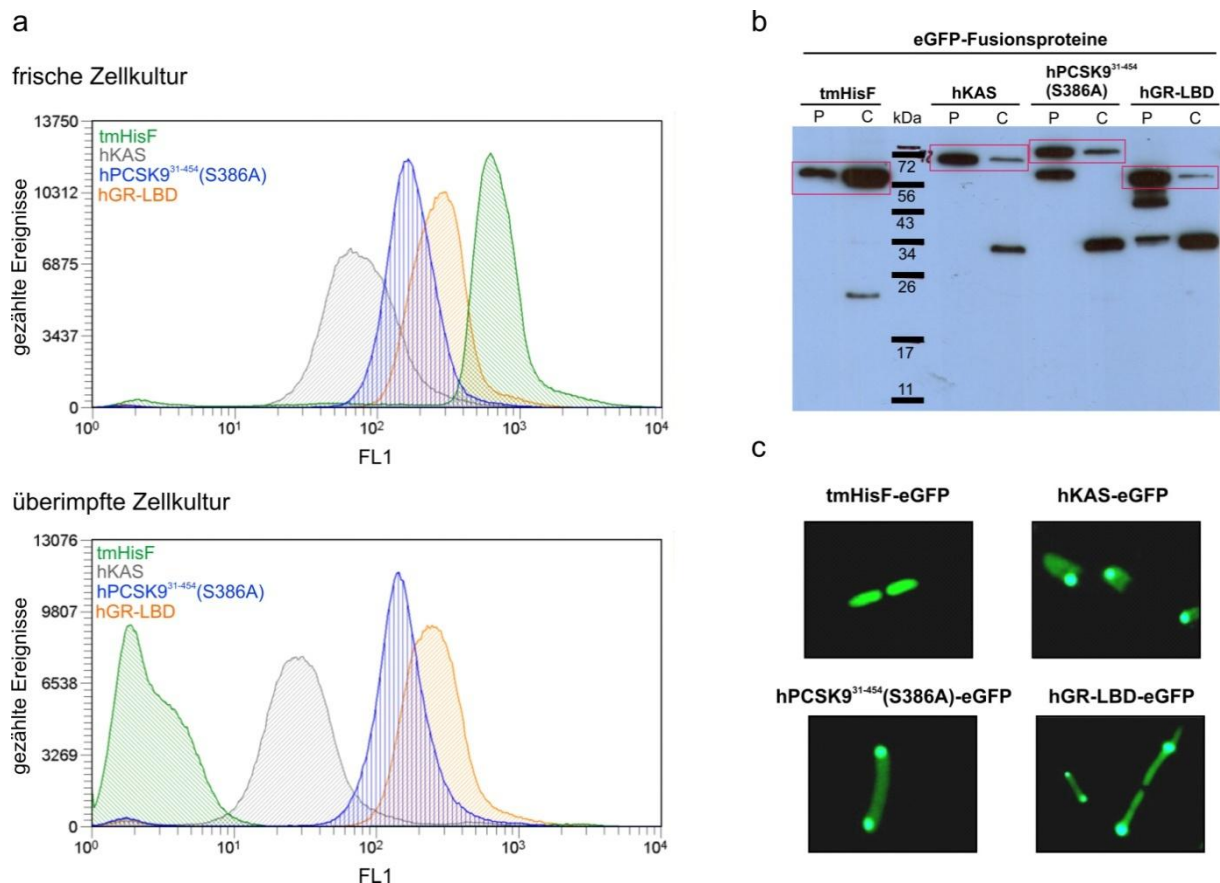


Abbildung 32: Analyse der stark konstitutiven Expression von egfp-fusionierten Genen durch pHCE13a.

Die Herstellung der verschiedenen eGFP-Fusionsproteine erfolgte in *E. coli* BL21(DE3), 6 h bei 37 °C.

a Analytische FACS-Histogramme der verschiedenen eGFP-Fusionsproteine. Es wurden jeweils 300.000 Ereignisse einer frischen und einer überimpften Zellkultur analysiert. **b** *Western-Blot* Detektion der erzeugten eGFP-Fusionsproteine durch Anti-GFP Antikörper. Die Bereiche mit den beobachteten bzw. erwarteten Banden der eGFP-Fusionsproteine sind durch rot umrandete Rechtecke markiert (tmHisF-eGFP: 57 kDa, hKAS-eGFP: 72 kDa, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP: 74 kDa, hGR-LBD-eGFP: 56,5 kDa); P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*). Die Abbau-Banden bei ~30 kDa könnten wildtypisches eGFP darstellen. Ein quantitativer Vergleich der Banden ist bei dieser Detektion nicht möglich, da die Spezifität des benutzten Antikörpers womöglich unterschiedlich ist für die verschiedenen eGFP-Abbauprodukte. **c** Fluoreszenz-Mikroskopie der eGFP-Fusionsprotein - herstellenden Zellen. Die Belichtungsdauer betrug 800 ms (tmHisF-eGFP), 3000 ms (hKAS-eGFP), 2000 ms (hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP) bzw. 1500 ms (hGR-LBD-eGFP).

Die Analyse der überimpften pHCE13a-transformierten Expressionskulturen zeigte allerdings, dass die durch ein bestimmtes eGFP-Fusionsprotein vermittelte Fluoreszenz nicht

konstant bleibt (Abbildung 32a), weswegen dieses Expressionssystem für die geplante gelenkte Evolution hoher Proteinlöslichkeit und Proteinstabilität ebenfalls ungeeignet ist. Zudem weisen die löslichen Zellfraktionen der hKAS-eGFP, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP herstellenden Zellen ein eGFP-beinhaltenes Abbauprodukt bei ca. 30 kDa auf (Abbildung 32b). Ein quantitativer Vergleich der Vollängen-Bande und dieser Abbaubande ist beim immunologischen Nachweis durch *Western-Blotting* leider nicht möglich, da die Spezifität des benutzten Antikörpers womöglich unterschiedlich ausfällt. Inwiefern solche Abbauprodukte die Analyse der eGFP-Fusionsprotein - vermittelten Fluoreszenz hier stören bleibt demnach unklar.

4.2.1.2.2 Induzierbare Expression von *egfp*-fusionierten Genen

4.2.1.2.2.1 eGFP-Fusionsproteine mit langem Linker (13 AS)

Bei induzierbarer Expression durch pER13a bzw. pER13b, (→ eGFP-Fusionsprotein mit GGRAGSAAGSGEF –Linker) findet sich die Positivkontrolle tmHisF-eGFP wie erwartet überwiegend in der löslichen Zellfraktion (Abbildung 33b), was zu einer hohen Fluoreszenz des Cytoplasmas führt (Abbildung 33a,c). hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP findet sich dagegen hauptsächlich in der unlöslichen Zellfraktion (Abbildung 33b) und die Zellen mit diesem Protein zeigen demnach auch eine niedrigere Fluoreszenz als *tmhisF-egfp* exprimierende Zellen (Abbildung 33a), wobei vor allem die Zell-Pole (= *inclusion bodies*) schwach fluoreszieren (Abbildung 33c). hGR-LBD-eGFP und hKAS-eGFP werden ausschließlich unlöslich hergestellt (Abbildung 33b), sind also komplett in *inclusion bodies* verpackt (Abbildung 33c), was zu einer noch geringen Fluoreszenz der Zellen führt (Abbildung 33a). Expression durch pER13a oder pER13b zeigte in allen Fällen keine signifikanten Unterschiede. Die Unterschiede zwischen den Fluoreszenzen der verschiedenen Konstrukte konnten bei neu transformierten Kulturen stets reproduziert werden. Die Absolutwerte der Fluoreszenz differierten lediglich aufgrund der unterschiedlichen täglichen Kalibrierung des MoFlo[®] Zytometers, weswegen im Folgenden alle Analysen innerhalb eines gezeigten FACS-Histogramms stets parallel aufgenommen wurden. Frische Transformation von Zellen mit den induzierbaren Vektoren stellt somit sicher, dass stets vergleichbare Fluoreszenz erhalten wird und keine Schwankungen wie bei pHCE13a (4.2.1.2.1) auftreten können. Zudem spiegelt sich hier die unterschiedliche Löslichkeit der Positivkontrolle tmHisF-eGFP, der Negativkontrolle hKAS-eGFP und der potentiellen Zielproteine hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP sehr gut in der Fluoreszenz der Expressionskulturen *in vivo* wider. Aus diesen Gründen

wurden im weiteren Verlauf dieser Arbeit nur noch pER13a und pER13b als *egfp*-Fusionsvektoren benutzt.

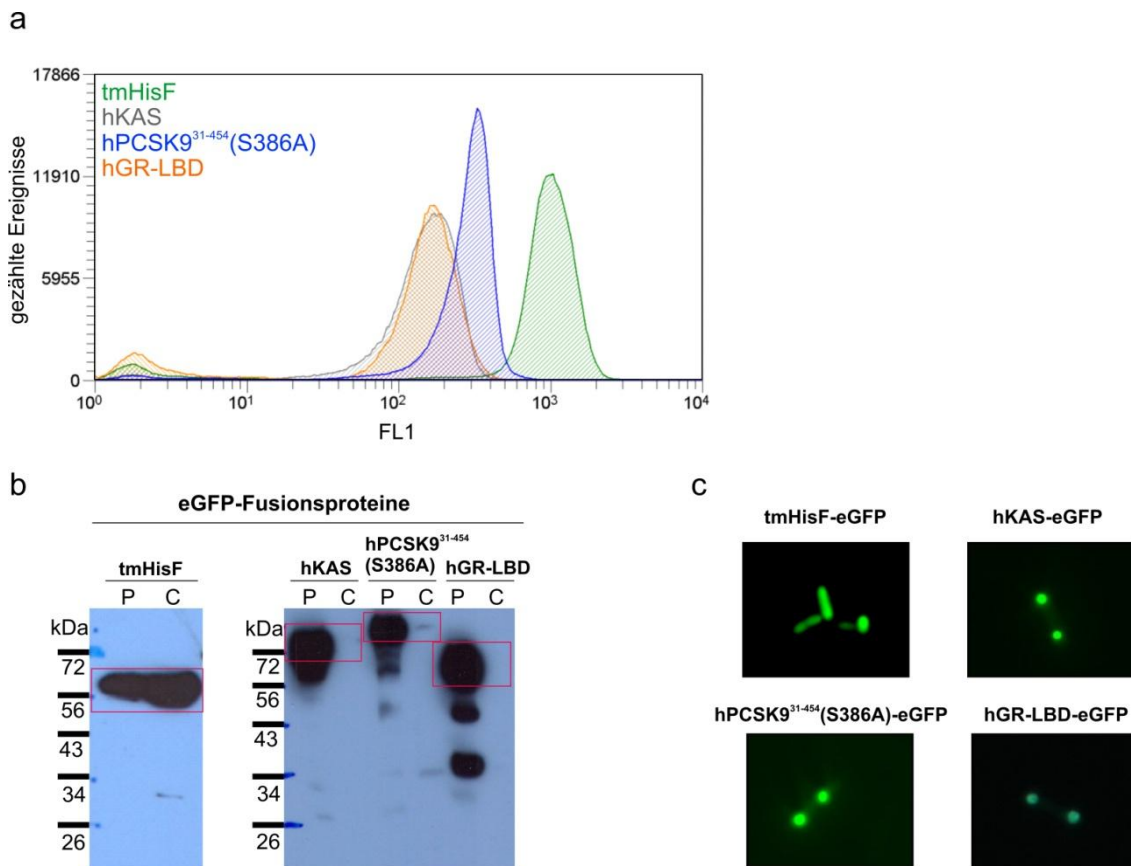


Abbildung 33: Analyse der induzierbaren Expression von *egfp*-fusionierten Genen durch pER13a.

Die Herstellung der verschiedenen eGFP-Fusionsproteine erfolgte in *E. coli* BL21(DE3), 4 h bei 37 °C.

a Analytische FACS-Histogramme der verschiedenen eGFP-Fusionsproteine. Es wurden jeweils 300.000 Ereignisse analysiert. **b** *Western-Blot* Detektion der erzeugten eGFP-Fusionsproteine durch Anti-GFP Antikörper. Die Bereiche mit den beobachteten bzw. erwarteten Banden der eGFP-Fusionsproteine sind durch rot umrandete Rechtecke markiert (tmHisF-eGFP: 57 kDa, hKAS-eGFP: 72 kDa, hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP: 74 kDa, hGR-LBD-eGFP: 56,5 kDa); P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*). **c** Fluoreszenz-Mikroskopie der eGFP-Fusionsprotein - herstellenden Zellen. Die Belichtungsdauer betrug 60 ms (tmHisF-eGFP), 200 ms (hKAS-eGFP) bzw. 300 ms (hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP und hGR-LBD-eGFP).

Es muss allerdings erwähnt werden, dass auch bei diesem Expressionssystem die eGFP-Fusionsproteine laut *Western-Blot* Analyse zum Teil proteolytisch abgebaut werden (Abbildung 33b). Inwieweit die Beurteilung der Proteinlöslichkeit *in vivo* mit Hilfe der Zell-Fluoreszenz durch diese Proteolyse-Anfälligkeit gestört wird, wird in 4.2.1.2.2.2 und 4.2.3 näher erläutert.

Die Analyse der potentiellen Zielproteine in diesem induzierbaren Expressionssystem zeigte, dass Zellen mit hGR-LBD-eGFP eine niedrigere Fluoreszenz als Zellen mit hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP aufweisen. hGR-LBD-eGFP besitzt somit ein größeres Potential für Verbesserungen hinsichtlich der erzeugten Fluoreszenz *in vivo*, weswegen sich dieses Protein besser für die geplante gelenkte Evolution hoher Löslichkeit und Stabilität eignet als

hPCSK9³¹⁻⁴⁵⁴(S386A)-eGFP. Eine nähere Charakterisierung von hGR-LBD-eGFP bzw. dessen Optimierung durch gelenkte Evolution wird in 4.2.3 bzw. 4.3 gezeigt.

4.2.1.2.2.2 eGFP-Fusionsproteine mit kurzem Linker (3 AS)

Die in Abbildung 32b bzw. 33b dokumentierte Anfälligkeit der eGFP-Fusionsproteine für proteolytischen Abbau könnte die Analyse der Proteinlöslichkeit *in vivo* stören, da lösliche eGFP-beinhaltende Abbauprodukte zu einer hohen Hintergrund-Fluoreszenz führen könnten. Um diese Artefakt-Quelle auszuschließen, wurde versucht den Protease-anfälligen internen Linker zu kürzen.

Die Eigenschaften von tmHisF-eGFP mit verkürztem Linker (Gly-Ser-Gly) wurde nach Transformation von *E. coli* BL21(DE3) Zellen mit pER13a_GSG-tmhisF (4.2.1.1) und anschließender Genexpression bei 37 °C (3.2.8.1) durch Durchflusszytometrie (3.2.9) analysiert, wobei keinerlei Fluoreszenz der Zellen mehr festgestellt werden konnte (Abbildung 34). Die Fluoreszenz von eGFP bildet sich demnach bei zu starker Verkürzung des Linkers nicht mehr aus, da vermutlich die korrekte Faltung von eGFP durch ein zu nah fusioniertes Protein behindert wird. Dieser verkürzte Linker stellt somit keine praktikable Alternative zu dem in pTNA13a, pHCE13a und pER13a bzw. pER13b kodierten GGRAGSAAGSGEF -Linker dar.

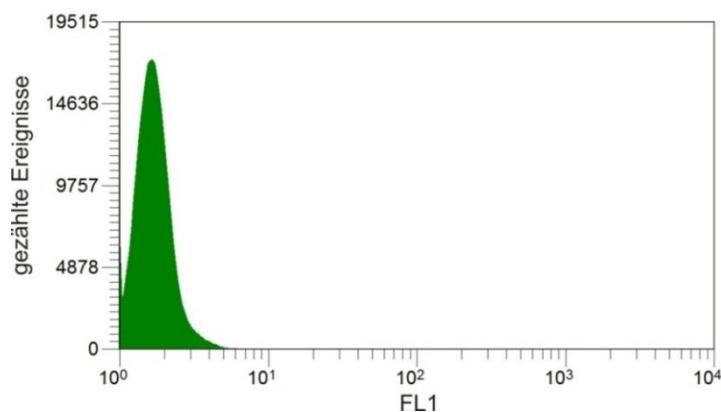


Abbildung 34: Analytisches FACS-Histogramm von tmHisF-eGFP(GSG).

Die Expression von tmhisF-egfp(GSG) erfolgte in *E. coli* BL21(DE3) für 4 h bei 37 °C. Es wurden 300.000 Ereignisse analysiert. Die Zellen zeigen ein charakteristisches Histogramm nicht-fluoreszierender Zellen.

4.2.2 Exemplarische Anreicherung fluoreszierender Zellen durch Durchflusszytometrie

Die Funktionsweise und Effektivität des zur Sortierung fluoreszierender Zellen verwendeten MoFlo[®] Zytometers wurde durch eine exemplarische Anreicherung getestet. Aus einer Mischung unterschiedlicher Zellen sollte hierbei ein im Unterschuss vorhandener Anteil an stark fluoreszierenden Zellen angereichert werden. Hierzu wurde ein Überschuss an *E. coli* BL21(DE3) Zellen, welche hkas-egfp 4 h bei 37 °C exprimierten (schwache Fluoreszenz) mit

einem geringen Anteil an Zellen, welche *tmhisF-egfp* 4 h bei 37 °C exprimierten (starke Fluoreszenz), versetzt. Das Zell-Gemisch wurde mit Hilfe von Durchflusszytometrie (3.2.9) analysiert bzw. sortiert (Abbildung 35). Nach Festlegung eines Sortier-Bereichs (SB-1) in welchem sich laut Kontrolle die Zellen mit *tmHisF-eGFP* befinden sollten, wurden diese in zwei Runden angereichert. In Runde 1 wurden $3,4 \cdot 10^7$ Zellen des Gemisches analysiert, wobei $1 \cdot 10^5$ (0,3 %) Zellen sortiert wurden. Diese Zellen wurden ohne erneute Anzucht direkt nochmals sortiert. Bei diesem *Resort* wurden von $2,1 \cdot 10^4$ analysierten Zellen, $1 \cdot 10^4$ Zellen sortiert, deren Plasmide nach Anzucht üN bei 37 °C in LB_{Kan}-Medium präpariert wurden. Anschließend wurde eine Retransformation von *E. coli* BL21(DE3) Zellen mit dem gewonnenen Plasmid-Gemisch durchgeführt, gefolgt von der Expression der angereicherten *egfp*-Fusionsgene 4 h bei 37 °C. Die FACS-Analyse dieser Anreicherung 1 zeigte, dass bereits 35 % der Zellen eine Fluoreszenz in SB-1 aufwiesen (Abbildung 35). Das *tmhisF-egfp* Fusionskonstrukt konnte demnach nach einer Anreicherungsrunde bereits 116-fach angereichert werden.

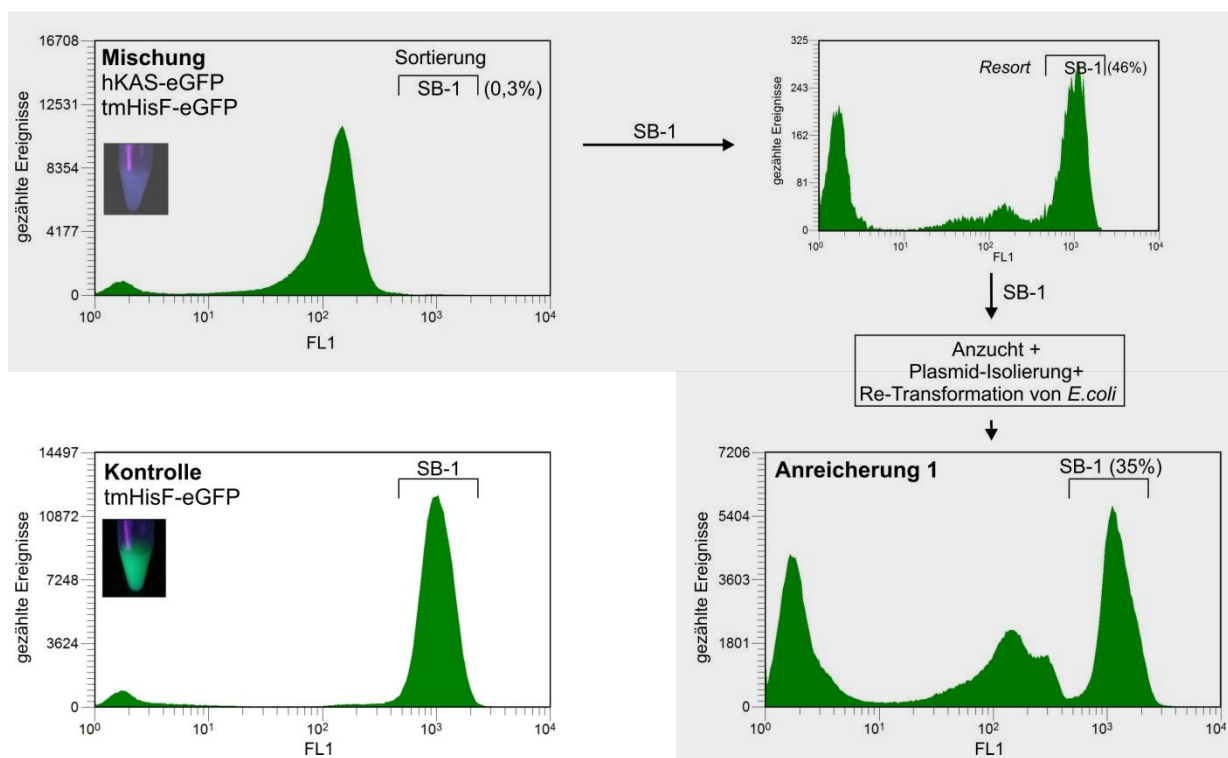


Abbildung 35: Exemplarische Anreicherung stark fluoreszierender Zellen durch Durchflusszytometrie.

Die Herstellung der eGFP-Fusionsproteine erfolgte in mit pER13a-Konstrukten transformierten *E. coli* BL21(DE3) Zellen, 4 h bei 37 °C nach Induktion mit je 0,5 mM IPTG.

Weiß hinterlegt: Zellen die ausschließlich *tmhisF-egfp* exprimieren (Zellsuspension zeigt deutliche Fluoreszenz nach Anregung bei 302 nm) dienten als Kontrolle und zur Festlegung des Sortier-Bereichs SB-1.

Grau hinterlegt (von links oben nach rechts unten): Aus einer Mischung von *hkas-egfp* und *tmhisF-egfp* exprimierenden Zellen (Zellsuspension zeigt keine Fluoreszenz nach Anregung bei 302 nm) wurden die Zellen mit *tmHisF-eGFP* angereichert. Auf die Sortierung von $1 \cdot 10^5$ aus $3,4 \cdot 10^7$ analysierten Zellen (0,3 %) folgte ein *Resort* von 10.000 aus 21.000 Zellen (46 %). Diese Zellen wurden üN bei 37 °C in LB-Selektivmedium angezogen, woraufhin die enthaltenen Plasmide isoliert werden konnten. Bei der FACS-Analyse dieser so erhaltenen Anreicherung 1 befanden sich 35 % der Zellen in SB-1.

Nach einer zweiten solchen Anreicherungsrunde befanden sich 68 % der Zellen in SB-1, was einem Anreicherungsfaktor von 227 entspricht (Abbildung 36). Einige dieser Zellen wurden mit Hilfe der Cyclone Sortiereinheit auf LB-Selektivagarplatten vereinzelt. Da die Expression von *tmhisF* durch *pER13a_tmhisF* auch ohne Induktion schwach erfolgt und zu einer starken Fluoreszenz führt (Dr. Jörg Claren, mündliche Mitteilung), konnten mit Hilfe einer UV-Lampe (Anregung mit 302 nm) alle so sortierten Zellen direkt als mit *pER13a_tmhisF* transformierte Klone identifiziert werden (Abbildung 36).

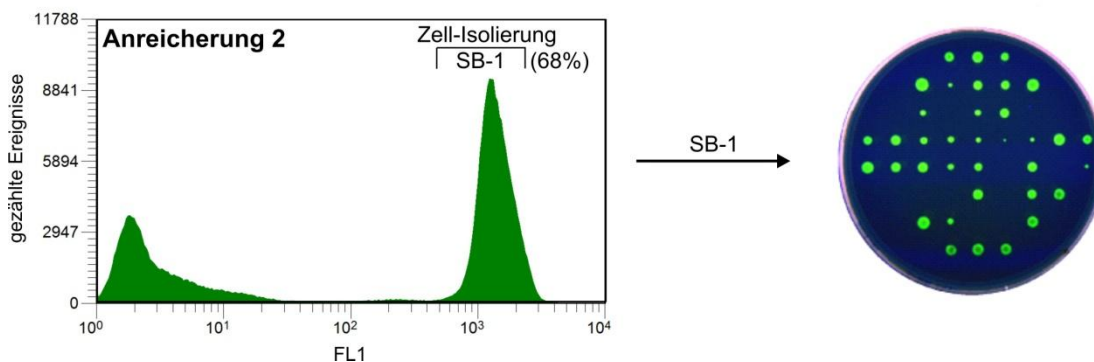


Abbildung 36: Zweite Runde der exemplarischen Anreicherung.

Nach Wiederholung des in 3.2.9 beschriebenen und in Abbildung 35 gezeigten Vorgehens wurde Anreicherung 2 erhalten. Hier befinden sich nun bereits 68 % aller Zellen in SB-1. Auf LB-Selektivagarplatten vereinzelt Zellen aus SB-1 von Anreicherung 2 zeigten nach Anregung bei 302 nm deutliche Fluoreszenz, woraus geschlossen werden kann, dass es sich bei allen vereinzelt Zellen um mit *pER13a_tmhisF* und nicht mit *pER13a_hkas* transformierte Klone handelt. Die gewünschte Anreicherung des *tmhisF-egfp* Fusionskonstrukts war somit erfolgreich.

Die in 3.2.9 erläuterte und hier näher vorgestellte Durchflusszytometrie ist demnach sehr effektiv. Bereits nach wenigen Runden kann eine starke Anreicherung und Isolierung der gewünschten Zellen bzw. *egfp*-Fusionskonstrukte erfolgen, wobei eine sehr hohe Anzahl an Zellen (10^7 - 10^8) innerhalb kurzer Zeit (bei 10000 eps: ca. 20 – 200 min) analysiert werden kann. Diese Methode kann somit zum Hochdurchsatz-*Screening egfp*-fusionierter Genbanken bei der gelenkten Evolution von Proteinen hin zu hoher Löslichkeit und Stabilität (siehe 3.2.8.1, 4.3) angewendet werden.

4.2.3 Nähere Charakterisierung von hGR-LBD-eGFP

hGR-LBD bietet als Zielprotein für die geplante gelenkte Evolution hoher Proteinlöslichkeit einige Vorteile. Zum Einen können die Bedingungen für die Herstellung von hGR-LBD in *E. coli* durch Zugabe des stabilisierenden hGR-Liganden Dexamethason (siehe 1.7, Abbildung 8) variiert werden, wodurch der Selektionsdruck bei der gelenkten Evolution unterschiedlich justiert werden kann. Zum Anderen existiert bereits eine optimierte hGR-LBD

Variante (hGR-LBD(F602S); Bledsoe *et al.*, 2002) die bei den folgenden Experimenten als Positivkontrolle benutzt werden kann. Bisher gelang die Herstellung von hGR-LBD im löslichen Zellextrakt von *E. coli* nur unter folgenden Bedingungen: Mutation von Phenylalanin 602 zu Serin (Bledsoe *et al.*, 2002) oder Tyrosin (Biggadike *et al.*, 2008), Fusion an die Glutathion-S-Transferase und Inkubation der exprimierenden Zellen bei niedrigen Temperaturen in Anwesenheit von Dexamethason.

Um die Auswirkung der F602S Mutation auf hGR-LBD-eGFP zu untersuchen wurde ausgehend von pER13b *hgr-lbd-egfp* bzw. *hgr-lbd(F602S)-egfp* (Konstrukte wurden von Dr. Ralf Thoma zur Verfügung gestellt) unter verschiedenen Bedingungen in *E. coli* BL21(DE3) exprimiert (3.2.8.1). Die Fluoreszenz der Expressionskulturen wurde mit Hilfe von Durchflusszytometrie (3.2.9) analysiert. Nach Expression bei 30 °C in Anwesenheit von Dexamethason zeigten Zellen mit hGR-LBD(F602S)-eGFP eine höhere Fluoreszenz als Zellen mit hGR-LBD-eGFP (Abbildung 37a). Dies ist bei 37 °C ohne Zugabe von Dexamethason in geringerem Maße ebenfalls der Fall (Abbildung 37b). Die durch F602S induzierte verbesserte Löslichkeit von hGR-LBD kann demnach mit Hilfe von eGFP als Reporter *in vivo* dargestellt werden. Die in Abbildung 32b bzw. 33b beobachtete Protease-Anfälligkeit der eGFP-Fusionsproteine stört hier somit offensichtlich nicht die Beurteilung der Menge an löslichem Protein *in vivo* über die Analyse der Zell-Fluoreszenz. Neue Mutationen in hGR-LBD die zu einer verbesserten löslichen Produktion des Proteins in *E. coli* führen, sollten also mit Hilfe von eGFP als Reporter identifiziert werden können.

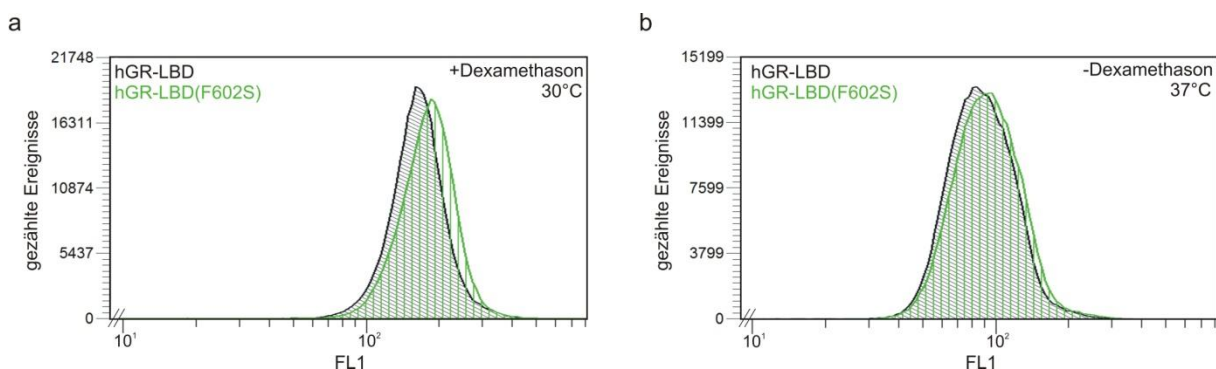


Abbildung 37: Analytische FACS-Histogramme von Zellen mit eGFP-fusioniertem hGR-LBD (schwarz) und hGR-LBD(F602S) (grün).

Es wurden je 300.000 Ereignisse analysiert.

a Die Expression der *egfp*-fusionierten Gene erfolgte in *E. coli* BL21(DE3), 6 h bei 30 °C nach Zugabe von 0,5 mM IPTG und 500 μ M Dexamethason.

b Die Expression der *egfp*-fusionierten Gene erfolgte in *E. coli* BL21(DE3), 4 h bei 37 °C nach Zugabe von 0,5 mM IPTG ohne Dexamethason.

Aufgrund der Ergebnisse dieser Vorarbeiten soll nun im Folgenden die gelenkte Evolution von hGR-LBD unter beiden genannten Bedingungen durchgeführt werden (4.3). Beim *Screening* unter moderaten Bedingungen (30 °C in Anwesenheit von Dexamethason) könnten hGR-LBD Varianten gefunden werden, die eine ähnliche Verbesserung wie hGR-LBD(F602S) zeigen und weiterhin abhängig sind von der Anwesenheit des Liganden. Beim *Screening* unter strengeren Bedingungen (37 °C ohne Dexamethason) könnten hGR-LBD Varianten gefunden werden, die eine stärkere Verbesserung zeigen und unabhängig vom Liganden löslich in *E. coli* hergestellt werden können.

4.3 Erhöhung der Löslichkeit und der Stabilität von hGR-LBD durch gelenkte Evolution mit Hilfe von eGFP als Reporter

Wie in 4.2 gezeigt wurde, kann sich die Löslichkeit eines Proteins *in vivo* in der Fluoreszenz eines C-terminal fusionierten eGFPs widerspiegeln. Wie bereits in anderen Arbeiten beschrieben (Pedelacq *et al.*, 2002; Yang *et al.*, 2002; Yang *et al.*, 2003) soll nun mit Hilfe dieses Reportersystems die Löslichkeit und Stabilität eines unlöslichen Proteins durch Zufallsmutagenese gefolgt von *Screening* nach verbesserten Varianten („gelenkte Evolution“) so weit erhöht werden, dass weitergehende biophysikalische Untersuchungen (z.B. Röntgenkristallstrukturanalyse) des Proteins ermöglicht bzw. vereinfacht werden. Bei dieser Methode der Proteinstabilisierung werden vom Zielprotein keinerlei strukturelle Daten benötigt. Darüber hinaus muss für das *Screening*-Verfahren auch keine spezifische Aktivität beim Zielprotein vorhanden sein, weshalb diese Methode für alle denkbaren Polypeptide anwendbar ist.

Aufgrund der in 4.2 gezeigten Vorarbeiten wurde beschlossen bei der Durchführung diesen Verfahrens pER13a bzw. pER13b als *egfp*-Expressionsvektoren, Durchflusszytometrie als Hochdurchsatz-*Screening*-Verfahren und hGR-LBD als Zielprotein zu benutzen. Durch Zufallmutagenese *in vitro* mittels *error prone* PCR (*epPCR*) sollte ein *hgr-lbd* Genrepertoire erzeugt und mit *egfp* in pER13b fusioniert werden. Nach Transformation eines *E. coli* Expressionsstamms mit dem Plasmid-Repertoire sollten dann stark fluoreszierende Zellen mit Hilfe des MoFlo[®] Zytometers angereichert werden. Nach Abtrennung von eGFP sollten die angereicherten hGR-LBD Varianten gereinigt, bezüglich ihrer Stabilität und Löslichkeit untersucht und der Einfluß einzelner Mutationen analysiert werden (Abbildung 38).

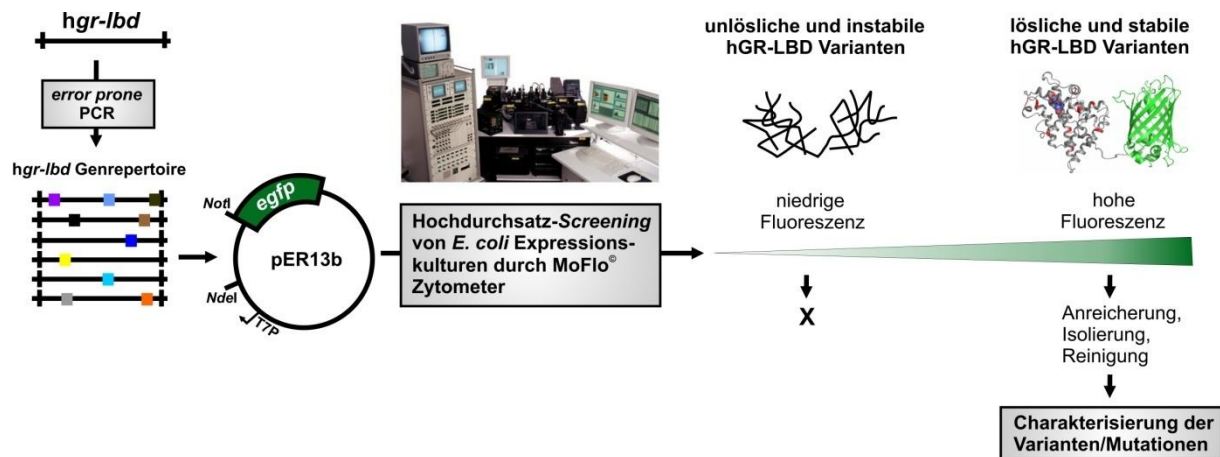


Abbildung 38: Strategie zur Erhöhung der Löslichkeit und Stabilität von hGR-LBD durch gelenkte Evolution.

Durch Zufallsmutagenese via *error prone PCR* und Fusion von *hgr-lbd* mit *egfp* werden *hgr-lbd-egfp* Fusionskonstrukte hergestellt, die in *hgr-lbd* randomisiert sind. Nach Transformation von *E. coli* BL21(DE3) mit dem Plasmid-Repertoire werden die Fusionsgene exprimiert, was das *Screening* auf hohe Fluoreszenz ermöglicht. Es werden nur solche Zellen angereichert, die eine lösliche hGR-LBD-eGFP Variante herstellen. Die hGR-LBD Varianten werden anschließend mit und ohne eGFP weiter charakterisiert, wobei die enthaltenen Mutationen auch vereinzelt untersucht werden.

Aufnahme des MoFlo[®] Zytometers aus http://citometria.blogspot.com/2007_04_01_archive.html.

4.3.1 Herstellung von *hgr-lbd* Genbanken in pER13b

Um eine hohe Diversität an *hgr-lbd* Mutanten *screenen* zu können, wurden zwei verschiedene Genbanken in pER13b hergestellt (3.3.9), wobei jeweils das komplette *hgr-lbd* Gen durch *epPCR* randomisiert wurde (3.3.3.3). Beide Genbanken sollten eine unterschiedliche Mutationsfrequenz aufweisen, um gleichzeitig hGR-LBD Varianten mit wenigen und vielen Aminosäureaustauschen untersuchen zu können. Hierfür wurden zwei unterschiedliche *epPCR*-Protokolle verwendet (bei Genbank 1 bzw. 2: *epPCR*-1 bzw. -2, siehe 3.3.3.3). In beiden Fällen wurde das Genrepertoire mit Hilfe der vektorspezifischen Primer 5'pER13 und 3'pER13 aus dem Templat pER13a_*hgr-lbd* amplifiziert und über die Restriktionsschnittstellen *NdeI* und *NotI* in pER13b eingefügt. Die Präparationen von Genbank 1 bzw. 2, die nach dem in 3.3.9 beschriebenen Protokoll hergestellt wurden, enthielten $3,0 \cdot 10^5$ bzw. $1,2 \cdot 10^6$ Kolonien auf je 40 großen LB-Selektivagarplatten. Kolonie-PCR (3.3.2) von je 20 Klonen ergab, dass die Ligationseffizienz bei 100 % bzw. 95 % lag. Daraus errechnen sich Genbank-Größen von $3,0 \cdot 10^5$ (Genbank 1) bzw. $1,1 \cdot 10^6$ (Genbank 2) unabhängigen Varianten. Zur Bestimmung der Anzahl und Verteilung der Mutationen wurden jeweils zehn Klone vollständig sequenziert und mit der *hgr-lbd* Ausgangssequenz verglichen (Tabelle 9).

Tabelle 9: Anzahl und Verteilung der Nukleotid-Austausche bei zufällig ausgewählten Klonen der *hgr-lbd* Genbank 1 und 2.

Es wurden jeweils 10 Klone sequenziert. Ts: Transitionen, Tv: Transversionen. Bei einem Klon in Genbank 2 wurde zusätzlich eine Deletion festgestellt.

	<i>hgr-lbd</i> Genbank 1	<i>hgr-lbd</i> Genbank 2
Transitionen	73,2%	45,5%
T → C	22,1%	27,3%
C → T	8,1%	0,0%
A → G	40,7%	9,1%
G → A	2,3%	9,1%
Transversionen	26,8%	54,5%
A → T	10,5%	18,1%
T → A	5,8%	27,3%
G → T	2,3%	0,0%
T → G	1,2%	0,0%
A → C	2,3%	0,0%
C → A	3,5%	9,1%
G → C	0,0%	0,0%
C → G	1,2%	0,0%
Ts/Tv	2,7	0,8
AT-Austausche	82,6%	81,8%
GC-Austausche	17,4%	18,2%
Mutationen/Gen	7,2 (2-13)	1,8 (1-3)
Größe der Genbank	~ 3,0*10 ⁵	~ 1,1*10 ⁶

Dabei zeigten beide Genbanken wie gewünscht eine unterschiedlich starke Mutationsfrequenz von 7,2 (2-13) bzw. 1,8 (1-3) Mutationen pro Sequenz. Bei beiden Genbanken überwiegen die AT-Austausche im Vergleich zu GC-Austauschen, was bei der verwendeten MnCl₂-induzierten *ep*PCR die Regel ist (Vanhercke *et al.*, 2005). Das Verhältnis von Transitionen zu Transversionen ist bei Genbank 2 ausgeglichen (Ts/Tv: 0,8), während Genbank 1 einen Überschuss an Transitionen (Ts/Tv: 2,7) zeigt.

Die Wahrscheinlichkeit, dass eine Aminosäure ausgetauscht wird, beträgt an Position eins eines Codons 96 %, an Position zwei 99 % und an Position drei 32 %. Durchschnittlich ergibt dies eine Wahrscheinlichkeit von 75 %, dass ein Basenaustausch den Austausch einer Aminosäure zur Folge hat. Das bedeutet, dass mit den 7,2 bzw. 1,8 Nukleotidaustauschen pro Gen in der *hgr-lbd* Genbank 1 bzw. 2 im Durchschnitt 5,4 bzw. 1,3 Aminosäureaustausche erzeugt würden.

4.3.2 Screening der Genbanken

4.3.2.1 Screening bei 37 °C in Abwesenheit eines Liganden

Die in 4.3.1 hergestellten *hgr-lbd* Genbanken wurden bei 37 °C in Abwesenheit eines Liganden nach verbesserten Varianten durchsucht. Hierzu wurden *E. coli* BL21(DE3) Zellen wie in 3.2.8.1 beschrieben mit den *hgr-lbd* Genbanken transformiert und angezogen, wobei die Expression unter den genannten Bedingungen mit 0,5 mM IPTG induziert wurde. Die Anzahl der mit Genbank 1 bzw. 2 transformierten und angeimpften Zellen betrug $4,4 \cdot 10^7$ bzw. $3,1 \cdot 10^7$ und überstieg damit die Größe der Genbank 1 ($3,0 \cdot 10^5$) bzw. 2 ($1,1 \cdot 10^6$) um den Faktor 147 bzw. 28.

Die die Genbanken 1 und 2 exprimierenden Zellen wurden gemischt und anschließend durch Durchflusszytometrie (3.2.9) analysiert und sortiert. Analytische FACS-Histogramme von Zellen mit den hGR-LBD Banken und Zellen mit hGR-LBD Wildtyp (jeweils fusioniert an eGFP) zeigen, dass die Mehrzahl der in den Banken enthaltenen hGR-LBD Varianten unter den angelegten Bedingungen zu einer niedrigeren Fluoreszenz führen als der Wildtyp (Abbildung 39).

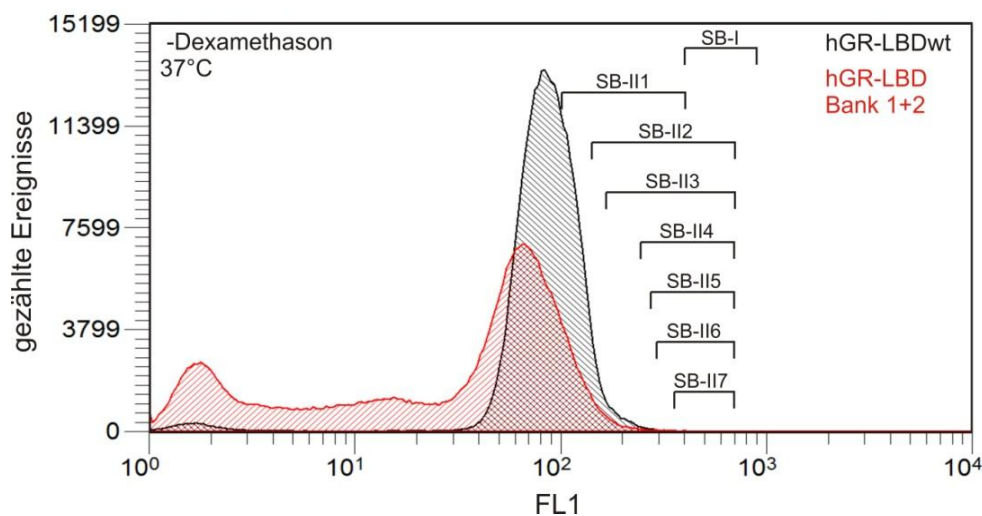


Abbildung 39: Analytische FACS-Histogramme von Zellen mit hGR-LBDwt (schwarz) und hGR-LBD Bank 1+2 (rot), nach Expression bei 37 °C ohne Zugabe eines Liganden.

Die Expression der *egfp*-fusionierten Gene aus pER13b erfolgte jeweils in *E. coli* BL21(DE3), für 4 h bei 37 °C. Es wurden je 300.000 Ereignisse analysiert. Die einzelnen Sortier-Bereiche der Versuchsreihe I (SB-I) bzw. der Anreicherungsrounden von Versuchsreihe II (SB-II-1-7) sind durch Klammern markiert.

Zellen mit löslicheren Varianten sollten stärker fluoreszieren als Zellen, die das wildtypische Protein herstellen, weshalb diese schrittweise mittels Durchflusszytometrie (3.2.9) angereichert wurden. Hierfür wurden zwei verschiedene Versuchsreihen angesetzt. In Versuchsreihe I wurden insgesamt drei Anreicherungsrounden bei konstant hohem Sortier-Bereich (SB-I, FL1: $4 \cdot 10^2 - 9 \cdot 10^2$) durchgeführt (Abbildung 39, Tabelle 10). Bei Versuchsreihe II wurden sieben Anreicherungsrounden durchgeführt, wobei zunächst von

einem relativ niedrigen Sortier-Bereich ausgegangen wurde (SB-III, FL1: $1 \cdot 10^2 - 4 \cdot 10^2$), welcher dann schrittweise erhöht wurde (Abbildung 39, Tabelle 10). Der Sortier-Bereich wurde hierbei stets so gewählt, dass bei jeder Runde die 0,5-5,5 % hellsten Zellen angereichert wurden (Tabelle 10). In beiden Versuchsreihen wurden bei jeder Anreicherungsrunde (Sortierung+Resort, siehe Abbildung 35) die sortierten Zellen in 5 ml LB-Medium aufgefangen und anschließend über Nacht bei 37 °C in 50 ml LB_{Kan}-Medium angezogen (3.2.9). Daraus konnte die Plasmid-DNA präpariert werden (3.3.7.1), welche zur Transformation frisch hergestellter elektrokompenter *E. coli* BL21(DE3) Zellen (3.2.4/3.2.6) benutzt wurde. So konnte die Expression der soeben sortierten Varianten wieder induziert werden (3.2.8.1) und die nächste Anreicherungsrunde durchgeführt werden (3.2.9). Die Anzahl der mit den Anreicherungen transformierten Zellen wurde stets über die Ausplattierung von Verdünnungsreihen auf LB_{Kan}-Agarplatten bestimmt (3.2.8.1). Es wurde so sichergestellt, dass die Anzahl der Transformanten die Anzahl an zuvor sortierten Zellen immer mindestens um den Faktor zehn überstieg und somit alle sortierten Varianten in die nächste Anreicherungsrunde überführt wurden. In Tabelle 10 sind die Ergebnisse der Anreicherungsunden von Versuchsreihe I und II zusammengefasst.

Tabelle 10: FACS-Anreicherungen von *hgr-lbd* Genbank 1+2 exprimierenden *E. coli* BL21(DE3) Zellen bei 37 °C ohne Zugabe eines Liganden.

GB: Genbank, A: Anreicherung, SB: Sortier-Bereich. In Versuchsreihe I wurden aufgrund des hohen Sortier-Bereichs so wenige Zellen sortiert, dass kein *Resort* möglich war. Bei Versuchsreihe II konnte stets ein *Resort* durchgeführt werden. Bei Versuchsreihe II wurde bei den Anreicherungen A3-A7 der vorhandene Anteil an Klonen mit korrektem pER13b_*hgr-lbd* Plasmid über Kolonie-PCR mit den Primern 5'pER13 und 3'pER13 bestimmt (Amplifikationsprodukt: ~1100 bp).

Analysierte Probe	Sortierung		Resort		resultierende Präparation	positiv (Kolonie-PCR)	
	Sortier-Bereich	analysierte Ereignisse	sortierte Ereignisse	analysierte Ereignisse			sortierte Ereignisse
Versuchsreihe I							
GB1+2	SB-I	$1,0 \cdot 10^8$	$3,2 \cdot 10^4$ (0,03%)	-	-	A1	-
A1	SB-I	$1,2 \cdot 10^8$	$8,2 \cdot 10^3$ (0,006%)	-	-	A2	-
A2	SB-I	$8,0 \cdot 10^7$	$4,0 \cdot 10^4$ (0,04%)	-	-	A3	-
Versuchsreihe II							
GB1+2	SB-III	$6,3 \cdot 10^7$	$3,5 \cdot 10^6$ (5,5%)	$3,6 \cdot 10^6$	$2,2 \cdot 10^6$ (61,1%)	A1	-
A1	SB-II2	$2,7 \cdot 10^7$	$1,0 \cdot 10^6$ (3,7%)	$1,0 \cdot 10^6$	$4,1 \cdot 10^5$ (41,0%)	A2	-
A2	SB-II3	$4,2 \cdot 10^7$	$1,2 \cdot 10^6$ (2,9%)	$1,1 \cdot 10^6$	$6,9 \cdot 10^5$ (62,7%)	A3	10/10 (100%)
A3	SB-II4	$1,0 \cdot 10^8$	$1,2 \cdot 10^6$ (1,2%)	$1,1 \cdot 10^6$	$3,2 \cdot 10^5$ (29,1%)	A4	10/10 (100%)
A4	SB-II5	$1,3 \cdot 10^8$	$1,0 \cdot 10^6$ (0,8%)	$9,1 \cdot 10^5$	$3,4 \cdot 10^5$ (37,4%)	A5	9/10 (90%)
A5	SB-II6	$1,3 \cdot 10^8$	$6,5 \cdot 10^5$ (0,5%)	$6,6 \cdot 10^5$	$1,9 \cdot 10^5$ (28,8%)	A6	20/20 (100%)
A6	SB-II7	$1,0 \cdot 10^8$	$4,6 \cdot 10^5$ (0,5%)	$4,6 \cdot 10^5$	$1,2 \cdot 10^5$ (26,1%)	A7	20/20 (100%)

Bei Versuchsreihe I konnte aufgrund der geringen Anzahl an sortierten Zellen kein *Resort* durchgeführt werden. Die sortierten Zellen wurden deshalb stets direkt in LB-Medium aufgefangen, in LB-Selektivmedium angeimpft und üN bei 37 °C neu angezogen, woraufhin die enthaltene Plasmid-DNA präpariert werden konnte (3.3.7.1).

Die FACS-Analyse der so gewonnenen Anreicherungen zeigte, dass keine stark fluoreszierenden Zellen erhalten worden sind (Abbildung 40). Ein zu Beginn zu hoch gewählter Sortier-Bereich, in dem sich unter 0,1 % der Genbank exprimierenden Zellen befinden, führt demnach zu keinem Erfolg.

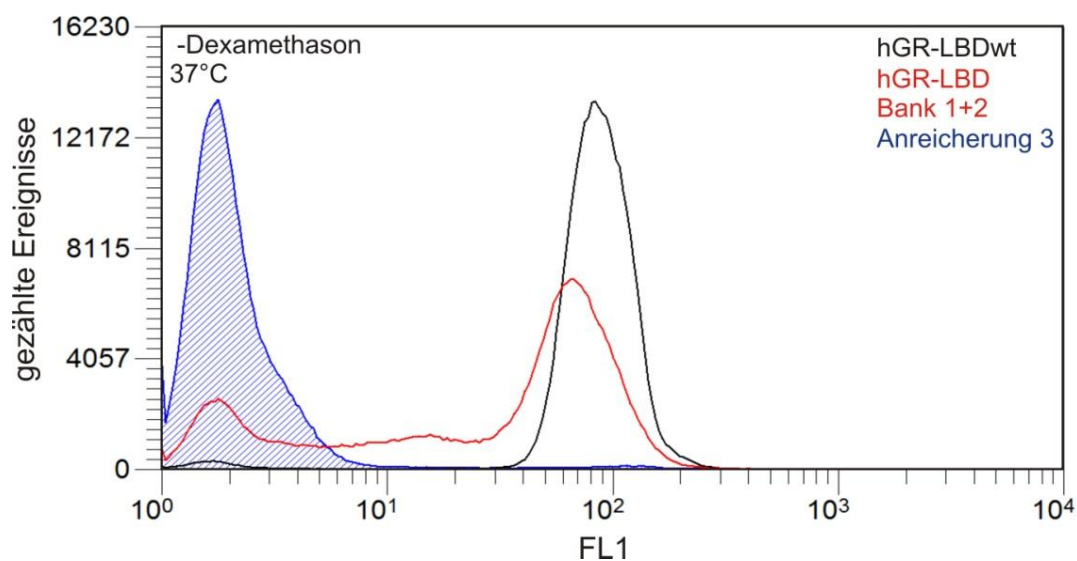


Abbildung 40: Sortierung bei 37 °C ohne Zugabe eines Liganden. Analytisches FACS-Histogramm der Anreicherungsrunde 3 der Versuchsreihe I.

Die Expression erfolgte jeweils in *E. coli* BL21(DE3), für 4 h bei 37 °C. Es wurden je 300.000 Ereignisse analysiert. Die Zellen in Anreicherung 3 zeigen keine Fluoreszenz.

Bei der Versuchsreihe II führten die großzügiger gewählten Sortier-Bereiche (0,5-5,5 % der hellsten Zellen) zu einer besseren Anreicherung stark fluoreszierender Zellen (Abbildung 41). Bei diesem Vorgehen wurden zudem in jeder Runde sehr viel mehr Zellen aussortiert im Vergleich zu Versuchsreihe I, weswegen stets ohne erneute Anzucht direkt nochmals sortiert werden konnte („*Resort*“), was die Qualität der Auslese weiter erhöhte. Die Analyse der retransformierten Anreicherungen zeigte allerdings, dass in allen Anreicherungen neben den stärker fluoreszierenden Zellen auch sehr viele nicht-fluoreszierende Zellen auftreten (Abbildung 41).

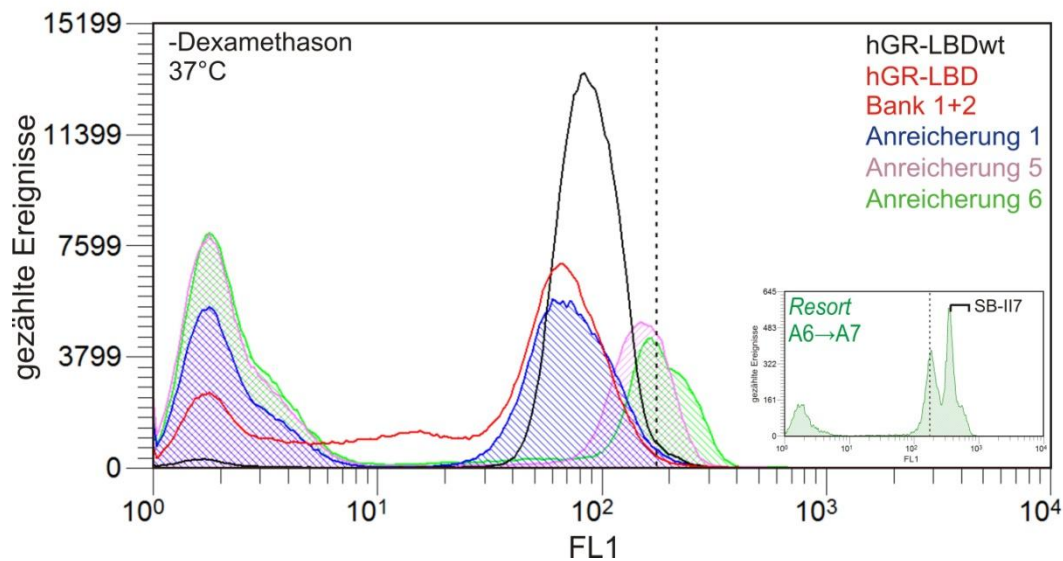


Abbildung 41: Sortierung bei 37 °C ohne Zugabe eines Liganden. Analytisches FACS-Histogramm der Anreicherungsrounds 1, 5 und 6 der Versuchsreihe II.

Die Expression erfolgte jeweils in *E. coli* BL21(DE3), für 4 h bei 37 °C. Es wurden je 300.000 Ereignisse analysiert. Das Inlet zeigt den *Resort* nach der Sortierung von Anreicherung 6 (A6), welcher zur Anreicherung 7 (A7) führte. Das Maximum des Peaks von Anreicherung 6 ist in beiden Histogrammen als gestrichelte Linie markiert.

Bei Versuchsreihe II wurden ab Anreicherungsrounde 3 mit Hilfe der Cyclone Sortiereinheit (3.2.9) im Zuge des *Resorts* einzelne Klone auf LB_{Kan} -Agarplatten sortiert und nach Inkubation bei 37 °C über Nacht mittels Kolonie-PCR (3.3.2) mit den vektorspezifischen und *hgr-lbd*-flankierenden Primern 5'pER13 und 3'pER13 untersucht. Hierbei sollte überprüft werden, ob die sortierten Klone das korrekte pER13b-Plasmid mit dem eingesetzten *hgr-lbd* Volllängenfragment besitzen. Bei früheren Arbeiten wurde beobachtet, dass es beim *Screening* nach stabilisierten Zielproteinvarianten mit Hilfe von C-terminal fusionierten Reporterproteinen zum Verlust des Zielgens im Vektor kommen kann, aufgrund von Rekombinationsereignissen oder Klonierungsartefakten (Seitz, 2006). Dadurch wird nur das wildtypische Reporterprotein hergestellt und ein starkes falsch positives Signal im *Screening* erzeugt. Dies war bei den sortierten Zellen der Versuchsreihe II nicht der Fall, da 90 – 100 % aller untersuchten Zellen bei der Kolonie-PCR das korrekte Amplifikationsprodukt (~1100 bp) zeigten (Tabelle 10). Sequenzierung von jeweils 5 Klonen der Anreicherungen 4 – 7 der Versuchsreihe II zeigte jedoch, dass alle sortierten Klone einen pER13b_ *hgr-lbd* Vektor besitzen, in welchem die *hgr-lbd* Sequenz eine Deletion der Position 199 aufweist (*hgr-lbd*_del, Abbildung 42). Diese Deletion führt zu einer Verschiebung des Leserasters, welche 22 Basen später zu einem neuen Stopp-Codon führt. Es handelt sich also bei den angereicherten *hgr-lbd-egfp* Varianten um falsch positive Mutanten, welche eine inkorrekte Expression des *egfp*-fusionierten *hgr-lbd* Gens zur Folge haben. Wahrscheinlich führt der

Translations-Abbruch in *hgr-lbd* zum Teil zu einem neuen Translationsstart an verschiedenen Stellen stromabwärts des neuen Stoppcodons. Dies würde zur Translation von eGFP führen, welches N-terminal mit unterschiedlich kurzen Teilen von hGR-LBD fusioniert ist. Diese kleinen Fusionspartner stören die eGFP-Faltung wenig und es kommt zu hoher Fluoreszenz aufgrund von löslichem, nahezu wildtypischem eGFP. Dies ist wohl nicht bei allen Zellen, welche mit pER13b_*hgr-lbd_del-egfp* transformiert sind, der Fall, was durch die hohe Anzahl an nicht-fluoreszierenden Zellen bei der Analyse der retransformierten Anreicherungen (Abbildung 41) deutlich wird.



Abbildung 42: Quelle falsch positiver Klone beim Screening bei 37 °C ohne Zugabe eines Liganden. Sequenzausschnitt des *hgr-lbd*-Gens von Base 192 bis 226.

Der Leserahmen ist durch Unterstreichungen markiert. In *hgr-lbd_del* wird durch die Basendeletion an Position 199 eine Verschiebung des Leserasters induziert, die zu einem neuen TGA-Stopp-Codon nach Base 220 führt.

Es konnten somit unter den angelegten Bedingungen (Expression bei 37 °C ohne Dexamethason) keine verbesserten hGR-LBD Varianten angereichert werden. Da in der Literatur bereits beschrieben wurde, dass eine niedrige Expressionstemperatur und die Zugabe des stabilisierenden Liganden Dexamethason die Löslichkeit von hGR-LBD in *E. coli* erhöhen kann (Bledsoe *et al.*, 2002), sollte das *Screening* unter solch moderateren Bedingungen (30 °C in Anwesenheit von Dexamethason) erfolgen, wodurch möglicherweise verbesserte Varianten eher identifiziert werden können.

4.3.2.2 Screening bei 30 °C in Anwesenheit von Dexamethason

Die in 4.3.1 hergestellten *hgr-lbd* Genbanken wurden bei 30 °C in Anwesenheit von 500 µM Dexamethason nach verbesserten Varianten durchsucht. Hierzu wurden *E. coli* BL21(DE3) Zellen wie in 3.2.8.1 beschrieben mit den *hgr-lbd* Genbanken transformiert und angezogen, wobei die Expression unter den genannten Bedingungen mit 0,5 mM IPTG induziert wurde. Die Anzahl der mit Genbank 1 bzw. 2 transformierten und angeimpften Zellen betrug jeweils $2,0 \cdot 10^7$, und überstieg damit die Größe der Genbank 1 ($3,0 \cdot 10^5$) bzw. 2 ($1,1 \cdot 10^6$) um den Faktor 67 bzw. 18.

Die die Genbanken 1 und 2 exprimierenden Zellen wurden gemischt und anschließend mittels Durchflusszytometrie (3.2.9) analysiert und sortiert. Analytische FACS-Histogramme von Zellen mit den hGR-LBD Banken und Zellen mit hGR-LBD Wildtyp (jeweils fusioniert an eGFP) zeigen, dass in den Banken viele Varianten enthalten sind, die zu einer deutlich niedrigeren Fluoreszenz führen. Daneben zeigt der Hauptpeak der Genbanken-exprimierenden Zellen jedoch eine leicht Verschiebung zu höherer Fluoreszenz im Vergleich zum Wildtyp (Abbildung 43).

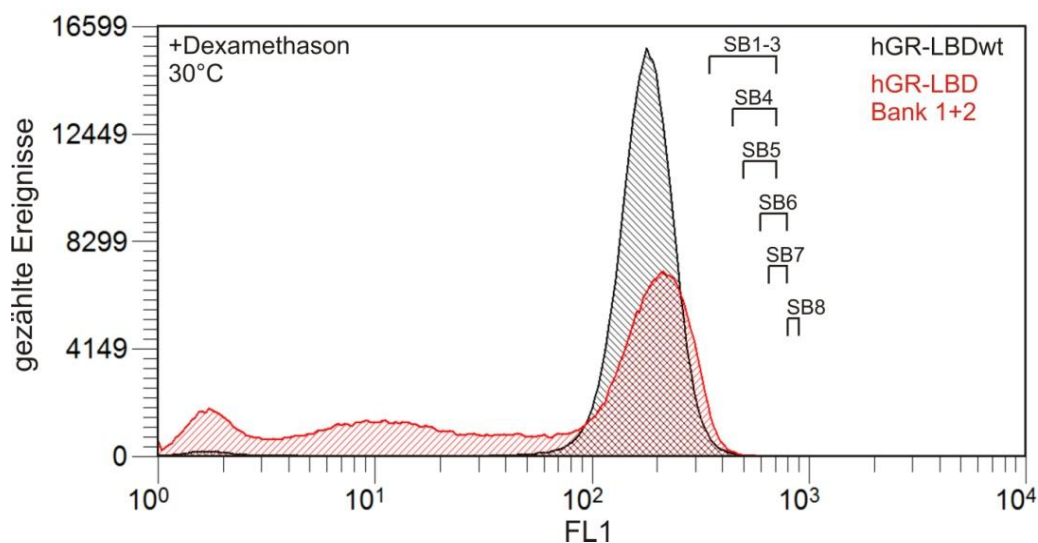


Abbildung 43: Analytische FACS-Histogramme von Zellen mit hGR-LBDwt (schwarz) und hGR-LBD Bank 1+2 (rot), nach Expression bei 30 °C in Anwesenheit von Dexamethason.

Die Expression der *egfp*-fusionierten Gene aus pER13b erfolgte jeweils in *E. coli* BL21(DE3), für 6 h bei 30 °C in Anwesenheit von 500 µM Dexamethason. Es wurden je 300.000 Ereignisse analysiert. Die Sortier-Bereiche der einzelnen Anreicherungsrounden (SB1-8) sind durch Klammern markiert.

Die stärker fluoreszierenden Zellen wurden wieder schrittweise mittels Durchflusszytometrie (3.2.9) angereichert. Insgesamt wurden acht Anreicherungsrounden durchgeführt, wobei der gewählte Sortier-Bereich ähnlich zu Versuchsreihe II in 4.3.2.1 zu Beginn nah am Hauptpeak der Genbanken angelegt wurde (SB1-3, FL1: $3,5 \cdot 10^2 - 7,0 \cdot 10^2$). Nach drei Anreicherungsrounden, bei denen der Anteil an in SB1-3 befindlichen Zellen deutlich anstieg (3,8 % → 17,5 %), wurde der Sortier-Bereich von Anreicherungsrunde zu Anreicherungsrunde schrittweise erhöht, wobei bei jeder Runde die 1-4 % hellsten Zellen sortiert wurden (Abbildung 43, Tabelle 11). Nach jeder Anreicherungsrunde (Sortierung +Resort; siehe Abbildung 35) wurden die sortierten Zellen über Nacht bei 37 °C in 50 ml LB_{Kan}-Medium angezogen (3.2.9). Nach der Präparation der Plasmid-DNA (3.3.7.1) wurden frisch hergestellte elektrokompente *E. coli* BL21(DE3) Zellen (3.2.4) damit transformiert (3.2.6). So konnte die Expression der soeben sortierten Varianten wieder induziert (3.2.8.1)

und die nächste Anreicherungsrunde durchgeführt werden (3.2.9). Die Anzahl der mit den Anreicherungen transformierten Zellen wurde stets über die Ausplattierung von Verdünnungsreihen auf LB_{Kan}-Agarplatten bestimmt (3.2.8.1). In allen Fällen überstieg die Anzahl der Transformanten die Anzahl an zuvor sortierten Zellen mindestens um den Faktor zehn. Damit wurde sichergestellt, dass alle sortierten Varianten in die nächste Anreicherungsrunde überführt wurden. In Tabelle 11 sind die Ergebnisse der Anreicherungsrounden 1-8 zusammengefasst.

Tabelle 11: FACS-Anreicherungen von *hgr-lbd* Genbank 1+2 exprimierenden *E. coli* BL21(DE3) Zellen bei 30 °C in Anwesenheit von Dexamethason.

GB: Genbank, A: Anreicherung, SB: Sortier-Bereich. Es konnte stets ein *Resort* durchgeführt werden. Der in den Anreicherungen A3-A8 vorhandene Anteil an Klonen mit korrektem pER13b_*hgr-lbd* Plasmid wurde über Kolonie-PCR mit den Primern 5'pER13 und 3'pER13 bestimmt (Amplifikationsprodukt: ~1100 bp).

Analysierte Probe	Sortier-Bereich	Sortierung		Resort		resultierende Präparation	positiv (Kolonie-PCR)
		analysierte Ereignisse	sortierte Ereignisse	analysierte Ereignisse	sortierte Ereignisse		
GB1+2	SB1-3	5,0*10 ⁷	1,9*10 ⁶ (3,8%)	1,8*10 ⁶	9,2*10 ⁵ (51,1%)	A1	-
A1	SB1-3	5,0*10 ⁷	3,0*10 ⁶ (6,0%)	3,4*10 ⁶	1,5*10 ⁶ (44,1%)	A2	-
A2	SB1-3	2,0*10 ⁷	3,5*10 ⁶ (17,5%)	3,5*10 ⁶	2,2*10 ⁶ (62,9%)	A3	8/10 (80%)
A3	SB4	1,0*10 ⁸	3,3*10 ⁶ (3,3%)	3,5*10 ⁶	2,2*10 ⁶ (62,9%)	A4	10/10 (100%)
A4	SB5	1,0*10 ⁸	1,9*10 ⁶ (1,9%)	1,9*10 ⁶	8,8*10 ⁵ (46,3%)	A5	18/20 (90%)
A5	SB6	1,2*10 ⁸	2,6*10 ⁶ (2,2%)	3,1*10 ⁶	1,2*10 ⁶ (38,7%)	A6	10/10 (100%)
A6	SB7	3,7*10 ⁷	1,0*10 ⁶ (2,7%)	9,5*10 ⁵	4,9*10 ⁵ (51,6%)	A7	1/80 (1,2%)
A7	SB8	7,0*10 ⁷	1,3*10 ⁶ (1,9%)	1,3*10 ⁶	4,4*10 ⁵ (33,8%)	A8	0/1000 (0%)

Die Anreicherung stark fluoreszierender Zellen bei 30 °C in Anwesenheit von Dexamethason erbrachte sehr viel homogenere Zell-Populationen als dies bei der Anreicherung bei 37 °C ohne Liganden der Fall war (4.3.2.1). So zeigte die Analyse der retransformierten Anreicherungen stets ein homogenes zu höherer Fluoreszenz verschobenes FACS-Histogramm (Abbildung 44). Bei allen Anreicherungen treten so gut wie keine nicht-fluoreszierenden Zellen auf.

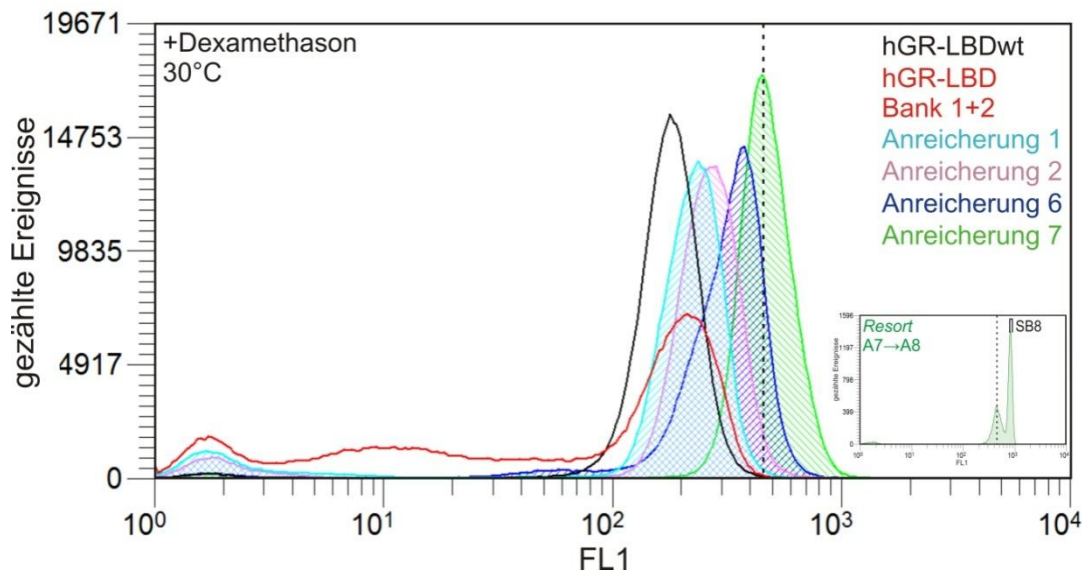


Abbildung 44: Sortierung bei 30 °C in Anwesenheit von Dexamethason. Analytisches FACS-Histogramm der Anreicherungsrounden 1, 2, 6 und 7.

Die Expression erfolgte jeweils in *E. coli* BL21(DE3), für 6 h bei 30 °C in Anwesenheit von 500 µM Dexamethason. Es wurden je 300.000 Ereignisse analysiert. Das Inlet zeigt den *Resort* nach der Sortierung von Anreicherung 7 (A7), welcher zur Anreicherung 8 (A8) führte. Das Maximum des Peaks von Anreicherung 7 ist in beiden Histogrammen als gestrichelte Linie markiert.

Mit Hilfe der Cyclone Sortiereinheit (3.2.9) wurden bei jeder Anreicherungsrunde im Zuge des *Resorts* einzelne Klone auf LB_{Kan}-Agarplatten sortiert und nach Inkubation bei 37 °C über Nacht mittels Kolonie-PCR (3.3.2) mit den vektorspezifischen und *hgr-lbd*-flankierenden Primern 5'pER13 und 3'pER13 untersucht, wodurch das in pER13b-Plasmid inserierte *hgr-lbd* Volllängenfragment detektiert werden sollte. Bis einschließlich Anreicherung 6 zeigten 90 – 100 % aller analysierten Klone das korrekte Amplifikationsprodukt (Tabelle 11). Bei Anreicherung 7 war dagegen nur noch einer von 80 analysierten Klonen positiv (1,2 %). Bei Anreicherung 8 wurden 1000 Klone analysiert, von denen keiner ein positives Resultat zeigte (Tabelle 11, Abbildung 45). Bei Anreicherungen in einem zu hohen Fluoreszenz-Bereich scheinen sich demnach Artefakte mit verkürztem *hgr-lbd* Anteil zu akkumulieren. Die Sequenzierung einiger dieser Kolonie-PCR - negativen Klone bestätigte, dass bei den dort vorhandenen pER13b_ *hgr-lbd* -Plasmiden die *hgr-lbd* Sequenz komplett oder zum Teil (verkürzt auf 3-250 bp) verloren ging, wodurch bei der Expression maximal nur ein kleiner Teil von hGR-LBD N-terminal an eGFP fusioniert vorlag. Diese kleinen Fusionspartner stören die eGFP-Faltung offensichtlich nicht wesentlich, weswegen es zu hoher Fluoreszenz kommt, was zu einem starken falsch positiven Signal im *Screening* führt. Solche verkürzten Artefakte beim *Screening* stabilisierter Proteinvarianten mit Hilfe von C-terminal fusionierten Reportproteinen wurden bereits in früheren Arbeiten (Seitz, 2006) beobachtet und entstehen womöglich durch Rekombinationsereignisse oder Klonierungsartefakte. Neben der in 4.3.2.1 beobachteten Leseraster-Verschiebung durch neu eingeführte Stoppcodone stellen somit

solche verkürzten Vektor-Artefakte eine zweite Quelle für falsch positive Varianten beim *Screening* dar.

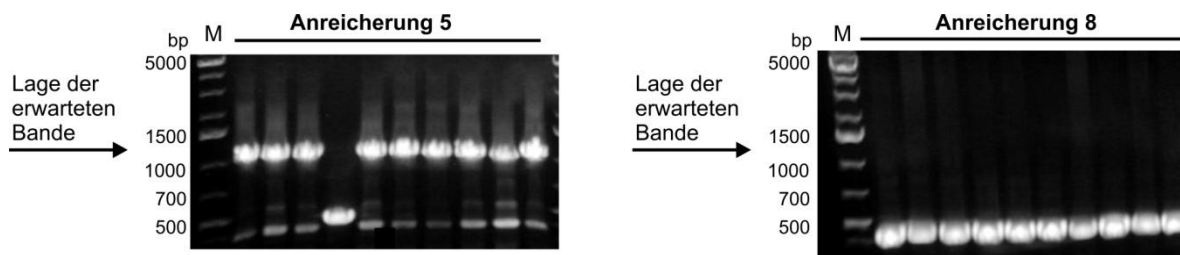


Abbildung 45: Quelle falsch positiver Klone beim *Screening* bei 30 °C in Anwesenheit von Dexamethason. Ergebnis analytischer Kolonie-PCR von je 10 Klonen aus Anreicherung 5 bzw. 8 (1 %iges Agarosegel).

M: DNA-Längenstandard (GeneRuler™ 1 kb Plus (MBI FERMENTAS)). Die erwartete Größe des Amplifikationsproduktes durch die Primer 5'pER13 / 3'pER13 und eines intakten pER13b_ *hgr-lbd* Konstrukts als Templat liegt bei 1107 bp. Bei Anreicherung 5 bzw. 8 zeigen 90 % bzw. 0 % der analysierten Klone dieses Ergebnis. Die negativen Klone zeigen ein verkürztes Amplifikationsprodukt, was auf den teilweisen oder vollständigen Verlust des *hgr-lbd* Inserts aus dem Vektor hindeutet. Sequenzierung einiger negativen Klone zeigte, dass nur noch 3-250 bp von *hgr-lbd* (774 bp) sromaufwärts von *egfp* in pER13b vorhanden waren.

Bei der Sequenzierung der multiplen Klonierungsstelle des pER13b_ *hgr-lbd* Plasmids der 18 Kolonie-PCR - positiven Klone aus Anreicherung 5 und des einen positiven Klons aus Anreicherung 7 wurden vier unterschiedliche, intakte *hgr-lbd-egfp* Varianten identifiziert (A5-1 - A5-4) (Tabelle 12; Nukleotidsequenzen: siehe Anhang 7.3), wobei die Variante A5-1 am häufigsten gefunden wurde (14/19 Klone). Die vier Varianten tragen 8-11 Nukleotid- bzw. 5-9 Aminosäureaustausche (Tabelle 12). Sie müssen demnach aus der *hgr-lbd* Genbank 1 stammen, da dort eine hohe Mutationsfrequenz eingestellt wurde (siehe Tabelle 9). Dies deutet darauf hin, dass Varianten aus Bank 2 mit wenigen Austauschen zu keiner erhöhten Fluoreszenz führten. Dieses Ergebnis reiht sich in die bereits mehrfach in der Literatur beschriebene Beobachtung ein, dass verbesserte Proteinvarianten in Banken mit erhöhter Mutationsfrequenz vermehrt gefunden werden können (Zaccolo & Gherardi, 1999; Daugherty *et al.*, 2000). Dies wird damit erklärt, dass von Genbanken mit einer niedrigen Mutationsfrequenz nur wenige unterschiedliche Polypeptidsequenzen kodiert werden, wodurch die tatsächliche Bank-Größe stark eingeschränkt wird (Drummond *et al.*, 2005). Einerseits weist Genbank 1 auf Genebene weniger unabhängige Varianten auf als Genbank 2 ($3,0 \cdot 10^5$ vs. $1,1 \cdot 10^6$). Andererseits werden allerdings von Genbank 1, aufgrund der höheren Mutationsfrequenz (7,2 vs. 1,8 Austausche/Gen), mehr einzigartige Polypeptidsequenzen kodiert. Somit ist in Bank 1 eine erhöhte Diversität von hGR-LBD vorhanden, wodurch die Identifikation von verbesserten Varianten eher möglich wird.

Tabelle 12: Screening bei 30 °C in Anwesenheit von Dexamethason. Isolierte hGR-LBD Varianten aus Anreicherung 5, 7 und 8.

Insgesamt wurden 18 sortierte Klone aus Anreicherung 5, und ein sortierter Klon aus Anreicherung 7, jeweils mit vollständigem *hgr-lbd* Insert, sequenziert. Die Variante A8 wurde durch OE-PCR mit den Primern 5'*hgr-lbd_OE1* und 3'*hgr-lbd_OE1* aus Anreicherung 8 erhalten. Die auch in A8 gefundenen Mutationen von hGR-LBD A5-1 sind grau hinterlegt. Die in rot markierten 14 der insgesamt 30 verschiedenen Mutationen werden in 4.3.4 näher besprochen.

	hGR-LBD A5-1	hGR-LBD A5-2	hGR-LBD A5-3	hGR-LBD A5-4	hGR-LBD A8
isoliert aus					
Anreicherung 5	13/18 (72,2%)	3/18 (16,7%)	1/18 (5,5%)	1/18 (5,5%)	-
Anreicherung 7	1/1 (100%)	-	-	-	-
OE-PCR aus Anreicherung 8	-	-	-	-	√
Mutationen					
stille Mutationen	5	2	3	1	3
AS-Mutationen	5	9	7	7	7
	I629T	L670P	L544F	S555T	M560K
	T668A	K703M	I559N	T561S	A605V
	L697P	E705G	V575A	L563P	I629T
	L741M	Q710R	M593V	H645R	T668A
	M752T	L721P	N619D	K695E	L697P
		F737Y	V702A	Q738H	L741M
		L741M	S765P	F774S	M752T
		E748K			
		N759D			

Die Retransformation der hGR-LBD Varianten A5-1 - A5-4 in *E. coli* BL21(DE3) mit anschließender Expression bei 30 °C in Anwesenheit von Dexamethason (3.2.8.1) und Analyse der Zell-Fluoreszenz durch Durchflusszytometrie bestätigte, dass die vier isolierten *hgr-lbd-egfp* Varianten A5-1 - A5-4 zu einer höheren Fluoreszenz führen als der Wildtyp (Abbildung 46).

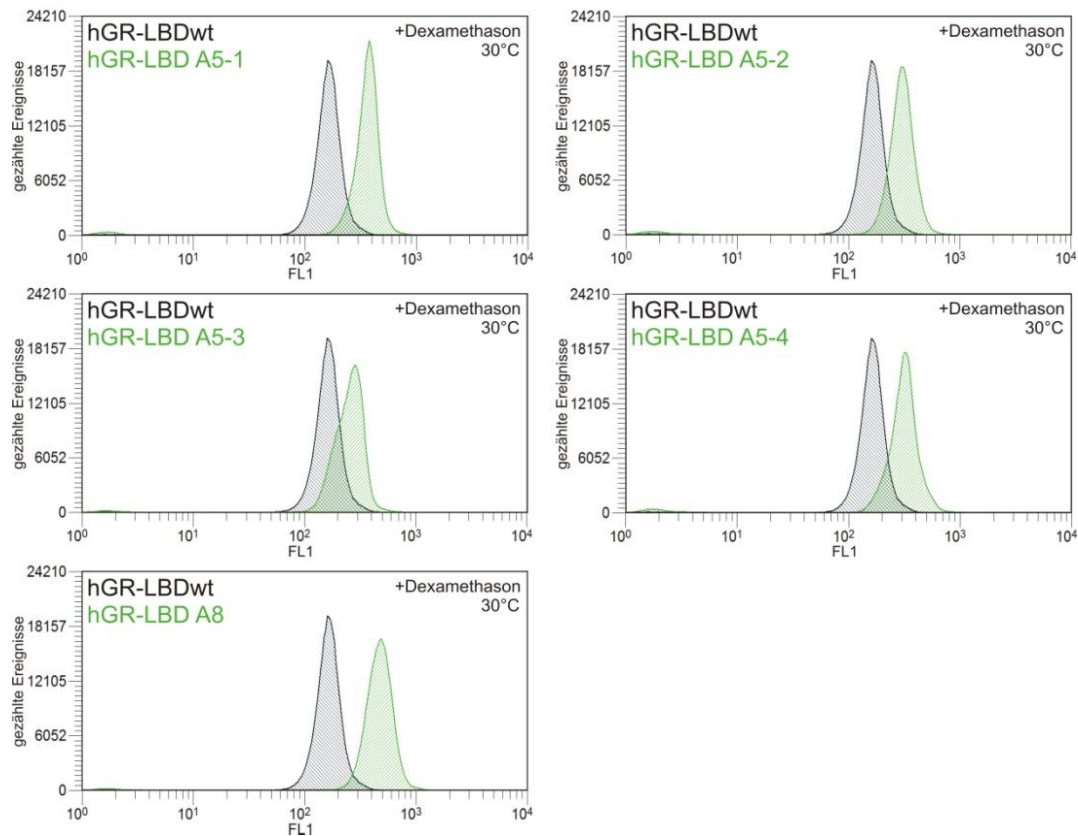


Abbildung 46: Analytische FACS-Histogramme von Zellen mit hGR-LBDwt (schwarz) und hGR-LBD A5-1 - A5-4 bzw. A8 (grün), bei 30 °C in Anwesenheit von Dexamethason.

Die Expression der *egfp*-fusionierten Gene erfolgte jeweils in *E. coli* BL21(DE3), für 6 h bei 30 °C in Anwesenheit von 500 µM Dexamethason. Es wurden jeweils 300.000 Ereignisse analysiert.

Die potentiell vorhandenen positiven *hgr-lbd* Varianten in der von Vektor-Artefakten dominierten Anreicherung 8 (A8) (siehe Abbildung 45) wurden angereichert, indem das *hgr-lbd* Volllängenfragment aus der A8-Plasmidpräparation durch *Overlap Extension* - PCR (OE-PCR, 3.3.3.2) re-amplifiziert wurde. Hierbei wurden zwei verschiedene Ansätze durchgeführt, bei denen unterschiedliche zum Zielgen komplementäre Primerpaare (5'*hgr-lbd*_OE1 / 3'*hgr-lbd*_OE1 und 5'*hgr-lbd*_OE2 / 3'*hgr-lbd*_OE2) benutzt wurden. Dies sollte sicherstellen, dass auch Mutationen gefunden werden können, die im Bereich der Primer liegen. Nach der Re-Klonierung der PCR-Produkte über die Restriktionsschnittstellen *NdeI* und *NotI* in pER13a wurden zehn der so erhaltenen Varianten in *E. coli* BL21(DE3) bei 30 °C in Anwesenheit von Dexamethason exprimiert (3.2.8.1) und die Fluoreszenz dieser Expressionskulturen durch Durchflusszytometrie (3.2.9) analysiert (nicht gezeigt). Dabei produzierte eine hGR-LBD Mutante (hGR-LBD A8) eine deutlich erhöhte Fluoreszenz im Vergleich zum Wildtyp und den hGR-LBD A5-1 - A5-4 Varianten (Abbildung 46). Die Sequenzierung von *hgr-lbd* A8 ergab, dass diese Variante neben Austauschen aus

hgr-lbd A5-1 zusätzlich Mutationen aufweist, welche zu den Aminosäureaustauschen M560K und A605V führen (Tabelle 12). Durch den PCR-Amplifikationsschritt kam es wohl zu einer Vermischung der in den ursprünglichen Varianten vorhandenen Mutationen, wodurch das wiederholte Auftreten der *hgr-lbd* A5-1 Mutationen erklärt werden kann. Die deutlich höhere Fluoreszenz durch hGR-LBD-eGFP A8 im Vergleich zu hGR-LBD-eGFP A5-1 zeigt, dass die beiden neu identifizierten Mutationen M560K+A605V zu einer sehr starken Verbesserung führen (Tabelle 12, Abbildung 46).

Durch das *Screening* unter moderaten Bedingungen (30 °C in Anwesenheit von Dexamethason) konnten somit fünf verschiedene hGR-LBD Varianten mit insgesamt 30 unterschiedlichen Mutationen identifiziert werden, welche als eGFP-Fusionsprotein *in vivo* eine höhere Fluoreszenz erzeugen als hGR-LBDwt. Im Folgenden wurde untersucht, inwiefern diese durch die Mutationen induzierte verbesserte Fluoreszenz eine erhöhte Löslichkeit und Stabilität von hGR-LBD widerspiegelt.

4.3.3 Charakterisierung der angereicherten hGR-LBD Varianten

4.3.3.1 Analyse der Löslichkeit der angereicherten hGR-LBD Varianten *in vivo* mit Hilfe der Chloramphenicol-Acetyltransferase als Reporter

Um zu überprüfen, ob die verbesserte Fluoreszenz der sortierten Varianten auf eine veränderte Wechselwirkung mit eGFP zurückgeht und somit möglicherweise ein falsch positives Ergebnis darstellt, sollte die Löslichkeit *in vivo* mit Hilfe eines anderen Reporterproteins nochmals überprüft werden. Hierbei wurde die Chloramphenicol-Acetyltransferase (CAT) benutzt, die bereits in anderen Arbeiten (Maxwell *et al.*; 1999; Sieber *et al.*, 2001; Seitz, 2006) erfolgreich als Faltungsreporter eingesetzt wurde. Zunächst wurde ein neuer *cat*-Fusionsvektor pER13cat (2.8.1.2.3) konstruiert, welcher analog zum *egfp*-Fusionsvektor pER13a aufgebaut ist. Hierzu wurde die kodierende Sequenz für CAT (inklusive des Linkers GGRLEAAR) mit den Primern 5' *cat*(*NotI*) und 3' *cat*(*XhoI*) aus dem Templat pCFN1 (2.8.3, Maxwell *et al.*, 1999) mittels PCR (3.3.1) amplifiziert. Das erhaltene Produkt konnte anschließend über die Restriktionsschnittstellen *NotI* und *XhoI* in pER13a eingefügt werden, wodurch *egfp* durch *cat* ersetzt wurde. Anschließend wurden *tmhisF* (Positivkontrolle), *hgr-lbdwt*, *hgr-lbd*(F602S) und die Varianten *hgr-lbd* A5-1 - A5-4 über die Restriktionsschnittstellen *NdeI* und *NotI* jeweils aus pER13a bzw. pER13b in pER13cat umklontiert.

Frisch hergestellte elektrokompente (3.2.4) *E. coli* T7 Express Zellen (2.7) wurden mit den verschiedenen pER13cat-Konstrukten transformiert (3.2.6). Nach Induktion der Expression mittels 0,5 mM IPTG in Anwesenheit von Dexamethason wurden sie auf LB-Agarplatten mit unterschiedlichen Konzentrationen an Chloramphenicol (0 – 800 µg/ml) ausplattiert und das Wachstum bei 30 °C für 85 h verfolgt (3.2.8.2). Die LB-Agarplatten enthielten kein Dexamethason, da dieses in Ethanol gelöst ist, was in den Platten zu einer Hemmung des Kolonie-Wachstums führte (Daten nicht gezeigt). In Abbildung 47 ist das Wachstumsverhalten bei einigen Chloramphenicol-Konzentrationen exemplarisch dargestellt und in Tabelle 13 sind alle Ergebnisse zusammengefasst.

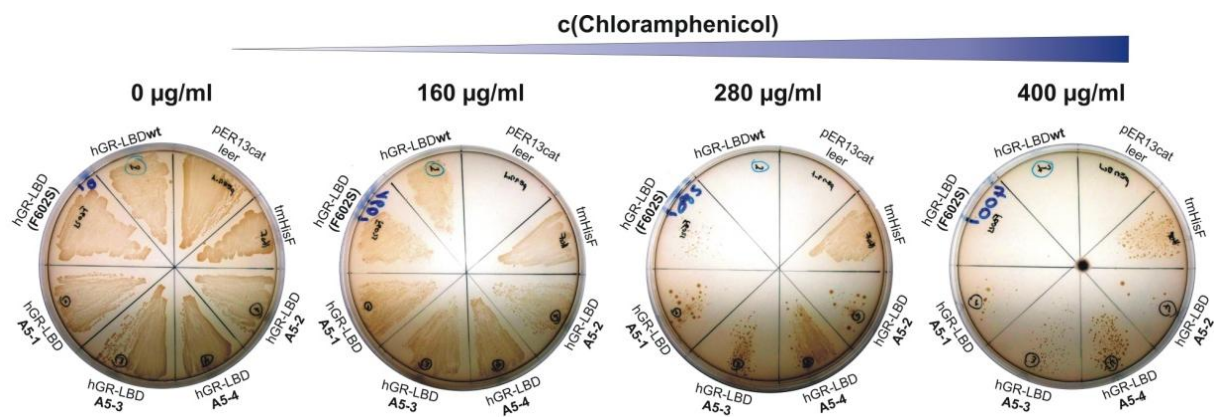


Abbildung 47: Wachstum von transformierten *E. coli* T7 Express Zellen auf LB-Platten in Anwesenheit unterschiedlicher Konzentrationen an Chloramphenicol.

Mit leerem pER13cat-Vektor transformierte Zellen können nur bei 0 µg/ml Chloramphenicol wachsen. Zellen mit sehr löslichem tmHisF-CAT zeigen starkes Wachstum bei allen Chloramphenicol-Konzentrationen. Die durch hGR-LBD-CAT Varianten vermittelte Chloramphenicol-Resistenz unterscheidet sich folgendermaßen: WT < F602S < A5-2 < A5-1 < A5-3 < A5-4.

Tabelle 13: Wachstum von transformierten *E. coli* T7 Express Zellen auf LB-Platten in Anwesenheit unterschiedlicher Konzentrationen an Chloramphenicol.

LB-Platten ohne Chloramphenicol enthielten 75 µg/ml Kanamycin. ++++ bis +: relatives Maß für die Anzahl und Größe der Kolonien; K.: gezählte Einzelkolonien; -: kein deutliches Wachstum nach 85 h bei 30 °C.

c(Chloramphenicol) [µg/ml]	pER13cat								
	leer	tmhis F	hgr-lbd						
			wt	F602S	A5-1	A5-2	A5-3	A5-4	
0	++++	++++	++++	++++	++++	++++	++++	++++	++++
40	6 K.	++++	++++	++++	++++	++++	++++	++++	++++
80	4 K.	++++	++++	++++	++++	++++	++++	++++	++++
120	2 K.	++++	++++	++++	++++	++++	++++	++++	++++
160	-	++++	++++	++++	++++	++++	++++	++++	++++
200	-	++++	++++	++++	++++	++++	++++	++++	++++
240	-	++++	+++	+++	+++	+++	+++	+++	++++
280	-	++++	-	++	+++	+++	+++	+++	++++
320	-	++++	-	>100 K.	++	+	++	++	+++
360	-	+++	-	3 K.	40K.	8K.	>100 K.	+++	+++
400	-	+++	-	-	40K.	10K.	>100 K.	+++	+++
600	-	>100 K.	-	-	2 K.	-	-	>100 K.	+++
800	-	>100 K.	-	-	-	-	-	-	-

Die Negativ- bzw. Positivkontrolle pER13cat_leer bzw. pER13cat_tmhisF, bei der kein Protein bzw. das sehr lösliche und thermostabile tmHisF als CAT-Fusion in den Zellen hergestellt wird, führt zu nahezu keinem Wachstum bzw. zu Wachstum bis 800 µg/ml Chloramphenicol (Abbildung 47, Tabelle 13). Der neu hergestellte pER13cat-Vektor eignet sich demnach als *cat*-Fusionsvektor für die Analyse der Proteinlöslichkeit *in vivo*.

Die Proteine hGR-LBDwt-CAT / hGR-LBD(F602S)-CAT / hGR-LBD A5-1-CAT / hGR-LBD A5-2-CAT / hGR-LBD A5-3-CAT / hGR-LBD A5-4-CAT befähigen die Zellen unter den genannten Bedingungen auf Medium mit bis zu 240 µg/ml / 320 µg/ml / 400 µg/ml / 400 µg/ml / 400 µg/ml / 600 µg/ml Chloramphenicol zu wachsen (Abbildung 47, Tabelle 13). Somit zeigen *E. coli* T7 Express Zellen, welche die aus dem eGFP-Reportersystem erhaltenen hGR-LBD Varianten A5-1 - A5-4 oder die bekannte Variante hGR-LBD(F602S) (Bledsoe *et al.*, 2002) als CAT-Fusionsproteine herstellen eine höhere Chloramphenicol-Resistenz als Zellen mit hGR-LBDwt-CAT.

Die verbesserte Löslichkeit *in vivo* von hGR-LBD(F602S) bzw. von hGR-LBD A5-1 - A5-4 kann also mit Hilfe der Chloramphenicol-Acetyltransferase als C-terminal fusioniertes Reporterprotein ebenfalls dargestellt werden. Dadurch kann ausgeschlossen werden, dass die verbesserte Fluoreszenz der in 4.3.2.2 isolierten eGFP-fusionierten hGR-LBD Varianten auf eGFP-abhängige Effekte zurückgeht.

4.3.3.2 Expression der angereicherten *hgr-lbd* Varianten in Abwesenheit von Reporterprotein

Um hGR-LBDwt, hGR-LBD(F602S) und die in 4.3.2.2 isolierten hGR-LBD Varianten A5-1 - A5-4 und A8 ohne eGFP oder CAT- Fusion herstellen zu können, wurden die kodierenden Sequenzen mit Hilfe der Restriktionsschnittstellen *NdeI* und *NotI* aus pER13b bzw. pER13a in pER14 (2.8.1.2.2) umklont. Bei pER14-Konstrukten wird nach Genexpression den Proteinen C-terminal ein His₆-Tag angefügt. Probeexpressionen der hergestellten pER14_ *hgr-lbd* Konstrukte in *E. coli* BL21(DE3) bei 20 °C in Anwesenheit von Dexamethason (3.4.1.1) zeigten, dass keine der hGR-LBD Varianten in der löslichen Zellfraktion der Zellen zu finden war (Daten nicht gezeigt). Da in der Literatur bereits beschrieben wurde, dass auch die bekannte hGR-LBD(F602S) Variante nur als Glutathion-S-Transferase (GST) –Fusionsprotein löslich hergestellt werden kann (Bledsoe *et al.*, 2002), wurden alle Varianten in den *his₆-gst*-Fusionsvektor pERN7 (2.8.1.2.4, Abbildung 11) umklont. Hierzu wurden die kodierenden Sequenzen mit den Primern 5'pER13 und 3'*hgr-lbd* (*NotI*/Stop) mittels PCR (3.3.1) amplifiziert, wobei die jeweiligen pER13a_ *hgr-lbd* bzw.

pER13b_*hgr-lbd*-Konstrukte als Templat dienten. Die erhaltenen Produkte konnten anschließend über die Restriktionsschnittstellen *NdeI* und *NotI* in pERN7 eingefügt werden. Daneben wurden neue *hgr-lbd* Konstrukte in pERN7 hergestellt, bei denen der bekannte Austausch für F602S zusätzlich in die Sequenzen der sortierten Varianten eingebracht wurde. Dies erfolgte über Megaprimer-PCR (3.3.3.1). Der Megaprimer wurde mit einem mutagenen Primer (3'*hgr-lbd*(F602S) bzw. 3'*hgr-lbd*(F602S+A605V)), und pERN7up mittels PCR (3.3.1) amplifiziert, wobei jeweils pERN7_*hgr-lbd* A5-1 - A5-4 bzw. pERN7_*hgr-lbd* A8 als Templat diente. Das Vollängen-Gen wurde anschließend jeweils durch den aus einem Agarosegel aufgereinigten (3.3.5) Megaprimer und 3'*hgr-lbd* (*NotI*/Stop) mit Hilfe desselben Templats wie zuvor durch PCR amplifiziert. Mittels der Restriktionsenzyme *NdeI* und *NotI* konnten die erhaltenen mutagenisierten Gene wieder in pERN7 kloniert werden.

Die so hergestellten pERN7_*hgr-lbd* Konstrukte, bei denen die in Tabelle 12 (siehe 4.3.2.2) aufgelisteten *hgr-lbd* Sequenzen mit und ohne die für F602S kodierende Mutation vorliegen, wurden in *E. coli* BL21(DE3) bei 20 °C in Anwesenheit von Dexamethason exprimiert (3.4.1.1). Die lösliche Zellfraktion wurde anschließend über Nickel-Sepharose gereinigt, um eindeutige Aussagen über die Menge an löslichem Ziel-Protein treffen zu können. Diese Analysen zeigten eindeutig, dass das wildtypische Protein und alle fünf neuen hGR-LBD Varianten A5-1 - A5-4 und A8 im Gegensatz zur bekannten Variante hGR-LBD(F602S) unter diesen Bedingungen vollständig unlöslich in *E. coli* hergestellt werden. Auch die Einführung von F602S in die neuen Varianten führte zu keiner Verbesserung (Abbildung 48). Vielmehr wurde der vorteilhafte Effekt dieser Mutation durch die anderen Austausche aufgehoben.

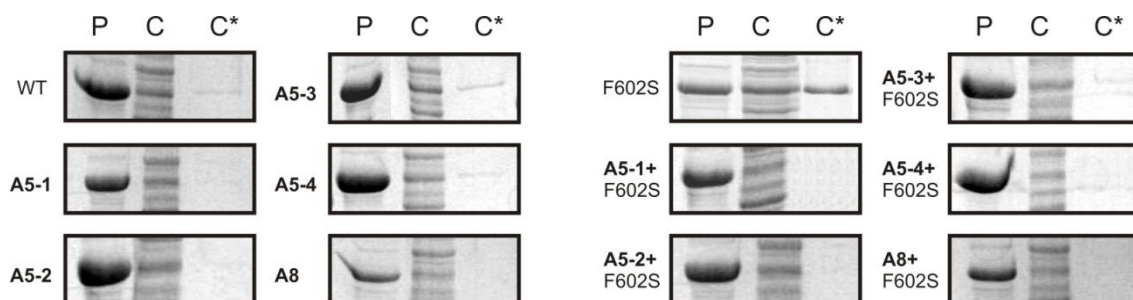


Abbildung 48: Analytische Herstellung His₆-GST-fusionierter hGR-LBD Varianten in *E. coli* BL21(DE3), dokumentiert durch SDS-PAGE (12,5 %).

P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*), C*: über Nickel-Sepharose gereinigte lösliche Zellfraktion. Die Expression erfolgte jeweils in *E. coli* BL21(DE3), üN bei 20 °C nach Zugabe von 500 µM Dexamethason. Es ist nur der 45-66 kDa Bereich des SDS-PA Gels gezeigt, in welchem sich His₆-GST-hGR-LBD (Molekulargewicht: ~58 kDa) befindet. Bei der Darstellung von Variante A5-3 wurden zwei SDS-Gele zusammenggefügt. hGR-LBD(F602S) zeigt im Gegensatz zum Wildtyp und allen anderen Varianten eine deutliche Bande in C*.

Die durch eGFP- (4.3.2.2) bzw. CAT-Fusion (4.3.3.1) dokumentierte verbesserte Löslichkeit der sortierten hGR-LBD Varianten kann somit nicht mehr beobachtet werden, sobald die C-terminal fusionierten Reporterproteine entfernt werden: weder mit C-terminalem His₆-Tag noch mit N-terminalem His₆-GST-Tag können die Varianten löslich in *E. coli* hergestellt werden. Die hohe Anzahl an Aminosäureaustauschen (5-9) in den fünf sortierten hGR-LBD Varianten könnte eine Ursache für dieses Problem darstellen. Die vorliegenden Ergebnisse lassen vermuten, dass zumindest einige der Austausche in hGR-LBD zu einer Erhöhung der eGFP-Fluoreszenz bzw. der CAT-Aktivität führen, ohne die Löslichkeit der jeweiligen Fusionsproteine zu verbessern. Alternativ könnte ein Teil der Mutationen zu einer Verbesserung der Löslichkeit der eGFP- bzw. CAT-Fusionsproteine führen, welche jedoch an den Kontext der beiden Reporter gebunden ist und sich somit im His₆-GST-hGR-LBD Fusionsprotein nicht ausprägt. Um solche Mutationen auszusortieren und um Mutationen aus hGR-LBD A5-1 - A5-4 und A8 identifizieren zu können, welche ebenfalls in hGR-LBD ohne fusioniertes Reporterprotein zu erhöhter Löslichkeit führen, wurden im Folgenden 14 der 30 Mutationen einzeln analysiert.

4.3.4 Einzel-Analyse von Mutationen aus hGR-LBD A5-1 - A5-4 und hGR-LBD A8

Vierzehn der insgesamt 30 Austausche wurden anhand der von Bledsoe *et al.* (2002) gelösten Röntgenkristallstruktur von hGR-LBD(F602S) (pdb 1m2z) für eine nähere Analyse ausgewählt (Abbildung 49).

Die Nukleotid-Austausche, welche für die Aminosäureaustausche L544F, I559N, M560K, A605V, N619D, H645R, K695E, V702A, E705G, Q710R, Q738H, M752T und N759D kodieren, wurden durch Megaprimer-PCR (3.3.3.1) einzeln in *hgr-lbd* eingeführt. Die Megaprimer wurden mit einem mutagenen Primer (3′*hgr-lbd* (L544F), 3′*hgr-lbd* (I559N), 3′*hgr-lbd* (M560K), 3′*hgr-lbd* (A605V), 3′*hgr-lbd* (N619D), 3′*hgr-lbd* (H645R), bzw. 5′*hgr-lbd* (K695E), 5′*hgr-lbd* (V702A), 5′*hgr-lbd* (E705G), 5′*hgr-lbd* (Q710R), 5′*hgr-lbd* (Q738H), 5′*hgr-lbd* (M752T), 5′*hgr-lbd* (N759D)) und pERN7up bzw. 3′*hgr-lbd* (*NotI*/Stop) mittels PCR (3.3.1) amplifiziert, wobei jeweils pERN7_ *hgr-lbd* als Templat diente. Das Vollängen-Gen wurde anschließend jeweils durch den aus einem Agarosegel aufgereinigten (3.3.5) Megaprimer und 3′*hgr-lbd* (*NotI*/Stop) bzw. pERN7up mit Hilfe desselben Templats wie zuvor durch PCR hergestellt. Zur Klonierung der Einzelvariante *hgr-lbd*(F774S) wurde das Vollängen direkt mit Hilfe der Primer pERN7up und 3′*hgr-lbd* (F774S/*NotI*) durch PCR amplifiziert. Mittels der Restriktionsenzyme *NdeI* und *NotI* konnten die erhaltenen mutagenisierten Gene wieder in pERN7 kloniert werden.

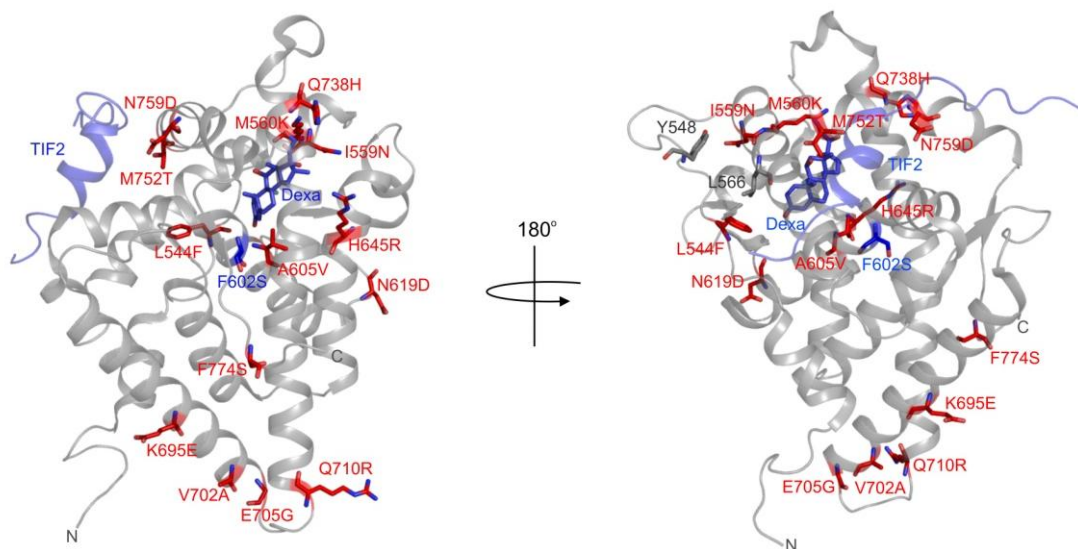


Abbildung 49: Lokalisation der 14 ausgewählten Einzel-Mutationen aus den sortierten hGR-LBD Varianten A5-1 - A5-4 und A8 (rot) in der Kristallstruktur von hGR-LBD(F602S) (pdb 1m2z) (grau).

Die bekannte Mutation F602S, gebundenes TIF2 Coaktivator-Peptid und Dexamethason sind blau dargestellt. Die gezeigten 14 Mutationen (mit Hilfe von PyMol in pdb 1m2z eingefügt) von hGR-LBD A5-1 - A5-4 und A8 wurden aus folgenden Gründen zur näheren Analyse ausgewählt: M560 und A605 sind an der Ligand-Bindung beteiligt; M752 und N759 sind Teil von α -Helix 12, deren Interaktionspartner (TIF2) bei heterologer Expression in *E. coli* fehlt; N619, H645, K695, Q738 und F774 sind zum Teil Lösungsmittel-zugänglich und stellen somit potentielle Ziele zur Oberflächenoptimierung dar; V702, E705 und Q710 befinden sich einem relativ unstrukturierten Bereich des Proteins (Loop zwischen α -Helix 9 und α -Helix 10 zeigt hohe B-Faktoren in der Struktur); L544 bzw. I559 liegen in Loop 540–554 bzw. α -Helix 3 – durch Mutation zu Phenylalanin (L544F) bzw. Asparagin (I559N) könnte möglicherweise die benachbarte α -Helix 3 bzw. Loop 540-554 durch van-der-Waals Wechselwirkungen bzw. H-Brücken zu L566 bzw. Y548 (nur rechts dargestellt) stabilisiert werden.

Die so hergestellten pERN7_*hgr-lbd* Konstrukte wurden für Probeexpressionen der *hgr-lbd* Einzelvarianten in *E. coli* BL21(DE3) bei 20 °C in Anwesenheit von Dexamethason benutzt (3.4.1.1). Die Gene *hgr-lbdwt* bzw. *hgr-lbd*(F602S) wurden als Negativ- bzw. Positivkontrolle für lösliche Protein-Expression in *E. coli* parallel untersucht. Die Analyse der über Nickel-Sepharose gereinigten löslichen Zellfraktion der Expressionskulturen zeigte, dass A605V (aus hGR-LBD A8), V702A (aus hGR-LBD A5-3), E705G (aus hGR-LBD A5-2) und M752T (aus hGR-LBD A5-1 bzw. A8) ähnlich wie F602S zu einer verbesserten Löslichkeit von His₆-GST fusioniertem hGR-LBD in *E. coli* führen (Abbildung 50).

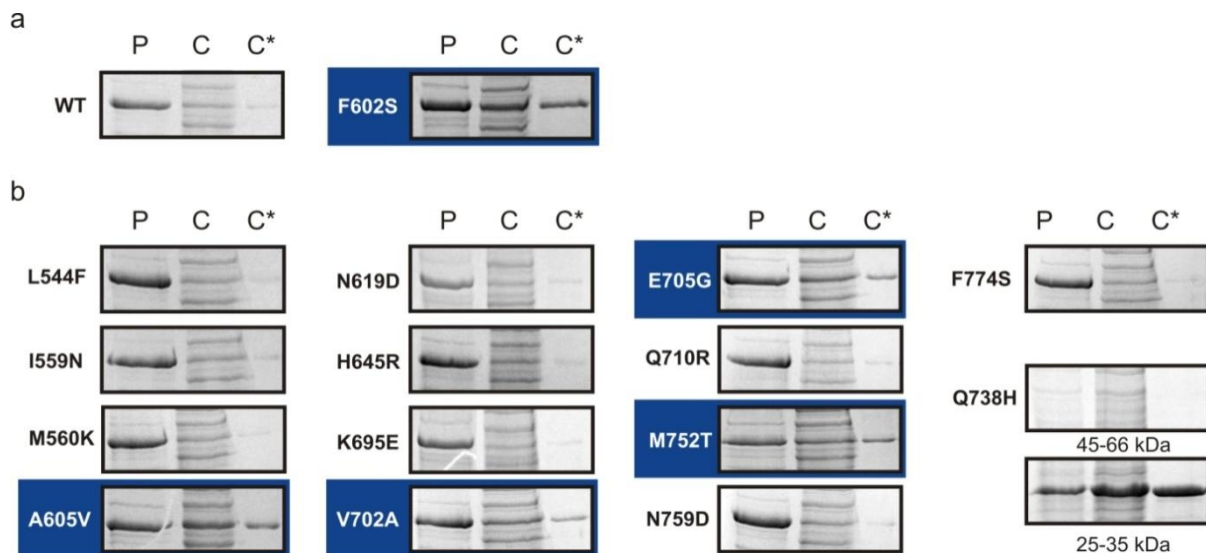


Abbildung 50: Analytische Herstellung His₆-GST-fusionierter hGR-LBD Einzelvarianten in *E. coli* BL21(DE3), dokumentiert durch SDS-PAGE (12,5 %).

P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*), C*: über Nickel-Sepharose gereinigte lösliche Zellfraktion. Die Expression erfolgte jeweils in *E. coli* BL21(DE3), üN bei 20 °C in Anwesenheit von 500 µM Dexamethason. Es wurde stets die gleiche Anzahl an Zellen ($2 \cdot 10^9$) analysiert. Es ist in der Regel nur der 45-66 kDa Bereich des SDS-PA Gels gezeigt, in welchem sich His₆-GST-hGR-LBD (Molekulargewicht: ~58 kDa) befindet. Im Fall von hGR-LBD(Q738H) ist zusätzlich der 25-35 kDa Bereich dargestellt. Löslich hergestellte hGR-LBD Einzelvarianten sind blau hinterlegt.

a hGR-LBDwt und hGR-LBD(F602S) sind als Negativ- bzw. Positivkontrolle für Löslichkeit *in vivo* gezeigt.

b 14 der 30 Mutationen von hGR-LBD A5-1 - A5-4 und A8 wurden als hGR-LBD Einzelvarianten analysiert. Bei der Darstellung von Variante hGR-LBD(K695E) wurden zwei SDS-PA Gele zusammengefügt. Bei der hGR-LBD(Q738H)-Präparation ist kein Volllängenprotein, aber ein lösliches Abbauprodukt (Molekulargewicht: ~30 kDa) vorhanden. Dies entspricht vermutlich zum großen Teil dem His₆-GST Protein. hGR-LBD(A605V), hGR-LBD(V702A), hGR-LBD(E705G) und hGR-LBD(M752T) zeigen ähnlich wie die bekannte Variante hGR-LBD(F602S) eine deutliche Bande in C*.

Die anderen Austausche führen zu keiner Erhöhung der Löslichkeit des Proteins. hGR-LBD(Q738H) wird zudem proteolytisch abgebaut. Das entstandene Abbauprodukt zeigt ein Molekulargewicht von ca. 30 kDa und ist zum Teil löslich. Das His₆-GST-hGR-LBD Fusionsprotein scheint hier demnach in 2 Teile (hGR-LBD: 30 kDa; His₆-GST: 28 kDa) gespalten zu werden, wobei sich His₆-GST in der löslichen Zellfraktion befindet. Die durch Q738H induzierte Proteolyse-Anfälligkeit könnte eine weitere Ursache für ein falsch positives Signal im eGFP-Reportersystem darstellen. In diesem Fall wäre die Verkürzung des Fusionsproteins *in vivo* zu nahezu wildtypischem Reporterprotein nicht auf Genebene durch eingeführte Stoppcodone (siehe 4.3.2.1, Abbildung 42) oder Genverlust (siehe 4.3.2.2, Abbildung 45) bedingt, sondern erfolgte auf Proteinebene. Somit ist das Q738H-enthaltende Protein hGR-LBD-eGFP A5-4 vermutlich eine falsch positiv identifizierte Variante, die intrinsisch zur *in vivo* Abspaltung von eGFPwt und somit zu hoher Fluoreszenz führt. Dies würde erklären, weshalb keine der untersuchten Einzelaustausche aus hGR-LBD A5-4 zu einer erhöhten Löslichkeit führte.

Zusammenfassend kann festgestellt werden, dass mit Hilfe der in 4.3.2.2 sortierten hGR-LBD-eGFP Varianten vier Aminosäureaustausche (A605V, V702A, E705G und M752T) identifiziert werden konnten, welche die lösliche Herstellung von hGR-LBD in *E. coli* und damit die Reinigung des Proteins erlauben. Inwiefern diese Mutationen additiv wirken und welchen konkreten Einfluss sie auf die eGFP-vermittelte Fluoreszenz, die Löslichkeit, die konformationelle Stabilität und die Struktur von hGR-LBD haben, wird im folgenden Kapitel 4.3.5 näher erläutert.

Von den beiden Mutationen M560K und A605V, welche im eGFP-Reportersystem für die starke Fluoreszenz-Erhöhung von Variante A5-1 nach A8 verantwortlich sind (Tabelle 12, Abbildung 46), führt in isolierter Form nur A605V zu einer erhöhten Löslichkeit. Um den durch M560K vermittelten Effekt in hGR-LBD-eGFP zu verstehen, wurde sein Einfluss auf die oben genannten Eigenschaften ebenfalls genauer analysiert.

4.3.5 Einfluss von M560K, F602S, A605V, V702A, E705G und M752T auf hGR-LBD und hGR-LBD-eGFP

Die Effekte der Mutationen F602S (Bledsoe *et al.*, 2002) bzw. M560K, A605V, V702A, E705G und M752T (in dieser Arbeit identifiziert: 4.3.2.2 / 4.3.4) auf hGR-LBD und hGR-LBD-eGFP wurden nun im Detail analysiert. Hierzu wurden ausgehend von den Einzelvarianten die in Tabelle 14 gezeigten Kombinationsvarianten hergestellt und charakterisiert.

Hierbei sollte untersucht werden, ob die für die Löslichkeit von His₆-GST-hGR-LBD vorteilhaften Mutationen (Abbildung 50) in hGR-LBD-eGFP zu einer reproduzierbaren Verbesserung der Fluoreszenz *in vivo* führen. Da die in 4.3.2.2 sortierten hGR-LBD Varianten A5-1 - A5-4 bzw. A8 als His₆-GST-Fusionsproteine keine verbesserte Löslichkeit zeigten (4.3.3.2, Abbildung 48), sollte die Untersuchung der hGR-LBD-eGFP Varianten mit den oben genannten Mutationen nochmals die Qualität von eGFP als Faltungsreporter *in vivo*, abseits der beim *Screening* auftretenden falsch positiven Artefakte (Abbildung 42, 45), beleuchten. Durch den Vergleich der Eigenschaften der identischen hGR-LBD Einzel- und Kombinationsvarianten bezüglich Fluoreszenz (als eGFP-Fusionsprotein), Löslichkeit (als His₆-GST-Fusionsprotein) und Stabilität (ohne Fusionsprotein) sollte die Aussagekraft des Reportersystems genau charakterisiert werden.

Darüber hinaus sollte durch diese Analyse überprüft werden, ob die positiven Mutationen additiv wirken und ob durch deren Kombination eine hGR-LBD Variante hergestellt werden

kann, die eine sehr stark erhöhte Löslichkeit bzw. Stabilität besitzt und somit für biophysikalische Untersuchungen wie z.B. die Röntgenkristallstrukturanalyse optimiert ist.

Tabelle 14: hGR-LBD Einzel- und Kombinationsvarianten zur Analyse von M560K, F602S, A605V, V702A, E705G und M752T.

F602S (Bledsoe *et al.*, 2002) bzw. A605V, V702A, E705G und M752T verbessern die lösliche Herstellung von His₆-GST-hGR-LBD in *E. coli* (4.3.4). M560K führt zusammen mit A605V zu erhöhter Fluoreszenz von hGR-LBD-eGFP *in vivo* (4.3.2.2).

hGR-LBD Einzelvarianten	hGR-LBD Kombinationsvarianten
M560K	A605V+M752T
F602S	A605V+M752T+V702A
A605V	A605V+M752T+E705G
V702A	A605V+M752T+V702A+E705G
E705G	F602S+A605V
M752T	F602S+A605V+M752T
	F602S+A605V+M752T+M560K
	F602S+A605V+M752T+V702A+E705G

4.3.5.1 Klonierung von *hgr-lbd* Einzel- und Kombinationsvarianten in pER13a und pERN7

Die Klonierung der *hgr-lbd* Einzelvarianten in pERN7 wurde bereits in 4.3.4 beschrieben. Die Klonierung der entsprechenden pER13a_*hgr-lbd* Konstrukte wurde analog hierzu durchgeführt, wobei jedoch pER13a_*hgr-lbd* bei den jeweiligen Amplifikationsschritten als Templat diente und 5'pER13 bzw. 3'pER13 als genflankierender Primer bei der Megaprimer-PCR bzw. der finalen Amplifikation des Volllängengens benutzt wurden.

Die *hgr-lbd* Kombinationsvarianten wurden hergestellt, indem sukzessive zusätzliche Mutationen in pERN7_*hgr-lbd*(F602S) bzw. pER13a_*hgr-lbd*(F602S) und pERN7_*hgr-lbd*(A605V) bzw. pER13a_*hgr-lbd*(A605V) durch Megaprimer-PCR (3.3.3.1) eingeführt wurden. Der Megaprimer wurde hierbei stets mit einem mutagenen Primer, welcher einen oder mehrere Austausche beinhaltete (3'*hgr-lbd* (M560K), 3'*hgr-lbd* (A605V), 3'*hgr-lbd* (F602S+A605V), 5'*hgr-lbd* (V702A), 5'*hgr-lbd* (E705G), 5'*hgr-lbd* (V702A+E705G)) und einem genflankierenden Primer (bei pER13a: 5'pER13 bzw. 3'pER13; bei pERN7: pERN7up bzw. T7 Terminator) mittels PCR (3.3.1) amplifiziert, wobei zuerst die oben genannten Vektoren und dann sukzessive die pERN7_*hgr-lbd* bzw. pER13a_*hgr-lbd* Konstrukte mit den zuvor eingefügten Austauschen als Templat dienten. Die Volllängen-Gene wurden anschließend mit Hilfe des aus einem Agarosegel aufgereinigten (3.3.5) Megaprimers und des entsprechenden genflankierenden Primers (bei pER13a: 5'pER13 bzw. 3'pER13; bei pERN7:

pERN7up bzw. T7 Terminator) mit Hilfe desselben Templats wie zuvor durch PCR amplifiziert. Mittels der Restriktionsenzyme *NdeI* und *NotI* konnten die erhaltenen mutagenisierten Gene jeweils in pERN7 bzw. pER13a kloniert werden und dienten dann als Templat für die nächste Megaprimer-PCR, bei welcher ein neuer Austausch hinzugefügt werden konnte. Nach jeder Klonierung wurde die Nukleotidsequenz durch Sequenzierung mit vektorspezifischen Primern (bei pER13a: 5'pER13 bzw. 3'pER13; bei pERN7: pERN7up bzw. T7 Terminator) bestätigt.

4.3.5.2 Analyse der Fluoreszenz von *E. coli* durch die eGFP-fusionierten hGR-LBD Einzel- und Kombinationsvarianten

Der spezifische Einfluss von M560K, F602S, A605V, V702A, E705G bzw. M752T auf die Fluoreszenz *in vivo* von hGR-LBD-eGFP wurde nun im Detail analysiert. Hierzu wurden die in 4.3.5.1 hergestellten pER13a_hgr-lbd Einzel- und Kombinationsvarianten in frisch hergestellte *E. coli* BL21(DE3) Zellen (3.2.4) transformiert (3.2.6), woraufhin die Expression bei 30 °C nach Zugabe von 500 µM Dexamethason für 6 h mit 0,5mM IPTG induziert wurde (3.2.8.1). Die Fluoreszenz der Expressionskulturen wurde mit Hilfe von Fluoreszenz-Mikroskopie (3.2.10) und Durchflusszytometrie (3.2.9) charakterisiert. Um auch sehr kleine Unterschiede in der Fluoreszenz eindeutig bestimmen zu können, wurde bei der FACS-Analyse jede Probe dreimal ausgezählt. Nach Bestimmung der mittleren Fluoreszenz der Proben (FL1>10) wurde nur der Medianwert der drei Analysen und das dazugehörige FACS-Histogramm weiter verwendet.

Die Analyse der verschiedenen Expressionskulturen zeigte, dass die Einführung von F602S, A605V, V702A, E705G bzw. M752T in hGR-LBD-eGFP jeweils zu einer Erhöhung der Fluoreszenz führt, wobei die positiven Effekte additiv wirken (Abbildung 52a). Durch Kombination der neu identifizierten Austausche A605V, V702A, E705G und M752T, welche die lösliche Herstellung von hGR-LBD in *E. coli* verbessern (4.3.4.), konnte so die mittlere Fluoreszenz (mf) der Expressionskulturen von 83,8 auf 122,9 relative Fluoreszenz-Einheiten erhöht werden (Abbildung 51a, 52a). Die zusätzliche Einführung der bekannten F602S Mutation (Bledsoe *et al.*, 2002) in die Kombinationsvariante führte zu einer weiteren Fluoreszenz-Steigerung (mf = 132,8, Abbildung 51b). Die stärkere Fluoreszenz der Zellen kann in beiden Varianten auf eine Erhöhung der Fluoreszenz in der löslichen Zellfraktion zurückgeführt werden, was durch ein stärker fluoreszierendes Cytoplasma dokumentiert wird (Abbildung 51a,b). Somit spiegelt hier die Fluoreszenz des C-terminal fusionierten Reporters eGFP die erhöhte Löslichkeit des Proteins sehr gut wider. Im Gegensatz dazu führt die

Einführung von M560K in hGR-LBD-eGFP zwar auch zu einer Verstärkung der Zell-Fluoreszenz auf $mf = 148,0$ (einzeln) bzw. $mf = 150,6$ (in Kombination mit F602S+A605V+M752T), allerdings durch Erhöhung der Fluoreszenz der *inclusion bodies*, welche an den Zell-Polen abgelagert werden (Abbildung 51c,d). Dies erklärt weshalb M560K einerseits in hGR-LBD-eGFP A8 zusammen mit A605V zu einer starken Erhöhung der Fluoreszenz führt (4.3.2.2), andererseits aber die lösliche Herstellung von hGR-LBD in *E. coli* nicht verbessert (4.3.4). Somit repräsentiert die Mutation M560K neben Verkürzungen auf Genebene (siehe 4.3.2.1, 4.3.2.2) oder Proteinebene (siehe 4.3.4) eine weitere Ursache falsch positiver Artefakte im eGFP-Reportersystem. Hierbei kommt es nicht durch lösliches eGFPwt, sondern durch erhöhte Fluoreszenz des unlöslichen hGR-LBD-eGFP Proteins zu einem falsch positiven Signal. Die Ursache für diesen Effekt ist unklar, wobei in der Literatur unterschiedlich starke Aktivität von Proteinen in *inclusion bodies* bereits mehrfach beschrieben worden ist (Ventura & Villaverde, 2006).

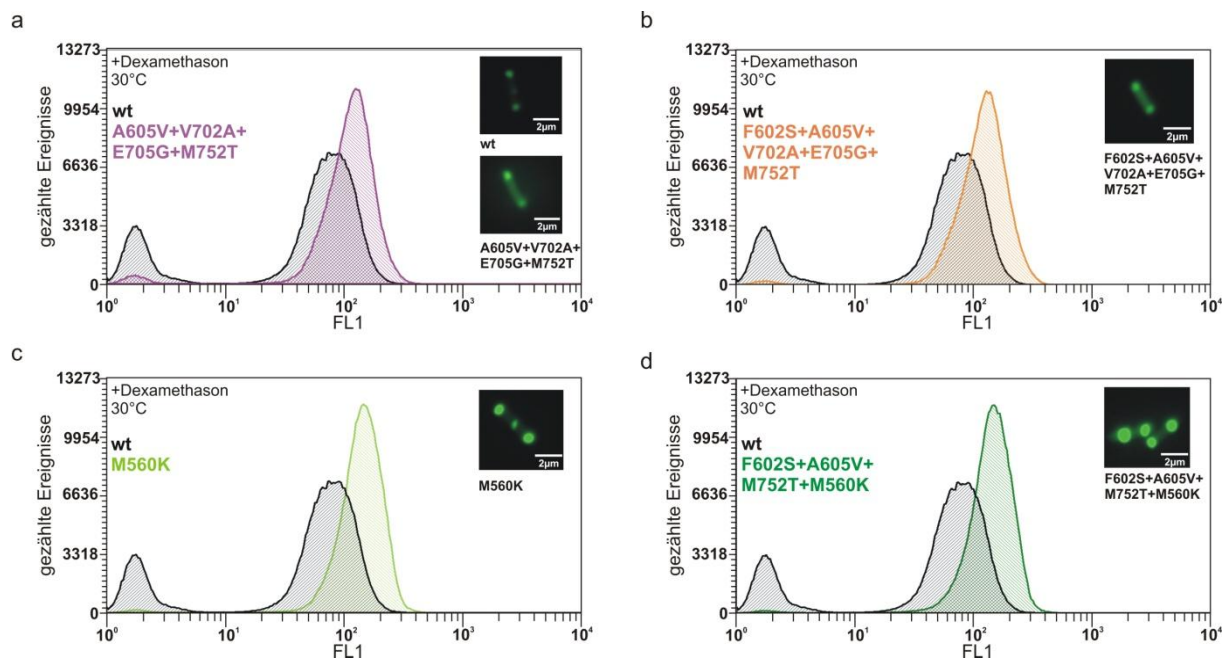


Abbildung 51: FACS-Analyse und Fluoreszenz-Mikroskopie von Zellen mit hGR-LBDwt (schwarz; mittlere Fluoreszenz (mf): 83,8) im Vergleich zu Zellen mit hGR-LBD Einzel- und Kombinationsvarianten, nach Expression bei 30 °C in Anwesenheit von Dexamethason.

Die Expression der *egfp*-fusionierten Gene erfolgte jeweils in *E. coli* BL21(DE3), für 6 h bei 30 °C in Anwesenheit von 500 µM Dexamethason. Bei der Durchflusszytometrie wurden jeweils 300.000 Ereignisse analysiert. Die Fluoreszenz-Mikroskopie wurde bei allen Expressionskulturen an einem Zeiss Axiovert 200M Mikroskop bei konstanter Belichtungsdauer (100 ms) durchgeführt.

hGR-LBD(A605V+V702A+E705G+M752T) (a, magenta) bzw. hGR-LBD(F602S+A605V+V702A+E705G+M752T) (b, orange) zeigen eine erhöhte mittlere Fluoreszenz (mf) von 122,9 bzw. 132,8 im Vergleich zu hGR-LBDwt ($mf = 83,8$) durch eine stärkere Fluoreszenz im Cytoplasma (= lösliche Zellfraktion). Die Einführung der Mutation M560K führt vereinzelt (c, hellgrün), aber auch in Kombination mit vorteilhaften Mutationen (d, dunkelgrün) zu einer hohen mittleren Fluoreszenz von 148,0 bzw. 150,6 durch eine Steigerung der Fluoreszenz in den *inclusion bodies* (= unlösliche Zellfraktion).

Anders als bei der Verkürzung des Zielproteins auf Genebene (4.3.2.1, 4.3.2.2) oder Proteinebene (4.3.4) können solche Artefakte nicht durch Überprüfung der korrekten Länge des Gens (Abbildung 45) bzw. des Proteins (Abbildung 50) erkannt werden. Um sie dennoch zügig identifizieren zu können, sollten bei zukünftigen Experimenten die sortierten Varianten sofort nochmals in *E. coli* exprimiert und mit Hilfe von Fluoreszenz-Mikroskopie die Lokalisation der erhöhten Fluoreszenz (Cytoplasma = lösliches Protein; Zell-Pole = unlösliche *inclusion bodies*) überprüft werden.

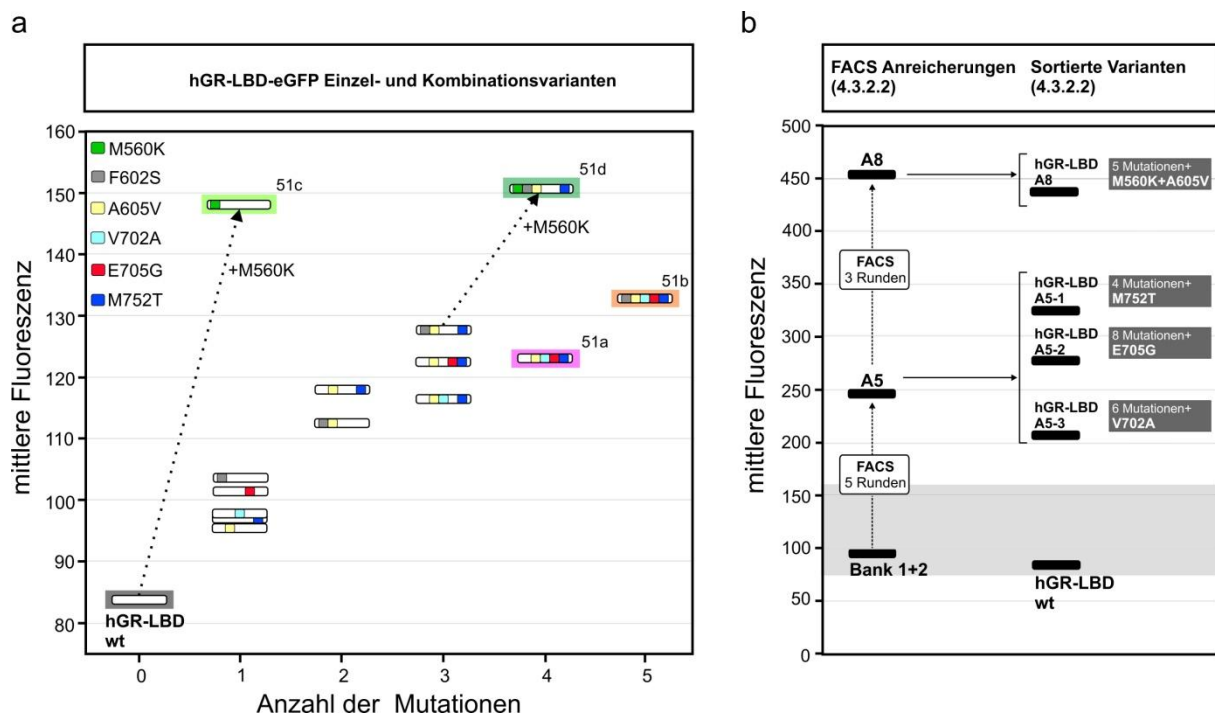


Abbildung 52: Mittlere Fluoreszenz von *E. coli* in der Durchflusszytometrie durch (a) gezielt hergestellte Einzel- und Kombinationsvarianten von hGR-LBD-eGFP und (b) durch FACS angereicherte Populationen (links) bzw. sortierte Varianten (rechts).

Die Expression der *egfp*-fusionierten Gene erfolgte jeweils in *E. coli* BL21(DE3), für 6 h bei 30 °C nach Zugabe von 500 µM Dexamethason. Es wurden jeweils 300.000 Ereignisse analysiert. Alle hier gezeigten Werte sind Medianwerte von drei Messungen und wurden parallel bei konstanter Kalibrierung des MoFlo[®] Zytometers aufgenommen.

a Mittlere Fluoreszenz von *E. coli*, vermittelt durch hGR-LBD-eGFP Einzel- und Kombinationsvarianten. Die in den FACS-Histogrammen in Abbildung 51 gezeigten Varianten sind farbig unterlegt bzw. mit der entsprechenden Abbildungsnummerierung versehen. Die Austausche die zu einer tatsächlich verbesserten löslichen Herstellung von hGR-LBD in *E. coli* führen (4.3.4) - F602S (Bledsoe *et al.*, 2002) und A605V, V702A, E705G und M752T (in dieser Arbeit identifiziert: 4.3.2.2, 4.3.4) - führen zu einer additiven Verbesserung der hGR-LBD-eGFP vermittelten Fluoreszenz *in vivo*. Die falsch positive Mutation M560K führt zu einer ungewöhnlich hohen Fluoreszenz der *inclusion bodies* (Abbildung 51c,d), was durch gepunktete Pfeile angedeutet ist.

b Mittlere Fluoreszenz von *E. coli*, vermittelt durch FACS-Anreicherungen bzw. sortierte hGR-LBD-eGFP Varianten aus 4.3.2.2. Die FACS-Anreicherungen bzw. die sortierten hGR-LBD-eGFP Varianten A5-1 - A5-3 und A8 aus denen die Austausche M560K, A605V, V702A, E705G und M752T identifiziert wurden, führen zu deutlich höherer Fluoreszenz als die vereinzelt bzw. neu kombinierten Austausche (in a gezeigter Fluoreszenzbereich ist grau hinterlegt).

Um die hGR-LBD-eGFP Einzel- und Kombinationsvarianten mit den sortierten hGR-LBD-eGFP Varianten A5-1 - A5-3 und A8, aus denen die untersuchten Mutationen identifiziert wurden, vergleichen zu können, wurden diese Varianten, die hGR-LBD Banken 1+2 und die daraus resultierten hGR-LBD-eGFP Anreicherungen parallel mit den Einzel- und Kombinationsvarianten auf gleiche Weise bei konstanter Kalibrierung des MoFlo[®] Zytometers analysiert (Abbildung 52b). Dieser Vergleich macht deutlich, dass die ursprüngliche Zusammensetzung der Mutationen in den sortierten hGR-LBD-eGFP Varianten A5-1 - A5-3 und A8 zu einer wesentlich höheren Fluoreszenz führt, als dies bei Vereinzeln bzw. Neu-Kombination der positiven Mutationen A605V, V702A, E705G und M752T in hGR-LBD-eGFP der Fall ist (Abbildung 52). Dieses Phänomen wird vermutlich durch synergistische Wechselwirkungen der Mutationen in hGR-LBD-eGFP A5-1 - A5-3 und A8 (siehe Tabelle 12), sowie durch Mutationen wie M560K, welche zu einer falsch positiven Verstärkung des Fluoreszenz-Signals führen, hervorgerufen.

4.3.5.3 Quantitative Analyse der löslichen Herstellung der hGR-LBD Einzel- und Kombinationsvarianten in *E. coli*

Die lösliche Herstellung der N-terminal mit His₆-GST fusionierten hGR-LBD Einzel- und Kombinationsvarianten in *E. coli* wurde mit Hilfe von SDS-PAGE (3.5.2) analysiert. Hierzu wurde *E. coli* BL21(DE3) mit den entsprechenden pERN7-Konstrukten transformiert (3.2.3) und Probeexpressionen bei 20 °C in Anwesenheit von Dexamethason durchgeführt (3.4.1.1). Die Menge an löslichem His₆-GST-hGR-LBD konnte anschließend über densitometrische Analyse der Proteinbanden der gereinigten löslichen Zellfraktion auf SDS-PA Gelen quantifiziert werden (3.5.3).

Durch die bekannte Mutation F602S (Bledsoe *et al.*, 2002) bzw. durch die neu identifizierten Mutationen A605V, V702A, E705G und M752T (4.3.2.2, 4.3.4) wird eine 3- bis 6-fach höhere Menge an Protein in der gereinigten löslichen Zellfraktion erhalten im Vergleich zum wildtypischen Protein (Abbildung 53).

Darüber hinaus zeigt die Analyse der hGR-LBD Kombinationsvarianten, dass diese positiven Aminosäureaustausche additiv die Löslichkeit verbessern. Der lösliche Anteil an Protein kann so bis auf ca. 90 % (F602S+A605V+M752T+V702A+E705G) gesteigert werden, wobei sich dann in der gereinigten löslichen Zellfraktion eine ca. 28-fach höhere Menge des Proteins befindet, verglichen mit hGR-LBDwt (Abbildung 53). Die Kombination nur der neu identifizierten Aminosäureaustausche A605V, V702A, E705G und M752T führt zu einer ca.

19-fachen Steigerung der löslichen Proteinmenge. Durch die neu entdeckten Mutationen allein kann somit die Löslichkeit des Proteins ebenfalls sehr stark erhöht werden, wodurch erstmals die Produktion und Untersuchung von hGR-LBD ohne eine Mutation von F602 möglich wird.

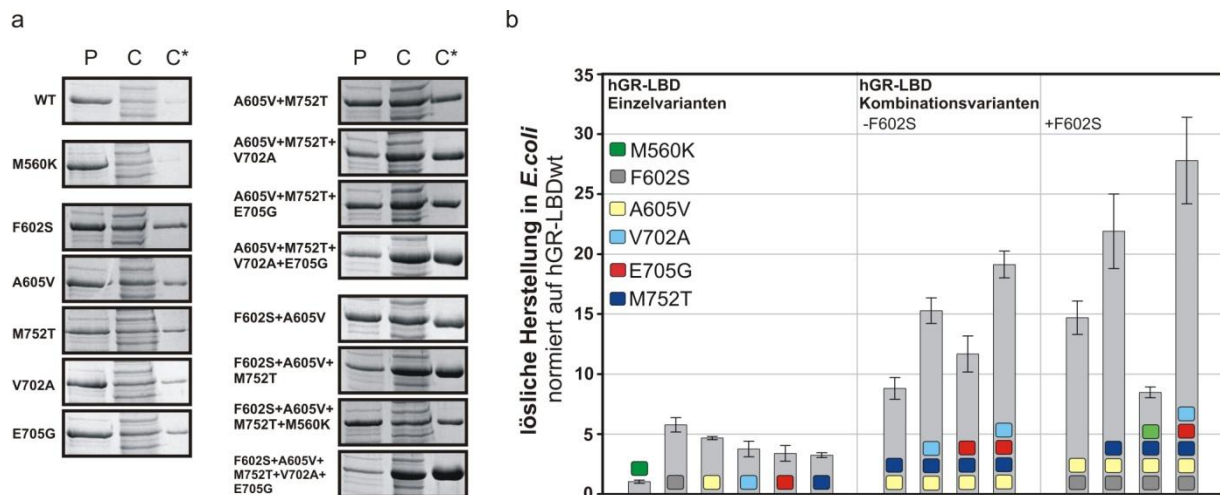


Abbildung 53: Quantitative Analyse der löslichen Herstellung His₆-GST-fusionierter hGR-LBD Einzel- und Kombinationsvarianten in *E. coli* BL21(DE3), dokumentiert durch SDS-PAGE (12,5 %).

P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*), C*: über Nickel-Sepharose gereinigte lösliche Zellfraktion. Die Expression erfolgte jeweils in *E. coli* BL21(DE3), üN bei 20 °C in Anwesenheit von 500 µM Dexamethason. Es wurde stets die gleiche Anzahl an Zellen ($2 \cdot 10^9$) analysiert.

a Es ist lediglich der 45-66 kDa Bereich des SDS-PA Gels gezeigt, in welchem sich His₆-GST-hGR-LBD (Molekulargewicht: ~58 kDa) befindet. Die Mutationen F602S (Bledsoe *et al.*, 2002) bzw. A605V, V702A, E705G und M752T (in dieser Arbeit identifiziert: 4.3.4) führen vereinzelt (Zellfraktionen entsprechen den in Abbildung 50 gezeigten) und additiv zu einer Verbesserung der löslichen Herstellung von His₆-GST-hGR-LBD in *E. coli*.

b Quantifizierung der gereinigten löslichen His₆-GST-hGR-LBD Einzel- und Kombinationsvarianten. Die in 3.5.3 beschriebene Quantifizierung der C*-Proteinbanden durch OptiQuant (Version 02.50) wurde bei jeder Variante von je drei Präparationen mit je mindestens zwei verschiedenen SDS-PA Gelen durchgeführt, was die Bestimmung der gezeigten Mittelwerte und Standardabweichungen ermöglichte.

Die aus dem *Screening* (4.3.2.2) falsch positiv identifizierte Mutation M560K führt wie bereits beschrieben (4.3.4) vereinzelt zu keiner Verbesserung der löslichen Herstellung von His₆-GST-hGR-LBD. Darüber hinaus wird die Löslichkeit einer verbesserten Variante (F602S+A605V+M752T) sogar deutlich verschlechtert sobald M560K hinzugefügt wird (Abbildung 53).

Die so quantifizierte Menge an löslich hergestelltem hGR-LBD (fusioniert an His₆-GST) in *E. coli* bei den verschiedenen Einzel- und Kombinationsvarianten korreliert mit der durch die entsprechenden hGR-LBD Varianten (fusioniert an eGFP) erzeugten mittleren Fluoreszenz von *E. coli* (4.3.5.2) (Abbildung 54). Dieses Ergebnis macht deutlich, dass eGFP tatsächlich sehr gut geeignet ist die unterschiedliche Löslichkeit eines Proteins *in vivo* darzustellen,

wodurch es sich anbietet dieses Protein weiterhin als Reporter für die gelenkte Evolution hoher Proteinlöslichkeit zu verwenden. Allerdings muss berücksichtigt werden, dass es zu artifiziellen Signalen kommen kann. So stellen die hGR-LBD Varianten mit der falsch positiven Mutation M560K, welche zur Erhöhung der hGR-LBD-eGFP *-inclusion body* Fluoreszenz führt, Ausreißer im Korrelationsdiagramm dar (Abbildung 54).

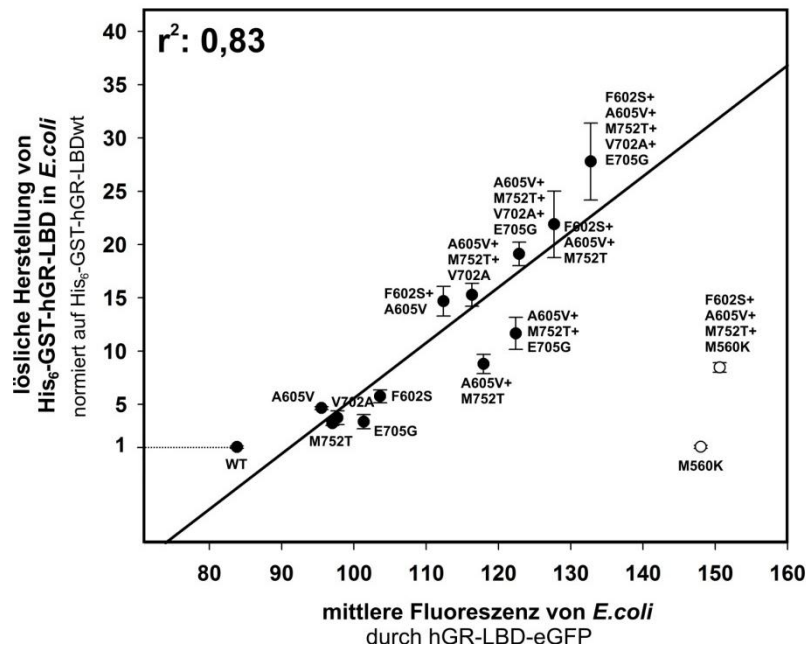


Abbildung 54: Korrelation zwischen der hGR-LBD-eGFP vermittelten Fluoreszenz und der Menge an löslich hergestelltem His₆-GST-hGR-LBD in *E. coli*.

Die durch hGR-LBD-eGFP vermittelte mittlere Fluoreszenz von *E. coli* wurde durch Durchflusszytometrie analysiert (siehe 4.3.5.2, Abbildung 52). Die Menge an löslich hergestelltem His₆-GST-hGR-LBD wurde durch densitometrische Auswertung von Probeexpressionen quantifiziert (Abbildung 53). Gefüllte Kreise: Bei den untersuchten hGR-LBD Einzel- und Kombinationsvarianten besteht eine Korrelation zwischen der Fluoreszenz von hGR-LBD-eGFP *in vivo* und der Löslichkeit des entsprechenden hGR-LBD Proteins (fusioniert an His₆-GST) *in vivo* (linearer Fit: $r^2 = 0,83$). Leere Kreise: hGR-LBD Varianten mit der Mutation M560K stellen Ausreißer dar, da bei ihnen die Fluoreszenz der unlöslichen Fusionsprotein enthaltenden *inclusion bodies* stark erhöht ist (siehe 4.3.5.2, Abbildung 51c,d).

4.3.5.4 Reinigung der hGR-LBD Einzel- und Kombinationsvarianten

Um die biochemischen und biophysikalischen Eigenschaften der in Tabelle 14 aufgelisteten hGR-LBD Einzel- und Kombinationsvarianten näher untersuchen zu können, wurden diese im präparativen Maßstab gereinigt (Ausbeuten: siehe Anhang 7.4, Tabelle A1). Die Herstellung der verschiedenen an His₆-GST fusionierten hGR-LBD Varianten erfolgte in mit den entsprechenden pERN7_hgr-lbd Plasmiden transformierten *E. coli* BL21(DE3) Zellen über Nacht bei 20 °C nach Zugabe von Dexamethason wie in 3.4.1.2 beschrieben. In Abbildung 55 ist als Beispiel die Reinigung von hGR-LBD(M560K+F602S+A605V+M752T) dargestellt.

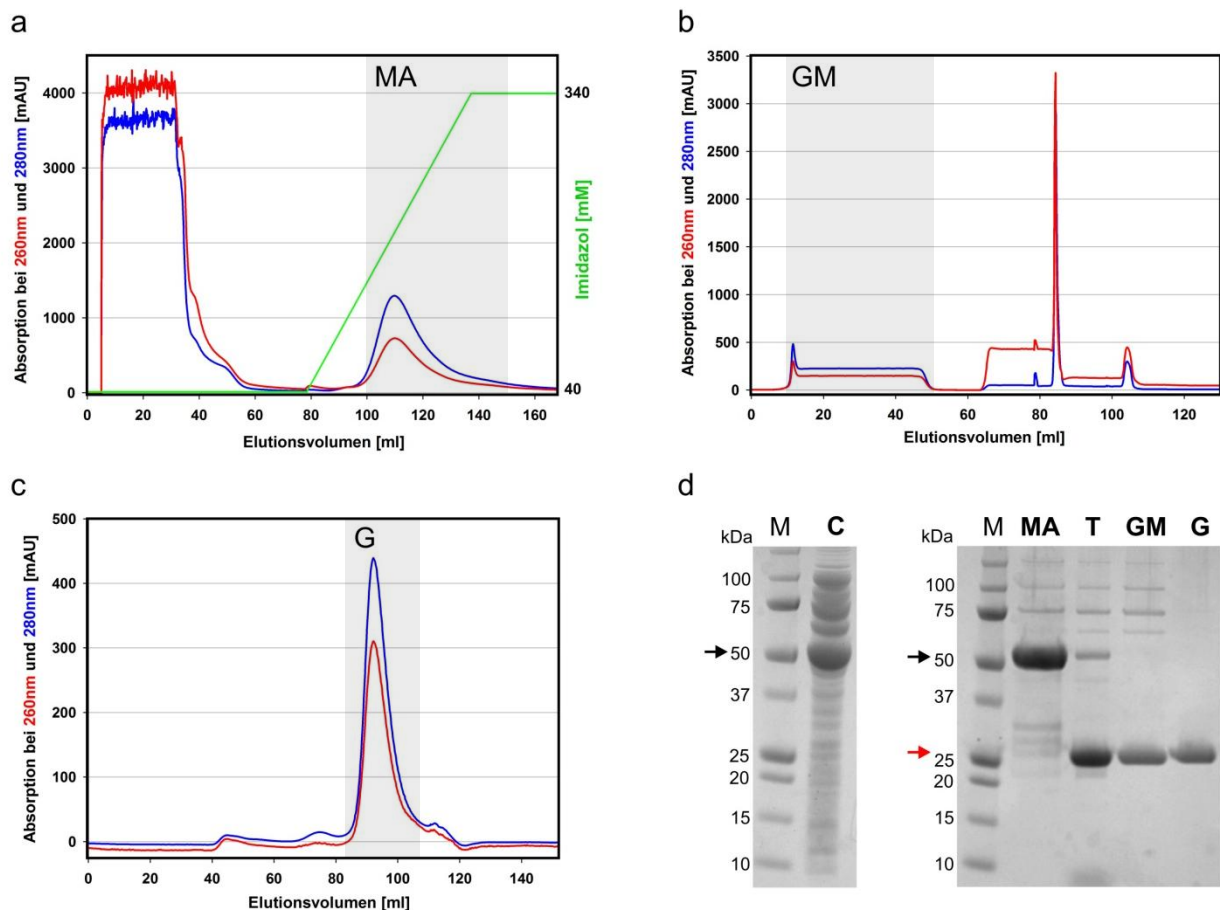


Abbildung 55: Reinigung des rekombinant hergestellten hGR-LBD(M560K+F602S+A605V+M752T) Proteins aus dem löslichen Zellextrakt von *E. coli* BL21(DE3).

Gesammelte Fraktionen sind in den Elutionsprofilen jeweils grau hinterlegt.

a Elutionsprofil der Reinigung über Metallchelate-Affinitätschromatographie, aufgenommen über die Messung der Absorption bei 260 nm (rot) bzw. 280 nm (blau). Die Elution erfolgte über einen Gradienten von 40 – 340 mM Imidazol.

b Elutionsprofil der Reinigung über gekoppelte Glutathion- und Metallchelate-Affinitätschromatographie, aufgenommen über die Messung der Absorption bei 260 nm (rot) bzw. 280 nm (blau). hGR-LBD befindet sich im Durchfluss (- 50 ml). Die Elution von gebundenem His₆-GST durch Imidazol und L-Glutathion (- 110 ml) erfolgt zur Reinigung der Säulen.

c Elutionsprofil der finalen Reinigung von hGR-LBD über präparative Gelfiltration, aufgenommen über die Messung der Absorption bei 260 nm (rot) bzw. 280 nm (blau).

d SDS-PAGE (gezeigt sind NuPAGE® Novex Bis-Tris Gele (4-12 %), Invitrogen) der verschiedenen Reinigungsschritte. M: Markerproteine (Precision Standard Protein All Blue, Bio-Rad); C: lösliche Zellfraktion (*crude extract*); MA: nach Metallchelate-Affinitätschromatographie (vgl. a); T: nach Thrombin-Verdau; GM: nach gekoppelter Glutathion- und Metallchelate-Affinitätschromatographie (vgl. b); G: nach präparativer Gelfiltration (vgl. c). Bei den SDS-PA Gelen sind die Proteinbanden von His₆-GST-hGR-LBD (~58 kDa) bzw. hGR-LBD (~30 kDa) mit einem schwarzen bzw. rotem Pfeil markiert.

Die Reinigung der rekombinanten Proteine aus dem löslichen Zellextrakt erfolgte zunächst über Metallchelate-Affinitätschromatographie (3.4.2.1, Abbildung 55a), wobei die His₆-GST-hGR-LBD Proteine jeweils bei ca. 150 – 340 mM Imidazol eluierten. Nach Abspaltung des N-terminal fusionierten His₆-GST durch Thrombin (3.4.2.2), erfolgte dessen Abtrennung aus der Proteinpräparation über eine gekoppelte Glutathion- und Metallchelate-Affinitätschromatographie (3.4.2.3, Abbildung 55b).

Anschließend wurde der gesammelte Durchfluss vereinigt und auf eine präparative Gelfiltrationssäule gegeben, um letzte Verunreinigungen zu entfernen (3.4.2.4, Abbildung 55c). Alle Reinigungsschritte mit Ausnahme der finalen Gelfiltration wurden in Anwesenheit von Dexamethason durchgeführt. Die letztendlich erhaltenen hGR-LBD Proteine in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9 wurden in flüssigen Stickstoff eingetropfelt und bei -80 °C gelagert (3.4.2.7).

Die einzelnen Reinigungsschritte wurden mit Hilfe von SDS-PAGE (3.5.2, Abbildung 55d) und zum Teil über analytische *reversed-phase* HPLC (3.5.6, Abbildung 56) verfolgt. Hierbei zeigten die final erhaltenen Proteinpräparationen stets einen sehr hohen Reinheitsgrad. Dies ist in Abbildung 56 am Beispiel von hGR-LBD(A605V+M752T) gezeigt.

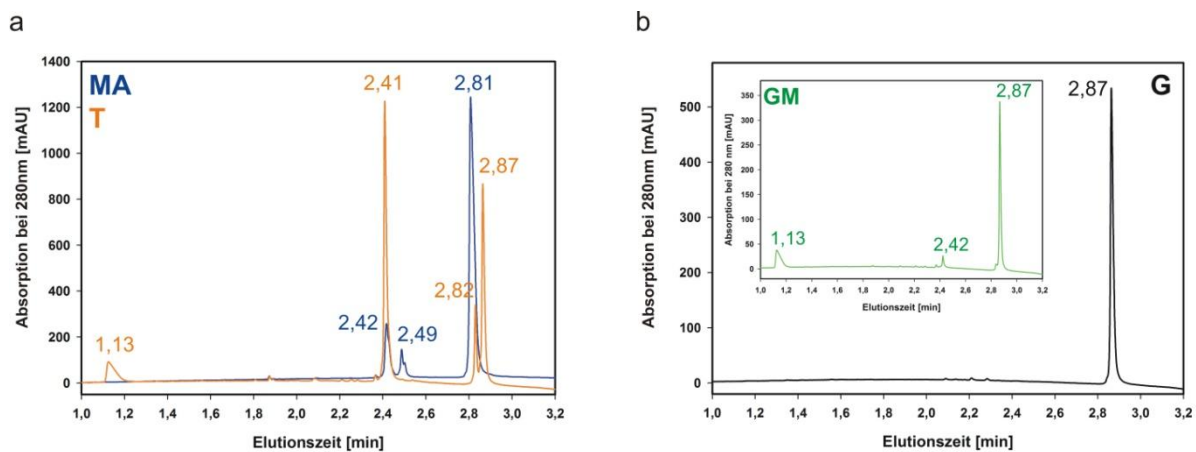


Abbildung 56: Reversed-phase HPLC Analyse der einzelnen Reinigungsschritte des rekombinant hergestellten hGR-LBD(A605V+M752T) Proteins.

Es ist jeweils die Analyse der vereinigten Fraktionen der einzelnen Schritte der Proteinreinigung (siehe Abbildung 55) durch eine C8(Poroshell 2.1x50 mm) -Säule gezeigt. Die Elution erfolgte jeweils mit einem Acetonitril-Gradienten (10-60 %). Zum besseren Vergleich der Elutionsprofile von MA und T, wurden die Signale von T mit 2,5 multipliziert.

a MA: nach Metallchelate-Affinitätschromatographie (2,42 min / 2,49 min: Proteinverunreinigungen; 2,81 min: His₆-GST-hGR-LBD); T: nach Thrombin-Verdau (1,13 min: CHAPS; 2,41 min: His₆-GST; 2,82 min: His₆-GST-hGR-LBD; 2,87 min: hGR-LBD)

b GM: nach gekoppelter Glutathion- und Metallchelate-Affinitätschromatographie (1,13 min: CHAPS; 2,42 min: Proteinverunreinigung; 2,87 min: hGR-LBD); G: nach finaler präparativer Gelfiltration (2,87 min: hGR-LBD).

Alle hGR-LBD Einzel- und Kombinationsvarianten mit den vorteilhaften Mutationen F602S bzw. A605V, V702A, E705G oder M752T konnten so aufgereinigt werden. Die Ausbeuten der Einzelvarianten lagen hierbei im Bereich von 0,1 – 0,2 mg pro Gramm Nassgewicht der *E. coli* – Expressionskultur (~ 0,4 – 0,9 mg Protein / Liter Expressionskultur) (siehe Anhang 7.4, Tabelle A1). Durch Kombination der neu identifizierten Austausche A605V, V702A, E705G und M752T konnte die Ausbeute auf bis zu 3,5 mg Protein / g Zellen (~ 16 mg Protein / Liter Expressionskultur) erhöht werden. Bei Reinigung der löslichsten Variante

hGR-LBD(F602S+A605V+V702A+E705G+M752T) wurde 5,3 mg Protein / g Zellen (~ 24 mg Protein / Liter Expressionskultur) erhalten. Die in dieser Arbeit gefundenen Mutationen ermöglichen somit zum ersten Mal die Reinigung von hGR-LBD in sehr großen Mengen, wodurch weitere Untersuchungen deutlich vereinfacht bzw. erst möglich werden. So kann nun zum Beispiel hGR-LBD in sehr hohen Konzentrationen für die Röntgenkristallstrukturanalyse des Proteins mit unterschiedlichen Liganden eingesetzt werden.

4.3.5.5 Konformationelle Stabilität der hGR-LBD Einzel- und Kombinationsvarianten

Um den Einfluss der Mutationen M560K, F602S, A605V, V702A, E705G und M752T auf die konformationelle Stabilität von hGR-LBD zu untersuchen, wurden die in Tabelle 14 aufgelisteten hGR-LBD Einzel- und Kombinationsvarianten in Anwesenheit von gebundenem Agonisten (Dexamethason) bzw. Antagonisten (Mifepriston) thermisch aufgefaltet (3.5.10) (Abbildung 57). In Abwesenheit eines Liganden ist hGR-LBD sehr instabil und kann nicht aufgereinigt bzw. untersucht werden (siehe 4.3.5.7). Aus den Auffaltungskurven konnte für jede hGR-LBD Variante eine apparente Schmelztemperatur (T_M^{app} = Temperatur bei der die Hälfte des Proteins in nicht-nativem Zustand vorliegt) bestimmt werden (3.5.10), welche als operationelles Maß für die Proteinstabilität diente (Tabelle 15). Da das wildtypische hGR-LBD Protein nicht aufgereinigt werden kann und somit auch die Bestimmung seines T_M^{app} – Wertes nicht möglich war, wurden die Effekte der einzelnen Mutationen aus dem Vergleich der verschiedenen Einzel- und Kombinationsvarianten abgeleitet.

Die gemessenen bzw. abgeleiteten T_M^{app} -Werte zeigten, dass die vorteilhaften Mutationen F602S (Bledsoe *et al.*, 2002) bzw. A605V, V702A, E705G und M752T (4.3.2.2, 4.3.4), welche bereits die Fluoreszenz *in vivo* von hGR-LBD-eGFP (4.3.5.2) und die lösliche Herstellung von hGR-LBD (4.3.5.3, 4.3.5.4) deutlich verbessern, vereinzelt und additiv auch die konformationelle Stabilität von hGR-LBD erhöhen. Hierbei führt die Kombination der im Rahmen dieser Arbeit identifizierten Austausch (A605V+V702A+E705G+M752T) zu einer Stabilisierung der Agonist-Konformation von +8,3 °C und der Antagonist-Konformation von +8,2 °C. Der bekannte Austausch F602S stabilisiert die Agonist- bzw. die Antagonist-Konformation des Proteins um +4,5 °C bzw. um +5,0 °C.

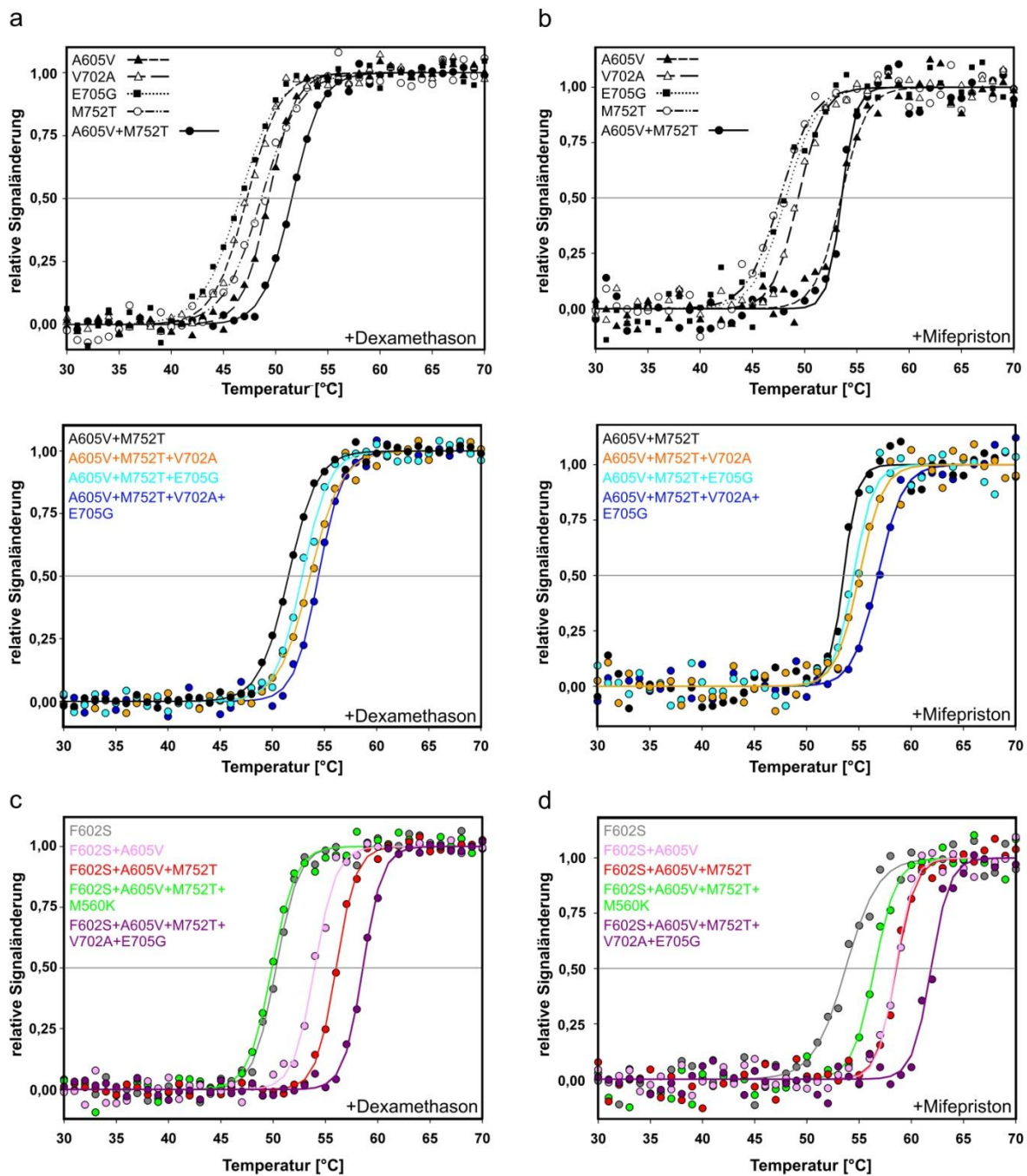


Abbildung 57: Thermische Auffaltungen der hGR-LBD Einzel- und Kombinationsvarianten.

Die Auffaltungen aller Proteine erfolgte mit einer Rate von 1 °C/min von 30 °C bis 70 °C in 25 mM Tris/Cl, 0,1 M NaCl, 2 mM TCEP, 10 % Glycerin (pH 7,9 bei RT; pH 7,8 - 6,6 bei 30 - 70 °C, da $\Delta pK_a/^\circ\text{C}(\text{Tris}) = -0,028$, Mandaji *et al.*, 2009) und wurde durch CD Spektroskopie bei 220 nm verfolgt.

Zu den Ansätzen wurde zusätzlich entweder der Agonist Dexamethason (200 μM ; Proteinkonzentration: 10 μM) (a/c) oder der Antagonist Mifepriston (50 μM ; Proteinkonzentration: 2,5 μM) (b/d) zugegeben. Aufgrund der niedrigen Löslichkeit von Mifepriston in wässrigem Puffer wurde in diesen Ansätzen eine niedrigere Proteinkonzentration eingesetzt, um die vollständige Sättigung des Proteins mit dem Liganden garantieren zu können ($K_D(\text{Mifepriston}) < 1 \text{ nm}$; Jung-Testas *et al.*, 1983). Um das so verschlechterte CD-Signal - Rausch Verhältnis des Proteins zu optimieren wurden die Auffaltungen mit Mifepriston mindestens dreimal durchgeführt und gemittelt.

a/b Auffaltungskurven von hGR-LBD Einzel- und Kombinationsvarianten mit Aminosäuremutationen, welche im Rahmen dieser Arbeit identifiziert wurden (A605V, V702A, E705G und M752T). Die Auffaltungen wurden in Anwesenheit des Agonisten Dexamethason (a) oder des Antagonisten Mifepriston (b) durchgeführt.

c/d Auffaltungskurven von hGR-LBD Einzel- und Kombinationsvarianten, welche F602S (Bledsoe *et al.*, 2002) enthalten. Die Auffaltungen wurden in Anwesenheit des Agonisten Dexamethason (c) oder des Antagonisten Mifepriston (d) durchgeführt.

Tabelle 15: Apparente T_M -Werte (T_M^{app}) der hGR-LBD Einzel- und Kombinationsvarianten.

Die durch die eingeführten Mutationen induzierten Veränderungen des T_M^{app} von hGR-LBD wurde durch den Vergleich der T_M^{app} -Werte verschiedener hGR-LBD Varianten ermittelt (verwendete Varianten sind in Klammern angegeben). Falls der Effekt einer Mutation von mehreren Vergleichen abgeleitet werden konnte, wird ein Mittelwert und die Maximal-Abweichung von diesem angegeben.

	T_M^{app}	
	+Agonist (200 μ M Dexamethason)	+Antagonist (50 μ M Mifepriston)
hGR-LBD Varianten ohne F602S		
¹ A605V	49,3°C	53,5°C
² M752T	48,6°C	47,6°C
³ V702A	47,2°C	49,4°C
⁴ E705G	46,6°C	48,2°C
⁵ A605V+M752T	51,5°C	53,5°C
⁶ A605V+M752T+V702A	53,6°C	55,0°C
⁷ A605V+M752T+E705G	52,9°C	54,5°C
⁸ A605V+M752T+V702A+E705G	54,4°C	56,8°C
hGR-LBD Varianten mit F602S		
⁹ F602S	50,3°C	53,7°C
¹⁰ F602S+A605V	54,0°C	58,5°C
¹¹ F602S+A605V+M752T	56,0°C	58,5°C
¹² F602S+A605V+M752T+M560K	49,9°C	56,5°C
¹³ F602S+A605V+M752T+V702A+E705G	58,6°C	61,9°C
Veränderungen gegenüber hGR-LBDwt		
F602S ^(10-1, 11-5, 13-8)	+4,5°C (\pm 0,3°C)	+5,0°C (\pm 0,1°C)
M560K ⁽¹²⁻¹¹⁾	-6,1°C	-2,0°C
A605V ^(5-2, 10-9)	+3,3°C (\pm 0,4°C)	+5,4°C (\pm 0,6°C)
V702A ^(6-5, 8-7)	+1,8°C (\pm 0,3°C)	+1,9°C (\pm 0,4°C)
E705G ^(7-5, 8-6)	+1,1°C (\pm 0,3°C)	+1,4°C (\pm 0,4°C)
M752T ^(5-1, 11-10)	+2,1°C (\pm 0,1°C)	\pm 0,0°C (\pm 0,0°C)
A605V+M752T+V702A+E705G ⁽¹³⁻⁹⁾	+8,3°C	+8,2°C

Während F602S, A605V, V702A bzw. E705G hGR-LBD sowohl mit gebundenem Dexamethason (Agonist) als auch mit gebundenem Mifepriston (Antagonist) stabilisieren, führt M752T nur in Anwesenheit des Agonisten zu einer Erhöhung des T_M^{app} -Wertes. Das *Screening* der hGR-LBD-eGFP Banken, durch welches die neuen positiven Mutationen identifiziert worden sind, erfolgte in Anwesenheit von Dexamethason (4.3.2.2). Es bestätigt sich hier somit einmal mehr die Beobachtung, dass bei gelenkter Evolution von Proteinen im Labor in allerster Linie solche Mutationen gefunden werden, welche unter den angelegten Bedingungen den untersuchten Effekt bewirken („*you get what you screen for*“; Arnold & Moore, 1997). Somit ist es eher erstaunlich, dass A605V, V702A und E705G beide Konformationen des Proteins stabilisieren. Mögliche Ursachen für die durch die einzelnen Mutationen vermittelte Stabilisierung von hGR-LBD werden in 4.3.5.8 näher besprochen.

Die falsch positive Mutation M560K, welche zu hoher Fluoreszenz von hGR-LBD-eGFP *inclusion bodies* in *E. coli* (4.3.5.2) jedoch zu keiner erhöhten Löslichkeit von His₆-GST-hGR-LBD führt (4.3.5.3), destabilisiert hGR-LBD deutlich um -6,1 °C (Agonist-Konformation) bzw. -2,0 °C (Antagonist-Konformation). Die erhöhte Fluoreszenz von *inclusion bodies* in *E. coli* stellt in diesem Fall somit keine erhöhte „konformationelle Qualität“ des aggregierten Proteins dar, wie dies in anderen Publikationen angedeutet wird (Garcia-Fruitos *et al.*, 2007).

Abgesehen von M560K korrelieren die ermittelten T_M^{app} -Werte der verschiedenen hGR-LBD Einzel- und Kombinationsvarianten gut mit der durch die entsprechenden hGR-LBD-eGFP Varianten erzeugten mittleren Fluoreszenz von *E. coli* (4.3.5.2) (Abbildung 58). Die Fluoreszenz des C-terminal fusionierten eGFP-Reporters *in vivo* spiegelt in diesem Fall somit nicht nur die Löslichkeit der hGR-LBD Varianten (siehe 4.3.5.3, Abbildung 54) sondern auch deren konformationelle Stabilität wider.

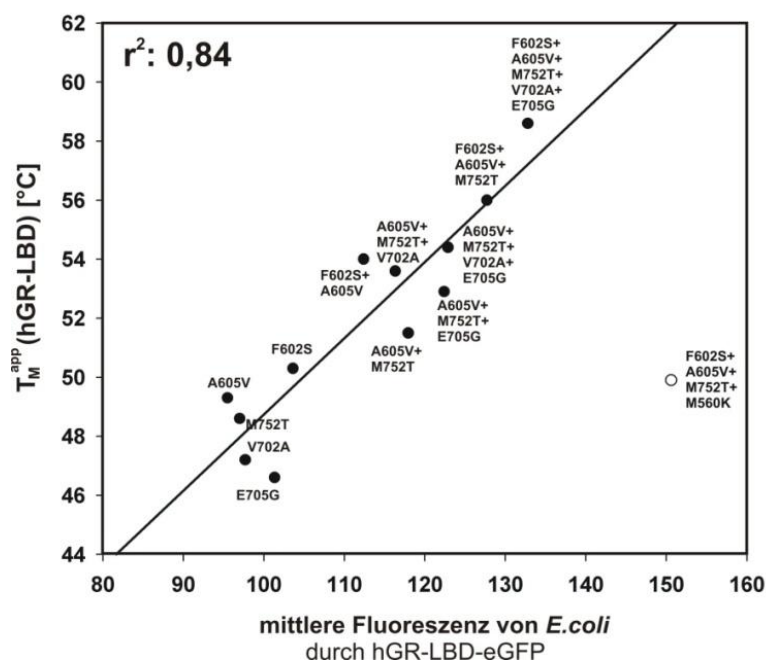


Abbildung 58: Korrelation zwischen der hGR-LBD-eGFP vermittelten Fluoreszenz in *E. coli* und der konformationellen Stabilität von gereinigtem hGR-LBD (Agonist-Konformation).

Die durch hGR-LBD-eGFP vermittelte mittlere Fluoreszenz von *E. coli* nach Genexpression in Anwesenheit von Dexamethason wurde durch Durchflusszytometrie analysiert (siehe 4.3.5.2, Abbildung 52). Die Stabilität der Agonist-Konformation der hGR-LBD Varianten wurde durch thermische Auffaltung in Anwesenheit von Dexamethason bestimmt, woraus apparente T_M -Werte abgeleitet werden konnten (Abbildung 57, Tabelle 15). Gefüllte Kreise: Bei den untersuchten hGR-LBD Einzel- und Kombinationsvarianten besteht eine Korrelation zwischen der Fluoreszenz *in vivo* von hGR-LBD-eGFP und der Stabilität des entsprechenden hGR-LBD Proteins (linearer Fit: $r^2 = 0,84$). Leerer Kreis: die hGR-LBD Variante mit der Mutation M560K stellt einen Ausreißer dar, da bei ihr die Fluoreszenz der unlöslichen Fusionsprotein enthaltenden *inclusion bodies* stark erhöht ist (siehe 4.3.5.2, Abbildung 51), welche aus instabilen und aggregierten Proteinspezies bestehen.

Zusammenfassend kann somit festgestellt werden, dass die Fluoreszenz von eGFP *in vivo* sehr gut als Reporter für die Löslichkeit und die konformationelle Stabilität eines N-terminal fusionierten Zielproteins verwendet werden kann. Allerdings müssen beim *Screening* von eGFP-fusionierten Banken die bereits beschriebenen Ursachen für falsch positive Artefakte (Verkürzungen auf Genebene durch neu eingeführte Stoppcodone bzw. Verlust des Zielgens: siehe 4.3.2.1 bzw. 4.3.2.2; Verkürzungen auf Proteinebene durch neu eingeführte Proteolyse-Anfälligkeit: siehe 4.3.4; Erhöhung der *inclusion body*-Fluoreszenz: siehe: 4.3.5.2) berücksichtigt werden und solche Klone nach der FACS-Anreicherung von fluoreszierenden *E. coli* Zellen durch Kolonie-PCR (3.3.2), Sequenzierung (3.3.10), Probeexpression (3.4.1.1) und Fluoreszenz-Mikroskopie (3.2.10) ausgeschlossen werden.

4.3.5.6 Weitere biophysikalische Charakterisierung von hGR-LBD

Der Einfluss der im Rahmen dieser Arbeit identifizierten und kombinierten Mutationen A605V, V702A, E705G und M752T (4.3.2.2, 4.3.4), welche die lösliche Herstellung von hGR-LBD in *E. coli* (4.3.5.3, 4.3.5.4) und dessen konformationelle Stabilität erhöhen (4.3.5.5), auf die Löslichkeit *in vitro*, die Sekundär- bzw. Tertiärstruktur und den Oligomerisierungsgrad von hGR-LBD wird im Folgenden dargestellt. Da das wildtypische hGR-LBD Protein nicht aufgereinigt werden kann und somit für weitergehende Untersuchungen nicht zur Verfügung steht, diente die bekannte Variante hGR-LBD(F602S) (Bledsoe *et al.*, 2002) jeweils als Referenzprotein. Durch vergleichende Analyse mit hGR-LBD(F602S+A605V+V702A+E705G+M752T) konnte der Einfluss der neuen Austausche auf die oben genannten biophysikalischen Eigenschaften von hGR-LBD untersucht werden.

4.3.5.6.1 Analyse der Löslichkeit *in vitro*

Um die Löslichkeit von hGR-LBD näher zu charakterisieren wurden die Proteine in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 50 µM Dexamethason, pH 7,9 mit Ammoniumsulfat fraktionell präzipitiert (3.5.11) (Abbildung 59). Hierbei zeigte sich, dass die Kombination der Mutationen A605V, V702A, E705G und M752T zu einer erhöhten Löslichkeit *in vitro* führt. So fällt hGR-LBD(F602S+A605V+V702A+E705G+M752T) erst bei höheren Ammoniumsulfat-Konzentrationen aus als hGR-LBD(F602S). Um 50 % des Proteins zu präzipitieren ist unter den genannten Bedingungen bei hGR-LBD(F602S+A605V+V702A+E705G+M752T) eine Ammoniumsulfat-Konzentration von 0,98 M nötig, im Vergleich zu 0,77 M bei hGR-LBD(F602S). Die erhöhte Löslichkeit

von hGR-LBD in *E. coli* durch A605V, V702A, E705G und M752T (siehe 4.3.5.3) wird demnach nicht durch extrinsische Faktoren (z.B. verbesserte Interaktion mit Chaperonen), sondern durch eine intrinsisch verbesserte Hydratisierung des Moleküls bedingt.

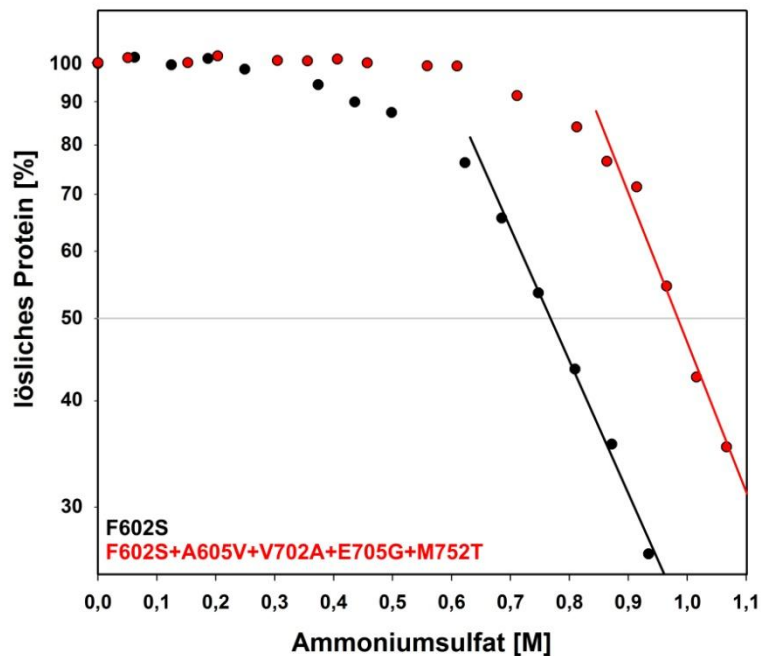


Abbildung 59: Löslichkeit von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) *in vitro*.

Die Proteine wurden jeweils mit einer Konzentration von 5 μM in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 50 μM Dexamethason, pH 7,9 bei RT fraktionell mit steigenden Konzentrationen an Ammoniumsulfat präzipitiert. Die Konzentration des im Überstand verbleibenden, löslichen Proteins wurde über Fluoreszenz-Spektroskopie anhand einer zuvor bestimmten Eichgerade (siehe Anhang 7.2) bestimmt. Bei jeder Probe wurde von einer 1:8 Verdünnung des löslichen Überstands (in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9) nach Anregung bei 280 nm die Emission bei 342 nm für 1 min gemessen und gemittelt (Schichtdicke der Küvette: 1 cm; Photomultiplier-Spannung: 845 V). Die für hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) ermittelten $[L_{1/2}]^{\text{app}}$ -Werte betragen 0,77 M bzw. 0,98 M.

Die so verbesserte Löslichkeit von hGR-LBD in wässrigem Puffer optimiert das Protein auch für weitere biochemische und biophysikalischen Untersuchungen, da es zum Einen höher einkonzentriert, und zum Anderen es auch bei höheren Salz-Konzentrationen noch in Lösung gehalten werden kann. Beides kann auch bei der Kristallisation des Proteins für die Röntgenstrukturanalyse von Vorteil sein.

4.3.5.6.2 Analyse des Oligomerisierungsgrades durch analytische Gelfiltration

Proteine mit bekanntem Molekulargewicht wurden wie in 3.5.5 beschrieben auf eine Superdex-75 Gelfiltrationssäule gegeben, wodurch eine Kalibrierungskurve erstellt werden konnte (siehe Anhang 7.5). Die anschließende Analyse von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) zeigte, dass beide Proteine ein identisches

Elutionsprofil aufweisen, dessen dominierender Hauptpeak einem Monomer entspricht (Abbildung 60a). Die Mutationen A605V, V702A, E705G und M752T haben somit keinen Einfluss auf den Oligomerisierungsgrad des Proteins.

Bei dem vor dem Hauptpeak eluierenden Protein handelt es sich vermutlich um unspezifische, lösliche Aggregate und nicht um definierte hGR-LBD Oligomere, da bei 10-facher Verdünnung der aufgetragenen Probe von 14 μM (Auftrag: 0,21 mg in 500 μl Pufferlösung) auf 1,4 μM (Auftrag: 0,02 mg in 500 μl Pufferlösung) der relative Anteil dieser Spezies in der Proteinlösung nicht geringer wurde (Abbildung 60b). Da die auf eine Superdex-75 geladenen Proben beim Gelfiltrationslauf ca. 1:10 verdünnt werden (Schwabe, 2007) kann somit zumindest in einem Konzentrationsbereich von ca. 0,14 - 1,4 μM eine definierte Oligomerisierung von hGR-LBD unter den genannten Bedingungen ausgeschlossen werden.

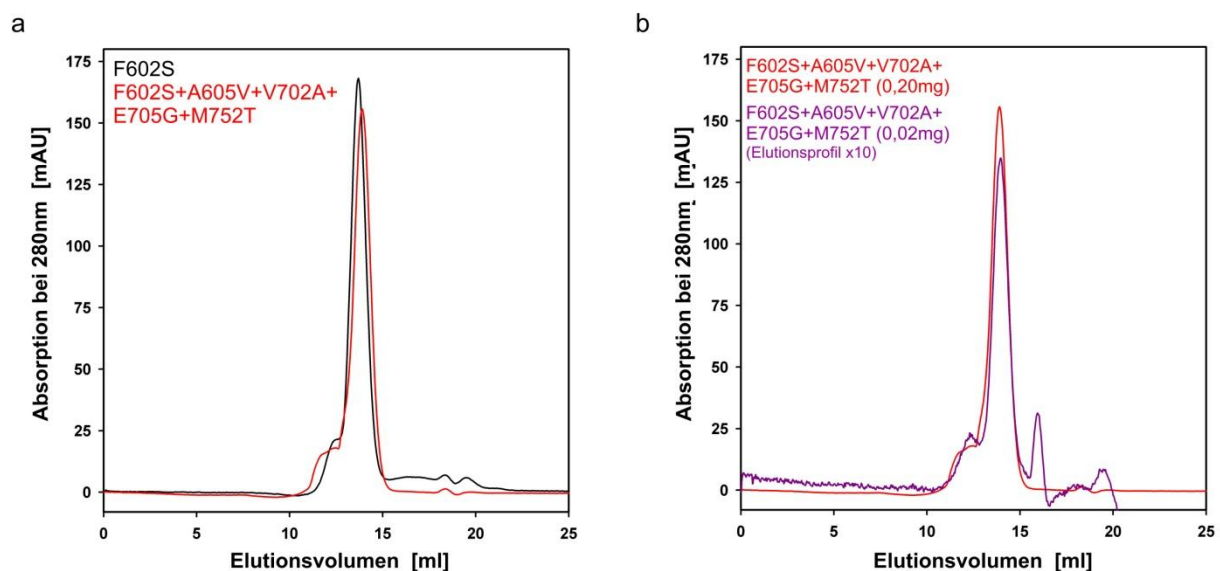


Abbildung 60: Analytische Gelfiltration von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T).

Die Elution erfolgte jeweils in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 20 μM Dexamethason, pH 7,9 bei RT und wurde bei 280 nm verfolgt.

a Es wurden je 0,21 mg von hGR-LBD(F602S) bzw. hGR-LBD(F602S+A605V+V702A+E705G+M752T) auf eine Superdex-75 Säule (Säulenvolumen = 24 ml) aufgetragen. Die Elutionsprofile von hGR-LBD(F602S) bzw. hGR-LBD(F602S+A605V+V702A+E705G+M752T) zeigen Hauptpeaks bei 13,7 bzw. 13,9 ml, was einem apparenten Molekulargewicht von 19,6 bzw. 17,9 kDa und damit am ehesten einem Monomer (MW(hGR-LBD: ~ 30 kDa)) entspricht. Die unmittelbar vor den Hauptpeaks eluierenden Nebenpeaks deuten auf einen für beide Proteine identisch geringen Anteil an löslichen Aggregaten hin.

b Auftrag von 0,21 bzw. 0,02 mg hGR-LBD(F602S+A605V+V702A+E705G+M752T) auf eine Superdex-75 Säule (Säulenvolumen = 24 ml). Um den relativen Anteil der enthaltenen Proteinspezies vergleichen zu können wurde das Elutionsprofil des 0,02 mg-Auftrages mit 10 multipliziert. Die resultierenden Elutionsprofile gleichen einander.

Analytische Ultrazentrifugation von hGR-LBD(A605V+M752T) und hGR-LBD(F602S+A605V+M752T) (in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 15 μM Dexamethason bzw. 15 μM Mifepriston, pH 7,9; durchgeführt von Eric

Kusznir und Dr. Francis Müller; Hoffmann-LaRoche, *Pharma Research Discovery*, Basel) zeigte ebenfalls, dass hGR-LBD in monomerer Form vorliegt (Daten nicht gezeigt).

Die hier vorgestellten Ergebnisse stehen in Widerspruch zu den veröffentlichten Daten von Bledsoe *et al.* (2002). Dort wurde mit Hilfe von analytischer Ultrazentrifugation gezeigt, dass hGR-LBD(F602S) in dem untersuchten Konzentrationsbereich dimerisiert ($K_D = 1,5 \mu\text{M}$). Allerdings handelte es sich bei dem dort untersuchten Protein um TIF2-gebundenes hGR-LBD(F602S). Insgesamt deuten somit die Ergebnisse dieser Arbeit darauf hin, dass die Dimerisierung von hGR-LBD(F602S) durch die Bindung des Coaktivator-Peptids TIF2 induziert werden könnte.

4.3.5.6.3 Spektroskopische Analyse

Um den Einfluss der Mutationen A605V, V702A, E705G und M752T auf die Sekundär- bzw. Tertiärstruktur von hGR-LBD zu untersuchen, wurden CD-Spektren im Fern-UV Bereich bzw. Fluoreszenzemissionsspektren von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) sowohl mit gebundenem Agonisten Dexamethason als auch mit gebundenem Antagonisten Mifepriston aufgenommen.

Die Fern-UV CD-Spektren aller untersuchten Proteine waren nahezu identisch und zeigten ein typisches Spektrum eines durch α -Helices dominierten Proteins mit Minima bei 208 bzw. 220 nm (Schmid, 1997) (Abbildung 61).

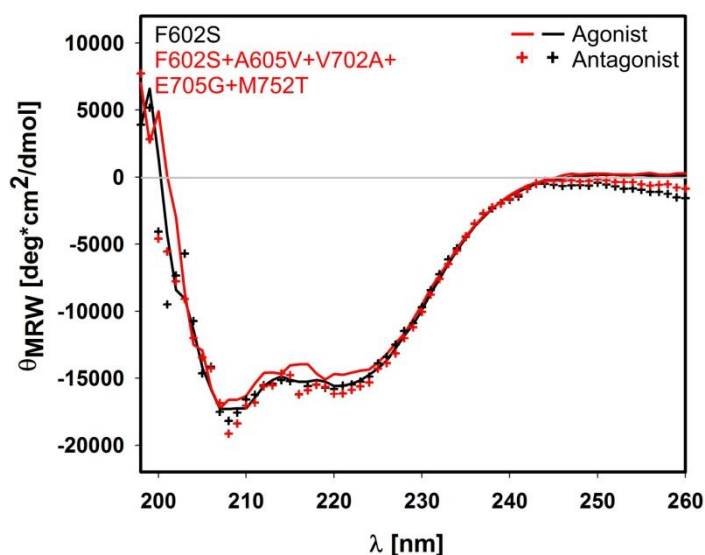


Abbildung 61: Fern-UV CD-Spektren von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T).

Die Spektren wurden in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9 bei RT aufgenommen. Die Schichtdicke der Küvette betrug 1 mm. Die eingesetzten Konzentrationen betragen 5 μM Protein + 50 μM Dexamethason (Agonist-Konformation) bzw. 2,5 μM Protein + 50 μM Mifepriston (Antagonist-Konformation). Die gezeigten Spektren sind jeweils das Mittel aus 10 übereinandergelegten Spektren.

Die mit Hilfe des Programms K2D (<http://www.embl-heidelberg.de/~andrade/k2d/>) aus dem Fern-UV CD-Spektrum von hGR-LBD(F602S) bestimmten Sekundärstruktur-Anteile des Proteins (α -Helices: 55 %, β -Stränge: 9 %, *random coil*: 36 %) decken sich sehr gut mit den Werten, welche aus der veröffentlichten Kristallstruktur von hGR-LBD(F602S) (Bledsoe *et al.*, 2002) abgeleitet werden können (α -Helices: 63 %, β -Stränge: 5 %, *random coil*: 32 %).

Zusammenfassend kann somit festgestellt werden, dass die Sekundärstruktur von hGR-LBD durch Fern-UV CD-Spektroskopie gut dargestellt werden kann und dass diese weder durch die eingeführten Mutationen noch durch eine Liganden-bedingte Konformationsänderung verändert wird.

Die Fluoreszenzemissionsspektren von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) zeigten ebenfalls keine Unterschiede, was darauf hindeutet, dass die neu eingeführten Mutationen A605V, V702A, E705G und M752T auch die globale Tertiärstruktur des Proteins nicht verändern (Abbildung 62).

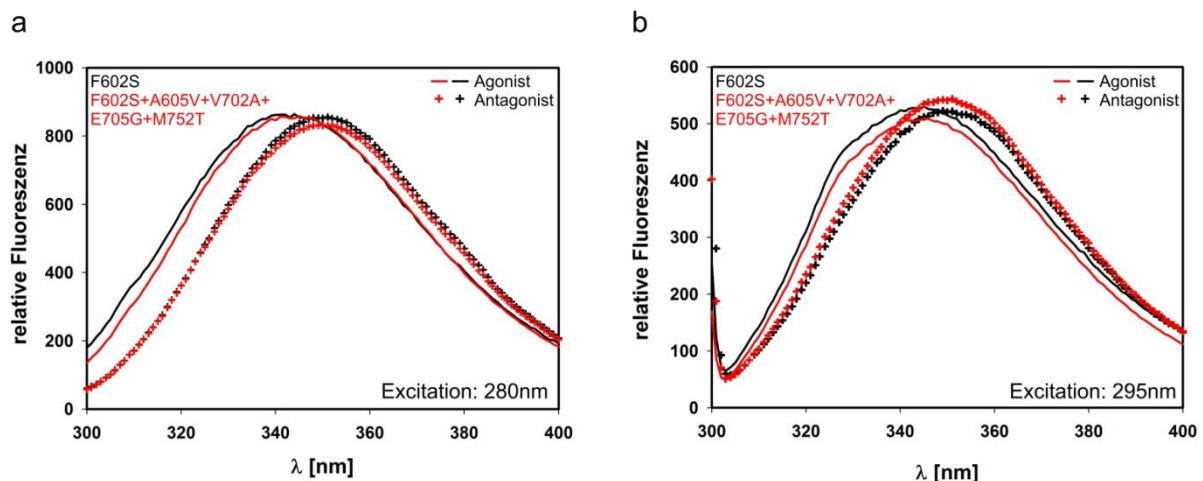


Abbildung 62: Fluoreszenzemissionsspektren von hGR-LBD(F602S) und hGR-LBD(F602S+A605V+V702A+E705G+M752T) nach Anregung bei 280 nm (a) bzw. 295 nm (b).

Die Spektren wurden in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9 bei RT aufgenommen. Die Schichtdicke der Küvette betrug 1 cm. Die eingesetzten Konzentrationen betragen 2,5 μ M Protein + 50 μ M Dexamethason (Agonist-Konformation) bzw. 2,5 μ M Protein + 50 μ M Mifepriston (Antagonist-Konformation) (Photomultiplier-Spannung: 750 V). Die gezeigten Spektren sind jeweils das Mittel aus 10 übereinandergelegten Spektren. Das Maximum der Fluoreszenzemission der Agonist-Konformation bzw. der Antagonist-Konformation liegt für beide Proteine bei ca. 345 nm bzw. 351 nm.

Die Fluoreszenzemissionsspektren der Agonist-Konformation und der Antagonist-Konformation unterscheiden sich allerdings deutlich (Abbildung 62). Beide Spektren zeigen ein Emissionsmaximum im langwelligen Bereich, was auf eine hohe Lösungsmittel-Zugänglichkeit der fünf enthaltenen Tryptophane hindeutet. Allerdings zeigen die Spektren der durch Mifepriston induzierten Antagonist-Konformation ein Emissionmaximum bei

höheren Wellenlängen (~ 351 nm) im Vergleich zu der durch Dexamethason induzierten Agonist-Konformation (~ 345 nm). Dieses Emissionsmaximum von über 350 nm deutet darauf hin, dass bei der Antagonist-Konformation des Proteins alle Tryptophane komplett Lösungsmittel-zugänglich sind. Die verschiedenen hGR-LBD Konformationen können somit durch Fluoreszenz-Spektroskopie voneinander unterschieden werden, was für weitergehende Untersuchungen (z.B. Analyse der Mifepriston-Bindung) ausgenutzt werden könnte.

4.3.5.7 Reinigung von Ligand-freiem hGR-LBD

Bei den bisher beschriebenen Reinigungen von hGR-LBD wurde sowohl bei der heterologen Expression des entsprechenden Gens in *E. coli* als auch bei den einzelnen Reinigungsschritten des Proteins jeweils der stabilisierende Ligand Dexamethason zugegeben. Erst beim finalen Reinigungsschritt (präparative Gelfiltration; siehe 3.4.2.4) wurde das Protein in einen Ligand-freien Puffer (25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9) überführt. Massenspektrometrische Untersuchungen dieser gereinigten hGR-LBD Präparationen (durchgeführt von Dr. Stefan Müller; Hoffmann-La Roche, *Pharma Research Discovery*, Basel) zeigten jedoch, dass darin immer noch ca. äquimolare Mengen an Dexamethason vorhanden waren (Daten nicht gezeigt). Um die vollständige Sättigung von hGR-LBD mit Mifepriston bei den Untersuchungen der Antagonist-Konformation des Proteins (thermische Auffaltungen: siehe 4.3.5.5; Spektroskopische Analyse: siehe 4.3.5.6.3) garantieren zu können, wurde deshalb stets ein hoher Überschuss an zugegebenem Mifepriston verwendet. Mifepriston bindet zudem stärker als Dexamethason an hGR-LBD (Jung-Testas & Baulieu, 1983), weshalb davon ausgegangen werden kann, dass bei dieser Vorgehensweise das Protein zu 100 % in der Antagonist-Konformation vorliegt, obwohl noch etwas Dexamethason in der Probenlösung vorhanden ist.

Um das Protein effektiv mit anderen, evtl. schwächer bindenden Liganden sättigen und um die Bindung von Dexamethason quantifizieren zu können, würde man allerdings eine komplett Ligand-freie hGR-LBD Präparation benötigen. Um eine solche Präparation von hGR-LBD zu erhalten, sollte die Expression und Reinigung des Proteins in vollständiger Abwesenheit von Dexamethason durchgeführt werden. Die bisher in der Literatur beschriebene reinigbare hGR-LBD Variante (hGR-LBD(F602S)) kann ohne Zugabe von Dexamethason nicht löslich in *E. coli* hergestellt werden (Bledsoe *et al.*, 2002). Um zu überprüfen, ob dies möglicherweise mit den in dieser Arbeit erzeugten stabilisierten hGR-LBD Einzel- und Kombinationsvarianten (Tabelle 14) möglich ist, wurden diese Proteine im analytischen Maßstab in *E. coli* in Abwesenheit von Dexamethason hergestellt (3.4.1.1). Die

optimierten hGR-LBD Varianten zeigten ohne Zugabe von Dexamethason eine deutlich erniedrigte Löslichkeit in *E. coli*, wobei die hGR-LBD Variante mit allen fünf stabilisierenden Mutationen (F602S+A605V+V702A+E705G+M752T) am stärksten in der löslichen Zellfraktion zu Finden war (Daten nicht gezeigt).

Die Ausbeute der Reinigung von hGR-LBD(F602S+A605V+V702A+E705G+M752T) in Abwesenheit von Dexamethason im präparativen Maßstab betrug dennoch lediglich 0,007 mg Protein pro Gramm Nassgewicht der *E. coli* – Expressionskultur (~ 0,03 mg Protein / Liter Expressionskultur). Es wurde somit eine 755-fach erniedrigte Ausbeute im Vergleich zur Reinigung desselben Proteins bei Zugabe von Dexamethason (Ausbeute: 5,29 mg pro Gramm Nassgewicht der *E. coli* – Expressionskultur) erhalten. Das Ligand-frei präparierte Protein wurde hinsichtlich seiner Sekundär- bzw. Tertiärstruktur und seiner konformationellen Stabilität untersucht (Abbildung 63).

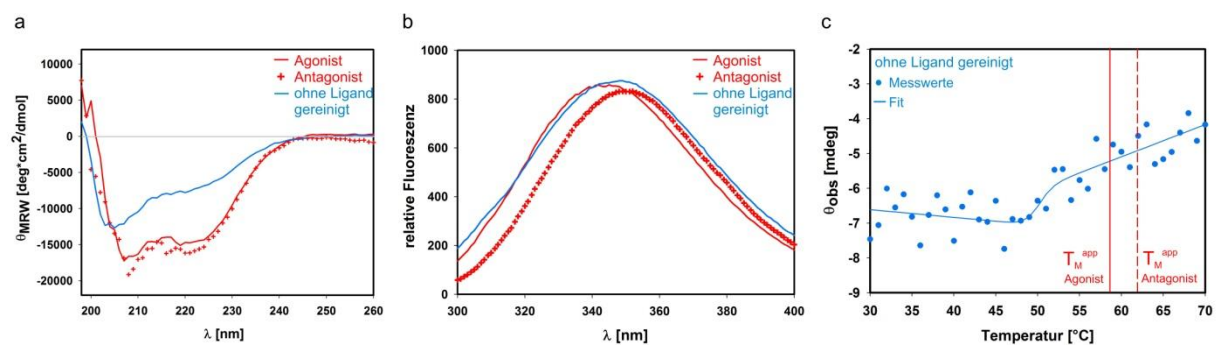


Abbildung 63: Charakterisierung von Ligand-freiem hGR-LBD(F602S+A605V+V702A+E705G+M752T).
a Fern-UV CD-Spektrum des Ligand-freien Proteins (blau) im Vergleich zu den Spektren der Agonist- bzw. Antagonist-Konformation des Proteins (rot) (siehe Abbildung 61). Das Spektrum wurde in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9 bei RT aufgenommen, wobei das Mittel aus 60 übereinandergelegten Spektren gezeigt ist. Die Schichtdicke der Küvette betrug 1 mm. Die eingesetzte Proteinkonzentration betrug 5 μ M.
b Fluoreszenzemissionsspektrum des Ligand-freien Proteins (blau) im Vergleich zu den Spektren der Agonist- bzw. Antagonist-Konformation des Proteins (rot) (siehe Abbildung 62a) nach Anregung bei 280 nm. Das Spektrum wurde in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, pH 7,9 bei RT aufgenommen, wobei das Mittel aus 10 übereinandergelegten Spektren gezeigt ist. Die Schichtdicke der Küvette betrug 1 cm. Die eingesetzte Proteinkonzentration betrug 2,5 μ M (Photomultiplier-Spannung: 750 V). Das Maximum der Fluoreszenzemission liegt bei 348 nm (Agonist-Konformation: 345 nm; Antagonist-Konformation: 352 nm).
c Die thermische Auffaltung des Ligand-freien Proteins erfolgte mit einer Rate von 1 °C/min von 30 °C bis 70 °C in 25 mM Tris/Cl, 0,1 M NaCl, 2 mM TCEP, 10 % Glycerin (pH 7,9 bei RT; pH 7,8 - 6,6 bei 30 - 70 °C, da $\Delta pK_a/^\circ C(\text{Tris}) = -0,028$, Mandaji *et al.*, 2009) und wurde durch CD Spektroskopie bei 220 nm verfolgt. Der T_M^{app} -Wert der Agonist- bzw. Antagonist-Konformation (58,6 °C bzw. 61,9 °C; Tabelle 15) des Proteins ist durch eine durchgezogene bzw. gestrichelte rote Linie markiert.

Das Fern-UV CD-Spektrum unterscheidet sich deutlich vom Spektrum der Agonist- bzw. Antagonist-Konformation des Proteins (Abbildung 63a). Die Analyse der Sekundärstruktur-Anteile mit Hilfe von K2D (<http://www.embl-heidelberg.de/~andrade/k2d/>) deutet darauf hin, dass weniger α -helikale Bereiche und mehr β -Stränge bzw. *random coil*-Bereiche als bei in

Anwesenheit von Liganden gereinigtem Protein vorhanden sind (α -Helices: 22 % vs. 55 %, β -Stränge: 22 % vs. 9 %, *random coil*: 56 % vs. 36 %) (vgl. 4.3.5.6.3). Das Fluoreszenzemissionspektrum nach Anregung mit 280 nm zeigte ein Maximum bei 348 nm, was zwischen dem Maximum der Agonist- (345 nm) und der Antagonist-Konformation (351 nm) liegt (Abbildung 63b). Die thermische Auffaltung des Ligand-freien Proteins verfolgt durch das CD-Signal bei 220 nm zeigt im Gegensatz zu den thermischen Auffaltungen der Agonist- bzw. Antagonist-Konformation des Proteins (siehe 4.3.5.5, Abbildung 57) einen wenig kooperativen Verlauf, was die Bestimmung eines T_M^{app} -Wertes verhinderte und gegen eine kompakte Struktur spricht (Abbildung 63c).

Zusammenfassend kann somit festgestellt werden, dass hGR-LBD in Abwesenheit eines Liganden offensichtlich keine definierte Struktur ausbilden kann. Das Vorhandensein von etwas Sekundärstruktur und vermutlich wenig Tertiärstruktur deutet darauf hin, dass es sich hierbei vielmehr um ein *Molten Globule*-artiges Molekül handelt.

Bereits in früheren Arbeiten wurde beschrieben, dass der Glucocorticoid-Rezeptor nur assoziiert an einen spezifischen Chaperon-Komplex (Hsp90 und andere Co-Chaperone: siehe 1.7, Abbildung 7) fähig ist Liganden zu binden (Bresnick *et al.*, 1989; Dittmar & Pratt, 1997; Dittmar *et al.*, 1997), was darauf hindeutet, dass das Apoprotein sich nicht korrekt falten kann. Die Präparation eines isolierten, korrekt gefalteten hGR-LBD Proteins erscheint somit nicht möglich. Alternativ zur Reinigung von hGR-LBD ohne Ligand, könnte allerdings versucht werden den Hsp90-stabilisierten hGR-LBD - Hetero-Komplex heterolog aus *E. coli* aufzureinigen und zu rekonstituieren. Ein so präparierter Ligand-Bindungs-fähiger hGR-LBD - Hetero-Komplex würde die Situation *in vivo* widerspiegeln und Bindungsstudien mit verschiedenen Liganden *in vitro* ermöglichen.

4.3.5.8 Strukturelle Analyse der vorteilhaften Mutationen in hGR-LBD

4.3.5.8.1 A605V in hGR-LBD

A605V wurde in der am stärksten fluoreszierenden hGR-LBD-eGFP Variante A8 identifiziert (4.3.2.2; Tabelle 12) und induziert isoliert in hGR-LBD-eGFP ebenfalls eine deutlich erhöhte Fluoreszenz *in vivo* (4.3.5.2). Dieser Aminosäureaustausch führt von allen in dieser Arbeit gefundenen positiven Mutationen in hGR-LBD zur stärksten Erhöhung der Löslichkeit (4.3.5.3) und Stabilität (4.3.5.5).

Der veröffentlichten Röntgenkristallstruktur von hGR-LBD(F602S) kann entnommen werden, dass Alanin 605 neben 21 anderen Aminosäuren an der Bindung von Dexamethason beteiligt ist (Bledsoe *et al.*, 2002) (Abbildung 64).

Die Seitenkette von A605 bildet van-der-Waals Wechselwirkungen mit dem Steroid-Grundgerüst des Liganden aus. Die Superpositionierung von hGR-LBD und der Ligand-Bindungsdomäne des humanen Progesteron-Rezeptors (hPR-LBD) zeigt, dass die Ligand-Bindungstasche der beiden verwandten Proteine (Sequenz-Identität: 53,1 %) stark konserviert ist: 17 der 22 Aminosäuren sind identisch (Abbildung 64). Allerdings liegt bei hPR-LBD an der entsprechenden Position zu A605 von hGR-LBD ein Valin (V760) vor.

Auch bei anderen humanen Nuklearhormon-Rezeptoren (siehe multiples Sequenzalignment: Anhang 7.6; Abbildung A4) befindet sich an dieser Position oftmals eine größere hydrophobe Aminosäure (z.B.: Androgen-Rezeptor α : Valin; Estrogen-Rezeptor α : Methionin). Da die Aminosäure an dieser Position über unspezifische van-der-Waals Interaktionen mit dem Liganden interagiert, kann davon ausgegangen werden, dass eine größere hydrophobe Aminosäure die Bindung stärkt. So wurde auch in der Arbeit von Ricketson *et al.* (2007), in welcher Varianten von GR aus der Ratte mit erhöhter Dexamethason-Empfindlichkeit in *S. cerevisiae* evolviert wurden, der zu A605V äquivalente Austausch A623V als vorteilhafte Mutation identifiziert.

Die in dieser Arbeit durch gelenkte Evolution identifizierte Mutation A605V in hGR-LBD scheint demnach die Protein-Ligand Wechselwirkung von homologen Rezeptoren an dieser Position zu imitieren, wodurch es zu einer Verstärkung der Bindung kommt. Da Dexamethason und Mifepriston in derselben Bindungstasche von hGR-LBD gebunden werden (Kauppi *et al.*, 2003) und beide ein Steroid-Grundgerüst besitzen, ist es nicht verwunderlich, dass sowohl die Agonist- als auch die Antagonist-Konformation von hGR-LBD durch A605V stabilisiert wird (siehe 4.3.5.5). Weshalb sich beim wildtypischen hGR-LBD an dieser Position ein Alanin befindet ist unklar, wobei möglicherweise die schwächere Ligand-Bindung eine Rolle bei der intramolekularen Kommunikation zwischen der Ligand-Bindungstasche und anderen funktionellen Teilen des Glucocorticoid-Rezeptors (AF-1, AF-2, siehe Abbildung 8) spielen könnte.

Größere hydrophobe Aminosäuren als Valin an Position 605 könnten zu einer noch stärkeren Interaktion mit dem Liganden und damit zu einer weiteren Stabilisierung von hGR-LBD führen. Die Modellierung verschiedener hydrophober Aminosäuren an Position 605 der veröffentlichten hGR-LBD(F602S) Struktur (pdb 1m2z; Bledsoe *et al.*, 2002) *in silico* zeigte jedoch, dass lediglich Valin gut positioniert werden kann. Bei größeren Resten kommt es bei den Modellierungen zu einer Überschneidung der van-der-Waals Radien der Seitenketten und des Liganden, was zu einer Abstoßung des Liganden führen würde (Abbildung 65a). Um diese Hypothese zu testen, wurden hGR-LBD(F602S)-Varianten hergestellt, bei denen

Alanin 605 zu Leucin, Isoleucin, Phenylalanin oder Methionin mutiert vorliegt. Die entsprechenden Nukleotid-Austausche wurden über Megaprimer-PCR (3.3.3.1) einzeln in *hgr-lbd*(F602S) eingeführt. Der Megaprimer wurde mit einem mutagenen Primer (3′*hgr-lbd* (F602S+A605L), 3′*hgr-lbd* (F602S+A605I) 3′*hgr-lbd* (F602S+A605F) 3′*hgr-lbd* (F602S+A605M)) und pERN7up mittels PCR (3.3.1) amplifiziert, wobei pERN7_ *hgr-lbd*(F602S) als Templat diente. Das Vollängen-Gen wurde anschließend jeweils durch den aus einem Agarosegel aufgereinigten (3.3.5) Megaprimer und 3′*hgr-lbd* (*NotI*/Stop) mit Hilfe desselben Templats wie zuvor durch PCR amplifiziert. Mittels der Restriktionsenzyme *NdeI* und *NotI* konnten die erhaltenen mutagenisierten Gene wieder in pERN7 kloniert werden. Die so hergestellten pERN7_ *hgr-lbd* Konstrukte wurden für Probeexpressionen in *E. coli* BL21(DE3) bei 20 °C in Anwesenheit von Dexamethason benutzt (3.4.1.1). Zur Kontrolle wurden *hgr-lbd*(F602S) und *hgr-lbd*(F602S+A605V) mit Hilfe der entsprechenden pERN7-Konstrukte parallel exprimiert. Die Analyse der über Nickel-Sepharose gereinigten löslichen Zellfraktion der Expressionskulturen machte deutlich, dass die größeren hydrophoben Aminosäuren an Position 605 zu einer verringerten Löslichkeit des Proteins führen (Abbildung 65b).

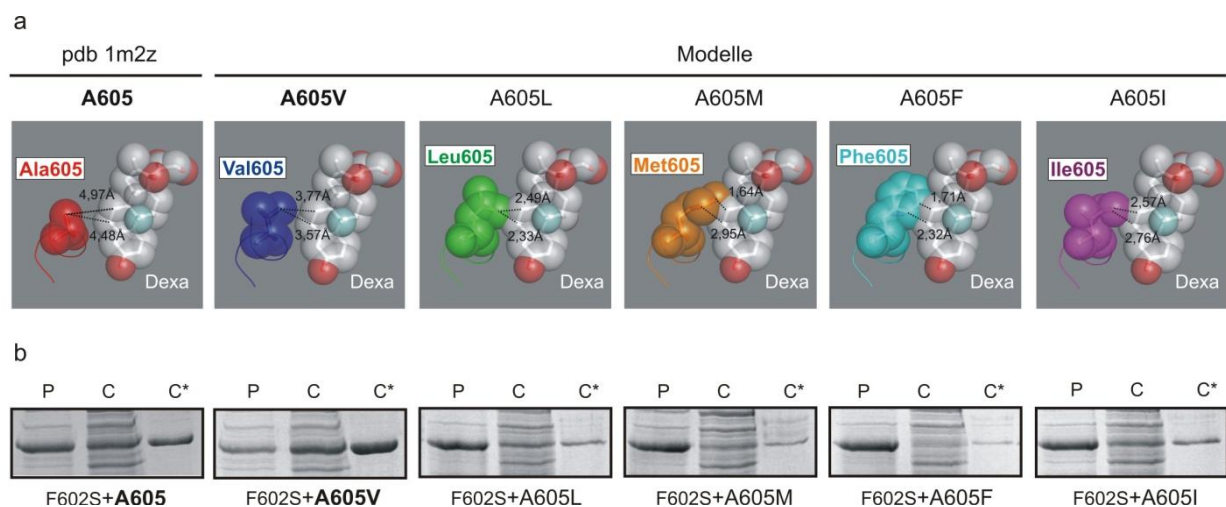


Abbildung 65: Analyse von hGR-LBD(F602S) mit unterschiedlichen Mutationen von A605.

a Die Mutagenese von A605 in hGR-LBD(F602S) (pdb 1m2z; Bledsoe *et al.*, 2002) wurde *in silico* mit Hilfe von PyMol (<http://www.pymol.org/>) durchgeführt. Es wurde jeweils das Rotamer der Aminosäuren-Seitenkette modelliert, welches in Richtung des gebundenen Liganden zeigt. Im Gegensatz zu Leucin, Methionin, Phenylalanin und Isoleucin, deren van-der-Waals Radien mit dem von Dexamethason kollidieren, kann Valin in einem sehr günstigen Abstand positioniert werden.

b Analytische Herstellung His₆-GST-fusionierter hGR-LBD(F602S) Varianten mit mutagenisiertem A605 in *E. coli*, dokumentiert durch SDS-PAGE (12,5 %). P: unlösliche Zellfraktion (Pellet), C: lösliche Zellfraktion (*crude extract*), C*: über Nickel-Sepharose gereinigte lösliche Zellfraktion. Die Expression erfolgte jeweils in *E. coli* BL21(DE3), üN bei 20 °C nach Zugabe von 500 µM Dexamethason. Es ist lediglich der 45-66 kDa Bereich des SDS-PA Gels gezeigt, in welchem sich His₆-GST-hGR-LBD (Molekulargewicht: ~58 kDa) befindet. Im Gegensatz zu der in 4.3.2.2 / 4.3.4 identifizierten positiven Mutation A605V führt A605L, A605M, A605F und A605I zu einer verringerten löslichen Herstellung des Proteins.

Der im FACS-*Screening* isolierte Aminosäureaustausch A605V scheint somit die optimale Mutation an dieser Stelle zu sein. Das neu eingeführte Valin wird vermutlich exakt in eine zwischen Protein und Ligand existierende Lücke positioniert, wodurch es zu einer erhöhten Stabilität des Ligand-gebundenen hGR-LBD Proteins kommt.

4.3.5.8.2 Röntgenkristallstrukturanalyse von mGR-LBD(F608S), mGR-LBD(F608S+A611V) und mGR-LBD(A611V+V708A+E711G)

Die in dieser Arbeit identifizierten vorteilhaften Mutationen in hGR-LBD sollten dazu benutzt werden, um die Löslichkeit und die Stabilität des Proteins so stark zu erhöhen, dass die Strukturanalyse durch Röntgenkristallographie vereinfacht bzw. optimiert werden kann.

Da sich GR-LBD aus der Maus (mGR-LBD) etwas besser kristallisieren lässt als das humane Protein (Dr. Ralf Thoma, mündliche Mitteilung), wurde mGR-LBD (Sequenz-Identität zu hGR-LBD: 95,3 %, Alignment mit hGR-LBD: siehe Anhang 7.6, Abbildung A4) für die Strukturaufklärung benutzt, wobei die Aminosäuren-Kennzeichnung im Vergleich zu hGR-LBD um sechs Positionen verschoben ist (F602S, A605V, V702A, E705G, M752T in hGR-LBD entspricht F608S, A611V, V708A, E711G, M758T in mGR-LBD). Die im Folgenden gezeigten röntgenkristallographischen Untersuchungen wurden von Dr. Ralf Thoma, Brigitte D'Arcy, Dr. Jörg Benz, Martine Stihle und Dr. Guillaume Schoch (F. Hoffmann – LaRoche, *Pharma Research Discovery*, Basel) durchgeführt.

Die vergleichende Analyse von mGR-LBD(F608S), mGR-LBD(F608S+A611V) und mGR-LBD(A611V+V708A+E711G) in Anwesenheit von Dexamethason und TIF2 Coaktivator-Peptid (Sequenz: E/A-NALLRYLLD) zeigte, dass die Varianten mit den in dieser Arbeit identifizierten Mutationen tatsächlich deutlich besser streuende Kristalle lieferten. So konnte die Struktur von mGR-LBD(F608S+A611V) bzw. mGR-LBD(A611V+V708A+E711G) mit einer Auflösung von 1,6 Å bzw. 1,5 Å aufgeklärt werden, während bei mGR-LBD(F608S) nur eine Auflösung von 1,9 Å erhalten wurde (Abbildung 66a). Die Strukturen von mGR-LBD(F608S+A611V) und mGR-LBD(A611V+V708A+E711G) besitzen damit auch die bis dato beste Auflösung von GR-LBD Strukturen überhaupt (pdb 1m2z: 2,5 Å, pdb 1nhz: 2,3 Å, pdb 3cld: 2,84 Å, pdb 3bqd: 2,5 Å, pdb 3e7c: 2,15 Å), wodurch eine detailliertere Analyse der Struktur des Proteins möglich wird.

Die vierte identifizierte vorteilhafte Mutation aus dieser Arbeit M758T (= M752T in hGR-LBD) wurde nicht in die zur Strukturaufklärung benutzten Proteine mit eingeführt, da diese Mutation die Bindung des Coaktivator-Peptids TIF2 verschlechtert (siehe 4.3.5.8.4), was für die Kristallisation von Nachteil ist.

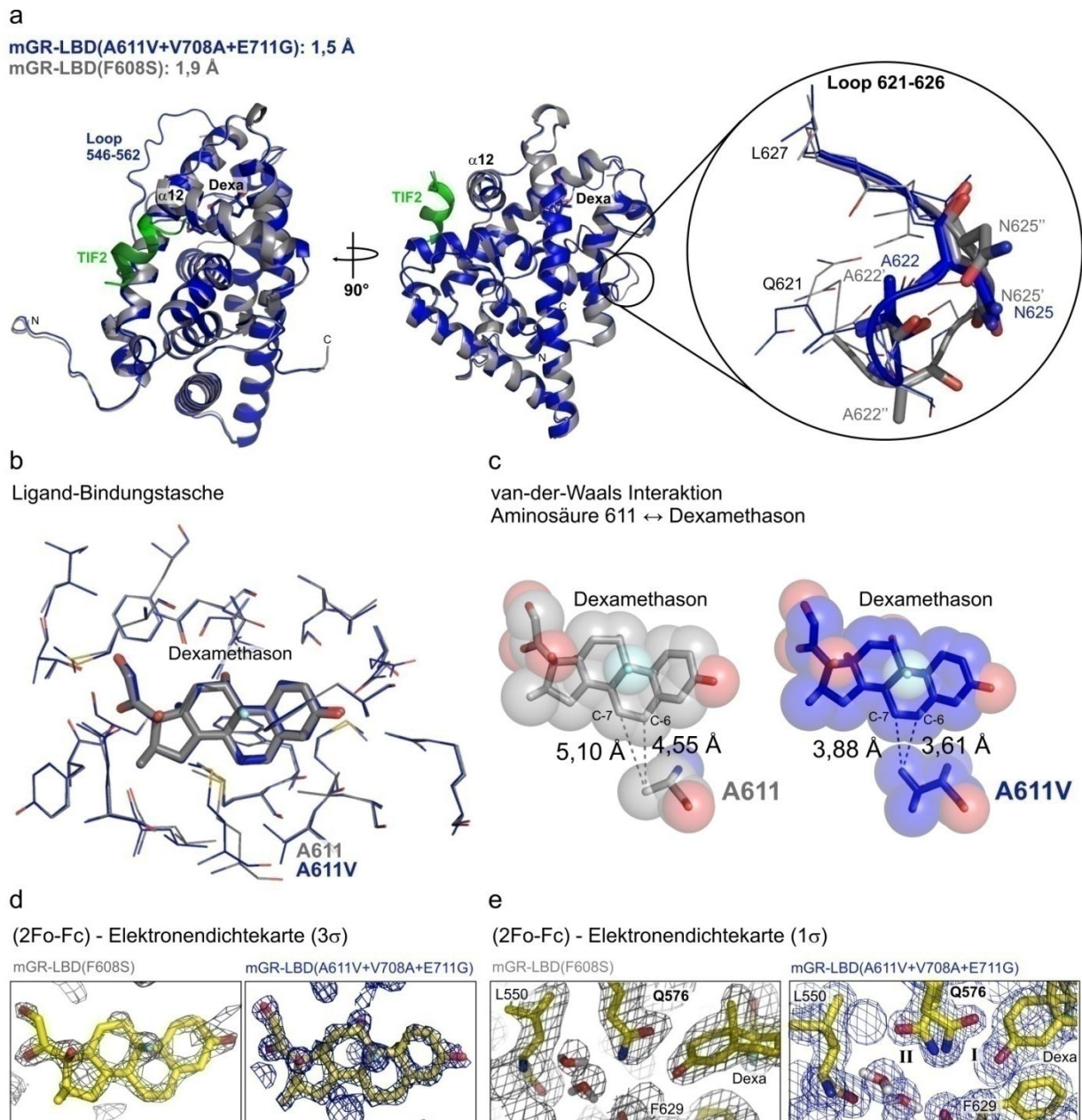


Abbildung 66: Röntgenkristallstruktur von mGR-LBD(A611V+V708A+E711G) bzw. mGR-LBD(F608S) mit gebundenem Dexamethason und TIF2 Coaktivator-Peptid bei einer Auflösung von 1,5 bzw. 1,9 Å.

a Die Strukturen von mGR-LBD(F608S) und mGR-LBD(A611V+V708A+E711G) konnten mit einem RMSD von 0,32 Å (258 C α -Atome mit Hilfe von DeepView verglichen) superpositioniert werden. Loop 546-562 ist in der mGR-LBD(A611V+V708A+E711G)-Struktur im Gegensatz zur mGR-LBD(F608S)-Struktur komplett aufgelöst. Loop 621-626 zeigt im Kristall von mGR-LBD(F608S) zwei Konformationen (grau und blau; exemplarisch sind A622 und N625 in den beiden unterschiedlichen Konformationen gezeigt und mit ' bzw. '' markiert), während bei mGR-LBD(A611V+V708A+E711G) nur eine Konformation (blau) vorhanden ist.

b Die Ligand-Bindungstaschen von mGR-LBD(F608S) (grau) und mGR-LBD(A611V+V708A+E711G) (blau) (22 Aminosäuren analog zu hGR-LBD: siehe Abbildung 64) zeigen bei Superpositionierung durch PyMol (<http://www.pymol.org/>) einen RMSD von nur 0,17 Å (162 Atome verglichen).

c Der stabilisierende Effekt von A611V kann durch optimierte van-der-Waals Wechselwirkungen der Seitenkette mit den C-6 und C-7 Atomen des gebundenen Steroid-Liganden erklärt werden. Gezeigt sind die Abstände zwischen C-6 bzw. C-7 des Liganden und dem terminalen C-Atom der A611(V)-Seitenkette.

d Optimierte Elektronendichte der stabilisierten mGR-LBD(A611V+V708A+E711G) – Struktur am Beispiel des gebundenen Liganden Dexamethason (gezeigt ist ein Konturniveau von 3 σ).

e Die höher aufgelöste Struktur von mGR-LBD(A611V+V708A+E711G) zeigt, dass Q576 zwei Konformationen annehmen kann (Häufigkeiten: I=70 %, II=30 %) (gezeigt ist ein Konturniveau von 1 σ).

Ein Vergleich der Strukturen zeigt, dass Loop 546-562, welcher in der mGR-LBD(F608S)-Struktur nicht komplett aufgelöst ist, durch die optimierte Elektronendichte von mGR-LBD(A611V+V708A+E711G) vollständig dargestellt werden kann (Abbildung 66a). Der Vergleich von mGR-LBD(F608S) mit mGR-LBD(F608S+A611V) bzw. mGR-LBD(A611V+V708A+E711G) zeigt zudem, dass bei den Strukturen der stabilisierten Varianten die Seitenkette von Cystein 628 (= C622 in hGR-LBD) (nicht gezeigt) und der gesamte Loop 621-626 (= Loop 615-620 in hGR-LBD) nur eine der zwei in mGR-LBD(F608S) beobachteten Konformationen aufweist (Abbildung 66a). A611V führt demnach auch zur Stabilisierung von weiter entfernten Bereichen des Proteins. Der Vergleich der mGR-LBD Strukturen mit der kürzlich veröffentlichten Kristallstruktur des intakten und DNA-gebundenen hRXR α -PPAR γ Nuklearhormonrezeptor-Komplexes (pdb 3dzy; Chandra *et al.*, 2008) legt nahe, dass Loop 621-626 möglicherweise eine wichtige Rolle bei der Interaktion zwischen der Ligand-Bindungsdomäne und der DNA-Bindungsdomäne des Glucocorticoid-Rezeptors spielen könnte (siehe Anhang 7.7; Abbildung A5). Weitergehende Untersuchungen *in vivo* bezüglich der Interaktion zwischen DNA-Bindungsdomäne und Ligand-Bindungsdomäne des wildtypischen und des stabilisierten mGR könnten zeigen, ob die vorhandene Flexibilität von Loop 621-626 funktionell von Vor- oder Nachteil für das Protein ist.

Die Analyse der mGR-LBD Strukturen mit A611V unterstützen desweiteren die in 4.3.5.8.1 getroffene Vermutung, dass A611V (= A605V in hGR-LBD) das Protein durch eine optimierte van-der-Waals Interaktion mit dem Liganden stabilisiert (Abbildung 66c). So positioniert sich das neu eingeführte Valin tatsächlich perfekt in die Lücke zwischen Proteinrückgrat und gebundenem Dexamethason. Hierbei wird die restliche Ligand-Bindungstasche des Proteins nicht verändert (Abbildung 66b), weshalb sich diese stabilisierte mGR-LBD Variante weiterhin für die Ko-Kristallisierung mit anderen potentiellen Liganden (z.B. neu identifizierte Wirkstoff-Leitstrukturen mit pharmazeutischer Relevanz) eignen würde, obwohl eine Ligand-bindende Aminosäure mutiert ist.

Die optimierte Elektronendichte der mGR-LBD Strukturen mit A611V bzw. A611V+V708A+E711A macht weitere atomare Details sichtbar. So kann zum Beispiel der Ligand Dexamethason deutlich besser beschrieben werden (Abbildung 66d). Desweiteren wird aufgrund der optimierten Elektronendichte eine zweite Konformation des an der Ligand-Bindung beteiligten Glutamins 576 (= Q570 in hGR-LBD) erkennbar (Häufigkeit: 30 %; Abbildung 66e), wobei diese bereits bei den veröffentlichten Strukturen von hGR-LBD mit gebundenem Deacylcortivazol bzw. mit einem gebundenem Nicht-Steroid-Agonisten

(pdb 3bqd bzw. pdb 3e7c) beschrieben worden ist. Die bevorzugte Konformation von Q576 scheint somit vom Liganden abhängig zu sein.

Mit der hoch aufgelösten Struktur von mGR-LBD(A611V+V708A+E711G) ist zum ersten Mal eine Struktur der Agonist-Konformation von GR-LBD ohne eine Mutation von F602 / 608 (hGR-LBD / mGR-LBD) zu Serin (Bledsoe *et al.*, 2002; Suino-Powell *et al.*, 2008) oder Tyrosin (Biggadike *et al.*, 2008; Madauss *et al.*, 2008) verfügbar. Durch Vergleich der Strukturen von mGR-LBD(A611V+V708A+E711G) und mGR-LBD(F608S) kann der stabilisierende Effekt der F608S Mutation erklärt werden. Global betrachtet führt F608S zu keinen größeren Veränderungen innerhalb von mGR-LBD. Allerdings können aufgrund des neu eingeführten hydrophilen Serin-Restes zusätzliche Wassermoleküle in das Protein eindringen, wodurch das benachbarte Serin 679 seine Konformation ändert und ein stabilisierendes Wasserstoffbrückennetzwerk innerhalb des Proteins etabliert wird (Abbildung 67). Von Kauppi *et al.* (2003) wurde eine ähnliche Wirkungsweise von F602S in der Antagonist-Konformation von hGR-LBD beobachtet.

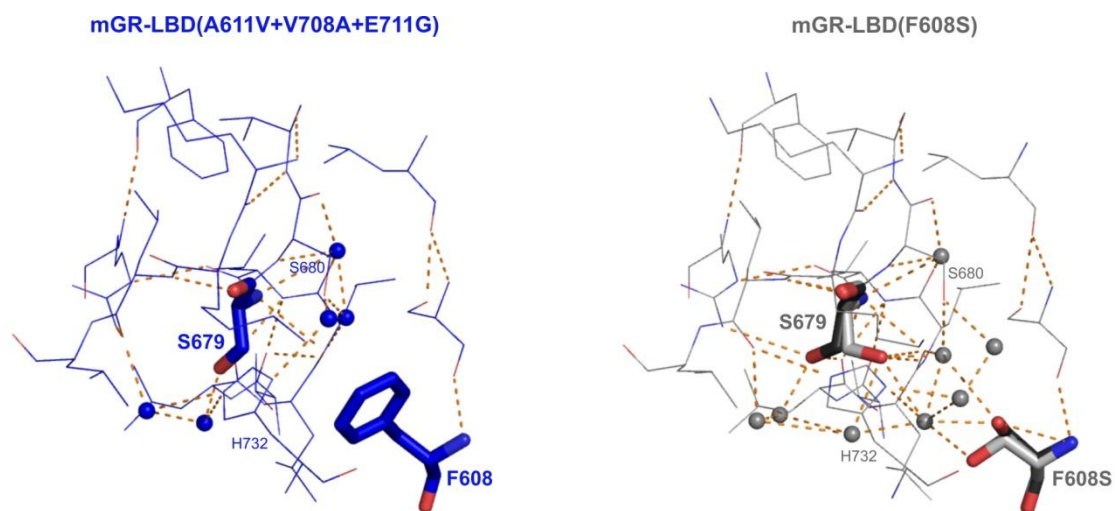


Abbildung 67: Etablierung eines Wasserstoffbrückennetzwerkes in mGR-LBD durch F608S.

Gezeigt sind jeweils alle Atome in mGR-LBD(A611V+V708A+E711G) bzw. mGR-LBD(F608S), welche weniger als 6 Å von S679 entfernt sind. Wasser-Moleküle sind als Kugeln dargestellt. Wasserstoffbrücken-Bindungen im gezeigten Bereich sind als orange gestrichelte Linien dargestellt. Bis auf S679 bzw. Aminosäure 608 sind die Konformationen aller gezeigten Aminosäureseitenketten in beiden Strukturen identisch.

Die mGR-LBD(A611V+V708A+E711G)-Struktur zeigt die wildtypische Situation an dieser Stelle. In mGR-LBD(F608S) wird durch F608S und S679 (zeigen jeweils zwei alternative Konformationen im Verhältnis 1:1), S680, H732 und mehrere Wasser-Moleküle ein internes Wasserstoffbrückennetzwerk ausgebildet.

Zusammenfassend kann somit festgestellt werden, dass die in dieser Arbeit identifizierten Aminosäureaustausche in GR-LBD eine große Hilfe bei röntgenkristallographischen Untersuchungen des Proteins darstellen. Die neu identifizierten stabilisierenden Mutationen können offensichtlich in verschiedenen orthologen GR-LBD Proteinen ihre Wirkung entfalten, wodurch deren Strukturaufklärung vereinfacht wird. Darüber hinaus ermöglicht die

Variation der verschiedenen additiv wirkenden Mutationen die Herstellung einer großen Bandbreite an stabilen GR-LBD-Varianten, welche zur Ko-Kristallisation mit verschiedenen Liganden verwendet werden können. Dadurch kann die Strukturaufklärung von GR-LBD mit potentiell pharmazeutisch relevanten Kleinmolekülen erleichtert werden.

4.3.5.8.3 V702A (V708A) und E705G (E711G) in hGR-LBD (mGR-LBD)

Die Austausch V702A bzw. E705G wurde in der stark fluoreszierenden hGR-LBD-eGFP Varianten A5-3 bzw. A5-2 identifiziert (4.3.2.2; Tabelle 12). Beide Mutationen induzieren isoliert in hGR-LBD-eGFP eine deutlich erhöhte Fluoreszenz *in vivo* (4.3.5.2) und verbessern die Löslichkeit (4.3.5.3) und die Stabilität (4.3.5.5) von hGR-LBD einzeln und additiv.

Die Reste V702 (V708 in mGR-LBD) und E705 (E711 in mGR-LBD) sind strukturell benachbart und liegen weit entfernt von jeglicher funktionellen Stelle des Proteins in einem Bereich, welcher vermutlich bei der Agonist- und der Antagonist-Konformation des Proteins identisch strukturiert ist. Somit ist es nicht verwunderlich, dass V702A und E705G die unterschiedlichen Konformationen von hGR-LBD gleichermaßen stabilisieren (4.3.5.5).

Ein Vergleich der mGR-LBD(F608S+A611V) und der mGR-LBD(A611V+V708A+E711G) - Struktur zeigt, dass die Mutationen V708A und E711G die lokale Konformation des Proteins nicht merklich beeinflussen (Abbildung 68a).

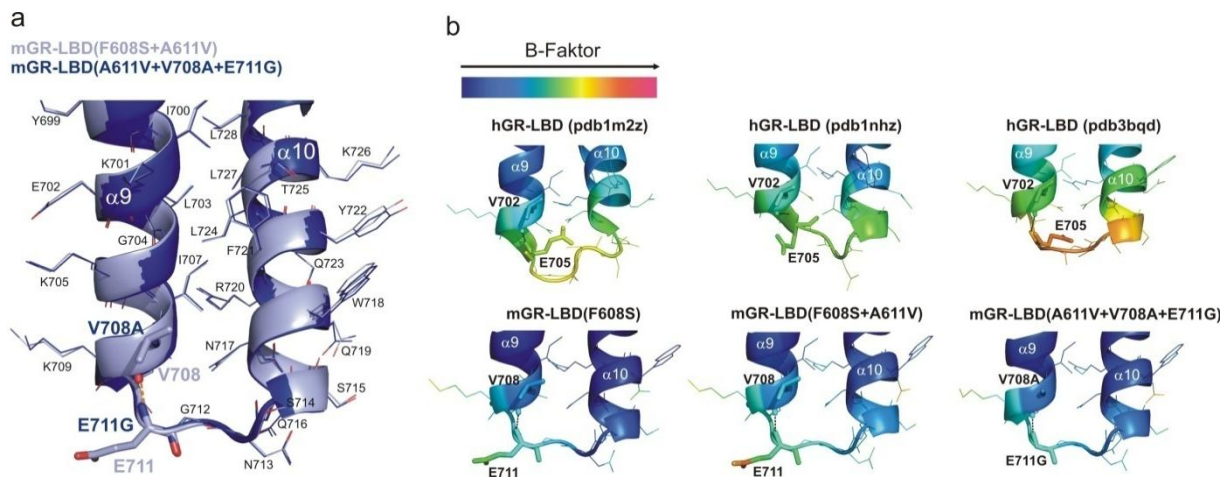


Abbildung 68: Umgebung von V702(V708) und E705(E711) in hGR-LBD (mGR-LBD).

a V708A und E711G induzieren keine strukturellen Veränderungen in mGR-LBD. So konnten die gezeigten AS 699-728 von mGR-LBD(F608S+A611V) und mGR-LBD(A611V+V708A+E711G) durch PyMol (<http://www.pymol.org/>) mit einem RMSD von 0,14 Å (218 Atome) superpositioniert werden. V708(A) liegt am C-terminalen Ende von α -Helix 9 und ragt ins Solvent. Die Aminogruppe von E711(G) bildet mit der Carbonylgruppe von V708(A) die letzte Wasserstoffbrückenbindung zwischen α -Helix 9 und Loop $\alpha_9\alpha_{10}$ aus (orange markiert, Distanz N-O: 2,9 Å).

b Erhöhte B-Faktoren und unterschiedliche Konformationen des Loops $\alpha_9\alpha_{10}$ in den gelösten Röntgenkristallstrukturen von hGR-LBD (pdb 1m2z; pdb 1nhz, pdb 3bqd) und mGR-LBD(F608S), mGR-LBD(F608S+A611V) bzw. mGR-LBD(A611V+V708A+E711G) weisen auf eine hohe Flexibilität dieses Bereiches hin.

V702 (V708) befindet sich am C-Terminus von α -Helix 9, wobei die Seitenkette der Aminosäure Lösungsmittel-zugänglich ist. Die Hydratisierung des Proteins sollte somit hier durch eine weniger hydrophobe Aminosäure optimiert werden, was die verbesserte Löslichkeit von hGR-LBD durch V702A erklären könnte. Darüber hinaus stabilisiert das neu eingeführte Alanin vermutlich α -Helix 9, da im Allgemeinen die Ausbildung von α -Helices durch Alanine stark begünstigt wird (relative Stabilisierung einer α -helikalen Konformation nach O'Neil & DeGrado (1990): -0,77 kcal (Ala) vs. -0,14 kcal/mol (Val)). E705 (E711) liegt am N-terminalen Ende von Loop $\alpha_9\alpha_{10}$, der α -Helix 9 und α -Helix 10 verbindet, wobei zwischen der Aminogruppe von E705 (E711) und der Carbonylgruppe von V702 (V708) eine Wasserstoffbrückenbindung ausgebildet wird (Abbildung 68).

In allen Röntgenkristallstrukturen von hGR-LBD und mGR-LBD zeigt Loop $\alpha_9\alpha_{10}$ erhöhte B-Faktoren im Vergleich zur restlichen Struktur, was auf eine hohe Flexibilität hindeutet (Abbildung 68b). So sind in den verschiedenen Röntgenkristallstrukturen auch unterschiedliche Konformationen von Loop $\alpha_9\alpha_{10}$ erkennbar, bei denen E705 (E711) entweder im Protein vergraben (pdb 1m2z, pdb 3bqd) oder komplett Lösungsmittel-zugänglich ist (pdb 1nhz, mGR-LBD - Strukturen) (Abbildung 68b). Durch Entfernung der dynamischen Seitenkette von E711 durch E711G wird demnach die Flexibilität des Proteins in diesem Bereich gesenkt, wodurch vermutlich die korrekte Ausbildung des Helix-Loop-Helix Motivs begünstigt und das Protein stabilisiert wird.

4.3.5.8.4 M752T

M752T wurde in den stark fluoreszierenden hGR-LBD-eGFP Varianten A5-1 und A8 identifiziert (4.3.2.2; Tabelle 12). Diese Mutation induziert isoliert in hGR-LBD-eGFP ebenfalls eine erhöhte Fluoreszenz *in vivo* (4.3.5.2) und verbessert die Löslichkeit (4.3.5.3) und die Stabilität (4.3.5.5) von hGR-LBD. Allerdings konnte gezeigt werden, dass durch M752T lediglich die Agonist-Konformation stabilisiert wird und die Mutation keinen Einfluss auf die Stabilität der Antagonist-Konformation des Proteins hat (4.3.5.5).

M752 liegt in α -Helix 12 („AF-2“ = *activation function 2*) von hGR-LBD, welche für die Erkennung und Bindung von Coaktivatoren und Corepressoren über ein LXXLL-Motiv verantwortlich ist (siehe 1.7; Abbildung 8). Falls in der menschlichen Zelle Ligand-gebundenes hGR ohne Cofaktor vorliegt, wird α -Helix 12 durch Bindung von Hsp90 stabilisiert (Fang *et al.*, 2006).

Die Röntgenkristallstrukturen der Agonist-Konformation von GR-LBD mit gebundenem TIF2 Coaktivator-Peptid zeigen, dass M752 an der Bindung des Coaktivators beteiligt ist. Hierbei wird durch die Seitenketten von M752 und mehreren Leucinen aus TIF2 ein hydrophobes Netzwerk ausgebildet (Abbildung 69).

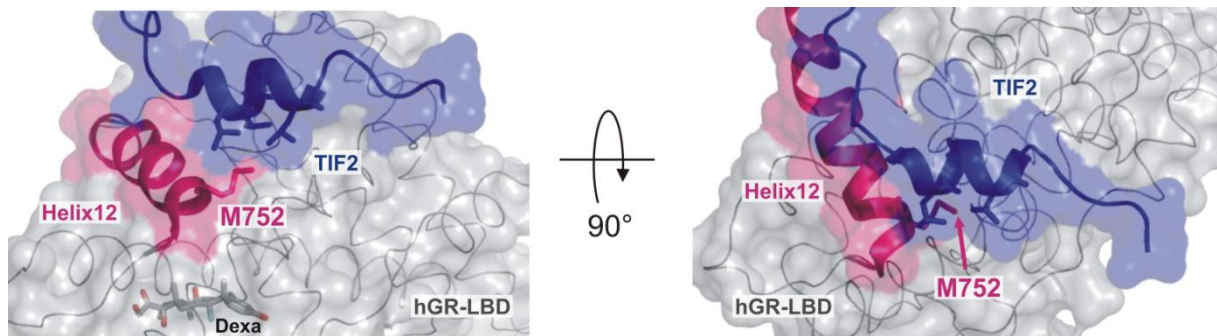


Abbildung 69: Bindung von Coaktivator-Peptid TIF2 an α -Helix 12 von hGR-LBD (Agonist-Konformation).

Gezeigt ist die Röntgenkristallstruktur von hGR-LBD mit gebundenem Dexamethason und TIF2 (pdb 1m2z; Bledsoe *et al.*, 2002). M752 ist an der Bindung von TIF2 beteiligt und bildet mit drei Leucinen von TIF2 (NLLRYLLD) ein hydrophobes Netzwerk aus. M752 ist somit bei TIF2-Bindung komplett vom Lösungsmittel abgeschirmt.

Die in dieser Arbeit gezeigten Untersuchungen von hGR-LBD wurden stets in Abwesenheit von TIF2 und Hsp90 durchgeführt, wodurch α -Helix 12 keinen stabilisierenden Partner besitzt und vermutlich Lösungsmittel-zugänglich vorliegt. Dadurch ist es nicht verwunderlich, dass bei der Sortierung optimierter hGR-LBD Varianten, eine Mutation in diesem Bereich gefunden worden ist. Vermutlich wird durch das hydrophilere Threonin an Position 752 die Hydratisierung der Lösungsmittel-zugänglichen α -Helix 12 verbessert.

Da M752 an der Interaktion zwischen TIF2 und hGR-LBD beteiligt ist, könnten Mutationen dieser Aminosäure die Bindung des Coaktivator-Peptids an das Protein beeinflussen. Mit Hilfe von Oberflächenplasmonresonanz-Experimenten (durchgeführt von Andrea Wiget, F. Hoffmann – LaRoche, *Pharma Research Discovery*, Basel) konnte gezeigt werden, dass durch M752T die Bindungsaffinität von TIF2 an hGR-LBD (Agonist-Konformation) tatsächlich um mindestens eine Größenordnung schlechter wird. Im Gegensatz dazu wird die Bindung des Corepressor-Peptids NCoR an hGR-LBD (Antagonist-Konformation) durch M752T nicht beeinträchtigt (Tabelle 16).

Tabelle 16: Bindungsaffinität von TIF2 Coaktivator- bzw. NCoR Corepressor-Peptid an hGR-LBD Varianten, ermittelt durch Oberflächenplasmonresonanz (BIAcoreTM).

Die Messungen in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 0,005 % P20, pH 7,9 bei RT, wurden von Andrea Wiget (F.Hoffmann – LaRoche, *Pharma Research Discovery*, Basel) durchgeführt.

hGR-LBD Variante	+Agonist (Dexamethason)	+Antagonist (Mifepriston)
	K_D (TIF2)	K_D (NCoR)
A605V	0,6 μ M	0,9 μ M
A605V+M752T	7,0 μ M	2,7 μ M
F602S	2,5 μ M	5,1 μ M
F602S+A605V+M752T	38,9 μ M	3,1 μ M

Für weitere röntgenkristallographische Untersuchungen von hGR-LBD Varianten mit Mutationen aus dieser Arbeit sind somit M752T-enhaltende Varianten lediglich für die Ko-Kristallisation mit NCoR geeignet. Die Ko-Kristallisation von hGR-LBD (Agonist-Konformation) und dem Coaktivator-Peptid TIF2 würde dagegen durch die M752T-bedingte Erniedrigung der Bindungsaffinität erschwert werden.

5 Abschließende Diskussion

5.1 Rationales Design einer hydrophilen Proteinoberfläche

Der gefaltete Zustand eines Proteins wird von einer Hydrat-Hülle umgeben, welche Wasserstoffbrückenbindungen zur Molekül-Oberfläche ausbildet und somit das Protein in Lösung hält. Durch Optimierung dieser Wechselwirkungen sollte die Löslichkeit eines Proteins verbessert werden können. Dies gelang bereits in mehreren publizierten Arbeiten, bei denen mit Hilfe von zielgerichteter Mutagenese hydrophobe Aminosäuren auf der Proteinoberfläche durch hydrophilere Reste ersetzt wurden (Dale *et al.*, 1994; Dyda *et al.*, 1994; Jenkins *et al.*, 1995; Malissard & Berger, 2001; Das & Georgiadis, 2001; Mosavi & Peng, 2003; Avramopoulou *et al.*, 2004; Trevino *et al.*, 2007). Darüber hinaus kann solch ein Austausch von exponierten hydrophoben Aminosäuren ebenfalls zu einer Erhöhung der konformationellen Stabilität eines Proteins führen. Im Grunde sollte eine an der Oberfläche lokalisierte hydrophobe Aminosäure im entfalteten und gefalteten Zustand des Proteins in gleicher Weise Lösungsmittel-exponiert vorliegen, wodurch sie keinen Einfluss auf die konformationelle Stabilität des Proteins (ΔG_{DN} ; siehe 1.1.1, 1.1.2; Formel 1, 2) hätte, da die Entropie des Wassers bei beiden Zuständen um den gleichen Betrag gesenkt würde. Allerdings kann es vorkommen, dass ein hydrophober Rest im entfalteten Zustand des Proteins insgesamt weniger Lösungsmittel-exponiert vorliegt als im nativen Zustand, wodurch der entfaltete Zustand begünstigt und die konformationelle Stabilität des Proteins gesenkt wird („reverser hydrophober Effekt“; Pakula & Sauer, 1990). In Übereinstimmung mit dieser Überlegung konnte bereits in mehreren Arbeiten durch den Austausch von hydrophoben Resten auf der Proteinoberfläche eine Stabilisierung eines Proteins erzielt werden (Tamura & Sturtevant, 1995; Strub *et al.*, 2004).

Aufbauend auf diesen Ergebnissen sollte versucht werden durch die Erhöhung der Anzahl an exponierten hydrophilen Aminosäureresten, die Löslichkeit und Stabilität von hPCSK9³¹⁻⁴⁵⁴(S386A) zu verbessern. Da zu diesem Zeitpunkt des Projekts noch keine experimentell ermittelten Strukturen zur Verfügung standen, wurden mehrere Homologie-Modelle mit Hilfe unterschiedlicher Templat-Strukturen und Modellierungs-Programme erstellt. Die Überprüfung dieser Modelle mit den mittlerweile verfügbaren Röntgenkristallstrukturen des Proteins (Cuningham *et al.*, 2007; Hampton *et al.*, 2007; Piper *et al.*, 2007) zeigte, dass die exponierten Aminosäuren gut vorhergesagt wurden (Tabelle 6). Die auf Grundlage dieser Modelle hergestellte Mutante hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt}, bei der 15 exponierte Reste (Trp, Phe, Tyr, Val, Leu, Thr) durch hydrophilere Aminosäuren (His, Glu,

Ser) ausgetauscht wurden, konnte im Vergleich zum Wildtyp mit einer ca. 5-fach höheren Ausbeute aus der löslichen Zellfraktion von *E. coli* aufgereinigt werden. Darüber hinaus wurde bei fraktionierten Präzipitationsexperimenten hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} erst bei höheren Ammoniumsulfat-Konzentrationen ausgefällt als das wildtypische Protein (Abbildung 29). Es kommt demnach durch die erhöhte Anzahl an polaren Resten auf der Oberfläche wie erwartet zu einer verbesserten Hydratisierung des Moleküls, wodurch die Löslichkeit *in vivo* und *in vitro* erhöht wird. Andererseits zeigte sich allerdings, dass die Mutationen zu einer Destabilisierung des Proteins führten, da hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt} bereits bei niedrigeren Harnstoff-Konzentrationen denaturierte als das wildtypische Protein (Abbildung 30). Man muss demnach davon ausgehen, dass einige der hydrophoben Aminosäuren auf der Oberfläche des wildtypischen hPCSK9³¹⁻⁴⁵⁴(S386A) Proteins zu dessen Stabilität beitragen. Diese Beobachtung widerspricht der generellen Annahme, dass lediglich hydrophobe Aminosäurenreste im Proteinkern aufgrund des hydrophoben Effekts einen signifikanten Beitrag zur konformationellen Stabilität eines Proteins haben. Allerdings wurde bereits in mehreren anderen Arbeiten beschrieben, dass hydrophobe Aminosäuren der Proteinoberfläche durch Interaktionen mit hydrophoben Teilen benachbarter polarer Reste einen wichtigen Beitrag zur konformationellen Stabilität eines Proteins leisten können (van den Burg *et al.*, 1994; Frigerio *et al.*, 1996). Darüber hinaus kann nicht ausgeschlossen werden, dass bei der hohen Anzahl an eingeführten Mutationen in hPCSK9³¹⁻⁴⁵⁴(S386A)^{opt}, es zum Teil zu destabilisierenden Wechselwirkungen zwischen den neuen Resten und Aminosäuren der Umgebung kommen kann (z.B. durch repulsive Kräfte aufgrund von kollidierenden van-der-Waals Radii oder elektrostatische Abstoßung durch neu eingeführte Glutamate).

Nichtsdestotrotz kann abschließend festgestellt werden, dass der beschriebene Ansatz eine erfolgsversprechende Vorgehensweise zur Erhöhung der Löslichkeit eines Proteins darstellt. Wie bereits erwähnt wurde, gibt es hierfür auch bereits zahlreiche Beispiele in der Literatur. Welche Konsequenzen die eingeführten hydrophilen Reste auf der Oberfläche allerdings auf die konformationelle Stabilität des Proteins haben ist ungewiss und muss, wie gezeigt wurde, im Einzelfall genau überprüft werden.

5.2 Erhöhung der Löslichkeit und Stabilität von Proteinen durch gelenkte Evolution mit Hilfe von eGFP als Faltungsreporter und Durchflusszytometrie als Hochdurchsatz *Screening*-Verfahren

Um Eigenschaften von Proteinen, wie z.B. die Stabilität oder auch die katalytische Aktivität, zu verändern, haben sich neben rationalem Design in den letzten Jahren immer mehr auch verschiedene Methoden der gelenkten Evolution etabliert. Hierbei werden die Prozesse, welche in der Natur zur Anpassung von Proteinen an unterschiedliche Bedingungen verwendet werden, im Labor imitiert (Arnold, 1998). Mit Hilfe solcher Ansätze konnten mittlerweile in vielen Fällen Proteine für die unterschiedlichsten Zwecke maßgeschneidert werden. Hierbei gelang beispielsweise die Variation der Substratspezifität (Fasan *et al.*, 2008) oder Enantioselektivität (Reetz *et al.*, 2004) von Enzymen, die Erhöhung der Toleranz von Proteinen gegenüber organischen Lösungsmitteln (Moore & Arnold, 1996), die Verbesserung schwacher enzymatischer Aktivitäten (Röthlisberger *et al.*, 2008) oder die Erhöhung der Stabilität von Proteinen (Kather *et al.*, 2008).

Mit Hilfe der mittlerweile vielfältigen molekularbiologischen Werkzeuge können die Mechanismen der Evolution – Diversifizierung durch Mutagenese gefolgt von Selektion – im Detail im Labor nachempfunden werden. So kann zunächst das für das Zielprotein kodierende Gen durch molekularbiologische Methoden (z.B. *error prone* PCR - Leung *et al.*, 1989; DNA-*Shuffling* - Stemmer *et al.*, 1994a,b oder PCR mit degenerierten Primern - Neylon, 2004; Reetz *et al.*, 2008) oder chemische Gen-Synthese (kommerziell erhältlich durch z.B. Sloning BioTechnology GmbH (Puchheim); GeneArt AG (Regensburg): siehe Scott *et al.*, 2009) mit globaler oder lokaler Randomisierung hergestellt werden. Entscheidend bei der gelenkten Evolution ist es anschließend, mit Hilfe einer geeigneten Selektions- oder *Screening*-Methode die gewünschten Proteinvarianten aus dem großen Repertoire an erzeugten Mutanten, zuverlässig aussortieren zu können. So sollte mit Hilfe des gewählten Selektions- oder *Screening*-Verfahrens eine hohe Anzahl an Varianten analysierbar sein, da bereits nach Einführung weniger Mutationen in ein Protein eine extrem hohe Diversität erzeugt wird. So entstehen bei Gen-Synthese mit kompletter Randomisierung (Kodierung aller 20 proteinogenen Aminosäuren) von zwei, drei bzw. n Positionen $4 \cdot 10^2$, $8 \cdot 10^4$ bzw. 20^n unterschiedliche Varianten. Auch durch *error prone* PCR hergestellte Genbanken, deren Größe durch die Transformationseffizienz der zur Präparation benutzten *E. coli* Zellen beschränkt wird, können über 10^6 Varianten beinhalten (Tobias, 2003). Zur Untersuchung einer so großen Anzahl an Varianten sind Selektions-Methoden sehr günstig. Hierbei werden Funktionen *in vivo* analysiert, bei denen das Zielprotein auf die Fähigkeit getestet werden

kann, metabolische Defekte zu komplementieren (Jürgens *et al.*, 2000) oder Toxine abzubauen (Hoseki *et al.*, 1999). Da somit nur solche Mutanten überleben sollten, welche die gewünschten Eigenschaften zeigen, können innerhalb eines Experiments sehr viele Varianten untersucht werden. Da allerdings nur relativ wenige Proteinfunktionen an das Überleben eines Organismus gekoppelt werden können, stehen für viele Fragestellungen keine Selektions-Methoden zur Verfügung. In solchen Fällen muss auf *Screening*-Verfahren zurückgegriffen werden. Hierbei werden die Eigenschaften aller in der Bank enthaltenen Varianten einzeln analysiert und bewertet, was den ganzen Prozess deutlich aufwendiger macht und in der Regel verlangsamt. *Screening*-Verfahren beruhen oftmals auf der Detektion eines kolorimetrischen oder fluoreszierenden Signals mit Hilfe von Mikrotiterplatten-Lesern, wobei in einem vernünftigen Zeitrahmen lediglich $10^2 - 10^4$ Varianten untersucht werden können (Jäckel *et al.*, 2008). Die Etablierung von Hochdurchsatz-Verfahren zum *Screening* von Banken würde also die Effizienz von gelenkter Evolution deutlich steigern. So gewann die Durchflusszytometrie, bei der mehr als 10^4 Ereignisse/sek analysiert werden können, als *Screening*-Verfahren bei der gelenkten Evolution von enzymatischen Aktivitäten in letzter Zeit immer mehr an Bedeutung (Bershtein & Tawfik, 2008; Yang & Withers, 2009). Da mit Hilfe dieser Methode lediglich Zellen sortiert werden können, wurden mehrere Systeme zur Kopplung der Zell-Fluoreszenz an die gewünschte Enzym-Funktion etabliert. Beispielsweise gelang dies indem der Zell-Export eines fluoreszierenden Substrates durch die gewünschte Enzymaktivität verhindert wurde (Aharoni *et al.*, 2006) oder die fluoreszierenden Produkte auf der Zell-Oberfläche fixiert wurden (Becker *et al.*, 2007; Lipovsek *et al.*, 2007). Eine weitere Möglichkeit zur Kopplung des Geno- und Phänotyps von Enzym-Varianten stellen die *in vitro* Kompartimentierung der Enzym-produzierenden Zelle (Aharoni *et al.*, 2005) oder ein *in vitro* Transkriptions/Translationssystem (Mastrobattista *et al.*, 2005) mit dem fluoreszierenden Enzym-Produkt in Wasser-Öl-Wasser Emulsions-Tröpfchen dar, welche anschließend mit Durchflusszytometrie analysiert und sortiert werden können (Miller *et al.*, 2006).

Im Rahmen dieser Arbeit wurde nun versucht Durchflusszytometrie als Hochdurchsatz *Screening*-Methode bei der gelenkten Evolution von Proteinen hin zu hoher Löslichkeit und Stabilität zu verwenden. Als Reporterprotein *in vivo* wurde hierbei ein C-terminal an das Zielprotein fusioniertes eGFP benutzt. Die Verwendung von eGFP als Faltungsreporter *in vivo* ist bereits mehrfach in der Literatur beschrieben, wobei in der Regel verschiedene eGFP-Fusionsproteinvarianten über die Fluoreszenz von auf Agar-Platten gewachsenen *E. coli* Kolonien analysiert wurden. Hierbei konnten in der Regel lediglich $1 - 4 \cdot 10^4$ Varianten

untersucht werden (Waldo *et al.*, 1999; Kawasaki & Inagaki, 2001; Pedelacq *et al.*, 2002; Li *et al.*, 2009). Mit Hilfe von Durchflusszytometrie könnte im Gegensatz dazu ein Vielfaches an Zellen schnell und zuverlässig analysiert werden. Diese Vorgehensweise wurde bisher erst in einer veröffentlichten Arbeit zur Identifikation einer löslicheren Variante der TEV-Protease benutzt (van den Berg *et al.*, 2006). Hierbei konnte die Reinigungs-Ausbeute des Proteins aus der löslichen Zellfraktion von *E. coli* ca. 5-fach gesteigert werden. Die Stabilität der erhaltenen Proteinvariante wurde nicht näher charakterisiert.

Zur Etablierung gelenkter Evolution hoher Proteinlöslichkeit und Proteinstabilität mit Hilfe von eGFP als Faltungsreporter und Durchflusszytometrie als *Screening*-Verfahren wurde in der vorliegenden Arbeit ein leistungsfähiges MoFlo[®] Zytometer verwendet, mit dessen Hilfe bis zu 70.000 bzw. 20.000 Zellen/sek analysiert bzw. sortiert werden können. Zunächst wurden unterschiedliche Expressionsbedingungen der eGFP-Fusionsgene getestet. Hierbei zeigte sich, dass das von Waldo *et al.* (1999) beschriebene System (induzierbare Expression mit Hilfe der T7-Polymerase) die optimale Expressionsbedingung darstellte. Durch keines der untersuchten eGFP-fusionierten Proteine hKAS, hPCSK9³¹⁻⁴⁵⁴(S386A) und hGR-LBD konnte bei starker oder schwacher konstitutiver Expression eine konstante Zell-Fluoreszenz erhalten werden, weswegen sich diese Systeme nicht zur schrittweisen Anreicherung von stark fluoreszierenden Zellen eigneten (Abbildung 31, 32).

Bei den Expressions-Tests wurde zum Teil proteolytischer Abbau der eGFP-Fusionsproteine *in vivo* beobachtet, was möglicherweise die Analyse der Proteinlöslichkeit *in vivo* stören könnte, da lösliche eGFP-beinhaltende Abbauprodukte zu einer hohen Hintergrund-Fluoreszenz führen könnten. Um die Proteolyse-Anfälligkeit zu unterbinden, wurde exemplarisch der flexible Linker zwischen dem Zielprotein tmHisF und eGFP versuchsweise von 13 auf 3 Aminosäuren verkürzt. Die Länge von Linkern zwischen zwei Domänen kann allerdings erheblichen Einfluss auf deren Faltung und Funktion ausüben (Wriggers *et al.*, 2005). So zeigte die Fusionsproteinvariante mit verkürztem Linker keine Fluoreszenz mehr (Abbildung 34), weshalb der ursprünglich benutzte GGRAGSAAGSGEF-Linker beibehalten wurde.

Trotz der beobachteten Proteolyse-Anfälligkeit konnte nach induzierter Genexpression die unterschiedliche Löslichkeit verschiedener Proteine (Positivkontrolle: tmHisF; Negativkontrolle: hKAS) (Abbildung 33) bzw. Proteinvarianten (hGR-LBD, hGR-LBD(F602S)) (Abbildung 37) sensitiv mit Hilfe des MoFlo[®] Zytometers abgebildet werden. Anreicherungstests mit dem Gerät zeigten zudem, dass stark fluoreszierende Zellen mit einem löslichen eGFP-Fusionsprotein aus einem zuvor festgelegten Sortier-Bereichs nach bereits

einer FACS-Runde über 100-fach angereichert werden konnten (Abbildung 35). Als Zielprotein zur gelenkten Evolution hoher Löslichkeit und Stabilität wurde die Ligand-Bindungsdomäne des humanen Glucocorticoid-Rezeptors (hGR-LBD), welche ohne stabilisierende Mutation und His₆-Glutathion-S-Transferase (His₆-GST)-Tag, nicht löslich in *E. coli* hergestellt werden kann (Bledsoe *et al.*, 2002) ausgewählt. Dieses Protein schien für die geplante Optimierung geeignet, da es fusioniert an eGFP nur eine sehr niedrige Fluoreszenz *in vivo* zeigte, wodurch ein großes Potential für Verbesserungen vorlag (Abbildung 37). Darüber hinaus ist mit F602S bereits eine stabilisierende Mutation in hGR-LBD (Bledsoe *et al.*, 2002) bekannt, welche als Positivkontrolle benutzt werden konnte (Abbildung 37).

Mit Hilfe des MoFlo[®] Zytometers gelang die Anreicherung stark fluoreszierender Zellen aus *E. coli* Kulturen, welche verschiedene *hgr-lbd-egfp* Genbanken exprimierten (Abbildung 44). Bei näherer Analyse der angereicherten Zellen und der darin enthaltenen Konstrukte, konnten neben falsch positiven Ursachen für hohe Zell-Fluoreszenz – ähnlich wie in Cabantous *et al.* (2008) beschrieben: Verkürzungen auf Genebene durch Deletion im Zielgen gefolgt von einem Stoppcodon oder kompletter Verlust des Zielgens (siehe 4.3.2.1, 4.3.2.2; Abbildung 42, 45) führt zu alternativer Translation von stark fluoreszierendem eGFPwt – insgesamt fünf hGR-LBD-eGFP Varianten (A5-1 - A5-4, A8) mit 30 Aminosäuremutationen identifiziert werden (Tabelle 12). Einerseits zeigten diese Varianten bei Retransformation der entsprechenden Konstrukte in *E. coli* eine erhöhte Fluoreszenz *in vivo* (Abbildung 46), andererseits wurden sie allerdings überraschenderweise in Abwesenheit von eGFP als His₆-GST-Fusionsproteine nicht löslicher in *E. coli* hergestellt als wildtypisches hGR-LBD (Abbildung 48). Daraus kann geschlossen werden, dass vermutlich einige der Mutationen zwar die Fluoreszenz von fusioniertem eGFP erhöhen, gleichzeitig aber einen negativen Effekt auf die Löslichkeit von His₆-GST-hGR-LBD ausüben. Der Einfluss der Mutationen scheint demnach zu einem gewissen Grad vom Fusionspartner abhängig zu sein. Die hGR-LBD Fusionspartner eGFP und His₆-GST unterscheiden sich in ihrer Lokalisation (eGFP: C-terminal; His₆-GST: N-Terminal) und in ihrem Oligomerisierungsgrad (eGFP = Monomer – Tsien *et al.*, 1998; GST = Homodimer - McTigue *et al.*, 1995; Waugh, 2005), weshalb solche Effekte nicht ausgeschlossen werden können. Daneben konnten Mutationen in hGR-LBD-eGFP identifiziert werden, welche zu einem falsch positiven Artefakt und nicht zu einer verbesserten Löslichkeit von hGR-LBD-eGFP führen. Zum Einen führt Q738H in hGR-LBD zu einer erhöhten Proteolyse-Anfälligkeit (Abbildung 50), wodurch es bei der Herstellung von hGR-LBD-eGFP vermutlich zu einer vermehrten Abspaltung des stark fluoreszierenden

eGFPwt kommt. Zum Anderen kommt es bei Einführung von M560K in hGR-LBD-eGFP zu einer starken Erhöhung der Fluoreszenz der unlöslichen Fusionprotein enthaltenden *inclusion bodies* in *E. coli* (Abbildung 51). Anders als in der Literatur zum Teil beschrieben wird (Garcia-Fruitos *et al.*, 2007) ist dies in diesem Fall aber kein Signal für „konformationelle Qualität“, da eindeutig gezeigt werden konnte, dass M560K zu einer Erniedrigung der Löslichkeit und der konformationellen Stabilität von hGR-LBD führt (Abbildung 53, 57). Insgesamt konnten somit fünf unterschiedliche Quellen falsch positiver Artefakte bei Benutzung von eGFP als Faltungsreporter *in vivo* identifiziert werden. Die Ursachen für diese Artefakte und die in dieser Arbeit beschriebenen Methoden zu deren Identifikation sind in Tabelle 17 zusammengefasst.

Tabelle 17: Falsch positive Artefakte beim Screening nach stark fluoreszierenden eGFP-Fusionsproteinen *in vivo*.

Die falsch positiven Artefakte werden jeweils in den angegebenen Kapiteln näher erläutert.

Artefakt auf Genebene	Ursache	Konsequenz	Identifikation
1. Stoppcodon im Zielgen (siehe 4.3.2.1)	Deletion durch <i>error prone</i> PCR	hohe Fluoreszenz durch Translation von <i>egfp</i> ohne/ mit stark verkürztem Fusionspartner	Analytisches FACS-Histogramm zeigt heterogene Zell-Population (Abb. 41)
2. Verkürztes Zielgen (siehe 4.3.2.2)	Rekombination/ Klonierungsartefakt	hohe Fluoreszenz durch Translation von <i>egfp</i> ohne/ mit stark verkürztem Fusionspartner	Kolonie-PCR nach FACS-Anreicherung: Amplifikationsprodukt ist zu kurz (Abb. 45)
Artefakt auf Proteinebene			
3. Verkürztes Protein (siehe 4.3.4)	Erhöhte Proteolyse-Anfälligkeit durch eingeführte Mutation	hohe Fluoreszenz durch eGFP ohne Fusionspartner	SDS-PAGE nach Probeexpression des Zielproteins zeigt deutliche Abbaubande (Abb. 50)
4. Stark fluoreszierende <i>inclusion bodies</i> (siehe 4.3.5.2)	Effizientere Ausbildung des eGFP-Chromophors durch eingeführte Mutation	überproportional starke Fluoreszenz in der unlöslichen Zellfraktion	Fluoreszenz-Mikroskopie: starke Fluoreszenz an den Zell-Polen (Abb. 51)
5. Keine Korrelation zwischen Löslichkeit des Zielproteins und eGFP-Fluoreszenz (siehe 4.3.3.2)	Erhöhung der eGFP-Fluoreszenz durch Mutation im Zielprotein	keine erhöhte Löslichkeit der Mutante nach Entfernung von eGFP	Probeexpression der sortierten Variante ohne eGFP: keine verbesserte Löslichkeit (Abb. 48)

Nichtsdestotrotz konnten nach Vereinzelung von 14 Mutationen aus den fünf hGR-LBD-eGFP Varianten vier Aminosäureaustausche (A605V, V702A, E705G, M752T) identifiziert werden, welche ähnlich wie die in der Literatur beschriebene Mutation F602S (Bledsoe *et al.*, 2002) zu einer erhöhten löslichen Herstellung von His₆-GST-hGR-LBD führen (Abbildung 50). Wie sich herausstellte, können diese Mutationen sowohl die Löslichkeit (Abbildung 53) und die konformationelle Stabilität (Abbildung 57) von hGR-LBD als auch die Fluoreszenz *in vivo* von hGR-LBD-eGFP (Abbildung 52) additiv verbessern. Die ermittelte Löslichkeit bzw. Stabilität von den hGR-LBD Varianten mit diesen Mutationen in unterschiedlicher Kombination korrelierte hierbei eindeutig mit der durch das MoFlo[®] Zytometer ermittelten Zell-Fluoreszenz durch die entsprechenden hGR-LBD-eGFP Varianten (Abbildung 54, 58). Zusammenfassend kann somit festgestellt werden, dass mit Hilfe des MoFlo[®] Zytometers eine hohe Anzahl an Proteinvarianten zügig analysiert und sortiert werden konnte. Aufgrund der beobachteten Korrelation zwischen der Fluoreszenz *in vivo* der eGFP-fusionierten Proteine und der Löslichkeit bzw. Stabilität der Proteine in Abwesenheit von eGFP empfiehlt sich dieses Vorgehen auch zur Verbesserung anderer Proteine. Bei der Anreicherung und Analyse von fluoreszierenden Zellen bzw. eGFP-Fusionsproteinen müssen allerdings die in Tabelle 17 beschriebenen falsch positiven Artefakte berücksichtigt werden. Eine optimale Vorgehensweise zur zügigen Identifikation vorteilhafter Mutationen auf Grundlage der in dieser Arbeit gewonnen Erkenntnisse ist in Abbildung 70 gezeigt.

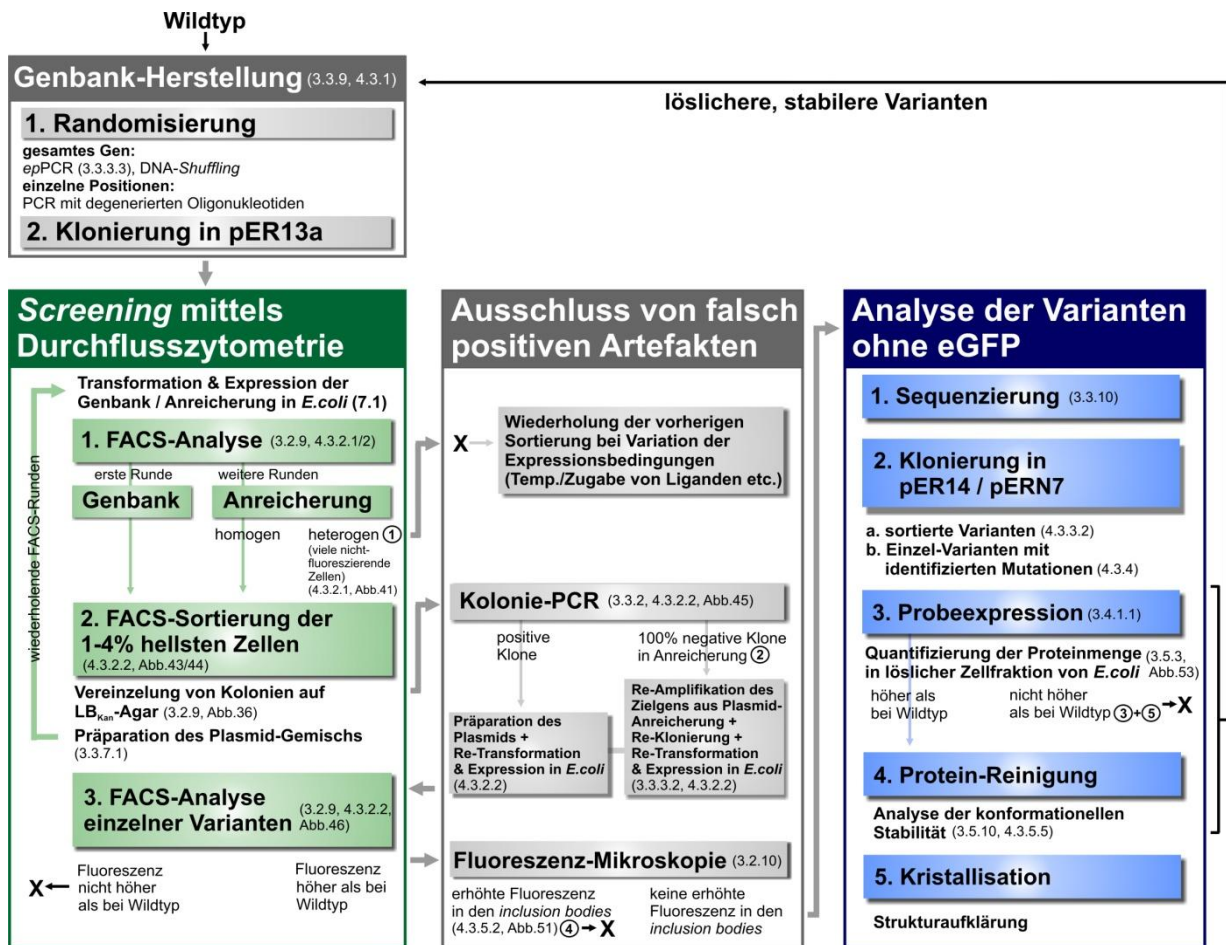


Abbildung 70: Schematische Darstellung des Vorgehens zur Verbesserung der Löslichkeit und Stabilität eines Proteins mit Hilfe von eGFP als Faltungsreporter und Durchflusszytometrie als Hochdurchsatz Screening-Methode.

Falsch positive Artefakte sind mit der in Tabelle 17 gezeigten Nummerierung markiert. Die Verwerfung einer Variante ist mit einem X gekennzeichnet.

Nach Randomisierung des für den Wildtyp oder für eine verbesserte Variante des Zielproteins kodierenden Gens können mehrere FACS-Anreicherungsrounds stark fluoreszierender Mutanten *in vivo* durchgeführt werden. Nach Ausschluss der falsch positiven Artefakte auf Genebene - Stoppcodon im Zielgen (vgl. Tabelle 17: Artefakt 1) bzw. Gen-Verlust (vgl. Tabelle 17: Artefakt 2) - sollten einzelne Mutanten nochmals hinsichtlich ihrer Fluoreszenz *in vivo* mittels Durchflusszytometrie getestet werden. Nach Ausschluss falsch positiver Artefakte durch stark fluoreszierende *inclusion bodies* (vgl. Tabelle 17: Artefakt 4) mittels Fluoreszenz-Mikroskopie können isolierte Mutanten bzw. Einzel-Varianten der durch Sequenzierung identifizierten Mutationen, nach Klonierung der entsprechenden Gene in pER14 bzw. pERN7, in Abwesenheit von eGFP analysiert werden. Nach Probeexpression der entsprechenden Konstrukte in *E. coli* kann die Menge an hergestelltem Zielprotein in der löslichen Zellfraktion quantifiziert werden. Dadurch können optimierte Mutanten identifiziert und weitere Artefakte auf Proteinebene - verkürzte Proteine durch gesteigerte Proteolyse-Anfälligkeit (vgl. Tabelle 17: Artefakt 3) bzw. Erhöhung der eGFP-Fluoreszenz durch Mutation im Zielprotein (vgl. Tabelle 17: Artefakt 5) - ausgeschlossen werden. Mit Hilfe der für die Löslichkeit vorteilhaften Mutationen kann das Zielprotein gereinigt werden, woraufhin die konformationelle Stabilität charakterisiert werden kann. Löslichere und stabilere Proteinmutanten können anschließend zur Strukturaufklärung mittels Röntgenkristallographie verwendet werden. Zur weiteren Verbesserung der Zielproteinvarianten kann nach Randomisierung der kompletten Sequenz bzw. einzelner Positionen eine neue Screening-Runde begonnen werden.

Um hGR-LBD weiter zu verbessern, könnten neue *hgr-lbd* Genbanken hergestellt und nach optimierten Varianten durchsucht werden, bei denen mit Hilfe von degenerierten Primern die für die Positionen F602, A605, V702, E705 und M752 kodierenden Codone komplett randomisiert wurden. Dadurch könnte diejenige hGR-LBD Mutante identifiziert werden, welche die optimalen Austausche an diesen sensitiven Stellen aufweist.

Um in Zukunft die beschriebenen falsch positiven Artefakte zu vermeiden, könnte alternativ zu eGFP das in der Literatur beschriebene Split-GFP (Cabantous *et al.*, 2005; Cabantous & Waldo, 2006) als Faltungsreporter-System benutzt werden. Hierbei wird das Zielprotein lediglich mit β -Strang 11 einer Superfolder GFP-Variante (Pedelacq *et al.*, 2006) fusioniert, wobei das restliche Superfolder-GFP koexprimiert wird. Hierdurch würden weniger Wechselwirkungen zwischen Zielprotein und Faltungsreporter bestehen, wodurch eGFP-abhängige Mutationen (vgl. Tabelle 17: Artefakt 5) vermieden werden könnten. Auch die Ausbildung stark fluoreszierender *inclusion bodies* (vgl. Tabelle 17: Artefakt 4) erscheint bei diesem System unwahrscheinlich, wodurch zumindest zwei der fünf beschriebenen falsch positiven Artefakte (Tabelle 17) ausgeschlossen werden könnten.

5.3 Löslichere und stabilere Mutanten der Ligand-Bindungsdomäne des Glucocorticoid-Rezeptors

Die Ligand-Bindungsdomäne des humanen Glucocorticoid-Rezeptors (hGR-LBD) ist Angriffspunkt vieler Medikamente mit z.B. anti-inflammatorischer Wirkung wie Cortisol, Dexamethason oder Prednisolon, welche spezifisch an das Protein binden und seine Aktivität als Transkriptionsfaktor beeinflussen (Golan, 2005). hGR-LBD stellt somit ein interessantes Zielprotein in der pharmazeutischen Forschung dar. Die Strukturaufklärung dieses Proteins mit gebundenen neuen Ligand-Leitstrukturen ist hierbei ein wichtiger Zwischenschritt bei der Etablierung neuer Wirkstoffe (Danley *et al.*, 2006). Zur Erzeugung von Proteinkristallen für diesen Zweck sind homogene, hochkonzentrierte Proteinlösungen nötig (Durbin & Feher, 1994). Bedauerlicherweise kann allerdings wildtypisches hGR-LBD nicht heterolog in *E. coli* hergestellt werden, da sich hierbei das Protein komplett in der unlöslichen Zellfraktion befindet. Erst mit Hilfe der rational eingeführten Mutation F602S (Bledsoe *et al.*, 2002; Kauppi *et al.*, 2003; Suino-Powell *et al.*, 2008) bzw. F602Y (Biggadike *et al.*, 2008; Madauss *et al.*, 2008) konnte hGR-LBD so weit stabilisiert werden, dass sowohl die Reinigung als auch die Röntgenkristallstrukturanalyse des Proteins mit verschiedenen Liganden möglich wurde. Allerdings kann hGR-LBD(F602S) lediglich mit einer Ausbeute von ca. 0,20 mg Protein / g Biomasse aus *E. coli* aufgereinigt werden (siehe Anhang 7.4; Tabelle A1), was die

Proteinherstellung zur Kristallisation sehr aufwendig macht. Die Identifikation anderer Mutationen, welche die Löslichkeit und Stabilität von hGR-LBD weiter erhöhen, würde somit die heterologe Herstellung in *E. coli* bzw. die Röntgenkristallstrukturanalyse deutlich erleichtern. Darüber hinaus könnte durch den Vergleich von verschiedenen Proteinvarianten der Einfluss der stabilisierenden Mutationen im Detail abgeschätzt und nur solche Varianten zur Ko-Kristallisation mit neu identifizierten Liganden benutzt werden, bei denen keine zu starken strukturellen Veränderungen durch die Mutationen induziert werden.

Mit der in 4.3 gezeigten und in 5.2 zusammengefassten Vorgehensweise konnten in der vorliegenden Arbeit insgesamt vier vorteilhafte Mutationen – A605V, V702A, E705G und M752T – in hGR-LBD identifiziert werden, welche vereinigt zu einer ca. 26-fachen Steigerung der Reinigungsausbeute des Proteins aus der löslichen Zellfraktion von *E. coli* führen (siehe Anhang 7.4; Tabelle A1) und erstmals die Reinigung des Proteins ohne die Mutation von F602 ermöglichten. Fraktionelle Präzipitationsexperimente mit Ammoniumsulfat zeigten, dass bei Einführung dieser Mutationen auch die Löslichkeit *in vitro* von hGR-LBD deutlich verbessert wird (Abbildung 59). Darüber hinaus führen A605V, V702A und E705G vereinzelt und additiv zu einer höheren konformationellen Stabilität sowohl der Agonist- als auch der Antagonist-Konformation. M752T stabilisiert exklusiv die Agonist-Konformation, während die Stabilität der Antagonist-Konformation nicht beeinflusst wird (Abbildung 57; Tabelle 15). Bei Vereinigung aller vier Mutationen erhöht sich die apparente Schmelztemperatur (T_M^{app}) beider Konformationen um ca. +8 °C (Abbildung 57; Tabelle 15), ohne dass der Oligomerisierungsgrad, die Sekundär- oder die Tertiärstruktur signifikant verändert wird (Abbildung 60, 61, 62). Aufgrund der verbesserten Stabilität und Löslichkeit *in vivo* bzw. *in vitro* wird die Herstellung hoch konzentrierter, homogener Proteinpräparationen vereinfacht, was einen günstigen Ausgangspunkt für die Erzeugung von Proteinkristallen mit hoher Qualität darstellt (Durbin & Feher, 1994).

So konnte die Röntgenkristallstrukturanalyse von GR-LBD aus der Maus (mGR-LBD) nach Einführung der entsprechenden Mutationen zu A605V, V702A und E705G (A611V, V708A und E711G) deutlich optimiert werden (siehe 4.3.5.8.2). Die Struktur von mGR-LBD(A611V+V708A+E711G) mit einer Auflösung von 1,5 Å stellt die bis dato am höchsten aufgelösete Struktur von GR-LBD Strukturen überhaupt dar, liefert neue Einblicke in atomare Details (Abbildung 66) und ist die einzig verfügbare GR-LBD Struktur ohne eine Mutation von F608 (Abbildung 67). Somit konnte in diesem Fall, ähnlich wie in anderen Arbeiten (Dyda *et al.*, 1994; Jenkins *et al.*, 1995; Das & Georgiadis, 2001; Pedelacq *et al.*, 2002; Yang *et al.* 2002; Yang *et al.* 2003; Serrano-Vega *et al.*, 2008; Warne *et al.*, 2008), die

Strukturaufklärung eines ursprünglich problematischen Proteins durch Einführung von Mutationen, welche die Löslichkeit bzw. Stabilität erhöhen, signifikant erleichtert bzw. verfeinert werden.

Durch Vergleich der Strukturen von mGR-LBD(F608S) (Auflösung: 1,9 Å) und mGR-LBD(A611V+V708A+E711G) können die stabilisierenden Effekte der einzelnen Mutationen erklärt werden. So wird durch F608S in der Agonist-Konformation des Proteins ein stabilisierendes Wasserstoffbrückennetzwerk etabliert, ähnlich wie dies bereits von Kauppi *et al.* (2003) für die Antagonist-Konformation beschrieben wurde (Abbildung 67). A611V (= A605V in hGR-LBD) ist Teil der Ligand-Bindungstasche (Abbildung 64) und optimiert den van-der-Waals Kontakt zwischen Protein und Ligand (Abbildung 66c), ohne dabei die Konformation der anderen Ligand-bindenden Reste zu beeinflussen (Abbildung 66b). Hierdurch kann der stabilisierende Effekt von A611V erklärt werden, da die konformationelle Stabilität eines Proteins bei affinerer Bindung eines Liganden generell durch die resultierende stärkere Populierung des nativen Zustandes erhöht wird (Brandts & Lin, 1990; Waldron & Murphy, 2003; Matulis *et al.*, 2005). Die Zugabe von stark bindenden Liganden ist eine weit verbreitete Vorgehensweise um Proteine so weit zu stabilisieren, dass sie in großen Mengen aufgereinigt und zur Kristallisation verwendet werden können (Elleby *et al.*, 2004; Vedadi *et al.*, 2006; Hassel *et al.*, 2006). V708A und E711G (=V702A und E705G in hGR-LBD) sind beide Oberflächen-exponiert und liegen in einer flexiblen Region zwischen α -Helix 9 und 10 von GR-LBD, welche bei allen veröffentlichten Kristallstrukturen hohe B-Faktoren aufweist (Abbildung 68). Wie in der Arbeit von Reetz *et al.* (2006) gezeigt wurde, kann die Optimierung von Bereichen mit hohen B-Faktoren durch Mutagenese zu einer deutlichen Stabilisierung eines Proteins führen. Neben einer verbesserten Hydratisierung und α -Helix-Ausbildung durch V708A (siehe 4.3.5.8.3), erfolgt die Stabilisierung des Proteins vermutlich durch Erniedrigung der Seitenketten-Entropie von E711G, wodurch womöglich die korrekte Orientierung von α -Helix 9 und 10 gefördert wird. Die Absenkung der Oberflächen-Entropie kann zudem zu einer Erhöhung der Kristallisationstendenz des Proteins führen (Derewenda & Vekilov, 2005). Der Effekt von M758T (= M752T in hGR-LBD) konnte strukturell nicht erklärt werden, da diese Mutation wegen ihrem negativen Einfluss auf die Bindung des Coaktivator-Peptids TIF2 an die Agonist-Konformation (siehe Tabelle 16) nicht in das zur Strukturaufklärung benutzte Protein mit eingefügt wurde. Bei zukünftiger Strukturaufklärung der Antagonist-Konformation von hGR-LBD mit dem Corepressor-Peptid NCoR können allerdings hGR-LBD Varianten mit M752T verwendet werden, da die Bindung von NCoR durch die Mutation nicht beeinträchtigt wird. M752T liegt in der für die

Coaktivator-Bindung verantwortlichen α -Helix 12 von hGR-LBD (Abbildung 69). Bei der heterologen Expression in *E. coli* und bei den biophysikalischen Charakterisierungen des Proteins in dieser Arbeit wurde kein Coaktivator zugesetzt, wodurch α -Helix 12 Lösungsmittel-zugänglich vorliegt. Somit kann zumindest vermutet werden, dass das neu eingeführte, hydrophile Threonin an Position 752, unter den angelegten Bedingungen, zu einer verbesserten Hydratisierung des Proteins führt und damit die beobachtete Stabilisierung bewirkt.

Zusammenfassend kann festgestellt werden, dass die in dieser Arbeit identifizierten Aminosäureaustausche in GR-LBD sehr nützlich zur röntgenkristallographischen Strukturanalyse des Proteins sind. Die Variation der vorteilhaften Mutationen ermöglicht die Herstellung verschiedener GR-LBD-Mutanten mit erhöhter Kristallisationstendenz, welche in Zukunft zur Ko-Kristallisation mit neu entwickelten, pharmazeutisch relevanten Liganden verwendet werden können.

6 Literaturverzeichnis

Abifadel, M., M. Varret, J. P. Rabes, D. Allard, K. Ouguerram, M. Devillers, C. Cruaud, S. Benjannet, L. Wickham, D. Erlich, A. Derre, L. Villeger, M. Farnier, I. Beucler, E. Bruckert, J. Chambaz, B. Chanu, J. M. Lecerf, G. Luc, P. Moulin, J. Weissenbach, A. Prat, M. Krempf, C. Junien, N. G. Seidah and C. Boileau (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 34(2): 154-156.

Aharoni, A., G. Amitai, K. Bernath, S. Magdassi and D. S. Tawfik (2005) High-throughput screening of enzyme libraries: thiolactonases evolved by fluorescence-activated sorting of single cells in emulsion compartments. *Chem Biol* 12(12): 1281-1289.

Aharoni, A., K. Thieme, C. P. Chiu, S. Buchini, L. L. Lairson, H. Chen, N. C. Strynadka, W. W. Wakarchuk and S. G. Withers (2006) High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat Methods* 3(8): 609-614.

Alborn, W. E., G. Cao, H. E. Careskey, Y. W. Qian, D. R. Subramaniam, J. Davies, E. M. Conner and R. J. Konrad (2007) Serum proprotein convertase subtilisin kexin type 9 is correlated directly with serum LDL cholesterol. *Clin Chem* 53(10): 1814-1819.

Alexandrescu, A. T., P. A. Evans, M. Pitkeathly, J. Baum and C. M. Dobson (1993) Structure and dynamics of the acid-denatured molten globule state of alpha-lactalbumin: a two-dimensional NMR study. *Biochemistry* 32(7): 1707-1718.

Allard, D., S. Amsellem, M. Abifadel, M. Trillard, M. Devillers, G. Luc, M. Krempf, Y. Reznik, J. P. Girardet, A. Fredenrich, C. Junien, M. Varret, C. Boileau, P. Benlian and J. P. Rabes (2005) Novel mutations of the PCSK9 gene cause variable phenotype of autosomal dominant hypercholesterolemia. *Hum Mutat* 26(5): 497.

Almlof, T., A. E. Wallberg, J. A. Gustafsson and A. P. Wright (1998) Role of important hydrophobic amino acids in the interaction between the glucocorticoid receptor tau 1-core activation domain and target factors. *Biochemistry* 37(26): 9586-9594.

Anderson, D. E., W. J. Bechtel and F. W. Dahlquist (1990) pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* 29(9): 2403-2408.

Arakawa, T. and S. N. Timasheff (1985) Theory of protein solubility. *Methods Enzymol* 114: 49-77.

Arnold, F. H. (1998) When blind is better: protein design by evolution. *Nat Biotechnol* 16(7): 617-618.

Arnold, F. H. (1998) Design by directed evolution. *Acc. Chem. Res.* 31: 125-131.

Arnold, F. H. and J. C. Moore (1997) Optimizing industrial enzymes by directed evolution. *Adv Biochem Eng Biotechnol* 58: 1-14.

Avramopoulou, V., A. Mamalaki and S. J. Tzartos (2004) Soluble, oligomeric, and ligand-binding extracellular domain of the human alpha7 acetylcholine receptor expressed in yeast: replacement of the hydrophobic cysteine loop by the hydrophilic loop of the ACh-binding protein enhances protein solubility. *J Biol Chem* 279(37): 38287-38293.

Baltrusch, S. and S. Lenzen (2008) Monitoring of glucose-regulated single insulin secretory granule movement by selective photoactivation. *Diabetologia* 51(6): 989-996.

Bandeiras, T. M., R. C. Hillig, P. M. Matias, U. Eberspaecher, J. Fanghanel, M. Thomaz, S. Miranda, K. Crusius, V. Putter, P. Amstutz, M. Gulotti-Georgieva, H. K. Binz, C. Holz, A. A. Schmitz, C. Lang, P. Donner, U. Egner, M. A. Carrondo and B. Muller-Tiemann (2008) Structure of wild-type Plk-1 kinase domain in complex with a selective DARPIn. *Acta Crystallogr D Biol Crystallogr* 64(Pt 4): 339-353.

Baneyx, F. and M. Mujacic (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* 22(11): 1399-1408.

Barnard, E., N. V. McFerran, A. Trudgett, J. Nelson and D. J. Timson (2008) Development and implementation of split-GFP-based bimolecular fluorescence complementation (BiFC) assays in yeast. *Biochem Soc Trans* 36(Pt 3): 479-482.

Beato, M. and J. Klug (2000) Steroid hormone receptors: an update. *Hum Reprod Update* 6(3): 225-236.

Becker, G. W. and H. M. Hsiung (1986) Expression, secretion and folding of human growth hormone in *Escherichia coli*. Purification and characterization. *FEBS Lett* 204(1): 145-150.

Becker, J. and E. A. Craig (1994) Heat-shock proteins as molecular chaperones. *Eur J Biochem* 219(1-2): 11-23.

- Becker, S., A. Michalczyk, S. Wilhelm, K. E. Jaeger and H. Kolmar** (2007) Ultrahigh-throughput screening to identify *E. coli* cells expressing functionally active enzymes on their surface. *Chembiochem* 8(8): 943-949.
- Beckman, R. A., A. S. Mildvan and L. A. Loeb** (1985) On the fidelity of DNA replication: manganese mutagenesis in vitro. *Biochemistry* 24(21): 5810-7.
- Benjannet, S., D. Rhainds, R. Essalmani, J. Mayne, L. Wickham, W. Jin, M. C. Asselin, J. Hamelin, M. Varret, D. Allard, M. Trillard, M. Abifadel, A. Tebon, A. D. Attie, D. J. Rader, C. Boileau, L. Brisette, M. Chretien, A. Prat and N. G. Seidah** (2004) NARC-1/PCSK9 and its natural mutants: zymogen cleavage and effects on the low density lipoprotein (LDL) receptor and LDL cholesterol. *J Biol Chem* 279(47): 48865-48875.
- Bershtein, S. and D. S. Tawfik** (2008) Advances in laboratory evolution of enzymes. *Curr Opin Chem Biol* 12(2): 151-158.
- Biggadike, K., R. K. Bledsoe, A. M. Hassell, B. E. Kirk, I. M. McLay, L. M. Shewchuk and E. L. Stewart** (2008) X-ray crystal structure of the novel enhanced-affinity glucocorticoid agonist fluticasone furoate in the glucocorticoid receptor-ligand binding domain. *J Med Chem* 51(12): 3349-3352.
- Binz, H. K., P. Amstutz, A. Kohl, M. T. Stumpp, C. Briand, P. Forrer, M. G. Grutter and A. Plückthun** (2004) High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol* 22(5): 575-582.
- Blaber, M., X. J. Zhang and B. W. Matthews** (1993) Structural basis of amino acid alpha helix propensity. *Science* 260(5114): 1637-1640.
- Blackwell, J. R. and R. Horgan** (1991) A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Lett* 295(1-3): 10-12.
- Bledsoe, R. K., V. G. Montana, T. B. Stanley, C. J. Delves, C. J. Apolito, D. D. McKee, T. G. Consler, D. J. Parks, E. L. Stewart, T. M. Willson, M. H. Lambert, J. T. Moore, K. H. Pearce and H. E. Xu** (2002) Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. *Cell* 110(1): 93-105.
- Bosshard, H. R., D. N. Marti and I. Jelesarov** (2004) Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. *J Mol Recognit* 17(1): 1-16.
- Bowden, G. A., A. M. Paredes and G. Georgiou** (1991) Structure and morphology of protein inclusion bodies in *Escherichia coli*. *Biotechnology (N Y)* 9(8): 725-730.
- Bradford, M. M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248-254.
- Brandts, J. F. and L. N. Lin** (1990) Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry* 29(29): 6927-6940.
- Bresnick, E. H., F. C. Dalman, E. R. Sanchez and W. B. Pratt** (1989) Evidence that the 90-kDa heat shock protein is necessary for the steroid binding conformation of the L cell glucocorticoid receptor. *J Biol Chem* 264(9): 4992-4997.
- Britton, K. L., P. J. Baker, K. M. Borges, P. C. Engel, A. Pasquo, D. W. Rice, F. T. Robb, R. Scandurra, T. J. Stillman and K. S. Yip** (1995) Insights into thermal stability from a comparison of the glutamate dehydrogenases from *Pyrococcus furiosus* and *Thermococcus litoralis*. *Eur J Biochem* 229(3): 688-695.
- Bujard, H., R. Gentz, M. Lanzer, D. Stueber, M. Mueller, I. Ibrahimi, M. T. Haeuptle and B. Dobberstein** (1987) A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. *Methods Enzymol* 155: 416-433.
- Bukau, B. and A. L. Horwich** (1998) The Hsp70 and Hsp60 chaperone machines. *Cell* 92(3): 351-366.
- Cabantous, S., Y. Rogers, T. C. Terwilliger and G. S. Waldo** (2008) New molecular reporters for rapid protein folding assays. *PLoS ONE* 3(6): e2387.
- Cabantous, S., T. C. Terwilliger and G. S. Waldo** (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol* 23(1): 102-107.
- Cabantous, S. and G. S. Waldo** (2006) In vivo and in vitro protein solubility assays using split GFP. *Nat Methods* 3(10): 845-854.
- Cacace, M. G., E. M. Landau and J. J. Ramsden** (1997) The Hofmeister series: salt and solvent effects on interfacial phenomena. *Q Rev Biophys* 30(3): 241-277.

- Cameron, J., O. L. Holla, J. K. Laerdahl, M. A. Kulseth, T. Ranheim, T. Rognes, K. E. Berge and T. P. Leren** (2008) Characterization of novel mutations in the catalytic domain of the PCSK9 gene. *J Intern Med* 263(4): 420-431.
- Carrio, M. M., J. L. Corchero and A. Villaverde** (1998) Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria. *FEMS Microbiol Lett* 169(1): 9-15.
- Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher** (1994) Green fluorescent protein as a marker for gene expression. *Science* 263(5148): 802-805.
- Chan, J. C., D. E. Piper, Q. Cao, D. Liu, C. King, W. Wang, J. Tang, Q. Liu, J. Higbee, Z. Xia, Y. Di, S. Shetterly, Z. Arimura, H. Salomonis, W. G. Romanow, S. T. Thibault, R. Zhang, P. Cao, X. P. Yang, T. Yu, M. Lu, M. W. Retter, G. Kwon, K. Henne, O. Pan, M. M. Tsai, B. Fuchslocher, E. Yang, L. Zhou, K. J. Lee, M. Daris, J. Sheng, Y. Wang, W. D. Shen, W. C. Yeh, M. Emery, N. P. Walker, B. Shan, M. Schwarz and S. M. Jackson** (2009) A proprotein convertase subtilisin/kexin type 9 neutralizing antibody reduces serum cholesterol in mice and nonhuman primates. *Proc Natl Acad Sci U S A*.
- Chan, M. K., S. Mukund, A. Kletzin, M. W. Adams and D. C. Rees** (1995) Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* 267(5203): 1463-1469.
- Chandra, V., P. Huang, Y. Hamuro, S. Raghuram, Y. Wang, T. P. Burris and F. Rastinejad** (2008) Structure of the intact PPAR-gamma-RXR-alpha nuclear receptor complex on DNA. *Nature*: 350-356.
- Cheng, L., J. Fu, A. Tsukamoto and R. G. Hawley** (1996) Use of green fluorescent protein variants to monitor gene transfer and expression in mammalian cells. *Nat Biotechnol* 14(5): 606-609.
- Cherezov, V., D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka and R. C. Stevens** (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 318(5854): 1258-1265.
- Chester, N. and D. R. Marshak** (1993) Dimethyl sulfoxide-mediated primer Tm reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Anal Biochem* 209(2): 284-290.
- Chopra, A. K., A. R. Brasier, M. Das, X. J. Xu and J. W. Peterson** (1994) Improved synthesis of Salmonella typhimurium enterotoxin using gene fusion expression systems. *Gene* 144(1): 81-85.
- Chyan, C. L., C. Wormald, C. M. Dobson, P. A. Evans and J. Baum** (1993) Structure and stability of the molten globule state of guinea-pig alpha-lactalbumin: a hydrogen exchange study. *Biochemistry* 32(21): 5681-5691.
- Cline, J., J. C. Braman and H. H. Hogrefe** (1996) PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res* 24(18): 3546-3551.
- Cohen, J., A. Pertsemlidis, I. K. Kotowski, R. Graham, C. K. Garcia and H. H. Hobbs** (2005) Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 37(2): 161-165.
- Cormack, B. P., R. H. Valdivia and S. Falkow** (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173(1 Spec No): 33-38.
- Cornvik, T., S. L. Dahlroth, A. Magnusdottir, M. D. Herman, R. Knaust, M. Ekberg and P. Nordlund** (2005) Colony filtration blot: a new screening method for soluble protein expression in Escherichia coli. *Nat Methods* 2(7): 507-509.
- Cramer, A., E. A. Whitehorn, E. Tate and W. P. Stemmer** (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol* 14(3): 315-319.
- Cubitt, A. B., R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross and R. Y. Tsien** (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* 20(11): 448-455.
- Cunningham, D., D. E. Danley, K. F. Geoghegan, M. C. Griffor, J. L. Hawkins, T. A. Subashi, A. H. Varghese, M. J. Ammirati, J. S. Culp, L. R. Hoth, M. N. Mansour, K. M. McGrath, A. P. Seddon, S. Shenolikar, K. J. Stutzman-Engwall, L. C. Warren, D. Xia and X. Qiu** (2007) Structural and biophysical studies of PCSK9 and its mutants linked to familial hypercholesterolemia. *Nat Struct Mol Biol* 14(5): 413-419.
- Dahlman-Wright, K., H. Baumann, I. J. McEwan, T. Almlof, A. P. Wright, J. A. Gustafsson and T. Hard** (1995) Structural characterization of a minimal functional transactivation domain from the human glucocorticoid receptor. *Proc Natl Acad Sci U S A* 92(5): 1699-1703.
- Dahlroth, S. L., P. Nordlund and T. Cornvik** (2006) Colony filtration blotting for screening soluble expression in Escherichia coli. *Nat Protoc* 1(1): 253-258.

- Dale, G. E., C. Broger, H. Langen, A. D'Arcy and D. Stüber** (1994) Improving protein solubility through rationally designed amino acid replacements: solubilization of the trimethoprim-resistant type S1 dihydrofolate reductase. *Protein Eng* 7(7): 933-939.
- Danley, D. E.** (2006) Crystallization to obtain protein-ligand complexes for structure-aided drug design. *Acta Crystallogr D Biol Crystallogr* 62(Pt 6): 569-575.
- Das, D. and M. M. Georgiadis** (2001) A directed approach to improving the solubility of Moloney murine leukemia virus reverse transcriptase. *Protein Sci* 10(10): 1936-1941.
- Daugherty, P. S., G. Chen, B. L. Iverson and G. Georgiou** (2000) Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. *Proc Natl Acad Sci U S A* 97(5): 2029-2034.
- Day, P. W., S. G. Rasmussen, C. Parnot, J. J. Fung, A. Masood, T. S. Kobilka, X. J. Yao, H. J. Choi, W. I. Weis, D. K. Rohrer and B. K. Kobilka** (2007) A monoclonal antibody for G protein-coupled receptor crystallography. *Nat Methods* 4(11): 927-929.
- De Giorgi, F., M. Brini, C. Bastianutto, R. Marsault, M. Montero, P. Pizzo, R. Rossi and R. Rizzuto** (1996) Targeting aequorin and green fluorescent protein to intracellular organelles. *Gene* 173(1 Spec No): 113-117.
- Derewenda, Z. S. and P. G. Vekilov** (2006) Entropy and surface engineering in protein crystallization. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1): 116-124.
- Dieken, E. S. and R. L. Miesfeld** (1992) Transcriptional transactivation functions localized to the glucocorticoid receptor N terminus are necessary for steroid induction of lymphocyte apoptosis. *Mol Cell Biol* 12(2): 589-597.
- Dill, K. A.** (1990) Dominant forces in protein folding. *Biochemistry* 29(31): 7133-7155.
- Dittmar, K. D., D. R. Demady, L. F. Stancato, P. Krishna and W. B. Pratt** (1997) Folding of the glucocorticoid receptor by the heat shock protein (hsp) 90-based chaperone machinery. The role of p23 is to stabilize receptor.hsp90 heterocomplexes formed by hsp90.p60.hsp70. *J Biol Chem* 272(34): 21213-21220.
- Dittmar, K. D. and W. B. Pratt** (1997) Folding of the glucocorticoid receptor by the reconstituted Hsp90-based chaperone machinery. The initial hsp90.p60.hsp70-dependent step is sufficient for creating the steroid binding conformation. *J Biol Chem* 272(20): 13047-13054.
- Dobson, C. M.** (1994) Protein folding. Solid evidence for molten globules. *Curr Biol* 4(7): 636-640.
- Dolgikh, D. A., L. V. Abaturov, Brazhnikov, O. Lebedev Iu and N. Chirgadze Iu** (1983) [Acid form of carbonic anhydrase: "molten globule" with a secondary structure]. *Dokl Akad Nauk SSSR* 272(6): 1481-1484.
- Dolgikh, D. A., R. I. Gilmanishin, E. V. Brazhnikov, V. E. Bychkova, G. V. Semisotnov, S. Venyaminov and O. B. Ptitsyn** (1981) Alpha-Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett* 136(2): 311-315.
- Dower, W. J., J. F. Miller and C. W. Ragsdale** (1988) High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res* 16(13): 6127-6145.
- Drew, D. E., G. von Heijne, P. Nordlund and J. W. de Gier** (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. *FEBS Lett* 507(2): 220-224.
- Drummond, D. A., B. L. Iverson, G. Georgiou and F. H. Arnold** (2005) Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J Mol Biol* 350(4): 806-816.
- Durbin, S. D. and G. Feher** (1996) Protein crystallization. *Annu Rev Phys Chem* 47: 171-204.
- Dyda, F., A. B. Hickman, T. M. Jenkins, A. Engelman, R. Craigie and D. R. Davies** (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* 266(5193): 1981-1986.
- Eckert, K. A. and T. A. Kunkel** (1990) High fidelity DNA synthesis by the Thermus aquaticus DNA polymerase. *Nucleic Acids Res* 18(13): 3739-3744.
- Eijsink, V. G., A. Bjork, S. Gaseidnes, R. Sirevag, B. Synstad, B. van den Burg and G. Vriend** (2004) Rational engineering of enzyme stability. *J Biotechnol* 113(1-3): 105-120.
- Eijsink, V. G., S. Gaseidnes, T. V. Borchert and B. van den Burg** (2005) Directed evolution of enzyme stability. *Biomol Eng* 22(1-3): 21-30.
- Elcock, A. H.** (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J Mol Biol* 284(2): 489-502.

- Elleby, B., S. Svensson, X. Wu, K. Stefansson, J. Nilsson, D. Hallen, U. Oppermann and L. Abrahmsen** (2004) High-level production and optimization of monodispersity of 11 β -hydroxysteroid dehydrogenase type 1. *Biochim Biophys Acta* 1700(2): 199-207.
- Ellis, R. J. and A. P. Minton** (2003) Cell biology: join the crowd. *Nature* 425(6953): 27-28.
- Fahey, R. C., J. S. Hunt and G. C. Windham** (1977) On the cysteine and cystine content of proteins. Differences between intracellular and extracellular proteins. *J Mol Evol* 10(2): 155-160.
- Fang, L., D. Ricketson, L. Getubig and B. Darimont** (2006) Unliganded and hormone-bound glucocorticoid receptors interact with distinct hydrophobic sites in the Hsp90 C-terminal domain. *Proc Natl Acad Sci U S A* 103(49): 18487-18492.
- Fasan, R., Y. T. Meharena, C. D. Snow, T. L. Poulos and F. H. Arnold** (2008) Evolutionary history of a specialized p450 propane monooxygenase. *J Mol Biol* 383(5): 1069-1080.
- Frigerio, F., I. Margarit, R. Nogarotto, V. de Filippis and G. Grandi** (1996) Cumulative stabilizing effects of hydrophobic interactions on the surface of the neutral protease from *Bacillus subtilis*. *Protein Eng* 9(5): 439-445.
- Fromant, M., S. Blanquet and P. Plateau** (1995) Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem* 224(1): 347-353.
- Galon, J., D. Franchimont, N. Hiroi, G. Frey, A. Boettner, M. Ehrhart-Bornstein, J. J. O'Shea, G. P. Chrousos and S. R. Bornstein** (2002) Gene profiling reveals unknown enhancing and suppressive actions of glucocorticoids on immune cells. *FASEB J* 16(1): 61-71.
- Garcia-Fruitos, E., M. Martinez-Alonso, N. Gonzalez-Montalban, M. Valli, D. Mattanovich and A. Villaverde** (2007) Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. *J Mol Biol* 374(1): 195-205.
- Garuti, R., C. Jones, W. P. Li, P. Michaely, J. Herz, R. D. Gerard, J. C. Cohen and H. H. Hobbs** (2005) The modular adaptor protein autosomal recessive hypercholesterolemia (ARH) promotes low density lipoprotein receptor clustering into clathrin-coated pits. *J Biol Chem* 280(49): 40996-41004.
- Georgiou, G. and P. Valax** (1996) Expression of correctly folded proteins in *Escherichia coli*. *Curr Opin Biotechnol* 7(2): 190-197.
- Giguere, V., S. M. Hollenberg, M. G. Rosenfeld and R. M. Evans** (1986) Functional domains of the human glucocorticoid receptor. *Cell* 46(5): 645-652.
- Golan D. E.** (2005) Principles of Pharmacology. The Pathophysiologic Basis of Drug Therapy. 2nd Edition. Chapter 27 - Pharmacology of the Adrenal Cortex.
- Goldstein, J. L. and M. S. Brown** (1973) Familial hypercholesterolemia: identification of a defect in the regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity associated with overproduction of cholesterol. *Proc Natl Acad Sci U S A* 70(10): 2804-2808.
- Goldstein, J. L. and M. S. Brown** (2001) Molecular medicine. The cholesterol quartet. *Science* 292(5520): 1310-1312.
- Goloubinoff, P., A. A. Gatenby and G. H. Lorimer** (1989) GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in *Escherichia coli*. *Nature* 337(6202): 44-47.
- Graham, M. J., K. M. Lemonidis, C. P. Whipple, A. Subramaniam, B. P. Monia, S. T. Crooke and R. M. Crooke** (2007) Antisense inhibition of proprotein convertase subtilisin/kexin type 9 reduces serum LDL in hyperlipidemic mice. *J Lipid Res* 48(4): 763-767.
- Grebenok, R. J., E. Pierson, G. M. Lambert, F. C. Gong, C. L. Afonso, R. Haldeman-Cahill, J. C. Carrington and D. W. Galbraith** (1997) Green-fluorescent protein fusions for efficient characterization of nuclear targeting. *Plant J* 11(3): 573-586.
- Gribenko, A. V., M. M. Patel, J. Liu, S. A. McCallum, C. Wang and G. I. Makhatadze** (2009) Rational stabilization of enzymes by computational redesign of surface charge-charge interactions. *Proc Natl Acad Sci U S A* 106(8): 2601-2606.
- Griswold, K. E., N. S. Aiyappan, B. L. Iverson and G. Georgiou** (2006) The evolution of catalytic efficiency and substrate promiscuity in human theta class 1-1 glutathione transferase. *J Mol Biol* 364(3): 400-410.
- Griswold, K. E., Y. Kawarasaki, N. Ghoneim, S. J. Benkovic, B. L. Iverson and G. Georgiou** (2005) Evolution of highly active enzymes by homology-independent recombination. *Proc Natl Acad Sci U S A* 102(29): 10082-10087.

- Gustafsson, E., C. Forsberg, K. Haraldsson, S. Lindman, L. Ljung and C. Furebring** (2009) Purification of truncated and mutated Chemotaxis Inhibitory Protein of *Staphylococcus aureus*--an anti-inflammatory protein. *Protein Expr Purif* 63(2): 95-101.
- Hagihara, Y., S. Mine and K. Uegaki** (2007) Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region. *J Biol Chem* 282(50): 36489-36495.
- Haigh, S. E., G. Twig, A. A. Molina, J. D. Wikstrom, M. Deutsch and O. S. Shirihai** (2007) PA-GFP: a window into the subcellular adventures of the individual mitochondrion. *Novartis Found Symp* 287: 21-36; discussion 36-46.
- Hampton, E. N., M. W. Knuth, J. Li, J. L. Harris, S. A. Lesley and G. Spraggon** (2007) The self-inhibited structure of full-length PCSK9 at 1.9 Å reveals structural homology with resistin within the C-terminal domain. *Proc Natl Acad Sci U S A* 104(37): 14604-14609.
- Hanahan, D.** (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166(4): 557-580.
- Hanakam, F., R. Albrecht, C. Eckerskorn, M. Matzner and G. Gerisch** (1996) Myristoylated and non-myristoylated forms of the pH sensor protein hisactophilin II: intracellular shuttling to plasma membrane and nucleus monitored in real time by a fusion with green fluorescent protein. *EMBO J* 15(12): 2935-2943.
- Hard, T., E. Kellenbach, R. Boelens, B. A. Maler, K. Dahlman, L. P. Freedman, J. Carlstedt-Duke, K. R. Yamamoto, J. A. Gustafsson and R. Kaptein** (1990) Solution structure of the glucocorticoid receptor DNA-binding domain. *Science* 249(4965): 157-160.
- Hassell, A. M., G. An, R. K. Bledsoe, J. M. Bynum, H. L. Carter, 3rd, S. J. Deng, R. T. Gampe, T. E. Grisard, K. P. Madauss, R. T. Nolte, W. J. Rocque, L. Wang, K. L. Weaver, S. P. Williams, G. B. Wisely, R. Xu and L. M. Shewchuk** (2007) Crystallization of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 63(Pt 1): 72-79.
- Hauptmann, P. and L. Lehle** (2008) Kex1 protease is involved in yeast cell death induced by defective N-glycosylation, acetic acid, and chronological aging. *J Biol Chem* 283(27): 19151-19163.
- Heery, D. M., E. Kalkhoven, S. Hoare and M. G. Parker** (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* 387(6634): 733-736.
- Heim, R., A. B. Cubitt and R. Y. Tsien** (1995) Improved green fluorescence. *Nature* 373(6516): 663-664.
- Heim, R., D. C. Prasher and R. Y. Tsien** (1994) Wavelength mutations and posttranslational autooxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* 91(26): 12501-12504.
- Heim, R. and R. Y. Tsien** (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol* 6(2): 178-182.
- Heitzer, M. D., I. M. Wolf, E. R. Sanchez, S. F. Witchel and D. B. DeFranco** (2007) Glucocorticoid receptor physiology. *Rev Endocr Metab Disord* 8(4): 321-330.
- Hendsch, Z. S. and B. Tidor** (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci* 3(2): 211-226.
- Hengel, R. L. and J. K. Nicholson** (2001) An update on the use of flow cytometry in HIV infection and AIDS. *Clin Lab Med* 21(4): 841-856.
- Hennig, M., R. Sterner, K. Kirschner and J. N. Jansonius** (1997) Crystal structure at 2.0 Å resolution of phosphoribosyl anthranilate isomerase from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability. *Biochemistry* 36(20): 6009-6016.
- Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen and L. R. Pease** (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77(1): 51-59.
- Höcker, B., J. Claren and R. Sterner** (2004) Mimicking enzyme evolution by generating new $(\beta\alpha)_8$ -barrels from $(\beta\alpha)_4$ -half-barrels. *Proc Natl Acad Sci U S A* 101(47): 16448-16453.
- Höcker, B., A. Lochner, T. Seitz, J. Claren and R. Sterner** (2009) High-Resolution Crystal Structure of an Artificial $(\beta\alpha)_8$ -Barrel Protein Designed from Identical Half-Barrels. *Biochemistry* 48 (6): 1145-1147.
- Hockney, R. C.** (1994) Recent developments in heterologous protein production in *Escherichia coli*. *Trends Biotechnol* 12(11): 456-463.

- Horvath, G., S. Young and E. Latz** (2009) Toll-like receptor interactions imaged by FRET microscopy and GFP fragment reconstitution. *Methods Mol Biol* 517: 33-54.
- Horwich, A. L.** (2004) Chaperoned protein disaggregation--the ClpB ring uses its central channel. *Cell* 119(5): 579-581.
- Hoseki, J., T. Yano, Y. Koyama, S. Kuramitsu and H. Kagamiyama** (1999) Directed evolution of thermostable kanamycin-resistance gene: a convenient selection marker for *Thermus thermophilus*. *J Biochem* 126(5): 951-956.
- Hu, C. D., Y. Chinenov and T. K. Kerppola** (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol Cell* 9(4): 789-798.
- Hu, C. D., A. V. Grinberg and T. K. Kerppola** (2006) Visualization of protein interactions in living cells using bimolecular fluorescence complementation (BiFC) analysis. *Curr Protoc Cell Biol* Chapter 21: Unit 21 23.
- Hu, C. D. and T. K. Kerppola** (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat Biotechnol* 21(5): 539-545.
- Humphreys, D. P., S. P. Heywood, L. M. King, L. C. Bowering, J. P. Turner and S. E. Lane** (2004) Engineering of *Escherichia coli* to improve the purification of periplasmic Fab' fragments: changing the pI of the chromosomally encoded PhoS/PstS protein. *Protein Expr Purif* 37(1): 109-118.
- Hunte, C., J. Koepke, C. Lange, T. Rossmannith and H. Michel** (2000) Structure at 2.3 Å resolution of the cytochrome bc(1) complex from the yeast *Saccharomyces cerevisiae* co-crystallized with an antibody Fv fragment. *Structure* 8(6): 669-684.
- Ibrahim, S. F. and G. van den Engh** (2007) Flow cytometry and cell sorting. *Adv Biochem Eng Biotechnol* 106: 19-39.
- Ikemura, H., H. Takagi and M. Inouye** (1987) Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli*. *J Biol Chem* 262(16): 7859-7864.
- Inoue, H., H. Nojima and H. Okayama** (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96(1): 23-28.
- Jäckel, C., P. Kast and D. Hilvert** (2008) Protein design by directed evolution. *Annu Rev Biophys* 37: 153-173.
- Jaenicke, R.** (1999) Stability and folding of domain proteins. *Prog Biophys Mol Biol* 71(2): 155-241.
- Jenkins, T. M., A. B. Hickman, F. Dyda, R. Ghirlando, D. R. Davies and R. Craigie** (1995) Catalytic domain of human immunodeficiency virus type 1 integrase: identification of a soluble mutant by systematic replacement of hydrophobic residues. *Proc Natl Acad Sci U S A* 92(13): 6057-6061.
- Jiang, Y., A. Lee, J. Chen, V. Ruta, M. Cadene, B. T. Chait and R. MacKinnon** (2003) X-ray structure of a voltage-dependent K⁺ channel. *Nature* 423(6935): 33-41.
- Johnson, R. L.** (1993) Flow cytometry. From research to clinical laboratory applications. *Clin Lab Med* 13(4): 831-852.
- Johnson, W. C., Jr.** (1990) Protein secondary structure and circular dichroism: a practical guide. *Proteins* 7(3): 205-214.
- Jones, P. L. and Y. B. Shi** (2003) N-CoR-HDAC corepressor complexes: roles in transcriptional regulation by nuclear hormone receptors. *Curr Top Microbiol Immunol* 274: 237-268.
- Jürgens, C., A. Strom, D. Wegener, S. Hettwer, M. Wilmanns and R. Sterner** (2000) Directed evolution of a (beta alpha)₈-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc Natl Acad Sci U S A* 97(18): 9925-9930.
- Jung-Testas, I. and E. E. Baulieu** (1983) Inhibition of glucocorticosteroid action in cultured L-929 mouse fibroblasts by RU 486, a new anti-glucocorticosteroid of high affinity for the glucocorticosteroid receptor. *Exp Cell Res* 147(1): 177-182.
- Kaneko, H., H. Minagawa and J. Shimada** (2005) Rational design of thermostable lactate oxidase by analyzing quaternary structure and prevention of deamidation. *Biotechnol Lett* 27(22): 1777-1784.
- Karshikoff, A. and R. Ladenstein** (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem Sci* 26(9): 550-556.
- Kather, I., R. P. Jakob, H. Dobbek and F. X. Schmid** (2008) Increased folding stability of TEM-1 beta-lactamase by in vitro selection. *J Mol Biol* 383(1): 238-251.

- Kauppi, B., C. Jakob, M. Farnegardh, J. Yang, H. Ahola, M. Alarcon, K. Calles, O. Engstrom, J. Harlan, S. Muchmore, A. K. Ramqvist, S. Thorell, L. Ohman, J. Greer, J. A. Gustafsson, J. Carlstedt-Duke and M. Carlquist** (2003) The three-dimensional structures of antagonistic and agonistic forms of the glucocorticoid receptor ligand-binding domain: RU-486 induces a transconformation that leads to active antagonism. *J Biol Chem* 278(25): 22748-22754.
- Kauzmann, W.** (1959) Some factors in the interpretation of protein denaturation. *Adv Protein Chem* 14: 1-63.
- Kawasaki, M. and F. Inagaki** (2001) Random PCR-based screening for soluble domains using green fluorescent protein. *Biochem Biophys Res Commun* 280(3): 842-844.
- Ke, A. and C. Wolberger** (2003) Insights into binding cooperativity of MATA1/MATalpha2 from the crystal structure of a MATA1 homeodomain-maltose binding protein chimera. *Protein Sci* 12(2): 306-312.
- Kelly, C. A., M. Nishiyama, Y. Ohnishi, T. Beppu and J. J. Birktoft** (1993) Determinants of protein thermostability observed in the 1.9-A crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*. *Biochemistry* 32(15): 3913-3922.
- Kino, T., M. U. De Martino, E. Charmandari, M. Mirani and G. P. Chrousos** (2003) Tissue glucocorticoid resistance/hypersensitivity syndromes. *J Steroid Biochem Mol Biol* 85(2-5): 457-467.
- Knapp, S., R. Ladenstein and E. A. Galinski** (1999) Extrinsic protein stabilization by the naturally occurring osmolytes beta-hydroxyectoine and betaine. *Extremophiles* 3(3): 191-198.
- Kobe, B., R. J. Center, B. E. Kemp and P. Pombourios** (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *Proc Natl Acad Sci U S A* 96(8): 4319-4324.
- Kohl, A., P. Amstutz, P. Parizek, H. K. Binz, C. Briand, G. Capitani, P. Forrer, A. Plückthun and M. G. Grutter** (2005) Allosteric inhibition of aminoglycoside phosphotransferase by a designed ankyrin repeat protein. *Structure* 13(8): 1131-1141.
- Koide, A. and S. Koide** (2007) Monobodies: antibody mimics based on the scaffold of the fibronectin type III domain. *Methods Mol Biol* 352: 95-109.
- Korndörfer, I., B. Steipe, R. Huber, A. Tomschy and R. Jaenicke** (1995) The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J Mol Biol* 246(4): 511-521.
- Kumar, R., I. V. Baskakov, G. Srinivasan, D. W. Bolen, J. C. Lee and E. B. Thompson** (1999) Interdomain signaling in a two-domain fragment of the human glucocorticoid receptor. *J Biol Chem* 274(35): 24737-24741.
- Kumar, R., R. Betney, J. Li, E. B. Thompson and I. J. McEwan** (2004) Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIIF. *Biochemistry* 43(11): 3008-3013.
- Kumar, R. and E. B. Thompson** (2005) Gene regulation by the glucocorticoid receptor: structure: function relationship. *J Steroid Biochem Mol Biol* 94(5): 383-394.
- Kumar, R., D. E. Volk, J. Li, J. C. Lee, D. G. Gorenstein and E. B. Thompson** (2004) TATA box binding protein induces structure in the recombinant glucocorticoid receptor AF1 domain. *Proc Natl Acad Sci U S A* 101(47): 16425-16430.
- Kuwajima, K.** (1996) The molten globule state of alpha-lactalbumin. *FASEB J* 10(1): 102-109.
- Kuwajima, K., K. Nitta, M. Yoneyama and S. Sugai** (1976) Three-state denaturation of alpha-lactalbumin by guanidine hydrochloride. *J Mol Biol* 106(2): 359-373.
- Kwon, H. J., T. A. Lagace, M. C. McNutt, J. D. Horton and J. Deisenhofer** (2008) Molecular basis for LDL receptor recognition by PCSK9. *Proc Natl Acad Sci U S A* 105(6): 1820-1825.
- Laemmli, U. K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227(5259): 680-685.
- Lagace, T. A., D. E. Curtis, R. Garuti, M. C. McNutt, S. W. Park, H. B. Prather, N. N. Anderson, Y. K. Ho, R. E. Hammer and J. D. Horton** (2006) Secreted PCSK9 decreases the number of LDL receptors in hepatocytes and in livers of parabiotic mice. *J Clin Invest* 116(11): 2995-3005.
- Lambert, G., F. Charlton, K. A. Rye and D. E. Piper** (2009) Molecular basis of PCSK9 function. *Atherosclerosis* 203(1): 1-7.

- LaVallie, E. R., Z. Lu, E. A. Diblasio-Smith, L. A. Collins-Racie and J. M. McCoy** (2000) Thioredoxin as a fusion partner for production of soluble recombinant proteins in *Escherichia coli*. *Methods Enzymol* 326: 322-340.
- Le Gouill, C., J. L. Parent, M. Rola-Pleszczynski and J. Stankova** (1994) Analysis of recombinant plasmids by a modified alkaline lysis method. *Anal Biochem* 219(1): 164.
- Lehmann, M., C. Loch, A. Middendorf, D. Studer, S. F. Lassen, L. Pasamontes, A. P. van Loon and M. Wyss** (2002) The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng* 15(5): 403-411.
- Lesley, S. A., J. Graziano, C. Y. Cho, M. W. Knuth and H. E. Klock** (2002) Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Eng* 15(2): 153-160.
- Leung D.W., Chen E., Goeddel D.V.** (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* 1 (1): 11-15.
- Li, J., C. Tumanut, J. A. Gavigan, W. J. Huang, E. N. Hampton, R. Tumanut, K. F. Suen, J. W. Trauger, G. Spraggon, S. A. Lesley, G. Liao, D. Yowe and J. L. Harris** (2007) Secreted PCSK9 promotes LDL receptor degradation independently of proteolytic activity. *Biochem J* 406(2): 203-207.
- Li, P., H. Guan, J. Li and Z. Lin** (2009) Heterologous expression, purification, and characterization of cytochrome P450sca-2 and mutants with improved solubility in *Escherichia coli*. *Protein Expr Purif* 65(2): 196-203.
- Lipovsek, D., E. Antipov, K. A. Armstrong, M. J. Olsen, A. M. Klibanov, B. Tidor and K. D. Wittrup** (2007) Selection of horseradish peroxidase variants with enhanced enantioselectivity by yeast surface display. *Chem Biol* 14(10): 1176-1185.
- Listwan, P., T. C. Terwilliger and G. S. Waldo** (2009) Automated, high-throughput platform for protein solubility screening using a split-GFP system. *J Struct Funct Genomics* 10(1): 47-55.
- Littlechild, J. A., J. Guy, S. Connelly, L. Mallett, S. Waddell, C. A. Rye, K. Line and M. Isupov** (2007) Natural methods of protein stabilization: thermostable biocatalysts. *Biochem Soc Trans* 35(Pt 6): 1558-1563.
- Liu, Y., A. Manna, R. Li, W. E. Martin, R. C. Murphy, A. L. Cheung and G. Zhang** (2001) Crystal structure of the SarR protein from *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 98(12): 6877-6882.
- Llopis, J., J. M. McCaffery, A. Miyawaki, M. G. Farquhar and R. Y. Tsien** (1998) Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc Natl Acad Sci U S A* 95(12): 6803-6808.
- Lochner, A.** (2008) Stabilisierung und Aktivierung eines künstlichen $(\beta\alpha)_8$ -Fass Proteins. Diplomarbeit, Universität Regensburg.
- Loinder, K. and M. Soderstrom** (2004) Functional analyses of an LXXLL motif in nuclear receptor corepressor (N-CoR). *J Steroid Biochem Mol Biol* 91(4-5): 191-196.
- Lonard, D. M. and B. W. O'Malley** (2006) The expanding cosmos of nuclear receptor coactivators. *Cell* 125(3): 411-414.
- Lu, N. Z. and J. A. Cidlowski** (2006) Glucocorticoid receptor isoforms generate transcription specificity. *Trends Cell Biol* 16(6): 301-307.
- Luisi, B. F., W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto and P. B. Sigler** (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* 352(6335): 497-505.
- Lund, P. A.** (2001) Microbial molecular chaperones. *Adv Microb Physiol* 44: 93-140.
- Madauss, K. P., R. K. Bledsoe, I. McLay, E. L. Stewart, I. J. Uings, G. Weingarten and S. P. Williams** (2008) The first X-ray crystal structure of the glucocorticoid receptor bound to a non-steroidal agonist. *Bioorg Med Chem Lett* 18(23): 6097-6099.
- Makhatadze, G. I., V. V. Loladze, D. N. Ermolenko, X. Chen and S. T. Thomas** (2003) Contribution of surface salt bridges to protein stability: guidelines for protein engineering. *J Mol Biol* 327(5): 1135-1148.
- Makrides, S. C.** (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* 60(3): 512-538.
- Malissard, M. and E. G. Berger** (2001) Improving solubility of catalytic domain of human beta-1,4-galactosyltransferase 1 through rationally designed amino acid replacements. *Eur J Biochem* 268(15): 4352-4358.

- Mandaji, M., G. Rubensam, R. B. Hoff, S. Hillebrand, E. Carrilho and T. L. Kist** (2009) Sample stacking in CZE using dynamic thermal junctions I. Analytes with low dpKa/dT crossing a single thermally induced pH junction in a BGE with high dpH/dT. *Electrophoresis* 30(9): 1501-1509.
- Mandy, F. F.** (2004) Twenty-five years of clinical flow cytometry: AIDS accelerated global instrument distribution. *Cytometry A* 58(1): 55-56.
- Mansfeld, J., G. Vriend, B. W. Dijkstra, O. R. Veltman, B. Van den Burg, G. Venema, R. Ulbrich-Hofmann and V. G. Eijsink** (1997) Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J Biol Chem* 272(17): 11152-11156.
- Martin, A., F. X. Schmid and V. Sieber** (2003) Proside: a phage-based method for selecting thermostable proteins. *Methods Mol Biol* 230: 57-70.
- Martin, A., V. Sieber and F. X. Schmid** (2001) In-vitro selection of highly stabilized protein variants with optimized surface. *J Mol Biol* 309(3): 717-726.
- Mastrobattista, E., V. Taly, E. Chanudet, P. Treacy, B. T. Kelly and A. D. Griffiths** (2005) High-throughput screening of enzyme libraries: in vitro evolution of a beta-galactosidase by fluorescence-activated sorting of double emulsions. *Chem Biol* 12(12): 1291-1300.
- Matulis, D., J. K. Kranz, F. R. Salemme and M. J. Todd** (2005) Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* 44(13): 5258-5266.
- Matsumura, M., G. Signor and B. W. Matthews** (1989) Substantial increase of protein stability by multiple disulphide bonds. *Nature* 342(6247): 291-293.
- Matthews, B. W., H. Nicholson and W. J. Becktel** (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci U S A* 84(19): 6663-6667.
- Maxwell, K. L., A. K. Mittermaier, J. D. Forman-Kay and A. R. Davidson** (1999) A simple in vivo assay for increased protein solubility. *Protein Sci* 8(9): 1908-1911.
- McEwan, I. J., A. P. Wright, K. Dahlman-Wright, J. Carlstedt-Duke and J. A. Gustafsson** (1993) Direct interaction of the tau 1 transactivation domain of the human glucocorticoid receptor with the basal transcriptional machinery. *Mol Cell Biol* 13(1): 399-407.
- McNutt, M. C., T. A. Lagace and J. D. Horton** (2007) Catalytic activity is not required for secreted PCSK9 to reduce low density lipoprotein receptors in HepG2 cells. *J Biol Chem* 282(29): 20799-20803.
- McTigue, M. A., D. R. Williams and J. A. Tainer** (1995) Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from *Schistosoma japonica* and its complex with the leading antischistosomal drug praziquantel. *J Mol Biol* 246(1): 21-27.
- Merz, A., M. C. Yee, H. Szadkowski, G. Pappenberger, A. Cramer, W. P. Stemmer, C. Yanofsky and K. Kirschner** (2000) Improving the catalytic activity of a thermophilic enzyme at low temperatures. *Biochemistry* 39(5): 880-889.
- Miller, O. J., K. Bernath, J. J. Agresti, G. Amitai, B. T. Kelly, E. Mastrobattista, V. Taly, S. Magdassi, D. S. Tawfik and A. D. Griffiths** (2006) Directed evolution by in vitro compartmentalization. *Nat Methods* 3(7): 561-570.
- Miyawaki, A., O. Griesbeck, R. Heim and R. Y. Tsien** (1999) Dynamic and quantitative Ca²⁺ measurements using improved cameleons. *Proc Natl Acad Sci U S A* 96(5): 2135-2140.
- Moore, J. C. and F. H. Arnold** (1996) Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat Biotechnol* 14(4): 458-467.
- Morise, H., O. Shimomura, F. H. Johnson and J. Winant** (1974) Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13(12): 2656-2662.
- Mosavi, L. K. and Z. Y. Peng** (2003) Structure-based substitutions for increased solubility of a designed protein. *Protein Eng* 16(10): 739-745.
- Muldoon, R. R., J. P. Levy, S. R. Kain, P. A. Kitts and C. J. Link, Jr.** (1997) Tracking and quantitation of retroviral-mediated transfer using a completely humanized, red-shifted green fluorescent protein gene. *Biotechniques* 22(1): 162-167.
- Mullis, K. B. and F. A. Faloona** (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* 155: 335-350.
- Myers, J. K. and C. N. Pace** (1996) Hydrogen bonding stabilizes globular proteins. *Biophys J* 71(4): 2033-2039.

- Naureckiene, S., L. Ma, K. Sreekumar, U. Purandare, C. F. Lo, Y. Huang, L. W. Chiang, J. M. Grenier, B. A. Ozenberger, J. S. Jacobsen, J. D. Kennedy, P. S. DiStefano, A. Wood and B. Bingham** (2003) Functional characterization of Narc 1, a novel proteinase related to proteinase K. *Arch Biochem Biophys* 420(1): 55-67.
- Neylon, C.** (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res* 32(4): 1448-1459.
- Nicholson, H., D. E. Anderson, S. Dao-pin and B. W. Matthews** (1991) Analysis of the interaction between charged side chains and the alpha-helix dipole using designed thermostable mutants of phage T4 lysozyme. *Biochemistry* 30(41): 9816-9828.
- Nicholson, H., W. J. Becktel and B. W. Matthews** (1988) Enhanced protein thermostability from designed mutations that interact with alpha-helix dipoles. *Nature* 336(6200): 651-656.
- Nishida, M., M. Cadene, B. T. Chait and R. MacKinnon** (2007) Crystal structure of a Kir3.1-prokaryotic Kir channel chimera. *EMBO J* 26(17): 4005-4015.
- Nozaka, M., K. Kuwajima, K. Nitta and S. Sugai** (1978) Detection and characterization of the intermediate on the folding pathway of human alpha-lactalbumin. *Biochemistry* 17(18): 3753-3758.
- O'Neil, K. T. and W. F. DeGrado** (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* 250(4981): 646-651.
- Ostermeier, C., S. Iwata, B. Ludwig and H. Michel** (1995) Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. *Nat Struct Biol* 2(10): 842-846.
- Pace, C. N.** (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131: 266-280.
- Pace, C. N.** (1990) Conformational stability of globular proteins. *Trends Biochem Sci* 15(1): 14-17.
- Pace, C. N., R. W. Alston and K. L. Shaw** (2000) Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci* 9(7): 1395-1398.
- Pace, C. N., E. J. Hebert, K. L. Shaw, D. Schell, V. Both, D. Krajeikova, J. Sevcik, K. S. Wilson, Z. Dauter, R. W. Hartley and G. R. Grimsley** (1998) Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. *J Mol Biol* 279(1): 271-286.
- Pace, C. N., B. A. Shirley, M. McNutt and K. Gajiwala** (1996) Forces contributing to the conformational stability of proteins. *FASEB J* 10(1): 75-83.
- Pace C.N. and Scholtz J.M.** (1997) Measuring the conformational stability of a protein. *In* Proofs for Protein Structure: A practical approach. Chapter 12, pp 299-321. Creighton T.E., IRL Press, Oxford, 2ed.
- Pace, C. N., S. Trevino, E. Prabhakaran and J. M. Scholtz** (2004) Protein structure, stability and solubility in water and other solvents. *Philos Trans R Soc Lond B Biol Sci* 359(1448): 1225-1234; discussion 1234-1225.
- Pace, C. N., F. Vajdos, L. Fee, G. Grimsley and T. Gray** (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* 4(11): 2411-2423.
- Pakula, A. A. and R. T. Sauer** (1990) Reverse hydrophobic effects relieved by amino-acid substitutions at a protein surface. *Nature* 344(6264): 363-364.
- Park, S. W., Y. A. Moon and J. D. Horton** (2004) Post-transcriptional regulation of low density lipoprotein receptor protein by proprotein convertase subtilisin/kexin type 9a in mouse liver. *J Biol Chem* 279(48): 50630-50638.
- Parsell, D. A., A. S. Kowal, M. A. Singer and S. Lindquist** (1994) Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* 372(6505): 475-478.
- Patterson, G. H., S. M. Knobel, W. D. Sharif, S. R. Kain and D. W. Piston** (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* 73(5): 2782-2790.
- Patterson, G. H. and J. Lippincott-Schwartz** (2004) Selective photolabeling of proteins using photoactivatable GFP. *Methods* 32(4): 445-450.
- Pedelacq, J. D., S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo** (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 24(1): 79-88.

- Pedelacq, J. D., E. Piltch, E. C. Liang, J. Berendzen, C. Y. Kim, B. S. Rho, M. S. Park, T. C. Terwilliger and G. S. Waldo** (2002) Engineering soluble proteins for structural genomics. *Nat Biotechnol* 20(9): 927-932.
- Pedone, E., M. Saviano, M. Rossi and S. Bartolucci** (2001) A single point mutation (Glu85Arg) increases the stability of the thioredoxin from *Escherichia coli*. *Protein Eng* 14(4): 255-260.
- Perl, D., U. Mueller, U. Heinemann and F. X. Schmid** (2000) Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat Struct Biol* 7(5): 380-383.
- Piper, D. E., S. Jackson, Q. Liu, W. G. Romanow, S. Shetterly, S. T. Thibault, B. Shan and N. P. Walker** (2007) The crystal structure of PCSK9: a regulator of plasma LDL-cholesterol. *Structure* 15(5): 545-552.
- Poo, H., J. J. Song, S.-P. Hong, Y.-H. Choi, S. W. Yun, J.-H. Kim, S. C. Lee, S.-G. Lee and M. H. Sung** (2002) Novel high-level constitutive expression system, pHCE vector, for a convenient and cost-effective soluble production of human tumor necrosis factor- α . *Biotechnology Letters* 24(14): 1185-1189.
- Presley, J. F., N. B. Cole, T. A. Schroer, K. Hirschberg, K. J. Zaal and J. Lippincott-Schwartz** (1997) ER-to-Golgi transport visualized in living cells. *Nature* 389(6646): 81-85.
- Ptitsyn, O. B., R. H. Pain, G. V. Semisotnov, E. Zerovnik and O. I. Razgulyaev** (1990) Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett* 262(1): 20-24.
- Puchkaev, A. V., L. S. Koo and P. R. Ortiz de Montellano** (2003) Aromatic stacking as a determinant of the thermal stability of CYP119 from *Sulfolobus solfataricus*. *Arch Biochem Biophys* 409(1): 52-58.
- Ramsden, W.** (1902). Some new properties of urea. *J Physiol* 28: xxiii-xxvii.
- Rasmussen, S. G., H. J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. Schertler, W. I. Weis and B. K. Kobilka** (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* 450(7168): 383-387.
- Reetz, M. T., C. Torre, A. Eipper, R. Lohmer, M. Hermes, B. Brunner, A. Maichele, M. Bocola, M. Arand, A. Cronin, Y. Genzel, A. Archelas and R. Furstoss** (2004) Enhancing the enantioselectivity of an epoxide hydrolase by directed evolution. *Org Lett* 6(2): 177-180.
- Reetz, M. T., J. D. Carballera and A. Vogel** (2006) Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew Chem Int Ed Engl* 45(46): 7745-7751.
- Reetz, M. T., D. Kahakeaw and R. Lohmer** (2008) Addressing the numbers problem in directed evolution. *ChemBiochem* 9(11): 1797-1804.
- Reichardt, H. M., F. Tronche, S. Berger, C. Kellendonk and G. Schutz** (2000) New insights into glucocorticoid and mineralocorticoid signaling: lessons from gene targeting. *Adv Pharmacol* 47: 1-21.
- Rhen, T. and J. A. Cidlowski** (2005) Antiinflammatory action of glucocorticoids--new mechanisms for old drugs. *N Engl J Med* 353(16): 1711-1723.
- Ricketson, D., U. Hostick, L. Fang, K. R. Yamamoto and B. D. Darimont** (2007) A conformational switch in the ligand-binding domain regulates the dependence of the glucocorticoid receptor on Hsp90. *J Mol Biol* 368(3): 729-741.
- Röthlisberger, D., O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik and D. Baker** (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453(7192): 190-195.
- Roman, L. J., E. A. Sheta, P. Martasek, S. S. Gross, Q. Liu and B. S. Masters** (1995) High-level expression of functional rat neuronal nitric oxide synthase in *Escherichia coli*. *Proc Natl Acad Sci U S A* 92(18): 8428-8432.
- Roodveldt, C., A. Aharoni and D. S. Tawfik** (2005) Directed evolution of proteins for heterologous expression and stability. *Curr Opin Struct Biol* 15(1): 50-56.
- Rosenbaum, D. M., V. Cherezov, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, X. J. Yao, W. I. Weis, R. C. Stevens and B. K. Kobilka** (2007) GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* 318(5854): 1266-1273.
- Rosky, P. J.** (2008) Protein denaturation by urea: slash and bond. *Proc Natl Acad Sci U S A* 105(44): 16825-16826.
- Sachdev, D. and J. M. Chirgwin** (1998) Order of fusions between bacterial and mammalian proteins can determine solubility in *Escherichia coli*. *Biochem Biophys Res Commun* 244(3): 933-937.

- Saerens, D., K. Conrath, J. Govaert and S. Muyldermans** (2008) Disulfide bond introduction for general stabilization of immunoglobulin heavy-chain variable domains. *J Mol Biol* 377(2): 478-488.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich** (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839): 487-491.
- Sambrook, J., E. E. Fritsch and T. Maniatis** (1989) Molecular cloning: a laboratory manual. *New York, Cold Spring Harbour*.
- Sancho, J., M. Bueno, L. A. Campos, J. Fernandez-Recio, M. P. Irun, J. Lopez, C. Machicado, I. Pedrosa and M. Toja** (2002) The 'relevant' stability of proteins with equilibrium intermediates. *ScientificWorldJournal* 2: 1209-1215.
- Sanger, F., S. Nicklen and A. R. Coulson** (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74(12): 5463-5467.
- Santoro, M. M. and D. W. Bolen** (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27(21): 8063-8068.
- Sarkar, C. A., I. Dodevski, M. Kenig, S. Dudli, A. Mohr, E. Hermans and A. Plückthun** (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci U S A* 105(39): 14808-14813.
- Sarkar, G. and S. S. Sommer** (1990) The "megaprimer" method of site-directed mutagenesis. *Biotechniques* 8(4): 404-407.
- Sasaki, D. T., E. H. Tichenor, F. Lopez, J. Combs, N. Uchida, C. R. Smith, W. Stokdijk, M. Vardanega, A. M. Buckle, B. Chen and et al.** (1995) Development of a clinically applicable high-speed flow cytometer for the isolation of transplantable human hematopoietic stem cells. *J Hematother* 4(6): 503-514.
- Sastry, M. S., K. Korotkov, Y. Brodsky and F. Baneyx** (2002) Hsp31, the Escherichia coli yedU gene product, is a molecular chaperone whose activity is inhibited by ATP at high temperatures. *J Biol Chem* 277(48): 46026-46034.
- Savory, J. G., B. Hsu, I. R. Laquian, W. Giffin, T. Reich, R. J. Hache and Y. A. Lefebvre** (1999) Discrimination between NL1- and NL2-mediated nuclear localization of the glucocorticoid receptor. *Mol Cell Biol* 19(2): 1025-1037.
- Schein, C. H.** (1990) Solubility as a function of protein structure and solvent components. *Biotechnology (N Y)* 8(4): 308-317.
- Scheinman, R. I., A. Gualberto, C. M. Jewell, J. A. Cidlowski and A. S. Baldwin, Jr.** (1995) Characterization of mechanisms involved in transrepression of NF-kappa B by activated glucocorticoid receptors. *Mol Cell Biol* 15(2): 943-953.
- Schmid F.X.** (1997) Optical spectroscopy to characterize protein conformation and conformational changes. *In* Proofs for Protein Structure: A practical approach. Chapter 11, pp 259-295. Creighton T.E., IRL Press, Oxford, 2ed.
- Schorl, C. and J. M. Sedivy** (2007) Analysis of cell cycle phases and progression in cultured mammalian cells. *Methods* 41(2): 143-150.
- Schwabe, A.** (2007) Mutationsanalysen der Komplexbildung von Tryptophan-Synthasen aus *Thermotoga maritima*. Diplomarbeit, Universität Regensburg.
- Schwehm, J. M., C. A. Fitch, B. N. Dang, E. B. Garcia-Moreno and W. E. Stites** (2003) Changes in stability upon charge reversal and neutralization substitution in staphylococcal nuclease are dominated by favorable electrostatic effects. *Biochemistry* 42(4): 1118-1128.
- Schweizer, A., H. Roschitzki-Voser, P. Amstutz, C. Briand, M. Gulotti-Georgieva, E. Prenosil, H. K. Binz, G. Capitani, A. Baici, A. Plückthun and M. G. Grutter** (2007) Inhibition of caspase-2 by a designed ankyrin repeat protein: specificity, structure, and inhibition mechanism. *Structure* 15(5): 625-636.
- Scott, C., C. J. Jackson, C. W. Coppin, R. G. Mourant, M. E. Hilton, T. D. Sutherland, R. J. Russell and J. G. Oakshott** (2009) Catalytic improvement and evolution of atrazine chlorohydrolase. *Appl Environ Microbiol* 75(7): 2184-2191.
- Sears, L. E., L. S. Moran, C. Kissinger, T. Creasey, H. Perry-O'Keefe, M. Roskey, E. Sutherland and B. E. Slatko** (1992) CircumVent thermal cycle sequencing and alternative manual and automated DNA sequencing protocols using the highly thermostable VentR (exo-) DNA polymerase. *Biotechniques* 13(4): 626-633.
- Seidah, N. G., S. Benjannet, L. Wickham, J. Marcinkiewicz, S. B. Jasmin, S. Stifani, A. Basak, A. Prat and M. Chretien** (2003) The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation. *Proc Natl Acad Sci U S A* 100(3): 928-933.

- Seidah, N. G. and A. Prat** (2007) The proprotein convertases are potential targets in the treatment of dyslipidemia. *J Mol Med* 85(7): 685-696.
- Seitz, T.** (2006) Stabilisierung eines künstlich erzeugten $(\beta\alpha)_8$ -Barrel Proteins durch gelenkte Evolution. Diplomarbeit, Universität Regensburg.
- Seitz, T., M. Bocola, J. Claren and R. Sterner** (2007) Stabilisation of a $(\beta\alpha)_8$ -barrel protein designed from identical half barrels. *J Mol Biol* 372(1): 114-129.
- Sennhauser, G., P. Amstutz, C. Briand, O. Storchenegger and M. G. Grutter** (2007) Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biol* 5(1): e7.
- Serrano, L., M. Bycroft and A. R. Fersht** (1991) Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles. *J Mol Biol* 218(2): 465-475.
- Serrano, L., J. L. Neira, J. Sancho and A. R. Fersht** (1992) Effect of alanine versus glycine in alpha-helices on protein stability. *Nature* 356(6368): 453-455.
- Serrano-Vega, M. J., F. Magnani, Y. Shibata and C. G. Tate** (2008) Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci U S A* 105(3): 877-882.
- Shaner, N. C., G. H. Patterson and M. W. Davidson** (2007) Advances in fluorescent protein technology. *J Cell Sci* 120(Pt 24): 4247-4260.
- Sharp, P. A., B. Sugden and J. Sambrook** (1973) Detection of two restriction endonuclease activities in Haemophilus parainfluenzae using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* 12(16): 3055-3063.
- Shimomura, O.** (2005) The discovery of aequorin and green fluorescent protein. *J Microsc* 217(Pt 1): 1-15.
- Shimomura, O., F. H. Johnson and Y. Saiga** (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusa, Aequorea. *J Cell Comp Physiol* 59: 223-239.
- Sieber, V.** (2003) Selection for soluble proteins via fusion with chloramphenicol acetyltransferase. *Methods Mol Biol* 230: 45-55.
- Sieber, V., C. A. Martinez and F. H. Arnold** (2001) Libraries of hybrid proteins from distantly related sequences. *Nat Biotechnol* 19(5): 456-460.
- Sieber, V., A. Plückthun and F. X. Schmid** (1998) Selecting proteins with improved stability by a phage-based method. *Nat Biotechnol* 16(10): 955-960.
- Siegel, M. S. and E. Y. Isacoff** (1997) A genetically encoded optical probe of membrane voltage. *Neuron* 19(4): 735-741.
- Simpson, E. R., J. K. Meldrum, R. Bofill, M. D. Crespo, E. Holmes and M. S. Searle** (2005) Engineering enhanced protein stability through beta-turn optimization: insights for the design of stable peptide beta-hairpin systems. *Angew Chem Int Ed Engl* 44(31): 4939-4944.
- Slatko, B. E.** (1994) Thermal cycle dideoxy DNA sequencing. *Methods Mol Biol* 31: 35-45.
- Smith, D. B.** (2000) Generating fusions to glutathione S-transferase for protein studies. *Methods Enzymol* 326: 254-270.
- Smoak, K. A. and J. A. Cidlowski** (2004) Mechanisms of glucocorticoid receptor signaling during inflammation. *Mech Ageing Dev* 125(10-11): 697-706.
- Smyth, D. R., M. K. Mrozkiewicz, W. J. McGrath, P. Listwan and B. Kobe** (2003) Crystal structures of fusion proteins with large-affinity tags. *Protein Sci* 12(7): 1313-1322.
- Spiro, K.** (1900). Über die Beeinflussung der Eiweisscoagulation durch stickstoffhaltige Substanzen. *Z Physiol Chem* 30: 182-199.
- Stark, G. R.** (1965) Reactions of cyanate with functional groups of proteins. 3. Reactions with amino and carboxyl groups. *Biochemistry* 4(6): 1030-1036.
- Stemmer, W. P.** (1994a) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci U S A* 91(22): 10747-10751.
- Stemmer, W. P.** (1994b) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370(6488): 389-391.

- Sterner, R. and W. Liebl** (2001) Thermophilic adaptation of proteins. *Crit Rev Biochem Mol Biol* 36(1): 39-106.
- Stickle, D. F., L. G. Presta, K. A. Dill and G. D. Rose** (1992) Hydrogen bonding in globular proteins. *J Mol Biol* 226(4): 1143-1159.
- Straight, A. F., W. F. Marshall, J. W. Sedat and A. W. Murray** (1997) Mitosis in living budding yeast: anaphase A but no metaphase plate. *Science* 277(5325): 574-578.
- Strub, C., C. Alies, A. Lougarre, C. Ladurantie, J. Czaplicki and D. Fournier** (2004) Mutation of exposed hydrophobic amino acids to arginine to increase protein stability. *BMC Biochem* 5: 9.
- Studier, F. W. and B. A. Moffatt** (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189(1): 113-130.
- Studier, F. W., A. H. Rosenberg, J. J. Dunn and J. W. Dubendorff** (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185: 60-89.
- Stüber D., Matile H., Garotta G.** (1990) System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies, and structure-function analysis. *Immunol Methods* 4: 121-152
- Suino-Powell, K., Y. Xu, C. Zhang, Y. G. Tao, W. D. Tolbert, S. S. Simons, Jr. and H. E. Xu** (2008) Doubling the size of the glucocorticoid receptor ligand binding pocket by deacylcortivazol. *Mol Cell Biol* 28(6): 1915-1923.
- Sun, D. P., U. Sauer, H. Nicholson and B. W. Matthews** (1991) Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. *Biochemistry* 30(29): 7142-7153.
- Tamakoshi, M., Y. Nakano, S. Kakizawa, A. Yamagishi and T. Oshima** (2001) Selection of stabilized 3-isopropylmalate dehydrogenase of *Saccharomyces cerevisiae* using the host-vector system of an extreme thermophile, *Thermus thermophilus*. *Extremophiles* 5(1): 17-22.
- Tamura, A. and J. M. Sturtevant** (1995) A thermodynamic study of mutant forms of Streptomyces subtilisin inhibitor. III. Replacements of a hyper-exposed residue, Met73. *J Mol Biol* 249(3): 646-653.
- Taylor, G., M. Hoare, G.R. Gray, and F.A.O. Martson** (1986) Size and density of inclusion bodies. *Biotechnology (N Y)* 4: 553-557.
- Tigerstrom, A., F. Schwarz, G. Karlsson, M. Okvist, C. Alvarez-Rua, D. Maeder, F. T. Robb and L. Sjolín** (2004) Effects of a novel disulfide bond and engineered electrostatic interactions on the thermostability of azurin. *Biochemistry* 43(39): 12563-12574.
- Timasheff, S. N.** (1992) Water as ligand: preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry* 31(41): 9857-9864.
- Timasheff, S. N.** (1993) The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annu Rev Biophys Biomol Struct* 22: 67-97.
- Timasheff, S. N.** (1995) Solvent stabilization of protein structure. *Methods Mol Biol* 40: 253-269.
- Timasheff & Arakawa** (1997) Stabilization of protein structure by solvents. In *Proofs for Protein Structure: A practical approach*. Chapter 14, pp 349-364. Creighton T.E., IRL Press, Oxford, 2ed.
- Tobias, A. V.** (2003) Preparing libraries in *Escherichia coli*. *Methods Mol Biol* 231: 11-16.
- Trevino, S. R., J. M. Scholtz and C. N. Pace** (2007) Amino acid contribution to protein solubility: Asp, Glu, and Ser contribute more favorably than the other hydrophilic amino acids in RNase Sa. *J Mol Biol* 366(2): 449-460.
- Trevino, S. R., J. M. Scholtz and C. N. Pace** (2008) Measuring and increasing protein solubility. *J Pharm Sci* 97(10): 4155-4166.
- Tsai, M. J. and B. W. O'Malley** (1994) Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu Rev Biochem* 63: 451-486.
- Tsien, R. Y.** (1998) The green fluorescent protein. *Annu Rev Biochem* 67: 509-544.
- Tsien, R. Y., B. J. Bacsikai and S. R. Adams** (1993) FRET for studying intracellular signalling. *Trends Cell Biol* 3(7): 242-245.

- Tsuji, T., R. Nakagawa, N. Sugimoto and K. Fukuhara** (1987) Characterization of disulfide bonds in recombinant proteins: reduced human interleukin 2 in inclusion bodies and its oxidative refolding. *Biochemistry* 26(11): 3129-3134.
- Tsumoto, K., D. Ejima, I. Kumagai and T. Arakawa** (2003) Practical considerations in refolding proteins from inclusion bodies. *Protein Expr Purif* 28(1): 1-8.
- Vallejo, L. F. and U. Rinas** (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* 3(1): 11.
- van den Berg, S., P. A. Lofdahl, T. Hard and H. Berglund** (2006) Improved solubility of TEV protease by directed evolution. *J Biotechnol* 121(3): 291-298.
- van den Burg, B., B. W. Dijkstra, G. Vriend, B. Van der Vinne, G. Venema and V. G. Eijnsink** (1994) Protein stabilization by hydrophobic interactions at the surface. *Eur J Biochem* 220(3): 981-985.
- van den Burg, B. and V. G. Eijnsink** (2002) Selection of mutations for increased protein stability. *Curr Opin Biotechnol* 13(4): 333-337.
- Vanhercke, T., C. Ampe, L. Tirry and P. Denolf** (2005) Reducing mutational bias in random protein libraries. *Anal Biochem* 339(1): 9-14.
- Vedadi, M., F. H. Niesen, A. Allali-Hassani, O. Y. Fedorov, P. J. Finerty, Jr., G. A. Wasney, R. Yeung, C. Arrowsmith, L. J. Ball, H. Berglund, R. Hui, B. D. Marsden, P. Nordlund, M. Sundstrom, J. Weigelt and A. M. Edwards** (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* 103(43): 15835-15840.
- Vega, G. L. and S. M. Grundy** (1986) In vivo evidence for reduced binding of low density lipoproteins to receptors as a cause of primary moderate hypercholesterolemia. *J Clin Invest* 78(5): 1410-1414.
- Veith, D. and M. Veith** (2005) Biologie fluoreszierender Proteine. *Biologie Unserer Zeit* 6(35): 394-404.
- Ventura, S. and A. Villaverde** (2006) Protein quality in bacterial inclusion bodies. *Trends Biotechnol* 24(4): 179-185.
- Villaverde, A. and M. M. Carrio** (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* 25(17): 1385-1395.
- Wahlberg, E., C. Lendel, M. Helgstrand, P. Allard, V. Dinibas-Renqvist, A. Hedqvist, H. Berglund, P. A. Nygren and T. Hard** (2003) An affibody in complex with a target protein: structure and coupled folding. *Proc Natl Acad Sci U S A* 100(6): 3185-3190.
- Waldburger, C. D., J. F. Schildbach and R. T. Sauer** (1995) Are buried salt bridges important for protein stability and conformational specificity? *Nat Struct Biol* 2(2): 122-128.
- Waldo, G. S.** (2003) Improving protein folding efficiency by directed evolution using the GFP folding reporter. *Methods Mol Biol* 230: 343-359.
- Waldo, G. S., B. M. Standish, J. Berendzen and T. C. Terwilliger** (1999) Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* 17(7): 691-695.
- Waldron, T. T. and K. P. Murphy** (2003) Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. *Biochemistry* 42(17): 5058-5064.
- Warne, T., M. J. Serrano-Vega, J. G. Baker, R. Moukhametzianov, P. C. Edwards, R. Henderson, A. G. Leslie, C. G. Tate and G. F. Schertler** (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454(7203): 486-491.
- Warnmark, A., J. A. Gustafsson and A. P. Wright** (2000) Architectural principles for the structure and function of the glucocorticoid receptor tau 1 core activation domain. *J Biol Chem* 275(20): 15014-15018.
- Watson, J. V.** (1999) The early fluidic and optical physics of cytometry. *Cytometry* 38(1): 2-14; discussion 11.
- Waugh, D. S.** (2005) Making the most of affinity tags. *Trends Biotechnol* 23(6): 316-320.
- Weast R.C.** (1982-1983) CRC Handbook of Chemistry and Physics. CRC Press. 63rd edition: D-231
- Weir, M. P. and J. Sparks** (1987) Purification and renaturation of recombinant human interleukin-2. *Biochem J* 245(1): 85-91.
- Werner, S. and M. Bronnegard** (1996) Molecular basis of glucocorticoid-resistant syndromes. *Steroids* 61(4): 216-221.

- Whitford D.** (2005) Proteins. Structure and Function. 1st Edition.
- Wigley, W. C., R. D. Stidham, N. M. Smith, J. F. Hunt and P. J. Thomas** (2001) Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein. *Nat Biotechnol* 19(2): 131-136.
- Williams, J. C., J. P. Zeelen, G. Neubauer, G. Vriend, J. Backmann, P. A. Michels, A. M. Lambeir and R. K. Wierenga** (1999) Structural and mutagenesis studies of leishmania triosephosphate isomerase: a point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power. *Protein Eng* 12(3): 243-250.
- Williams, S. P. and P. B. Sigler** (1998) Atomic structure of progesterone complexed with its receptor. *Nature* 393(6683): 392-396.
- Wilson, G. G. and N. E. Murray** (1991) Restriction and modification systems. *Annu Rev Genet* 25: 585-627.
- Wriggers, W., S. Chakravarty and P. A. Jennings** (2005) Control of protein functional dynamics by peptide linkers. *Biopolymers* 80(6): 736-746.
- Yang, F., L. G. Moss and G. N. Phillips, Jr.** (1996) The molecular structure of green fluorescent protein. *Nat Biotechnol* 14(10): 1246-1251.
- Yang, J. K., M. S. Park, G. S. Waldo and S. W. Suh** (2003) Directed evolution approach to a structural genomics project: Rv2002 from Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* 100(2): 455-460.
- Yang, J. K., H. J. Yoon, H. J. Ahn, B. I. Lee, S. H. Cho, G. S. Waldo, M. S. Park and S. W. Suh** (2002) Crystallization and preliminary X-ray crystallographic analysis of the Rv2002 gene product from Mycobacterium tuberculosis, a beta-ketoacyl carrier protein reductase homologue. *Acta Crystallogr D Biol Crystallogr* 58(Pt 2): 303-305.
- Yang, T. T., L. Cheng and S. R. Kain** (1996) Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res* 24(22): 4592-4593.
- Yang, G. and S. G. Withers** (2009) Ultrahigh-Throughput FACS-Based Screening for Directed Enzyme Evolution. *Chembiochem*.
- Yokoe, H. and T. Meyer** (1996) Spatial dynamics of GFP-tagged proteins investigated by local fluorescence enhancement. *Nat Biotechnol* 14(10): 1252-1256.
- Zaccolo, M. and E. Gherardi** (1999) The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J Mol Biol* 285(2): 775-783.
- Zhang, X., W. Meining, M. Fischer, A. Bacher and R. Ladenstein** (2001) X-ray structure analysis and crystallographic refinement of lumazine synthase from the hyperthermophile Aquifex aeolicus at 1.6 Å resolution: determinants of thermostability revealed from structural comparisons. *J Mol Biol* 306(5): 1099-1114.
- Zhao, Z., Y. Tuakli-Wosornu, T. A. Lagace, L. Kinch, N. V. Grishin, J. D. Horton, J. C. Cohen and H. H. Hobbs** (2006) Molecular characterization of loss-of-function mutations in PCSK9 and identification of a compound heterozygote. *Am J Hum Genet* 79(3): 514-523.
- Zhou, Y., J. H. Morais-Cabral, A. Kaufman and R. MacKinnon** (2001) Chemistry of ion coordination and hydration revealed by a K⁺ channel-Fab complex at 2.0 Å resolution. *Nature* 414(6859): 43-48.

7 Anhang

7.1 Protokoll zur Zell-Sortierung mittels Durchflusszytometrie

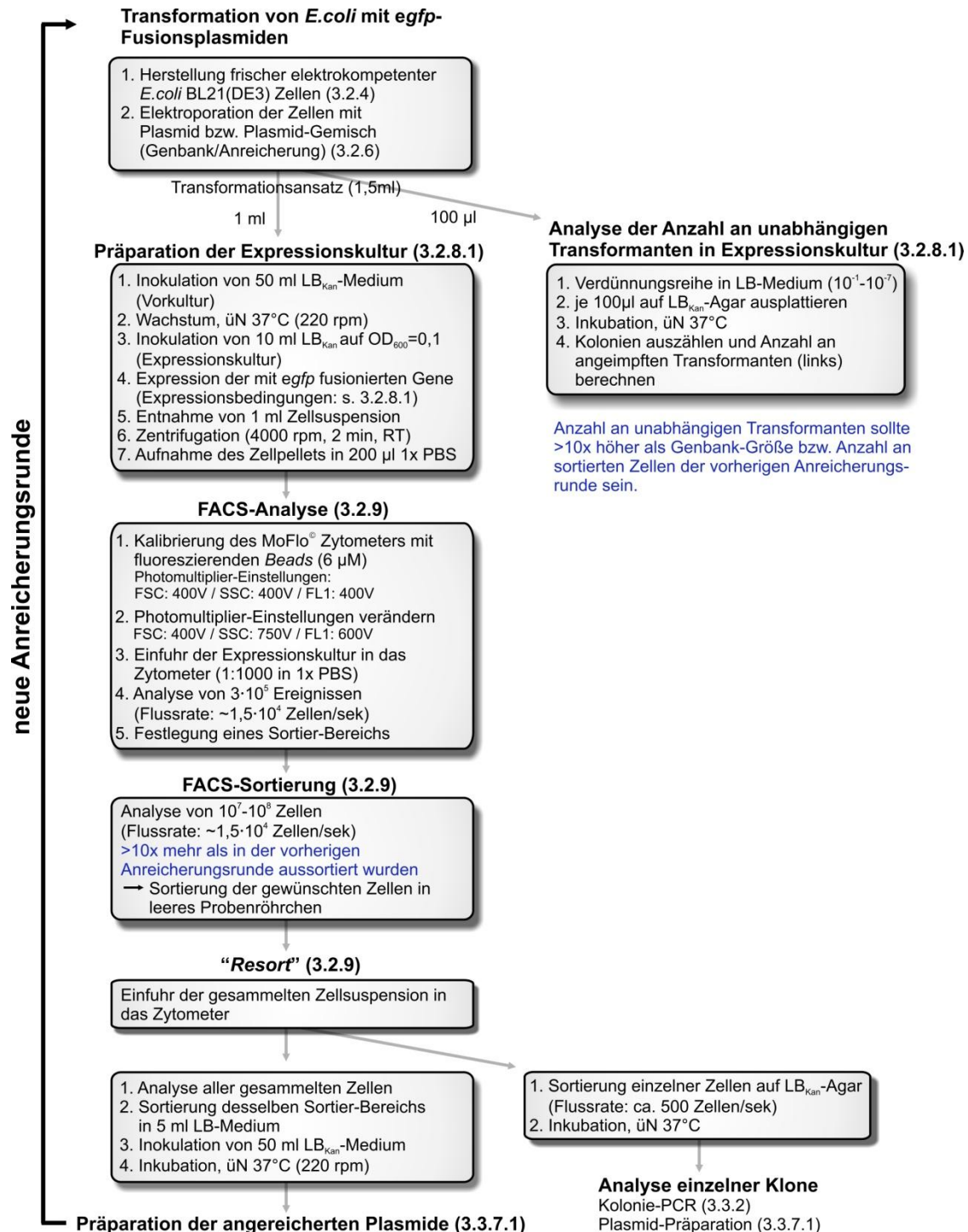


Abbildung A1: Schematische Darstellung des Vorgehens bei einer FACS-Sortierung.

Maßnahmen um sicherzustellen, dass bei der FACS-Analyse alle Varianten der Genbank bzw. der vorherigen Anreicherungsrunde berücksichtigt werden, sind blau dargestellt.

7.2 Bestimmung der Proteinkonzentration durch Fluoreszenz-Spektroskopie

Nach fraktionierter Präzipitation mit Ammoniumsulfat (3.5.11) wurde die Proteinkonzentration im löslichen Überstand über Fluoreszenz-Spektroskopie ermittelt. Um sicherzustellen, dass in dem verwendeten Konzentrationsbereich die Fluoreszenz-Emission linear proportional zur Proteinkonzentration ist und somit ein innerer Filter-Effekt ausgeschlossen werden kann, wurden folgende Eichgeraden erstellt.

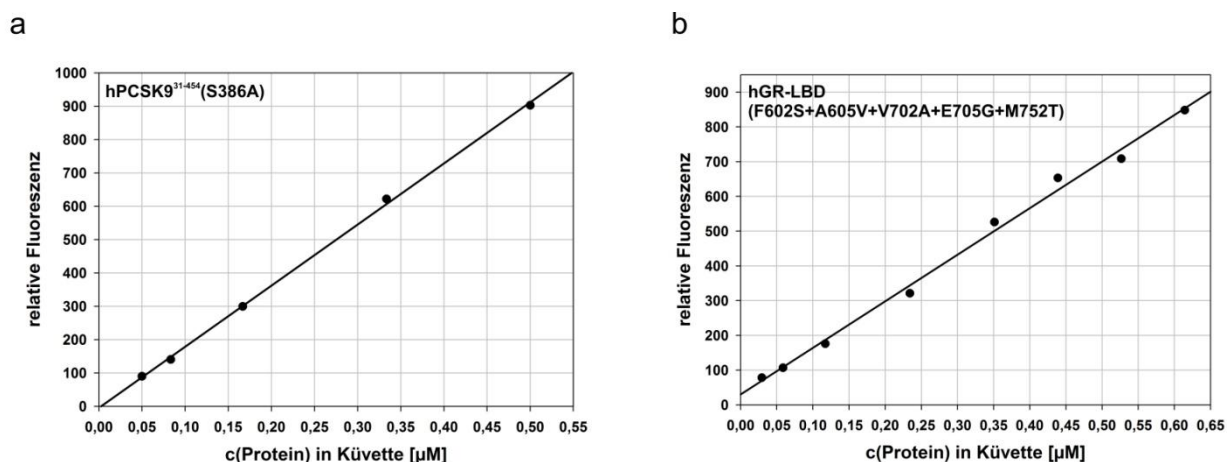


Abbildung A2: Eichkurven zur Bestimmung der Proteinkonzentration durch Fluoreszenz-Spektroskopie.

a Die Emission von hPCSK9³¹⁻⁴⁵⁴ (S386A) bei 344 nm nach Anregung mit 280 nm ist im 4.1.4.1 untersuchten Bereich linear proportional zur Konzentration an eingesetztem Protein. Die Messungen erfolgten in 50 mM Tris/Cl, 5 mM DTT, pH 8,0 bei RT. Bei jeder Probe wurde die Emission 1 min lang gemessen und gemittelt (Schichtdicke der Küvette: 1 cm; Photomultiplier-Spannung: 910 V).

b Die Emission von hGR-LBD(F602S+A605V+V702A+E705G+M752T) bei 342 nm nach Anregung mit 280 nm ist im 4.3.5.6.1 untersuchten Bereich linear proportional zur Konzentration an eingesetztem Protein. Die Messungen erfolgten in 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 50 µM Dexamethason, pH 7,9 bei RT. Bei jeder Probe wurde die Emission 1 min lang gemessen und gemittelt (Schichtdicke der Küvette: 1 cm; Photomultiplier-Spannung: 845 V).

7.3 Multiples Sequenzalignment der mittels FACS isolierten hGR-LBD Varianten

Die verwendete Ligand-Bindungs Domäne (LBD) des humanen Glucocorticoid-Rezeptors (hGR) besteht aus den Aminosäuren 521-777 des Proteins. Die folgenden hGR-LBD Varianten sind im Vergleich zum Wildtyp gezeigt (siehe Tabelle 12 in 4.3.2.2):

- A5-1 – A5-4: Sequenzen der vier sortierten Varianten aus Anreicherung 5 bzw.7
- A8: Sequenz aus Anreicherung 8 (re-amplifiziert durch OE-PCR)

Stille Mutationen sind gelb, Nukleotid-Austausche die zu Aminosäuren-Austausche führen sind grün markiert. Es ist sowohl die Nukleotid- als auch die Aminosäure-Sequenz gezeigt. Stopp-Codone sind durch * markiert.

Gen (1557+)	10	20	30	40	50	60	70	80	90	100
AS	521	525	530	535	540	545	550			
Wildtyp	ATGGTTCCTGCAACGTTACCACAACCTACCCCTACCCCTGGTGTCACTGTTGGAGGTTATTGAACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCC									
A5-1	ATGGTTCCTGCAACGTTACCACAACCTACCCCTACCCCTGGTGTCCGCTGTTGGAGGTTATTGAACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCC									
A5-2	ATGGTTCCTGCAACGTTACCACAACCTACCCCTACCCCTGGTGTCACTGTTGGAGGTTATTGAACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCC									
A5-3	ATGGTTCCTGCAACGTTACCACAACCTACCCCTACCCCTGGTGTCACTGTTGGAGGTTATTGAACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCC									
A5-4	ATGGTTCCTGCAACGTTACCACAACCTACCCCTACCCCTGGTGTCACTGTTGGAGGTTATTGAACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCC									
A8	ATGGTTCCTGCAACGTTACCACAACCTACCCCTACCCCTGGTGTCACTGTTGGAGGTTATTGAACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCC									

Gen (1557+)	110	120	130	140	150	160	170	180	190	200
AS	555	560	565	570	575	580	585			
Wildtyp	CAGACTCAACTTGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTATTGCAGCAGTAAAATGGGCAAAGGCAATACCAGGTTTCAGGAA									
A5-1	CAGACTCAACTTGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTATTGCAGCAGTAAAATGGGCAAAGGCAATACCAGGTTTCAGGAA									
A5-2	CAGACTCAACTTGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTATTGCAGCAGTAAAATGGGCAAAGGCAATACCAGGTTTCAGGAA									
A5-3	CAGACTCAACTTGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTATTGCAGCAGTAAAATGGGCAAAGGCAATACCAGGTTTCAGGAA									
A5-4	CAGACTCAACTTGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTATTGCAGCAGTAAAATGGGCAAAGGCAATACCAGGTTTCAGGAA									
A8	CAGACTCAACTTGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTATTGCAGCAGTAAAATGGGCAAAGGCAATACCAGGTTTCAGGAA									

Gen (1557+)	210	220	230	240	250	260	270	280	290	300
AS	590	595	600	605	610	615				
Wildtyp	CTTACACCTGGATGACCAAATGACCCTACTGCAGTACTCCTGGATGTTTCTTATGGCATTGCTCTGGGGTGGAGATCATATAGACAATCAAGTGCAAAC									
A5-1	CTTACACCTGGATGACCAAATGACCCTACTGCAGTACTCCTGGATGTTTCTTATGGCATTGCTCTGGGGTGGAGATCATATAGACAATCAAGTGCAAAC									
A5-2	CTTACACCTGGATGACCAAATGACCCTACTGCAGTACTCCTGGATGTTTCTTATGGCATTGCTCTGGGGTGGAGATCATATAGACAATCAAGTGCAAAC									
A5-3	CTTACACCTGGATGACCAAATGACCCTACTGCAGTACTCCTGGATGTTTCTTATGGCATTGCTCTGGGGTGGAGATCATATAGACAATCAAGTGCAAAC									
A5-4	CTTACACCTGGATGACCAAATGACCCTACTGCAGTACTCCTGGATGTTTCTTATGGCATTGCTCTGGGGTGGAGATCATATAGACAATCAAGTGCAAAC									
A8	CTTACACCTGGATGACCAAATGACCCTACTGCAGTACTCCTGGATGTTTCTTATGGCATTGCTCTGGGGTGGAGATCATATAGACAATCAAGTGCAAAC									

Gen (1557+) 310 320 330 340 350 360 370 380 390 400
AS 620 625 630 635 640 645 650
Wildtyp CTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCTCTGAGT
L L C F A P D L I I N E Q R M T L P C M Y D Q C K H M L Y V S S E
A5-1 CTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCTCTGAGT
L L C F A P D L I I N E Q R M T L P C M Y D Q C K H M L Y V S S E
A5-2 CTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCTCTGAGT
L L C F A P D L I I N E Q R M T L P C M Y D Q C K H M L Y V S S E
A5-3 CTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCTCTGAGT
L L C F A P D L I I N E Q R M T L P C M Y D Q C K H M L Y V S S E
A5-4 CTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCTCTGAGT
L L C F A P D L I I N E Q R M T L P C M Y D Q C K H M L Y V S S E
A8 CTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCTCTGAGT
L L C F A P D L I I N E Q R M T L P C M Y D Q C K H M L Y V S S E

Gen (1557+) 410 420 430 440 450 460 470 480 490 500
AS 655 660 665 670 675 680 685
Wildtyp TACACAGGCTTCAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTTCAGTTCCCTAAGGACGGTCTGAAGAGCCAAAGAGCTATT
L H R L Q V S Y E E Y L C M K T L L L L S S V P K D G L K S Q E L F
A5-1 TACACAGGCTTCAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTTCAGTTCCCTAAGGACGGTCTGAAGAGCCAAAGAGCTATT
L H R L Q V S Y E E Y L C M K T L L L L S S V P K D G L K S Q E L F
A5-2 TACACAGGCTTCAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTTCAGTTCCCTAAGGACGGTCTGAAGAGCCAAAGAGCTATT
L H R L Q V S Y E E Y L C M K T L L L L S S V P K D G L K S Q E L F
A5-3 TACACAGGCTTCAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTTCAGTTCCCTAAGGACGGTCTGAAGAGCCAAAGAGCTATT
L H R L Q V S Y E E Y L C M K T L L L L S S V P K D G L K S Q E L F
A5-4 TACACAGGCTTCAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTTCAGTTCCCTAAGGACGGTCTGAAGAGCCAAAGAGCTATT
L H R L Q V S Y E E Y L C M K T L L L L S S V P K D G L K S Q E L F
A8 TACACAGGCTTCAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTTCAGTTCCCTAAGGACGGTCTGAAGAGCCAAAGAGCTATT
L H R L Q V S Y E E Y L C M K T L L L L S S V P K D G L K S Q E L F

Gen (1557+) 510 520 530 540 550 560 570 580 590 600
AS 690 695 700 705 710 715
Wildtyp TGATGAAATTAGAATGACCTACATCAAGAGCTAGGAAAAGCCATTGTCAAGAGGGAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTTATCAACTGACA
D E I R M T Y I K E L G K A I V K R E G N S S Q N W Q R F Y Q L T
A5-1 TGATGAAATTAGAATGACCTACATCAAGAGCTAGGAAAAGCCATTGTCAAGAGGGAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTTATCAACTGACA
D E I R M T Y I K E L G K A I V K R E G N S S Q N W Q R F Y Q L T
A5-2 TGATGAAATTAGAATGACCTACATCAAGAGCTAGGAAAAGCCATTGTCAAGAGGGAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTTATCAACTGACA
D E I R M T Y I K E L G K A I V K R E G N S S Q N W Q R F Y Q L T
A5-3 TGATGAAATTAGAATGACCTACATCAAGAGCTAGGAAAAGCCATTGTCAAGAGGGAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTTATCAACTGACA
D E I R M T Y I K E L G K A I V K R E G N S S Q N W Q R F Y Q L T
A5-4 TGATGAAATTAGAATGACCTACATCAAGAGCTAGGAAAAGCCATTGTCAAGAGGGAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTTATCAACTGACA
D E I R M T Y I K E L G K A I V K R E G N S S Q N W Q R F Y Q L T
A8 TGATGAAATTAGAATGACCTACATCAAGAGCTAGGAAAAGCCATTGTCAAGAGGGAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTTATCAACTGACA
D E I R M T Y I K E L G K A I V K R E G N S S Q N W Q R F Y Q L T

Gen (1557+) 610 620 630 640 650 660 670 680 690 700
AS 720 725 730 735 740 745 750
Wildtyp AAACCTTGGATTCTATGCATGAAGTGGTTGAAAATCTCCTTAACTATTGCTTCCAAACATTTTGGATAAGACCATGAGTATTGAATCCCCGAGATGT
K L L D S M H E V V E N L L N Y C F Q T F L D K T M S I E F P E M
A5-1 AAACCTTGGATTCTATGCATGAAGTGGTTGAAAATCTCCTTAACTATTGCTTCCAAACATTTTGGATAAGACCATGAGTATTGAATCCCCGAGATGT
K L L D S M H E V V E N L L N Y C F Q T F L D K T M S I E F P E M
A5-2 AAACCTTGGATTCTATGCATGAAGTGGTTGAAAATCTCCTTAACTATTGCTTCCAAACATTTTGGATAAGACCATGAGTATTGAATCCCCGAGATGT
K L L D S M H E V V E N L L N Y C F Q T F L D K T M S I E F P E M
A5-3 AAACCTTGGATTCTATGCATGAAGTGGTTGAAAATCTCCTTAACTATTGCTTCCAAACATTTTGGATAAGACCATGAGTATTGAATCCCCGAGATGT
K L L D S M H E V V E N L L N Y C F Q T F L D K T M S I E F P E M
A5-4 AAACCTTGGATTCTATGCATGAAGTGGTTGAAAATCTCCTTAACTATTGCTTCCAAACATTTTGGATAAGACCATGAGTATTGAATCCCCGAGATGT
K L L D S M H E V V E N L L N Y C F Q T F L D K T M S I E F P E M
A8 AAACCTTGGATTCTATGCATGAAGTGGTTGAAAATCTCCTTAACTATTGCTTCCAAACATTTTGGATAAGACCATGAGTATTGAATCCCCGAGATGT
K L L D S M H E V V E N L L N Y C F Q T F L D K T M S I E F P E M

Gen (1557+)	710	720	730	740	750	760	770
AS	755	760	765	770	775		
Wildtyp	TAGCTGAAATCATCACC AATCAGATACCAA AATATTCAAATGGA AATATCAAAAAA CTTCTGTTTCATCAA AAGTGA						
	L A E I I T N Q I P K Y S N G N I K K L L F H Q K *						
A5-1	TAGCTGAAATCATCACC AATCAGATACCAA AATATTCAAATGGA AATATCAAAAAA CTTCTGTTTCATCAA AAGTGA						
	L A E I I T N Q I P K Y S N G N I K K L L F H Q K *						
A5-2	TAGCTGAAATCATCACC AATCAGATACCAA AATATTCAAATGGA AATATCAAAAAA CTTCTGTTTCATCAA AAGTGA						
	L A E I I T H Q I P K Y S N G N I K K L L F H Q K *						
A5-3	TAGCTGAAATCATCACC AATCAGATACCAA AATATTCAAATGGA AATATCAAAAAA CTTCTGTTTCATCAA AAGTGA						
	L A E I I T N Q I P K Y E N G N I K K L L F H Q K *						
A5-4	TAGCTGAAATCATCACC AATCAGATACCAA AATATTCAAATGGA AATATCAAAAAA CTTCTGTTTCATCAA AAGTGA						
	L A E I I T N Q I P K Y S N G N I K K L L S H Q K *						
A8	TAGCTGAAATCATCACC AATCAGATACCAA AATATTCAAATGGA AATATCAAAAAA CTTCTGTTTCATCAA AAGTGA						
	L A E I I T N Q I P K Y S N G N I K K L L F H Q K *						

7.4 Molekulargewichte, spezifische Extinktionskoeffizienten bei 280 nm und Ausbeuten der aufgereinigten hGR-LBD Einzel- und Kombinationsvarianten

Alle aufgereinigten hGR-LBD (= Aminosäure 521-777 von hGR) Einzel- und Kombinationsvarianten sind N-terminal um drei Aminosäuren (GSH) erweitert, welche nach der Abspaltung des N-terminal fusionierten His₆-GST durch Thrombin zurückbleiben. Daneben besitzen alle Varianten die gleiche Anzahl an Tryptophanen (5) und Tyrosinen (12), wodurch sie denselben molaren Extinktionskoeffizienten aufweisen ($\epsilon_{280} = 45380 \text{ M}^{-1}\text{cm}^{-1}$). Das Molekulargewicht, der spezifische Extinktionskoeffizient ($^{0,1\%}A_{280}$) und die jeweiligen Reinigungs-Ausbeuten der Proteine sind in Tabelle A1 gezeigt.

Tabelle A1: Molekulargewichte, molare Extinktionskoeffizienten und Ausbeuten der in dieser Arbeit aufgereinigten hGR-LBD Einzel- und Kombinationsvarianten.

hGR-LBD	MW [Da]	$^{0,1\%}A_{280}$ [cm ² /mg]	Ausbeute pro g Biomasse
F602S	30193	1,50	0,20 mg
A605V	30281	1,50	0,12 mg
V702A	30225	1,50	0,11 mg
E705G	30181	1,50	0,06 mg
M752T	30223	1,50	0,10 mg
A605V+M752T	30251	1,50	0,35 mg
A605V+M752T+V702A	30223	1,50	3,55 mg
A605V+M752T+E705G	30179	1,50	2,39 mg
A605V+M752T+V702A+E705G	30151	1,51	2,20 mg
F602S+A605V	30221	1,50	0,74 mg
F602S+A605V+M752T	30191	1,50	2,06 mg
F602S+A605V+M752T+M560K	30188	1,50	0,17 mg
F602S+A605V+M752T+V702A+E705G	30091	1,51	5,29 mg

7.5 Kalibrierung der Superdex-75 (16/60) Gelfiltrationssäule

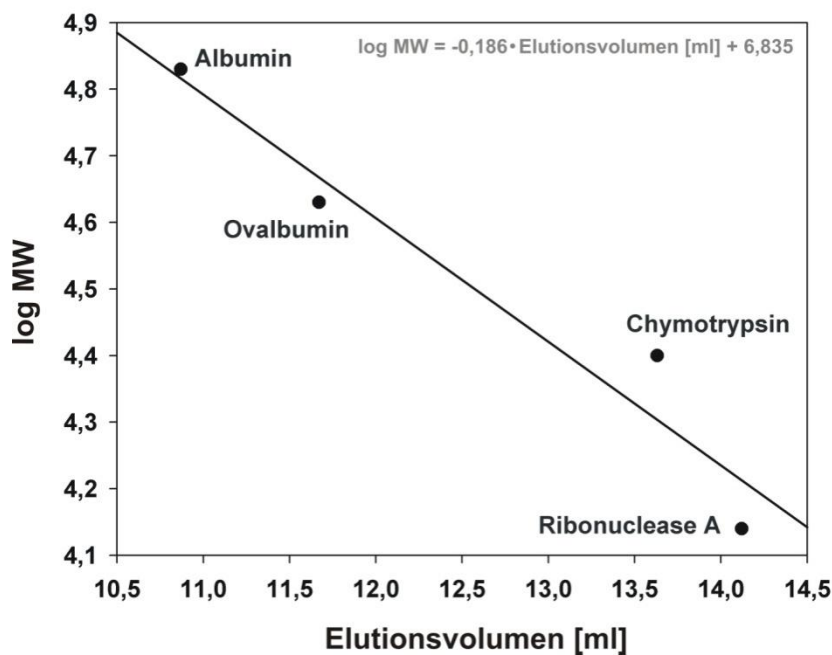


Abbildung A3: Kalibrierung der Superdex S-75 Gelfiltrationssäule.

Verschiedene Proteine mit bekanntem Molekulargewicht (Tabelle A2) wurden bei RT mit 25 mM Tris/Cl, 100 mM NaCl, 2 mM TCEP, 10 % Glycerin, 20 μ M Dexamethason, pH 7,9 auf die Säule aufgetragen und mit einer Flussrate von 0,5 ml/min eluiert. Dargestellt ist die Auftragung des Logarithmus des Molekulargewichts (log MW) gegen das Elutionsvolumen in ml.

Tabelle A2: Zur Kalibrierung der Gelfiltrationssäule Superdex S-75 verwendete Proteine.

Protein	MW [g/mol]	log MW	Elutionsvolumen [ml]
Albumin	67000	4,83	10,87
Ovalbumin	43000	4,63	11,67
Chymotrypsin	25000	4,40	13,63
Ribonuklease A	13700	4,14	14,12

7.6 Multiples Sequenzalignment hGR-LBD - homologer Proteine

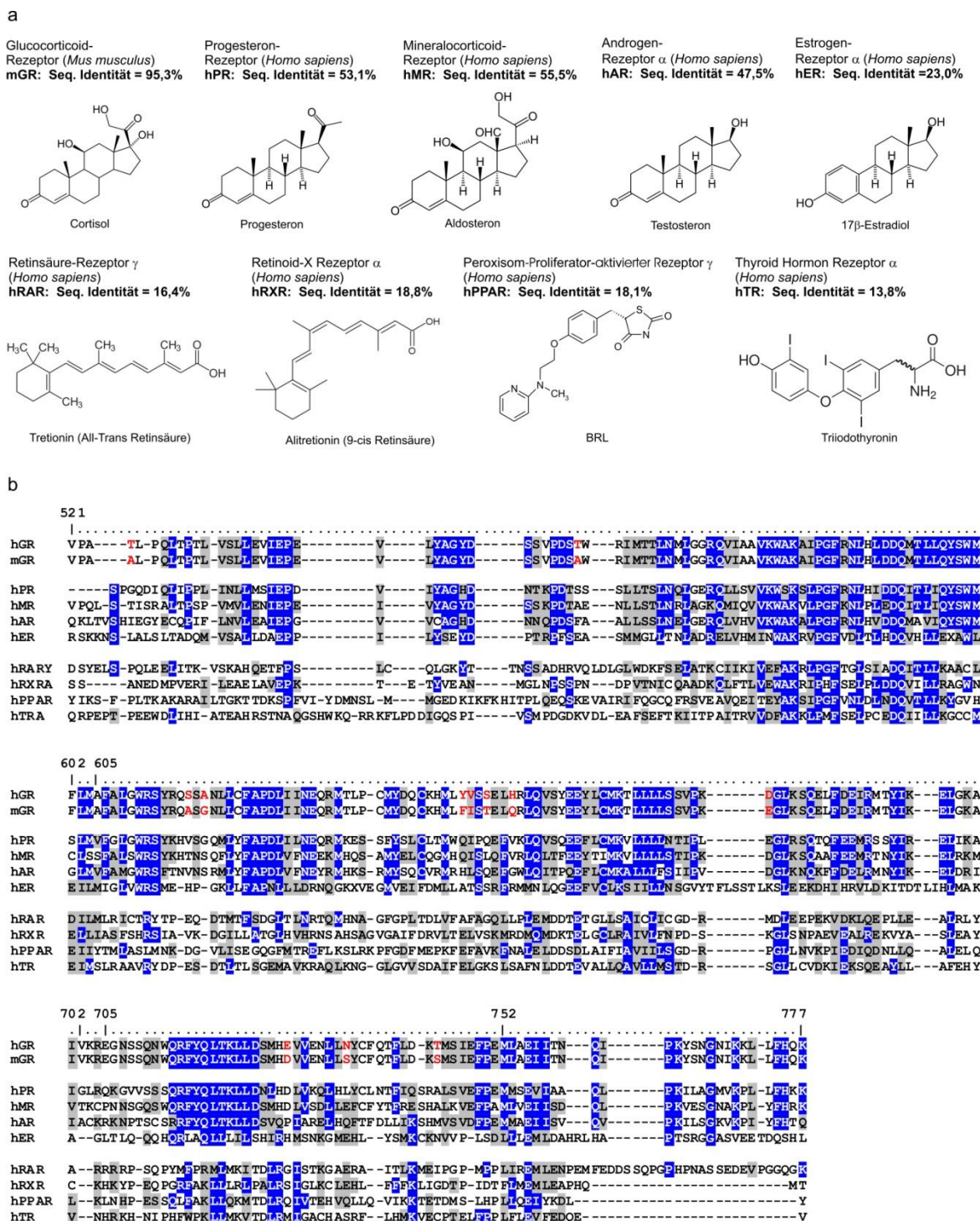


Abbildung A4: Multiples Sequenzalignment hGR-LBD – homologer Ligand-Bindungs Domänen verschiedener Nuklearhormon-Rezeptoren, erstellt durch MCOffee.

Homepage: <http://www.igs.cnrsrms.fr/Tcoffee/tcoffee.cgi/index.cgi?stage1=1&daction=MCOFFEE::Regular>

a Homologe Ligand-Bindungs Domänen von hGR-LBD mit exemplarischen Liganden. Die Sequenz-Identität mit hGR-LBD ist jeweils angegeben.

b Die in dieser Arbeit besprochenen Reste F602, A605, V702, E705 und M752 von hGR-LBD sind gekennzeichnet. Übereinstimmungen bzw. konservative Substitutionen sind blau bzw. grau hinterlegt (verwendete Substitutionsmatrix: Blosum-62; Threshold: 50 %). Unterschiede zwischen hGR-LBD und mGR-LBD sind rot markiert.

7.7 Superpositionierung von mGR-LBD mit hRXR α – hPPAR γ Nuklearhormonrezeptor-Komplex

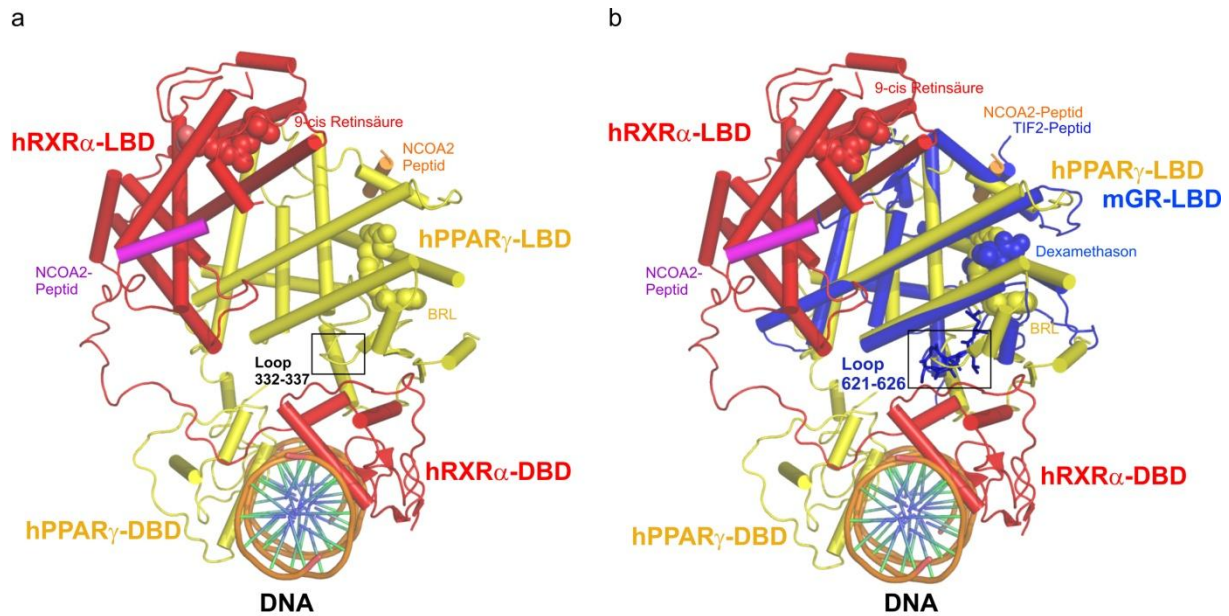


Abbildung A5: Superpositionierung von mGR-LBD mit hRXR α – hPPAR γ Nuklearhormonrezeptor-Komplex (pdb 3dzy; Chandra *et al.*, 2008), erstellt durch PyMol (<http://www.pymol.org/>).

DBD: DNA-bindende Domäne; LBD: Ligand-Bindungsdomäne.

a Veröffentlichte Struktur des hRXR α – hPPAR γ Nuklearhormonrezeptors mit gebundenen Liganden (9-cis Retinsäure und BRL) bzw. NCOA2-Peptid und gebundener DNA (pdb 3dzy). **b** Die C α -Atome von mGR-LBD(A611V+V708A+E711G) und der Ligand-Bindungsdomäne von hPPAR γ konnten mit einem RMSD von 1,58 Å (158 C α -Atome) superpositioniert werden. Loop 621-626 aus mGR-LBD(A611V+V708A+E711G) überlagert mit Loop 332-337 aus hPPAR γ , welcher mit der DNA-bindenden Domäne von hRXR α interagiert.

7.8 Teilpublikationen zur Stabilisierung und Strukturaufklärung des künstlich hergestellten $(\beta\alpha)_8$ -Barrel Proteins HisF-C*C

Im Rahmen dieser Dissertation wurde neben der in Kapitel 4 und 5 beschriebenen Optimierung der Proteine hPCSK9³¹⁻⁴⁵⁴(S386A) und hGR-LBD auch das artifizielle $(\beta\alpha)_8$ -Barrel Protein HisF-C*C, welches aus der duplizierten und *in tandem* fusionierten C-terminalen Hälfte von tmHisF (HisF-C) besteht (Höcker *et al.*, 2004), so weit stabilisiert, dass die Strukturaufklärung mittels Röntgenkristallographie möglich wurde.

Die erzielten Ergebnisse werden in diesem Kapitel anhand von zwei Teilpublikationen gezeigt. Im Folgenden wird kurz erläutert, welcher Beitrag zu diesen Publikationen im Rahmen der vorliegenden Dissertation geleistet wurde.

1.

Seitz, T., M. Bocola, J. Claren and R. Sterner (2007)

Stabilisation of a $(\beta\alpha)_8$ -barrel protein designed from identical half barrels. *J Mol Biol* 372(1): 114-129.

Die Selektion der Varianten HisF-C*C 1-4 (Abbildung 2) und deren Charakterisierung durch chemische Denaturierung und analytische Gelfiltration (Abbildung 3) ist bereits in Seitz (2006) dokumentiert.

Die Quantifizierung der Löslichkeit *in vivo* der Varianten HisF-C*C 1-4 (Abbildung 4), sowie die Herstellung und vollständige Charakterisierung von HisF-C*C Δ , HisF-C*C Δ +D135G, HisF-C*C Δ +T142S, HisF-C*C Δ +Y143H, HisF-C*C Δ +V234M bzw. HisF-C**C (Abbildung 4, 5, 6) erfolgte im Rahmen der vorliegenden Dissertation.

Die Modellierung von HisF-C**C (Abbildung 7) wurde von Dr. Marco Bocola durchgeführt.

2.

Höcker, B., A. Lochner, T. Seitz, J. Claren and R. Sterner (2009)

High-Resolution Crystal Structure of an artificial $(\beta\alpha)_8$ -Barrel Protein designed from Identical Half-Barrels. *Biochemistry* 48 (6): 1145–1147.

Die Herstellung und Charakterisierung von HisF-C***C (Abbildung 1) erfolgte im Rahmen der Diplomarbeit von Adriane Lochner (Lochner, 2008). Diese Arbeit wurde während der vorliegenden Dissertation von mir angeleitet, wobei die in Abbildung 1c gezeigten chemischen Denaturierungen von HisF, HisF-C**C und HisF-C***C von mir reproduziert wurden, um Mittelwerte und Standardabweichungen der jeweiligen $[D_{1/2}]^{app}$ -Werte bestimmen zu können.

Die Röntgenkristallstrukturanalyse von HisF-C***C wurde von Dr. Birte Höcker (MPI für Entwicklungsbiologie, Tübingen) durchgeführt (Abbildung 2, 3, S1, S2).

doi:10.1016/j.jmb.2007.06.036

J. Mol. Biol. (2007) 372, 114–129

JMB

Available online at www.sciencedirect.com

 ScienceDirect

Stabilisation of a $(\beta\alpha)_8$ -Barrel Protein Designed from Identical Half Barrels

Tobias Seitz, Marco Bocola, Jörg Claren and Reinhard Sterner*

Institute of Biophysics and
Physical Biochemistry
University of Regensburg
Universitätsstrasse 31
D-93053 Regensburg, Germany

It has been suggested that the common $(\beta\alpha)_8$ -barrel enzyme fold has evolved by the duplication and fusion of identical $(\beta\alpha)_4$ -half barrels, followed by the optimisation of their interface. In our attempts to reconstruct these events *in vitro* we have previously linked in tandem two copies of the C-terminal half barrel HisF-C of imidazole glycerol phosphate synthase from *Thermotoga maritima* and subsequently reconstituted in the fusion construct HisF-CC a salt bridge cluster present in wild-type HisF. The resulting recombinant protein HisF-C^{*C}, which was produced in an insoluble form and unfolded with low cooperativity at moderate urea concentrations has now been stabilised and solubilised by a combination of random mutagenesis and selection *in vivo*. For this purpose, *Escherichia coli* cells were transformed with a plasmid-based gene library encoding HisF-C^{*C} variants fused to chloramphenicol acetyltransferase (CAT). Stable and soluble variants were identified by the survival of host cells on solid medium containing high concentrations of the antibiotic. The selected HisF-C^{*C} proteins, which were characterised *in vitro* in the absence of CAT, contained eight different amino acid substitutions. One of the exchanges (Y143C) stabilised HisF-C^{*C} by the formation of an intermolecular disulfide bond. Three of the substitutions (G245R, V248M, L250Q) were located in the long loop connecting the two HisF-C copies, whose subsequent truncation from 13 to 5 residues yielded the stabilised variant HisF-C^{*C} Δ. From the remaining substitutions, Y143H and V234M were most beneficial, and molecular dynamics simulations suggest that they strengthen the interactions between the half barrels by establishing a hydrogen-bonding network and an extensive hydrophobic cluster, respectively. By combining the loop deletion of HisF-C^{*C} Δ with the Y143H and V234M substitutions, the variant HisF-C^{**C} was generated. Recombinant HisF-C^{**C} is produced in soluble form, forms a pure monomer with its tryptophan residues shielded from solvent and unfolds with similar cooperativity as HisF. Our results show that, starting from two identical and fused half barrels, few amino acid exchanges are sufficient to generate a highly stable and compact $(\beta\alpha)_8$ -barrel protein with wild-type like structural properties.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: protein design; enzyme evolution; $(\beta\alpha)_8$ -barrel; half barrel; stability screen

*Corresponding author

Abbreviations used: AdoMet, S-adenosyl-L-methionine; CAT, chloramphenicol acetyltransferase; MD, molecular dynamics; GdmCl, guanidinium chloride.

E-mail address of the corresponding author:
Reinhard.Sterner@biologie.uni-regensburg.de

Introduction

It has been suggested that complex contemporary protein folds have evolved by the association and fusion of small polypeptide fragments.^{1,2} The internal symmetry observed in a number of enzymes suggests that identical polypeptide segments first assembled to homo-oligomers, followed by the covalent linkage of the fragments as a consequence of gene duplication and fusion.^{3–5} The $(\beta\alpha)_8$ (or

TIM) barrel is the most common enzyme fold, which currently contains about 400 identified members listed in the SCOP database†, and therefore is ideally suited for the study of enzyme evolution.^{6–11} The canonical $(\beta\alpha)_8$ -barrel consists of eight $(\beta\alpha)$ units, each of which comprises a β -strand and a subsequent α -helix which are connected by a $\beta\alpha$ -loop; the individual units are linked by $\alpha\beta$ -loops.^{12–14} The eight β -strands form a central parallel β -sheet (the barrel), which is surrounded by the eight α -helices. Within the central β -barrel, the residues are forming four layers, which are ordered perpendicular to the barrel axis and are numbered from the C- to the N-terminal barrel face. Each layer consists of residues from the odd (1, 3, 5, 7) or the even (2, 4, 6, 8) numbered strands, which generates a 4-fold symmetry. The active site of all known $(\beta\alpha)_8$ -barrel enzymes is harboured by the C-terminal ends of the β -strands and by the $\beta\alpha$ -loops.

The apparent modularity of the $(\beta\alpha)_8$ -barrel suggests that it has evolved by the duplication, fusion and recombination of fragments, which comprised either a single primordial $(\beta\alpha)$ element, or an even numbered multiple thereof. This hypothesis has been supported by protein fragmentation and folding studies.¹¹ For example, the observation that many $(\beta\alpha)_8$ -barrels contain a conserved Gly-X-Asp motif in the loops $\alpha1\beta2$, $\alpha3\beta4$, $\alpha5\beta6$, and $\alpha7\beta8$ but not in the loops $\alpha2\beta3$, $\alpha4\beta5$, and $\alpha6\beta7$ implies that $(\beta\alpha)_2$ -units might constitute the smallest evolving entity.¹⁰ Along these lines, the members of the S-adenosyl-L-methionine (AdoMet) radical protein family have three putative architectures, which comprise four, six and eight $(\beta\alpha)$ -units.¹⁵ The four or six N-terminal units contain the elements for radical generation, whereas the remainder of the protein is crucial for substrate binding. These findings suggest that AdoMet proteins might have evolved by the stepwise fusion of $(\beta\alpha)_2$ units, starting from a $(\beta\alpha)_4$ -half barrel to a $(\beta\alpha)_6$ -three-quarter to a $(\beta\alpha)_8$ -full barrel structure. Along the same lines, fragmentation studies with the α -subunit of tryptophan synthase are compatible with a 6 (4+2) + 2 folding mechanism.^{16,17} Moreover, the phosphoribosyl anthranilate isomerase (PRAI) from yeast forms a stable fragment consisting of the N-terminal $(\beta\alpha)_{1-6}$ units, which associates with the unstructured $(\beta\alpha)_{7-8}$ unit to a functional complex *in vitro*.¹⁸ However, out of three fragment combinations of PRAI from *Escherichia coli* co-expressed *in vivo*, only the non-covalent complex between $(\beta\alpha)_{1-4}$ and $(\beta\alpha)_{5-8}$ yielded a functional enzyme.¹⁹ This finding is in favour of a “4+4” folding and evolution mechanism, which has also been supported by studies on triosephosphate isomerase from chicken and rabbit muscle.^{20,21}

Although the mechanistic details appear to be more complex than suggested by such a simple model,^{22,23} the analysis of the enzymes *N'*-(5'-phosphoribosyl)formimino]-5-aminoimidazole-4-

carboxamide ribonucleotide isomerase (HisA) and imidazole glycerol phosphate synthase (HisF) has provided ample evidence for the crucial role of $(\beta\alpha)_4$ -half barrels in the folding and evolution of $(\beta\alpha)_8$ -barrel enzymes.¹¹ The amino acid sequences and X-ray structures of HisA and HisF from *Thermotoga maritima*, which catalyse two successive transformations of similar bisphosphorylated substrates within histidine biosynthesis, show an internal 2-fold symmetry (Figure 1(a)).^{24,25} The pairs of N-terminal halves (HisA-N and HisF-N) that consist of the first four units ($(\beta\alpha)_{1-4}$) and the pairs of the C-terminal halves (HisA-C and HisF-C) that consist of the last four units ($(\beta\alpha)_{5-8}$) show sequence identities between 16 and 26% and root-mean-square (rms) deviations of their main-chain non-hydrogen bonds between only 1.4 Å and 2.1 Å, respectively. When produced separately, HisF-N and HisF-C form folded and mainly homo-dimeric proteins, which upon co-expression *in vivo* or joint refolding *in vitro* assemble to the non-covalent hetero-dimeric HisF-NC complex that displays wild-type catalytic activity.⁹ Remarkably, four of the five $(\beta\alpha)$ units of several members of the flavodoxin-like fold family show striking similarities to the four $(\beta\alpha)$ units of HisF-N and HisF-C.²⁶ Moreover, the catalytically crucial N-terminal $(\beta\alpha)_4$ -half barrels of eukaryotic and prokaryotic phosphoinositide-specific phospholipases C (PI-PLC) superimpose with rms deviations of 1.85 Å for 104 equivalent C $^{\alpha}$ -atoms.²⁷ Taken together, these findings suggest that $(\beta\alpha)_4$ -half barrels and presumably also $(\beta\alpha)_2$ -quarter barrels are independently evolving units. This conclusion would be further strengthened if stable and functional $(\beta\alpha)_8$ -barrels could be generated by the fusion and recombination of existing half barrels.

Along these lines, we recently generated and characterised the stable chimeric $(\beta\alpha)_8$ -barrels HisAF and HisFA, in which the N-terminal half of HisA was linked to the C-terminal half of HisF, and *vice versa*.²⁸ Moreover, we duplicated and fused in tandem the *hisF-C* gene, yielding the HisF-CC protein (Figure 1(b)). Using rational protein design, the interface between the two identical halves in HisF-CC was stabilised by reconstituting a salt bridge cluster, which is conserved in all known HisF proteins. Although the resulting construct HisF-C*C containing the mutations A124R and A220K (Figure 1(b)) is a predominantly monomeric and a more compact protein than HisF-CC, it was produced in *E. coli* as inclusion bodies, had a slight tendency to aggregate, and differed in its spectroscopic properties from wild-type HisF.²⁸ To further increase the stability and solubility of HisF-C*C, we have now applied random mutagenesis followed by selection *in vivo* using a chloramphenicol acetyltransferase (CAT) reporter system.²⁹ The combination of several beneficial mutations identified by this approach allowed us to generate the HisF-C**C protein (Figure 1(c)), which can be produced in soluble form in *E. coli*, forms a homogeneous and compact monomer in gel filtration experiments, shows an identical fluorescence spectrum as wild-

† scop.mrc-lmb.cam.ac.uk/scop/

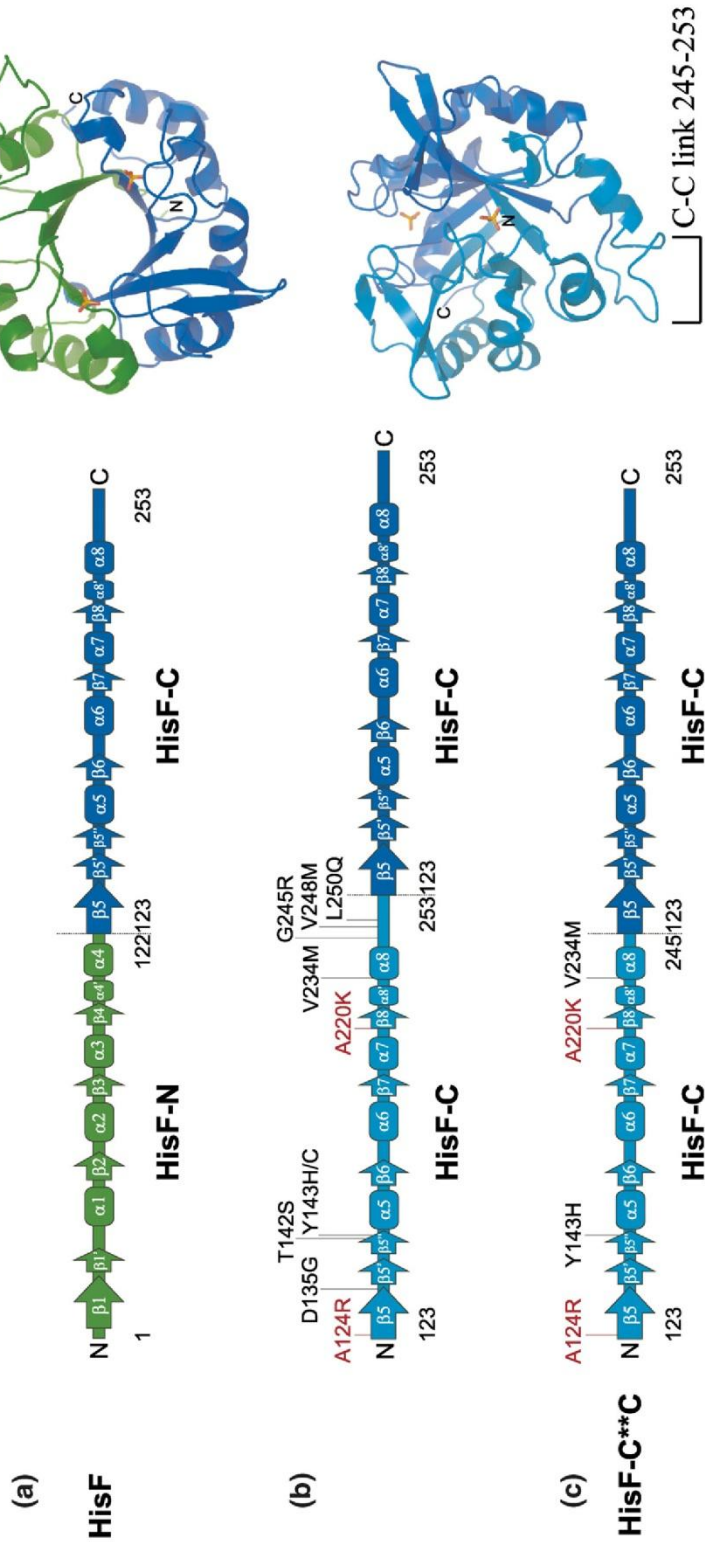


Figure 1. Design of the $(\beta\alpha)_8$ -barrel protein HisF-C**C from two identical and fused HisF-C half barrels. (a) Secondary structure elements and tertiary structure (PDB code 1thf) of wild-type HisF (ribbon diagram showing a view onto the C-terminal face of the central β -barrel). The N-terminal half barrel HisF-N $[(\beta\alpha)_{1-4}]$ is shown in green, and the C-terminal half barrel HisF-C $[(\beta\alpha)_{5-8}]$ is shown in blue. The two phosphate ions bound to HisF-N and HisF-C are depicted as tetrahedrons. The N- and C-termini of the polypeptide chain are marked. (b) The fusion of two HisF-C units (the N-terminal one depicted in cyan, the C-terminal one depicted in black) yielded HisF-C**C. The two exchanges A124R and A220K leading to the reconstitution of the native salt bridge cluster in HisF-C**C are shown in red, and the exchanges introduced by random mutagenesis and selected *in vivo* in this work are shown in black. A model of the HisF-C**C scaffold (ribbon diagram showing a side view onto the central β -barrel) with the long loop linking the two HisF-C units being highlighted is shown on the right. The two phosphate ions bound to the N and C-terminal half barrels are depicted as tetrahedrons. The N- and C-termini of the polypeptide chain are marked. (c) The optimised variant HisF-C**C contains a shortened loop and four amino acid substitutions.

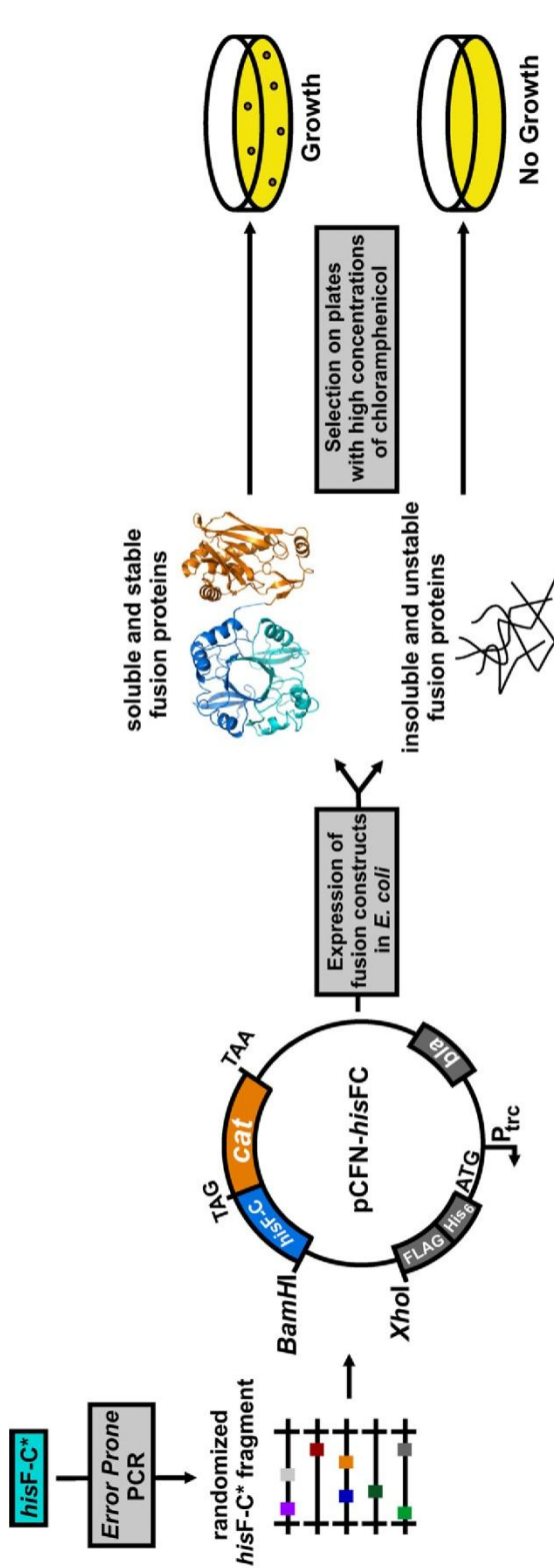


Figure 2. Scheme of the experimental protocol to isolate stabilized HisF-C**C*-CAT fusion proteins by random mutagenesis and selection *in vivo*.²⁹ See the text for details.

type HisF, unfolds with high cooperativity, and contains an intact phosphate binding site. Molecular dynamics simulations provide a structural explanation for the observed stability/solubility increase (HisF-CC < HisF-C*C < HisF-C**C < HisF), which seems to be due to the formation of beneficial electrostatic interactions and the optimisation of van der Waals interactions at the interface of the two fused half barrels.

Results

Random mutagenesis and selection of improved HisF-C*C variants

The method used to increase the stability of HisF-C*C is based on the observation that cells expressing fusions of an unstable and insoluble protein to CAT exhibit decreased resistance to chloramphenicol compared to fusions with stable proteins (Figure 2).²⁹ To generate improved variants of HisF-C*C, the N-terminal half of the *hisF-C*C* gene (*hisF-C**) was subjected to random mutagenesis using error-prone PCR. The resulting *hisF-C** mixture was ligated into a modified pCFN vector, upstream of an unchanged wild-type *hisF-C* fragment fused to the *cat* gene. The resulting plasmid repertoire was first amplified in the *E. coli* strain XL1-BlueMrf', and then used to transform *E. coli* JM101 cells which were streaked on medium plates containing chloramphenicol at concentrations of 0 $\mu\text{g/ml}$ (control), 300 $\mu\text{g/ml}$, 350 $\mu\text{g/ml}$ and 400 $\mu\text{g/ml}$. The sequencing of ten *hisF-C*C* gene variants isolated from the control plate showed that the error-prone PCR had, on average, introduced two nucleotide exchanges. However, colony PCR showed that all 35 clones grown on plates with the highest chloramphenicol concentration of 400 $\mu\text{g/ml}$ were false-positives, which had lost the major part of their *hisF-C*C* insert. Moreover, from the approximately 300 colonies grown in the presence of 300 and 350 $\mu\text{g/ml}$ chlor-

amphenicol, only 17 allowed to amplify a PCR product of the expected length. The sequencing of these clones yielded eleven different *hisF-C*C* gene variants, which were sub-cloned without the *cat* gene into the expression plasmid pET24a (+) such that a C-terminal His₆-tag was added to the recombinant protein. The HisF-C*C variants were produced by heterologous gene expression in *E. coli* BL21 (DE3) and were found mainly in the insoluble fraction of the cell extract. Nine variants were solubilised in guanidinium chloride, refolded in phosphate buffer and characterised. Two variants, HisF-C*C 2 and HisF-C*C 4, could be purified from the soluble fraction of the cell extract using metal chelate affinity chromatography.

Stability and solubility of selected variants

From the 11 purified HisF-C*C variants, seven turned out to be no more stable in the absence of CAT than the parent HisF-C*C protein (data not shown) and were therefore not further investigated. The remaining four variants (HisF-C*C 1-4), which contained between one and four amino acid exchanges, unfolded at higher concentrations of urea and with a higher cooperativity than HisF-C*C (Table 1; Figure 3(a)). The stability of HisF-C*C 4, which carries, besides two other replacements, the Y143C exchange, was decreased in the presence of 1 mM of the reducing agent dithiothreitol (DTT). This finding suggested that HisF-C*C 4 can form a stabilizing intermolecular disulfide bond under oxidizing conditions. This hypothesis was confirmed by analytical gel filtration, which demonstrated that HisF-C*C 4 is mainly dimeric in the absence of DTT but forms a pure monomer in its presence (Figure 3(b)). The variant HisF-C*C 2 is exclusively monomeric and elutes from the gel filtration column as a symmetrical peak (Figure 3(b)). In contrast, the elution peaks of HisF-C*C, HisF-C*C 1 and HisF-C*C 3 are much less symmetrical, suggesting that a considerable fraction of these proteins forms ill-defined oligomers. In

Table 1. Results of urea induced unfolding of HisF-C*C variants

Variant	Amino acid exchanges in N-terminal half (HisF-C*)	$D_{1/2}$ (M) ^a	m^{app} (kJ/mol*M) ^b
HisF-C*C		4.3	3.7
HisF-C*C 1	V234M+G245R	4.3	5.0
HisF-C*C 2	D135G+T142S+Y143H+L250Q	4.7	6.3
HisF-C*C 3	V234M	4.4	4.1
HisF-C*C 4	Y143C+V234M+V248M	4.8 (5.4) ^c	6.7 (6.5) ^c
HisF-C*C Δ	Short $\alpha_8\beta_5$ -loop	4.6	5.8
HisF-C*C Δ +D135G	Short $\alpha_8\beta_5$ -loop+ D135G	4.5	4.9
HisF-C*C Δ +T142S	Short $\alpha_8\beta_5$ -loop+ T142S	4.5	5.6
HisF-C*C Δ +Y143H	Short $\alpha_8\beta_5$ -loop+ Y143H	4.9	8.4
HisF-C*C Δ +V234M	Short $\alpha_8\beta_5$ -loop+ V234M	4.8	7.5
HisF-C**C	Short $\alpha_8\beta_5$ -loop+ Y143H+ V234M	5.2 (4.8) ^d	11.6 (10.6) ^d

^a $D_{1/2}$: urea concentration at which 50% of the protein is non-native.

^b m^{app} (apparent m -value): operational measure for the cooperativity of unfolding, obtained by analyzing normalized unfolding traces with the two-state model.⁶⁶

^c The values of HisF-C*C 4 were determined in the presence of 1 mM DTT (and in its absence).

^d The values of HisF-C**C were determined in 50 mM potassium phosphate (pH 7.5) (and in 50 mM Tris-HCl (pH 7.5)).

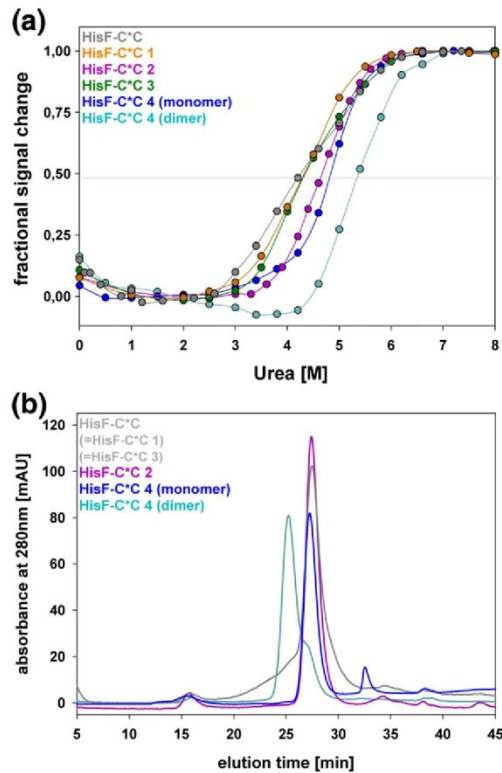


Figure 3. Stability and homogeneity of selected variants HisF-C*C 1-4. (a) Traces of reversal chemical unfolding. Proteins at a concentration of 5 μ M were incubated with the given concentrations of urea in 50 mM potassium phosphate (pH 7.5) at 25 °C. Unfolding was followed by the decrease of the fluorescence emission at 320 nm after excitation at 280 nm, and the signals obtained after equilibration were normalized. The $D_{1/2}$ (urea) and apparent m -values of the variants are listed in Table 1. (b) Elution profiles of analytical gel filtration. Each protein (0.15–0.2 mg) was applied on a Superdex 75 (Amersham) column equilibrated with 50 mM potassium phosphate (pH 7.5), containing 300 mM potassium chloride, at 25 °C. The elution times of the main peaks of HisF-C*C and HisF-C*C 2 correspond to the monomers, and the shape of the flanking region of HisF-C*C indicates the population of ill-defined higher oligomers. The HisF-C*C 4 variant was assayed both in the presence of 1 mM DTT (monomer populated), and in its absence (disulfide-linked dimer populated).

accordance with these findings, more than 10% of recombinant HisF-C*C 2 and HisF-C*C 4 but less than 3% of recombinant HisF-C*C, HisF-C*C 1 and HisF-C*C 3 are produced in soluble form in *E. coli* (Figure 4). These data demonstrate that the HisF-C*C 2 and HisF-C*C 4 variants are significantly more stable and soluble than their parent HisF-C*C, whereas the substitutions present in the HisF-C*C 1 and HisF-C*C 3 result in only minor improvements.

Three out of the eight residue substitutions found in HisF-C*C 1-4 are located in the loop connecting the two half barrels, which corresponds to the

C-terminal extension of HisF-C (Figure 1(b)). This finding suggested that this long stretch of wild-type sequence is destabilising. The loop was therefore shortened from 13 to five residues (Figure 1(c)), and the stability and solubility of the resulting variant HisF-C*C Δ was investigated. Moreover, HisF-C*C Δ was used as a scaffold to test the effect of the substitutions D135G, T142S, Y143H and V234M individually. As deduced from urea-induced unfolding, the stabilizing effect of the exchanges is as follows: Δ +D135G \approx Δ +T142S $<$ Δ $<$ Δ +V234M $<$ Δ +Y143H (Table 1; Figure 5(a)). The effect of these substitutions on the percentage of recombinant protein found in the soluble fraction of the cell extract is somewhat different: Δ +D135G $<$ Δ \approx Δ +T142S \approx Δ +V234M $<$ Δ +Y143H (Figure 4). These results show that the shortening of the loop between the two HisF-C half barrels improves both the stability and solubility of HisF-C*C. From the individual substitutions, D135G and T142S are neutral at best and were isolated only accidentally in the context of HisF-C*C 2. The exchanges Y143H and V234M are increasing the stability, whereas Y143H is by far the most crucial substitution for the improvement of solubility.

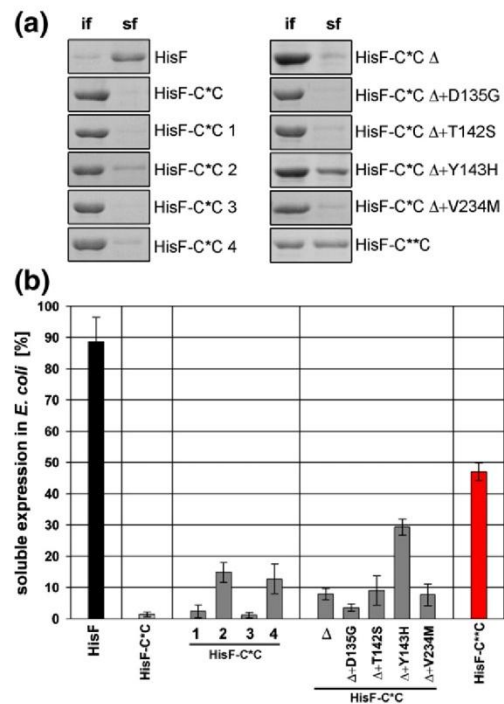


Figure 4. Solubility analysis of selected and designed HisF-C*C variants. (a) Insoluble (if) and soluble (sf) fractions of the different variants as found in cell extracts of the *E. coli* expression strain BL 21(DE3) (SDS-PAGE (12.5% acrylamide) stained with Coomassie blue). (b) Percentage of the proteins produced in soluble form (as deduced from the analysis of three different clones of each variant). The protein bands were scanned, and their intensities were determined using Optiquant (version 2.5).

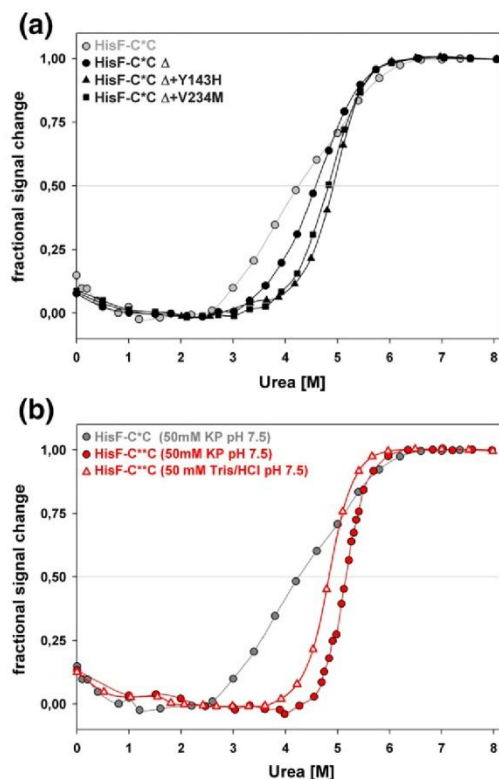


Figure 5. Stability of designed HisF-C*C variants. (a) The unfolding traces of the variants HisF-C*C Δ , HisF-C*C Δ +Y143H and HisF-C*C Δ +V234M are plotted in comparison to HisF-C*C. Proteins at a concentration of 5 μ M were incubated with the given concentrations of urea in 50 mM potassium phosphate (KP) (pH 7.5), at 25 $^{\circ}$ C. (b) To test the stabilising effect of bound phosphate ions, HisF-C**C was unfolded in 50 mM KP (pH 7.5) and in 50 mM Tris-HCl (pH 7.5). Reversible unfolding was followed by the decrease of the fluorescence emission at 320 nm after excitation at 280 nm, and the signals obtained after equilibration were normalized. The $D_{1/2}$ (urea) and apparent m -values of the variants are listed in Table 1.

Generation and characterisation of HisF-C**C

In order to produce an even more beneficial HisF-C*C variant, the Y143H and V234M substitutions were combined on the HisF-C*C Δ scaffold. The resulting variant HisF-C**C was produced to 50% as a soluble protein (Figure 4) that unfolds at higher concentrations of urea than any of the originally isolated HisF-C*C variants (Table 1). Its stability is elevated in phosphate compared to Tris buffer (Figure 5(b)), indicating that it possesses two phosphate binding sites as found in wild-type HisF (Figure 1). In order to compare the stability of HisF-C**C with that of HisF, which cannot be completely denatured in urea, unfolding was induced by the stronger denaturant guanidinium chloride. The measurements showed that HisF-C**C is less stable than wild-type HisF but unfolds with

higher cooperativity (Figure 6(a)), indicating that the newly generated variant is as compact as the native protein. In accordance with this finding, HisF-C**C and HisF elute from the gel filtration column as equally homogeneous monomers, whose peaks are much sharper than that of HisF-C*C (Figure 6(b)).

To compare the tertiary structures of HisF-C*C and HisF-C**C with that of wild-type HisF, fluorescence emission and near-UV circular dichroism spectra were recorded. The fluorescence spectra are dominated by the single tryptophan residue 156 in helix α 5 of HisF, which exists as two copies in HisF-C*C and HisF-C**C. The structural environment of its indole side-chain is hydrophobic in wild-type HisF (Ala106, Val107, and Ile113 from helix α 4 of the adjacent HisF-N half barrel) as well as in HisF-C*C and HisF-C**C (Phe 227, His228, Val234/Met234 from helix α 8 of the adjacent HisF-C half barrels), and should therefore be shielded from solvent in an intact ($\beta\alpha$)₈-barrel structure. The fluorescence emission maxima of wild-type HisF and HisF-C**C lie at 322 and 323 nm, respectively, whereas the emission maximum of HisF-C*C lies at 327 nm (Figure 6(c)). These results demonstrate that the two tryptophan 156 residues in HisF-C**C are equally well shielded from solvent as the single tryptophan 156 in wild-type HisF. In contrast, the red-shifted emission maximum of HisF-C*C shows that at least one of its tryptophan residues must be surrounded by a more polar environment. Remarkably, the gel filtration profiles and fluorescence emission spectra of HisF-C*C Δ , HisF-C*C Δ +Y143H and HisF-C*C Δ +V234M are identical to those of HisF-C*C (data not shown), demonstrating that the combination of the beneficial substitutions is required to completely shield the tryptophan residues. The pronounced near-UV CD spectra of HisF-C*C and HisF-C**C have a similar shape as the spectrum of wild-type HisF (Figure 6(d)), showing that the aromatic residues are immobilised in an asymmetric environment in both halves. This result proves that HisF-C*C and HisF-C**C do not form "molten globules" but adopt a well-defined tertiary structure.

The unfolding of a great number of proteins with known X-ray structure has shown that the amount of protein surface exposed to solvent (Δ ASA) correlates with the denaturation m -value.³⁰ Moreover, Δ ASA also correlates with the number of residues of a globular protein. As a consequence, the measured apparent m -values of the various HisF-C*C constructs allowed us to estimate the corresponding Δ ASA values. The HisF-C*C and HisF-C*C Δ scaffolds contain 274 and 266 residues, corresponding to calculated Δ ASA values of 24,575 and 23,831 \AA^2 , respectively. The apparent m -values measured for most HisF-C*C variants indicate that their Δ ASA values are smaller by a factor of 2–4, suggesting that the folded states of these proteins do not completely shield their hydrophobic interior from solvent. This result indicates, together with the gel filtration and fluorescence data (Figure 6(b) and (c)), that monomeric HisF-C*C is in equilibrium with

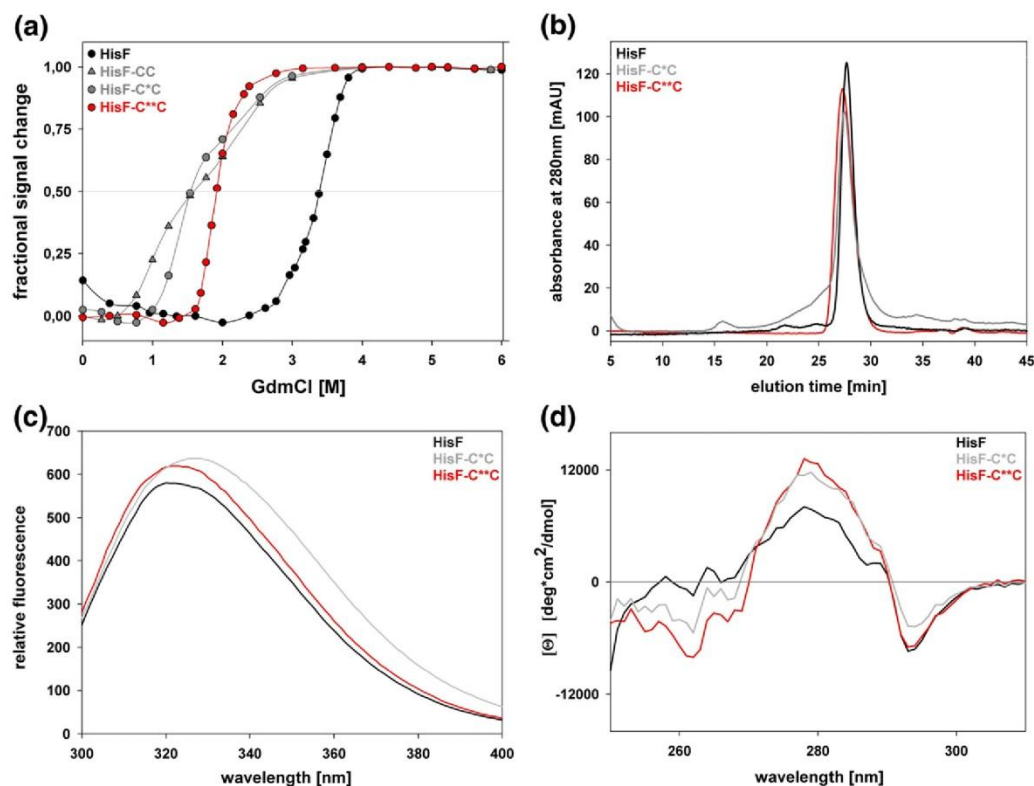


Figure 6. Results of the stepwise optimisation of fused HisF-C half barrels in comparison to wild-type HisF. (a) Traces of reversible chemical unfolding. Proteins at a concentration of 5 μM were incubated with the given concentrations of GdmCl in 50 mM potassium phosphate (pH 7.5), at 25 $^{\circ}\text{C}$. Unfolding was followed by the decrease of the fluorescence emission at 320 nm after excitation at 280 nm, and the signals obtained after equilibration were normalised. The transition midpoints of both HisF-CC and HisF-C*C lie at 1.5 M GdmCl; the multi-phasic unfolding traces did not allow to determine apparent m -values. The transition midpoints of HisF-C**C and wild-type HisF lie at 1.9 M and 3.4 M GdmCl, respectively; the unfolding of these variants is highly cooperative with apparent m -values of 17.3 $\text{kJ mol}^{-1} \text{M}^{-1}$ for HisF-C**C and 14.2 $\text{kJ mol}^{-1} \text{M}^{-1}$ for wild-type HisF. (b) Elution profiles of analytical gel filtration. Each protein (0.15–0.2 mg) was applied on a Superdex 75 (Amersham) column equilibrated with 50 mM potassium phosphate (pH 7.5), containing 300 potassium chloride, at 25 $^{\circ}\text{C}$. The elution times of the main peaks correspond to the monomers, and the shape of the flanking region of HisF-C*C indicates the population of ill-defined higher oligomers. (c) Fluorescence emission spectra. The spectra were monitored after excitation at 280 nm of 5 μM protein in 50 mM potassium phosphate (pH 7.5), at 25 $^{\circ}\text{C}$. The measured emission intensity of HisF, which is mainly caused by its single Trp156, was multiplied by a factor of 2 in order to compensate for the presence of two Trp156 residues in HisF-C*C and HisF-C**C. The emission maxima are as follows: HisF: 322(\pm 2) nm; HisF-C**C: 323(\pm 2) nm; HisF-C*C: 327(\pm 1) nm. The values given are the mean and standard deviation of three independent measurements. (d) Near-UV CD spectra. The spectra shown are the mean of ten individual spectra, which were recorded at 25 $^{\circ}\text{C}$ in 50 mM potassium phosphate (pH 7.5).

a small fraction of a partially unfolded and aggregation-prone conformation. In contrast, the deduced ΔASA value of HisF-C**C of 21,782 \AA^2 indicates that this variant is a compactly folded protein with a buried hydrophobic core, again in accordance with the results from analytical gel filtration and fluorescence emission spectroscopy (Figure 6(b) and (c)).

Molecular dynamics simulations

In order to gain insight into the mechanism of stabilisation caused by the various amino acid substitutions, the crystal structure of HisF (PDB

code 1thf)²⁵ was used as the basis to construct models of duplicated and fused half barrels. To start with, two HisF-C units were covalently linked, and the resulting HisF-CC variant was then relaxed by 2 ns of molecular dynamics (MD) simulation at 300K using the MAB-force field as implemented in MOLOC[‡].³¹ Remarkably, the presence of two bound phosphate ions in the simulations, one in each half barrel as observed in the crystal structure of HisF (Figure 1), stabilised HisF-CC as well as the other variants. Simulations performed without bound phosphate

[‡] <http://www.moloc.ch>

led to the collapse of the phosphate binding site and a partially distorted protein backbone. Starting from HisF-CC, structures of the variants HisF-C*C, HisF-C*C Δ , and HisF-C**C were stepwise calculated and relaxed by MD, which resulted in the model displayed in Figure 7(a). The gain in stability caused by the exchanges introduced at the half barrel interface ("interaction energy") was analysed using a pseudo-thermodynamic cycle. To this end, the energies of the two isolated, solvated $(\beta\alpha)_4$ -half barrels were subtracted from the energy of the intact $(\beta\alpha)_8$ -barrel structure.³² The calculated interaction energy of HisF-CC (440(\pm 10) kJ/mol) is smaller than

that of HisF-C*C (565(\pm 10) kJ/mol), confirming the stabilizing effect of the reconstituted wild-type salt bridge cluster in layer 4 of the central β -barrel of HisF-C*C²⁸ (Figure 7(b)). Moreover, the deduced interaction energy of 125 kJ/mol for this cluster is reasonable for a network of four charged residues.³³ The calculated interaction energy of HisF-C*C Δ (557(\pm 10) kJ/mol) is identical within error to that of HisF-C*C, indicating that the shortening of the connecting loop between the two half barrels does not influence their interaction. In contrast, the presence of the two exchanges Y143H and V234M at the half barrel interface of HisF-C**C leads to an

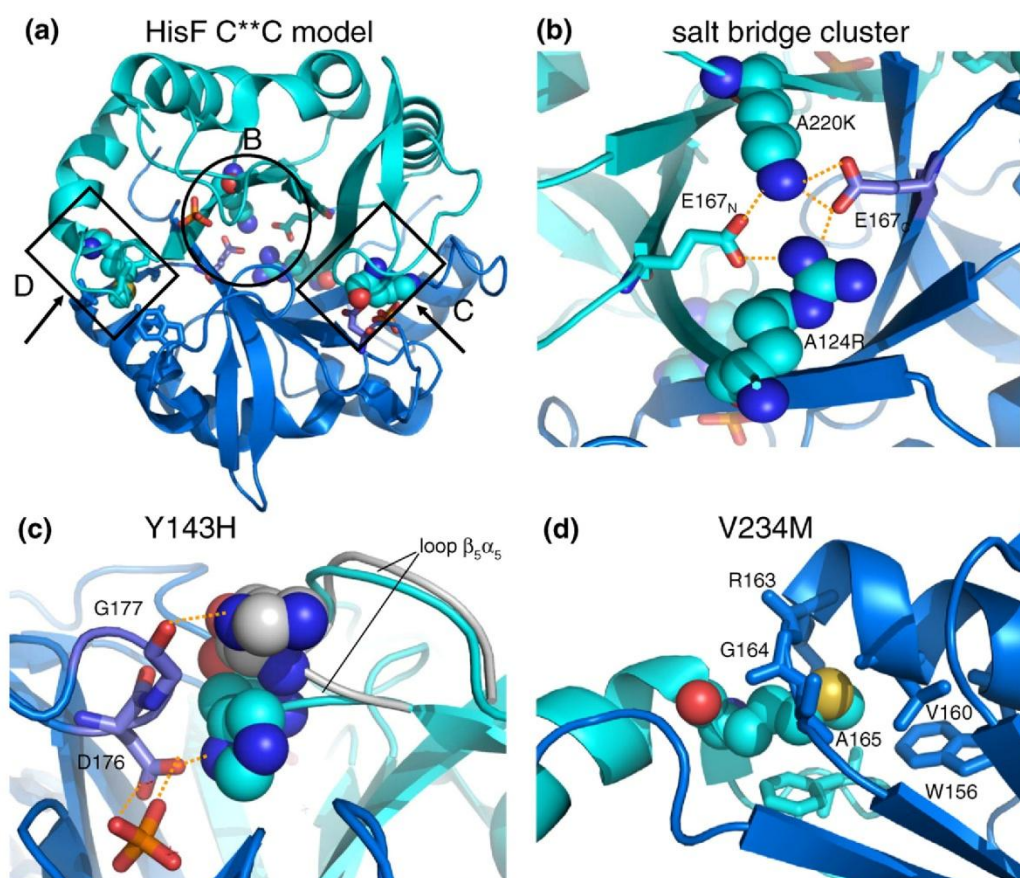


Figure 7. Structural model of HisF-C**C based on MD simulations. (a) A top view onto the C-terminal face of the central β -barrel is shown as a ribbon diagram. The introduced stabilizing residues are highlighted as van der Waals-spheres, and other important residues are shown as sticks. The two phosphate ions bound to the N- and C-terminal half barrels are depicted as tetrahedrons. The colour code is the same as for Figure 1(c); the mutated N-terminal half barrel is shown in cyan, and the wild-type C-terminal half barrel is shown in blue. The arrows indicate the direction of the side views shown in (c) and (d). (b) The reconstituted salt bridge cluster forming layer 4 of the central β -barrel is composed of N- and C-terminal glutamate residues (E167_N, E167_C) as well as arginine and lysine residues introduced into the N-terminal half (A124R and A220K). H-bonds are represented by dotted orange lines. (c) The histidine residue introduced into the N-terminal half barrel (Y143H) forms a H-bond network with Asp176 and the phosphate ligand of the C-terminal half barrel. In a possible alternative conformation, His143 forms a H-bond with Gly177 and is solvent exposed. The corresponding alternative closed and open conformations of loop $\beta_5\alpha_5$ (residues 141–149) and the carbon atoms of the imidazole side-chain of His143 in the two conformations are shown in cyan and grey, respectively. (d) The buried methionine residue introduced into the N-terminal half barrel (V234M) is protruding into a hydrophobic pocket of the C-terminal half barrel, where it comes in van der Waals contact with Trp156, Val160, Arg163, Gly164 and Ala165.

increase by 42 kJ/mol in the calculated interaction energy to 599(±10) kJ/mol. The histidine residue at position 143 is located in loop β 5 α 5 near the phosphate-binding site of the C-terminal half barrel. During the molecular dynamics relaxation a hydrogen bond network between His143, the bound phosphate, the conserved crystal water molecules next to them and Asp176 was established, which provides about 30 kJ/mol of interaction energy. However, the simulations suggest that an alternative conformation exists in which the imidazole side-chain of His143 is solvent-exposed and forms an H-bond with the carbonyl oxygen of Gly177 (Figure 7(c)). The introduced methionine at position 234 is in hydrophobic contact with several residues of the C-terminal half barrel (Val160, Arg163, Gly164, Ala165; Figure 7(d)). The energy decomposition suggests mainly hydrophobic interactions of about 12 kJ/mol contributed by the van der Waals energy of the MAB force field.³⁴ The calculated interaction energy between HisF-N and HisF-C within wild-type HisF is 638(±10) kJ/mol. Taken together, the results of the MD simulations qualitatively reflect the experimentally determined stability increase from HisF-CC *via* HisF-C*C and HisF-C**C to HisF, indicating that the beneficial amino acid substitutions can be interpreted as an improved interaction between the fused half barrels.

Discussion

Structurally defined and stable domains with a given function constitute the basic building blocks of proteins.^{35,36} It follows that the duplication, fusion and recombination of domain-encoding genes are the evolutionary mechanisms used by nature to generate complex multi-domain proteins.³⁷ When analysing the evolution of the versatile and ubiquitous ($\beta\alpha$)₈-barrel proteins, we noticed that some representatives of this highly symmetrical fold consist of sub-domains comprising ($\beta\alpha$)₄-half barrels.⁹ In an attempt to reconstruct the evolution of the ($\beta\alpha$)₈-barrel HisF from the ($\beta\alpha$)₄-half barrel HisF-C, we earlier generated and characterised the HisF-CC construct which consists of two identical half barrels linked by a Gly-Ser-Gly linker. Using rational design, HisF-CC was optimised to HisF-C*C by reconstituting the native salt bridge cluster which forms layer 4 at the N-terminal face of the central β -barrel of HisF.²⁸ Since model building did not point to other beneficial amino acid exchanges to be introduced by rational design, we used a combination of random mutagenesis and selection *in vivo* to further stabilise HisF-C*C.

Pitfalls of the selection system

Various directed evolution methods for increasing protein solubility and stability have been described.^{38–42} We chose an approach in which a mixture of *hisF-C*C* genes with a randomised N-terminal half was fused to the *cat* gene, and

selected for HisF-C*C-CAT fusion proteins which allowed their host cells to grow in the presence of high concentrations of chloramphenicol.^{29,43,44} Although the original vector pCFN1²⁹ was modified such that without ligation of the randomised N-terminal *hisF-C** fragment a frameshift was generated, 100% of the variants selected in the presence of 400 μ g/ml and about 95% of the variants selected in the presence of 350 μ g/ml and 300 μ g/ml chloramphenicol were false-positives. These clones carried plasmids with extensively truncated *hisF-C*C* mutants in frame with the *cat* gene, generating CAT proteins with short N-terminal extensions that did not compromise its high solubility. We speculate that these deletions have mainly occurred during the cloning procedure of the *hisF-C** library because cells transformed with the empty pCFN-*hisFC* vector (Figure 2) were highly sensitive to chloramphenicol. Probably the false-positives were generated by unspecific digestions of the used restriction enzymes XhoI and BamHI, and were strongly enriched as a result of the high selection pressure. However, we cannot exclude that some of the observed deletions were caused by recombination events in *E. coli*. The sequencing of a number of false-positive clones showed that the inserts were diverse, which precluded the identification and subsequent removal of unwanted restriction “hot spots” by site directed mutagenesis. Moreover, from the 11 selected unique variants with a complete *hisF-C*C* insert, in the absence of CAT seven proteins did not show a higher stability or solubility than their parent HisF-C*C. The reliability of screening methods using fused reporter proteins frequently suffers from unexpected interactions between reporter and target, which can influence the outcome of the assay. In the case of CAT, which is an obligate trimer, the linkage to a target protein might lead in certain cases to higher-order complexes with unpredictable effects on the solubility of the fusion constructs.^{40,42}

Structural basis of stabilisation

From the true-positive variants, HisF-C*C 1 and 3 were slightly and HisF-C*C 2 and 4 were considerably more stable and soluble than HisF-C*C (Table 1; Figures 3 and 4). These proteins altogether contained eight different amino acid exchanges of which three (G245R, V248M, L250Q) were located in the long $\alpha\beta$ loop connecting the two HisF-C units (Figure 1(b)). The shortening of this loop from 13 to five residues in HisF-C*C Δ increased the stability and the solubility of the protein significantly (Table 1; Figures 4 and 5). This result is in line with the observation that $\alpha\beta$ loops of ($\beta\alpha$)₈-barrels are generally short and contribute to the conformational stability of the fold.¹¹ Along the same lines, a number of proteins from hyperthermophiles carry shorter loops compared to their homologues from mesophiles, which causes a decrease of the entropy difference between the unfolded and the folded state that results in an increase of the stabilisation free energy.^{45,46} Remarkably, the deletion of another two

residues resulting in a loop containing only the introduced Gly-Ser-Gly linker decreased stability and solubility (data not shown), possibly due to the introduction of conformational strain.

The tyrosine residue at position 143 was replaced in the two best variants, with cysteine in HisF-C^{*}C 4 and with histidine in HisF-C^{*}C 2 (Table 1). The substitution Y143C in HisF-C^{*}C 2 leads to the formation of an intermolecular disulfide bridge (Figure 3(b)). However, the cytoplasm of *E. coli* is reducing and HisF-C^{*}C 2 is more stable than HisF-C^{*}C also in the presence of DTT (Figure 3(a)). It is therefore unclear whether the cysteine was formed already in the host cell during the *in vivo* selection or only upon contact of the protein with air oxygen during its purification. However, the X-ray structures of a number of cytoplasmic enzymes from hyperthermophilic Archaea suggest that cysteine residues can contribute to extreme protein thermostability of intracellular proteins,^{47–50} an hypothesis which is supported by the computational analysis of microbial genomes.^{51,52} The introduction of new S-S bonds⁵³ is a widely used protein stabilisation strategy, and the beneficial effect of S-S bonds is ascribed to the decrease of the entropy of the unfolded state.^{54–58} In addition, the detected disulfide bond in HisF-C^{*}C 2 could also stabilise the protein by increasing its association state from the monomer to the dimer. Along these lines, a number of $(\beta\alpha)_8$ -barrels and other proteins from hyperthermophiles are higher-order oligomers than their counterparts from mesophiles.^{59–61} The structural basis of the stabilising effect of the Y143H exchange is less clear. Although the MD simulations suggest that the introduced imidazole ring leads to the formation of a hydrogen bonding network (Figure 7(c)), alternative mechanisms are feasible. Since the imidazole side-chain of histidine is more polar than the aromatic ring of tyrosine, its higher solvation energy could stabilise the native state. Along these lines, position 143 is located in the long loop connecting strand β_5 with helix α_5 , which was highly flexible in the simulations and might easily become solvent-exposed. Accordingly, we have frequently observed in the simulations geometries in which the imidazole ring is not in contact to Asp176 and the bound phosphate ion but forms an H-bond with the carbonyl group of Gly177 at the surface of the protein (Figure 7(c)). In accordance with this finding, the Y143H exchange leads to a strong improvement of the solubility of HisF-C^{*}C (Figure 4). The substitution V234M was found in three of the four improved HisF-C^{*}C variants and is the single exchange present in HisF-C^{*}C 3 (Table 1). The MD simulations suggest that the long and linear hydrophobic methionine side-chain protrudes into a patch of hydrophobic residues of the adjacent half barrel (Figure 7(d)). This newly formed hydrophobic cluster is well shielded from solvent, which explains why HisF-C^{*}C 3 and HisF-C^{*}C Δ +V234M are no more soluble than HisF-C^{*}C and HisF-C^{*}C Δ , respectively (Figure 4). However, HisF-C^{**}C is more soluble than HisF-C^{*}C Δ +Y143H, suggesting

that the effect of the V234M exchange can be influenced by the protein scaffold in which it is introduced. Along the same lines, the increase of the apparent *m*-value caused by the V234M substitution is somewhat more pronounced in the background of HisF-C^{*}C Δ +Y143H than in the background of HisF-C^{*}C Δ and much stronger than in the background of HisF-C^{*}C (Table 1). The basis of this context dependence is difficult to elucidate in the absence of highly resolved X-ray structures of the different variants. One can speculate that the presence of the Y143H exchange and the shortening of the loop induce a better defined local structure in which methionine 234 can unfurl its stabilizing effect more efficiently.

Implications for the evolution of $(\beta\alpha)_8$ -barrels

The structural superposition of the wild-type HisF half barrels HisF-N and HisF-C reveals that about 100 of the approximately 120 aligned residues are different.²⁵ It is therefore remarkable that four amino acid substitutions in the N-terminal half of HisF-CC were sufficient to generate the HisF-C^{**}C protein, whose tertiary structure must be highly similar to that of HisF. Moreover, neither the Y143H nor the V234M substitution in the N-terminal half of HisF-C^{**}C reconstituted residues found at the equivalent positions in HisF-N. These findings strengthen the hypothesis that half barrels are flexible and independently evolving structural entities, which can be mixed and matched by nature to generate the highly symmetrical and catalytically versatile and ubiquitous TIM-barrel fold.^{10,11}

Earlier attempts to design $(\beta\alpha)_8$ -barrels from individual $(\beta\alpha)$ units mainly generated molten globules that contained a considerable amount of secondary structure but did not fold into well-defined tertiary structures.¹¹ However, recently an idealised $(\beta\alpha)_8$ -barrel with native-like properties was generated on the basis of a two-step computational approach.⁶² Since the central eight-stranded β -sheet was constructed on the basis of a 4-fold symmetry, the successful design supports the hypothesis that $(\beta\alpha)_2$ units constitute the smallest independently evolving units within $(\beta\alpha)_8$ -barrels.^{10,15} However, it has not been reported whether this protein is a monomer with a tertiary structure that is equally well defined and compact as that of HisF-C^{**}C.

In HisF, the catalytically essential residues aspartate 11 in β -strand 1 and aspartate 130 in β -strand 5 are related by a 2-fold symmetry axis along the central barrel axis and perfectly superimpose when HisF-N and HisF-C are structurally aligned.^{9,25,63} HisF-C^{**}C carries two structurally equivalent aspartate residues in its two β -strands 5, and both chemical unfolding and MD simulations suggest that HisF-C^{**}C possesses intact sites to bind the two terminal phosphate moieties of the HisF substrate *N'*-[(5'-phosphoribulose)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (PRFAR) (Figures 5(b) and 7(a)). Nevertheless, HisF-C^{**}C did not show measurable catalytic

activity, even at a protein concentration of 50 μM (data not shown). Probably, minor perturbations of the active site geometry of HisF-C^{*}C prevent the productive binding and turnover of PRFAR, which shall now be established by further rounds of random mutagenesis and selection *in vivo* supported by molecular dynamics studies *in silico*.

Materials and Methods

Construction of a *hisF-C**C gene library with randomised N-terminal half

The chloramphenicol-acetyltransferase (CAT) fusion vector pCFN-*hisFC* (Figure 2) was derived from pCFN1.²⁹ This vector includes phage fl and pBR322 origins of replication, the ampicillin resistance gene (*bla*) and a strong isopropyl- β -thiogalactoside (IPTG)-inducible promoter (*P*_{lac}). Sequences inserted into the multiple cloning site (MCS) of pCFN1 between the restriction sites BglII and XbaI lead to the expression of proteins fused at their N-terminus with a His₆-tag and the FLAG epitope, and at their C-terminus with CAT. An amber stop-codon is inserted between the XbaI site and the *cat* gene such that an amber suppressor *E. coli* strain is needed for expression of the full-length fusion gene. pCFN-*hisFC* was constructed in two steps. First the 5' BglII restriction site of the MCS was changed into a XhoI site by using the QuikChange™ Site-Directed Mutagenesis Kit (Stratagene) with the complementary oligonucleotides 5'-CTACAAG-GACGATGACGACTCGAGCTCACGTCTAGA-3' and 5'-GCCGAGCCTACTCTAGACGTGAGCTCGAGTCGTCATC-3' as overlapping primers,⁶⁴ yielding pCFN1(XhoI) (new restriction site underlined). Next *hisF-C* was amplified by PCR using a *hisF-C* containing pCFN1 vector (pCFN1-*hisFC*) as template and the oligonucleotides 5'-ATACTCGAGTTGGGATCCGGTCAGGCCGTTGTC-3' with XhoI and BamHI sites (underlined) as 5'-primer, and 5'-CTCGCGAGCCTACTCTAGAAA-3' with a XbaI site (underlined) as 3'-primer. The amplified *hisF-C* fragment was cloned into pCFN1(XhoI) using XhoI and XbaI restriction sites, introducing a new BamHI site between XhoI and *hisF-C* and yielding pCFN-*hisFC*. In order to avoid expression of HisF-C - CAT fusion proteins through religated vectors in the upcoming selection, a frame shift between the coding region of FLAG and the BamHI site was introduced.

The *hisF-C** gene, which is identical to *hisF-C* except for the substitutions A124R and A220K,²⁸ was subjected to error-prone PCR using the GeneMorph® II Random Mutagenesis Kit (Stratagene) under conditions which should introduce about two nucleotide exchanges per gene. To this end, the oligonucleotides 5'-AAGAAGGA-GATATACTCGAGCCAGCGC-3' and 5'-TTGTCGACG-GAGCTCGAATTCGGATCC-3' were used as 5' and 3' primers, respectively. These contained XhoI and BamHI restriction sites (underlined) and reconstituted the correct reading frame for the translation of the *hisF-C**C-cat fusion construct after cloning of the randomised *hisF-C** fragments into pCFN-*hisFC*.

The plasmids obtained were used to transform *E. coli* XL1-BlueMrf' cells (Stratagene) by electroporation, which were plated on medium containing ampicillin to select for plasmid uptake. As estimated from the number of grown colonies and the ligation efficiency tested by colony PCR, the *hisF-C**C-cat gene repertoire contained approximately

7.4×10^4 different variants. The grown colonies were scratched off the plates and resuspended, followed by the isolation of the pCFN-*hisFC* C^{*}C plasmid library. Sequencing of several clones grown under such non-selective conditions showed that Mutazyme® II DNA Polymerase had inserted two mutations per gene on average, as intended, with a range from zero to seven exchanges. The distribution of mutations showed a bias towards GC over AT base-pair exchanges (70% versus 30%), and of transitions over transversions (75% versus 25%).

In vivo selection for soluble HisF-C^{*}C variants

The amber-suppressor *E. coli* strain JM101⁶⁵ was transformed by electroporation with the pCFN-*hisFC* C^{*}C gene library, as well as with the controls pCFN-*hisFC* containing "wild-type" *hisF-C** and pCFN-*hisFC* containing no insert. In each case, expression was induced immediately after transformation by adding 200 $\mu\text{g}/\text{ml}$ of IPTG. Following shaking of the transformants for 2 h at 37 °C, the cells were streaked on plates containing 200 $\mu\text{g}/\text{ml}$ IPTG and different concentrations of chloramphenicol (0–400 $\mu\text{g}/\text{ml}$), and incubated at 37 °C for 16 h. Cells transformed with "wild-type" pCFN-*hisFC* C^{*}C were able to grow on plates with up to 250 $\mu\text{g}/\text{ml}$ chloramphenicol, whereas cells transformed with empty pCFN-*hisFC* vector grew only in the absence of the antibiotic, due to the introduced frame shift between the coding region of FLAG and the BamHI site mentioned above. From the pCFN-*hisFC* C^{*}C gene library several colonies grown in the absence of chloramphenicol, and the >330 colonies grown on plates containing 300, 350 and 400 $\mu\text{g}/\text{ml}$ chloramphenicol were picked and subjected to plasmid isolation. Colony PCR showed that 97% of these colonies were false-positives, which carried plasmids with truncated inserts that had lost the major part of the *hisF-C**C sequence and thus contained essentially only the *cat* gene.

Cloning of selected and designed *hisF-C**C variants into pET24a(+)

In order to produce the selected HisF-C^{*}C variants in the absence of CAT as well as the N-terminal His₆- and FLAG-tags, the *hisF-C** inserts from the 11 true positive colonies grown in the presence of high chloramphenicol concentrations were sub-cloned from pCFN-*hisFC* C^{*}C into pET24a(+) vectors containing the wild-type *hisF-C* gene (pET24a(+)-*hisFC*₂).²⁸ To this end, the *hisF-C** inserts were cut out from pCFN-*hisFC* C^{*}C with the help of the restriction enzymes XhoI and BamHI. The resulting fragment was amplified by PCR using the oligonucleotides 5'-ATACATATGCAGCGCGTTGTCGTGGCGATA-3' with a NdeI site (underlined) as 5'-primer and 5'-TCGACGGAGTCTAGATTCCGATCCCAACCC-3' with a BamHI site (underlined) as 3'-primer. The amplification product was ligated with pET24-*hisFC*₂, yielding the various pET24a(+)-*hisFC* C^{*}C variants. In these constructs, the genes of *hisF-C** and *hisF-C* are connected by the BamHI site in such a way that in the HisF-C^{*}C proteins both halves are linked by a Gly-Ser-Gly stretch.

For construction of the variant HisF-C^{*}C Δ in which the $\alpha\beta$ loop connecting the N- and C-terminal halves was shortened from 13 to 5 residues, *hisF-C** was amplified by PCR using an isolated *hisF-C** fragment as template. The oligonucleotide 5'-ATACATATGCAGCGCGTTGTCGTGGCGATA-3' with a NdeI site (underlined) was used as 5'-primer and the oligonucleotide 5'-

ATAGACGGATCCGCCGTGTTTTTGGAGGTACTCTTT-CAGTTC-3' with a BamHI site (underlined) was used as 3'-primer. The amplified fragment *hisF-C** Δ was cloned into pET24a(+)-*hisFC*₂, yielding pET24a(+)-*hisFC*^{*} Δ . As a consequence, the halves connecting $\alpha\beta$ -loop in HisF-C^{*} Δ ends with Gly245 (Figure 1(c)), and the Gly-Ser-Gly stretch connecting the two halves was preserved.

To introduce the single amino acid exchanges D135G, T142S, Y143H and V234M into the N-terminal half of HisF-C^{*} Δ the *hisF-C** sequence was amplified by PCR using an isolated *hisF-C** fragment as template. The oligonucleotides 5'-GATATACATATGCAGCG-CGTTGTTCGTGGCGATAGATGCAAAAAGAGTG-**G**GTGGAGAGTTC-3' (for generation of *hisF-C**-D135G), 5'-GATATACATATGCAGCGCGTTGTTCGTGGCGATA-GATGCAAAAAGAGTGGATGGAGAGTTCATGGTCTT-CAGCTACTCCGAAAAGAAG-3' (for generation of *hisF-C**-T142S), 5'-GATATACATATGCAGCG-CGTTGTTCGTGGCGATAGATGCAAAAAGAGTG-GATGGAGAGTTCATGGTCTTCCACTCCG-GAAAGAAG-3' (for generation of *hisF-C**-Y143H) and 5'-ATACATATGCAGCGCGTTGTTCGTGGCGATA-3' (for generation of *hisF-C**-V234M) were used as 5'-primers. The oligonucleotides 5'-ATAGACGGATCCGCCGTGTTTTTGGAGGTACTCTTT-CAGTTC-3' (for generation of *hisF-C**-D135G, *hisF-C**-T142S and *hisF-C**-Y143H) and 5'-ATAGACGGATCCGCCGTGTTTTTGGAGGTACTCTTT-CAGTTCATGTC-3' (for generation of *hisF-C**-V234M) were used as 3'-primer. As a consequence, at the 5' terminus a NdeI-site (underlined) and at the 3' terminus a BamHI site (underlined) were introduced together with the specific codon exchange (bold). The 3'-primers are designed in such a way that the amplified *hisF-C** fragment is shortened like in *hisF-C** Δ . Thus, the N-terminal half of each single variant ends at Gly245, and the two halves are linked by a Gly-Ser-Gly stretch. The cloning of the amplification products into pET24-*hisFC*₂ yielded pET24a(+)-*hisF C** Δ +D135G, pET24a(+)-*hisF C** Δ +T142S, pET24a(+)-*hisF C** Δ +Y143H and pET24a(+)-*hisF C** Δ +V234M.

For construction of *hisF-C** Δ in pET24a(+), *hisF-C** was amplified by PCR using an isolated *hisF-C** fragment as template and the oligonucleotides mentioned above which introduce the Y143H and the V234M exchanges as well as the loop shortening. The cloning of the amplification fragment into pET24a(+)-*hisFC*₂ yielded pET24a(+)-*hisF C** Δ . All constructs were sequenced entirely to exclude inadvertent mutations.

Heterologous expression of *hisF-C** Δ variants and protein purification

Heterologous expression of the *hisF-C** Δ variants was performed in *E. coli* BL21(DE3) (Stratagene) transformed with the various pET24a(+) plasmids containing the encoding sequences. To this end, four liters of Luria Broth (LB) media supplemented with 75 μ g/ml of kanamycin (for maintenance of pET24a(+)) were inoculated with a pre-culture and incubated at 37 °C. After an OD₆₀₀ of 0.6 was reached, expression was induced by adding 1 mM IPTG, and growth continued for another 4 h. Cells were harvested by centrifugation (Sorvall/RC5B, GS3 rotor, 15 min, 4000 rpm at 4 °C), washed with 10 mM potassium phosphate (pH 7.5), and centrifuged again. The cells were resuspended in 140 ml of the same buffer, lysed by sonification (Branson Sonifier W-250D; 2 \times 2 min in 15 s intervals, 50% pulse, 0 °C), and centrifuged again (Sorvall/RC5B, SS34 rotor, 30 min, 13,000 rpm at 4 °C) to separate the soluble from the insoluble fraction of the cell extract.

The variants HisF-C^{*} Δ 2 and 4, as well as HisF-C^{*} Δ , Δ +D135G, Δ +T142S, Δ +Y143H, Δ +V234M were purified from the soluble cell fraction. To this end, the extract was loaded onto a nickel Sepharose column (HisTrap FF crude 5 ml; GE Healthcare) which had been equilibrated with 50 mM potassium phosphate, 300 mM NaCl, 1 mM imidazole (pH 7.5). The column was washed with the equilibration buffer, and the bound His₆-tagged protein was eluted by applying a linear gradient of 1 mM–300 mM imidazole. Fractions with pure protein were pooled and dialysed extensively against 50 mM potassium phosphate (pH 7.5). In the case of the HisF-C^{*} Δ 4, the dialysis buffer contained either 1 mM dithiothreitol (DTT) to avoid disulfide-linkage *via* Y143C or no DTT to allow for cystine formation. According to SDS-PAGE (12.5% (w/v) acrylamide), all proteins were pure >95%.

The variants HisF-C^{*} Δ 1 and 3 were purified from the insoluble cell fraction, which was washed with 10 mM potassium phosphate (pH 7.5), resuspended in 80 ml of the same buffer containing 6 M guanidinium chloride (GdmCl) and incubated for 30 min at room temperature. Following the addition of 80 ml of 1 M GdmCl and further incubation for 30 min the suspension was centrifuged (Sorvall/RC5B, SS34 rotor, 30 min, 13,000 rpm at 4 °C). The supernatant was mixed with 280 ml of 2 M GdmCl and then dialysed extensively against 50 mM potassium phosphate (pH 7.5), to induce protein refolding. The solution was centrifuged (Sorvall/RC5B, SS34 rotor, 30 min, 13,000 rpm at 4 °C) to separate protein aggregates. Most refolded HisF-C^{*} Δ variants were pure to >95%. Otherwise, minor impurities were removed by applying nickel Sepharose chromatography as described above.

Wild-type HisF, HisF-C and HisF-CC were produced in *E. coli* BL21(DE3) as mentioned above and purified as described.²⁸

Analysis of protein solubility

HisF and the HisF-C^{*} Δ variants were produced in 50 ml cultures of *E. coli* BL21(DE3) cells as described above. From each culture a 2 ml aliquot was centrifuged (Eppendorf 5415R, 1 min, 13,000 rpm at 4 °C). The cells were resuspended in 1 ml of 100 mM potassium phosphate, 2 mM EDTA (pH 7.5), and lysed by sonification (Branson Sonifier W-250D, 45 sec, 15% pulse, 0 °C). The soluble and insoluble fractions of 100 μ l of the cell extract were separated. The two preparations were supplemented with polyacrylamide gel loading buffer and applied at different dilutions on 12.5% SDS-polyacrylamide gels. The protein bands were subsequently stained with Coomassie blue and their intensities were quantified using OptiQuant (version 02.50). The protein bands shown in Figure 4(a) represent undiluted samples of soluble and insoluble cell fractions. To avoid quantification errors caused by a non-linear relationship between the staining signal and the amount of protein, protein bands of the soluble and insoluble fractions with similar staining intensity but different degrees of dilution were compared, followed by back calculation to the undiluted solutions. The data shown in Figure 4(b) are the result of three independent experiments, which allowed us to determine mean values and standard deviations.

Analytical methods

Analytical gel filtration was performed by using a calibrated Superdex 75 column (Amersham). In each case 0.15–0.2 mg of the protein was applied on the column and

eluted with a flow rate of 0.5 ml/min in 50 mM potassium phosphate, 300 mM potassium chloride (pH 7.5), at 25 °C. HisF-C*₄ was analysed either in the presence or absence of 1 mM DTT. Fluorescence emission spectra of HisF, HisF-C*₄ and HisF-C**₄ were measured at 25 °C with a Cary Eclipse spectrophotometer (Varian) using 5 μM protein dissolved in 50 mM potassium phosphate (pH 7.5), and an excitation wavelength of 280 nm. The generated emission is mainly caused by the single tryptophan 156 of HisF which exists as two copies in HisF-C*₄ and HisF-C**₄, due to its location in the C-terminal half of HisF. Near-UV CD spectra were recorded at 25 °C with an Aviv CD spectrometer 62-DS (*d*=1 cm) using 68 μM HisF, 42 μM HisF-C*₄ and 54 μM HisF-C**₄ dissolved in 50 mM potassium phosphate (pH 7.5). The measured intensities were normalized to molar ellipticities.

Protein unfolding induced by urea or GdmCl was followed in 50 mM potassium phosphate (pH 7.5), by the decrease of the fluorescence intensity at 320 nm after excitation at 280 nm. The signal was measured after different time intervals until no further change was observed to ensure that the equilibrium was reached. The obtained intensities were normalised and plotted as fractional change of the native signal. This allowed us to determine the midpoint of unfolding $D_{1/2}$ [M], which represents the concentration of urea where 50% of the protein had non-native structure and served as an operational measure for conformational stability. Moreover, the unfolding traces were analysed according to the two-state model⁶⁶ to determine an apparent *m*-value, which served as an operational measure for the cooperativity of the unfolding process. All proteins could be refolded after complete unfolding in 8 M urea or 6 M GdmCl by removing the denaturant *via* dilution or dialysis against 50 mM potassium phosphate (pH 7.5). To investigate the stabilizing effect of phosphate on HisF-C**₄ the unfolding of this protein was also measured in 50 mM Tris-HCl (pH 7.5). To completely remove phosphate ions, purified HisF-C**₄ was unfolded by 8 M urea and refolded by extensive dialysis against 50 mM Tris-HCl (pH 7.5).

The ammonia dependent HisF activity was measured as described.⁹ Δ ASA (\AA^2) values, which denote the difference of solvent exposed surface between the folded and unfolded states were calculated as described.³⁰

Molecular dynamics simulations

The structural and energetic properties of the various proteins were analysed by molecular dynamics calculations using the molecular modelling program MOLOC³¹ and the force field MAB³⁴ with implicit solvation. The constructs were built on the basis of the published X-ray structure of HisF (PDB code 1thf)²⁵ by a 180° rotation around the barrel axis of the C-terminal half barrel and its subsequent fitting onto the N-terminal half barrel using DeepView 3.7§. Subsequently, first the loops were relaxed and later the whole structure by minimization and 2 ns of molecular dynamics at 300 K. Starting from HisF-CC, models for HisF-C*₄, HisF-C*₄ Δ, and HisF-C**₄ were constructed and relaxed by MD. For lack of a reliable method to calculate the free energy of the unfolded state, "interaction energies" between the linked half barrels were calculated using a pseudo-thermodynamic circle.³²

The energetic properties were analysed after an equilibration phase of 500 ps.

Acknowledgements

We thank Adriane Lochner for experimental help, and Dr. Birte Höcker for discussion and comments on the manuscript. This work was supported by grants of the Deutsche Forschungsgemeinschaft to M.B. (SPP 1170:BO 3066/1-1) and R.S. (STE 891/4-3, 4-4).

References

- Lupas, A. N., Ponting, C. P. & Russell, R. B. (2001). On the evolution of protein folds: are similar motifs in different protein folds the result of convergence, insertion, or relics of an ancient peptide world? *J. Struct. Biol.* **134**, 191–203.
- Riechmann, L. & Winter, G. (2006). Early protein evolution: building domains from ligand-binding polypeptide segments. *J. Mol. Biol.* **363**, 460–468.
- Marcotte, E. M., Pellegrini, M., Thompson, M. J., Yeates, T. O. & Eisenberg, D. (1999). A combined algorithm for genome-wide prediction of protein function. *Nature*, **402**, 83–86.
- Marcotte, E. M., Pellegrini, M., Ng, H. L., Rice, D. W., Yeates, T. O. & Eisenberg, D. (1999). Detecting protein function and protein-protein interactions from genome sequences. *Science*, **285**, 751–753.
- Yadid, I. & Tawfik, D. S. (2007). Reconstruction of functional β -propeller lectins via homo-oligomeric assembly of shorter fragments. *J. Mol. Biol.* **365**, 10–17.
- Gerlt, J. A. & Raushel, F. M. (2003). Evolution of function in (β/α)₈-barrel enzymes. *Curr. Opin. Chem. Biol.* **7**, 252–264.
- Glasner, M. E., Gerlt, J. A. & Babbitt, P. C. (2007). Mechanisms of protein evolution and their application to protein engineering. *Adv. Enzymol. Relat. Areas Mol. Biol.* **75**, 193–239.
- Henn-Sax, M., Höcker, B., Wilmanns, M. & Sterner, R. (2001). Divergent evolution of ($\beta\alpha$)₈-barrel enzymes. *Biol. Chem.* **382**, 1315–1320.
- Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A. & Sterner, R. (2001). Dissection of a ($\beta\alpha$)₈-barrel enzyme into two folded halves. *Nature Struct. Biol.* **8**, 32–36.
- Nagano, N., Orengo, C. A. & Thornton, J. M. (2002). One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.* **321**, 741–765.
- Sterner, R. & Höcker, B. (2005). Catalytic versatility, stability, and evolution of the ($\beta\alpha$)₈-barrel enzyme fold. *Chem. Rev.* **105**, 4038–4055.
- Pujadas, G. & Palau, J. (1999). TIM barrel fold: structural, functional and evolutionary characteristics in natural and designed molecules. *Biologica (Bratislava)*, **54**, 231–253.
- Vega, M. C., Lorentzen, E., Linden, A. & Wilmanns, M. (2003). Evolutionary markers in the (β/α)₈-barrel fold. *Curr. Opin. Chem. Biol.* **7**, 694–701.
- Wierenga, R. K. (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Letters*, **492**, 193–198.

§ <http://www.expasy.org/spdbv/>

15. Nicolet, Y. & Drennan, C. L. (2004). AdoMet radical proteins - from structure to evolution - alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucl. Acids Res.* **32**, 4015–4025.
16. Zitzewitz, J. A., Gualfetti, P. J., Perkons, I. A., Wasta, S. A. & Matthews, C. R. (1999). Identifying the structural boundaries of independent folding domains in the α subunit of tryptophan synthase, a β/α barrel protein. *Protein Sci.* **8**, 1200–1209.
17. Zitzewitz, J. A. & Matthews, C. R. (1999). Molecular dissection of the folding mechanism of the α subunit of tryptophan synthase: an amino-terminal autonomous folding unit controls several rate-limiting steps in the folding of a single domain protein. *Biochemistry*, **38**, 10205–10214.
18. Eder, J. & Kirschner, K. (1992). Stable substructures of eightfold $\beta\alpha$ -barrel proteins: fragment complementation of phosphoribosylanthranilate isomerase. *Biochemistry*, **31**, 3617–3625.
19. Soberon, X., Fuentes-Gallego, P. & Saab-Rincon, G. (2004). *In vivo* fragment complementation of a (β/α)₈ barrel protein: generation of variability by recombination. *FEBS Letters*, **560**, 167–172.
20. Bertolaet, B. L. & Knowles, J. R. (1995). Complementation of fragments of triosephosphate isomerase defined by exon boundaries. *Biochemistry*, **34**, 5736–5743.
21. Pan, H., Raza, A. S. & Smith, D. L. (2004). Equilibrium and kinetic folding of rabbit muscle triosephosphate isomerase by hydrogen exchange mass spectrometry. *J. Mol. Biol.* **336**, 1251–1263.
22. Silverman, J. A. & Harbury, P. B. (2002). The equilibrium unfolding pathway of a (β/α)₈ barrel. *J. Mol. Biol.* **324**, 1031–1040.
23. Gu, Z., Zitzewitz, J. A. & Matthews, C. R. (2007). Mapping the structure of folding cores in TIM barrel proteins by hydrogen exchange mass spectrometry: the roles of motif and sequence for the indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*. *J. Mol. Biol.* **368**, 582–594.
24. Thoma, R., Schwander, M., Liebl, W., Kirschner, K. & Sterner, R. (1998). A histidine gene cluster of the hyperthermophile *Thermotoga maritima*: sequence analysis and evolutionary significance. *Extremophiles*, **2**, 379–389.
25. Lang, D., Thoma, R., Henn-Sax, M., Sterner, R. & Wilmanns, M. (2000). Structural evidence for evolution of the β/α barrel scaffold by gene duplication and fusion. *Science*, **289**, 1546–1550.
26. Höcker, B., Schmidt, S. & Sterner, R. (2002). A common evolutionary origin of two elementary enzyme folds. *FEBS Letters*, **510**, 133–135.
27. Heinz, D. W., Essen, L. O. & Williams, R. L. (1998). Structural and mechanistic comparison of prokaryotic and eukaryotic phosphoinositide-specific phospholipases C. *J. Mol. Biol.* **275**, 635–650.
28. Höcker, B., Claren, J. & Sterner, R. (2004). Mimicking enzyme evolution by generating new ($\beta\alpha$)₈-barrels from ($\beta\alpha$)₄-half barrels. *Proc. Natl Acad. Sci. USA*, **101**, 16448–16453.
29. Maxwell, K. L., Mittermaier, A. K., Forman-Kay, J. D. & Davidson, A. R. (1999). A simple *in vivo* assay for increased protein solubility. *Protein Sci.* **8**, 1908–1911.
30. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
31. Gerber, P. R. & Müller, K. (1995). MAB, a generally applicable molecular force field for structure modeling in medicinal chemistry. *J. Comput. Aided Mol. Des.* **9**, 251–268.
32. van Gunsteren, W. F. & Berendsen, H. J. (1987). Thermodynamic cycle integration by computer simulation as a tool for obtaining free energy differences in molecular chemistry. *J. Comput. Aided Mol. Des.* **1**, 171–176.
33. Halgren, T. A. (1996). Merck Molecular Force Field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comp. Chem.* **17**, 520–552.
34. Gerber, P. R. (1998). Charge distribution from a simple molecular orbital type calculation and non-bonding interaction terms in the force field MAB. *J. Comput. Aided Mol. Des.* **12**, 37–51.
35. Vogel, C., Bashton, M., Kerrison, N. D., Chothia, C. & Teichmann, S. A. (2004). Structure, function and evolution of multidomain proteins. *Curr. Opin. Struct. Biol.* **14**, 208–216.
36. Orengo, C. A. & Thornton, J. M. (2005). Protein families and their evolution - a structural perspective. *Annu. Rev. Biochem.* **74**, 867–900.
37. Vogel, C., Teichmann, S. A. & Pereira-Leal, J. (2005). The relationship between domain duplication and recombination. *J. Mol. Biol.* **346**, 355–365.
38. Arnold, F. H. & Georgiou, G. (2003). Directed enzyme evolution - screening and selection methods. In *Methods in Molecular Biology* (Walker, J. M., ed), p. 230, Humana Press, Totowa, NJ.
39. Eijsink, V. G., Gaseidnes, S., Borchert, T. V. & van den Burg, B. (2005). Directed evolution of enzyme stability. *Biomol. Eng.* **22**, 21–30.
40. Hart, D. J. & Tarendeau, F. (2006). Combinatorial library approaches for improving soluble protein expression in *Escherichia coli*. *Acta Crystallog. sect. D*, **62**, 19–26.
41. Roodveldt, C., Aharoni, A. & Tawfik, D. S. (2005). Directed evolution of proteins for heterologous expression and stability. *Curr. Opin. Struct. Biol.* **15**, 50–56.
42. Waldo, G. S. (2003). Genetic screens and directed evolution for protein solubility. *Curr. Opin. Chem. Biol.* **7**, 33–38.
43. Sieber, V., Martinez, C. A. & Arnold, F. H. (2001). Libraries of hybrid proteins from distantly related sequences. *Nature Biotechnol.* **19**, 456–460.
44. Sieber, V. (2003). Selection for soluble proteins via fusion with chloramphenicol acetyltransferase. *Methods Mol. Biol.* **230**, 45–55.
45. Yano, J. K. & Poulos, T. L. (2003). New understandings of thermostable and peizostable enzymes. *Curr. Opin. Biotechnol.* **14**, 360–365.
46. Miyazono, K., Sawano, Y. & Tanokura, M. (2005). Crystal structure and structural stability of acylphosphatase from hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Proteins: Struct. Funct. Genet.* **61**, 196–205.
47. Toth, E. A., Worby, C., Dixon, J. E., Goedken, E. R., Marqusee, S. & Yeates, T. O. (2000). The crystal structure of adenylosuccinate lyase from *Pyrobaculum aerophilum* reveals an intracellular protein with three disulfide bonds. *J. Mol. Biol.* **301**, 433–450.
48. Chiu, H. J., Johnson, E., Schröder, I. & Rees, D. C. (2001). Crystal structures of a novel ferric reductase from the hyperthermophilic archaeon *Archaeoglobus fulgidus* and its complex with NADP⁺. *Structure*, **9**, 311–319.
49. Karlström, M., Stokke, R., Steen, I. H., Birkeland, N. K. & Ladenstein, R. (2005). Isocitrate dehydrogenase from the hyperthermophile *Aeropyrum pernix*:

- X-ray structure analysis of a ternary enzyme-substrate complex and thermal stability. *J. Mol. Biol.* **345**, 559–577.
50. Arai, R., Kukimoto-Niino, M., Kuroishi, C., Bessho, Y., Shirouzu, M. & Yokoyama, S. (2006). Crystal structure of the probable haloacid dehalogenase PH0459 from *Pyrococcus horikoshii* OT3. *Protein Sci.* **15**, 373–377.
 51. Mallick, P., Boutz, D. R., Eisenberg, D. & Yeates, T. O. (2002). Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. *Proc. Natl Acad. Sci. USA*, **99**, 9679–9684.
 52. Beeby, M., O'Connor, B. D., Ryttersgaard, C., Boutz, D. R., Perry, L. J. & Yeates, T. O. (2005). The genomics of disulfide bonding and protein stabilization in thermophiles. *PLoS Biol.* **3**, e309.
 53. Dombkowski, A. A. (2003). Disulfide by design: a computational method for the rational design of disulfide bonds in proteins. *Bioinformatics*, **19**, 1852–1853.
 54. Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl Acad. Sci. USA*, **84**, 6663–6667.
 55. Matsumura, M., Signor, G. & Matthews, B. W. (1989). Substantial increase of protein stability by multiple disulphide bonds. *Nature*, **342**, 291–293.
 56. Betz, S. F. (1993). Disulfide bonds and the stability of globular proteins. *Protein Sci.* **2**, 1551–1558.
 57. Mansfeld, J., Vriend, G., Dijkstra, B. W., Veltman, O. R., Van den Burg, B., Venema, G. *et al.* (1997). Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J. Biol. Chem.* **272**, 11152–11156.
 58. Ivens, A., Mayans, O., Szadkowski, H., Jürgens, C., Wilmanns, M. & Kirschner, K. (2002). Stabilization of a ($\beta\alpha$)₈-barrel protein by an engineered disulfide bridge. *Eur. J. Biochem.* **269**, 1145–1153.
 59. Sterner, R. & Liebl, W. (2001). Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* **36**, 39–106.
 60. Vieille, C. & Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
 61. Walden, H., Bell, G. S., Russell, R. J., Siebers, B., Hensel, R. & Taylor, G. L. (2001). Tiny TIM: a small, tetrameric, hyperthermostable triosephosphate isomerase. *J. Mol. Biol.* **306**, 745–757.
 62. Offredi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, A. S., Van de Weerd, C. *et al.* (2003). *De novo* backbone and sequence design of an idealized α/β -barrel protein: evidence of stable tertiary structure. *J. Mol. Biol.* **325**, 163–174.
 63. Beismann-Driemeyer, S. & Sterner, R. (2001). Imidazole glycerol phosphate synthase from *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex. *J. Biol. Chem.* **276**, 20387–20396.
 64. Höcker, B. (2002). *Experimentelle Rekonstruktion der Evolution von Proteinen am Beispiel zweier ($\beta\alpha$)₈-Barrel Enzyme aus der Histidin-Biosynthese*. PhD thesis, University of Cologne, Germany.
 65. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.
 66. Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.

Edited by F. Schmid

(Received 7 May 2007; received in revised form 5 June 2007; accepted 13 June 2007)
Available online 19 June 2007

Rapid Reports

High-Resolution Crystal Structure of an Artificial $(\beta\alpha)_8$ -Barrel Protein Designed from Identical Half-Barrels^{†,‡}Birte Höcker,^{*,§,||} Adriane Lochner,^{||,⊥} Tobias Seitz,[⊥] Jörg Claren,[⊥] and Reinhard Sterner^{*,⊥}

Max Planck Institute for Developmental Biology, Spemannstrasse 35, D-72076 Tübingen, Germany, and Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany

Received November 18, 2008; Revised Manuscript Received December 17, 2008

ABSTRACT: Ample evidence suggests that the ubiquitous $(\beta\alpha)_8$ -barrel enzyme fold has evolved by the duplication and fusion of an ancestral $(\beta\alpha)_4$ -half-barrel. To reconstruct this process in the laboratory with a model protein, we earlier fused two copies of the C-terminal half-barrel HisF-C of imidazole glycerol phosphate synthase (HisF) and stepwise stabilized the resulting HisF-CC construct. We now further increased its stability and solubility by introducing two additional amino acid exchanges, which allowed us to crystallize the resulting artificial $(\beta\alpha)_8$ -barrel protein HisF-C***C. The analysis of its X-ray structure at 2.1 Å resolution reveals a striking similarity to wild-type HisF, helps us to understand its improved stability, and provides further insights into the evolution of $(\beta\alpha)_8$ -barrel proteins.

The molecular evolution of new protein structures and functions is largely based on the duplication and fusion of existing genes and gene fragments (1, 2). This becomes particularly obvious in the case of protein folds with internal sequence and structural symmetries (3). A prominent example is the $(\beta\alpha)_8$ -barrel, which is the most frequent and most versatile enzyme fold (4, 5). The canonical $(\beta\alpha)_8$ -barrel consists of a central eight-stranded β -sheet, the barrel, which is surrounded by the eight α -helices. The active site is generally located at the C-terminal end of the barrel and in the $\beta\alpha$ -loops ("catalytic face"), whereas the N-terminal end of the barrel and the $\beta\alpha$ -loops are important for the conformational stability of the fold ("stability face") (6). The striking 2-fold sequence and structural symmetry of the $(\beta\alpha)_8$ -barrel enzymes *N*-[(5'-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (ProFAR) isomerase (HisA) and imidazole glycerol phosphate synthase

(HisF) suggests that these enzymes have evolved from a common $(\beta\alpha)_4$ -half-barrel ancestor by a series of gene duplication and fusion events (7). We have reconstructed this process in the laboratory by fusing two identical copies of HisF-C, which is the C-terminal half-barrel of HisF, and by stabilizing the resulting construct HisF-CC stepwise to HisF-C*C and HisF-C**C, using rational design and library selection (8, 9). We have now further stabilized HisF-C**C by the introduction of two additional residue exchanges, thus generating HisF-C***C.

The N-terminal half of HisF-C*C contains the amino acid exchanges A124_NR and A220_NK, which were introduced into HisF-CC to reconstitute a salt bridge cluster at the stability face of the central β -barrel of the wild-type HisF enzyme (8). The N-terminal half of HisF-C**C additionally contains a shortened loop connecting the two half-barrels and the two stabilizing exchanges Y143_NH and V234_NM, which were identified by random mutagenesis and selection *in vivo* (9). The high internal 2-fold symmetry suggested to us that the introduction of the corresponding exchanges Y143_CH and V234_CM into the C-terminal half of HisF-C**C might further improve the properties of the resulting HisF-C***C protein (Figure 1a).

Heterologous expression in *Escherichia coli* showed that the fraction of recombinant HisF-C***C that was produced in soluble form (84%) was much larger than the soluble fraction of HisF-C**C (51%) and almost as large as the soluble fraction of the wild-type HisF enzyme (90%) (Figure 1b). The purified HisF-C***C protein unfolds in guanidinium chloride with cooperativity comparable to that of but at denaturant concentrations slightly higher than that of HisF-C**C (Figure 1c). Moreover, HisF-C***C unfolds with cooperativity similar to that of albeit at denaturant concentrations significantly lower than that of wild-type HisF, which must be extraordinarily stable to avoid denaturation in its hyperthermophilic host, *Thermotoga maritima*.

The HisF-C***C protein was crystallized, and its X-ray structure was determined at 2.1 Å resolution by molecular replacement with the coordinates of the C-terminal half of wild-type HisF [residues 123–253, Protein Data Bank (PDB) entry 1thf] (Table S1) (7). HisF-C***C forms a $(\beta\alpha)_8$ -barrel structure with essentially the same topology and secondary structure content as HisF (Figure 2 and Figure S1). The areas around helices α_8 _N and α_8 _C are not well defined and could

[†] This work was supported by a grant from the Deutsche Forschungsgemeinschaft (STE 891/4-3) and institutional funds from the Max-Planck Society.

[‡] The atomic coordinates of HisF-C***C have been deposited in the Protein Data Bank as entry 2w6r.

^{*} To whom correspondence should be addressed. B.H.: telephone, +49-7071-601-322; fax, +49-7071-601-305; e-mail, birte.hoecker@tuebingen.mpg.de. R.S.: telephone, +49-941-943-3015; fax, +49-941-943-2813; e-mail, Reinhard.Sterner@biologie.uni-regensburg.de.

[§] Max Planck Institute for Developmental Biology.

^{||} These authors contributed equally to the experiments.

[⊥] University of Regensburg.

1146 *Biochemistry*, Vol. 48, No. 6, 2009

Rapid Reports

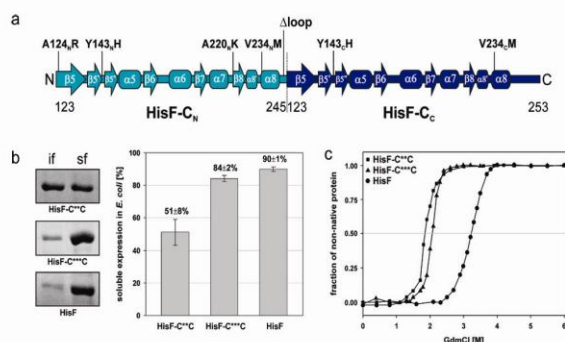


FIGURE 1: Design of the $(\beta\alpha)_8$ -barrel protein HisF-C***C from two identical and fused HisF-C half-barrels. (a) Secondary structure elements and stabilizing amino acid exchanges. The exchanges in HisF-C_N and the shortening of the loop connecting the two half-barrels were introduced in earlier work to stabilize HisF-CC via HisF-C**C to HisF-C***C (8, 9). The exchanges Y143_H and V234_M were introduced in this work to further stabilize HisF-C***C to HisF-C***C. (b) Solubility analysis of HisF-C***C, HisF-C**C, and HisF. In the left panel, insoluble (if) and soluble (sf) fractions of *E. coli* host cell extracts analyzed by SDS-PAGE (12.5% acrylamide) and staining with Coomassie blue. In the right panel, the percentage of the proteins produced in soluble form (as deduced from the analysis of three different clones of each variant). (c) Stability of HisF-C**C, HisF-C***C, and HisF in 50 mM potassium phosphate (pH 7.5) at 25 °C. Chemical unfolding was induced by incubation in guanidinium chloride (GdmCl) and monitored by the decrease in fluorescence intensity at 320 nm following excitation at 280 nm. The signals were normalized, and the following $D_{1/2}$ values (in molar GdmCl) were determined: HisF-C**C, 1.90 ± 0.03 M; HisF-C***C, 2.10 ± 0.02 M; HisF, 3.32 ± 0.07 M.

not be built satisfactorily into the density (Figures S1 and S2). All other parts of the structure though have a high resolution and can be compared to the parent HisF structure and an earlier calculated structural model of HisF-C**C (9).

The structure of HisF-C***C superimposes nicely with that of HisF. The overall root-mean-square deviation (rmsd) of the two proteins is 1.7 Å, and the rmsd values for the superimposed N-terminal and C-terminal halves are 1.6 and 0.6 Å, respectively. HisF-C***C shows an extremely high 2-fold symmetry, with a rmsd of only 0.6 Å for its two superimposed HisF-C_N and HisF-C_C halves, compared to a rmsd of 1.6 Å for the superimposed HisF-N and HisF-C halves of wild-type HisF (Figure 2). The two aspartate residues of HisF-C***C, D130_N and D130_C, which are

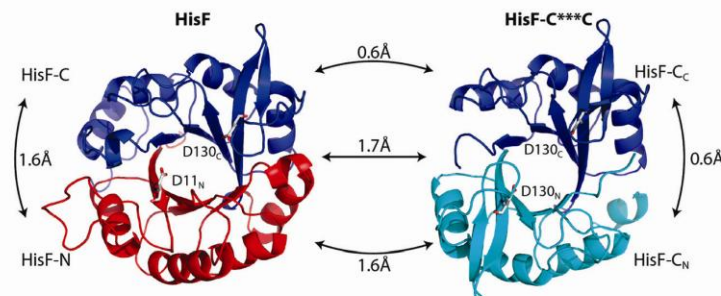


FIGURE 2: Ribbon diagrams of the crystal structures of HisF (PDB entry 1thf) and HisF-C***C. Top view onto the catalytic face (cp. Figure 3a). The functional aspartate residues D11_N and D130_C of HisF and the corresponding D130_N and D130_C residues of HisF-C***C are shown as sticks. The indicated root-mean-square deviations were calculated for the following pairs: HisF–HisF-C***C, 1.7 Å (184 superimposed C α atoms); N-terminal halves of the two proteins, 1.6 Å (89 C α atoms); C-terminal halves of the two proteins, 0.6 Å (103 C α atoms); the halves within HisF, 1.6 Å (89 C α atoms); the halves within HisF-C***C, 0.6 Å (103 C α atoms).

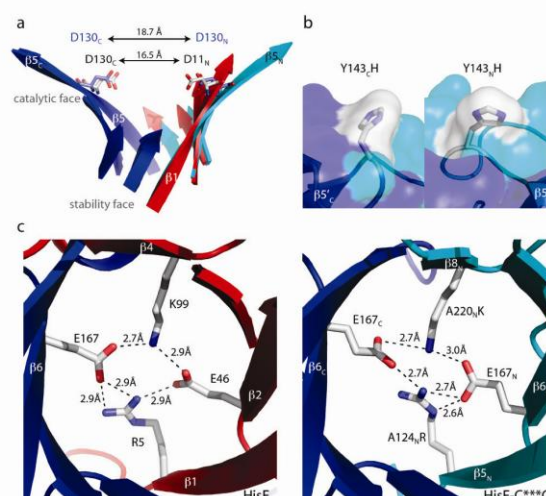


FIGURE 3: Important structural features of HisF-C***C in comparison to those of HisF. (a) The β -barrel of HisF-C***C opens up wider at the catalytic face than HisF, resulting in a distance increase in the functional aspartate residues of ~ 2 Å. (b) The exchanges Y143_NH and Y143_CH, which contribute to the solubility and stability of HisF-C***C, are solvent-exposed in the crystal structure. (c) The stabilizing salt bridge cluster of HisF is formed in a similar way in HisF-C***C through the exchanges A124_NR and A220_NK. Hydrogen bonds between the involved side chain residues, which are shown as sticks, are indicated by dashed lines up to a distance cutoff of 3.0 Å. Colors as in Figure 2.

located at the C-terminal ends of strands $\beta 5_N$ and $\beta 5_C$, respectively, correspond to catalytic aspartate residues D11 and D130 of HisF, which are located at the C-terminal ends of strands $\beta 1$ and $\beta 5$, respectively (Figure 2 and Figure S1).

We therefore tested whether HisF-C***C has HisF activity by incubating high concentrations of the protein (50 μ M) with the substrates ammonia and *N*'-[5'-(phosphoribulose)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (PRFAR). However, no production of ImGP could be detected in the spectroscopic HisF assay (10). This lack of measurable activity could be due to the increased distance of catalytic aspartates D130_N and D130_C in comparison to D11 and D130 (Figure 3a). Moreover, although the addition of orthophosphate stabilizes HisF-C***C, we did not find any evidence for the binding of the bisphosphorylated substrate analogue ProFAR (10). As the N-terminal phosphate binding site of HisF-C***C is slightly shifted with

Rapid Reports

respect to that of HisF (not shown), the proper anchoring of the corresponding phosphate group of ProFAR and PRFAR in the designed active site might be hampered, resulting in a low binding affinity.

The X-ray structure further allowed us to analyze the stabilizing effects of the amino acid exchanges that were introduced to generate HisF-C***C from HisF-CC. The histidine residues introduced by library selection [Y143_NH (9)] and site-directed mutagenesis [Y143_CH (this work)] point into the solvent (Figure 3b), which allows them to engage in favorable interactions with water, explaining why these substitutions increase both solubility and stability. Crystal contacts (not shown) appear to stabilize this conformation compared to an alternative one suggested by molecular dynamics simulation of HisF-C**C, in which the imidazole group of H143_N forms a hydrogen bond network with the carboxylate group of D176_C and the phosphate ion bound to the C-terminal half-barrel (9). The methionine residues introduced at position 234 are not resolved in the crystal structure, neither in the N-terminal nor in the C-terminal half. Their contribution to increased stability or solubility thus cannot be evaluated on the basis of this structure. A structural model of HisF-C**C, however, suggests that the long methionine side chain protrudes into a hydrophobic pocket of the adjacent half-barrel to form favorable van der Waals interactions (9).

Inspection of the structure shows that exchanges A124_NR and A220_NK at the N-terminal face of the central barrel result in the formation of a salt bridge cluster involving E167_N and E167_C, as expected (Figure 3c). This rationally designed salt bridge cluster strengthens the interaction between the N- and C-terminal half-barrel, explaining why the introduced exchanges are beneficial both for the compactness and for the stability of the protein (8).

Following the fusion of two identical HisF-C half-barrels, the shortening of the connecting loop and as little as six amino acid exchanges were sufficient to generate the highly stable protein HisF-C***C. The artificial ($\beta\alpha$)₈-barrel shows a high degree of structural similarity to the wild-type HisF enzyme. This finding is remarkable, given that the N-terminal half-barrels HisF-C_N and HisF-N share a degree of sequence identity of only ~25%, and demonstrates that ($\beta\alpha$)₈-barrels can be readily evolved in the laboratory from a half-barrel. Our results make it likely that similar gene duplication and fusion events have led to the generation of an ancestral ($\beta\alpha$)₈-barrel during natural evolution (6). Although the sequences of HisF-C and HisF-C***C certainly differ from those of the ancestral proteins due to genetic drift, they are good models for these proteins and the putative evolutionary events connecting them. It will now be interesting to test whether and which catalytic activities can be established on the HisF-C***C scaffold and how many mutations will be necessary to do so. The structure presented here shows how the barrel of HisF-C***C opens up wider than the barrel of HisF, resulting in a larger binding pocket (Figure 3a). This observation and the absence of detectable HisF activity

Biochemistry, Vol. 48, No. 6, 2009 1147

illustrate that a fine-tuning of the active site will be required, which might be hard to achieve by rational approaches and probably will require directed evolution techniques, i.e., random mutagenesis followed by selection or screening.

The regular order of the ($\beta\alpha$)-units has prompted experiments aimed at designing artificial ($\beta\alpha$)₈-barrel enzymes early on. However, the first attempts yielded molten globules with defined secondary structures at best (11–13). More recently, a ($\beta\alpha$)₈-barrel that consists of an idealized artificial backbone containing an eight-stranded β -sheet with 4-fold symmetry surrounded by eight α -helices was designed. An optimal amino acid sequence fitting the idealized backbone was then chosen with an automated selection algorithm (14). Although the designed protein is stable and appears to adopt a well-defined tertiary structure in solution, no high-resolution X-ray or NMR structure has been published until now. HisF-C***C is therefore the first artificial ($\beta\alpha$)₈-barrel for which a high-resolution structure is available, complementing structural analysis of an artificial nine-stranded barrel that was constructed by fusing HisF-C to parts of a flavodoxin-like ($\beta\alpha$)₅-protein (15).

SUPPORTING INFORMATION AVAILABLE

A description of the experimental procedures with supporting references, data collection and refinement statistics from the X-ray structure determination of HisF-C***C (Table S1), structure-based sequence alignment of HisF and HisF-C***C (Figure S1), and electron density maps around helices α _{8N} and α _{8C} of HisF-C***C (Figure S2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

1. Bashton, M., and Chothia, C. (2007) *Structure* 15, 85–99.
2. Chothia, C., Gough, J., Vogel, C., and Teichmann, S. A. (2003) *Science* 300, 1701–1703.
3. Yacid, I., and Tawfik, D. S. (2007) *J. Mol. Biol.* 365, 10–17.
4. Gerlt, J. A., and Raushel, F. M. (2003) *Curr. Opin. Chem. Biol.* 7, 252–264.
5. Wierenga, R. K. (2001) *FEBS Lett.* 492, 193–198.
6. Sterner, R., and Höcker, B. (2005) *Chem. Rev.* 105, 4038–4055.
7. Lang, D., Thoma, R., Henn-Sax, M., Sterner, R., and Wilmanns, M. (2000) *Science* 289, 1546–1550.
8. Höcker, B., Claren, J., and Sterner, R. (2004) *Proc. Natl. Acad. Sci. U.S.A.* 101, 16448–16453.
9. Seitz, T., Bocola, M., Claren, J., and Sterner, R. (2007) *J. Mol. Biol.* 372, 114–129.
10. Beismann-Driemeyer, S., and Sterner, R. (2001) *J. Biol. Chem.* 276, 20387–20396.
11. Beauregard, M., Goraj, K., Goffin, V., Heremans, K., Goormaghtigh, E., Ruysschaert, J. M., and Martial, J. A. (1991) *Protein Eng.* 4, 745–749.
12. Tanaka, T., Kuroda, Y., Kimura, H., Kidokoro, S., and Nakamura, H. (1994) *Protein Eng.* 7, 969–976.
13. Wodak, S. J., Lasters, I., Pio, F., and Claessens, M. (1990) *Biochem. Soc. Symp.* 57, 99–121.
14. Offredi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, A. S., Van de Weerd, C., Hoch, J. C., Prosperi, C., Francois, J. M., Mayo, S. L., and Martial, J. A. (2003) *J. Mol. Biol.* 325, 163–174.
15. Bharat, T. A., Eisenbeis, S., Zeth, K., and Höcker, B. (2008) *Proc. Natl. Acad. Sci. U.S.A.* 105, 9942–9947.

BI802125B

Supporting Information

High resolution crystal structure of an artificial ($\beta\alpha$)₈-barrel protein designed from identical half barrels

Birte Höcker, Adriane Lochner, Tobias Seitz, Jörg Claren, and Reinhard Sterner

Experimental procedures

Cloning of *hisF*-C***C into pET24a(+)

The *hisF*-C half barrel gene was amplified by PCR using the oligonucleotide 5'- GGT CGC GGA TCC GGT CAG GCC GTT GTC GTG GCG ATA GAT GCA AAA AGA GTG GAT GGA GAG TTC ATG GTC TTC ACC **CAC** TCC GGA AAG AAG-3' with a *Bam*HI site (underlined) as 5'-primer, and the oligonucleotide 5'-GTG GTG CTC GAG CAA CCC CTC CAG TCT CAC GTT CAC TCC GTG TTT TTT GAG GTA CTC TTT CAG TTC TCT **CAT** GTC GAT CTC TC-3' with a *Xho*I-site (underlined) as 3'-primer. Two new codons (bold letters) introduced the amino acid exchanges Y143H and V234M. Cloning of the amplification fragment into the pET24a(+) vector containing the 5'-terminal half of *hisF*-C***C (1) yielded pET24a(+)-*hisF*-C***C. The entire gene was sequenced to exclude inadvertent mutations.

Heterologous gene expression and purification of recombinant proteins

Heterologous expression of *hisF*-C***C was performed in *E. coli* BL21(DE3) (Stratagene) transformed with pET24a(+)-*hisF*-C***C, which results in the addition of a C-terminal hexahistidine tag to the recombinant protein. To this end, two liters of Luria Broth (LB) medium supplemented with 75 µg/ml kanamycin were inoculated with a pre-culture and incubated at

37°C. After an OD₆₀₀ of 0.6 was reached, expression was induced by adding 0.5 mM IPTG, and growth was continued for another 4 h. Cells were harvested by centrifugation (Sorvall/RC5B, GS3, 15 min, 4000 rpm, 4°C), washed with 10 mM potassium phosphate, pH 7.5, and centrifuged again. The cells were suspended in 70 ml of the same buffer, lysed by sonification (Branson Sonifier W-250D, 2 x 2 min in 15 sec intervals, 50% pulse, 0°C), and centrifuged again (Sorvall/RC5B, SS34, 30 min, 13.000 rpm, 4°C) to separate the soluble from the insoluble fraction of the cell extract. The soluble supernatant was loaded onto a nickel sepharose column (HisTrap FF crude 5 ml, GE Healthcare), which had been equilibrated with 50 mM potassium phosphate, 300 mM NaCl, 1 mM imidazole, pH 7.5. The column was washed with the equilibration buffer, and the bound His₆-tagged protein was eluted by applying a linear gradient of 1-700 mM imidazole. Fractions with pure protein were pooled and dialyzed extensively against 50 mM potassium phosphate, pH 7.5. According to SDS-PAGE (12.5 % acrylamide), HisF-C***C was more than 95 % pure.

Wild-type HisF and HisF-C**C were produced in *E. coli* BL21(DE3) and purified as described (1, 2).

Analysis of protein solubility

HisF, HisF-C**C, and HisF-C***C were produced in 50 ml cultures of *E. coli* BL21(DE3) cells as described above. From each culture a 2 ml aliquot was centrifuged (Eppendorf 5415R, 1 min, 13.000 rpm, 4°C). The cells were suspended in 1 ml of 100 mM potassium phosphate, 2 mM EDTA, pH 7.5, and lysed by sonification (Branson Sonifier W-250D, 45 sec, 15% pulse, 0°C). The soluble and insoluble fractions of 100 µl of the cell extract were separated. The two preparations were supplemented with polyacrylamide gel loading buffer and loaded at different dilutions on 12.5% SDS polyacrylamide gels. The protein bands were subsequently stained with Coomassie blue, and their intensities were quantified using OptiQuant (Version 02.50). The protein bands shown in Figure 1b (left panel) represent

undiluted samples of soluble and insoluble cell fractions. To avoid quantification errors caused by a non-linear relationship between the staining signal and the amount of protein, protein bands of the soluble and insoluble fractions with similar staining intensity but different degrees of dilution were compared, followed by back calculation to the undiluted solutions. The data shown in Figure 1b (right panel) are the result of three independent experiments, which allowed us to determine the shown mean values and standard deviations.

Analysis of protein stability

Protein unfolding induced by guanidinium chloride (GdmCl) was followed in 50 mM potassium phosphate, pH 7.5, at 25°C, by monitoring the decrease of the fluorescence intensity at 320 nm after excitation at 280 nm. The signal was measured after different time intervals until no further change was observed to ensure that the equilibrium was reached. The obtained intensities were normalized and plotted as fractional change of the native signal. The $D_{1/2}$ values (GdmCl concentration at which 50% of the native signal is lost) given in Figure 1c are the means and standard deviations of three independent experiments. Following complete unfolding in 6 M GdmCl, all proteins could be refolded by removing the denaturant via dilution or dialysis against 50 mM potassium phosphate, pH 7.5.

Mass spectrometry

Mass spectrometry was done by MALDI-TOF on a 4800 Plus MALDI-TOF/TOF-Analyzer (Applied Biosystems). The lyophilized samples were dissolved in 5 µl matrix solution (5 mg α -cyano-4hydroxy-cinnamic acid (CHCA) in 1ml H₂O) and spotted onto the MALDI target. For molecular weight measurements spectra were recorded in the linear mode, mass calibration was done using internal standards.

Crystallization, data collection, and refinement

Crystals were obtained by the sitting drop vapour diffusion method. Drops contained 0.4 μ l of a HisF-C***C protein solution mixed with 0.4 μ l of 0.1 M Tris-HCl pH 8.0, 12% w/v PEG 8000, and were equilibrated against 75 μ l of reservoir buffer. Following a short transfer into crystallization buffer with 25% PEG 400, the crystals were flash frozen in liquid nitrogen. Data of single crystals was collected at the synchrotron beamline PXII (Swiss Light Source, Villigen PSI, Switzerland) at 100K, recording single one-degree images on a MarCCD 225mm detector. The data was indexed, integrated, scaled with XDS, and converted with XDSCONV (3). The dataset had a resolution of 2.1 Å, cell dimensions were a=b=58 Å, c=131 Å, α = β =90°, and γ =120°, and the rhombohedral space group was P3₁21. The crystals contained one protein molecule in the asymmetric unit, yielding a V_M of 2.37 Å³/Da. Molecular replacement searches were performed with Molrep 9.4.09 (4) using the coordinates of the C-terminal half of HisF (PDB ID 1thf, residues 123-253). Model building was performed with the programs Coot (5) and PHASER (6). Refinement was performed with REFMAC 5.4.0066 (7), resulting in final R_{cryst} and R_{free} values of 24.2% and 28.5%. The data collection and refinement statistics are summarized in Table S1.

Structural superimposition

Structural superimpositions were performed using the program STAMP (8).

Supporting references

1. Seitz, T., Bocola, M., Claren, J., and Sterner, R. (2007) Stabilisation of a $(\beta\alpha)_8$ -barrel protein designed from identical half barrels, *J. Mol. Biol.* 372, 114-129.
2. Beismann-Driemeyer, S., and Sterner, R. (2001) Imidazole glycerol phosphate synthase from *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex, *J. Biol. Chem.* 276, 20387-20396.
3. Kabsch, W. (1988) Automatic indexing of rotation diffraction patterns, *J. Appl. Cryst.* 21, 67-72.
4. Vagin, A., and Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement, *Acta Crystallogr. Section D: Biol. Crystallogr.* 56, 1622-1624.
5. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. Section D: Biol. Crystallogr.* 60, 2126-2132.
6. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions, *Acta Crystallogr. Section D: Biol. Crystallogr.* 61, 458-464.
7. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. Section D: Biol. Crystallogr.* 53, 240-255.
8. Russell, R. B., and Barton, G. J. (1992) Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels, *Proteins* 14, 309-323.

Table S1: X-ray structure determination of HisF-C***C:

Data collection and refinement statistics.

Data collection

Wavelength [Å]	0.979
Space group	P3 ₁ 21
Resolution [Å]	40.0 – 2.1 (2.2 - 2.1)
Unique reflections	15628 (2379)
Redundancy	8.5 (7.6)
Completeness [%]	99.3 (96.6)
R _{merge} [%]	7.5 (88.2)
I/σ(I)	14.8 (2.4)
Wilson B-factor [Å ²]	51.3

Refinement statistics

Space group	P3 ₁ 21
Resolution [Å]	33.0 – 2.1 (2.2 - 2.1)
R _{cryst} [%]	24.2 (31.3)
R _{free} [%]	28.5 (38.4)
Mean B-value [Å ²]	47.2
r.m.s.d. of bond length [Å]	0.01
r.m.s.d. of angle [deg]	1.38
Number of water molecules	51

Model quality (Molprobit)

Residues in favoured region	204 (93.6%)
Residues in allowed region	217 (99.5%)
Residues in outlier region	1 (0.5%)

Figure S1: Structure-based sequence alignment of HisF and HisF-C***C. The colours of secondary structure elements are as in Figure 2. The aspartate residues which are shown as sticks in Figure 2 are in bold letters. Stabilizing residues that were introduced upon generating HisF-C***C from HisF-CC (cp. Figure 1a) are circled. Residues not resolved in the structure of HisF-C***C are in smaller letters. Mass spectrometry with HisF-C***C yielded a molecular weight of 29506 ± 6 Da, which is in excellent agreement with the calculated molecular weight for the protein (including His₆-tag) of 29559 Da. This result shows that the unresolved residues have not been degraded.

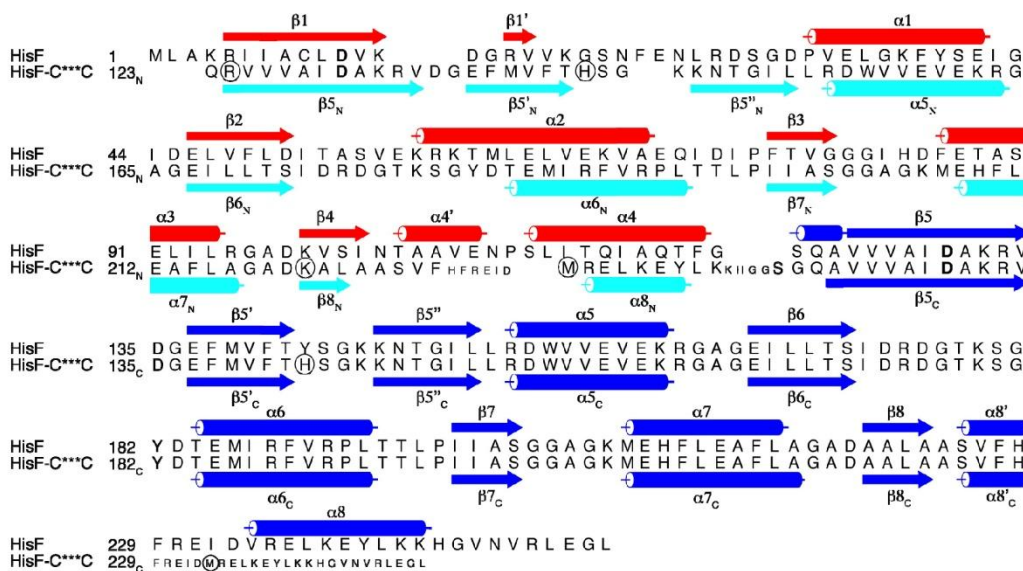
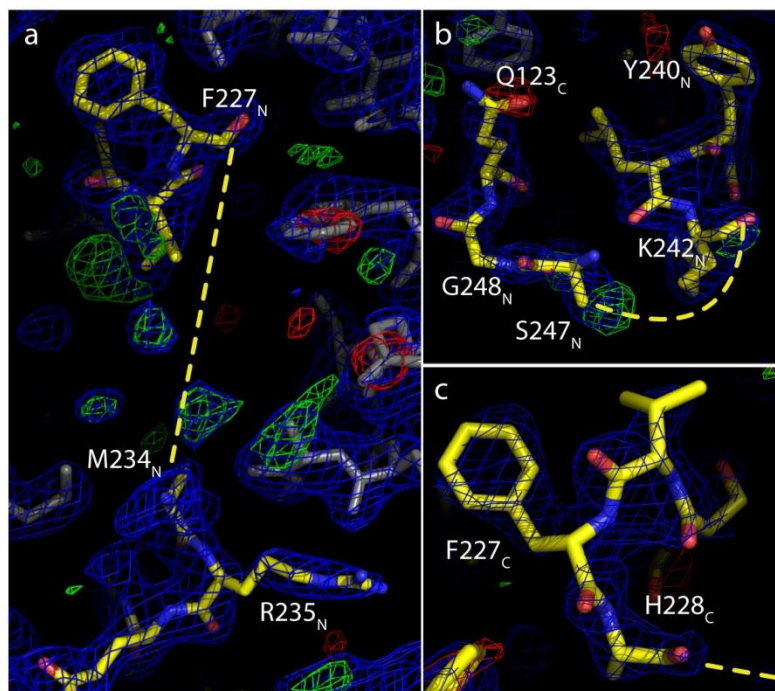


Figure S2: Electron density maps around helices α_{8N} and α_{8C} of HisF-C***C. (a) The sequence stretch between F227_N and M234_N (marked by the dashed line), which precedes helix α_{8N} , was not built. (b) The loop region between K242_N and S247_N, which follows helix α_{8N} , is not resolved. (c) The residues following H228_C, including those corresponding to the putative helix α_{8N} , could also not be placed.

Several residues in the regions of helices α_{8N} and α_{8C} are shown as yellow sticks, while residues from non-adjacent structural elements are shown in gray. The $2F_O-F_C$ map is shown in blue, the F_O-F_C map is shown in green (positive density) and in red (negative density).



8 Danksagung

An dieser Stelle möchte ich mich bei all denen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Zuallererst möchte ich mich herzlich bei meinem Doktorvater Prof. Dr. Reinhard Sterner für das spannende Thema und seine stets vorhandene Hilfsbereitschaft bedanken. Ich danke für sein großes Interesse, die vielen wissenschaftlichen Diskussionen und für das in mich gesetzte Vertrauen die verschiedenen Projekte sehr selbstständig bearbeiten zu können. Nicht zuletzt möchte ich mich auch für die Möglichkeit zur Teilnahme an zahlreichen interessanten Tagungen bedanken.

Herrn PD Dr. Winfried Hausner danke ich für die Übernahme des Zweitgutachtens.

Ein großes Dankeschön gilt den Mitarbeitern von F. Hoffmann – LaRoche (*Pharma Research Discovery*, Basel) ohne deren Unterstützung die Realisierung des hGR-Projektes nicht möglich gewesen wäre. Zuallererst möchte ich Dr. Ralf Thoma danken. Seine große Hilfsbereitschaft und die vielen intensiven Diskussionen mit ihm waren für das Gelingen dieses Projekts unabdingbar. Brigitte D’Arcy danke ich für ihre tatkräftige Unterstützung bei der Aufreinigung der ersten Proteinvarianten in Basel. Bei Dr. Guillaume Schoch, Dr. Jörg Benz und Martine Stihle möchte ich mich für die Aufklärung der wunderschönen Röntgenkristallstrukturen bedanken, welche wesentlich zum erfolgreichen Abschluss des Projekts beigetragen haben. Andrea Wiget, Dr. Stefan Müller, Eric Kusznir und Dr. Francis Müller danke ich für die von Ihnen durchgeführten BIAcoreTM-, Massenspektrometrie- bzw. AUZ-Analysen. Dr. Daniel Schlatter danke ich für die gute Zusammenarbeit während des hPCSK9-Projekts. Bei Dr. Armin Ruf und Dr. Michael Hennig möchte ich mich für ihr großes Interesse und für die interessanten Diskussionen bei unseren regelmäßigen Treffen bedanken.

Für technische Hilfe am MoFlo[®] Zytometer möchte ich mich herzlich bei Wolfgang Forster (Lehrstuhl Mikrobiologie, Universität Regensburg) bedanken. Desweiteren gebührt Dr. Stefan Becker und Prof. Dr. Kolmar (TU Darmstadt) großer Dank für die sehr hilfreiche Einführung in die Benutzung des Geräts zu Beginn des hGR-Projekts.

Besonderer Dank gilt auch Dr. Birte Höcker (Max-Planck Institut für Entwicklungsbiologie, Tübingen). Ihr Einsatz und ihre Hilfsbereitschaft waren unerlässlich für das Gelingen des HisF-C**C* – Projekts. An dieser Stelle möchte ich mich auch ganz herzlich bei meinen ehemaligen Kollegen Dr. Jörg Claren und Adriane Lochner bedanken, die an diesem Projekt maßgeblich beteiligt waren. Unsere Zusammenarbeit hat mir extrem viel Spaß gemacht. Man kann sich wohl keine bessere, produktivere und lustigere Arbeitsatmosphäre vorstellen, als wir drei sie hatten. Ich wünsche den Beiden hiermit alles erdenklich Gute für ihre neuen Tätigkeiten und hoffe dass sich unsere Wege auch in Zukunft kreuzen werden.

Ein weiteres Dankeschön gilt Sonja Fuchs, die mir im Laboralltag immer mit Hilfe zur Seite stand und mit ihrer unerschütterlich positiven Art einen stets unermüdlichen Arbeitseinsatz zeigte. Auch bei Christiane Endres, Jeanette Ueckert, Hermine Reisner und Barbara Keller möchte ich mich hiermit herzlich für die vielen kleinen und großen Hilfestellungen im Labor bedanken.

Auch allen anderen Mitarbeitern des Lehrstuhls Biochemie II danke ich für so manche Hilfe und die stets gute Arbeitsatmosphäre. Dr. Marco Bocola und Dr. Rainer Merkl danke ich für ihre Hilfe bei bioinformatischen Fragestellungen. Dankeschön auch an Dr. Patrick Babinger für seine kompetente Hilfe bei verschiedensten Problemen im Labor. Klaus Tiefenbach war stets eine große Hilfe bei unterschiedlichen biophysikalischen Fragen. Dr. Helmut Durchschlag danke ich für die aufschlussreichen Diskussionen über „echte“ Biochemie und andere Lebensbereiche. Bei Daniel Schneider möchte ich mich bedanken für die humorvolle und musikalische Bereicherung des Laboralltags und für die zahlreiche Kletterexkursionen inklusive der Einführung in die Knotenkunde. Für eine hilfsbereite und stets angenehme Zusammenarbeit möchte ich mich auch bei Dr. Sandra Schlee, Monika Meier, Felix List, Alexander Ehrmann, Susanne Dietrich, Harald Guldan, Linn Carstensen, Thomas Schwab, Bettina Sommer, Josef Sperl, Bernd Reisinger, Kathleen Burghardt, Daniel Roderer und David Peterhoff bedanken. Hermann Zellner und Dietmar Birzer danke ich für das gewährte Asyl während des Zusammenschreibens dieser Arbeit und für die tieferen Einblicke in PyMol.

Meinen Eltern und meinem Bruder Andreas mit Familie danke ich ganz herzlich für die Unterstützung während meines Studiums und meiner Promotion. Auch bei meinen Schwiegereltern und Virginia möchte ich mich für ihre Unterstützung bedanken.

Ganz besonders danke ich Hannah, die mich während des gesamten Studiums unterstützt hat und ohne deren Liebe und Geduld diese Arbeit nicht möglich gewesen wäre.

9 Eidesstattliche Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen und Abbildungen –, die anderen Werken in Wortlaut oder dem Sinn nach entnommen sind, in jedem Fall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen und Tagungsbeiträgen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. R. Sterner betreut worden.

Regensburg, den 02.11.2009

Tobias Seitz

Teilpublikationen

Seitz T., Bocola M., Claren J., Sterner R. (2007) Stabilisation of a $(\beta\alpha)_8$ -barrel protein designed from identical half barrels. *J Mol Biol* 372(1):114-29.

Höcker B., Lochner A., Seitz T., Claren J., Sterner R. (2009) High-resolution crystal structure of an artificial $(\beta\alpha)_8$ -barrel protein designed from identical half barrels. *Biochemistry* 48 (6): 1145–1147.

Tagungsbeiträge

Poster

Claren J., Seitz T., Höcker B., Sterner R. Evolution and design of $(\beta\alpha)_8$ -barrels from $(\beta\alpha)_4$ half-barrels. *19th Faltertage*, Halle (Saale), 2006

Fischer A., Seitz, T., Merkl R., Sterner R., Bocola M. Computational protein stabilization. *8th International Conference on Protein Stabilisation – ProStab*, Graz (Österreich), 2009

Höcker B., Lochner A., Seitz T., Claren J., Sterner R. High-resolution crystal structure of an artificial $(\beta\alpha)_8$ -barrel protein designed from identical half barrels. *VIII European Symposium of the Protein Society*, Zürich (Schweiz), 2009

Seitz T., Thoma R., D'Arcy B., Wiget A., Schoch G., Stihle M., Benz J., Hennig M., Sterner R. Enhancing the solubility and stability of the human glucocorticoid receptor ligand binding domain by library screening using green fluorescent protein as reporter. *VIII European Symposium of the Protein Society*, Zürich (Schweiz), 2009 / *Protein Structure Determination in Industry*, Basel (Schweiz), 2009

Carstensen L., Höcker B., Seitz T., Lochner A., Sterner R. Structure, stability and folding of an artificial $(\beta\alpha)_8$ -barrel protein designed from identical half barrels. *Structure and Stability of Biomacromolecules*, Košice (Slowakei), 2009

Vorträge

“Directed evolution to stabilize an artificially designed $(\beta\alpha)_8$ -barrel protein.”
DFG program SPP1170 Workshop on directed evolution, IU Bremen - 01/08/2006

“Stabilization of an artificially designed protein by directed evolution.”
2nd Symposium of RZBB Regensburg, Kostenz - 28/09/2006

„Stabilisierung eines neu erzeugten $(\beta\alpha)_8$ -Barrel Proteins durch gelenkte Evolution.“
26. Rabensteiner Kolleg, Pottenstein - 22/05/2008

„Protein stabilization by directed evolution.“
20th Faltertage, Lutherstadt-Wittenberg - 26/09/2008

„Erhöhung der Löslichkeit und Stabilität des humanen Glucocorticoid-Rezeptors durch gelenkte Evolution mit GFP als Reporter.“
27. Rabensteiner Kolleg, Pottenstein - 12/06/2009

„Directed evolution case story.“
Protein Structure Determination in Industry, Basel (Schweiz) - 10/11/2009