

Untersuchungen zur Evolution der $(\beta\alpha)_8$ -*barrel* Faltung aus $(\beta\alpha)_2$ -Modulen

Dissertation

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



vorgelegt von
Markus Richter
aus Kelheim

Mai 2008

Promotionsgesuch eingereicht am: 21.05.2008

Kolloquium am: 24.07.2008

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

Prüfungsausschuss:

Vorsitzender: Prof. Dr. Herbert Tschochner

1. Gutachter: Prof. Dr. Reinhard Sterner

2. Gutachter: PD Dr. Wolfram Gronwald

3. Prüfer: Prof. Dr. Dr. Hans Robert Kalbitzer

Die vorliegende Arbeit wurde in der Zeit von Mai 2004 bis Mai 2008 am Lehrstuhl Biochemie II des Institutes für Biophysik und physikalische Biochemie der Naturwissenschaftlichen Fakultät III – Biologie und Vorklinische Medizin der Universität Regensburg unter Leitung von Prof. Dr. Reinhard Sterner durchgeführt.

Inhaltsverzeichnis

Inhaltsverzeichnis	IV
Abbildungsverzeichnis	VIII
Tabellenverzeichnis	XI
Formelverzeichnis	XII
1 Einleitung	1
1.1 Evolution von Proteinen	1
1.2 Einteilung von Proteinstrukturen	2
1.3 Repeats und Symmetrien innerhalb von <i>folds</i>	4
1.3.1 Die all- α - <i>repeats</i>	4
1.3.2 Die all- β - <i>repeats</i>	5
1.3.3 Die gemischten $\alpha\beta$ - <i>repeats</i>	7
1.4 Die $(\beta\alpha)_8$ - <i>barrel</i> Faltung	9
1.4.1 Zweifache Symmetrie innerhalb der $(\beta\alpha)_8$ - <i>barrel</i> Faltung	11
1.4.2 Vierfache Symmetrie innerhalb der $(\beta\alpha)_8$ - <i>barrel</i> Faltung	15
1.5 Ziel der Arbeit	17
2 Material	18
2.1 Geräte	18
2.2 Chemikalien	20
2.3 Kits	20
2.4 Enzyme	20
2.5 Verbrauchsmaterialien	21
2.6 Bakterienstämme	22
2.7 Vektoren	23
2.7.1 pET-Vektoren	23
2.7.1.1 pET21a(+)	24
2.7.1.2 pET24a(+)	24
2.7.1.3 pET28a(+)	25
2.7.2 pTYB-Vektoren	25
2.7.2.1 pTYB1	26
2.8 Oligodesoxyribonukleotide	26
2.8.1 Primer zur Klonierung der HisF-Viertel	27
2.8.2 Primer zur Klonierung einer Faktor Xa-Schnittstelle zwischen HisF_N1 und HisF_N2 bzw. HisF_C1 und HisF_C2	27
2.8.3 Primer zur Klonierung von Varianten der HisF-Viertel	27
2.8.4 Primer zur Klonierung von HisF_N1N1 und HisF_C1C1	28
2.8.5 Primer zur Klonierung von HisF_N1abcd und HisF_C1abcd	28
2.8.6 Vektorspezifische Sequenzierprimer	28
2.9 Längenstandards	29
2.9.1 DNA-Längenstandards	29
2.9.2 Protein-Längenstandards	29
2.10 Nährmedien	31
2.11 Lösungen und Puffer	32
2.12 Software	36
3 Methoden	37
3.1 Mikrobiologische Methoden	37
3.1.1 Vorbereitung von Geräten und Lösungen	37
3.1.2 Anzucht und Lagerung von Bakterienstämmen	37

3.1.3	Transformation von <i>E. coli</i> Zellen	38
3.1.3.1	Präparation chemisch kompetenter Zellen (Inoue <i>et al.</i> , 1990)	38
3.1.3.2	Chemische Transformation (Inoue <i>et al.</i> , 1990).....	38
3.1.3.3	Bestimmung der Transformationseffizienz	38
3.2	Molekularbiologische Methoden	39
3.2.1	Isolierung von Plasmid-DNA im präparativen Maßstab	39
3.2.2	Bestimmung der DNA-Konzentration	39
3.2.3	Polymerasekettenreaktion (PCR).....	40
3.2.4	Kolonie-PCR (Insert-Screening) zur Überprüfung des Klonierungserfolgs..	42
3.2.5	<i>Splicing Overlap Extension-PCR</i> (SOE-PCR)	42
3.2.6	Spaltung von DNA mit Restriktionsendonukleasen (Wilson & Murray, 1991)..	43
3.2.7	Ligation von DNA-Enden.....	44
3.2.8	Agarosegelelektrophorese	44
3.2.9	Extraktion von DNA-Fragmenten aus Agarosegelen.....	44
3.2.10	DNA-Sequenzierung	45
3.3	Proteinchemische Methoden	45
3.3.1	Expression von Genen und Anreicherung von Proteinen in analytischen Mengen	45
3.3.2	Expression von Genen und Anreicherung von Proteinen in präparativen Mengen	46
3.3.3	Rückfaltung der Proteine aus <i>Inclusion bodies</i>	47
3.3.4	Größenausschlusschromatographie (SEC) im präparativen Maßstab	47
3.3.5	Rückfaltung der Proteine aus <i>Inclusion bodies</i> nach Immobilisierung auf einer Affinitätssäule	48
3.3.6	<i>Intein Mediated Purification with an Affinity Chitin-binding Tag</i> (IMPACT™).....	49
3.3.7	Enzymatische Spaltung der HisF-Hälften mit Faktor Xa.....	51
3.3.8	Dialyse von Proteinlösungen.....	51
3.3.9	Einkonzentrieren von Proteinlösungen und Lagerung der Proteine	51
3.3.10	Festphasenpeptidsynthese (Woolley & Merrifield, 1963)	52
3.4	Analytische Methoden.....	53
3.4.1	SDS-Polyacrylamid Gelelektrophorese (SDS-PAGE)	53
3.4.1.1	Tris-Glycin-HCl PAGE	53
3.4.1.2	Tris-Tricin PAGE (Schägger & von Jagow, 1987).....	54
3.4.2	SDS-Gelfärbung.....	55
3.4.2.1	Coomassie-Färbung von SDS-Gelen	55
3.4.2.2	Silberfärbung von SDS-Gelen (Heukeshoven & Dernick, 1988).....	55
3.4.3	Western Blot und Immunodetektion	55
3.4.4	Konzentrationsbestimmung von Proteinen.....	57
3.4.4.1	Konzentrationsbestimmung nach Bradford (1976)	57
3.4.4.2	Konzentrationsbestimmung mittels Absorptionsspektroskopie	57
3.4.5	Analytische Größenausschlusschromatographie (SEC)	59
3.4.6	Analytische Ultrazentrifugation	59
3.4.6.1	Sedimentationsgeschwindigkeitsläufe	59
3.4.6.2	Gleichgewichtsläufe.....	61
3.4.7	Massenspektrometrie	63
3.4.7.1	Elektrospray Ionisations Massenspektrometrie (ESI-MS).....	63
3.4.7.2	Matrix-assisted Laserdesorptions/Ionisations-Massenspektrometrie (MALDI-MS).....	64
3.4.8	Fluoreszenzemissionsspektroskopie.....	66
3.4.9	Circular-Dichroismus-Spektroskopie	66

3.4.10	Stabilitätsuntersuchungen mittels chemischer Denaturierung.....	68
4	Ergebnisse.....	70
4.1	Die vierfache Symmetrie des ($\beta\alpha$) ₈ - <i>barrel</i> Enzyms HisF	70
4.1.1	Struktur- und Sequenzähnlichkeiten der vier HisF-Viertel HisF_N1, HisF_N2, HisF_C1 und HisF_C2	71
4.1.2	Strukturbasiertes Sequenzalignment von HisF-Sequenzen	73
4.2	Expression, Reinigung und Charakterisierung der Hälften HisF_N und HisF_C	76
4.2.1	Expression und Anreicherung von HisF_N und HisF_C.....	76
4.2.2	Bestimmung von Homogenität und Assoziationszustand der Halb <i>barrel</i> über analytische Größenausschlusschromatographie (SEC).....	78
4.2.3	Spektroskopische Charakterisierung von HisF_N und HisF_C	80
4.2.3.1	Fern-UV CD-Spektroskopie	80
4.2.3.2	Fluoreszenzemissionsspektroskopie	82
4.2.4	Stabilitätsuntersuchung von HisF_N und HisF_C	83
4.3	Zusammenfassung bisheriger Ergebnisse zu den HisF-Vierteln	85
4.4	Das Viertel HisF_N1	86
4.4.1	Expression und Anreicherung von HisF_N1.....	86
4.4.2	Bestimmung der Homogenität und des Assoziationszustandes von HisF_N1	88
4.4.2.1	Analytische Gelfiltration	88
4.4.2.2	Chemisches <i>cross-linking</i>	89
4.4.2.3	Analytische Ultrazentrifugation	90
4.4.2.4	Massenspektroskopische Untersuchung von HisF_N1	92
4.4.3	Spektroskopische Charakterisierung von HisF_N1	93
4.4.3.1	Fern-UV CD-Spektroskopie	93
4.4.3.2	Fluoreszenzemissionsspektroskopie	96
4.4.4	Stabilitätsuntersuchung von HisF_N1	97
4.5	Stabilisierung von HisF_N1 mittels rationalem Design.....	98
4.5.1	Klonierung der HisF_N1-Varianten HisF_N1_E46Q und HisF_N1_R5A_E46Q	99
4.5.2	Reinigung der HisF_N1-Varianten	100
4.5.3	Bestimmung der Homogenität und des Assoziationszustandes von HisF_N1_E46Q und HisF_N1_R5A_E46Q	102
4.5.3.1	Analytische Gelfiltration	102
4.5.3.2	Chemisches <i>cross-linking</i>	103
4.5.3.3	Analytische Ultrazentrifugation von HisF_N1_E46Q und HisF_N1_R5A_E46Q	104
4.5.4	Spektroskopische Charakterisierung von HisF_N1_E46Q und HisF_N1_R5A_E46Q	107
4.5.4.1	Fern-UV CD-Spektroskopie	107
4.5.4.2	Fluoreszenzemissionsspektroskopie	110
4.5.5	Stabilitätsuntersuchung von HisF_N1_E46Q und HisF_N1_R5A_E46Q ...	110
4.6	Resümee zum Viertel HisF_N1 und den Varianten HisF_N1_E46Q und HisF_N1_R5A_E46Q	112
4.7	Das Viertel HisF_C1	113
4.7.1	Expression und Anreicherung von HisF_C1.....	113
4.7.2	Untersuchung der Homogenität und des Assoziationszustandes von HisF_C1 mittels analytischer Gelfiltration.....	114
4.7.3	Spektroskopische Charakterisierung von HisF_C1	115
4.7.3.1	Fern-UV CD-Spektroskopie	115

4.7.3.2	Fluoreszenzemissionsspektroskopie	116
4.7.4	Stabilitätsuntersuchung von HisF_C1	117
4.8	Stabilisierung von HisF_C1 mittels rationalem Design.....	118
4.8.1	Planung der Aminosäureaustausche in HisF_C1.....	118
4.8.2	Klonierung der Varianten von HisF_C1	119
4.8.3	Expression und Anreicherung der HisF_C1-Varianten.....	120
4.8.4	Untersuchung der Homogenität und des Assoziationszustandes der HisF_C1 Varianten	122
4.8.4.1	Analytische Gelfiltration	122
4.8.4.2	Chemisches <i>cross-linking</i>	124
4.8.4.3	Analytische Ultrazentrifugation	125
4.8.5	Spektroskopische Charakterisierung.....	128
4.8.5.1	Fern-UV CD-Spektroskopie	128
4.8.5.2	Fluoreszenzemissionsspektroskopie	130
4.8.6	Stabilitätsuntersuchungen der HisF_C1-Varianten	132
4.9	Resümee zum Viertel HisF_C1 und seinen Varianten	133
4.10	Die Viertel HisF_N2 und HisF_C2.....	134
4.10.1	Klonierung von HisF_N2 und HisF_C2 mit Intein als Fusionprotein und deren Reinigung	134
4.10.2	Klonierung einer Faktor Xa-Schnittstelle zwischen die Viertel HisF_N1 und HisF_N2 bzw. HisF_C1 und HisF_C2	136
4.10.3	Festphasenpeptidsynthese von HisF_N2.....	138
4.10.4	Strukturuntersuchungen an HisF_N2.....	140
4.10.5	Festphasenpeptidsynthese von HisF_C2.....	142
4.10.6	Strukturuntersuchungen an HisF_C2_S210C_F210W_A220Q.....	143
4.11	Erstellung eines Modells zur Evolution der ($\beta\alpha$) ₈ -barrel Faltung aus ($\beta\alpha$) ₂ - Modulen	145
4.12	Die duplizierten und fusionierten Viertel HisF_N1N1 und HisF_C1C1.....	146
4.12.1	Klonierung, Expression und Anreicherung von HisF_N1N1	146
4.12.2	Bestimmung der Homogenität und des Assoziationszustandes von HisF_N1N1	148
4.12.2.1	Analytische Gelfiltration	148
4.12.2.2	Analytische Ultrazentrifugation	149
4.12.3	Spektroskopische Charakterisierung von HisF_N1N1	150
4.12.3.1	Fern-UV CD-Spektroskopie	150
4.12.3.2	Fluoreszenzemissionsspektroskopie	151
4.12.4	Stabilitätsuntersuchung von HisF_N1N1.....	152
4.12.5	Resümee zur Charakterisierung von HisF_N1N1	153
4.12.6	Klonierung und Reinigung von HisF_C1C1.....	153
5	Diskussion	155
6	Zusammenfassung	166
7	Literaturverzeichnis	169
8	Abkürzungsverzeichnis	182
9	Anhang.....	186
10	Danksagung.....	188
	Erklärung.....	190

Abbildungsverzeichnis

Abbildung 1: Bänderdiagramme der drei wichtigsten Supersekundärstrukturelemente.	3
Abbildung 2: Bänderdiagramme der <i>all-α-repeats</i>	5
Abbildung 3: Bänderdiagramme der <i>all-β-repeats</i>	6
Abbildung 4: Bänderdiagramme der gemischten α/β - <i>repeats</i>	8
Abbildung 5: Die $(\beta\alpha)_8$ - <i>barrel</i> Faltung.	10
Abbildung 6: Superpositionierung von HisF_N und HisF_C.	12
Abbildung 7: Das Zweistufenmodell der Evolution von <i>hisA</i> und <i>hisF</i>	13
Abbildung 8: Darstellung der vier Schichten im zentralen β -Faltblatt von HisF.	15
Abbildung 9: Unterteilung des $(\beta\alpha)_8$ - <i>barrels</i> von HisF in $(\beta\alpha)_2$ -Module	16
Abbildung 10: <i>Dot-plot</i> von HisF mit dem Programm HHrep.	16
Abbildung 11: Schematische Darstellung des pET-Expressionssystems.	23
Abbildung 12: Schematische Darstellung des Vektors pET21a(+).	24
Abbildung 13: Schematische Darstellung des Vektors pET28a(+).	25
Abbildung 14: Schematische Darstellung des Vektors pTYB1.	26
Abbildung 15: Verwendete DNA-Größenstandards.	29
Abbildung 16: <i>Low Molecular Weight Marker</i>	30
Abbildung 17: <i>Small Fragment Marker</i>	30
Abbildung 18: <i>Splicing Overlap Extension-PCR</i> (SOE-PCR).	43
Abbildung 19: Schematische Darstellung der Chitinaffinitätsreinigung.	50
Abbildung 20: Zeitabhängige Änderung des Konzentrationsgradienten im Sedimentationsgleichgewichtsexperiment.	62
Abbildung 21: Aufteilung von HisF in die $(\beta\alpha)_2$ -Module (HisF_N1, HisF_N2, HisF_C1 und HisF_C2) und deren Superpositionierung.	71
Abbildung 22: Multiples Sequenz <i>alignment</i> der HisF-Viertel.	72
Abbildung 23: Strukturbasiertes Sequenz <i>alignment</i> von HisF.	74
Abbildung 24: Reinigung von HisF_N und HisF_C.	78
Abbildung 25: Analytische Gelfiltration von HisF_N und HisF_C.	79
Abbildung 26: Fern-UV CD-Spektrum von HisF_N und HisF_C.	80
Abbildung 27: Fluoreszenzemissionsspektren von HisF_N und HisF_C.	82
Abbildung 28: Entfaltung von HisF_N und HisF_C durch chemische Denaturierung.	84
Abbildung 29: Reinigung von HisF_N1.	87

Abbildung 30: Analytische Gelfiltration von HisF_N1.....	88
Abbildung 31: Strukturformel von Glutardialdehyd und schematische Darstellung der <i>cross-linking</i> -Reaktion.	89
Abbildung 32: Chemisches <i>cross-linking</i> von HisF_N1.	90
Abbildung 33: Analytische Ultrazentrifugation von HisF_N1.....	92
Abbildung 34: Fern-UV CD-Spektrum von HisF_N1.....	93
Abbildung 35: TFE-Titration von HisF_N1.	95
Abbildung 36: Fluoreszenzemissionsspektren von HisF_N1.	97
Abbildung 37: Entfaltung von HisF_N1 durch chemische Denaturierung.	98
Abbildung 38: Bänderdiagramm von <i>layer 4</i> des modellierten HisF_N1-Tetramers.	99
Abbildung 39: Reinigung von HisF_N1_E46Q und HisF_N1_R5A_E46Q.....	101
Abbildung 40: Analytische Gelfiltration der HisF_N1-Varianten.....	102
Abbildung 41: Chemisches <i>cross-linking</i> der HisF_N1-Varianten.....	103
Abbildung 42: Analytische Ultrazentrifugation der HisF_N1-Varianten.....	106
Abbildung 43: Fern-UV CD-Spektren der HisF_N1-Varianten.....	108
Abbildung 44: TFE-Titration der HisF_N1-Varianten.	108
Abbildung 45: Fluoreszenzemissionsspektren der HisF_N1-Varianten.	110
Abbildung 46: Entfaltung der HisF_N1-Varianten durch chemische Denaturierung	111
Abbildung 47: Reinigung von HisF_C1.....	113
Abbildung 48: Analytische Gelfiltration von HisF_C1.....	114
Abbildung 49: Fern-UV CD-Spektrum von HisF_C1.....	115
Abbildung 50: Fluoreszenzemissionsspektren von HisF_C1.....	117
Abbildung 51: Entfaltung von HisF_C1 durch chemische Denaturierung.	117
Abbildung 52: Bänderdiagramm von HisF_C1.....	119
Abbildung 53: Reinigung der HisF_C1-Varianten.	122
Abbildung 54: Analytische Gelfiltration der HisF_C1-Varianten.....	123
Abbildung 55: Chemisches <i>cross-linking</i> der HisF_C1-Varianten.....	125
Abbildung 56: Analytische Ultrazentrifugation der HisF_C1-Varianten.....	125
Abbildung 57: Fern-UV CD-Spektren der HisF_C1 Varianten.	128
Abbildung 58: TFE-Titration der HisF_C1-Varianten.	129
Abbildung 59: Fluoreszenzemissionsspektren der HisF_C1-Varianten.	131
Abbildung 60: Entfaltung der HisF_C1-Varianten durch chemische Denaturierung	132
Abbildung 61: Reinigungsversuch von HisF_N2 mit Intein als Fusionsprotein.	135

Abbildung 62: Reinigung von HisF_N1XaN2 mittels Rückfaltung auf einer HisTrap™-Säule.	137
Abbildung 63: Bänderdiagramm von HisF_N2.....	138
Abbildung 64: Festphasenpeptidsynthese von HisF_N2.	139
Abbildung 65: Absorptions- und fern-UV CD-Spektrum von HisF_N2_S101C_F120W.	140
Abbildung 66: Fluoreszenzemissionsspektrum von HisF_N2_S101C_F120W.....	141
Abbildung 67: Bänderdiagramm und Festphasenpeptidsynthese von HisF_C2.	142
Abbildung 68: Absorptionsspektren von HisF_C2_S210C_F210W_A220Q.	143
Abbildung 69: TFE-Titration von HisF_C2_S201C_F210W_A220Q.....	144
Abbildung 70: Evolutionsmodell zur Entstehung der $(\beta\alpha)_8$ -barrel Faltung aus $(\beta\alpha)_2$ -Modulen.....	146
Abbildung 71: Reinigung von HisF_N1N1.	147
Abbildung 72: Analytische Gelfiltration von HisF_N1N1.	148
Abbildung 73: Analytische Ultrazentrifugation von HisF_N1N1.	149
Abbildung 74: Fern-UV CD-Spektrum von HisF_N1N1.	150
Abbildung 75: Fluoreszenzemissionsspektren von HisF_N1N1.	152
Abbildung 76: Entfaltung von HisF_N1N1 durch chemisch Denaturierung.....	152
Abbildung 77: Reinigung von HisF_C1C1 mittels Rückfaltung auf einer HisTrap™-Säule.	154

Tabellenverzeichnis

Tabelle 1: Der <i>Low Molecular Weight Marker</i>	30
Tabelle 2: Der „ <i>small fragment marker</i> “	30
Tabelle 3: Zusammensetzung von Tris-Glycin-Gelen.....	54
Tabelle 4: Zusammensetzung von Tris-Tricin-Gelen.	54
Tabelle 5: Definition der HisF-Viertel.	70
Tabelle 6: Sequenzähnlichkeiten und <i>rmsd</i> -Werte der vier HisF-Viertel.....	72
Tabelle 7: Apparente und berechnete Molekulargewichte von HisF_N, HisF_C.	80
Tabelle 8: Sekundärstrukturelementvorhersage für HisF_N, HisF_C.	81
Tabelle 9: Sekundärstrukturelementvorhersage für HisF_N1.....	94
Tabelle 10: Sekundärstrukturelementvorhersage für HisF_N1 in Abhängigkeit von der TFE-Konzentration.	95
Tabelle 11: Apparente und berechnete Molekulargewichte der HisF_N1-Varianten.	103
Tabelle 12: Sekundärstrukturelementvorhersage für die HisF_N1-Varianten in Abhängigkeit von der TFE-Konzentration.	109
Tabelle 13: Apparentes und berechnetes Molekulargewichte von HisF_C1.....	114
Tabelle 14: Sekundärstrukturelementvorhersage für HisF_C1.....	116
Tabelle 15: Zur Herstellung der HisF_C1 Varianten verwendete Primer und Matrizen.	120
Tabelle 16: Reinigungserfolg und Ausbeuten der einzelnen HisF_C1-Varianten. ...	121
Tabelle 17: Apparente Molekulargewichte der HisF_C1-Varianten.	123
Tabelle 18: Sekundärstrukturelementvorhersage der HisF_C1 Varianten.....	129
Tabelle 19: Sekundärstrukturelementvorhersage für die HisF_C1-Varianten in Abhängigkeit von der TFE-Konzentration.	130
Tabelle 20: Sekundärstrukturelementvorhersage für HisF_N2_S101C_F120W.....	141
Tabelle 21: Sekundärstrukturelementvorhersage für HisF_C2_S201C_F210W _A220Q in Abhängigkeit von der TFE-Konzentration.	144
Tabelle 22: Sekundärstrukturelementvorhersage für HisF_N1N1.	151
Tabelle 23: Charakteristische Größen der vermessenen Proteine	186

Formelverzeichnis

Formel 1: Bestimmung der Transformationseffizienz.	39
Formel 2: Photometrische Bestimmung von DNA-Konzentrationen.	40
Formel 3: Berechnung der Schmelztemperatur eines Oligodesoxyribonukleotids....	41
Formel 4: Bestimmung der <i>Annealing</i> -Temperatur in einer PCR-Reaktion..	41
Formel 5: Bestimmung der Proteinkonzentration nach Bradford (1976).	57
Formel 6: Bestimmung des molaren Extinktionskoeffizienten.....	58
Formel 7: Bestimmung des spezifischen Extinktionskoeffizienten.....	58
Formel 8: Bestimmung der Proteinkonzentration mittels Absorptionsspektroskopie.	58
Formel 9: Berechnung des Sedimentationskoeffizienten.....	60
Formel 10: Berechnung des apparenten Sedimentationskoeffizienten s_{app}	60
Formel 11: Umrechnung von s_{app} auf Standardbedingungen ($s_{20,w}$).	61
Formel 12: Berechnung des Molekulargewichtes aus Sedimentationsgleichgewichtsläufen.	62
Formel 13: Berechnung der molaren Elliptizität pro Aminosäurerest.	67
Formel 14: Berechnung des Anteils an ungefaltetem Protein.....	68

1 Einleitung

1.1 *Evolution von Proteinen*

Neben Nukleinsäuren, Kohlenhydraten und Lipiden sind Proteine die elementaren Bausteine des Lebens, weshalb die Erforschung ihrer Evolution seit jeher großes wissenschaftliches Interesse fand. Durch die stetig wachsende Anzahl sequenzierter Genome wurde die Etablierung einer routinemäßigen Methode zur Annotation von Genen notwendig. Diese ist der Sequenzvergleich (Bork & Koonin, 1998). Mithilfe von sequenz- und strukturbasierten Vergleichen konnten auch Ähnlichkeiten zwischen den Genprodukten, also den Proteinen, aufgedeckt werden. Ein wichtiges Ergebnis dieser Untersuchungen war, dass Proteine meist aus einem definierten Satz Domänen aufgebaut sind. Unter einer Domäne versteht man eine strukturell abgeschlossene, oft auch funktionelle Einheit (Doolittle, 1995), die aus jeweils 100 bis 250 Aminosäuren aufgebaut ist. Obwohl viele Proteindomänen in Bakterien, Archaeen und Eukaryoten entstanden und sich anschließend durch Endosymbiose und lateralen Gentransfer auf alle anderen Reiche des Lebens verteilten (Söding & Lupas, 2003), wurde ein Basissatz von Proteindomänen bereits beim letzten gemeinsamen Vorläufer (LUCA; engl. „*last universal common ancestor*“) gefunden (Doolittle, 1999).

Proteinevolution wird durch verschiedene Mechanismen vorangetrieben. Als ein Prozess zur Entstehung der Proteinviefalt ist die Domänenkombination (Domänen-*shuffling*) zu nennen (Doolittle, 1995). Ein Beispiel stellt die Antikörper-Viefalt des menschlichen Immunsystems dar, welche durch fortwährende Neukombination von nur wenigen Domänen erzeugt wird. Auch Prokaryoten nutzen diesen Mechanismus, um ihre Oberflächenproteine zu verändern, mit dem Ziel das Immunsystem des Wirts zu umgehen. Ein zweiter wichtiger Mechanismus zur Entstehung der Proteinviefalt ist die Zirkularpermutation (Ponting & Russell, 1995; Lindqvist & Schneider, 1997), die für die Entstehung von etwa 400 der ca. 3000 Domänen in bekannten Proteinstrukturen verantwortlich gemacht wird (Jung & Lee, 2001).

Den wichtigsten Mechanismus zur Erzeugung der Proteinviefalt stellt jedoch zweifelsohne die Duplikation von Genen bzw. ganzen Genomen dar (Ohno, 1970). In

Prokaryoten sind etwa 70 %, in Eukaryoten sogar ca. 90 % der Proteindomänen durch Duplikationsereignisse entstanden (Chothia *et al.*, 2003). Die Duplikation von DNA-Abschnitten kann auf verschiedenen Ebenen erfolgen. Zum Beispiel ermöglicht die Duplikation eines gesamten Genoms die Erhöhung des Komplexitätsgrades eines Organismus (Wolfe & Shields, 1997; Vision *et al.*, 2000; McLaysaght *et al.*, 2002). Hingegen erlaubt die Duplikation von bereits bestehenden Operonen die Etablierung neuer Biosynthesewege. Dabei müssen die Produkte der duplizierten Gene neue Funktionen etablieren, um nicht inaktiviert zu werden (Lynch & Conery, 2000). Weiterhin bietet die Genduplikation mit anschließender Fusion die Grundlage zur Entstehung neuer (Einzel-)Proteine. So ist die ungleiche Rekombination von Genen (Marcotte *et al.*, 1999) für die Entstehung repetitiver Proteine verantwortlich (Marcotte *et al.*, 1999; Kajava, 2001). Extrembeispiele stellen das Titin und Nebulin aus Muskeln dar, die aus Hunderten von Immunglobulin-Domänen aufgebaut sind. Durch Punktmutationen, Insertionen und Deletionen können durch Genduplikation entstandene Proteine soweit verändert werden, dass auf Sequenzebene keine Ähnlichkeiten mehr bestehen, obwohl die Proteine noch die gleiche Funktion erfüllen. Auf Strukturebene hingegen sollten aus einem gemeinsamen Vorläufer entstandene Proteine dennoch Ähnlichkeiten aufweisen, da sich Strukturmerkmale im Verhältnis zu Sequenzmerkmalen nur langsam verändern (Chothia & Lesk, 1986). Deshalb ist neben Sequenzvergleichen eine auf dem Vergleich der Strukturen basierende Klassifizierung von Proteinen sinnvoll, um auf ihre evolutionäre Verwandtschaft schließen zu können.

1.2 Einteilung von Proteinstrukturen

Aminosäuresequenz und Struktur von Proteindomänen bilden die Grundlage für die Zuordnung von den Proteinen, die in den bekannten Datenbanken, wie SCOP („*structural classification of proteins*“; <http://scop.mrc-lmb.cam.ac.uk/scop/>; (Murzin *et al.*, 1995)) und CATH („*class, architecture, topology or fold, homologous superfamily*“; <http://www.cathdb.info>; (Orengo *et al.*, 1997)) hinterlegt sind.

Im Falle der SCOP-Datenbank werden die Proteine mit einer klar ersichtlichen evolutionären Verwandtschaft (Sequenzidentität $\geq 30\%$) in „Familien“

zusammengefasst. Mehrere Familien, deren Mitglieder untereinander eine geringe Sequenzidentität aber strukturelle und funktionelle Ähnlichkeiten aufweisen, bilden eine „Superfamilie“, wobei die einzelnen Vertreter Strukturähnlichkeiten von bis zu 50 % aufweisen (Orengo *et al.*, 2001; Redfern *et al.*, 2005). Dies ermöglicht die Ableitung eines *fingerprints* für entfernt verwandte Proteine (Orengo, 1999).

Auf der nächsten Hierarchiestufe werden mehrere „Superfamilien“ zu *folds* gruppiert, deren Mitglieder durch die gleiche Anordnung und Abfolge von den Sekundärstrukturelementen α -Helix, β -Strang und *random coil* gekennzeichnet sind. Die Gesamtzahl an unterschiedlichen *folds* beläuft sich auf mindestens 10000, wobei zwischen *unifolds*, *mesofolds* und *superfolds* unterschieden werden kann. Ein *unifold* ist jeweils nur bei den Mitgliedern einer „Familie“ zu finden. Dies trifft nach einem Modell von Coulson und Moulton aus dem Jahre 2002 auf die meisten *folds* zu. *Mesofolds* hingegen wurden in einer mittleren Anzahl von Familien gefunden. Nach Coulson und Moulton (2002) besitzen 80 % aller Sequenz-Familien einen von etwa 400 *folds*, die zum Großteil bekannt sind. *Superfolds* werden in vielen Proteinfamilien gefunden. Was ist neben dem gleichen verwandtschaftlichen Ursprung die Ursache der Ausbildung von wiederkehrenden *folds*? Die Stabilität und Faltungseffizienz einzelner Typen gehört nach Govindarajan & Goldstein (1996) und Salem *et al.* (1999) dazu. Ein weiterer Grund ist die bessere Tauglichkeit einiger *folds* als Grundgerüst („*scaffold*“) zur Ausbildung aktiver Zentren (Ponting & Russell, 2002). Einige *folds* bestehen aus einer begrenzten Anzahl kompakter, sich wiederholender Sekundärstrukturelemente, die als Supersekundärstrukturelemente bezeichnet werden (Holm & Sander, 1998). Die häufigsten Supersekundärstrukturelemente sind namentlich der $\beta\beta$ -*hairpin*, der $\alpha\alpha$ -*hairpin* und gemischte $\beta\alpha\beta$ -Elemente, die im Durchschnitt über 60 % aller *folds* bilden (Abbildung 1).



Abbildung 1: Bänderdiagramme der drei wichtigsten Supersekundärstrukturelemente.

- (A) In grün ist der $\beta\beta$ -*hairpin* dargestellt, die Pfeile stehen für β -Stränge.
- (B) In rot ist der $\alpha\alpha$ -*hairpin* dargestellt, die Schrauben symbolisieren α -Helices.
- (C) In blau ist das $\beta\alpha\beta$ -Element dargestellt, wobei auch hier die Pfeile den β -Strängen und die Schraube der α -Helix entsprechen. (Nach Söding *et al.*, 2003).

Die Hypothese, dass Proteindomänen aus einem kleinen Satz von Supersekundärstrukturelementen entstanden sind, wird durch diese Tatsache ebenso gestützt, wie durch die Annahme, dass in der evolutionären Frühphase der Peptidbiosynthese aufgrund des noch mangelhaften DNA-Reperatursystems vermutlich nur relativ kurze korrekte Polypeptidketten gebildet werden konnten (Lupas *et al.*, 2001). Dieser kleine Satz an Peptiddomänen bzw. kleinen Supersekundärstrukturelementen hatte dann die Möglichkeit, zu Proteindomänen zu oligomerisieren und zu fusionieren.

1.3 *Repeats und Symmetrien innerhalb von folds*

Als einfachste *repeats* gelten monotone Abfolgen eines der oben genannten Supersekundärstrukturelemente α -*hairpin*, β β -*hairpin* und β α β -Elemente, die meist in rinnenförmige Strukturen münden. Aber auch globuläre Proteine können durch die repetitive Abfolge eines der Supersekundärstrukturelemente gebildet werden (Söding & Lupas, 2003).

1.3.1 Die *all- α -repeats*

Der Tetratricopeptid (TRP)-ähnliche, der Armadillo- und der HEAT (nach den ersten vier Proteinen benannt, in denen der *repeat* zuerst beschrieben wurde)-*repeat* zählen zu den sogenannten *all- α* Strukturen (allein aus α -Helices aufgebaute Strukturen; Abbildung 2). Der TRP-ähnliche *repeat* wird durch zwei jeweils aus 34 Aminosäuren bestehenden α -Helices, die sich nebeneinander anlagern, gebildet (Sikorski *et al.*, 1990; Lamb *et al.*, 1995). Bei diesem sehr ursprünglichen *repeat*, der sich in Eukaryoten, Bakterien und Archaeen findet (Ponting *et al.*, 1999), wird aufgrund hoher Sequenzähnlichkeiten von einem gemeinsamen Vorläufer, d.h. divergenter Evolution ausgegangen. Die Hsp70-Chaperone (Scheufler *et al.*, 2000), die leichte Kette des Kinesin (Gindhart & Goldstein, 1996), die schwere Kette des Clathrin und die Aspartylphosphat-Phosphatase (Andrade *et al.*, 2000) sind Beispiele für das Vorhandensein TRP-ähnlicher *repeats* (Abbildung 2A).

Die Armadillo- (Peifer *et al.*, 1994) und HEAT-*repeats* (Andrade & Bork, 1995) bestehen aus drei bzw. zwei antiparallelen α -Helices mit jeweils 33 bzw. 47 Aminosäureresten (Abbildung 2B). Beide *repeats* weisen eine hohe Sequenz-ähnlichkeit auf, vor allem innerhalb der Reste im hydrophoben Kern (Andrade *et al.*, 2001). Dies deutet ebenfalls auf eine divergente Evolution der beiden *repeats* hin, die ausschließlich an Protein-Protein-Wechselwirkungen beteiligt und somit an Kontaktflächen von Proteinen zu finden sind.

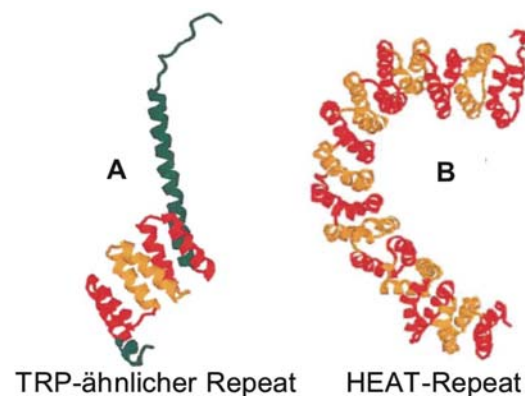


Abbildung 2: Bänderdiagramme von Proteinstrukturen aus *all- α -repeats*.

Die sich wiederholenden Einheiten sind abwechselnd in rot und orange dargestellt (Nach Andrade *et al.* 2001).

- (A) Ein Fragment der menschlichen Phosphatase 5 (PDB: 1a17), das dreimal den TetraRicopeptid (TRP)-*repeat* enthält (Das *et al.*, 1998).
- (B) Die PR65/A Untereinheit der menschlichen Proteinphosphatase 2A (PDB: Kette A von 1b3u), eines der vier Proteine, die dem HEAT-*repeat* seinen Namen verleihen (Groves *et al.*, 1999). Der Armadillo-*repeat* entspricht dem HEAT-*repeat* mit Ausnahme einer zusätzlichen α -Helix.

1.3.2 Die *all- β -repeats*

Sogenannte *all- β -repeats* (nur aus β -Strängen aufgebaut; Abbildung 3) bilden geschlossene globuläre Strukturen aus. Ein Vertreter dieser Klasse ist der WD40-Repeat (Neer *et al.*, 1994), der meist durch circa 40 Aminosäuren gebildet wird und immer ein konserviertes Tryptophan (W) sowie ein konserviertes Aspartat (D) enthält. Durch vier- bis achtmalige Wiederholung eines viersträngigen antiparallelen β -Faltblattes (= *blade*) werden die häufigsten β -Propeller Strukturen gebildet (Abbildung 3A), die intern symmetrisch sind. So zeigen Vertreter der 6-blättrigen β -Propeller-Faltung eine sechsfache interne Wiederholung des Faltblattes (McLachlan, 1987). Da es sich um eine geschlossene 3D-Struktur handelt, bei der N- und C-

Terminus interagieren, wird eine weitere Diversifikation durch Zirkularpermutation ermöglicht (Neer & Smith, 1996). In Versuchen mit Tachylectin-2, welches aufgrund seiner 5-*blade* β -Propeller Struktur eine fünffache Symmetrie besitzt, lagerten sich Peptide von durchschnittlich 100 Aminosäureresten Länge zu Homopentameren zusammen (Yadid & Tawfik, 2007). Allerdings entspricht das Molekulargewicht des in diesen Experimenten identifizierten Moduls dem von zwei *blades* (Modulen). Dies könnte ein Hinweis auf ein „verstecktes“ Modul sein. Es war nicht möglich, ein stabiles Fragment zu isolieren, das dem Molekulargewicht eines Einzelmoduls (ca. 50 Aminosäuren) entspricht.

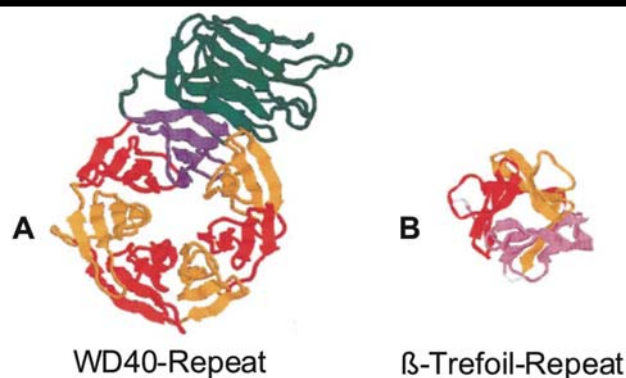


Abbildung 3: Bänderdiagramme von Proteinstrukturen aus *all*- β -repeats.

Die sich wiederholenden Einheiten sind abwechselnd in rot und orange dargestellt (Nach Andrade *et al.*, 2001).

- (A) WD40-repeat der Galaktoseoxidase (PDB: 1gof) aus *D. dendroides* (Ito *et al.*, 1991). Zu sehen ist die sechsfache Symmetrie innerhalb des β -Propeller.
- (B) β -Trefoil repeats des menschlichen Fibroblasten-Wachstumsfaktors (PDB: Kette A von 2afg) (Eriksson *et al.*, 1993), die für dreifache interne Symmetrie verantwortlich sind.

Eine zweite Klasse von *all*- β Strukturen bilden *repeats*, die zur β -Trefoil-Faltung (Murzin *et al.*, 1992) führen. Hier sind sechs zweisträngige β -*hairpins* aus jeweils etwa 40 Aminosäuren assembliert (Abbildung 3B). Drei der sechs β -*hairpins* lagern sich zu einer Fassstruktur (*barrel*) zusammen. Die restlichen drei bilden eine Kappe (Murzin *et al.*, 1992). Wie der Trypsin-Inhibitor aus der Sojabohne (McLachlan, 1979) und das Cytokin Interleukin 1 des Menschen (Priestle *et al.*, 1988) zeigen, besitzen β -Trefoil-Faltungs-Domänen meist eine dreifache interne Symmetrie. Dass eine β -Trefoil-Faltungs-Domäne isoliert stabil sein kann, wurde am Beispiel der β -Trefoil-Domäne der schweren Kette des Botulinustoxins gezeigt (Sharma *et al.*, 2006), die an Synaptotagmin, einem bekannten Rezeptor des Toxins, und an Ganglioside binden konnte. Das komplette Botulinustoxin besteht aus einer leichten und

schweren Kette, die über eine Disulfidbrücke verbunden sind und zusammen aus 1296 Aminosäuren bestehen. Die β -Trefoil Domäne startet bei Aminosäure 1092 und stellt somit einen Teil der C-terminalen Hälfte des gesamten Proteins dar.

1.3.3 Die gemischten $\alpha\beta$ -repeats

Eine der zwei gemischten $\alpha\beta$ -Strukturen ist der Ankyrin-*repeat* (ANK, Abbildung 4A). Er wurde benannt nach dem ersten Protein, in dem er beschrieben wurde (Lux *et al.*, 1990). Der ANK-*repeat* bildet mit Hilfe von 33 Aminosäureresten eine L-förmige Struktur, die aus zwei antiparallelen α -Helices gefolgt von einem β -*hairpin* (Gorina & Pavletich, 1996) besteht. Proteine, die diesen *repeat* besitzen, sind nicht auf eine Klasse fixiert. Dennoch ist er meist bei der Ausbildung von Protein-Protein-Wechselwirkungen zu finden (Sedgwick & Smerdon, 1999). Das Tumorsuppressorprotein p16^{INK4} (p16) ist das kleinste strukturell charakterisierte Protein mit ANK-*repeat*. Es besteht aus vier ANK-*repeats* mit N- und C-terminalen Fortsätzen, die in Lösung unstrukturiert vorliegen. Um zu überprüfen, ob ein einzelner 33 Aminosäuren langer *repeat* die kleinste strukturierte Einheit darstellt, wurde zuerst die C-terminale Hälfte von Tumorsuppressor p16 (bestehend aus *repeat* 3 und 4) charakterisiert (Zhang & Peng, 2000). Das Ergebnis war eine annähernd wildtypische Faltung. Ein weiteres Unterteilen in die Einzel-*repeats* hatte keinen Erfolg, die 33 Aminosäure-Einheiten wiesen in Lösung keinerlei Struktur auf. Dies veranlasste die Gruppe anzunehmen, dass die kleinste Domäne des ANK-*repeats* aus zwei Einzel-*repeats* besteht. Dies stützt die These, dass mehrere *repeats* benötigt werden, um eine gegenseitige Stabilisierung zu ermöglichen (McDonald & Peters, 1998).

Der zweite zur Klasse der gemischten $\alpha\beta$ -*repeats* gehörende Vertreter ist der Leucinreiche *repeat* (Abbildung 4B), der aus 20 Aminosäuren einen β -Strang mit antiparalleler α -Helix ausbildet (Kobe & Deisenhofer, 1994; Price *et al.*, 1998). Die repetitiven Einheiten lagern sich in einem Bogen aneinander, wodurch sich eine große Grube für Protein-Protein-Wechselwirkungen ergibt. Dieser *repeat* findet sich überwiegend bei Proteinen der Signaltransduktion und des DNA-Reparatursystems, sowie bei Transmembranrezeptoren. Als Beispiele sind Glykoproteine (Iozzo, 1998), TGF β (Hildebrand *et al.*, 1994) und Kollagen (Svensson *et al.*, 2000) zu nennen.

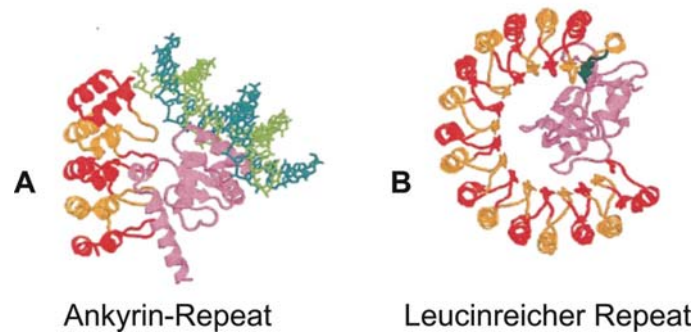


Abbildung 4: Bänderdiagramme von Proteinstrukturen aus gemischten α/β -Repeats.

Die sich wiederholenden Einheiten sind abwechselnd in rot und orange dargestellt (Nach Andrade *et al.* 2001).

- (A) Ein Fragment der β -Untereinheit des GA-bindenden Proteins (PDB: Kette B von 1awc) aus der Maus (Batchelor *et al.*, 1998) das den Ankyrin-repeat beinhaltet. Zu sehen ist der Komplex mit der α -Untereinheit und gebundener DNA (21 bp).
- (B) Leucinreicher repeat des Ribonukleaseinhibitors (PDB: Kette I von 1dfj) aus dem Schwein im Komplex mit der Ribonuklease (Kobe & Deisenhofer, 1995).

Neben den oben beschriebenen sechs Faltungstypen gibt es weitere weniger häufig vorkommende *repeats* mit sich wiederholenden Faltungseinheiten von 12 – 106 Aminosäuren Länge.

Weiterhin lassen sich strukturelle Homologien und damit Hinweise auf Genduplikation innerhalb von Proteinen nachweisen, die auf den ersten Blick nur aus einer Domäne zu bestehen scheinen (McLachlan, 1987). So zeigen die Sequenzen von Ferredoxinen (Fukuyama *et al.*, 1988) und Calcium-bindenden Domänen wie Parvalbumin (McLachlan, 1987), interne Wiederholungen, die sich in einer zweifachen Symmetrie der Struktur widerspiegeln. Ähnliches gilt für die Rossmann-Faltung (Rossmann *et al.*, 1974; McLachlan, 1987) oder γ -Crystallin-ähnliche Proteine (Bloemendal *et al.*, 2004).

Die Erhaltung solcher repetitiver Muster kann neben funktionellen Ursachen (McLachlan, 1987) auch in der wechselseitigen Stabilisierung von duplizierten und fusionierten Proteindomänen zur Erniedrigung der Translations- und Rotationsentropie des entfalteten Zustands begründet sein (Terwilliger, 1995).

Verschiedene Domänen eines Proteins können auch ohne auffällige interne Sequenz- und Strukturähnlichkeiten aus einem kleineren gemeinsamen Vorläuferprotein entstanden sein, nach dessen Duplikation und Fusion Änderungen in der Faltung zu einer optimalen Assoziation der beiden Kopien geführt haben (Kinch & Grishin, 2002). So wurde bei der künstlichen Herstellung eines N-terminalen

Fragments von Ubiquitin eine Dimerisierung beobachtet, bei der sich die α -Helices und β -Stränge so umordnen, dass die native Ubiquitin-Faltung nachgeahmt wird (Bolton *et al.*, 2001).

Auch bei der $(\beta\alpha)_8$ -*barrel* Faltung, eine auf den ersten Blick Eindomänen-Faltung, sind interne Symmetrien und ein modularer Aufbau aus $\beta\alpha$ -Elementen zu erkennen. Durch mehrfache Duplikation solcher Einheiten könnte dieser Faltungstyp entstanden sein.

1.4 Die $(\beta\alpha)_8$ -*barrel* Faltung

Bei der $(\beta\alpha)_8$ -*barrel* Faltung handelt es sich um einen der ältesten Faltungstypen in der Natur. Dies haben phylogenetische Untersuchungen zur Verteilung von Faltungstypen in essentiellen metabolischen Biosynthesewegen gezeigt (Caetano-Anolles *et al.*, 2007). Ungefähr 10% aller Polypeptidketten enthalten ein $(\beta\alpha)_8$ -*barrel*, welche damit die am häufigsten vorkommende Proteindomäne darstellt (Wierenga, 2001). Expressionsstudien von Hefegenomen zeigten, dass sogar 8 % aller transkribierten Einheiten für eine $(\beta\alpha)_8$ -*barrel* Domäne kodieren (Jansen & Gerstein, 2000). Die $(\beta\alpha)_8$ -*barrel* Faltung konnte erstmals bei der Triosephosphatisomerase (TIM) aus *Gallus gallus* nachgewiesen werden (Banner *et al.*, 1975). Daher stammt die alternative Bezeichnung TIM-*barrel*. Fast alle Proteine, die diesen Faltungstyp enthalten sind Enzyme. Mit Ausnahme der Ligasen sind sie in allen von der Enzymkommission (EC) definierten Enzymklassen vertreten (Nagano *et al.*, 2002). Einige von ihnen gehören zu den effizientesten Biokatalysatoren überhaupt. So ist die TIM eines der wenigen bekannten diffusionskontrollierten Enzyme (Knowles, 1991). Die Orotidin-5'-monophosphat (OMP)-Decarboxylase beschleunigt die von ihr katalysierte Reaktion um den Faktor 10^{17} und stellt damit das leistungsstärkste bekannte Enzym dar (Wierenga, 2001).

Ein kanonisches $(\beta\alpha)_8$ -*barrel* besteht aus mindestens 200 Aminosäuren und setzt sich aus acht Modulen zusammen, die jeweils einen β -Strang und eine α -Helix enthalten (Sterner & Höcker, 2005). Innerhalb eines Moduls wird der β -Strang von den Aminosäuren der N-terminalen Hälfte gebildet, die α -Helix von den C-terminalen. Dabei ist der C-Terminus des β -Stranges mit dem N-Terminus der α -Helix durch

einen $\beta\alpha$ -loop verknüpft. Die Module untereinander wiederum sind durch $\alpha\beta$ -loops verbunden, welche den C-Terminus der α -Helix von Modul n mit dem N-Terminus des β -Stranges von Modul n+1 verknüpfen. Dadurch ergibt sich insgesamt eine $\beta\alpha$ -Supersekundärstruktur (Abbildung 5A), ähnlich den $\beta\alpha\beta$ -Supersekundärstrukturelementen, die in 1.2 dargestellt wurden. In der Tertiärstruktur lagern sich die acht β -Stränge zu einem parallelen β -Faltblatt zusammen. Dieses zentrale Fass (*barrel*) hat im Falle der TIM aus *Trypanosoma brucei* einen Durchmesser von 7 Å (Wierenga, 2001). Die acht, meist amphipatischen α -Helices umgeben dieses Fass. Mit ihrer hydrophilen Seite stehen sie in Kontakt zum Lösungsmittel, während sie mit ihrer hydrophoben Seite mit den β -Strängen interagieren (Abbildung 5B).

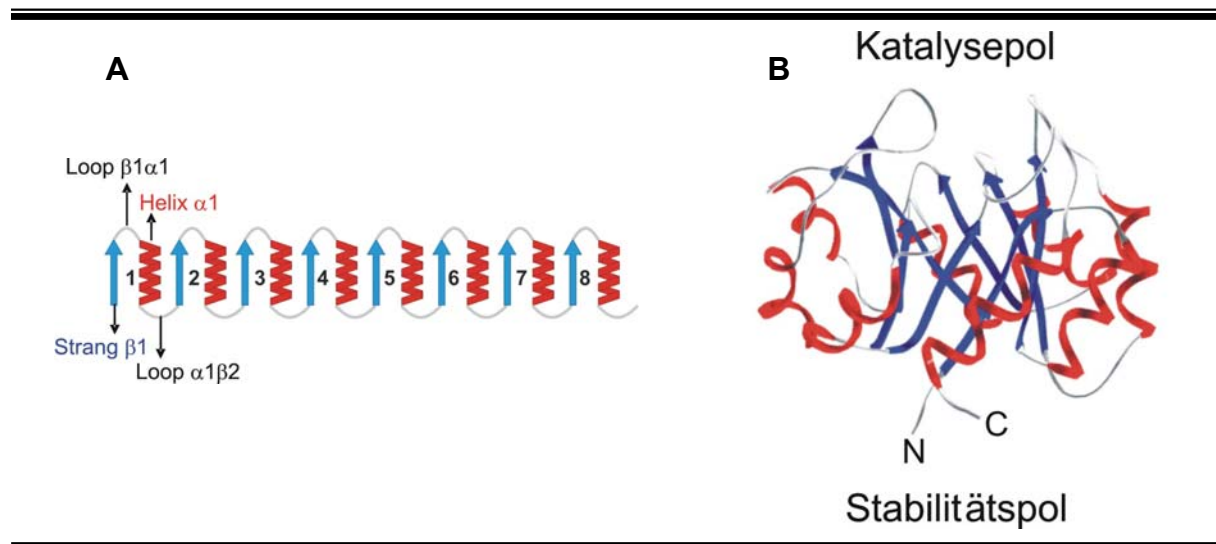


Abbildung 5: Die $(\beta\alpha)_8$ -barrel Faltung. β -Stränge sind in blau, α -Helices in rot dargestellt.

- (A) Die topologische Darstellung der rechtsgängigen $\beta\alpha$ -Supersekundärstruktur. Sie ist aus acht $\beta\alpha$ -Modulen aufgebaut, was zur pseudo-achtfachen Symmetrie innerhalb der Sekundärstruktur der $(\beta\alpha)_8$ -barrel Faltung führt.
- (B) Seitenansicht auf das Bänderdiagramm des $(\beta\alpha)_8$ -barrel Proteins Imidazolglycerolphosphat-Synthase (HisF) aus *Thermotoga maritima* (Lang *et al.*, 2000). Die katalytisch essentiellen Reste liegen an den C-terminalen Enden der β -Stränge und in den darauffolgenden $\beta\alpha$ -loops („Katalysepol“). Der Rest der Faltung, inklusive der gegenüberliegenden $\alpha\beta$ -loops, ist für die konformationelle Stabilität des Proteins verantwortlich („Stabilitätspol“) (Nach Sterner und Höcker, 2005).

Bei allen $(\beta\alpha)_8$ -barrel Enzymen ist eine Aufgabenteilung unterschiedlicher Bereiche innerhalb des Proteingerüsts konserviert. So befinden sich die katalytisch essentiellen Aminosäuren stets an den C-terminalen Enden der β -Stränge bzw. in den $\beta\alpha$ -loops. Die $\alpha\beta$ -loops dagegen sind zumeist unflexibler und kürzer als die $\beta\alpha$ -loops, was wichtig für die konformationelle Stabilität des Proteins ist (Urfer & Kirschner, 1992). Dadurch kann zwischen einem Katalysepol und einem

Stabilitätspol unterschieden werden (Abbildung 5B). $(\beta\alpha)_8$ -*barrel* Enzyme stellen somit ein gutes Strukturgerüst für Modifizierungen katalytischer Aktivitäten dar, da Aminosäuren am Katalysepol ohne Beeinträchtigung der Stabilität des Proteins ausgetauscht werden können. Dies könnte auch der Grund dafür sein, weshalb diese Struktur in der Natur so weit verbreitet ist und macht sie zu einem geeigneten Objekt zur Veränderung katalytischer Aktivitäten durch gerichtete Evolution im Labor (Sternier & Höcker, 2005).

Die große Anzahl verschiedener $(\beta\alpha)_8$ -*barrel* Proteine macht eine Kategorisierung notwendig. Eine erste Einteilung in vier verschiedene Strukturfamilien erfolgte von Faber und Petsko (1990). In den verschiedenen Datenbanken zur Erfassung von Proteinstrukturen (1.2) stellen die $(\beta\alpha)_8$ -*barrel* Proteine einen zentralen *fold* dar und werden anhand ihres Ähnlichkeitsgrades eingeteilt. In der SCOP-Datenbank sind mittlerweile 33 Superfamilien von $(\beta\alpha)_8$ -*barrel* Proteinen enthalten (Stand: Februar 2008).

1.4.1 Zweifache Symmetrie innerhalb der $(\beta\alpha)_8$ -*barrel* Faltung

Neben der Frage nach den Verwandtschaftsbeziehungen von $(\beta\alpha)_8$ -*barrel* Enzymen wird seit einiger Zeit die Entwicklung des $(\beta\alpha)_8$ -*barrel* Strukturgerüsts untersucht. Bis vor kurzem wurde angenommen, dass $(\beta\alpha)_8$ -*barrel* Proteine aus einer einzigen Domäne bestehen. Eine genaue Analyse der Sequenzen und Strukturen der beiden $(\beta\alpha)_8$ -*barrel* Proteine HisA und HisF – zwei Enzyme aus der Histidinbiosynthese, die zwei aufeinanderfolgende Syntheseschritte katalysieren (Henn-Sax *et al.*, 2001) – ergab jedoch starke Ähnlichkeiten zwischen den N-terminalen ($\beta\alpha$ -Module 1 - 4, im Folgenden bezeichnet als HisA_N bzw. HisF_N) und den C-terminalen Hälften ($\beta\alpha$ -Module 5 - 8, im Folgenden bezeichnet als HisA_C bzw. HisF_C) beider Enzyme. Strukturbasierte Sequenzalignments von HisA_N, HisF_N, HisA_C und HisF_C aus *Thermotoga maritima* brachten Sequenzidentitäten zwischen 15 und 26% zutage, die *rmsd*-Werte (vom engl. „root mean square deviation“) der C α -Atome lagen zwischen 1,55 und 2,10 Å (Lang *et al.*, 2000).

In Abbildung 6 ist exemplarisch die Superpositionierung von HisF_N und HisF_C dargestellt. Sie zeigt, dass die beiden essentiellen Aspartatreste am C-terminalen Ende von β -Strang 1 und 5 perfekt zur Deckung kommen, ebenso wie die beiden im

Kristall gebundenen Phosphatmoleküle, welche vermutlich die Positionen der zwei Phosphatgruppen des Substrates N-[5'-phosphoribosylformimino]-5-aminoimidazol-4-carboxamid-Ribonukleotid (PRFAR) (Abbildung 6) einnehmen (Höcker *et al.*, 2001).

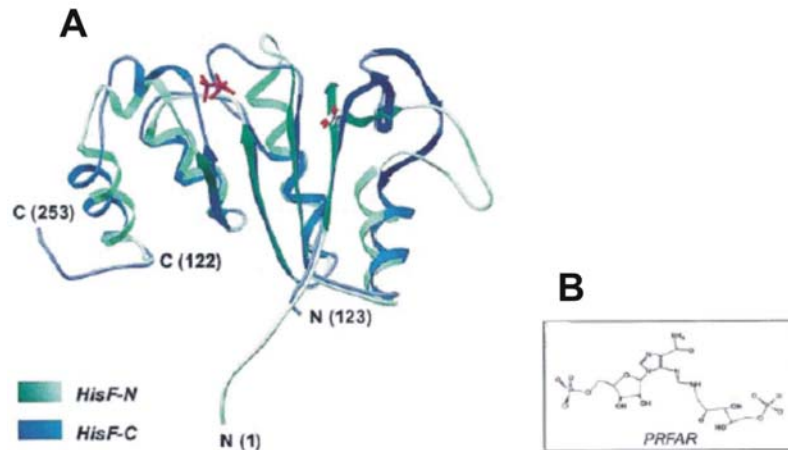


Abbildung 6: Superpositionierung von HisF_N und HisF_C (Nach Höcker *et al.*, 2001).

- (A) Die C α -Atome von HisF_N ($\beta\alpha_{1-4}$; blau) und HisF_C ($\beta\alpha_{5-8}$; grün) kommen mit einem *rmsd*-Wert von 1,58 Å zur Deckung. Die zwei essentiellen Reste Asp11 (am C-terminalen Ende von β -Strang 1) und Asp130 (am C-terminalen Ende von β -Strang 5) – die Seitenketten sind als *sticks* gezeigt – und die beiden gebundenen Phosphatmoleküle (violette Tetraeder) an den C-terminalen Ende von β -Strang 4 und 8 superpositionieren sehr gut.
- (B) Strukturformel des Substrates N-[5'-phosphoribosylformimino]-5-aminoimidazol-4-carboxamid-Ribonukleotid (PRFAR) von HisF.

Nach getrennter Expression in *E. coli* und Reinigung von HisF_N und HisF_C zeigen beide Hälften native Sekundär- und Tertiärstrukturen und bilden zur Stabilisierung ein Homodimer. Die Zusammenlagerung von HisF_N und HisF_C zu einem Heterodimer konnte sowohl *in vitro* nach gemeinsamer Rückfaltung der getrennt gereinigten Hälften, als auch *in vivo* nach Koexpression beider Gene beobachtet werden (Höcker *et al.*, 2001). Durch Fusion zweier identischer HisF_C Hälften erhielt man das artifizielle Protein HisF_CC, das hauptsächlich als Monomer vorlag und etwas stabiler war als HisF_C, wie Auffaltungskurven belegten. Durch rationales Design wurde in *layer* 4 von HisF_CC das Salzbrückencluster nachgebildet, das in HisF wesentlich zur Stabilität des Proteins beiträgt. Man erhielt das ausschließlich monomere Protein HisF_C*C, das zudem eine erhöhte Stabilität im Vergleich zu HisF_CC aufweist (Höcker *et al.*, 2004). Durch gerichtete Evolution und Selektion auf erhöhte Löslichkeit nach Fusion an die Chloramphenicolacetyltransferase (CAT) konnte HisF_C*C weiter zu HisF_C**C stabilisiert und als lösliches Protein aus *E.*

coli gereinigt werden (Seitz *et al.*, 2007). Durch rationales Design gelang eine nochmalige Stabilisierung zu HisF_C***C, das kristallisiert und dessen Röntgenstruktur mit einer Auflösung von 2,0 Å aufgeklärt werden konnte (Lochner, 2008). Diese Befunde stützen die von Fani *et al.* (1994) aufgestellte Theorie, dass der gemeinsame Vorläufer von HisF und HisA durch Duplikation und Fusion eines $(\beta\alpha)_4$ -barrel-Proteins entstanden ist (Abbildung 7) und bestätigen die Vermutung, dass zumindest diese beiden $(\beta\alpha)_8$ -barrel-Enzyme aus zwei Domänen bestehen.

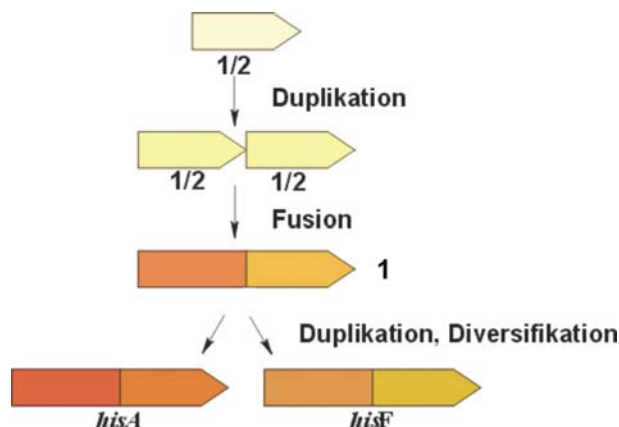


Abbildung 7: Das Zweistufenmodell der Evolution von *hisA* und *hisF* (Nach Fani *et al.*, 1994).

Ausgangspunkt ist ein DNA-Abschnitt, der für ein $(\beta\alpha)_4$ -Halbbarrel kodiert (1/2). Dieser konnte nach Duplikation zum Vollständigen fusionieren. Dieses entspricht einem ursprünglichem $(\beta\alpha)_8$ -barrel (1). Durch sich anschließende erneute Duplikation und Diversifikation sind die beiden Gene für HisF und HisA entstanden.

Weitere Untersuchungen ließen annehmen, dass es sich bei einem $(\beta\alpha)_4$ -barrel tatsächlich um eine unabhängig evolvierende Domäne handeln kann. So wurden bei einer umfassenden Datenbanksuche nach Proteinen mit strukturellen Ähnlichkeiten zu HisF_C und HisF_N mehrere Mitglieder der Flavodoxin-ähnlichen Faltungsfamilie identifiziert (Höcker *et al.*, 2002). Die Mitglieder dieser Familie treten entweder isoliert oder als Domäne in größeren Proteinen auf (Murzin *et al.*, 1995). Sie weisen eine $(\beta\alpha)_5$ -Faltung aus fünf parallelen β -Strängen und α -Helices auf, die durch *loops* miteinander verbunden sind. Vier der fünf $\beta\alpha$ -Einheiten entsprechen den vier Einheiten im $(\beta\alpha)_4$ -barrel, ein $\beta\alpha$ -Modul entspricht zwei kurzen zusätzlichen β -Strängen in HisF, welche in den $\beta\alpha$ -loops 1 und 5 liegen und ein kurzes antiparalleles β -Faltblatt bilden (Lang *et al.*, 2000). Die Superpositionierung der Strukturen von HisF_C mit der Flavodoxin-ähnlichen Domäne der Methylmalonyl-CoA-Mutase (MMCoA) aus *Propionibacterium shermanii* (PDB-Eintrag: 1req) ergab für die C α -Atome einen *rmsd*-Wert von nur 1,60 Å. Es spricht somit einiges dafür, dass $(\beta\alpha)_4$ -Halbbarrel und die Flavodoxin-ähnliche $(\beta\alpha)_5$ -Faltung einen gemeinsamen Ursprung besitzen. Es wird angenommen, dass die Abschirmung des hydrophoben Kerns vom Lösungsmittel während der Evolution entweder durch Dimerisierung zu einem $(\beta\alpha)_8$ -

barrel, wie bei den HisF-Hälften beobachtet, oder durch Einfügen eines fünften $\beta\alpha$ -Moduls, wie bei der Flavodoxin-ähnlichen $(\beta\alpha)_5$ -Faltung, erreicht wurde (Höcker *et al.*, 2002).

Mittlerweile wurden in verschiedenen Proteinen Hinweise auf $(\beta\alpha)_4$ -Halb*barrel* Subdomänen gefunden. Beispielsweise weisen die C-terminalen Hälften der $(\beta\alpha)_8$ -*barrel* Domänen von eukaryotischen und prokaryotischen Phosphoinositid-spezifischen Phospholipasen C (PI-PLC) keine Homologien auf, während ihre N-terminalen $(\beta\alpha)_4$ -Hälften, die alle katalytischen Reste enthalten, mit einem *rmsd*-Wert von nur 1,85 Å für 104 äquivalente C α -Atome superpositionieren (Heinz *et al.*, 1998). Bei den Mitgliedern der (AdoMet)-Radikal Protein-Superfamilie treten die drei Faltungstypen $(\beta\alpha)_4$, $(\beta\alpha)_6$ und $(\beta\alpha)_8$ auf. Basierend auf diesen Ergebnissen wird ein Evolutionsweg dieser Superfamilie aus einem $(\beta\alpha)_4$ -Halb*barrel* über ein $(\beta\alpha)_6$ -Dreiviertel*barrel* zu einem vollständigen $(\beta\alpha)_8$ -*barrel* vermutet (Nicolet & Drennan, 2004). Eine Studie, in der die beiden Hälften von TrpF aus *E. coli*, einem weiteren Vertreter mit $(\beta\alpha)_8$ -*barrel* Faltung, untersucht wurden, führte zu dem Ergebnis, dass die beiden Module nach Koexpression *in vivo* zur Komplementation eines *trpF*-defizienten Stammes fähig waren (Soberon *et al.*, 2004). Dieser Befund stützt die Hypothese der Evolution der $(\beta\alpha)_8$ -Faltung ausgehend von $(\beta\alpha)_4$ -Vorläufern ebenso wie die Tatsache, dass sich die beiden Hälften von TIM aus *Gallus gallus* nach gemeinsamer Rückfaltung zu einem katalytisch aktiven Protein zusammenlagern (Bertolaet & Knowles, 1995). Weitere Hinweise auf den modularen Aufbau von $(\beta\alpha)_8$ -Proteinen lieferten Untersuchungen zur Faltung und Entfaltung von TIM aus dem Hasenmuskel (Pan *et al.*, 2004).

Tm0979, ein Protein unbekannter Funktion aus *Thermotoga maritima*, besteht laut NMR-Strukturanalyse aus vier $\beta\alpha$ -Einheiten und weist an seinem viersträngigen parallelen β -Faltblatt eine hydrophobe Oberfläche auf, wodurch eine Homodimerbildung vermittelt wird (Gaspar *et al.*, 2005). Deshalb wird vermutet, dass es sich bei Tm0979, das sogar geringe Ähnlichkeit zu HisF aus *Thermotoga maritima* aufweist (72 Reste superpositionieren mit einem *rmsd*-Wert von 3,5 Å), um einen Nachkommen eines Halb*barrels* handeln könnte, von dem auch $(\beta\alpha)_8$ -*barrel* Proteine und Proteine der Flavodoxin-ähnlichen Faltung abstammen. Ob Tm0979 allerdings tatsächlich eine vollständige $(\beta\alpha)_4$ -Struktur aufweist, ist noch nicht vollständig geklärt, da andere NMR-Studien auf eine $\alpha(\beta\alpha)_3$ -Struktur hindeuten (Peti *et al.*, 2005).

1.4.2 Vierfache Symmetrie innerhalb der $(\beta\alpha)_8$ -barrel Faltung

Betrachtet man die Supersekundärstruktur eines $(\beta\alpha)_8$ -barrels, erkennt man eine so genannte pseudo-achtfache Symmetrie, die durch die acht $\beta\alpha$ -Module gegeben ist (vgl. Abbildung 5A). Auf der Ebene der Tertiärstruktur ergibt sich ein anderes Bild, das in Abbildung 8 am Beispiel von HisF gezeigt ist.

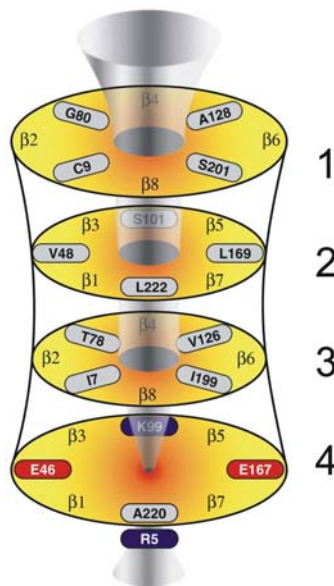


Abbildung 8: Schematische Darstellung der vier Schichten im zentralen β -Faltblatt von HisF.

Bei Ansicht des $(\beta\alpha)_8$ -barrels entlang der Längsachse werden vier Schichten (1, 2, 3, 4) sichtbar, die durch Aminosäureseitenketten der β -Stränge im Zentrum ausgebildet werden. Dabei ragen in *layer* 1 und 3 die Aminosäurereste der ungeradzahligen β -Stränge, in *layer* 2 und 4 die der geradzahligen β -Stränge ins Innere des barrels. Die entsprechenden Schichten sind in gelb und die Aminosäuren, aus denen sie aufgebaut sind im Einbuchstabencode mit der entsprechenden Positionsnummer dargestellt. In Schicht 4 wird durch die Reste ein Salzbrücken-Cluster ausgebildet, dazu ersetzt die positiv geladene Seitenkette von Arginin 5 (R5) aus β -Strang 1 den unpolaren aliphatischen Rest Alanin 220 (A220) aus β -Strang 8. (Nach Douangamath *et al.*, (2002))

Aminosäureseitenketten der β -Stränge, die das zentrale Fass bilden, ragen in den Innenraum des Fasses und bilden vier Schichten (*layer*) aus, die senkrecht zur Achse des barrels orientiert sind. *Layer* 1 und 3 werden von Aminosäureseitenketten der β -Stränge 1, 3, 5 und 7 gebildet, *layer* 2 und 4 von Seitenketten der β -Stränge 2, 4, 6 und 8. In *layer* 4 bilden die Reste ein Salzbrückencluster aus und anstelle des aliphatischen Alanin (A) 220 aus β 8 tritt Arginin (R) 5 aus β 1 in das Cluster ein. Insgesamt ergibt sich somit für das $(\beta\alpha)_8$ -barrel eine vierfache Symmetrie aus $(\beta\alpha)_2$ -Einheiten (Wierenga, 2001).

Da die Ausbildung dieser *layer* nur durch Oligomerisierung von $(\beta\alpha)_2$ -Einheiten ($N = 1, 2$) zustande kommen kann, liegt die Vermutung nahe, dass $(\beta\alpha)_8$ -Proteine aus $(\beta\alpha)_2$ -Vorläufereinheiten entstanden sein können. Entsprechend steht $N = 2$ für die bisher untersuchten Hälften, HisA_N, HisF_N, HisA_C und HisF_C.

Die vierfache Symmetrie von $(\beta\alpha)_8$ -barrel Enzymen wird durch die Beobachtung gestützt, dass 20% aller *loops* vor geradzahligen β -Strängen, d.h. am Stabilitätspol,

ein konserviertes GXD-Motiv enthalten (Nagano *et al.*, 2002). Dies trifft auch für die entsprechenden *loops* innerhalb der vier $(\beta\alpha)_2$ -Module von HisF (HisF_N1, HisF_N2 und HisF_C1 und HisF_C2) zu (Abbildung 9).

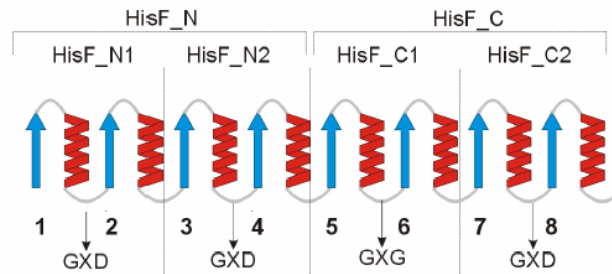


Abbildung 9: Schematische Unterteilung des $(\beta\alpha)_8$ -barrels von HisF in $(\beta\alpha)_2$ -Module

Das $(\beta\alpha)_8$ -barrel von HisF kann in zwei $(\beta\alpha)_4$ -Module (HisF_N und HisF_C) unterteilt werden, die sich wiederum aus jeweils zwei $(\beta\alpha)_2$ -Modulen (HisF_N1, HisF_N2 und HisF_C1, HisF_C2) zusammensetzen. Die $\alpha\beta$ -loops innerhalb der $(\beta\alpha)_2$ -Module weisen ein charakteristisches GXD- bzw. GXG-Motiv auf (Nagano *et al.*, 2002).

Fani *et al.* (2007) konnten mit Hilfe von Sequenzalignments allen bekannten HisA-Proteinsequenzen eine vierfache Symmetrie auf Primärstrukturebene zuweisen, wobei Sequenzähnlichkeiten von 13 – 58 % zwischen den HisA-Vierteln ermittelt wurden. Wenn man eine evolutionäre Verwandtschaft von HisA und HisF annimmt (Thoma *et al.*, 1998; Lang *et al.*, 2000; Henn-Sax *et al.*, 2001), sollte also auch HisF eine vierfache Symmetrie auf Sequenzebene besitzen.

Einen Hinweis auf die vierfache Sequenzsymmetrie lieferte eine Analyse mit dem Programm HHRep (Söding *et al.*, 2006), bei welcher die eingegebene Sequenz in ein Hidden-Markow-Modell (HMM) umgewandelt und paarweise mit sich selbst verglichen wird. Als Punktdiagramm (*Dot-plot*) angezeigt, lässt das Ergebnis als Erstes eine Diagonale erkennen, die dem Selbstvergleich entspricht (Abbildung 10).

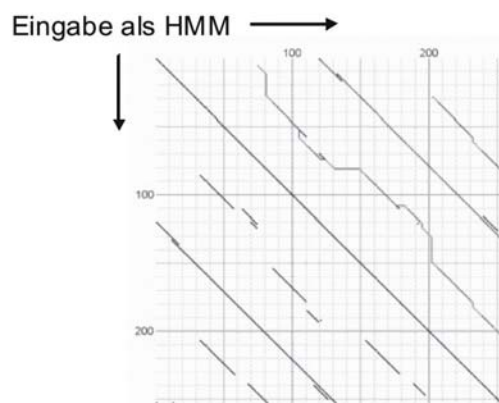


Abbildung 10: Dot-plot der Sequenzanalyse von HisF mit dem Programm HHrep (Nach Söding *et al.*, 2006).

Die Eingabesequenz wird zunächst in ein Hidden-Markow-Modell umgewandelt. Anschließend wird die Eingabe paarweise mit sich selbst verglichen. Der *Dot-plot* zeigt das Ergebnis an. Die Diagonale in der Mitte entspricht dem Selbstvergleich, die parallel verschobenen Diagonalen weisen weitere ähnliche Bereiche innerhalb der Eingabe aus. Für HisF ergibt sich demnach eine vierfache Symmetrie innerhalb der Sequenz.

Weitere Sequenzbereiche, die identisch bzw. ähnlich sind, werden als parallel verschobene Diagonalen angezeigt. Wie in Abbildung 10 ersichtlich, ergeben sich für HisF vier Diagonalen, was für die vierfache Ähnlichkeit bzw. Symmetrie der $(\beta\alpha)_2$ -Module spricht.

1.5 Ziel der Arbeit

Wie oben beschrieben, scheint die $(\beta\alpha)_8$ -barrel Faltung eine vierfache Symmetrie zu besitzen, d.h. sie besteht aus vier $(\beta\alpha)_2$ -Modulen. Legt man – wie in Abschnitt 1.2 – die Evolution von Proteindomänen aus sich wiederholenden Einheiten bzw. Supersekundärstrukturelementen zugrunde, dann ergeben sich mehrere Fragen:

1. Ist das $(\beta\alpha)_2$ -Modul die kleinste stabile Einheit, aus der sich HisF bzw. alle $(\beta\alpha)_8$ -barrel Proteine entwickelt haben?
2. Bildet ein $(\beta\alpha)_2$ -Modul ein stabiles Oligomer mit annähernd nativer Sekundär- und Tertiärstruktur, verglichen mit einem $(\beta\alpha)_8$ -barrel?

Um diese Fragen zu beantworten sollten die einzelnen Viertel von HisF (HisF_N1, HisF_N2, HisF_C1 und HisF_C2; Abbildung 9) heterolog in *E. coli* exprimiert, gereinigt und biophysikalisch charakterisiert werden. Diese Analysen sollen Aufschluss über Stabilität, Oligomerisierungsgrad und Struktur der $(\beta\alpha)_2$ -Module geben und mit den entsprechenden Kenngrößen des wildtypischen HisF bzw. der Hälften (HisF_N und HisF_C; Abbildung 9) verglichen werden.

3. Lassen sich die Module durch einzelne Aminosäureaustausche weiter stabilisieren und lässt sich evtl. sogar enzymatische Aktivität bei den Homooligomeren detektieren bzw. etablieren?
4. Welches der vier HisF-Viertel stellt möglicherweise das ursprünglichste der Module dar und kann somit als Vorläufereinheit identifiziert werden?

2 Material

2.1 Geräte

Absorptionsphotometer: Zweistrahl-UV/VIS-Spektralphotometer Cary Bio 100, Varian (Darmstadt); Ultrospec 10 cell density meter, Amersham Biosciences (Cambridge, UK); Biophotometer, Eppendorf (Hamburg)

Agarose-Gelelektrophorese-Apparatur: hauseigene Werkstatt

ÄKTAbasic™ 10 better: GE-Healthcare (Freiburg)

Mit folgender analytischer Säule:

Superdex 75 HR 10/300: GE Healthcare (Freiburg)

ÄKTAprime™ plus mit Rekorder Rec 102: GE Healthcare (Freiburg)

Mit folgender Säule:

HisTrap™ FF crude 1 ml: GE Healthcare (Freiburg)

Blotkammer: Hoefer SemiPhor semi-dry transfer unit Modell TE70, Hoefer Pharmacia Biotech Inc. (CA, USA)

Brutschrank: BD 115, Binder (Tuttlingen)

Circular Dichroismus-Spektralpharimeter: Jasco J-815 mit Peltier-Einheit PTC-423S/15, Jasco (Groß-Umstadt)

Dampfsterilisator: Autoklav München 25, MMM (München)

Digitalisierboard: Summa Sketch III Professional, GTCO CalComp GmbH (München)

Elektrophorese-Zubehör: Hoefer Mighty Small II, Hoefer Pharmacia Biotech (CA, USA) und Multi Gel Caster Gelgießstand, Amersham Biosciences (Cambridge, UK)

Filtrationsanlage: Filtereinheit Millex FG13, Ultrafree-20 Nanopore-Wasser-Anlage, Millipore (Schwalbach)

Fluoreszenzspektrophotometer: Fluoreszenz-Spektrometer Cary Eclipse, Varian (Darmstadt)

FPLC-System LCC-501 Plus, Fraktionssammler Frac 100 und Rekorder Rec 112: Amersham Pharmacia Biotech. (Freiburg)

Mit folgender präparativer Säule:

HiLoad 26/60 Superdex 75 prep grade: GE Healthcare (Freiburg)

Gasbrenner: Gasprofi 2SCS, Wild Tec GmbH (Göttingen)

Gefrierschrank, - 20 °C: Liebherr Premium, Liebherr (Nussbaumen)

Gefrierschrank, - 80 °C: MDF-K72V, Sanyo (Tokio, Japan)

Geldokumentationsanlage: Multi Doc-It Digital Imaging System, UVP Inc. (Upland, USA)

Glaswaren und -pipetten: Schott (Mainz), Fischer Scientific (Schwerte)

Heizblöcke: HBT-2 131, HLC (Bovenden); Thermomixer 5436, Eppendorf (Hamburg)

Magnetrührer (beheizbar): MR 3001, Heidolph (Kelheim)

Mikroliterpipetten: Research 0,5 – 1000 µl, Eppendorf (Hamburg)

Mikrowelle: HMT 842C, Bosch (Nürnberg)

Netzgerät: EPS 301, GE Healthcare (München)

PCR-Geräte: Mastercycler personal und Mastercycler epgradient S, Eppendorf (Hamburg)

Peristaltikpumpe: Miniplus 2, Gilson (Villiers le Bel, Frankreich)

pH-Meter: Inolab Level 1, Inolab (Weilheim)

Pipettus-accu: Hirschmann Laborgeräte (Eberstadt)

Quartzküvetten: 1 - 10 mm Schichtdicke, Hellma (Mülheim) und Starna (Pfungstadt)

Säulenmaterial zur Affinitätschromatographie mittels *gravity-flow*: Chitin *beads*, New England Biolabs GmbH (Frankfurt am Main)

Schüttelinkubatoren: Certomat H, S II und BS-1, Braun Biotech. Int. (Melsungen)

Schütteltische: GFL 3011 und 3013, GFL (Burgwedel)

Ultraschallgerät: Branson Sonifier 250 D mit ½“ Resonator und 1/8“ Mikrospritze, Heinemann (Schwäbisch Gmünd)

UV-Leuchttisch: Reprostar, Camag Chemie Erzeugnisse (Berlin)

Vakuumkonzentrator: Concentrator 5301, Eppendorf (Hamburg)

Vakuumpumpe: ME 2C, Vacuubrand (Wertheim)

Vortex Genie 2, Scientific Ind. (Bohemia, USA)

Waagen: MC1, Sartorius (Göttingen); H54AR, H20T, P1200 und P3000, Mettler (Gießen)

Wasserbad: Wasserbad 1012, GFL (Burgwedel)

Zentrifugen: Centrifuge 5415 D, Centrifuge 5415 R und Centrifuge 5810 R, Eppendorf (Hamburg); Sorvall RC-5C plus, Sorvall RC-5B und Sorvall RC-2B mit den Rotortypen SS34 und GS3, Du Pont Instruments (Bad Homburg)

2.2 *Chemikalien*

Alle für molekularbiologische Arbeiten notwendigen Chemikalien besaßen den höchstmöglichen Reinheitsgrad. Wenn nicht anders angegeben wurden die Chemikalien von Becton Dickinson GmbH (Heidelberg), Bio-Rad Laboratories (München), Biozym Scientific GmbH (Hess), Carl Roth GmbH & Co. KG (Karlsruhe), GE Healthcare (Freiburg), Gerbu Biochemicals GmbH (Gaiberg), KMF Laborchemie Handels GmbH (Lohmar), KPL (Gaithersburg, MD, USA), Merck KGaA (Darmstadt), MP Biomedicals (Eschwege), Pierce (Rockford, IL, USA), Roche Diagnostics GmbH (Mannheim), Serva (Heidelberg) oder Sigma Aldrich (Steinheim) bezogen.

2.3 *Kits*

D-Tube™ Dialyzer Midi MWCO 3,5: Calbiochem (San Diego, CA, USA)

D-Tube™ Electroelution Accessory Kit: Calbiochem (San Diego, CA, USA)

IMPACT™-CN: New England Biolabs GmbH (Frankfurt am Main)

LMW Gel Filtration Calibration Kit: GE Healthcare (Freiburg)

NucleoBond AX DNA-Extraction Kit: Macherey Nagel (Düren)

NucleoSpin Extract II: Macherey Nagel (Düren)

QIAprep Spin Miniprep Kit: QIAGEN (Hilden)

QIAquick Gel Extraction Kit: QIAGEN (Hilden)

QIAquick PCR-Purification Kit: QIAGEN (Hilden)

2.4 *Enzyme*

Faktor Xa-Protease (Restriction Grade Factor Xa Kit): Novagen (Darmstadt)

GoTaq™ DNA-Polymerase (Green GoTaq™ Reaction Buffer): Promega (Mannheim)

Lysozym (aus Hühnerei): Boehringer-Ingelheim (Mannheim)

Restriktionsendonukleasen: New England Biolabs GmbH (Frankfurt am Main)

Taq-Polymerase (Ready-Mix RED *Taq* PCR Reaction Mix): Sigma-Aldrich (Steinheim)

T4-Ligase: MBI Fermentas GmbH (St. Leon-Rot)

2.5 Verbrauchsmaterialien

Blot-Membran: Immobilon-P Transfer Membran, Millipore (Schwalbach)

Dialyseschläuche: Carl Roth GmbH & Co. KG (Karlsruhe), Pierce (Rockford, IL, USA) oder Spectrum Laboratories Inc. (Gardena, CA, USA)

Einwegspritzen: Omnifix® 2, 5 und 60 ml, Braun Biotech. (Melsungen)

Filterpapier: Whatman (Maidstone, UK)

Konzentratoren: Amicon Ultra 4, 15 (MWCO: 30, 10 oder 5 kDa) Centrifugal Filter Devices, Millipore (Schwalbach)

Kunststoffküvetten: ½ Mikroküvette, Ratiolab (Dreieich-Buchschlag)

Kunststoffröhrchen: Plastikröhrchen 15, 50 ml mit Schraubdeckel, Sarstedt (Nümbrecht)

Latex-Handschuhe: Safeskin, Kimberly-Clark (Roswell, USA)

Membranfilter: ME24, 0,2 µm Porengröße, Ø 47 mm, Schleicher & Schuell (Dassel)

NAP™-Säulen: NAP™-5, 10 und 25 Säulen, Amersham Biosciences (Mannheim)

Nitril-Handschuhe: Safeskin, Kimberly-Clark (Roswell, USA)

Parafilm „M“ Laboratory Film: Pechiney (Menasha, WI, USA)

Pasteurpipetten: Hirschmann (Ebermannstadt)

PCR-Gefäße: 0,2 ml, PeqLab (Erlangen)

Petrischalen: 94/16, Greiner bio-one (Nürtingen)

Pipettenspitzen: 10, 200 und 1000 µl, Sarstedt (Nümbrecht)

Reaktionsgefäße: 1,5 bzw. 2 ml, Sarstedt (Nümbrecht)

Spritzenfilter: 0,22 µm Porengröße, Renner GmbH (Daunstadt)

2.6 Bakterienstämme

***Escherichia coli* DH5 α** (Hanahan, 1983):

[F⁻, *endA1*, *hsdR17* (*r_k⁻m_k⁻*), *supE44*, *thi1*, *recA1*, *gyrA* (Nal^r), *relA1*, Δ (*lacZYAargF*) U169, Φ 80/*lacZ* Δ M15]

***E. coli* BL21 (DE3)** (Studier & Moffatt, 1986):

hsdS, *gal* [λ cl, *ts857*, *cnd1*, *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*]

***E. coli* BL21-CodonPlus (DE3)-RIPL** Stratagene (La Jolla, CA, USA):

BF⁻ *ompT* *hsdS*(*r_B⁻ m_B⁻*) *dcm*⁺ Tet^r *gal* λ (DE3) *endA* Hte [*argU proL Cam*^r] [*argU ileY leuW* Strep/Spec^r]

BL21 CodonPlus (DE3) RIPL enthält ein zusätzliches Plasmid mit Genen, welche für in *E. coli* seltene tRNAs kodieren (*argU* (AGA, AGG); *proL* (CCC); *ileY* (AUA); *leuW* (CUA)). Zusätzlich trägt das Plasmid das Gen für die Chloramphenicolacetyltransferase, die Resistenz gegen Chloramphenicol vermittelt.

***E. coli* RosettaTM (DE3)** Novagen (Darmstadt):

F⁻ *ompT* *hsdS_B*(*r_B⁻ m_B⁻*) *gal dcm* (DE3) pRARE (Cam^R)

RosettaTM (DE3) unterscheidet sich von BL21-CodonPlus (DE3)-RIPL durch ein zusätzliches Gen, das für die tRNA *glyT* (GGA) kodiert.

***E. coli* ER2566** New England Biolabs GmbH (Frankfurt am Main):

fhuA2 lacZ::T7 gene1 [*lon*] *ompT gal sulA11* R(*mcr-73::miniTn10--Tet^S*)² [*dcm*] R(*zgb-210::Tn10--Tet^S*) *endA1* Δ (*mcrC-mrr*)114::IS10

2.7 Vektoren

2.7.1 pET-Vektoren

Gene, die in die multiple Klonierungsstelle („multiple cloning site“, MCS) von pET-Vektoren (Novagene (Darmstadt), „*plasmid for expression by T7 RNA polymerase*“) kloniert wurden, werden von der DNA-Polymerase des Phagen T7 abgelesen (Studier & Moffatt, 1986). pET-Vektoren können daher nur in Bakterienstämmen abgelesen werden, die die entsprechende T7 Polymerase exprimieren (z.B. BL21(DE3)). Die Expression des Polymerasegens und des in das Plasmid klonierten Gens erfolgt unter Kontrolle eines *lacUV5*-Promotor-Operators und wird durch die Zugabe von Isopropyl- β -D-Thiogalaktopyranosid (IPTG) induziert. Der entsprechende *lac*-Repressor (*LacI*) ist auf dem Plasmid kodiert und wird konstitutiv exprimiert (Abbildung 11).

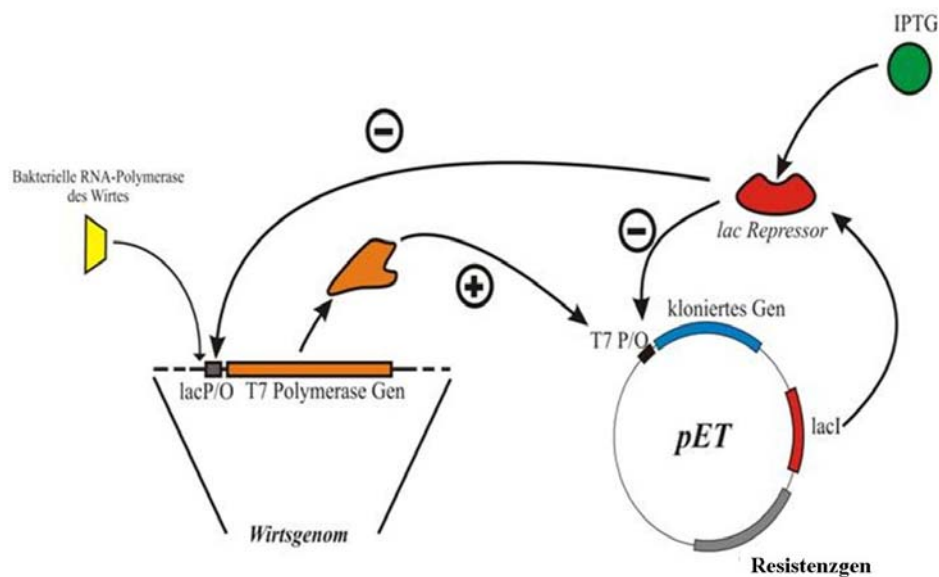


Abbildung 11: Schematische Darstellung des pET-Expressionssystems.

2.7.1.1 pET21a(+)

Der Vektor pET21a(+) enthält stromabwärts der multiplen Klonierungsstelle sechs für Histidin codierende Triplets (Abbildung 12). Diese führen nach der Genexpression zu einem Protein, das C-terminal einen Hexahistidin-*Tag*[®] (His₆-*Tag*) trägt, sofern das Stopcodon des eingefügten Gens bei der Klonierung deletiert wurde. Dadurch wird eine Reinigung mittels Metallchelataffinitätschromatographie möglich. Zur Selektion enthält pET21a(+) das offene Leseraster *bla*, das für die β -Lactamase codiert, die den transformierten Zellen Resistenz gegen das Antibiotikum Ampicillin verleiht.

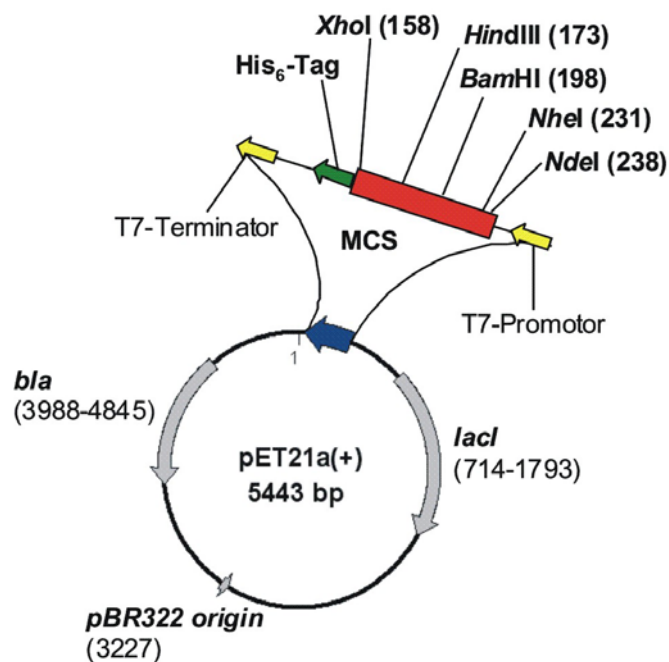


Abbildung 12: Schematische Darstellung des Vektors pET21a(+) mit den in dieser Arbeit genutzten Restriktionsenzymstichstellen in der *Multiple Cloning Site* (MCS).

2.7.1.2 pET24a(+)

Der pET24a(+)-Vektor entspricht im Aufbau dem Vektor pET21a(+) (Abbildung 12). Er enthält jedoch anstelle des β -Lactamase-Gens (*bla*) das für die Kanamycinnukleotidyltransferase (*kan*) kodierende Gen, wodurch den transformierten Zellen Kanamycin-Resistenz vermittelt wird.

2.7.1.3 pET28a(+)

Im Vektor pET28a(+) befindet sich stromaufwärts der MCS ein His₆-Tag gefolgt von einer Thrombinschnittstelle (Abbildung 13). Um den His₆-Tag nutzen zu können wurde das jeweilige Insert 5'-seitig über eine *Nde*I Schnittstelle in den entsprechenden Leserahmen kloniert. 3'-seitig wurde zur Klonierung z.B. die *Bam*HI-Schnittstelle verwendet und dabei ein Stopcodon eingefügt, wodurch der stromabwärts der MCS liegende His₆-Tag nicht translatiert wurde. Das Plasmid trägt als Selektionsmarker das Kanamycinresistenzgen.

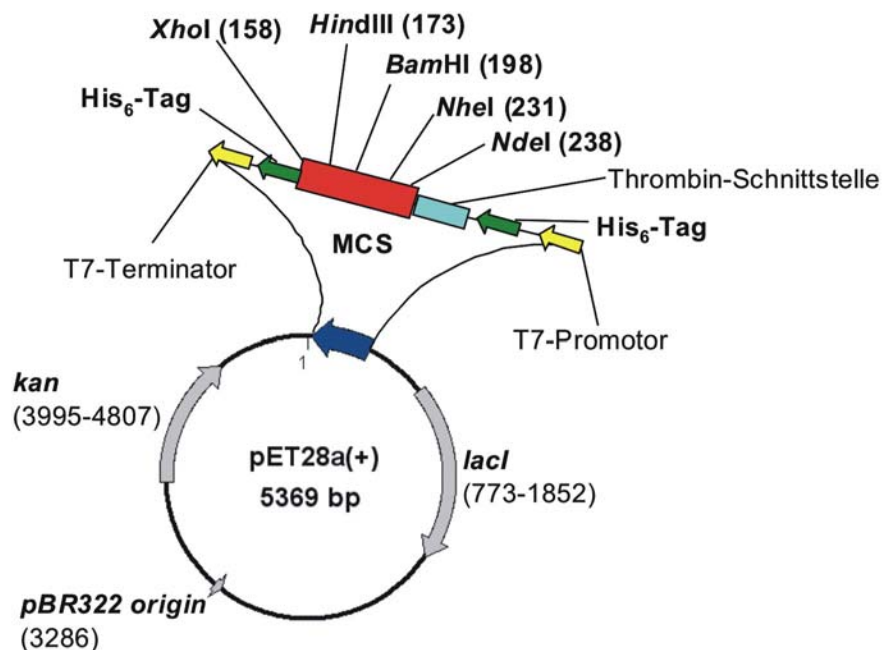


Abbildung 13: Schematische Darstellung des Vektors pET28a(+) mit den in dieser Arbeit genutzten Restriktionsenzymschnittstellen in der *Multiple Cloning Site* (MCS).

2.7.2 pTYB-Vektoren

pTYB-Vektoren wurden von New England Biolabs GmbH (Frankfurt/Main) bezogen und benutzen wie die pET-Vektoren (2.1.2.1) den T7/lac Promotor, um eine stringente Kontrolle der Fusionsgenexpression zu gewährleisten. Somit kann auch hier die Zielproteinexpression durch IPTG induziert werden. Die Expression erfolgt in dem *E. coli*-Stamm ER2566 (2.6), einem Derivat des BL21-Stammes.

2.7.2.1 pTYB1

pTYB1 (Abbildung 14) trägt unmittelbar stromabwärts der MCS das offene Leseraster für das Intein aus *Saccharomyces cerevisiae*. Dies ist gefolgt von dem Gen für die Chitinbindedomäne (CBD). Durch *SapI*-Klonierung des Zielgens am 3'-Ende lässt sich ein Fusionsprotein herstellen, das C-terminal mit Intein/CBD fusioniert ist, was die Reinigung des Zielproteins mittels einer Chitinaffinitätschromatographie erlaubt. Im Rahmen dieser Arbeit wurde jedoch 3'-seitig mit *XhoI* kloniert, wodurch C-terminal am Zielprotein nach Abtrennung des Intein/CBD-Proteins drei Aminosäurereste (Gly, Ser, Ser) verbleiben.

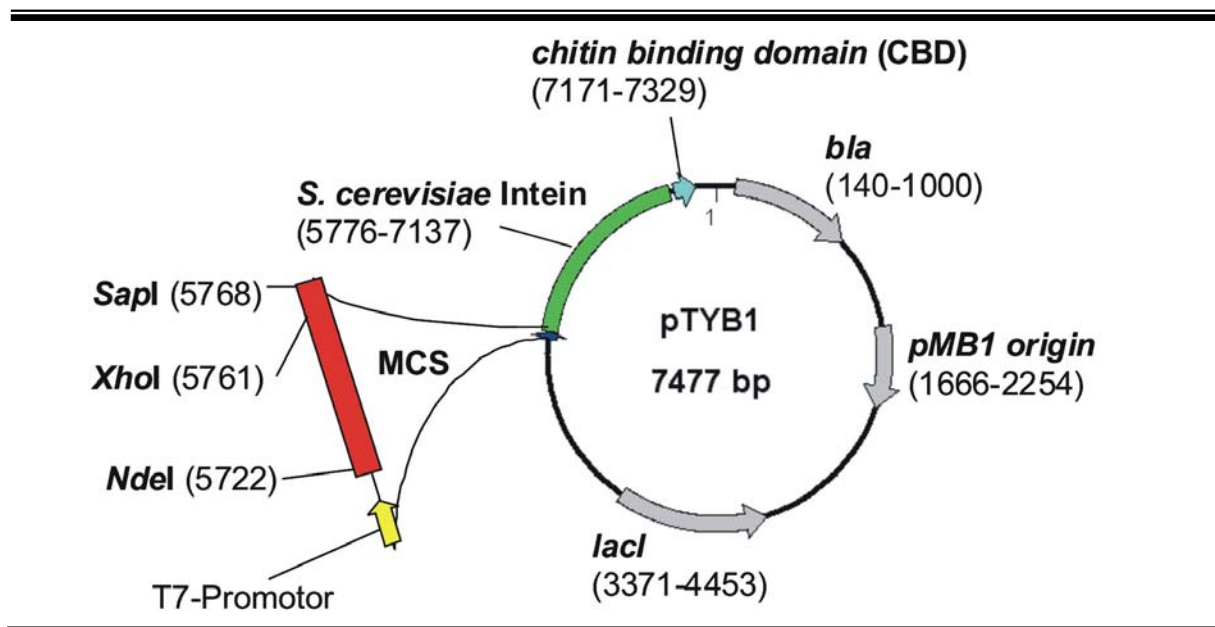


Abbildung 14: Schematische Darstellung des Vektors pTYB1 mit den in dieser Arbeit genutzten Restriktionsenzymschnittstellen in der *Multiple Cloning Site* (MCS).

2.8 Oligodesoxyribonukleotide

Alle aufgeführten Oligodesoxyribonukleotide wurden von der Firma Metabion (Martinsried) bezogen.

2.8.1 Primer zur Klonierung der HisF-Viertel

<i>hisF_N1_5'NdeI</i>	5'-AGC CAT ATG CTC GCT AAA AGA ATA ATC GCG-3'
<i>hisF_N1_3'BamHI</i>	5'-ATA GGA TCC TCA GTC GAT CTG CTC GGC CAC-3'
<i>hisF_C1_5'NdeI</i>	5'-ATA CAT ATG CAG GCC GTT GTC GTG GTG GCG ATA-3'
<i>hisF_C1_3'BamHI</i>	5'-ATA GGA TCC TCA AAG TGT GGT TAG TGG-3'
<i>hisF_N2_5'NdeI</i>	5'-CGC CAT ATG ATT CCG TTC ACT GTT GGA GGA-3'
<i>hisF_N2_3'intein</i>	5'-TTC GGA TCC GCT CTT CCG CAA CTC CCA A-3'
<i>hisF_C2_5'NdeI</i>	5'-ATA CAT ATG CTT CCC ATC ATC GCT TCC GGT GGT-3'
<i>hisF_C2_3'intein</i>	5'-TTC GGA TCC GCT CTT CCG CAC AAC CCC T-3'

2.8.2 Primer zur Klonierung einer Faktor Xa-Schnittstelle zwischen HisF_N1 und HisF_N2 bzw. HisF_C1 und HisF_C2

<i>hisF_N1XaN2.for</i>	5'-AGA TCG ACA TTG AGG GCC GCA TTC CGT T-3'
<i>hisF_N1XaN2.rev</i>	5'-AAC GGA ATG CGG CCC TCA ATG TCG ATC T-3'
<i>hisF_C1XaC2.for</i>	5'-TAA CCA CAA TCG AGG GTC GCC TTC CCA T-3'
<i>hisF_C1XaC2.rev</i>	5'-ATG GGA AGG CGA CCC TCG ATT GTG GTT A-3'

2.8.3 Primer zur Klonierung von Varianten der HisF-Viertel

5'N1_R5A	5'-CAT ATG CTC GCT AAA GCG ATA ATC GCG TGT CTC-3'
5'N1_E46Q	5'-GAA ATT GGA ATA GAC CAG CTC GTT TTT CTG GAT-3'
3'N1_E46Q	5'-ATC CAG AAA AAC GAG CTG GTC TAT TCC AAT TTC-3'
5'N1_F23W	5'-GTG AAG GGA ACG AAC TGG GAA AAC CTC AGG GAC-3'
3'N1_F23W	5'-GTC CCT GAG GTT TTC CCA GTT CGT TCC CTT CAC-3'
5'N2_F86W	5'-ATC CAC GAC TGG GAA ACG GCC-3'
3'N2_F86W	5'-GGC CGT TTC CCA GTC GTG GAT-3'
5'C1_A128C_NdeI	5'-ATT CAT ATG CAG GCC GTT GTC GTG TGT ATA GAT GCA-3'
5'C1_Y143H	5'-TTC ATG GTC TTC ACC CAT TCC GGA AAG AAG AAC-3'
3'C1_Y143H	5'-GTT CTT CTT TCC GGA ATG GGT GAA GAC CAT GAA-3'
5'C1_E167Q	5'-GGA GCA GGA CAA ATC CTG CTC-3'

3'C1_E167Q	5'-GAG CAG GAT TTG TCC TGC TCC-3'
5'C2_F227W	5'-CTT GCG GCT TCT GTC TGG CAC TTC AGA GAG ATC-3'
3'C2_F227W	5'-GAT CTC TCT GAA GTG CCA GAC AGA AGC CGC AAG-3'

2.8.4 Primer zur Klonierung von HisF_N1N1 und HisF_C1C1

5'N1_R5A_NheI	5'-CTA GCT AGC GCT AAA GCG ATA ATC-3'
---------------	---------------------------------------

2.8.5 Primer zur Klonierung von HisF_N1abcd und HisF_C1abcd

<i>hisF_N1-1_5'_NdeI</i>	5'-GGA AGC CAT ATG CTC GCT AAA AGA ATA ATC GCG-3'
<i>hisF_N1-1_3'_NheI</i>	5'-GGA ATA GCT AGC CTC GGC CAC CTT TTC GAC CAG-3'
<i>hisF_N1-2_5'_NheI</i>	5'-AGC GCT AGC GCT AAA AGA ATA ATC GCG TGT-3'
<i>hisF_N1-2_3'_BamHI</i>	5'-ATA GGA TCC CTC GGC CAC CTT TTC GAC CAG-3'
<i>hisF_N1-3_5'_BamHI</i>	5'-AGC GGA TCC GCT AAA AGA ATA ATC GCG TGT-3'
<i>hisF_N1-3_3'_HindIII</i>	5'-ATA AGA AGC TTG CTC GGC CAC CTT TTC GAC CAG-3'
<i>hisF_N1-4_5'_HindIII</i>	5'-CGC AGC CAA GCT TCT AAA AGA ATA ATC GCG TGT-3'
<i>hisF_N1-4_3'_XhoI</i>	5'-ATA CTC GAG CTC GGC CAC CTT TTC GAC CAG-3'
5'C1a_NdeI	5'-ATT CCA TAT GGC CGT TGT CGT-3'
3'C1a_NheI	5'-CGT ACC TAG CTA GCT GTG GTT A-3'
5'C1b_NheI	5'-CTA GCT AGC GCC GTT GTC GT-3'
3'C1b_BamHI	5'-TAC GCG GAT CCT GTG GTT AGT-3'
5'C1c_BamHI	5'-CGC GGA TCC GCC GTT GTC GT-3'
3'C1c_HindIII	5'-GCC TCC CAA GCT TTG TGG TTA G-3'
5'C1d_HindIII	5'-CCC AAG CTT GCC GTT GTC GT-3'
3'C1d_XhoI	5'-ATT CCG CTC GAG TGT GGT TAG-3'

2.8.6 Vektorspezifische Sequenzierprimer

T7-Promotor	5'-TAA TAC GAC TCA CTA TAG GG-3'
T7-Terminator	5'-GCT AGT TAT TGC TCA GCG G-3'

2.9 Längenstandards

2.9.1 DNA-Längenstandards

Die DNA-Längenstandards *Smart Ladder* von Eurogentec (Köln) (Abbildung 15A) und *Gene Ruler™* 1kb DNA Ladder der Fermentas GmbH (St. Leon-Rot) (Abbildung 15B) dienen sowohl als Längenstandard zur Größenbestimmung von DNA-Fragmenten, als auch zur Abschätzung von DNA-Mengen im Agarosegel (3.2.8).

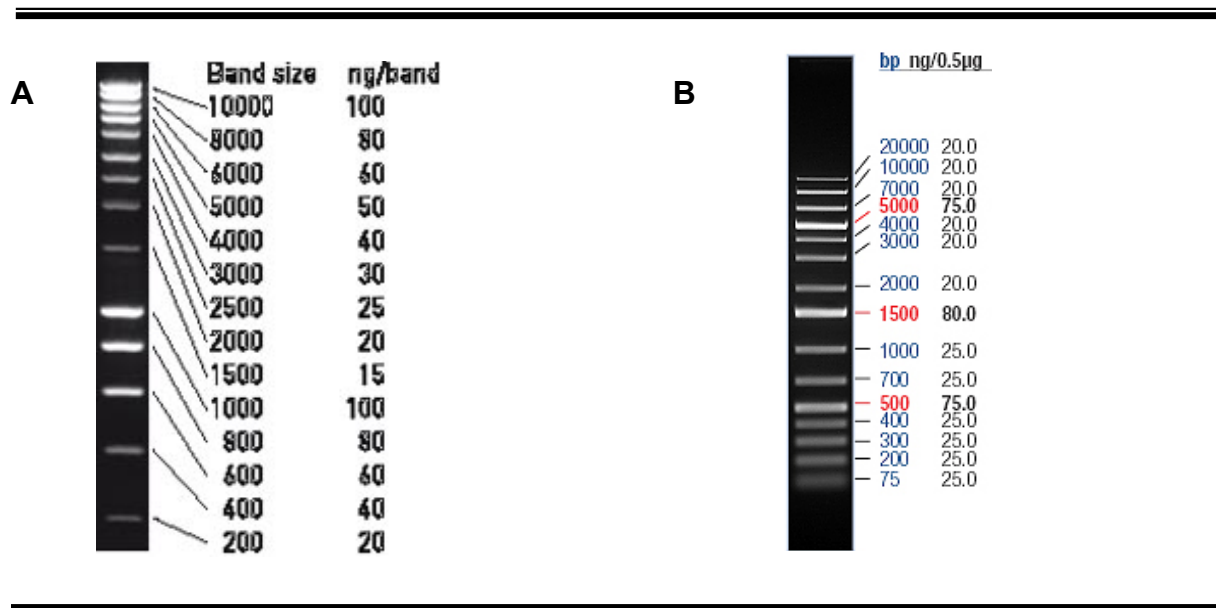


Abbildung 15: Die verwendeten DNA-Größenstandards.

(A) 1%iges Agarosegel mit 5 µl *Smart Ladder* von Eurogentec (Köln)

(B) 1%iges Agarosegel mit 0,5 µg *Gene Ruler™* von Fermentas (St. Leon-Rot)

Gezeigt sind die einzelnen Bandengrößen und die korrespondierenden DNA-Mengen.

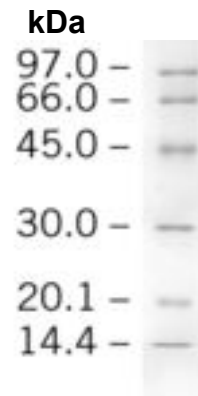
2.9.2 Protein-Längenstandards

Low Molecular Weight Marker: Amersham Biosciences (Mannheim)

Der *Low Molecular Weight Marker* (Tabelle 1; Abbildung 16) diene sowohl als Längenstandard zur Größenbestimmung von Proteinen, als auch zur Abschätzung der Proteinmengen in 12,5 bzw. 15 %igen SDS-Polyacrylamidgelen (3.4.1.1).

Tabelle 1: Der *Low Molecular Weight Marker*

Markerproteine	MW [Da]
Phosphorylase A	97000
BSA	66000
Ovalbumin	45000
Carboanhydratase	30000
Trypsin Inhibitor	20100
β -Lactalbumin	14400

Abbildung 16: 15%iges Polyacrylamidgel mit 5 μ l *Low Molecular Weight Marker*.

Angezeigt sind die Molekulargewichte der einzelnen Proteine in Kilo-Dalton (kDa).

Molecular Weight Marker, MW Range 2512-16949 Da, “small fragment marker”:

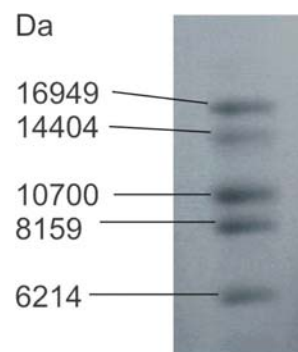
Amersham Biosciences (Mannheim)

Dieser Marker (Tabelle 2; Abbildung 17) wurde für die Größenbestimmung von besonders kleinen Proteinfragmenten verwendet, die auf Tris-Tricin-Gelen (3.4.1.2) mit 20% Polyacrylamid analysiert wurden.

Tabelle 2: Der „*small fragment marker*“

Markerproteine	MW [Da]
Globin	16949
Globin I+II	14404
Globin I+III	10700
Globin I	8159
Globin II	6214
Globin III *	2512

* Konnte auf den SDS-Gelen nicht identifiziert werden.

Abbildung 17: 20%iges Tris-Tricin-Gel mit 5 μ l *Small Fragment Marker*.

Angezeigt sind die Molekulargewichte der einzelnen Proteine in Dalton (Da).

2.10 *Nährmedien*

Alle Nährmedien wurden mit zweifach destilliertem Wasser hergestellt und zur Sterilisation 30 min bei 120 °C autoklaviert. Bei der Herstellung von Selektivmedien wurden die entsprechenden Antibiotika nach dem Autoklavieren in Form von tausendfach konzentrierten, sterilfiltrierten Stammlösungen zugegeben.

LB (Luria-Bertani (Sambrook, 1989))

1 % (w/v) Trypton

1% (w/v) NaCl

0,5 % (w/v) Hefeextrakt

SOB

2 % (w/v) Trypton

0,05 % NaCl

0,5 % Hefeextrakt

Nach dem Autoklavieren wurden dem Medium mit Hilfe von sterilen Stammlösungen 10 mM MgSO₄, 10 mM MgCl₂ und 2,5 mM KCl (Endkonzentrationen) zugesetzt.

Vollmedium mit Betain/Sorbitol (LB*)

1 % (w/v) Trypton

1% (w/v) NaCl

0,5 % (w/v) Hefeextrakt

2% (v/v) Glycerin

2,5 mM Betain

500 mM Sorbitol

LB-Agar

LB-Medium versetzt mit 1,5 % (w/v) Agar.

2.11 *Lösungen und Puffer*

Die Fertigung von Lösungen und Puffern, die nicht von Firmen bezogen wurden, ist im Folgenden beschrieben. Soweit nicht anders vermerkt, wurden wässrige Lösungen hergestellt und bei Raumtemperatur gelagert. Es wurde MILLIPORE-Wasser oder bidestilliertes Wasser verwendet.

1 % Agarose

1 % (w/v) Agarose wurde in 0,5 % (w/v) TBE aufgekocht und bei 60 °C gelagert.

Ampicillin-Stammlösung

Na⁺-Ampicillin-Salz wurde in Wasser in einer Endkonzentration von 150 mg/ml gelöst, sterilfiltriert und bei 4 °C gelagert.

Ammoniumperoxodisulfat (APS)-Stammlösung

APS wurde in Wasser in einer Endkonzentration 10 % (w/v) bzw. 250 mg/ml gelöst, sterilfiltriert und bei -20 °C gelagert.

Auffaltungspuffer

10 M Harnstoff gelöst in 50 mM Kaliumphosphat-Puffer pH 7,5

Aufschlusspuffer

50 mM Kaliumphosphat-Puffer, pH 7,5

Bradford-Reagenz

Chromatographie-Puffer

- Chitinaffinitätschromatographie

Säulenpuffer: 20 mM Tris/HCl, 500 mM NaCl, 1 mM EDTA, pH 8,0

Cleavage-Puffer: 20 mM Tris/HCl, 500 mM NaCl, 1 mM EDTA, 50 mM DTT,
pH 8,0

Regenerationspuffer: 0,3 M NaOH

- Größenausschlusschromatographie

50 mM Kaliumphosphat-Puffer, 300 mM KCl, pH 7,5

- **Rückfaltung auf der HisTrap™-Säule**

Bindepuffer: 20 mM Tris/HCl, 6 M GdmCl, 0,5 M NaCl, 5 mM Imidazol, pH 8,0

Waschpuffer: 20 mM Tris/HCl, 6 M Urea, 0,5 M NaCl, 5 mM Imidazol, pH 8,0

Rückfaltungspuffer: 20 mM Tris/HCl, 0,5 M NaCl, 5 mM Imidazol, pH 8,0

Elutionspuffer: 20 mM Tris/HCl, 0,5 M NaCl, 1 M Imidazol, pH 8,0

Coomassie-Färbelösung

50 % (v/v) Ethanol (96 %ig), 10 % (v/v) Essigsäure (100 %ig), 0,2 % (w/v) Coomassie Brilliantblau G250 wurden in Wasser gelöst, durch einen Papierfilter filtriert und lichtgeschützt bei Raumtemperatur gelagert.

Ethidiumbromid-Stammlösung

10 mg/ml Ethidiumbromid

Faktor Xa-Salzlösung (10x)

1 M NaCl, 10 mM CaCl₂

Glycerin (87%)

Isopropyl-β-D-thiogalaktopyranosid (IPTG)-Stammlösung

1 M IPTG wurden in Wasser gelöst, sterilfiltriert und bei -20 °C gelagert.

Kaliumphosphat-Puffer (KP)

1 M K₂HPO₄ und 1 M KH₂PO₄ wurden in Wasser gelöst und durch verschiedene Mischungsverhältnisse auf den jeweils gewünschten pH-Wert gebracht.

Kanamycin-Stammlösung

75 mg/ml Kanamycin-Sulfat wurden in Wasser gelöst, sterilfiltriert und bei 4 °C gelagert.

Lösungen für die Silberfärbung von Polyacrylamidgelen

- **Fixierlösung I:** 30% (v/v) Ethanol, 10% (v/v) Essigsäure
- **Fixierlösung II:** 30% (v/v) Ethanol, 0,5 M NaAcetat
0,1 g $\text{Na}_2\text{S}_2\text{O}_3$ und 1 ml Glutardialdehyd (25%) wurden frisch zu 50 ml Fixierlösung II zugeben.
- **Silbernitratlösung:** 0,1% AgNO_3 . Die Lagerung erfolgte lichtgeschützt bei Raumtemperatur.
- **Entwicklungslösung:** Eine wässrige Lösung von 2,5% Natriumcarbonat und 0,02% Formaldehyd. Die Lösung wurde vor jedem Gebrauch frisch angefertigt.

Lösungen für den Western Blot:

- **Transferpuffer:** 20 mM Tris, 200 mM Glycin, 20 % Methanol
- **Blockpuffer:** PBS-T mit Magermilchpulver (Endkonz.: 7 % (w/v))
- **PBS-T:** 1x PBS versetzt mit 0,1 % (v/v) Tween 20

PBS (10x)

80 g/l NaCl, 2 g/l KCl, 14,4 g/l $\text{Na}_2\text{HPO}_4 \times 2\text{H}_2\text{O}$, 2,4 g/l KH_2PO_4 . Der pH-Wert betrug 7,4.

Zellresuspensionspuffer zur Präparation von *Inclusion bodies*

20 mM Tris/HCl, pH 8,0

Puffer für die Rückfaltung von Proteinen aus *Inclusion bodies*

1, 2 und 6 M Guanidiniumhydrochlorid (GdmCl) wurden in 10 mM KP, pH 7,5 gelöst.

SDS-PAGE-Laufpuffer (1x)

25 mM Tris, 0,1 % (w/v) SDS, 0,2 M Glycin

SDS-PAGE-Probenpuffer (2x) mit bzw. ohne Mercaptoethanol

10 % (w/v) Glycerin, mit bzw. ohne 5 % (v/v) Mercaptoethanol, 2 % (w/v) SDS, 0,01 % (w/v) Bromphenolblau, 1,25 M Tris/HCl, pH 6,8

SDS-PAGE-Sammelgelpuffer

0,5 M Tris/HCl, 0,4 % (w/v) SDS, pH 6,8

SDS-PAGE-Trenngelpuffer

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

Saccharose-Farbmarker für die Agarose-Gelelektrophorese

60 % (w/v) Saccharose, 0,05 % (w/v) Bromphenolblau, 0,05 % (w/v) Xylencyanol FF
in 0,5x TBE Puffer

TBE-Puffer (5x)

445 mM Tris, 445 mM Borsäure, 12,5 mM EDTA

TFB-I Lösung

100 mM KCl, 50 mM MnCl₂, 30 mM KOAc, 10 mM CaCl₂, 15 % (v/v) Glycerin. Der pH-Wert wurde mit Hilfe von Essigsäure auf 5,8 – 6,2 eingestellt.

TFB-II Lösung

10 mM Tris/HCl, 10 mM KCl, 75 mM CaCl₂, 15 % (v/v) Glycerin. Der pH-Wert wurde mit Kaliumhydroxid auf 7,0 eingestellt.

Tris-Tricin-Gel-Puffer (3x)

3 M Tris, 0,3% (w/v) SDS, pH 8,45

Tris-Tricin-Laufpuffer I

0,1 M Tris, 0,1 M Tricin, 0,1% (w/v) SDS, pH 8,45

Tris-Tricin-Laufpuffer II

0,2 M Tris/HCl, pH 8,9

Waschpuffer für die Präparation von *Inclusion bodies*

20 mM Tris/HCl, 2 M Urea, 500 mM NaCl, 2% (v/v) Triton X-100, pH 8,0

2.12 *Software*

BioEdit	Version 7.0.0	Ibis Biosciences, (Carlsbad, CA, USA)
Blast		NCBI (National Center for Bioinformation)
Cary Eclipse	Version 1.1 (132)	Varian Australia Pty Ltd. 1999/2000
Cary Win UV	Version 3.0 (182)	Varian Australia Pty Ltd. 2002
ClustalW	Version 1.83	EMBL-EBI (European Molecular Biology Laboratory-European Bioinformatics Institute)
Corel Draw	Version 12.0	Corel Corporation
Dichroweb		WebServer unter http://www.cryst.bbk.ac.uk/cdweb/html/home.html
Doc-It	Version 2.3.0	UVP (Upland, USA)
Gene Runner	Version 3.05	Hastings Software Inc. 1994
JalView	Version 2.3	Barton Group
Jasco Manager	Version 1.54.03	Jasco (Groß-Umstadt)
MAMMOTH	Version 1.0	WebServer unter http://ub.cbm.uam.es/mammoth/pair/index3.php
Office 2003	Word, Excel, Power Point	Microsoft Corporation
Pov Ray™	Version v3.6 incl Win32	POV-team 1991-2002
SeqMan	Version 4.03	DNASTAR Inc.
Sigma Plot	Version 9.0	Systat Software Inc.
Sigma Scan	Version 2.0	Systat Software Inc.
Swiss PDB Viewer	Version v3.7.b	N. Guex 1995-1999
TCoffee	Version 5.31	WebServer unter http://www.igs.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi
Unicorn	Version 5.0	GE Healthcare (Freiburg)

3 Methoden

3.1 *Mikrobiologische Methoden*

3.1.1 Vorbereitung von Geräten und Lösungen

Hitzestabile Glaswaren, Lösungen und Medien wurden 30 min bei 121 °C und 2 bar autoklaviert. Hitzelabile Lösungen, die nicht durch Autoklavieren sterilisiert werden konnten, wurden durch einen Sterilfilter ME24 mit einer Porengröße von 0,2 µm filtriert, wodurch praktisch alle Mikroorganismen aus der Lösung entfernt wurden. Lösungen für die Säulenchromatographie wurden zusätzlich entgast.

3.1.2 Anzucht und Lagerung von Bakterienstämmen

E. coli Stämme wurden, wenn nicht anders angegeben, bei 37 °C in LB-Medium angezogen. Bei Stämmen mit Selektionsmarker wurde dem Medium das entsprechende Antibiotikum aus einer 1000-fach konzentrierten Stammlösung zugesetzt. Flüssigkulturen bis zu einem Volumen von 50 ml wurden bei 250 Upm schüttelnd inkubiert (bis zu 5 ml im Reagenzglas, darüber im Erlenmeyerkolben). 1000 ml-Flüssigkulturen wurden bei 115 Upm in einem 3000 ml Erlenmeyerkolben, teilweise mit Schikanen, angezogen. Das Wachstum der Zellen wurde durch Messung der optischen Dichte (OD₆₀₀) am Photometer gegen reines Medium verfolgt.

Zur Anzucht von Einzelkolonien wurde die Zellsuspension in gewünschten Verdünnungen auf den entsprechenden Festmedien ausplattiert, und ebenfalls bei 37 °C inkubiert.

Die dauerhafte Lagerung erfolgte in einer 1:1 Mischung mit Glycerin bei -80 °C (Glycerinkultur).

3.1.3 Transformation von *E. coli* Zellen

3.1.3.1 Präparation chemisch kompetenter Zellen (Inoue *et al.*, 1990)

Zur Präparation von chemisch kompetenten Zellen wurde 500 ml SOB-Medium in einem 1 l Erlenmeyerkolben mit Zellen des entsprechenden Stammes angeimpft und bei 37 °C und 250 Upm bis zu einer OD₆₀₀ von 0,4 – 0,7 angezogen. Die Kultur wurde 15 Minuten auf Eis abgekühlt und anschließend durch Zentrifugation (Eppendorf Centrifuge 5810 R, 4000 Upm, 10 Minuten, 4 °C) geerntet. Das Zellpellet wurde in 50 ml eiskalter TFB-I Lösung aufgenommen und anschließend erneut unter den gleichen Bedingungen zentrifugiert. Zum Schluss wurden die Zellen in 5 ml eiskalter TFB-II Lösung resuspendiert, in 100 µl-Aliquots in 1,5 ml Reaktionsgefäßen in flüssigem Stickstoff schockgefroren und bei -80 °C gelagert.

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

Zur Transformation wurden die kompetenten Zellen auf Eis aufgetaut und mit maximal 100 ng Plasmid-DNA oder 10 µl eines Ligationsansatzes (3.2.7) versetzt. Nach fünfminütiger Inkubation auf Eis wurde der Ansatz für 42 Sekunden einem Hitzeschock (42 °C) ausgesetzt, fünf Minuten auf Eis abgekühlt und anschließend nach Zugabe von 900 µl LB-Medium eine Stunde bei 37 °C zur Ausbildung der Antibiotikaresistenz inkubiert („kuren“). Danach wurden die Zellen in einer passenden Verdünnung auf Selektivmedium ausplattiert und über Nacht bei 37 °C inkubiert.

3.1.3.3 Bestimmung der Transformationseffizienz

Zur Bestimmung der Transformations-Effizienz T_E wurde von den transformierten Zellen nach dem „Kuren“ (3.1.4.2) eine Verdünnungsreihe erstellt und von jeder Verdünnung, sowie von der unverdünnten Probe 200 µl auf Selektivmedium

ausgestrichen. Aus der Anzahl gewachsener Kolonien wurde nach Formel 1 die Transformations-Effizienz ermittelt:

$$\frac{\text{Kolonien auf der Platte} \cdot \text{Verdünnungsfaktor}}{\mu\text{g zugegebener DNA}} = \text{Transformationseffizienz}$$

Formel 1: Bestimmung der Transformationseffizienz.

3.2 *Molekularbiologische Methoden*

3.2.1 Isolierung von Plasmid-DNA im präparativen Maßstab

Die Isolierung von Plasmid-DNA in präparativen Mengen erfolgte mit Hilfe des Nucleobond AX DNA-Extraktionskits (Machery-Nagel) nach dem Protokoll des Herstellers. Für 50 ml Kulturvolumen wurde die AX-100 Säule verwendet. Die von der Säule eluierte DNA wurde mit 0,7 Volumen Isopropanol gefällt und zentrifugiert (Eppendorf Centrifuge 5415 R, 13000 Upm, 1 Stunde, 4 °C). Das Pellet wurde anschließend mit 100 µl 70 % (v/v) Ethanol gewaschen und in 50 µl Wasser aufgenommen.

3.2.2 Bestimmung der DNA-Konzentration

Die DNA-Konzentration einer wässrigen Lösung lässt sich aus der Absorption bei 260 nm (A_{260}) nach dem Lambert-Beerschen Gesetz bestimmen. Die Messung erfolgte in einer Quartzküvette mit einer Schichtdicke von 1 cm gegen Wasser als Referenz. Eine A_{260} von 1 entspricht etwa einer Konzentration von 50 µg doppelsträngiger (ds) DNA pro ml Lösung.

Entsprechend berechnet sich die DNA-Konzentration wie folgt:

$$c_{dsDNA} (\mu g / ml) = A_{260} \cdot 50 \cdot f$$

c: Konzentration

f: Verdünnungsfaktor

A_{260} : Absorption bei 260 nm

Formel 2: Photometrische Bestimmung von DNA-Konzentrationen.

In der Regel wurde eine 100-fache Verdünnung der DNA-Lösung analysiert. Für die Messung wurde ein Spektrum von 240 bis 320 nm mit dem Cary Bio 100 Spektralphotometer (Varian) aufgenommen. Das Verhältnis von A_{260}/A_{280} ist ein Maß für die Reinheit der Lösung und erreicht bei proteinfreien Proben einen Wert von größer als 1,8.

Eine grobe Abschätzung der DNA-Konzentration erfolgte nach gelelektrophoretischer Auftrennung im Agarosegel durch Vergleich der Bandenintensität mit den Banden der eingesetzten DNA-Längenstandards (2.9.1).

3.2.3 Polymerasekettenreaktion (PCR)

Die Polymerasekettenreaktion (PCR) ist eine *in vitro* Methode zur enzymatischen Amplifikation spezifischer DNA-Fragmente (Mullis & Faloona, 1987; Saiki *et al.*, 1988). Sie ermöglicht die selektive Vervielfältigung eines DNA-Fragments aus einem komplexen DNA-Gemisch heraus. Die Reaktion basiert auf Anlagerung (*Annealing*) und Verlängerung (*Extension*) zweier Oligonukleotid-Primer, die revers und komplementär zu dem 5'- bzw. 3'-flankierenden Bereich einer doppelsträngigen Zielsequenz sind. Die Reaktion besteht aus aufeinander folgenden Amplifikationszyklen, die jeweils durch drei Phasen unterschiedlicher Temperatur charakterisiert sind. In der ersten Phase, der Denaturierungsphase, wird die Doppelstrang-DNA bei 95 °C in Einzelstrang-DNA überführt. Durch das anschließende Absenken auf die primerspezifische *Annealing*-Temperatur in der Anlagerungsphase kommt es zur Hybridisierung der Oligonukleotid-Primer mit den komplementären Zielregionen der einzelsträngigen Matrizen-DNA. In der Extensionsphase werden bei höherer Temperatur die Primer durch eine

thermostabile Polymerase komplementär zur Matrize verlängert. Da in jedem Zyklus die Zahl der DNA-Stränge verdoppelt wird, kommt es zu einem exponentiellen Anstieg der Fragmente der Zielsequenz.

Die PCR-Bedingungen wurden in Abhängigkeit von der Länge der zu amplifizierenden Sequenz und dem G/C-Gehalt der jeweiligen spezifischen Primer gewählt. Ein typischer Reaktionsansatz hatte ein Volumen von 50 µl und enthielt folgende Komponenten: 100 ng Matrizen-DNA, 100 pmol eines jeden Primers und Ready RED *Taq* PCR Reaction Mix mit MgCl₂ (Sigma-Adrich) bzw. Green Go *Taq*TM Reaction Buffer (Fermentas) nach Herstellerangaben. Zu Ansätzen mit Green Go *Taq*TM Reaction Buffer mussten noch dNTP's mit einer Endkonzentration von je 0,2 mM und 1 U Go *Taq*TM-Polymerase zugesetzt werden.

Nach einer initialen Denaturierungsphase für 3 Minuten bei 95 °C wurde der Ablauf von 30 Reaktionszyklus gestartet: Denaturierung für 45 s bei 95 °C, Primer-*Annealing* für 45 s bei der für das jeweilige Primerpaar charakteristischen Temperatur (siehe unten) und Primer-Extension für 1 min/kb bei 72 °C. Nach einer abschließenden Extension für 10 Minuten bei 72 °C wurde der Ansatz auf 8 °C abgekühlt.

Die Schmelztemperatur T_M eines Primers wurde nach einer empirischen Formel, die sowohl den relativen molaren GC-Gehalt (% GC), als auch die Länge der Oligonukleotide (n) berücksichtigt, berechnet:

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

Formel 3: Berechnung der Schmelztemperatur eines Oligodesoxyribonukleotids.

Die Schmelztemperatur (T_M) lag in der Regel zwischen 50 und 65 °C.

Aus den Schmelztemperaturen des Primerpaares (T_{M1} und T_{M2}) wurde die *Annealing*-Temperatur T_A wie folgt ermittelt:

$$T_A = \left[\frac{T_{M1} + T_{M2}}{2} \right] - 3^\circ C$$

Formel 4: Bestimmung der *Annealing*-Temperatur in einer PCR-Reaktion.

3.2.4 Kolonie-PCR (*Insert-Screening*) zur Überprüfung des Klonierungserfolgs

Zur schnellen Überprüfung des Erfolges einer Klonierung wurden Kolonien nach einer Transformation (3.1.3.2) mit Hilfe einer Kolonie-PCR auf das Vorhandensein eines Inserts überprüft. Dazu wurde mit einer sterilen Pipettenspitze etwas Zellmaterial der zu untersuchenden Kolonie entnommen und in einem vorbereiteten PCR-Ansatz (3.2.3) resuspendiert. Mit dem Rest der an der Pipettenspitze haftenden Zellen wurden 5 ml LB-Medium inokuliert, um gegebenenfalls das Plasmid isolieren zu können. Der PCR-Ansatz enthielt in der Regel Primer, die stromauf- und abwärts der MCS auf dem Plasmid lagen. In einem 5 min dauernden ersten Aufheizschrittes (95 °C) wurden die Zellen lysiert und gaben die Plasmide frei. Nach beendeter PCR wurde die Größe der erhaltenen PCR-Produkte durch Agarosegelelektrophorese (3.2.8) abgeschätzt.

3.2.5 *Splicing Overlap Extension-PCR* (SOE-PCR)

Die *Splicing Overlap Extension-PCR* ist eine *in vitro* Mutagenesetechnik, die in zwei Schritten zu einem gezielten Austausch eines oder mehrerer Nukleotide führt (Ho *et al.*, 1989). Dazu werden zwei zueinander komplementäre Mutationsprimer von etwa 20 bp Länge definiert. Diese tragen im zentralen Bereich den gewünschten Nukleotidaustausch. Im ersten Schritt werden in zwei getrennten PCR-Ansätzen (3.2.3) zwei Fragmente mit Hilfe der klonierten, zu verändernden DNA-Sequenz als Matrize, je eines Mutationsprimers und des entsprechenden vektorspezifischen stromauf- bzw. abwärts flankierenden Primers amplifiziert. Die beiden Fragmente teilen eine kurze überlappende Region, die den eingeführten Austausch enthält.

Diese Fragmente dienen im zweiten Schritt als Matrize, um mit Hilfe des stromauf- und abwärts Primers das komplette Gen zu amplifizieren (Abbildung 18). Dafür wurden die PCR-Produkte des ersten Schrittes auf ein Agarosegel aufgetragen, ausgeschnitten und aus dem Gel extrahiert (3.2.9). Je 4 µl dienten im zweiten Schritt als Matrize, in welchem die *Annealing*-Temperatur grundsätzlich 45 °C betrug.

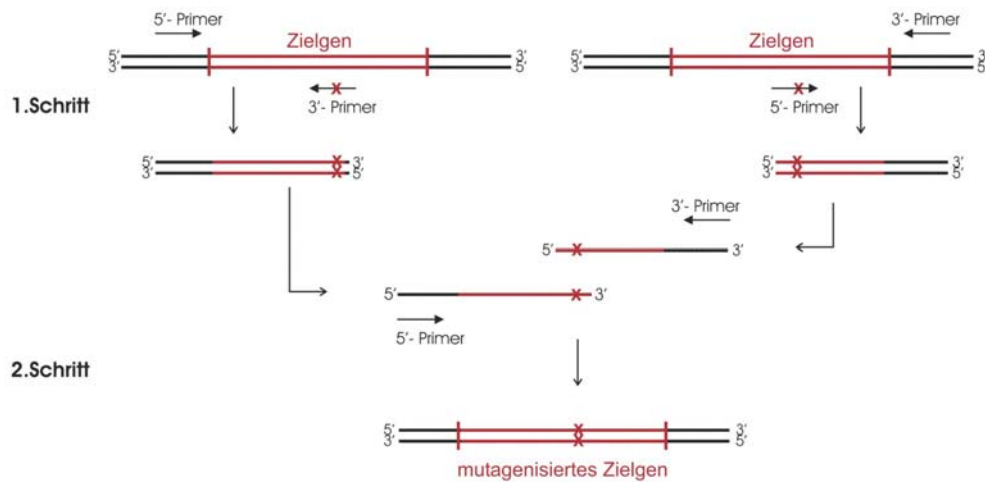


Abbildung 18: Schematische Darstellung des Ablaufes einer *Splicing Overlap Extension-PCR* (SOE-PCR).

1. Schritt: In zwei getrennten Ansätzen dient das Zielgen als Matrize. Die Amplifikation wird jeweils mit einem endständigen und einem Mutationsprimer durchgeführt.

2. Schritt: Die Reaktionsprodukte aus Schritt 1 überlappen in der Region, die den Austausch trägt und können so in einer zweiten PCR-Runde als Matrize eingesetzt werden. Das Produkt dieser Reaktion ist das mutagenisierte Zielgen.

3.2.6 Spaltung von DNA mit Restriktionsendonukleasen (Wilson & Murray, 1991)

DNA wurde mit Restriktionsendonukleasen unter den vom Hersteller empfohlenen Temperatur- und Pufferbedingungen gespalten. Für analytische Zwecke wurden in der Regel 200 ng DNA mit etwa 2 U Enzym gespalten, in präparativen Ansätzen 1 µg DNA durch mindestens 20 U Enzym. Für einen Doppelverdau mit zwei Enzymen wurde ein nach Herstellerangaben geeigneter Puffer gewählt. War dies nicht möglich, wurde ein sequentieller Verdau durchgeführt, wobei nach der ersten Spaltung eine Reinigung durch Gelextraktion (3.2.9) erfolgte. Der Erfolg einer Spaltung wurde mittels Agarosegelelektrophorese überprüft (3.2.8).

3.2.7 Ligation von DNA-Enden

Bei der Ligation von DNA-Fragmenten zu einem zirkulären Produkt wurde das Insertfragment in einem 5 – 8-fachen molaren Überschuss zum Vektorfragment eingesetzt. Beide Fragmente wurden zuvor zur Reinigung auf ein Agarosegel aufgetragen und extrahiert (3.2.9). Die Ligation erfolgte in T4-Ligase-Puffer mit 2 U T4-Ligase für 16 Stunden bei 16 °C in einem Gesamtvolumen von 20 µl.

3.2.8 Agarosegelelektrophorese

Die Auftrennung doppelsträngiger (ds) DNA erfolgte in Gelen von 1% (w/v) Agarose in 0,5 x TBE-Puffer. Die Agarose wurde in einem handelsüblichen Mikrowellengerät aufgekocht und mit 0,5 µg/ml Ethidiumbromid versetzt. Die Proben wurden mit mindestens 1/5 Volumen Saccharose-Farbmarker versetzt und nach Aushärtung des Gels und Zugabe des Laufpuffers (0,5 x TBE) in die Geltaschen gefüllt. Als Längenstandard diente die *Smart Ladder* (Eurogentec) bzw. der *Gene Ruler™* (Fermentas). Die Elektrophorese wurde bei einer Spannung von 190 V durchgeführt, die aufgetrennten DNA-Fragmente unter UV-Licht der Wellenlänge 302 nm (Sharp *et al.*, 1973) sichtbar gemacht und ggf. extrahiert (3.2.9).

3.2.9 Extraktion von DNA-Fragmenten aus Agarosegelen

Für eine Aufreinigung wurden die DNA-Fragmente mit einem Skalpell aus dem Agarosegel ausgeschnitten und entweder mit dem QIAquick Gel Extraction Kit (Quiagen) oder dem Nucleo Spin Extract Kit (Nucleobond) nach Herstellerangaben gereinigt.

3.2.10 DNA-Sequenzierung

Die Bestimmung der Nukleotidsequenzen von Genen erfolgte durch die Geneart AG (Biopark Regensburg). Es wurde mit dem Kettenabbruch-Verfahren von Sanger (1977) gearbeitet, bei dem verschiedene fluoreszenzmarkierte 2',3'-Dideoxynukleotide zusätzlich zu den 2'-Desoxynukleotiden in der *Thermal Cycle Sequencing* Reaktion (Sears *et al.*, 1992; Slatko, 1994) eingesetzt werden. Dadurch entstehen 3'-terminal fluoreszenzmarkierte Einzelstrang-Abbruchfragmente, die gelelektrophoretisch aufgetrennt werden. Als Matrize wurden 300 ng Plasmid-DNA eingesetzt. Die Primermenge betrug 10 pmol. Die Auswertung der Daten erfolgte mit den Programmen SeqMan, Gene Runner und BlastN (2.12).

3.3 Proteinchemische Methoden

3.3.1 Expression von Genen und Anreicherung von Proteinen in analytischen Mengen

Alle zu exprimierenden Gene lagen in Vektoren (pET-Serie bzw. pTYB) vor, die unter der Kontrolle des T7 Promotor- und *lac* Operator- Systems stehen (2.7.1).

Um zu untersuchen, ob sich die klonierten Gene überexprimieren lassen, wurden Expressionen im analytischen Maßstab (Probeexpressionen) durchgeführt. Dazu wurden mit einer frisch transformierten *E. coli* Kolonie 5 ml LB-Medium mit entsprechendem Antibiotikum inokuliert und bei 37 °C über Nacht schüttelnd inkubiert. Mit dieser Kultur wurden weitere 20 ml LB-Medium mit entsprechendem Antibiotikum angeimpft und bei 37 °C bis zum Erreichen einer Zelldichte von $OD_{600} = 0,8$ inkubiert. Dieser Ansatz wurde auf zwei Kolben aufgeteilt. Einer Hälfte wurde zur Induktion der Expression IPTG (Endkonz.: 1 mM) zugegeben, der anderen nicht. Nach Inkubation über Nacht wurden die Zellen durch Zentrifugation (Eppendorf Centrifuge 5810R, 4000 Upm, 10 min, 4 °C) pelletiert, in 1 ml Aufschlusspuffer (2.11) resuspendiert und mit einem Ultraschall Desintegrator (Branson Digital Sonifier, Model W-250 D; Amplitude 20 %, 2x 30 sek mit 1 min Pause) aufgeschlossen.

Die Proben wurden in einer Tischzentrifuge (Eppendorf 5415 D, 13000 Upm, 5 min, RT) zentrifugiert und der Überstand vollständig in ein neues Reaktionsgefäß überführt. Das erhaltene Pellet mit den unlöslichen Zellbestandteilen (P) wurde in 1 ml Aufschlusspuffer (2.11) resuspendiert, 100 µl davon in ein neues Eppendorf Reaktionsgefäß überführt und mit gleichem Volumen 2x SDS-Probenpuffer versetzt. Zur Untersuchung des löslichen Zellextrakts (*crude extract* oder Rohextrakt, C) wurden 100 µl des Überstands in ein neues Reaktionsgefäß überführt und mit gleichem Volumen 2x SDS-Probenpuffer versetzt.

Die beiden mit SDS-Probenpuffer versetzten Aliquots (P, C) wurden für 5 min bei 95 °C inkubiert und per SDS-PAGE (3.4.1) analysiert.

3.3.2 Expression von Genen und Anreicherung von Proteinen in präparativen Mengen

Um Proteinmengen im Milligramm-Maßstab zu reinigen, wurden Expressionen im präparativen Maßstab durchgeführt. Dazu wurden mit einer frisch transformierten *E. coli* Kolonie 50 ml LB-Medium mit entsprechendem Antibiotikum inokuliert und bei 37 °C über Nacht schüttelnd inkubiert. Mit dieser Kultur wurde 1 l LB-Medium mit entsprechenden Antibiotika angeimpft und bei 37 °C bis zum Erreichen einer Zelldichte von $OD_{600} = 0,8$ inkubiert. Zeigte die Probeexpression an, dass eine Induktion der Expression mit IPTG notwendig ist, wurde beim Erreichen der oben genannten Zelldichte IPTG (Endkonz.: 1 mM) zugegeben. Nach Wachstum über Nacht wurden die Zellen durch Zentrifugation geerntet (Sorvall RC-5B, GS3 Rotor, 4000 Upm, 15 min, 4 °C) und in 50 ml Aufschlusspuffer resuspendiert. Der Zellaufschluss erfolgte durch Sonifizieren mit dem Branson Digital Sonifier (Model W-250 D; Amplitude 50 %, 2 x 1 min mit jeweils 1 min Pause). Die Probe wurde während der gesamten Prozedur auf Eis gehalten, um eine übermäßige Erwärmung der Lösung zu vermeiden. Anschließend wurde ein Aliquot von 100 µl zur Analyse von Pellet und Rohextrakt (P/C) entnommen. Die restliche Probe wurde scharf zentrifugiert (Sorvall RC-5C oder RC-2B, Rotor SS34, 13000 Upm, 15 min) und der Überstand in 50 ml Plastikröhrchen überführt. Die Probe P/C wurde zentrifugiert (Eppendorf 5415 D, 13000 Upm, 10 min, RT), der Überstand (C) in ein neues

Eppendorfgefäß überführt und das resultierende Pellet (P) in 100 µl Aufschlusspuffer aufgenommen. Die Proben (P, C) wurden mit 2x SDS-Probenpuffer versetzt und mittels SDS-PAGE (3.4.1) analysiert.

3.3.3 Rückfaltung der Proteine aus *Inclusion bodies*

Befand sich das exprimierte Protein in der unlöslichen Fraktion (P) des Zellextraktes (*inclusion bodies*), wurde es durch Solubilisierung in chaotropher Lösung mit sich anschließender Rückfaltung zunächst schrittweise in eine lösliche Form überführt. Dazu wurde das Pellet (P) nach Waschen mit Aufschlusspuffer in 20 ml 6 M Guanidiniumhydrochlorid (GdmCl), 10 mM KP, pH 7,5 aufgenommen und 30 min bei RT unter ständigem Rühren inkubiert. Dann wurden 20 ml 1 M GdmCl, 10 mM KP, pH 7,5 zugegeben, nochmals 30 min bei RT unter ständigem Rühren inkubiert und anschließend abzentrifugiert (Sorvall RC-5C oder RC-2B, Rotor SS34, 13000 Upm, 15 min, 4 °C). Der Überstand wurde in ein neues Reaktionsgefäß überführt und mit 70 ml 2 M GdmCl, 10 mM KP, pH 7,5 gemischt. Daraufhin wurde die Lösung gegen 10 l 50 mM KP pH 7,5 dialysiert, wodurch das GdmCl praktisch vollständig entfernt und damit die Rückfaltung des Proteins induziert wurde. Der Dialysepuffer wurde zweimal nach je 24 h zur Einstellung des Diffusionsgleichgewichtes erneuert. Nach der Dialyse wurde der Inhalt des Dialyseschlauches abzentrifugiert (Sorvall RC-5C oder RC-2B, Rotor SS34, 13000 Upm, 20 min, 4 °C). Von Überstand (ÜD) und Pellet (PD) wurden Aliquots mit 2x SDS-Probenpuffer versetzt und mittels SDS-PAGE analysiert (3.4.1).

3.3.4 Größenausschlusschromatographie (SEC) im präparativen Maßstab

Die präparative SEC (vom engl. *size exclusion chromatography*) wurde in dieser Arbeit genutzt, um nach der Rückfaltung (3.3.3) der Proteine höhermolekulare Verunreinigungen zu entfernen.

Die verwendete Säule (HiLoad 26/60 Superdex 75 *prep grade*) mit einem Säulenvolumen von 320 ml besteht aus quervernetzter Agarose und Dextran mit

einer durchschnittlichen Partikelgröße von 34 µm. Das Trennprinzip ist das eines umgekehrten Molekularsiebs, in dem die Proteine aufgrund ihrer Größe unterschiedlich lange verweilen (Ausschluss-Chromatographie). Die präparative Gelfiltrationssäule ist an eine FPLC-Anlage angeschlossen und wird bei 4°C betrieben.

Die Probe wurde über einen Auftragsloop (10 ml) auf die mit Laufpuffer (50 mM KP, 300 mM KCl, pH 7,5) äquilibrierte Gelfiltrationssäule gegeben. Die Elution erfolgte durch Spülen der Säule mit 360 ml Laufpuffer bei einer Geschwindigkeit von 0,5 ml/min. Der Fraktionssammler wurde bei Beginn des Laufes gestartet und auf eine Fraktionsgröße von 4 ml eingestellt, d.h. insgesamt wurden 90 Fraktionen gesammelt. Die Analyse interessanter Fraktionen erfolgte mittels SDS-PAGE (3.4.1).

3.3.5 Rückfaltung der Proteine aus *Inclusion bodies* nach Immobilisierung auf einer Affinitätssäule

Diese Methode (Amersham Biosciences, 1999) wurde vor allem bei den Proteinen angewendet, die zwar in *Inclusion bodies* exprimiert werden, sich aber durch Rückfaltung (3.3.3) mit anschließender präparativer Gelfiltration (3.3.4) nicht rein darstellen lassen. Dazu mussten die Proteine einen His₆-Tag tragen. Dieser bindet auch in Anwesenheit von chaotropen Denaturierungsmitteln an die divalenten Metallionen einer HisTrap™-Säule. Dadurch lassen sich die gewünschten Proteine vereinzeln und Verunreinigungen auswaschen.

Zur Präparation der *Inclusion bodies* wurden die Zellen nach der Ernte in 40 ml Resuspensionspuffer (20 mM Tris/HCl, pH 8,0) aufgenommen, mit einer Spatelspitze Lysozym versetzt und 30 min auf Eis inkubiert. Anschließend wurden die Zellen mit dem Branson Digital Sonifier (Model W-250 D; Amplitude 40 %, 3 x 30 sek mit jeweils 1 min Pause) sonifiziert und abzentrifugiert (Sorvall RC-5C oder RC-2B, Rotor SS34, 13000 Upm, 10 min, 4 °C). Das Pellet wurde in 30 ml kaltem Waschpuffer (20 mM Tris/HCl, 2 M Harnstoff, 0,5 M NaCl, 2% Triton X-100, pH 8,0) resuspendiert und erneut wie oben beschrieben sonifiziert und abzentrifugiert. Nach einem weiteren Waschschriff wurden die *Inclusion bodies* in 50 ml Bindepuffer (20 mM Tris/HCl, 6 M GdmCl, 0,5 M NaCl, 5 mM Imidazol, pH 8,0) durch Rühren für 1 h bei RT solubilisiert

und wiederum wie oben zentrifugiert. Nachdem der Überstand durch einen Spritzenvorsatzfilter (0,22 µm) filtriert worden war, wurde er auf die HisTrap™-Säule (Säulenvolumen = 1 ml) geladen. Diese war an eine ÄKTA™prime plus-Anlage angeschlossen, die Flussrate betrug 0,5 – 1 ml/min mit einem maximalen Rückdruck von 0,5 MPa. Nachdem die gesamte Proteinlösung aufgetragen war, wurde mit weiteren drei Säulenvolumen (SV) Bindepuffer behandelt. Anschließend wurde über einen Stufengradienten auf 100 % Waschpuffer (20 mM Tris/HCl, 6 M Urea, 0,5 M NaCl, 5 mM Imidazol, pH 8,0) gewechselt und die Säule mit vier SV gewaschen. Die Rückfaltung erfolgte durch einen linearen Gradienten über acht SV auf 100 % Rückfaltungspuffer (20 mM Tris/HCl, 0,5 M NaCl, 5 mM Imidazol, pH 8,0), mit dem noch für weitere 10 SV gewaschen wurde. Danach konnten die Proteine mittels eines linearen Gradienten über 20 SV auf 100 % Elutionspuffer (20 mM Tris/HCl, 0,5 M NaCl, 1 M Imidazol, pH 8,0) von der Säule eluiert werden. Der Fraktionssammler wurde bei Beginn der Elution gestartet, wobei Fraktionen von 1 ml gesammelt wurden. Auch der anschließende Waschschrift mit 100 % Elutionspuffer über fünf SV wurde gesammelt. Die gesamte Reinigung konnte über den Leitfähigkeitsmesser und die Aufzeichnung der Absorption bei 280 nm verfolgt werden. Interessante Fraktionen wurden mittels SDS-PAGE (3.4.1) analysiert und solche mit sauberem Protein vereinigt und über Nacht gegen 10 l Puffer (50 mM KP, pH 7,5) bei 4 °C dialysiert (3.3.8).

3.3.6 *Intein Mediated Purification with an Affinity Chitin-binding Tag* (IMPACT™)

Bei der *Intein Mediated Purification with an Affinity Chitin-binding Tag* (IMPACT™) Methode macht man sich die induzierbare Selbst- oder Eigenspaltungsaktivität des Protein-*splicing*-Elements Intein zu nutze. Dadurch lässt sich das Zielprotein in nur einem Säulenreinigungsschritt gewinnen. Proteine, deren Gene in dem Vektor pTYB1 (2.7.2) exprimiert wurden, tragen C-terminal das *Saccharomyces cerevisiae* Intein als Fusionsprotein. Intein wiederum ist C-terminal eine Chitinbindedomäne (CBD) nachgelagert, durch die das zu reinigende Protein an eine Chitin-Säule binden kann. Durch die Selbstspaltungsaktivität des Inteins, die durch Dithiotreitol

(DTT) oder β -Mercaptoethanol induziert werden kann, werden Ziel- und Fusionsprotein voneinander getrennt. Die Intein-CBD-Einheit verbleibt auf der Säule und das Zielprotein kann eluiert werden (Abbildung 19).

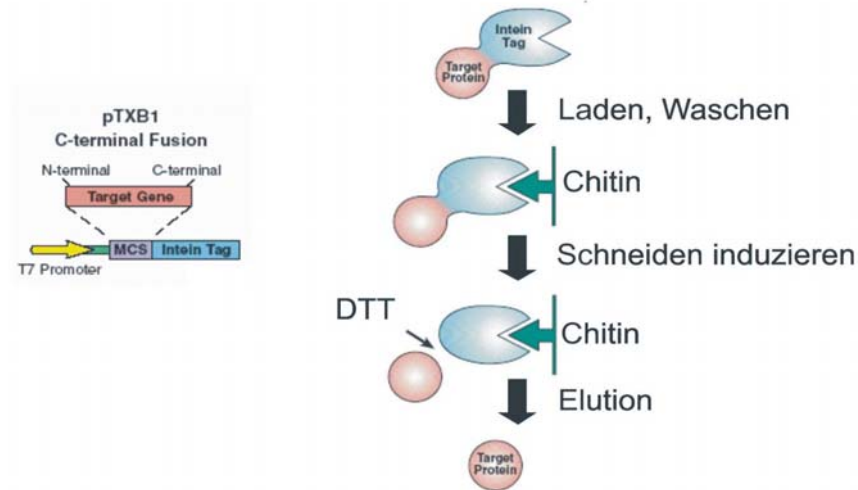


Abbildung 19: Schematische Darstellung der Chitinaffinitätsreinigung. Aus dem *Instruction Manual* für den IMPACT™ Kit (NewEnglandBiolabs).

Zellen, in denen das Fusionsprotein exprimiert wurde, wurden nach der Ernte und Aufnahme in 50 ml Säulenpuffer (20 mM Tris/HCl, 0,5 M NaCl, 1 mM EDTA, pH 8,0) wie in 3.3.2 beschrieben, einer Rohextraktpräparation unterzogen. Der Rohextrakt wurde nach Filtration (Spritzenvorsatzfilter 0,22 μ m) auf die nach Herstellerangaben vorbereitete und mit Säulenpuffer äquilibrierte Chitin-Säule (Säulenvolumen 20 ml) geladen. Die Reinigung erfolgte bei 4 °C durch Gravitationsfluss. Nach einem Waschschrift mit 10 SV wurde die Selbstspaltungsaktivität des Inteins mit Hilfe von 3 SV *Cleavage*-Puffer (20 mM Tris/HCl, 0,5 M NaCl, 1 mM EDTA, 50 mM DTT, pH 8,0) induziert. Die Reaktion lief bei 4 °C in 16 – 40 h vollständig ab. Danach wurde das Zielprotein durch Spülen der Säule mit Säulenpuffer eluiert. Dabei wurden (per Hand in Reaktionsgefäßen) 1 ml Fraktionen aufgefangen und mittels SDS-PAGE (3.4.1) analysiert. Nach jeder Anwendung wurde die Säule mit drei SV Regenerationspuffer (0,3 M NaOH) gespült, 30 min bei 4 °C inkubiert und danach noch einmal mit sieben SV Regenerationspuffer gespült. Der Regenerationspuffer wurde mit 20 SV Wasser entfernt und die Säule mit fünf SV Säulenpuffer äquilibriert. Danach war die Chitin-Säule wieder einsatzbereit.

3.3.7 Enzymatische Spaltung der HisF-Hälften mit Faktor Xa

Faktor Xa ist ein Blutgerinnungsfaktor, der mit hoher Spezifität die Aminosäuresequenz -Ile-Glu-Gly-Arg- erkennt und C-terminal davon schneidet.

Faktor Xa wurde im Rahmen dieser Arbeit genutzt, um HisF_N2 bzw. HisF_C2 von HisF_N1 bzw. HisF_C1 zu trennen. Dazu wurde mittels PCR (3.2.3) je eine Faktor Xa-Schnittstelle in die Hälften HisF_N und HisF_C eingefügt. Die resultierenden Proteine besaßen C-terminal einen His₆-Tag, wodurch nach Verdau mit Faktor Xa die Viertel HisF_N2 und HisF_C2 durch Ni-Chelat-Affinitätschromatographie angereichert werden konnten.

Für einen Verdauansatz wurden Faktor Xa und Substrat in einem molaren Verhältnis von 1:50 bis 1:100 gemischt. Die Inkubation erfolgte in 1x Faktor Xa Salzlösung für 18 h bei pH 8 und 22 °C. Die Analyse auf vollständige Spaltung erfolgte durch SDS-PAGE (3.4.1).

3.3.8 Dialyse von Proteinlösungen

Um einen Pufferwechsel zwischen zwei Reinigungsschritten zu vollziehen bzw. zur Rückfaltung von Proteinen aus der unlöslichen Fraktion, wurden die Proteinlösungen gegen ein mindestens 100-faches Volumen an gewünschtem Puffer über Nacht bei 4 °C dialysiert. Die Dialyseschläuche wurden zuvor in bidest. Wasser eingeweicht.

3.3.9 Einkonzentrieren von Proteinlösungen und Lagerung der Proteine

Proteinlösungen wurden durch Zentrifugation einkonzentriert. Dafür wurden, entsprechend der Menge der Proteinlösung und Größe des Proteins, *Amicon Centrifugal Devices* (MWCO: 5, 10 bzw. 30 kDa) der Firma Millipore nach Herstellerangaben verwendet.

Die konzentrierten Proteinlösungen wurden in flüssigen Stickstoff eingetropfelt und die gefrorenen Kügelchen bei -80 °C gelagert. Eine kurzfristige Lagerung erfolgte bei 4 °C auf Eis.

3.3.10 Festphasenpeptidsynthese (Woolley & Merrifield, 1963)

Bei der Festphasenpeptidsynthese (Woolley & Merrifield, 1963) wird das Peptid an einem polymeren Träger schrittweise durch Kondensationsreaktionen vom C-Terminus ausgehend aufgebaut. Die C-terminale Aminosäure des zu synthetisierenden Peptids wird dabei mit ihrer Carboxygruppe über eine Ankergruppe an ein polymeres Harz geheftet. Der N-Terminus der darauf folgenden Aminosäure wird entweder durch 9-Fluorenylmethoxycarbonyl (Fmoc) oder t-Butyloxycarbonyl (Boc) geschützt. Ihre Carbonylgruppe wird z.B. als Ester mit Hilfe von Kupplungsreagenzien aktiviert. Danach erfolgt die Kupplung an die freie Aminogruppe der ersten Aminosäure und die Abspaltung der Schutzgruppe. Dieser Zyklus wird so lange wiederholt bis das Peptid die gewünschte Länge erreicht hat. Dabei ist zu beachten, dass auch die reaktiven Seitenketten der sogenannten trifunktionellen Aminosäuren (Lys, Arg, His, Glu, Asp, Ser, Thr, Tyr, Cys) durch entsprechende reversible Schutzgruppen bei der Kupplung blockiert werden müssen, um falsche Verknüpfungen zu verhindern. Nach der Synthese wird die kovalente Bindung zwischen der C-terminalen Aminosäure und der Ankergruppe gespalten, wobei je nach Anker das Peptid als Säure oder Amid vorliegen kann. Die Seitenkettenschutzgruppen werden hierbei mitabgespalten. Die Analyse der Reinheit der synthetischen Proteine erfolgt über *Reversed Phase* (RP)-HPLC, die anschließende Charakterisierung durch Elektrospray-Ionisierung-Massenspektrometrie (ESI-MS).

Die Festphasenpeptidsynthese wurde im Rahmen dieser Arbeit zur Darstellung der Viertel HisF_N2 und HisF_C2 eingesetzt und freundlicherweise von der Arbeitsgruppe von Frau Dr. C. Cabrele (Institut für Organische Chemie, Universität Regensburg) in einem Syrol *Peptide Synthesizer* von MultiSynTech (Witten) nach der Fmoc-Methode durchgeführt.

3.4 *Analytische Methoden*

3.4.1 **SDS-Polyacrylamid Gelelektrophorese (SDS-PAGE)**

Proteine werden durch Bindung des anionischen Detergenz Natriumdodecylsulfat (SDS) denaturiert und negativ geladen. Dabei binden durchschnittlich 1,4 Aminosäurereste ein Molekül SDS. Im SDS-Protein-Komplex kann die Eigenladung des Proteins neben der hohen negativen Ladung des Detergenz vernachlässigt werden. Daher haben SDS-beladene Proteine identische Ladungs-Masse-Verhältnisse. In SDS-haltigen Polyacrylamidgelen wandern Proteine bei angelegter Spannung mit Geschwindigkeiten, die dem Logarithmus ihrer Masse umgekehrt proportional sind (Laemmli, 1970) zur Anode und werden somit allein aufgrund ihres Molekulargewichtes aufgetrennt. Das macht eine Größenabschätzung von Peptiden möglich. Dabei sinkt mit der Konzentration des Acrylamid und des Quervernetzers Tetramethylethylendiamin (TEMED) die Maschengröße.

Die Proben wurden 1:1 mit 2x SDS- Probenpuffer versetzt, 5 min bei 95 °C inkubiert und abzentrifugiert (Eppendorf 5415 D, 13.000 Upm, 1 min). Von jeder Probe wurden 15 – 30 µl auf ein Gel aufgetragen.

3.4.1.1 **Tris-Glycin-HCl PAGE**

Tris-Glycin-HCl PAGE wurde standardmäßig zur Analyse von Proteinen mit einem Molekulargewicht von über 14 kDa verwendet. Glycin ist hier das Terminationsion. Die genaue Zusammensetzung der Gele ist in Tabelle 3 aufgeführt. Es wurde eine Spannung von 45 mA angelegt, die Dauer der Elektrophorese betrug 30 min.

Tabelle 3: Zusammensetzung von Tris-Glycin-Gelen (12,5 bzw. 15% Polyacrylamid) für SDS-PAGE (Angaben in ml).

	Trenngel		Sammelgel	
Polyacrylamidkonz.	12,5%	15%	4%	4%
Anzahl der Gele	(13 Stück)	(5 Stück)	(13 Stück)	(5 Stück)
Trenngelpuffer	19,50	18,75	---	---
Sammelgelpuffer	---	---	7,38	1,87
Acrylamid/Bisacrylamid (37,5:1) 40%	26,20	18,75	5,90	1,405
H₂O bidest.	31,58	12,25	15,95	11,64
TEMED	0,089	0,03	0,029	0,025
10% APS	0,195	0,5	0,089	0,15

3.4.1.2 Tris-Tricin PAGE (Schägger & von Jagow, 1987)

Denaturierende Tris-Tricin-Polyacrylamidgele enthalten eine höhere Konzentration Polyacrylamid. Dadurch können Proteine mit einem Molekulargewicht von 1 bis 100 kDa aufgetrennt werden (Schägger & von Jagow, 1987). In diesem Falle wird Tricin anstelle von Glycin das Terminationsion, da sich Tricin durch eine höhere Auftrennungskapazität auszeichnet. Die Zusammensetzung der Gele ist in Tabelle 4 gezeigt. Tris-Tricin-Laufpuffer I wurde in die obere Elektrophoresekammer gefüllt, Tris-Tricin-Laufpuffer II in die untere. Die Laufzeit betrug ca. 1 h bei einer Spannung von 50 mA.

Tabelle 4: Zusammensetzung von fünf Tris-Tricin-Gelen (20% Polyacrylamid) für SDS-PAGE (Angaben in ml).

	Trenngel	Sammelgel
	20 % (5 Stück)	4 % (5 Stück)
3x Gelpuffer	13,0	4,1
Acrylamid/Bisacrylamid (37,5:1) 40%	19,5	1,6
H₂O bidest.	6,5	10,4
TEMED	0,013	0,013
APS (250 mg/ml)	0,13	0,13

3.4.2 SDS-Gelfärbung

3.4.2.1 Coomassie-Färbung von SDS-Gelen

Denaturierende Polyacrylamidgele wurden nach dem Lauf zur Anfärbung der Proteinbanden 15 min mit SDS-PAGE Färbelösung inkubiert. Durch mehrmaliges Aufkochen in Wasser (Mikrowellenherd, 900 W, 2 min) wurde der unspezifisch gebundene Farbstoff wieder entfernt.

3.4.2.2 Silberfärbung von SDS-Gelen (Heukeshoven & Dernick, 1988)

Im Vergleich zur Coomassie Färbung ist die Silberfärbung 100 bis 1000 mal sensitiver. Dadurch ist es möglich 0,1 – 1,0 ng eines Proteins in einer einzelnen Bande nachzuweisen. Der Färbungsprozess beruht auf den verschiedenen Oxidationszuständen, die die Silberionen, die an Seitenketten der Aminosäurenreste gebunden sind, annehmen können (Merril *et al.*, 1984).

Nach der Elektrophorese wurde das Polyacrylamidgel 30 min in Fixierlösung I, anschließend 30 – 60 min in Fixierlösung II unter leichtem Schütteln inkubiert. Nachdem dreimal 10 min mit bidestillierten Wasser gewaschen wurde, wurde das Gel in einer 0,1%igen Silbernitratlösung 30 min inkubiert. Die Banden wurden durch die Entwicklungslösung sichtbar gemacht. Dafür wurde das Gel unter leichtem Schütteln so lange in der Entwicklungslösung inkubiert bis sich der gewünschte Kontrast eingestellt hatte. Die Reaktion wurde durch Inkubation für einige Minuten in 1% Essigsäure gestoppt. Anschließend wurde das Gel mit bidestillierten Wasser gewaschen.

3.4.3 Western Blot und Immunodetektion

Der *Western Blot* dient der spezifischen Identifizierung und Quantifizierung von Proteinen in komplexen Proteingemischen. Dabei können einzelne Proteine in einer

Gesamt-Zellprotein-Präparation oder in Gewebe-Lysaten nachgewiesen werden. Nach Auftrennung der Proteine durch SDS-PAGE (3.4.1) werden sie auf eine Membran transferiert, dadurch immobilisiert und mit einem spezifischen mit einem Marker versehenen Antikörper inkubiert (Towbin *et al.*, 1979). Durch eine Nachweisreaktion (z.B. enzymatisch durch eine Phosphatase-Reaktion), die mit Hilfe des an den Antikörper gebundenen Markers ermöglicht wird, erfolgt die Detektion des Proteins.

Im Rahmen dieser Arbeit wurde der *Western Blot* vor allem nach schwachen Proteinexpressionen zum Nachweis der mit einem His₆-Tag versehenen Proteine eingesetzt. Nach Auftrennung des Proteingemisches mittels SDS-PAGE (3.4.1) erfolgte der *blot* auf eine PVDF Membran bei 2,5 mA/cm² in 60 Minuten in Transferpuffer im *semi-dry* Verfahren. Die Membran wurde zuvor für wenige Sekunden in 100 % Methanol aktiviert, in H₂O bidest. gewaschen und 5 min in Transferpuffer inkubiert. Im Folgenden der Aufbau einer Blot-Apparatur (Schichtung von unten nach oben): Anode, 3 Lagen auf Gelgröße zugeschnittenes und in Transferpuffer getränktes Whatman-Papier, PVDF Membran, SDS-Gel, 3 Lagen auf Gelgröße zugeschnittenes und in Transferpuffer getränktes Whatman-Papier, Kathode.

Nach Beendigung des Blots wurde die Membran reversibel mit Ponceau S Lösung angefärbt, um den Transfer der Proteine zu überprüfen und die Banden des Proteinstandards markieren zu können. Nach Entfärbung durch mehrmaliges Schwenken in 1x PBS erfolgte das Blockieren freier Bindestellen auf der Membran durch Inkubation in 20 ml Blockpuffer für 1 Stunde. Nach anschließender einstündiger Inkubation in 20 ml, den Antikörper (Anti-His₆-Peroxidase, Roche) in einer 1:500 Verdünnung enthaltenden, Blockpuffer wurde die Membran dreimal 10 min in je 20 ml PBS-T gewaschen. Die enzymatische Reaktion erfolgte durch kurzes Schwenken in ca. 7 ml TMB Membrane Peroxidase Substrat (KPL Inc., Gaithersburg, MD, USA). Durch Schwenken in bidestillierten Wasser wurde die Reaktion gestoppt, sobald Banden sichtbar wurden. Alle Arbeitsschritte erfolgten bei Raumtemperatur.

3.4.4 Konzentrationsbestimmung von Proteinen

3.4.4.1 Konzentrationsbestimmung nach Bradford (1976)

Der Bradford-Test dient der quantitativen Bestimmung des Proteingehalts von wässrigen Lösungen. Der in der Färbelösung enthaltene Farbstoff Coomassie Brilliant Blue G250 bindet an Proteine, wodurch sich sein Absorptionsmaximum von 465 nm nach 595 nm verschiebt (Bradford, 1976). Die Bindung erfolgt vorwiegend an kationische, nichtpolare und hydrophobe Reste. Bei einer Konzentration von 0,2 - 1,4 mg Protein pro ml Lösung ist die Absorption bei 595 nm annähernd linear. Für den Test wurden Proteinlösung und Bradford-Reagenz im Verhältnis 1:5 (v/v) gemischt. Nach 5 min Inkubation wurde die Absorption bei 595 nm gemessen. Als Referenz diente eine gleich behandelte Mischung von Reagenz und Puffer.

Anhand einer Kalibrierungskurve mit Rinderserumalbumin wurde die Proteinmenge nach folgender Gleichung bestimmt:

$$c = \frac{(A_{595} - b)}{m}$$

c: Proteinkonzentration [mg/ml];

A_{595} : Absorption bei 595 nm

b: y-Achsenabschnitt der Kalibrierungsgeraden

m: Steigung der Kalibrierungsgeraden

Formel 5: Bestimmung der Proteinkonzentration nach Bradford (1976).

3.4.4.2 Konzentrationsbestimmung mittels Absorptionsspektroskopie

Die aromatischen Aminosäuren Tryptophan und Tyrosin, sowie Cystine (Disulfidbrücken) absorbieren UV-Licht im Bereich von 250 – 300 nm. Nach Pace *et al.* (1995) kann aus der Aminosäuresequenz der molare Extinktionskoeffizient bei einer Wellenlänge von 280 nm (ϵ_{280}) berechnet werden:

$$\varepsilon_{280} = \sum \text{Trp} \times 5500 + \sum \text{Tyr} \times 1490 + \sum \text{Cystine} \times 125$$

Formel 6: Bestimmung des molaren Extinktionskoeffizienten.

Mit dessen Hilfe kann bei bekanntem Molekulargewicht der spezifische Extinktionskoeffizient ($^{0,1\%}A_{280}$) ermittelt werden:

$$^{0,1\%}A_{280} = \frac{\varepsilon_{280}}{M}$$

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

$^{0,1\%}A_{280}$: spezifischer Extinktionskoeffizient [$\text{cm}^2\text{mg}^{-1}$]

M: Molare Masse [g/mol]

Formel 7: Bestimmung des spezifischen Extinktionskoeffizienten.

Durch Einsetzen in das Lambert-Beersche-Gesetz kann somit die Proteinkonzentration (c) berechnet werden:

$$c = \frac{A_{280}}{^{0,1\%}A_{280} \times d}$$

c: Konzentration [mg/ml]

A_{280} : Absorption bei 280 nm

d: Schichtdicke der Küvette [cm]

$^{0,1\%}A_{280}$: spezifischer Extinktionskoeffizient [$\text{cm}^2\text{mg}^{-1}$]

Formel 8: Bestimmung der Proteinkonzentration mittels Absorptionsspektroskopie.

Um Informationen über den Zustand der Proteinlösung zu erhalten, wurden Spektren von 240 – 350 nm aufgenommen, wobei sich das Absorptionsmaximum eines Proteins bei 278 nm befindet. Das Verhältnis der Absorption von Protein und Nukleinsäuren (Quotient A_{280}/A_{250}) sollte mindestens 1,8 betragen und es sollte keine Absorption oberhalb von 300 nm (Hinweis auf Aggregatbildung) auftreten, da dies zu einer Verfälschung des Ergebnisses durch Lichtstreuung führt. 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.

3.4.5 Analytische Größenausschlusschromatographie (SEC)

Mit Hilfe der analytischen Gelfiltration kann das apparente Molekulargewicht von Proteinen und damit, bei bekanntem Molekulargewicht des Monomers, der Oligomerisierungszustand bestimmt werden. Bei globulären Proteinen besteht ein linearer Zusammenhang zwischen dem Logarithmus des Molekulargewichts und dem Elutionsvolumen. Das apparente Molekulargewicht eines Proteins kann daher durch Vergleich des Elutionsvolumens mit den Elutionsvolumina von Standard-Proteinen bestimmt werden. Weicht die Gestalt des Proteins von der globulären Form ab, so können sich jedoch erhebliche Differenzen zum tatsächlichen Molekulargewicht ergeben.

Zur Erstellung einer Kalibrierungsgeraden wurde eine analytische Superdex 75 HR 10/30 Säule (GE Healthcare, SV 24 ml) mit einem SV Gelfiltrationspuffer (50 mM KP, 300 mM KCl, pH 7,5) äquilibriert. Anschließend wurden die Proteine des *Low Molecular Weight* (LMW)-Gelfiltrationskalibrierungskits (2.3) nacheinander nach Anweisung des Herstellers auf die Säule aufgetragen und bei einer Flussrate von 0,5 ml/min und maximalem Rückdruck von 1,7 MPa eluiert. Die Kalibrierungsgerade wurde durch Auftragen des Logarithmus des theoretischen Molekulargewichts gegen die Elutionszeit erstellt (siehe Anhang). Nach der Kalibrierung wurde die Proteinprobe über den Probenloop mit einem Volumen von nicht mehr als 200 µl auf die Säule aufgetragen und unter gleichen Bedingungen wie die Standardproteine eluiert. Die Detektion erfolgte durch Messung der Absorption bei 260 nm und 280 nm. Das apparente Molekulargewicht wurde durch Vergleich mit der Kalibrierungsgeraden ermittelt.

3.4.6 Analytische Ultrazentrifugation

3.4.6.1 Sedimentationsgeschwindigkeitsläufe

Mittels analytischer Ultrazentrifugation kann der Sedimentationskoeffizient bestimmt werden, welcher für jedes Protein, abhängig von seinem Molekulargewicht, seiner Dichte und seiner Form, charakteristisch ist. Dazu wird die Ultrazentrifuge mit verhältnismäßig hohen Umdrehungszahlen betrieben, wodurch Proteine in den radial

ausgerichteten Sektorzellen bis zum Zellenboden wandern. Im Verlauf der Zentrifugation kommt es zur Ausbildung einer Trennlinie zwischen reinem Lösungsmittel ohne Makromoleküle und demjenigen Lösungsmittelanteil, in dem noch sedimentierende Proteine enthalten sind. Diese Grenzlinie wandert mit der Zeit, wobei ihre Wanderungsgeschwindigkeit die Sedimentationsgeschwindigkeit der Makromoleküle widerspiegelt. Die Sedimentation wird in regelmäßigen Zeitabständen über die Messung des Absorptionsverlaufes in der Sektorzelle verfolgt.

Der Sedimentationskoeffizient s entspricht dem Quotienten aus der Sedimentationsgeschwindigkeit und der Zentrifugalbeschleunigung:

$$s = \frac{\text{Sedimentationsgeschwindigkeit}}{\text{Zentrifugalbeschleunigung}} = \frac{dr/dt}{\omega^2 r} [S]$$

1S (Svedberg-Einheit) = $1 \cdot 10^{-13}$ s

r : Abstand des Meßpunktes zum Rotationszentrum [cm]

t : Zeit [s]

ω : Winkelgeschwindigkeit ($2 \cdot \pi \cdot \text{Upm}/60$) [s^{-1}]

Formel 9: Berechnung des Sedimentationskoeffizienten.

Die logarithmische Auftragung des Abstandes r gegen die Zeit t ergibt eine Gerade, aus deren Steigung der apparente Sedimentationskoeffizient (s_{app}) berechnet werden kann:

$$s_{app} = \frac{2,303}{\omega^2} \cdot \frac{d \log r}{dt}$$

Formel 10: Berechnung des apparenten Sedimentationskoeffizienten s_{app} .

Der Wert für den Quotienten $2,303/\omega^2$ lässt sich für die jeweilige Drehzahl aus Tabellen entnehmen, z.B. Chervenka (1973).

Die Umrechnung von s_{app} auf Standardbedingungen ($T = 293,15$ K, Lösungsmittel = Wasser) erfolgt nach Formel 11:

$$s_{20,w} = s_{app} \cdot \frac{\eta_{T,w}}{\eta_{20,w}} \cdot \frac{\eta_{T,LM}}{\eta_{T,w}} \cdot \frac{1 - \bar{v}_{20,w} \cdot \rho_{20,w}}{1 - \bar{v}_{T,LM} \cdot \rho_{T,LM}}$$

T: Temperatur [K]

η : relative Viskosität

\bar{v} : partielles spezifisches Volumen des Proteins [cm³/g]

ρ : Dichte des Lösungsmittels [g/cm³]

Formel 11: Umrechnung von s_{app} auf Standardbedingungen ($s_{20,w}$).

Die Werte für die Temperaturabhängigkeiten der Viskosität, des partiellen spezifischen Volumens und der Lösungsmitteldichte lassen sich ebenfalls aus Tabellen entnehmen.

$\frac{\eta_{T,LM}}{\eta_{T,w}}$ wurde als 1,015 und $\frac{1 - \bar{v}_{20,w} \cdot \rho_{20,w}}{1 - \bar{v}_{T,LM} \cdot \rho_{T,LM}}$ als 1 angenommen. Die Durchführung

erfolgte an einer Beckman Spinco Model E in einem An-G Rotor unter Verwendung von 12 mm Doppelsektorzellen mit Saphir-Fenstern. Die Zentrifugation erfolgte bei Raumtemperatur. Für die Bestimmung von T wurde die Temperatur im Gerät vor und nach dem Lauf bestimmt und gemittelt. Als Umdrehungsgeschwindigkeit wurden 52000 Upm gewählt. Die Detektion erfolgte bei 277 nm. Alle Proben wurden zuvor in den gleichen Puffer (50 mM KP pH 7,5) gebracht, der auch als Referenz diente, und zentrifugiert (Eppendorf 5415 D, 13000 Upm, 20 Minuten, 4 °C). Die Proteinkonzentration lag bei 0,2 mg/ml.

3.4.6.2 Gleichgewichtsläufe

Die Zentrifugation bei niedrigeren Umdrehungszahlen eröffnet die Möglichkeit, über das Sedimentationsgleichgewicht das Molekulargewicht zu bestimmen. In diesem Fall wurde die Methode des „*High-Speed Sedimentation Equilibrium* (HSSE)“ verwendet. Dabei kommt es nach genügend langer Laufzeit zu einem Gleichgewicht zwischen zwei entgegengesetzten Prozessen, der Sedimentation im Zentrifugalfeld und der Rückdiffusion der Teilchen. Analog zum Sedimentationsgeschwindigkeitslauf (3.4.6.1) wird der Absorptionsverlauf in der Sektorzelle in regelmäßigen Zeitabständen verfolgt. Abbildung 20 zeigt schematisch die Einstellung des Gleichgewichts.

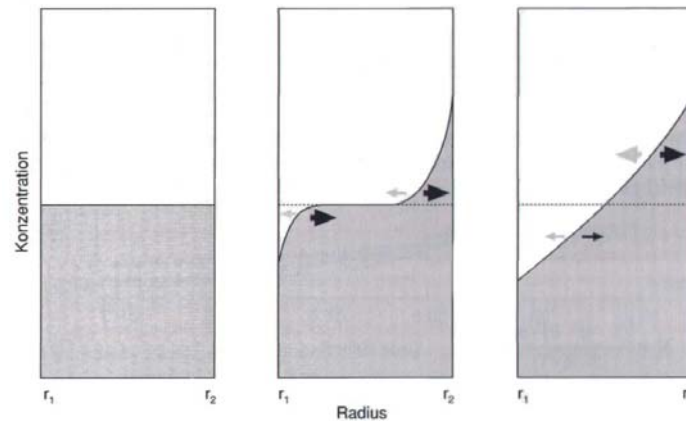


Abbildung 20: Zeitabhängige Änderung des Konzentrationsgradienten im Sedimentationsgleichgewichtsexperiment

Links: Beginn des Experiments; Mitte: In der ersten Phase sedimentiert gelöstes Material verstärkt zu Boden. Die Sedimentation überwiegt gegenüber der Diffusion (ungleich große Pfeile); rechts: Das Sedimentationsgleichgewicht ist (annähernd) erreicht (gleich große Pfeile). Abbildung aus Holtzhauer (1996).

Im Gleichgewichtszustand gilt:

$$M = \frac{2 \cdot RT}{(1 - \bar{v}_2 \rho_1) \cdot \omega^2} \cdot \frac{d \ln c_2}{dr^2}$$

M: molare Masse [g/mol]

R: allgemeine Gaskonstante = 8,31 J·mol⁻¹·K⁻¹

T: absolute Temperatur [K]

c₂: Konzentration des gelösten Proteins [mg/ml]

ρ₁: Massendichte des Lösungsmittels [g/ml]

\bar{v}_2 : partielles spezifisches Volumen des Proteins [ml/g]

r: Abstand des Meßpunktes zum Rotationszentrum [cm]

ω: Winkelgeschwindigkeit (2 * π * rpm/60) [s⁻¹]

Formel 12: Berechnung des Molekulargewichts aus Sedimentationsgleichgewichtsläufen.

Wie bei den Sedimentationsgeschwindigkeitsläufen erfolgte die Durchführung an einer Beckman Spinco Model E in einem An-G Rotor unter Verwendung von 12 mm Doppelsektorzellen mit Saphir-Fenstern bei Raumtemperatur, jedoch bei halber Füllhöhe. Auch hier wurde für die Bestimmung von T die Temperatur im Gerät vor und nach dem Lauf bestimmt und gemittelt. Die Läufe wurden bei 16000 Upm durchgeführt, die Detektion erfolgte bei 277 nm. Analog zu den Sedimentationsgeschwindigkeitsläufen wurden alle Proben zuvor in den gleichen Puffer (50 mM KP pH 7,5)

gebracht, der auch als Referenz diente, und zentrifugiert (Eppendorf 5415 D, 13000 Upm, 20 Minuten, 4 °C). Den Proben wurde 100 µM Pefabloc SC, 5 mM EDTA (Endkonzentration) zugegeben, um proteolytischen Abbau zu verhindern. Die Proteinkonzentration lag ebenfalls bei 0,2 mg/ml.

Die Auswertung erfolgte nach der *Meniscus-Depletion-Sedimentation-Equilibrium*-Methode (Yphantis, 1964) mit Hilfe eines selbstgeschriebenen Makros für das Programm Sigma-Plot (freundliche Gabe von Klaus-Jürgen Tiefenbach, (Tiefenbach, 1997)). Mit diesem wird die Basislinie in einem begrenzten Bereich um den Nullpunkt so lange variiert, bis die Punkte der Auftragung $\ln c$ gegen r^2 weitgehend auf einer idealen Geraden liegen. Das partielle spezifische Volumen wurde nach Cohn und Edsall aus der Aminosäuresequenz des jeweiligen Proteins berechnet (Cohn & Edsall, 1943; Durchschlag, 1986) und betrug 0,751 ml/g (HisF_N1), 0,7511 ml/g (HisF_N1_E46Q), 0,752 ml/g (HisF_N1_R5A_E46Q), 0,744 ml/g (HisF_C1_A128C_Y143H), 0,7456 ml/g (HisF_A128C_Y143H_E167Q) und 0,751 ml/g (HisF_N1N1).

3.4.7 Massenspektrometrie

Unter Massenspektrometrie versteht man eine Analysetechnik zur Bestimmung der Molekülmasse freier Ionen im Hochvakuum. Ein Massenspektrometer besteht aus einer Ionenquelle, in der aus einer Substanzprobe ein Strahl gasförmiger Ionen erzeugt wird, einem Massenanalysator, welcher die Ionen hinsichtlich des Masse/Ladungs-Quotienten (m/z) auftrennt, sowie einem Detektor, der ein Massenspektrum liefert. Aus diesem kann abgelesen werden, welche Ionen in welchen relativen Mengen gebildet wurden. Die massenspektrometrische Analyse der Proteine wurde von der Arbeitsgruppe Prof. Dr. Deutzmann (Lehrstuhl Biochemie I, Universität Regensburg) mit Hilfe einer Elektrospray-Ionisations (ESI)- und einer matrix-unterstützten Laserdesorptions/Ionisations (MALDI)-Anlage durchgeführt.

3.4.7.1 Elektrospray Ionisations Massenspektrometrie (ESI-MS)

Bei der Elektrospray-Ionisations-Massensspektrometrie (ESI-MS) kommt es zur Dispersion einer wässrigen Proteinlösung in sehr viele kleine geladene Tröpfchen (*Coulomb* Explosion) mit Hilfe eines elektrostatisch geladenen Feldes. Der Theorie

nach werden entweder Tröpfchen erzeugt, die nur noch ein Makromolekül enthalten und schließlich desolvatisiert werden oder es kommt zu einer Emission von freien Molekülonen aus sehr kleinen elektrisch geladenen Tröpfchen. Die Massenbestimmung beginnt mithilfe eines Quadrupols im Hochvakuum. Der Quadrupol besteht aus vier parallelen stabförmigen Metallelektroden, in deren Mitte ein kombiniertes Wechsel- und Gleichspannungsfeld aufgebaut wird. Dieses Feld können nur Ionen mit einem bestimmten m/z -Verhältnis auf einer stabilen oszillierenden Bahn durchlaufen, um am Ende auf einen Detektor zu treffen. Alle anderen Ionen mit einem abweichenden m/z -Verhältnis fliegen auf instabilen Bahnen und werden durch Kollision mit den Metallstäben gestoppt. Durch Variation des Feldes kann das m/z -Verhältnis, bei dem ein Durchqueren des Feldes möglich ist, festgelegt werden. Da es während des ESI-Prozesses zur Bildung von mehrfach geladenen Ionen kommt, findet man für Proteine entsprechend ihrer Masse eine Serie von Ionensignalen, die sich um jeweils eine Ladung unterscheiden. Das Molekulargewicht eines Molekülonen lässt sich aus der Anzahl der Ladungen und über das Masse-Ladungs-Verhältnis von im Spektrum aufeinanderfolgenden Molekülonen berechnen. Die Auswertung des Spektrums und die Berechnung der Molekulargewichte wurden von der Arbeitsgruppe Prof. Dr. Deutzmann durchgeführt. Da Zusätze wie Puffer, Salze und Detergenzien den Elektrospray Prozess stark stören, wurde die Probe vor der Massenspektrometrie über *Reversed Phase* (RP)-HPLC gereinigt.

3.4.7.2 Matrix-assisted Laserdesorptions/Ionisations-Massenspektrometrie (MALDI-MS)

Die matrix-assisted Laserdesorptions/Ionisations-Massenspektrometrie (MALDI-MS) wurde mit Hilfe einer Anlage von Applied Biosciences durchgeführt. Bei dieser Methode wird die Probe mit einem 1000 – 10000 fachen Überschuss geeigneter Matrixmoleküle kokristallisiert. Bei der Matrix handelt es sich meist um kleine organische Moleküle, die bei der eingestrahlten Laserwellenlänge eine starke Absorption zeigen. Der Beschuss der Probe nach Einbettung hat im Vergleich zum direkten Beschuss höhere Intensitäten der Analyt-Molekül-Ionen und die Verringerung des Anteils an Fragmentationen zur Folge, ist also matrix-unterstützt

(*matrix assisted*, MA). Im Hochvakuum der Ionenquelle des Massenspektrometers wird die kristalline Oberfläche der präparierten Probe einem intensiven Impuls kurzwelliger Laserstrahlen von wenigen Nanosekunden Dauer ausgesetzt (*Laserdesorption/Ionisation*, LDI). Die Einkopplung der für die Ionen notwendigen Energie erfolgt per UV-Bestrahlung über resonante elektronische Anregung der Matrixmoleküle in das π -Elektronensystem aromatischer Verbindungen. Theoretische Überlegungen lassen den Schluss zu, dass die zunächst in den Matrixmolekülen gespeicherte elektronische Anregungsenergie in extrem kurzen Zeiten in das Festkörnergitter relaxiert und dort eine starke Störung und Ausdehnung bewirkt. Es erfolgt dann, weit vor Erreichen eines thermischen Gleichgewichts, ein Phasenübergang der explosiv einen Teil der Festkörperoberfläche auflöst. Dabei werden neben Matrixmolekülen auch Probenmoleküle in die Gasphase freigesetzt. Der Probe gegenüber befindet sich im Abstand von wenigen Millimetern eine Elektrode, durch die ein elektrostatisches Feld von einigen 100 bis einigen 1000 V/mm erzeugt wird. Im Rahmen dieser Arbeit wurde eine Kathode verwendet, die positiv geladene Ionen von der Probenoberfläche in Richtung des Analysators beschleunigt. Die innerhalb des kurzen Laserimpulses gebildeten Ionen werden in der Quelle durch das elektrostatische Feld auf einige kilo-Elektronenvolt (keV) beschleunigt und durchlaufen nach dem Verlassen der Quelle eine feldfreie Driftstrecke, in der sie nach ihrem Masse/Ladungs- (m/z) Verhältnis aufgetrennt werden. Die Massenbestimmung erfolgt mithilfe eines Flugzeitmassenspektrometers in einem Hochvakuum über eine sehr genaue elektronische Messung der Zeit, die zwischen dem Start der Ionen in der Quelle bis zum Eintreffen am Detektor vergeht (TOF-Analysator, vom engl. *time of flight*). Da Ionen mit unterschiedlichen m/z -Werten – bei gleicher kinetischer Energie – in der Beschleunigungsstrecke der Quelle auf unterschiedliche Geschwindigkeiten gebracht wurden, kann bei bekannter Beschleunigungsspannung und Flugstrecke der Ionen in der feldfreien Driftstrecke durch die Messung der Flugzeit das m/z -Verhältnis bestimmt werden. In einem Spektrum der MALDI-MS mit Molekülen mit großem Molekulargewicht wie Proteinen, erhält man als Hauptpeak das einfach positiv und zweifach positiv geladene Molekülion $(M+H)^+$ und $(M+H)^{2+}$.

3.4.8 Fluoreszenzemissionsspektroskopie

Absorbiert ein Molekül Licht bestimmter Wellenlänge, so wird es von einem energetischen Grundzustand in einen angeregten Zustand transferiert. Die Rückkehr in den Grundzustand kann u.a. 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 größeren Wellenlängen, also einer Rotverschiebung entspricht.

Die Fluoreszenz von Proteinen beruht im Wesentlichen auf der aromatischen Aminosäure Tryptophan. Eine geringe Rolle spielt das Tyrosin. Aromatische Seitenketten haben bei einer Wellenlänge von etwa 280 nm ein Absorptionsmaximum. Nach Anregung bei 295 nm ist die Fluoreszenz ausschließlich auf Tryptophane zurückzuführen.

Die Fluoreszenz von Tryptophan zeigt eine Abhängigkeit von der Polarität der Umgebung. Eine unpolare Umgebung des Tryptophans führt zu einem Fluoreszenzemissionsmaximum bei ca. 320 – 330 nm. Im Vergleich dazu ist es in stark polarer Umgebung rotverschoben und findet sich bei einer Wellenlänge von ca. 350 nm (Schmid, 1997). Über die Messung von Fluoreszenzspektren kann also eine Aussage über die Lage eines Tryptophans in einem Protein gemacht und daraus wiederum auf seinen Faltungszustand geschlossen werden. Die Messungen wurden am Fluoreszenz-Spektrophotometer Cary Eclipse (Varian) durchgeführt. In einer Quartzküvette wurde die Proteinlösung mit einer Konzentration von 5 μ M in 50 mM KP, pH 7,5 vorgelegt. Nach Anregung bei 280 nm bzw. 295 nm wurde ein Emissionsspektrum von 300 nm bis 400 nm aufgenommen. Dabei wurden je zehn Messungen aufgenommen und daraus der Mittelwert gebildet.

3.4.9 Circular-Dichroismus-Spektroskopie

Die Circular-Dichroismus (CD)-Spektroskopie eignet sich besonders zur Untersuchung von Proteinkonformationen. Bei dieser Methode wird die ungleiche Absorption von links- und rechtsdrehendem polarisierten Licht optisch aktiver Substanzen ausgenutzt. Dabei sind die Chromophore entweder selbst asymmetrisch

oder liegen in einer asymmetrischen Umgebung vor. Die Absorption im Wellenlängenbereich von 170 – 250 nm (fern-UV CD) wird hauptsächlich von den Peptidbindungen bestimmt, sodass dieser Teil des Spektrums Aussagen über die Sekundärstruktur des Proteins ermöglicht. Die verschiedenen Sekundärstrukturelemente wie α -Helix, β -Faltblatt oder *random-coil* besitzen sehr distinkte fern-UV CD-Spektren (Johnson, 1990). Aus diesen verschiedenen Signalen setzt sich ein Proteinspektrum zusammen (Stark, 1965), wobei α -Helices die anderen Strukturelemente in der Signalintensität dominieren. Die Absorption im Bereich von 250 – 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.

Die CD-Messungen wurden an einem Jasco J-815 Spektralphotometer bei 25 °C durchgeführt. Die Schichtdicke der Küvette betrug 0,1 cm (fern-UV CD) oder 1 cm (nah-UV CD). Das CD-Signal, gemessen als Elliptizität Θ_{obs} , wurde in Abständen von 0,5 nm aufgenommen (Scan-Geschwindigkeit 50 nm/min, Mittelwertbildung aus fünf Einzelmessungen). Die Werte im fern-UV CD wurden als Θ_{MRW} angegeben, was der durchschnittlichen molaren Elliptizität pro Aminosäurerest entspricht. Die Umrechnung erfolgte nach Formel 13 (Schmid, 1997):

$$\Theta_{\text{MRW}} = \frac{\Theta_{\text{obs}} * \text{MRW}}{c * d} = \frac{\Theta_{\text{obs}} * \text{MW} * 100}{c * d * N_A}$$

Θ_{MRW} : durchschnittliche molare Elliptizität pro Aminosäure [deg* cm² /dmol]

Θ_{obs} : gemessene Elliptizität [mgrad]

MRW : mittleres Molekulargewicht der Aminosäuren [Da]

c : Konzentration der Proteinlösung [mg /ml]

d : Schichtdicke der Küvette [cm]

MW : Molekulargewicht des Proteins [g /mol]

N_A : Anzahl der Aminosäuren

Formel 13: Berechnung der molaren Elliptizität pro Aminosäurerest.

3.4.10 Stabilitätsuntersuchungen mittels chemischer Denaturierung

Um die Stabilität von Proteinen in Gegenwart von chaotropen Denaturierungsmitteln zu untersuchen, wurde natives Protein unterschiedlichen Harnstoff (Urea)-Konzentrationen ausgesetzt. Dazu wurden Messreihen mit gleich bleibender Proteinkonzentration aber steigender Harnstoffkonzentration hergestellt und bei 8 °C gelagert (Verdünnungspuffer: 50 mM KP, pH 7,5; Auffaltungspuffer: 10 M Harnstoff, 50 mM KP, pH 7,5). Die Messungen am Fluoreszenz- bzw. CD-Spektrometer wurden bei Raumtemperatur durchgeführt. Die Einstellung des Gleichgewichtes wurde durch mehrere Messungen zu unterschiedlichen Zeitpunkten verfolgt. Das Gleichgewicht galt als erreicht, wenn die Messwerte an zwei aufeinander folgenden Zeitpunkten unverändert blieben. Fluorimetrische Messungen dienten zur Verfolgung der Entfaltung der Tertiärstruktur. Dazu wurden zunächst Emissionsspektren von nativem und vollständig denaturiertem Protein aufgenommen und die Wellenlänge ermittelt, bei der die Emissionsdifferenz zwischen beiden Spektren am größten war (diese lag meist bei 320 nm). Bei eben dieser Wellenlänge wurde anschließend die Fluoreszenz über 60 s nach Anregung bei 280 nm gemessen. Um auch die Entfaltung der Sekundärstruktur untersuchen zu können, wurden die Proben im fern-UV Bereich bei 220 nm CD-spektroskopisch vermessen. Eine Messung dauerte 60 s (ein Messwert pro Sekunde), anschließend wurde der Mittelwert gebildet. Die Messwerte wurden als Funktion der Konzentration an Denaturierungsmittel aufgetragen.

Um die Auffaltungen der Tertiärstruktur- und Sekundärstruktur vergleichen zu können, wurden die Messpunkte normiert. Hierfür wurden die Signalwerte y_N von nativem (0 M Harnstoff) bzw. y_U vollständig denaturiertem (8 M Harnstoff) Protein als 0 %ige bzw. 100 %ige Auffaltung des Proteins definiert. Jeder Signalwert y der Messreihe lässt sich somit in den Anteil ungefaltetes Protein F_U von 0 (entspricht y_N) bis 1 (entspricht y_U) umrechnen.

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

Formel 14: Berechnung des Anteils an ungefaltetem Protein bei der chemischen Auffaltung und die damit verbundene Normierung auf 0 – 1.

Sind die über Fluoreszenz- bzw. CD-Spektroskopie verfolgten und normierten Auffaltungskurven deckungsgleich, kann die Denaturierung zumindest in erster Näherung mit dem Zwei-Zustands-Modell beschrieben werden. Dies erlaubt es theoretisch, durch Extrapolation der Übergangsregion der Auffaltung auf 0 M Denaturierungsmittel, die freie Energie der Auffaltung in Wasser ($\Delta G(\text{H}_2\text{O})$) zu berechnen (Pace & Scholtz, 1997). Dies war allerdings bei keiner der in dieser Arbeit gemessenen Auffaltungskurven möglich. Die Konzentration an Denaturierungsmittel, bei welcher das Protein zu gleichen Teilen gefaltet und entfaltet vorliegt, wird als $D_{1/2}$ -Wert definiert. Der $D_{1/2}$ -Wert wurde in der vorliegenden Arbeit als operationelles Maß der Protein-Stabilität verwendet.

4 Ergebnisse

4.1 Die vierfache Symmetrie des $(\beta\alpha)_8$ -barrel Enzyms HisF

Wie bereits unter 1.4.2 erörtert, findet sich innerhalb der $(\beta\alpha)_8$ -barrel Faltung von HisF eine vierfache Symmetrie, die bei Betrachtung des zentralen β -Faltblattes entlang seiner Längsachse deutlich wird. So sind vier Schichten erkennbar, die durch Aminosäurereste der inneren β -Stränge gebildet werden (Abbildung 8). Dadurch kann man ein Evolutionsmodell aufstellen, nach dem das $(\beta\alpha)_8$ -barrel durch Homooligomerisierung von vier zunächst identischen $(\beta\alpha)_2$ -Einheiten zustande gekommen ist. Auch auf Sequenzebene ist eine vierfache Symmetrie nachzuweisen. So findet man in den $\alpha\beta$ -loops von HisF (Abbildung 9) und vieler anderer $(\beta\alpha)_2$ -Module (Nagano *et al.*, 2002) ein wiederkehrendes GXD- bzw. GXG-Motiv, das nur durch Duplikation einer $(\beta\alpha)_2$ -Einheit erhalten geblieben sein kann. Einen weiteren Hinweis lieferte eine Analyse mit dem Programm HHrep (Söding *et al.*, 2006). Hier wird die Eingabesequenz zunächst in ein Hidden-Markow-Model umgewandelt und anschließend paarweise mit sich selbst verglichen. Für die Sequenz von HisF wurden vier Bereiche ermittelt, die Ähnlichkeiten zueinander aufweisen (Abbildung 10).

Unterteilt man HisF in vier $(\beta\alpha)_2$ -Einheiten (vgl. Abbildung 21), ergeben sich die folgenden Viertel: HisF_N1, HisF_N2, HisF_C1 und HisF_C2. Tabelle 5 gibt einen Überblick über ihre Größen und Grenzen.

Tabelle 5: Definition der HisF-Viertel.

Bezeichnung	Grenzen der Viertel (angegeben sind die jeweils erste und letzte Aminosäure in HisF)	Anzahl der Aminosäuren	$M_{\text{kalk.}}$ [Da]	Anzahl der Basenpaare
HisF_N1: $(\beta\alpha)_{1-2}$	1 – 74	74	8315	222
HisF_N2: $(\beta\alpha)_{3-4}$	75 – 122	48	4988	144
HisF_C1: $(\beta\alpha)_{5-6}$	123 – 196	74	8279	222
HisF_C2: $(\beta\alpha)_{7-8}$	197 – 254	58	6308	174

4.1.1 Struktur- und Sequenzähnlichkeiten der vier HisF-Viertel HisF_N1, HisF_N2, HisF_C1 und HisF_C2

Zerlegt man das Enzym HisF auf diesen Grundlagen strukturell in seine $(\beta\alpha)_2$ -Einheiten (Abbildung 21A) und superpositioniert diese mit Hilfe des Programm STAMP, erhält man eine gute Überlagerung der vier Einheiten (Abbildung 21B). Lediglich die Überlänge der *loops* $\beta 1/\alpha 1$ in HisF_N1 und $\beta 5/\alpha 5$ in HisF_C1 im Vergleich zu den entsprechenden *loops* $\beta 3/\alpha 3$ in HisF_N2 und $\beta 7/\alpha 7$ in HisF_C2 führen zu deutlichen Differenzen (vgl. Abbildung 23). Mit dem Programm MAMMOTH (Ortiz *et al.*, 2002) unter <http://ub.cbm.uam.es/mammoth/pair/index3.php> wurden die Strukturausschnitte für die vier Viertel aus der HisF-Struktur (1thf) paarweise superpositioniert. Die *rmsd*-Werte dieser Analyse sind in Tabelle 6 zusammengefasst.

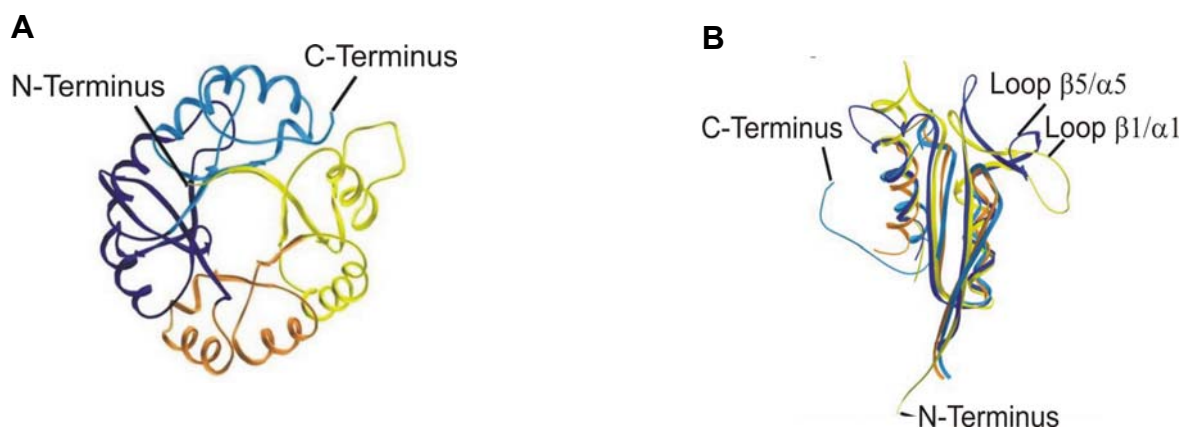


Abbildung 21: Bänderdiagramme von HisF zur Darstellung der Aufteilung in die $(\beta\alpha)_2$ -Module (HisF_N1, HisF_N2, HisF_C1 und HisF_C2) und deren Superpositionierung. (Nach Bosnali, 2003)

- (A) HisF_N1: $(\beta\alpha)_{1-2}$ (gelb), HisF_N2: $(\beta\alpha)_{3-4}$ (orange), HisF_C1: $(\beta\alpha)_{5-6}$ (dunkelblau), HisF_C2: $(\beta\alpha)_{7-8}$ (hellblau).
 (B) Gemeinsame Superpositionierung der C α -Atome aller Viertel durch STAMP

HisF wurde auch auf Aminosäuresequenzebene in die vier Viertel HisF_N1, HisF_N2, HisF_C1 und HisF_C2 aufgeteilt (vgl. Tabelle 5 und Abbildung 23). Mit diesen Fragmenten wurde ein multiples Sequenzalignment mithilfe des Programms Toffee (Notredame *et al.*, 2000; Poirot *et al.*, 2003) durchgeführt. In dem alignment werden die GXD- bzw. GXG-Motive, die jeweils in den $\alpha\beta$ -loops innerhalb eines

jeden $(\beta\alpha)_2$ -Moduls liegen (vgl. Abbildung 9), als konservierte Bereiche erkannt, was ihre hohe Sequenzähnlichkeit auf Primärstrukturebene verdeutlicht (Abbildung 22).

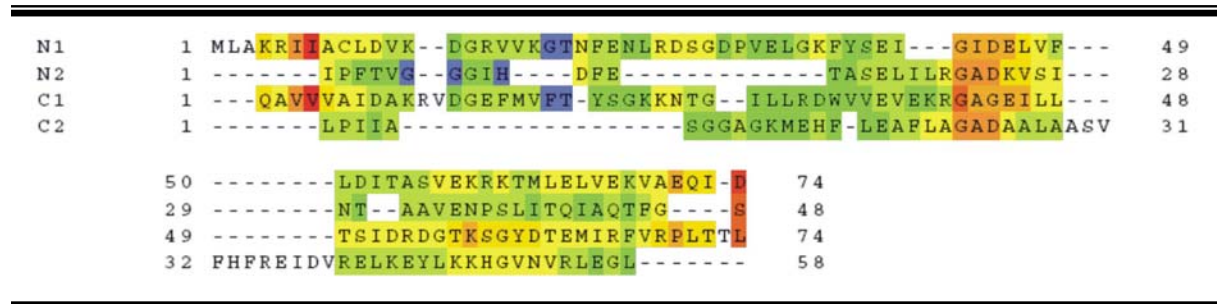


Abbildung 22: Multiples Sequenzalignment der HisF-Viertel erstellt mit dem Programm TCOffee. Multiples Sequenzalignment der HisF-Viertel (N1, N2, C1 und C2) erstellt mit dem Programm TCOffee. Die Farben geben die Ähnlichkeit der Aminosäuren an. Gelb, orange und rot spiegeln hohe, grün mittlere und blau geringe Ähnlichkeit wider. Der Bereich der vier GXD- bzw. GXG-Motive zeigt einen hohen Konservierungsgrad.

Weiterhin gibt das Programm TCOffee in einem sogenannten LOG-File die Ähnlichkeiten der alignierten Sequenzen im paarweisen Vergleich in Prozent an. Diese Ergebnisse und die *rmsd*-Werte aus der paarweisen Superpositionierung sind in Tabelle 6 zusammengefasst.

Tabelle 6: Sequenzähnlichkeiten der vier HisF-Viertel untereinander in Prozent, wie sie vom Programm TCOffee berechnet wurden (obere Zeile). *rmsd*-Werte in Å aus der paarweisen Superpositionierung der HisF-Viertel mit dem Programm MAMMOTH (untere Zeile).

	HisF_N1	HisF_N2	HisF_C1	HisF_C2
HisF_N1	-----	21	20	23
		4,0 (48 Ca)	3,56 (74 Ca)	3,12 (57 Ca)
HisF_N2	21	-----	23	18
	4,0 (48 Ca)		3,39 (48 Ca)	2,95 (48 Ca)
HisF_C1	20	23	-----	12
	3,56 (74 Ca)	3,39 (48 Ca)		3,98 (57 Ca)
HisF_C2	23	18	12	-----
	3,12 (57 Ca)	2,95 (48 Ca)	3,98 (57 Ca)	

HisF_N1 besitzt durchschnittlich ca. 20 % Ähnlichkeit zu den drei anderen Vierteln. Dies ist ein deutlicher Hinweis darauf, dass es sich bei HisF_N1 um das ursprünglichste Viertel handelt. HisF_N2 besitzt ebenfalls noch eine hohe Ähnlichkeit

zu den Vierteln der C-terminalen Hälfte. Die Viertel der C-terminalen Hälfte hingegen besitzen zueinander gerade eine Ähnlichkeit von 12 %. Die *rmsd*-Werte der vier Viertel liegen zwischen 3 bis 4 Å, wobei immer die Anzahl der C α -Atome des kleineren Viertels des jeweiligen paarweisen Vergleichs herangezogen wurde. Dies ist ein Hinweis auf die strukturelle Ähnlichkeit der vier HisF-Viertel. Um eine globalere Betrachtung der evolutionären Verwandtschaft der ($\beta\alpha$)₂-Einheiten von HisF und um zusätzliche Hinweise auf die vierfache Symmetrie zu erhalten, wurde ein strukturbasiertes Sequenzalignment von verschiedenen HisF-Sequenzen erstellt.

4.1.2 Strukturbasiertes Sequenzalignment von HisF-Sequenzen

Zur Erstellung eines strukturbasierten Sequenzalignments wurde über die Webseite des Sanger-Institutes (www.sanger.ac.uk) das vollständige Sequenzalignment aller annotierten HisF-Sequenzen (PFAM-Eintrag: PF00977) bezogen. Mit dem Programm Jalview wurde anschließend ein Stammbaum für alle im alignment enthaltenen Sequenzen erstellt. Die elf HisF-Sequenzen des Astes des Stammbaums, in welchen auch die HisF-Sequenz von *Thermotoga maritima* eingeordnet wurde, wurden mit dem Programm MAFFT (Kato et al., 2002; Kato et al., 2005) unter <http://align.bmr.kyushu-u.ac.jp/mafft/online/server/> zu einem neuen multiplen Sequenzalignment zusammengeführt. Ein strukturbasiertes Sequenzalignment wurde von drei HisF-Strukturen (HisF aus *Thermotoga maritima*, PDB-Eintrag: 1gpw Kette A; HisF aus *Thermus thermophilus*, PDB-Eintrag: 1ka9 Kette F; HisF aus *Pyrobaculum aerophilum*, PDB-Eintrag: 1h5y Kette B) mit dem Programm Salign unter <http://salilab.org/salign/> erstellt und anschließend mit dem gefilterten Sequenzalignment aller HisF-Sequenzen mit Hilfe des Programms Toffee verknüpft. Dadurch konnte für die restlichen Sequenzen eine Sekundärstrukturelementzuordnung erreicht werden. Die weitere Bearbeitung des alignments erfolgte mit dem Web-Programm ESPript (Gouet et al., 1999) unter <http://esript.ibcp.fr/ESPript/ESPript/>. In Abbildung 23 ist das Ergebnis gezeigt.

Die Sequenzen wurden, basierend auf der Einteilung von HisF aus *Thermotoga maritima*, in die vier Viertel (HisF_N1, HisF_N2, HisF_C1 und HisF_C2) eingeteilt. Dabei zeigt sich, dass die GXD- bzw. GXG-Motive (rot hinterlegt) durchweg in allen Sequenzen konserviert sind und in den $\alpha\beta$ -loops zu liegen kommen (vgl. Abbildung

9). Ebenso konserviert sind die Reste R5, E46, K99, und E167 (grün hinterlegt), die an der Ausbildung eines Salzbrückenclusters in *layer 4* beteiligt sind (vgl. Abbildung 8). Der Rest A220, dessen Funktion von R5 übernommen wird, zeigt einen geringeren Konservierungsgrad (ebenfalls grün hinterlegt). Da dieser Rest nicht zur Stabilisierung benötigt wird, kann er leichter mutiert werden.

In den Vierteln HisF_N1, HisF_N2 und HisF_C1 fällt im jeweils ersten β -Strang der entsprechenden $(\beta\alpha)_2$ -Module ein hoch konserviertes Cystein (gelb hinterlegt) an den Positionen 9 (HisF_N1), 78 (HisF_N2) und 124 (HisF_C1) auf. Im Viertel HisF_C2 wurde dieses Cystein durch die ebenfalls polare Aminosäure Serin ersetzt (ebenfalls gelb hinterlegt). Auf DNA-Ebene genügt ein einziger Nukleotidaustausch, um von dem Tripletts das für Cystein kodiert (TG (T bzw. C)) zu dem Tripletts das für Serin kodiert (AG (T bzw. C)) zu kommen. Diese Tatsache spricht für die Stabilität der Position.

Abbildung 23: Strukturbasiertes Sequenzalignment von HisF. (nächste Seite)

Aus den ersten drei Sequenzen wurde das strukturbasierte Sequenzalignment mit dem Programm Salign erstellt. Dazu wurden die Strukturen von HisF aus *Thermotoga maritima* (1gpw Kette A), *Thermus thermophilus* (1ka9 Kette F) und *Pyrobaculum aerophilum* (1h5y Kette B) herangezogen. Die restlichen zwölf Sequenzen (Namen der Organismen aus denen sie stammen sind als Kürzel angegeben und unten beschrieben) wurden aus einem Sequenzalignment aller bekannten HisF-Sequenzen (PFAM-Eintrag: PF00977) basierend auf den Ergebnissen eines Verwandtschaftsbaums, der mit dem Programm Jalview erstellt wurde, gefiltert. Anschließend wurde aus diesen Sequenzen mit dem Programm MAFFT ein multiples Sequenzalignment erstellt. Dieses und das strukturbasierte alignment wurden mit dem Programm Toffee kombiniert. Über dem Gesamtalignment sind die Sekundärstrukturelemente von HisF aus *Thermotoga maritima* gezeigt.

Das alignment wurde wie für HisF aus *Thermotoga maritima* beschrieben (vgl. Tabelle 5) in die vier Viertel HisF_N1, HisF_N2, HisF_C1 und HisF_C2 eingeteilt. Die vier GXD- bzw. GXG-Motive in den $\alpha\beta$ -loops der einzelnen $(\beta\alpha)_2$ -Einheiten (vgl. Abbildung 9) sind konserviert und rot hinterlegt. Die Reste R5, E46, K99 und E167, die an der Ausbildung eines Salzbrückenclusters in *layer 4* beteiligt sind (vgl. Abbildung 8), sind ebenfalls konserviert und grün hinterlegt. Der Rest A220, dessen Aufgabe von R5 übernommen wird, wurde im Lauf der Evolution mutiert (ebenfalls grün hinterlegt). In den Vierteln HisF_N1, HisF_N2 und HisF_C1 ist im jeweils ersten β -Strang eines jeden $(\beta\alpha)_2$ -Moduls ein Cystein konserviert (Positionen 9, 78, 124; gelb hinterlegt). Im Viertel HisF_C2 ist dieses Cystein durch die polare Aminosäure Serin ersetzt die ebenfalls konserviert an Position 201 liegt (ebenfalls gelb hinterlegt).

BUCAP: *Buchnera aphidicola* subsp. *Shizaphis graminum*

BUCAI: *Buchnera aphidicola* subsp. *Acyrtosiphon pisum*

BUCBP: *Buchnera aphidicola* subsp. *Baizongia pistaciae*

ANOGA: *Anopheles gambiae*

VIBVU: *Vibrio vulnificus*

HAEIN: *Haemophilus influenzae*

SHEON: *Shewanella oneidensis*

BLOFL: *Blochmannia floridanus*

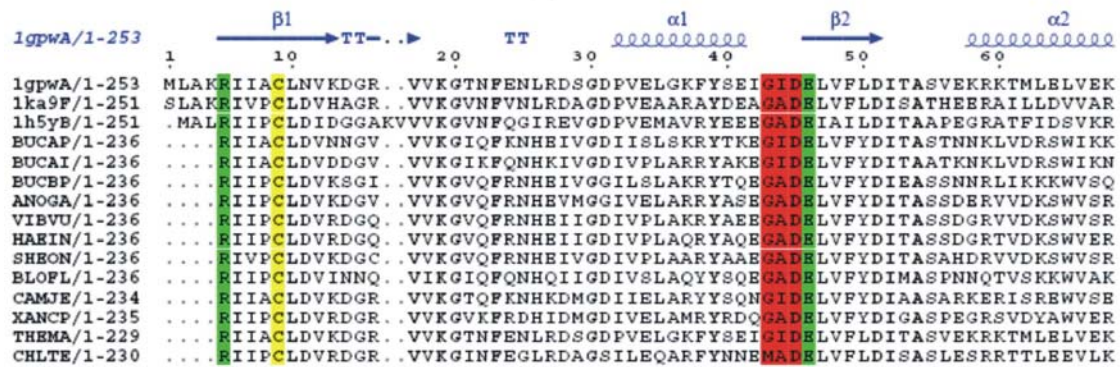
CAMJE: *Campylobacter jejuni*

XANCP: *Xanthomonas campestris*

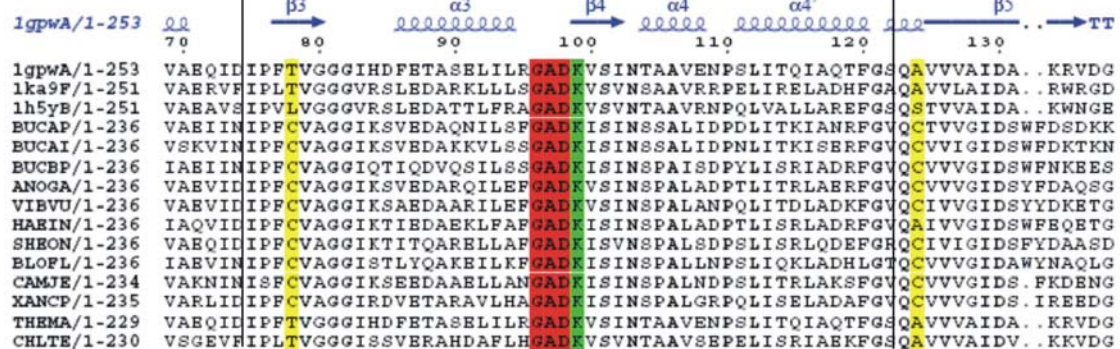
THEMA: *Thermotoga maritima*

CHLTE: *Chlorobium tepidum*

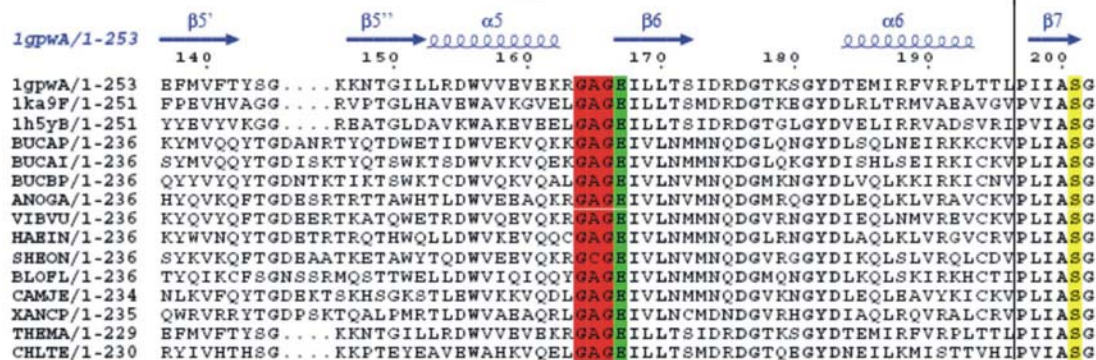
HisF_N1



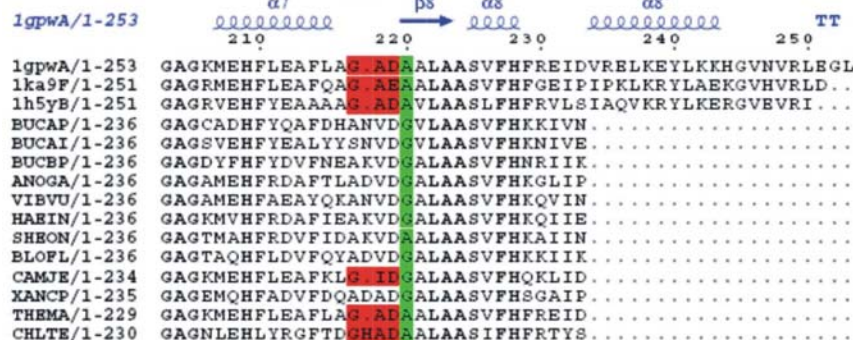
HisF_N2



HisF_C1



HisF_C2



4.2 **Expression, Reinigung und Charakterisierung der Hälften HisF_N und HisF_C**

Die beiden Hälften HisF_N (bestehend aus den Vierteln HisF_N1 und HisF_N2) und HisF_C (bestehend aus den Vierteln HisF_C1 und HisF_C2) (vgl. Abbildung 23) wurden schon ausführlich untersucht und charakterisiert (Höcker *et al.*, 2001). Dennoch gelang mit Hilfe eines effizienteren Reinigungsprotokolls neben der Ergebnisreproduktion ein Erkenntniszugewinn. Beide Proteine konnten von Birte Höcker durch Rückfaltung mit dem chaotropen Detergenz Guanidiniumhydrochlorid (GdmCl) aus dem unlöslichen Teil der Zellfraktion in die lösliche Fraktion der Dialyse überführt werden. Anschließend wurden HisF_N und HisF_C mit Hilfe einer Metallchelataffinitätssäule auf > 95 % Homogenität gereinigt. Im Rahmen des weiterentwickelten Reinigungsprotokolls konnten aus der Proteinlösung nach der Rückfaltung durch präparative Gelfiltration nicht erwünschte höhermolekulare Proteine und Aggregate der beiden Hälften entfernt werden.

4.2.1 **Expression und Anreicherung von HisF_N und HisF_C**

Das für HisF_N kodierende Gen *hisF_N* (Basenpaare 1 – 366 von *hisF*) lag bereits zu Beginn dieser Arbeit über *NdeI* und *XhoI* kloniert in dem Vektor pET21a (2.7.1.1) vor, wodurch das resultierende Protein C-terminal einen His₆-Tag trägt. Des Weiteren war zur spektroskopischen Markierung der Hälfte das Phenylalanin an Position 23 gegen Tryptophan ausgetauscht worden. Im Weiteren wird die Variante HisF_N_F23W als HisF_N bzw. N-terminale Hälfte bezeichnet. Das Gen *hisF_C* (Basenpaare 367 – 759 von *hisF*), das für HisF_C kodiert, lag ebenfalls über *NdeI* und *XhoI* kloniert in dem Vektor pET24a (2.1.7.2) vor, wobei das natürliche Stop-Codon von *hisF* entfernt wurde. Dadurch trägt das rekombinante Protein ebenfalls C-terminal einen His₆-Tag. Ein spektroskopischer Marker ist in HisF_C durch das wildtypische Tryptophan 156 gegeben. BL21(DE3)-Zellen wurden getrennt mit den Plasmiden transformiert (3.1.3.2). Über eine Probeexpression (3.3.1) konnte für jedes Protein ein Expressionsklon identifiziert werden, wobei die Zielproteine nach Induktion mit IPTG (Endkonz.: 1 mM) in der unlöslichen Fraktion (Pellet, P) zu finden

waren. Mit diesen Klonen wurde wie unter 3.3.2 beschrieben eine präparative Expression durchgeführt. Um HisF_N (14,4 kDa) und HisF_C (15,6 kDa) aus dem unlöslichen Teil des Zelllysates (P) anzureichern, wurden diese nach Solubilisierung mit Hilfe des chaotropen Detergenz Guanidiniumhydrochlorid (GdmCl) durch schrittweise Verdünnung des GdmCl und Dialyse gegen 50 mM KP, pH 7,5 rückgefaltet (3.3.5).

Anschließend befanden sich HisF_N und HisF_C zu ca. 50 % im Überstand der Dialyse (ÜD). Um zurückgebliebene Verunreinigungen zu entfernen, wurde die Lösung (bei einem Kulturvolumen von 1 l ca. 110 ml) mit Hilfe eines *Amicon Centrifugal Device* (MWCO: 10 kDa) auf ein Volumen von ca. 40 ml einkonzentriert (3.3.9) und in vier aufeinander folgenden Läufen über die präparative Gelfiltration (3.3.4) bei 4 °C zu einer Homogenität > 95 % weitergereinigt. Ausgewählte Fraktionen wurden mittels SDS-PAGE (3.4.1.1) analysiert und die sauberen Proteinfractionen anschließend vereinigt, über Nacht bei 4 °C gegen 50 mM KP, pH 7,5 dialysiert (3.3.8), einkonzentriert und bei -80 °C gelagert (3.3.9). Abbildung 24A zeigt exemplarisch die schrittweise Anreicherung von HisF_N. HisF trägt an Position 9 in der Aminosäuresequenz einen Cysteinrest (vgl. Abbildung 23). Über diesen kann HisF_N mit einem zweiten Molekül HisF_N unter oxidierenden Bedingungen eine Disulfidbrücke ausbilden. Daher ist auf dem SDS-Gel in Abwesenheit von β -Mercaptoethanol eine Bande auf der Höhe von ca. 30 kDa zu sehen, die dem Molekulargewicht eines HisF_N-Homodimers entspricht. Die Reinigung von HisF_C erfolgte analog. Das gereinigte Protein ist in Abbildung 24B zu sehen. Es konnten 22 mg HisF_N bzw. 38 mg HisF_C aus jeweils 1 l Kulturmedium gewonnen werden.

Sowohl HisF_N, als auch HisF_C zeigen für Proteine typische Absorptionsspektren (3.4.4.2, Abbildung 24C) mit einem Minimum bei 250 nm und einem Maximum bei ungefähr 278 nm. Die sehr geringe Absorption bei Wellenlängen ≥ 300 nm deutet auf die Abwesenheit von Aggregaten und damit auf eine gute Löslichkeit der Proteine hin. Die Parameter zu Berechnung der Proteinkonzentration aus dem Absorptionsspektrum sind Tabelle 23 im Anhang zu entnehmen.

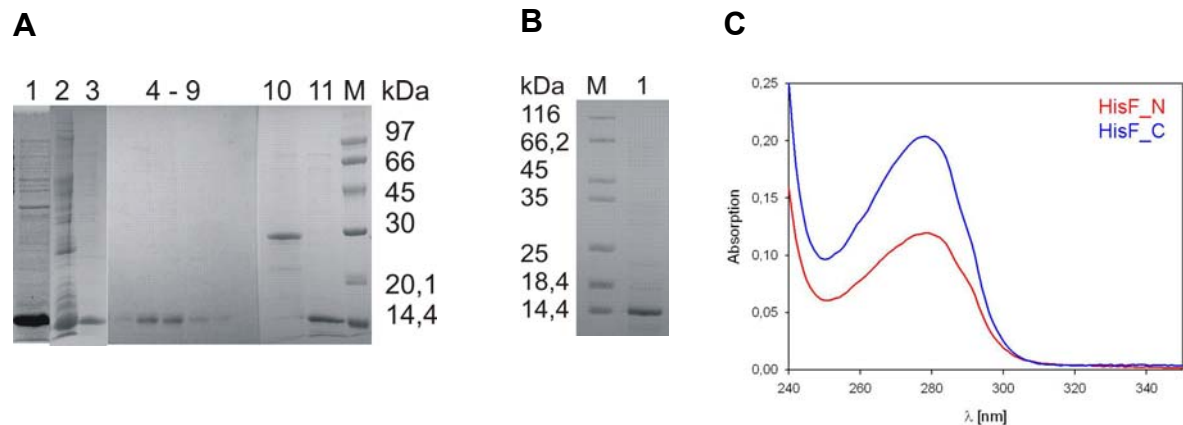


Abbildung 24: Reinigung von HisF_N und HisF_C, verfolgt mittels SDS-PAGE auf einem Tris-Glycin-Gel (12,5 % Polyacrylamid).

- (A) HisF_N findet sich nach Induktion mit 1 mM IPTG im unlöslichen Teil des Zelllysates (Spur 1). Nach Rückfaltung aus 6 M GdmCl finden sich ca. 50 % im löslichen Teil des Dialysats (Spur 3). Durch präparative Gelfiltration ließen sich weitere Verunreinigungen abtrennen (Spuren 4 - 9). Durch eine Disulfidbrücke über das Cysteine 9 entsteht unter oxidierenden Bedingungen ein HisF_N Homodimer (Spuren 10 + 11).
 (1: Pellet induziert; 2: unlöslicher Teil nach der Dialyse; 3: löslicher Teil nach der Dialyse; 4 - 9: Fraktionen nach der präparativen Gelfiltration; 10: ohne Mercaptoethanol; 11: mit Mercaptoethanol versetzt; M: *low molecular weight marker*).
- (B) HisF_C nach der Rückfaltung mit anschließender präparativer Gelfiltration.
 (1: gereinigtes HisF_C; M: *low molecular weight marker*)
- (C) Absorptionsspektren von gereinigtem HisF_N und HisF_C in 50 mM KP, pH 7,5 bei RT.
 HisF_N: $c = 0,23 \text{ mg/ml}$; $A_{280/250}$: 1,86
 HisF_C: $c = 0,31 \text{ mg/ml}$; $A_{280/250}$: 2,08

4.2.2 Bestimmung von Homogenität und Assoziationszustand der Halb-barrel über analytische Größenausschlusschromatographie (SEC)

Um die Homogenität und den Assoziationszustand von HisF_N und HisF_C zu bestimmen, wurde mit den unter 4.2.1 gereinigten Proteinen analytische Gelfiltration (Superdex 75-Säule) durchgeführt. Während der Elution wurde die Absorption bei 280 nm verfolgt (3.4.5). Abbildung 25 zeigt die Elutionsprofile von HisF_N und HisF_C im Vergleich zu HisF.

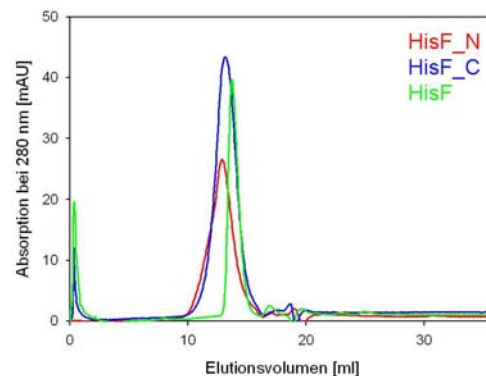


Abbildung 25: Analytische Gelfiltration von HisF_N und HisF_C im Vergleich zu HisF (Superdex 75-Säule).

Die Elution von HisF_N, HisF_C und - zum Vergleich - HisF wurde bei 280 nm verfolgt. Es wurden 100 μ l HisF_N ($c = 0,23$ mg/ml), 200 μ l HisF_C ($c = 0,31$ mg/ml) und 100 μ l HisF ($c = 0,3$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, 300 mM KCl pH 7,5.

Die apparenten Molekulargewichte konnten mit Hilfe einer Kalibrierungsgeraden (siehe Anhang) abgeschätzt werden. HisF_N zeigt einen relativ breiten, aber dennoch symmetrischen Peak (Abbildung 25), dessen apparentes Molekulargewicht eher dem des Trimers als dem des Dimers entspricht (Tabelle 7). Ein möglicher Grund ist die nicht ganz globuläre Form des Proteins, die zu einer Veränderung des Laufverhaltens während der Gelfiltration führen kann. Dass es sich nur um das Homodimer handeln kann, zeigt Abbildung 24A. Zwei HisF_N-Monomere sind durch das in HisF vorhandene Cystein 9 über eine Disulfidbrücke verknüpft.

Das Elutionsprofil von HisF_C weist ebenfalls einen symmetrischen Peak auf (Abbildung 25), dessen Elutionsvolumen dem Molekulargewicht des Dimers entspricht (Tabelle 7). Mithilfe der veränderten Reinigungsprozedur konnten somit die nach Metallchelataffinitätschromatographie noch vorhandenen Aggregate von HisF_N und HisF_C abgetrennt werden (Höcker *et al.*, 2001). Das Elutionsvolumen des zum Vergleich mit untersuchten HisF entspricht wie erwartet einem Monomer (Tabelle 7).

Tabelle 7: Aus analytischer Gelfiltration ermittelte apparente und für die Monomere berechnete Molekulargewichte, sowie daraus abgeleitete Assoziationszustände von HisF_N, HisF_C und HisF. Zusätzlich sind noch die Grenzen und die Anzahl der Aminosäuren der beiden Hälften angegeben (vgl. Abbildung 23).

Protein	Grenzen der Hälften (angegeben sind die jeweils erste und letzte Aminosäure in HisF)	Anzahl der Amino- säuren	MW _{kalk} [kDa]	MW _{app} [kDa]	Assoziations- zustand
HisF_N	1 – 122	122	14,4	40	Dimer
HisF_C	123 – 253	131	15,6	35,1	Dimer
HisF	-----	---	27,7	25,8	Monomer

4.2.3 Spektroskopische Charakterisierung von HisF_N und HisF_C

4.2.3.1 Fern-UV CD-Spektroskopie

Um Einblicke in die Sekundärstrukturen der HisF-Hälften zu bekommen, wurden Circular-Dichroismus-Spektren im fern-UV Bereich von 260 bis 180 nm aufgenommen (3.4.9). In Abbildung 26 sind die Spektren von HisF_N und HisF_C nach Umrechnung in molare Elliptizität pro Aminosäure (Θ_{MRW}) und zum Vergleich das von HisF gezeigt. Die Lage der beiden Minima (etwa 222 nm und 208 nm), des Nulldurchgangs (etwa 200 nm), und des Maximums (etwa 192 nm) deuten auf die Existenz α -helikaler Strukturen hin (Johnson, 1990).

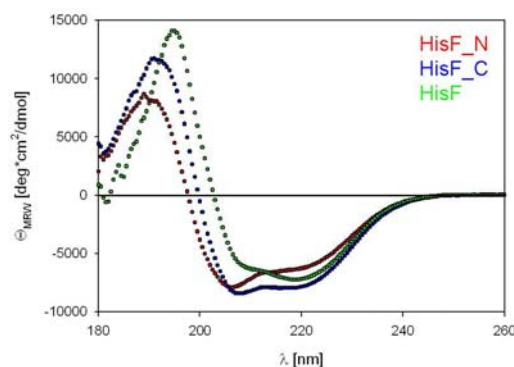


Abbildung 26: Fern-UV CD-Spektrum von HisF_N und HisF_C im Vergleich zu HisF.

Die Spektren wurden jeweils in einer 1 mm Quartzküvette mit 0,2 mg/ml Protein in 10 mM KP pH 7,5 aufgenommen. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C.

Mithilfe des Programms ContinLL (Provencher & Glockner, 1981; van Stokkum *et al.*, 1990) wurde auf dem Dichroweb-Server (Whitmore & Wallace, 2004; Whitmore & Wallace, 2008) eine Sekundärstrukturvorhersage mit drei unterschiedlichen Referenzproteindatensätzen (Lees *et al.*, 2006) durchgeführt, deren Ergebnisse in Tabelle 8 zusammengefasst sind.

Tabelle 8: Vorhergesagte Sekundärstrukturelementanteile von HisF_N, HisF_C und HisF (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Die kalkulierten Werte stammen aus den in 1thf abgelegten Sekundärstrukturanteilen der HisF-Struktur.

Sekundärstruk- turelement	HisF_N		HisF_C		HisF	
	Vorhersage aus CD	kalk. aus 1thf	Vorhersage aus CD	kalk. aus 1thf	Vorhersage aus CD	Struktur (1thf)
α -Helix	34	35	41	29	29	34
β -Strang	11	21	9	29	22	24
<i>random coil</i>	55	44	50	42	48	41

HisF_N zeigt zwar einen hohen Anteil α -helikaler Strukturen, der mit dem aus der Struktur von HisF (PDB-Eintrag: 1thf) berechneten Wert gut übereinstimmt (Tabelle 8), ist aber offensichtlich nicht in der Lage, den erwarteten β -Faltblatt-Anteil auszuprägen. Über 50 Prozent der Aminosäurereste der Polypeptidkette liegen nach der Vorhersage ungeordnet vor. Dagegen bilden sich laut Vorhersage in HisF_C mehr α -helikale Strukturen aus, als aus der Struktur von HisF (1thf) berechnet wurden (Tabelle 8). Der Anteil an β -Faltblatt ist mit 9 % entsprechend zu klein. 50 Prozent der Aminosäurereste der Polypeptidkette liegen nach der Vorhersage ungeordnet vor. In den früheren Studien wurden ebenfalls für HisF_N und HisF_C die Sekundärstrukturanteile mit den Programmen ContinLL, K2D und Selcon (Greenfield, 1996) vorhergesagt. Dabei zeigte sich, dass ca. 50 Prozent der Polypeptidketten der beiden Proteine an der Ausbildung von α -helikalen und β -Faltblatt Strukturen beteiligt sind, ohne dass eine genauere Differenzierung vorgenommen wurde. Die hier vorgestellten Ergebnisse sind somit mit den veröffentlichten Daten konsistent (Höcker *et al.*, 2001).

Zur Kontrolle des Programms wurden auch die Strukturanteile von HisF basierend auf dem fern-UV CD-Spektrum berechnet und mit den Daten aus 1thf verglichen

(Tabelle 8). Diese stimmen sehr gut überein, was die Glaubwürdigkeit der Vorhersagen stärkt.

4.2.3.2 Fluoreszenzemissionsspektroskopie

Zur Untersuchung der Tertiärstruktur von HisF_N und HisF_C wurden Fluoreszenzemissionsspektren aufgenommen (3.4.8). Beide Proteine enthalten jeweils einen Tryptophanrest und einen (HisF_N) bzw. drei (HisF_C) Tyrosinreste. Die Emissionsspektren wurden im Bereich von 300 – 400 nm aufgenommen. Die Anregungswellenlänge betrug 280 nm, wodurch die Emission sowohl der Tryptophan- als auch Tyrosinreste erfasst werden konnte. Kommt es zu einer Rotverschiebung des Emissionsmaximums nach Denaturierung mit 8 M Harnstoff, kann auf die Existenz von Tertiärstruktur unter nativen Pufferbedingungen geschlossen werden. Deshalb wurden zusätzlich Emissionsspektren nach vollständiger Denaturierung der beiden Hälften in 8 M Harnstoff aufgenommen. Die Emissionsmaxima von HisF_N und HisF_C unter nativen Pufferbedingungen liegen bei 341 nm bzw. 336 nm (Abbildung 27). Die in 8 M Harnstoff ermittelten Emissionsmaxima weisen eine Rotverschiebung um 10 – 18 nm auf (Abbildung 27).

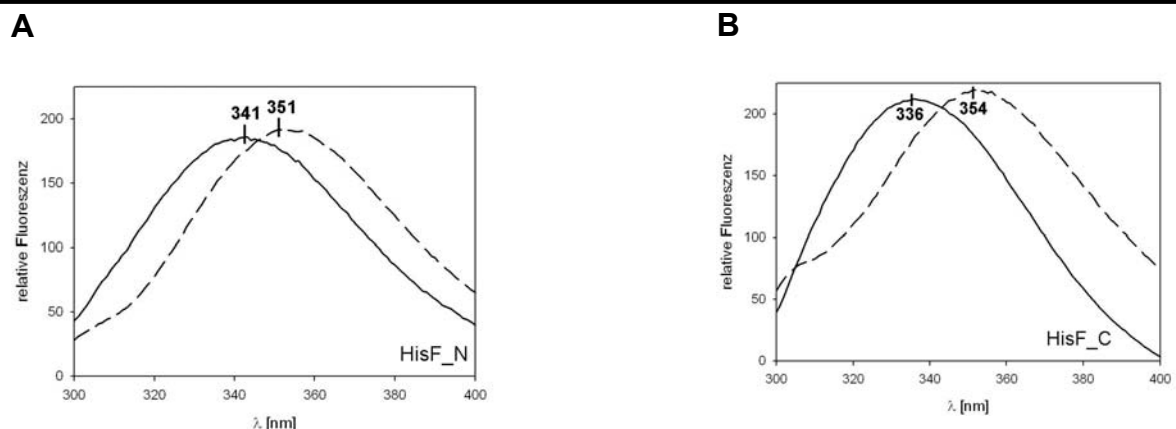


Abbildung 27: Fluoreszenzemissionsspektren von HisF_N und HisF_C im nativen und denaturierten Zustand.

- (A) Die Emissionsspektren von HisF_N unter nativen (50 mM KP, pH 7,5; durchgezogene Linie) und denaturierenden Pufferbedingungen (8 M Harnstoff, 50 mM KP, pH 7,5; gestrichelte Linie) nach Anregung bei 280 nm. Die Emissionsmaxima liegen bei 341 nm bzw. 351 nm.
- (B) Die Emissionsspektren von HisF_C unter nativen (50 mM KP, pH 7,5; durchgezogene Linie) und denaturierenden Pufferbedingungen (8 M Harnstoff, 50 mM KP, pH 7,5; gestrichelte Linie) nach Anregung bei 280 nm. Die Emissionsmaxima liegen bei 336 nm bzw. 354 nm.

Die Proteinkonzentrationen betrugen 5 μ M, die Schichtdicke der Küvette 1 cm. Die gezeigten Spektren sind das Mittel aus zehn Einzelmessungen bei 25 °C. Das bei höheren Wellenlängen liegende Emissionsmaximum von HisF_N deutet darauf hin, dass sein Tryptophanrest 23 weniger stark vom Lösungsmittel abgeschirmt ist als der Tryptophanrest 156 aus HisF_C.

Das künstlich eingeführte Tryptophan an Position 23 in HisF_N ist in dem langen $\beta 1/\alpha 1$ -loop lokalisiert (vgl. Abbildung 21 und 23) und liegt etwas lösungsmittel exponierter vor als das native Tryptophan 156 in α -Helix 5 von HisF_C, was durch das höhere Emissionsmaximum von 341 nm im Vergleich zu 336 nm deutlich wird. Dennoch muss durch die Ausbildung von Tertiär- und Quartärstruktur eine gewisse Abschirmung des Tryptophan 23 vom Lösungsmittel vorhanden sein, da sich das Maximum nach Inkubation des Proteins in 8 M Harnstoff auf 351 nm verschiebt.

Das Emissionsmaximum von HisF_C unter nativen Pufferbedingungen bei 336 nm (Abbildung 27) stimmt im Rahmen der Messgenauigkeit mit dem in der Literatur veröffentlichten Wert überein (Höcker *et al.*, 2004). Im vollständig entfalteten Zustand (in Gegenwart von 8 M Harnstoff) kommt es zu einer Rotverschiebung auf 354 nm. Dadurch wird deutlich, dass das natürliche Tryptophan an Position 156 stark vom Lösungsmittel abgeschirmt ist, was auf eine ausgeprägte Tertiärstruktur von HisF_C hindeutet.

4.2.4 Stabilitätsuntersuchung von HisF_N und HisF_C

Mit Hilfe von Auffaltungskurven in Abhängigkeit von steigender Konzentration des Detergenz Harnstoff (3.4.10) und sich anschließender spektroskopischer Analyse, konnten Aussagen über die Stabilität von HisF_N und HisF_C getroffen werden.

Die Auffaltung der Sekundärstruktur von HisF_N und HisF_C wurde am CD-Spektropolarimeter bei 220 nm verfolgt. Abbildung 28A zeigt die normierten Auffaltungskurven von HisF_N und HisF_C.

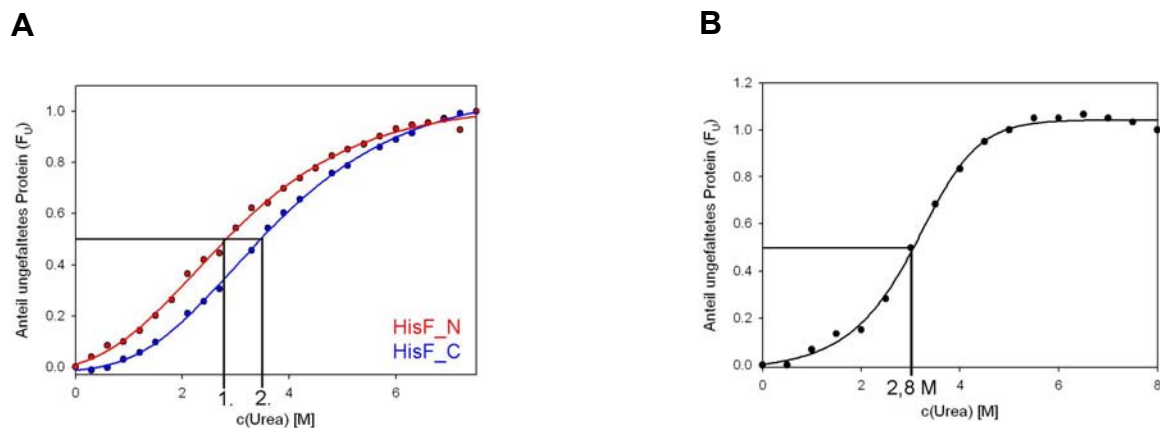


Abbildung 28: Entfaltung der Sekundärstruktur von HisF_N, HisF_C (A) und der Tertiärstruktur von HisF_C (B) durch chemische Denaturierung mit Harnstoff.

- (A) Auffaltung der Sekundärstruktur von HisF_N und HisF_C in 50 mM KP, pH 7,5 bei 25 °C, verfolgt am CD-Spektrometer bei 220 nm. Die Proteinkonzentration betrug jeweils 0,2 mg/ml, die Schichtdicke der Küvette 1 mm. Die $D_{1/2}$ -Werte sind durch die Hilfslinien eingezeichnet und liegen bei 2,8 M (1., HisF_N) bzw. 3,5 M (2., HisF_C) Harnstoff. Beide Proteine zeigen eine eher unkooperative Auffaltung ihrer Sekundärstrukturelemente.
- (B) Auffaltung der Tertiärstruktur von HisF_C in 50 mM KP, pH 7,5 bei 25 °C, verfolgt am Fluoreszenzspektrometer. Die Emission wurde nach Anregung bei 280 nm bei 320 nm gemessen. Die Proteinkonzentration betrug 5 μ M, die Schichtdicke der Küvette 1 cm. Der $D_{1/2}$ -Wert ist durch die Hilfslinien eingezeichnet und liegt bei 2,8 M Harnstoff. Die Tertiärstruktur von HisF_C entfaltet kooperativ.
- Die sigmoidalen Ausgleichslinien sollen das Auge des Betrachters leiten.

Man erkennt, dass die Auffaltung der Sekundärstrukturelemente von HisF_N und HisF_C eher unkooperativ verläuft (Abbildung 28A). Die Auffaltung setzt bereits bei niedrigen Konzentrationen an Denaturierungsmittel ein, im Fall von HisF_N bei 0,3 M Harnstoff und von HisF_C bei 0,9 M Harnstoff. Das bedeutet, dass sich die einzelnen Strukturelemente untereinander nur wenig stabilisieren. Der $D_{1/2}$ -Wert der Entfaltung, also die Harnstoff-Konzentration, bei der die Hälfte aller Proteinmoleküle entfaltet vorliegt, wurde als operationelles Maß für Stabilitätsunterschiede ermittelt. Im Falle von HisF_N liegt er bei 2,8 M Harnstoff, bei HisF_C kommt es zu einer Verschiebung auf 3,5 M Harnstoff. Dies deutet auf eine etwas höhere Stabilität der C-terminalen Hälfte hin.

Die chemische Entfaltung der Tertiärstruktur konnte nur von HisF_C am Fluoreszenzspektrometer bei einer Anregungswellenlänge von 280 nm und einer Emissionswellenlänge von 320 nm verfolgt werden. Da das Tryptophan 23 in HisF_N bereits im nativen Zustand lösungsmittel exponiert vorliegt (vgl. Abbildung 27) war es nicht möglich eine aussagekräftige Signaländerung während der Entfaltung zu messen. Nach Normierung der Messpunkte erhielt man eine Auffaltungskurve von HisF_C (Abbildung 28B). Die Tertiärstruktur von HisF_C entfaltet kooperativer als

seine Sekundärstruktur, was durch den steileren Übergangsbereich zum Ausdruck kommt. Der $D_{1/2}$ -Wert von 2,8 M Harnstoff stimmt mit den veröffentlichten Daten überein (Höcker *et al.*, 2004). Die recht hohe Stabilität und Kompaktheit von HisF_C gaben den Anstoß HisF_C für weitere Studien zur Evolution der $(\beta\alpha)_8$ -barrel Proteine zu verwenden (Höcker *et al.*, 2004; Seitz *et al.*, 2007).

4.3 Zusammenfassung bisheriger Ergebnisse zu den HisF-Vierteln

Bereits im Jahr 2003 wurde versucht, die HisF-Viertel einzeln zu exprimieren und zu charakterisieren, um Hinweise auf den modularen Aufbau des $(\beta\alpha)_8$ -barrels zu erhalten und die Evolutionshypothese mit zweifacher Duplikation und Fusion von $(\beta\alpha)_2$ -Einheiten zu bestätigen (Bosnali, 2003). Manal Bosnali war während ihrer Diplomarbeit in der Lage, HisF_N1 und HisF_C1 heterolog in *E. coli* herzustellen, sie aus der unlöslichen Zellfraktion rückzufalten, um sie anschließend mittels Anionenaustauscherchromatographie zu reinigen. Danach wurden die Viertel biophysikalisch charakterisiert.

Dabei konnte sie mittels analytischer Gelfiltration feststellen, dass HisF_N1 zwar zu einem kleinen Anteil Aggregate, aber mehr Oktamere und Tetramere in Lösung bildet. Des Weiteren konnte gezeigt werden, dass das Protein in Lösung Sekundärstruktur und geringfügig Tertiärstruktur besitzt. Allerdings war es schwierig, hier eine definierte Aussage zu treffen, da das einzige Tryptophan an Position 23 (künstlich eingeführt) in dem längeren *loop* $\beta 1/\alpha 1$ von vorneherein lösungsmittel exponiert vorliegt und somit bereits unter nativen Bedingungen das Emissionsmaximum bei 347 nm lag und sich unter denaturierenden Bedingungen um nur 5 nm auf 352 nm verschob. Auch HisF_C1 zeigte eine gewisse Tendenz zur Aggregation, ein geringer Anteil des Proteins eluierte jedoch mit dem apparenten Molekulargewicht des Monomers von der Gelfiltrationssäule. HisF_C1 besitzt ebenfalls in Lösung einen gewissen Anteil an Sekundär- und Tertiärstruktur und liegt nicht als reines Zufallsknäuel vor.

HisF_N2 und HisF_C2 konnten nur als Fusionsproteine mit Glutathion-S-Transferase (GST) hergestellt werden. Die GST-Fusionsproteine fanden sich im löslichen Teil des

Zellextraktes und wurden mittels Affinitätschromatographie gereinigt. Anschließend wurden HisF_N2 und HisF_C2 durch Proteolyse von GST abgetrennt. Analytische Gelfiltration der gereinigten Viertel ergab, dass HisF_N2 und HisF_C2 hauptsächlich als Monomere vorliegen. Fern-UV Circular-Dichroismus- und Fluoreszenzemissions-spektroskopie wiesen aber auf keine definierte Struktur hin. Eine Koexpression mit dem anderen Viertel der jeweiligen HisF-Hälfte (HisF_N1 bzw. HisF_C1) hatte sowohl von einem als auch von zwei verschiedenen Plasmiden aus keinen Erfolg. Die gemeinsame Rückfaltung der vier Viertel aus Guanidiniumhydrochlorid (GdmCl) gelang nicht, vermutlich wegen zu geringer Proteinkonzentrationen an HisF_N2 und HisF_C2.

4.4 Das Viertel HisF_N1

4.4.1 Expression und Anreicherung von HisF_N1

Das *hisF_N1*-Gen (Basenpaare 1 – 222 in *hisF*) lag bereits über *NdeI* und *BamHI* kloniert im Vektor pET21a (2.7.1.1) vor (Bosnali, 2003). Da HisF_N1 nur einen Tyrosinrest besitzt, war an Position 23 das Phenylalanin gegen Tryptophan zur Einführung eines spektroskopischen Markers in diesem Viertel ausgetauscht worden. Im Folgenden wird die Variante HisF_N1_F23W als HisF_N1 bezeichnet. Bei allen weiteren Varianten, die diesen Austausch enthalten, wird F23W ebenfalls nicht explizit in der Nomenklatur erwähnt. Eine Proteinpräparation im analytischen Maßstab (3.3.1) nach Kultivierung in *E. coli* BL21(DE3) über Nacht bei 37 °C zeigte, dass HisF_N1 vollständig im unlöslichen Teil des Zelllysates zu finden war. Eine Induktion mit IPTG (Endkonz.: 1 mM) war notwendig, um das Protein zu erzeugen. Für eine Anreicherung im Milligramm-Bereich wurde der unlösliche Teil des Zelllysates (*Inclusion bodies*) präpariert (3.3.2). Anschließend wurde HisF_N1 durch Rückfaltung (3.3.3) in den löslichen Teil der Dialyse überführt. Der Dialyse-Überstand wurde durch präparative Gelfiltration weitergereinigt (3.3.4). Über SDS-PAGE (3.4.1.2) wurden die Fraktionen ermittelt, in denen HisF_N1 sauber vorlag, diese vereinigt, über Nacht bei 4 °C gegen 50 mM KP, pH 7,5 dialysiert und mit Hilfe eines *Amicon Centrifugal Device* (MWCO: 5 kDa) einkonzentriert und bei -80 °C gelagert.

(3.3.9). Der Verlauf einer Reinigung von HisF_N1 ist in Abbildung 29A gezeigt. Das Absorptionsspektrum (Abbildung 29B) zur Ermittlung der Proteinkonzentration zeigt kein wesentliches Meßsignal oberhalb von 300 nm, was auf eine gute Löslichkeit von HisF_N1 hinweist. Insgesamt konnten 77 mg HisF_N1 aus einem Liter Kulturmedium erhalten werden. Wie bei HisF_N (vgl. Abbildung 24) bildet sich unter oxidierenden Bedingungen eine Disulfidbrücke zwischen zwei Cysteinresten an Position 9. Dies hat eine Homodimerbildung von HisF_N1 zur Folge (Abbildung 29A).

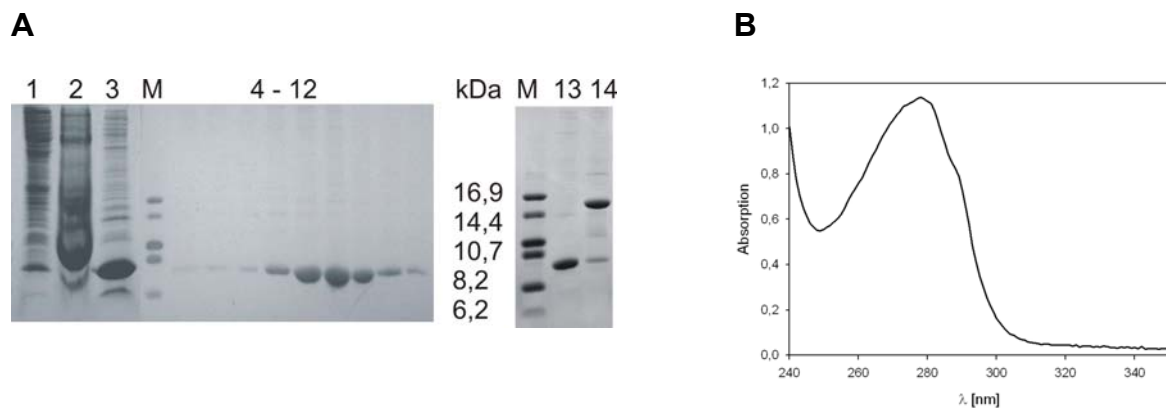


Abbildung 29: Reinigung von HisF_N1, verfolgt über die Analyse wichtiger Fraktionen mittels SDS-PAGE auf einem Tris-Tricin-Gel (20 % Polyacrylamid).

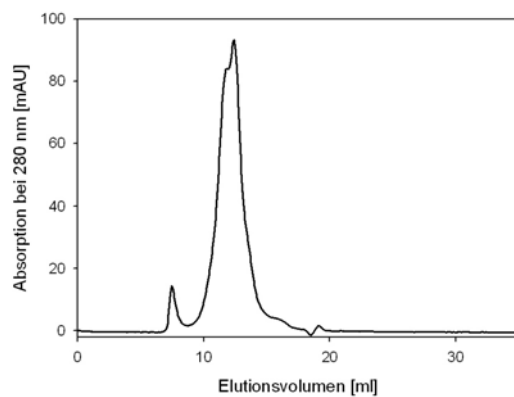
- (A) Tris-Tricin SDS-PAGE (20 % Polyacrylamid) relevanter Fraktionen der Reinigung von HisF_N1. Das Protein wird nach Induktion mit 1mM IPTG unlöslich (Spur 1) exprimiert, findet sich aber nach Rückfaltung zu ca. 50 % im löslichen Teil der Dialyse (Spur 3). Durch präparative Gelfiltration können höhermolekulare Verunreinigungen entfernt werden (Spuren 4 - 12), anschließend liegt das Protein mit einer Homogenität von > 95 % vor (Spur 13). Über eine Disulfidbrücke zwischen Cysteinen an Position 9 zweier Monomere bildet sich eine Homodimer (Spur 14).
 (1: unlöslicher Teil des Zellextrakts; 2: unlöslicher Teil nach der Dialyse; 3: löslicher Teil nach der Dialyse; 4 - 12: Fraktionen nach Elution von der Superdex 75-Säule; 13: mit Mercaptoethanol versetzt; 14: ohne Mercaptoethanol; M: *small fragment marker*)
- (B) Absorptionsspektrum von gereinigtem HisF_N1 in 50 mM KP, pH 7,5 bei RT.
 $c = 1,34 \text{ mg/ml}$; $A_{280/250}$: 2,02

Die Parameter zur Ermittlung der Proteinkonzentration aus dem Absorptionsspektrum sind Tabelle 23 im Anhang zu entnehmen.

4.4.2 Bestimmung der Homogenität und des Assoziationszustandes von HisF_N1

4.4.2.1 Analytische Gelfiltration

Um die Homogenität des gereinigten HisF_N1-Proteins zu untersuchen, wurde eine analytische Gelfiltration mit einer Superdex 75-Säule durchgeführt (3.4.5). Dazu wurden 100 µl der Lösung mit gereinigtem Protein (4.4.1) auf die Säule geladen und mit 50 mM KP, 300 mM KCl, pH 7,5 eluiert. Das Ergebnis ist in Abbildung 30 gezeigt.



	MW _{app}	MW _{kalk}	Assoziationszustand
HisF_N1	43 kDa	8,4 kDa	Pentamer

Abbildung 30: Analytische Gelfiltration von HisF_N1 (Superdex 75-Säule).

Die Elution von HisF_N1 wurde bei 280 nm verfolgt. Es wurden 100 µl HisF_N1 ($c = 1,34 \text{ mg/ml}$) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, 300 mM KCl, pH 7,5. Rechts ist das mit Hilfe der Kalibrierungsgeraden bestimmte apparente Molekulargewicht, das berechnete Molekulargewicht für das Monomer und der abgeleitete Assoziationszustand des Proteins angegeben.

Im Elutionsprofil von HisF_N1 erkennt man einen annähernd symmetrischen Hauptpeak mit einer Schulter bei kleinerem Elutionsvolumen. Durch Vergleich mit der Kalibrierungsgeraden (siehe Anhang) konnte für den Hauptpeak das Molekulargewicht eines theoretischen Pentamers ermittelt werden. Ein Pentamer ist aufgrund der Homodimerisierung unter oxidierenden Bedingungen (vgl. Abbildung 29A) nur schwer vorstellbar. Deshalb erscheint eine tetramere Anordnung von HisF_N1 wahrscheinlicher. Dass das apparente Molekulargewicht nicht genau mit dem für ein Tetramer berechneten Molekulargewicht übereinstimmt, liegt wahrscheinlich an der nicht ganz globulären Form von HisF_N1. Das Auftreten eines Nebenpeaks bei einem Elutionsvolumen von ca. 8 ml deutet darauf hin, dass unter

den gegebenen Bedingungen ein kleiner Teil von HisF_N1 in Form von Aggregaten vorliegt.

4.4.2.2 Chemisches *cross-linking*

Um die aus der analytischen Gelfiltration gewonnene Vermutung, dass HisF_N1 als Tetramer vorliegt (4.4.2.1) zu stützen, wurde ein chemisches *cross-linking* von HisF_N1 durchgeführt. Beim *cross-linking* werden über einen *linker*, in diesem Fall Glutardialdehyd, zwei Proteine kovalent miteinander verknüpft. Dabei reagiert der *linker* mit einem reaktiven Aminosäurerest unter Bildung einer kovalenten Bindung. Glutardialdehyd stellt einen eher unspezifischen *linker* da, der sowohl mit Lysinresten, als auch mit anderen freien Aminogruppen unter Bildung eines Imins reagieren kann. In wässrigen Lösungen mit neutralem pH-Wert kann Glutardialdehyd Polymere unterschiedlicher Länge bilden (Peters & Richards, 1977) (Abbildung 31).

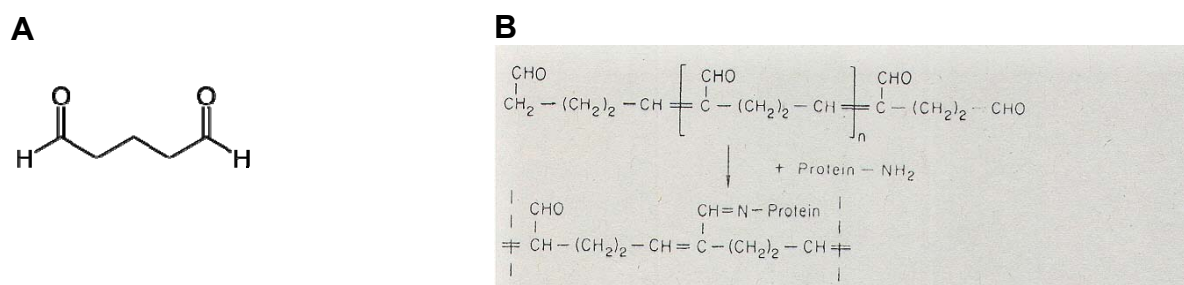


Abbildung 31: Strukturformel von Glutardialdehyd (A) und die schematische Darstellung der *cross-linking*-Reaktion (B).

- (A) Strukturformel des Glutardialdehyd, der im Rahmen dieser Arbeit als *cross-linker* verwendet wurde.
- (B) Schematische Darstellung der Reaktion die während des *cross-linking* abläuft. Glutardialdehyd bildet bei annähernd neutralem pH-Wert Polymere (oben). Die Aldehydgruppen reagieren mit freien Aminogruppen zur Schiff'schen-Base (Imin), die durch die konjugierte Doppelbindung stabilisiert wird (unten). (Nach Peters und Richards, 1977)

Das *cross-linking* wurde in einem Gesamtvolumen von 200 µl durchgeführt. Bei einer Proteinkonzentration von 0,2 mg/ml in 50 mM KP, pH 7,5 wurde der Ansatz für zwei Minuten mit Glutardialdehyd (Endkonz.: 0,2 % (v/v)) bei 25 °C inkubiert. Anschließend wurde die Reaktion mit NaBH₄ in NaOH (Endkonz.: 0,01 M) gestoppt und für weitere 20 min bei RT geschüttelt. Die Auswertung erfolgte mit Hilfe einer SDS-PAGE (3.4.1.2) und ist in Abbildung 32 gezeigt.

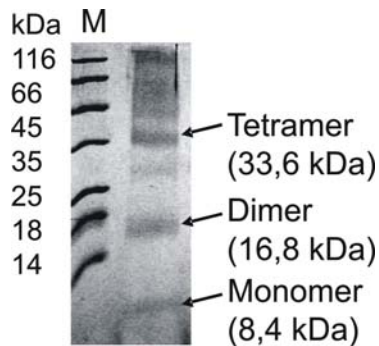


Abbildung 32: Chemisches *cross-linking* von HisF_N1 mit Glutardialdehyd, analysiert über SDS-PAGE auf einem Tris-Tricin-Gel (20% Polyacrylamid).

HisF_N1 wird zu einem Tetramer verlinkt. Durch eine unvollständige Reaktion sind auch noch Mono- und Dimere zu sehen.

Die Reaktion lief in einem Volumen von 200 μ l mit einer Proteinkonzentration von 0,2 mg/ml und 0,2% (v/v) Glutardialdehyd für 2 min ab. Anschließend wurde mit NaBH₄ in NaOH (Endkonz.: 0,01 M) für 20 min bei RT gestoppt. (M = *low molecular weight marker*)

Nach dem *cross-linking* sind sowohl Mono-, Di- und Tetramere, als auch höhere Oligomerisierungszustände auf dem SDS-Gel zu finden. Mono- und Dimere lassen sich zum Einen durch den unvollständigen Ablauf der chemischen Reaktion an einigen HisF_N1-Oligomermolekülen erklären. Zum Anderen sind sie ein Hinweis darauf, dass HisF_N1 nicht ausschließlich als Tetramer vorliegt. Die höheren Aggregate entstehen durch unterschiedlich lange Glutardialdehyd-Polymere. Dadurch wird auch eine Vernetzung über eine größere Distanz möglich (Peters & Richards, 1977). Da die Bande des Tetramers deutlich stärker als die Banden des Dimers und des Monomers ist, kann man davon ausgehen, dass sich HisF_N1 überwiegend zu einem Tetramer zusammenlagert.

4.4.2.3 Analytische Ultrazentrifugation

Weitere Hinweise auf den Oligomerisierungszustand von HisF_N1 in Lösung sollten mit Hilfe der analytischen Ultrazentrifugation erhalten werden. In Zusammenarbeit mit Dr. H. Durchschlag (Lehrstuhl Biochemie II der Universität Regensburg) wurden sowohl Sedimentationsgeschwindigkeits- als auch Gleichgewichtsläufe (3.4.6) mit einer Proteinkonzentration von 0,2 mg/ml in 50 mM KP, pH 7,5 durchgeführt. Das Protein wurde in einen Sektor der Doppelsektorzelle gefüllt, im zweiten Sektor befand sich Puffer als Referenz.

Der Sedimentationsgeschwindigkeitslauf wurde bei 21 °C und einer Drehzahl von 52000 rpm durchgeführt. Nach jeweils 16 min wurde der Absorptionsverlauf der gesamten Zelle bei 277 nm aufgenommen. Die Messung wurde nach vier Stunden

beendet. Nach Auswertung der Messdaten des Sedimentationsgeschwindigkeitslaufes (3.4.6.1) konnte ein Sedimentationskoeffizient von $2,16 \cdot 10^{-13}$ Sekunden = 2,16 Svedberg-Einheiten (S) ($1 \text{ S} = 10^{-13}$ Sekunden) bestimmt werden (Abbildung 33C). Die Sedimentationsgeschwindigkeit ist unter anderem abhängig von Größe und Form des untersuchten Stoffes. Ein Vergleich mit ideal globulären Proteinen, die einen annähernd identischen Sedimentationskoeffizienten besitzen (vgl. Tabellen in Smith (1970)), führte zu einer Molekulargewichtsabschätzung von 20 bis 40 kDa.

Der Gleichgewichtslauf (3.4.6.2) wurde bei 24 °C durchgeführt. Nach einer sogenannten *overspeed*-Phase (20000 rpm) von drei Stunden, in der sich das Protein schneller zum Zellenboden bewegt, wurde mit 16000 rpm weiterzentrifugiert. Nach 24 und 48 h wurde der Absorptionsverlauf der Zelle bei 277 nm gemessen. Da die beiden Absorptionsverläufe deckungsgleich waren (Abbildung 33A), kann man davon ausgehen, dass das Gleichgewicht erreicht war. Durch Logarithmieren des Absorptionsverlaufs und Auftragung gegen den Abstand vom Rotationszentrum (r^2) wurde die in Abbildung 33B gezeigte Gerade erhalten. Der Graph in Abbildung 33B enthält zwei Bereiche, durch die jeweils eine Ausgleichsgerade gelegt werden kann. Bereich 1 liegt zwischen 45 und 48 cm², Bereich 2 zwischen 48 und 51 cm². Da die Steigungen in beiden Bereiche nahezu identisch sind, ist die Berechnung des Molekulargewichtes über den gesamten Bereich sinnvoller. Auf dieser Grundlage lässt sich ein Molekulargewicht von 36,5 kDa berechnen, was im Rahmen der Meßgenauigkeit dem theoretischen Molekulargewicht des Tetramers von 33,6 kDa entspricht. Damit unterstützt die analytische Ultrazentrifugation die Ergebnisse aus der analytischen Gelfiltration und dem chemischen *cross-linking*, wonach sich vier HisF_N1-Monomere in Lösung zum Homotetramer zusammenlagern.

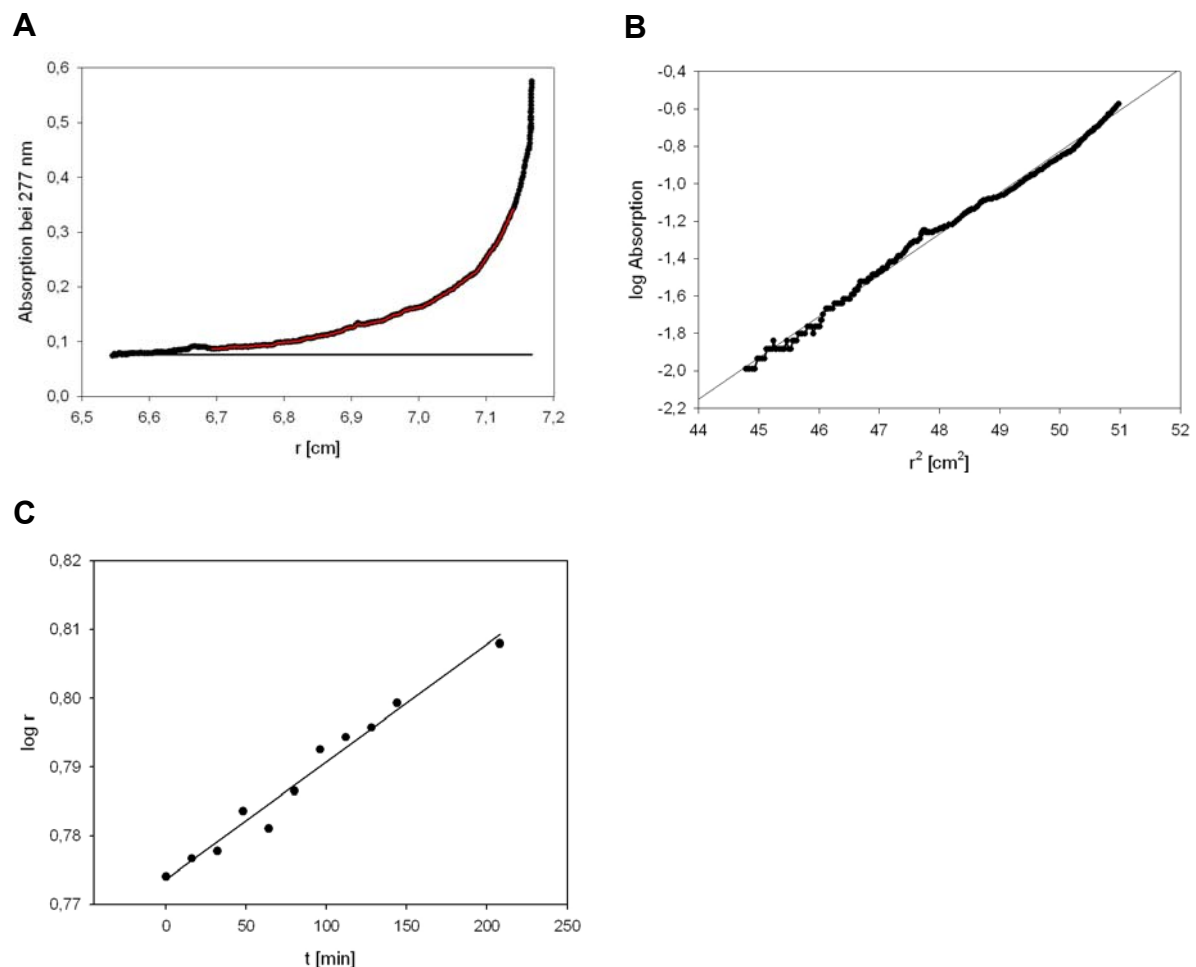


Abbildung 33: Auswertung der analytischen Ultrazentrifugation (Sedimentationsgeschwindigkeits- und Gleichgewichtslauf) von HisF_N1.

- (A) Absorption in der Messzelle im Gleichgewichtslauf von HisF_N1 (0,2 mg/ml in 50 mM KP, pH 7,5) bei 24 °C, verfolgt bei 277 nm nach 48 h. Die durchgezogene Linie stellt die Baseline dar, um die alle Absorptionswerte korrigiert wurden. Der rot markierte Bereich (Gesamtbereich) diente zur Erstellung der Geraden in (B).
- (B) Halblogarithmische Auftragung des Gesamtbereichs in (A). Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.2 ein Molekulargewicht von 36,5 kDa berechnet. Bei einem Molekulargewicht von 8,4 kDa für das Monomer, ist dies ein deutlicher Hinweis darauf, dass HisF_N1 in Lösung überwiegend als Homotetramer vorliegt.
- (C) Auswertung des Sedimentationsgeschwindigkeitslaufes von HisF_N1 (0,2 mg/ml in 50 mM KP, pH 7,5) bei 21 °C. Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.1 ein Sedimentationskoeffizient von 2,16 S berechnet. Ein Vergleich mit globulären Proteinen, die den gleichen Sedimentationskoeffizienten besitzen, deutet auf ein Molekulargewicht von 20 – 40 kDa hin.

4.4.2.4 Massenspektroskopische Untersuchung von HisF_N1

Die massenspektroskopische Untersuchung von HisF_N1 wurde von der Arbeitsgruppe Prof. Dr. Deutzmann (Lehrstuhl Biochemie I, Universität Regensburg)

durchgeführt. Das Elektrosprayionisations-Massenspektrum (ESI-MS; 3.4.7.1) von HisF_N1 (0,2 mg/ml in 50 mM KP, pH 7,5) lieferte ein Molekulargewicht von 16704 Da. Dieses entspricht genau dem berechneten Molekulargewicht des Dimers von 16704 Da. Offensichtlich wurden durch die vorgeschaltete *reversed phase* (RP)-HPLC die zum Tetramer assoziierten HisF_N1-Homodimere getrennt.

Aus dem Matrix-assisted Laserdesorptions/Ionisations-Massenspektrums (MALDI-MS; 3.4.7.2) konnte nach Laserbeschuss und Aufzeichnung des Massenspektrums ein Molekulargewicht von 8342 Da ermittelt werden. Dieser Wert entspricht dem Molekulargewicht des Monomers von 8352 Da. Dies bedeutet, dass durch die hochenergetische Strahlung die Disulfidbrücke zwischen den beiden HisF_N1-Monomeren im Homodimer gespalten wurde.

4.4.3 Spektroskopische Charakterisierung von HisF_N1

4.4.3.1 Fern-UV CD-Spektroskopie

Die Sekundärstruktur von HisF_N1 wurde mit Hilfe von fern-UV Circular-Dichroismus-Spektroskopie untersucht. In Abbildung 34 ist das fern-UV CD-Spektrum von HisF_N1 nach Umrechnung in molare Elliptizität pro Aminosäure (Θ_{MRW}) im Vergleich zu den Spektren von HisF_N und HisF gezeigt.

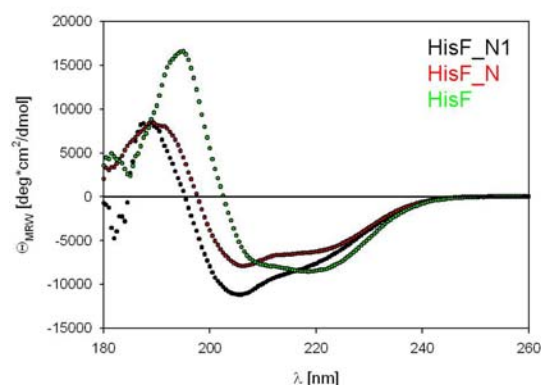


Abbildung 34: Fern-UV CD-Spektrum von HisF_N1 im Vergleich zu HisF_N und HisF.

Das Spektrum von HisF_N1 wurde in einer 1 mm Quartzküvette mit 0,2 mg/ml Protein in 50 mM KP, pH 7,5 aufgenommen. Das gezeigte Spektrum ist das Mittel aus fünf Einzelmessungen bei 25 °C. Die Daten von HisF_N und HisF, die unter identischen Bedingungen aufgenommen wurden, stammen aus Abbildung 26.

Die Auswertung der Sekundärstrukturelementanteile wurde auf dem Dichroweb-Server mit dem Programm ContinLL durchgeführt (4.2.3.1). Es wurden drei unterschiedliche Referenzproteindatensätze (Lees *et al.*, 2006) verwendet. Die Mittelung der drei Vorhersagen ist in Tabelle 9 zusammengefasst.

Tabelle 9: Vorhergesagte Sekundärstrukturelementanteile von HisF_N1 (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

Sekundärstrukturelement	HisF_N1	
	Vorhersage aus CD	kalk. aus 1thf
α -Helix	19	31
β -Strang	24	22
<i>random coil</i>	57	47

Die β -Faltblatt Anteile von HisF_N1 sind entsprechend der Berechnung voll ausgebildet. Die α -helikalen Anteile sind weniger stark ausgeprägt als im Kontext des wildtypischen HisF, weshalb insgesamt ca. 60 % der Polypeptidkette ungeordnet vorliegen.

Trifluorethanol (TFE), ein alkoholisches Kosolvens, unterstützt die Faltung von Sekundärstrukturelementen von Peptiden, indem es hydrophobe Wechselwirkungen destabilisiert, während es Wasserstoffbrücken stabilisiert (Buck, 1998). TFE wird häufig bei der Untersuchung von Peptiden eingesetzt, die nicht automatisch native Sekundärstrukturen annehmen. Um die Auswirkung von TFE auf die Sekundärstruktur von HisF_N1 zu untersuchen, wurde dieses mit steigenden Konzentrationen TFE versetzt und die fern-UV CD-Spektren gemessen. Es wurden zwei Stammlösungen von HisF_N1 (Stammlösung I: 0,2 mg/ml in 50 mM KP, pH 7,5; Stammlösung II: 0,2 mg/ml in 50 mM KP, pH 7,5 mit 80 % (v/v) TFE) hergestellt und daraus Lösungen mit verschiedenen TFE-Konzentrationen, bei gleich bleibender Proteinendkonzentration pipettiert. Als Referenz diente Puffer ohne TFE, da dieses

kein Signal im CD-Spektrum liefert. In Abbildung 35 ist das Ergebnis nach Umrechnung in molare Elliptizität pro Aminosäure (Θ_{MRW}) gezeigt.

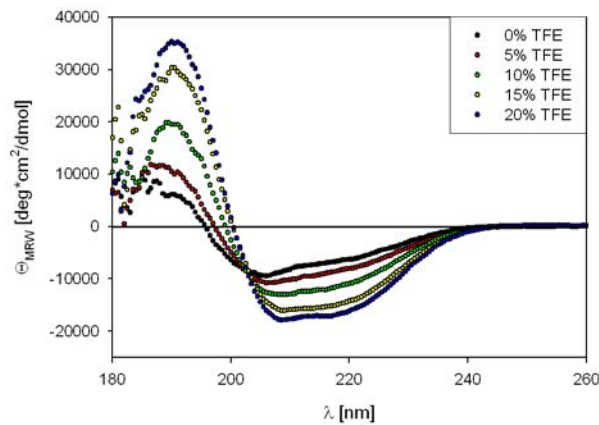


Abbildung 35: TFE-Titration von HisF_N1, verfolgt mittels CD-Spektroskopie im fern-UV Bereich.

Die fern-UV CD-Spektren wurden mit 0,2 mg/ml HisF_N1 (50 mM KP pH 7,5) und steigenden TFE-Konzentrationen [%] (v/v) aufgenommen. Die Schichtdicke der Küvette betrug 1 mm. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C.

Basierend auf den Spektren aus Abbildung 35 wurden mit dem Programm ContinLL auf dem Dichroweb-Server (4.2.3.1) die Sekundärstrukturelementanteile von HisF_N1 bei den verschiedenen TFE-Konzentrationen berechnet. Die ermittelten Werte sind in Tabelle 10 zusammengefasst.

Tabelle 10: Vorhergesagte Sekundärstrukturelementanteile von HisF_N1 in Abhängigkeit von der TFE-Konzentration (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

Konz. TFE (v/v) [%] Sekundär- strukturelement	0	5	10	15	20	kalk. aus 1thf
	Vorhersage aus CD					
α -Helix	20	24	30	43	45	31
β -Strang	20	22	26	16	18	22
<i>random coil</i>	60	54	44	41	37	47

Durch das Kosolvens TFE wird die Ausbildung von α -Helices in HisF_N1 deutlich gefördert, wodurch der Anteil an ungeordneter Struktur entsprechend zurückgeht. HisF_N1 nimmt bei einer TFE-Konzentration von ca. 10 % (v/v) die native Sekundärstrukturelementverteilung an, die anhand der Kristallstruktur von HisF (PDB-Eintrag: 1thf) für das Viertel berechnet wurde (Tabelle 10).

4.4.3.2 Fluoreszenzemissionsspektroskopie

Um Einblicke in die Tertiärstruktur von HisF_N1 zu erhalten, wurden Fluoreszenzemissionsspektren (3.4.8) aufgenommen (Abbildung 36). Das Emissionsmaximum von HisF_N1 nach Anregung bei 280 nm liegt bei 345 nm, nach vollständiger Denaturierung (8 M Harnstoff) kommt es zu einer Rotverschiebung auf 356 nm. Dies deutet daraufhin, dass HisF_N1 Tertiärstruktur besitzt, der einzige Tryptophanrest 23 im $\beta 1\alpha 1$ -loop ist im Vergleich zu HisF_N aber weniger stark vom Lösungsmittel abgeschirmt, da letzteres im nativen Zustand ein Emissionsmaximum bei 341 nm zeigt (vgl. Abbildung 27A). Wie bei HisF_N (vgl. Abbildung 27A) verschiebt sich das Emissionsmaximum von HisF_N1 nach Denaturierung in Harnstoff um ca. 10 nm.

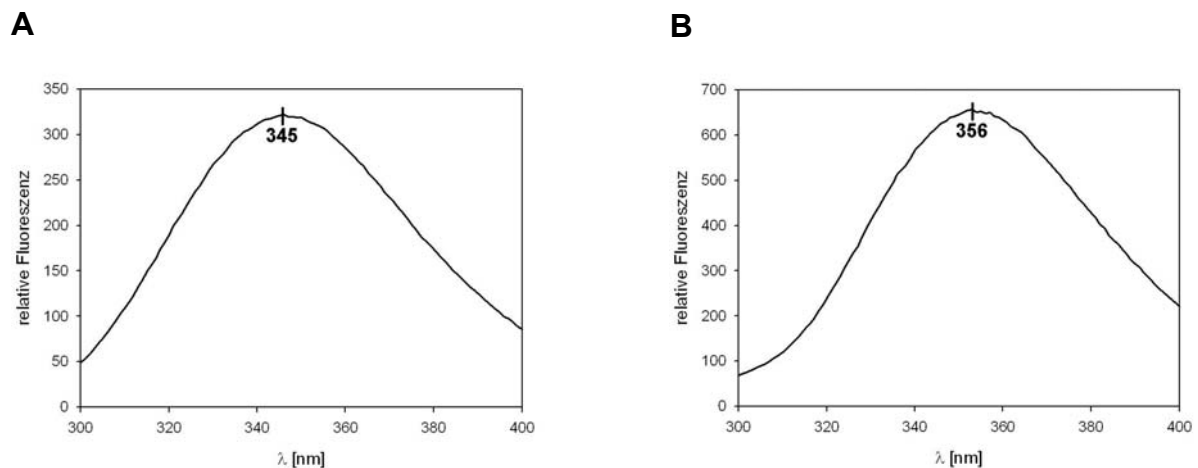


Abbildung 36: Fluoreszenzemissionsspektren von HisF_N1 im nativen und denaturierten Zustand.

- (A) Das Emissionsspektrum von HisF_N1 unter nativen Pufferbedingungen (50 mM KP, pH 7,5) nach Anregung bei 280 nm. Das Emissionsmaximum liegt bei 345 nm.
- (B) Das Emissionsspektrum von HisF_N1 unter denaturierenden Pufferbedingungen (8 M Harnstoff, 50 mM KP, pH 7,5) nach Anregung bei 280 nm. Das Emissionsmaximum liegt bei 356 nm.

Die Proteinkonzentrationen betrugen 5 μ M, die Schichtdicke der Küvette 1 cm. Die gezeigten Spektren sind das Mittel aus zehn Einzelmessungen bei 25 °C.

Die Emissionsspektren von HisF_N1 deuten im Vergleich zu HisF_N auf ein etwas weniger stark vom Lösungsmittel abgeschirmtes Tryptophan hin.

4.4.4 Stabilitätsuntersuchung von HisF_N1

Zur Bestimmung der Harnstoffkonzentration, bei der die Hälfte aller HisF_N1-Moleküle denaturiert vorliegen ($D_{1/2}$ -Wert), wurden chemische Auffaltungen (3.4.10) von HisF_N1 im fern-UV Bereich bei 220 nm am CD-Spektrometer vermessen. Eine fluorimetrische Messung der Auffaltung führte nicht zum Erfolg. Vermutlich lag dies am bereits im nativen Zustand zu lösungsmittel exponiert vorliegenden Tryptophan 23 (vgl. Abbildung 36). Abbildung 37 zeigt die normierte Auffaltungskurve von HisF_N1. Es ist eine wenig kooperative Entfaltung der Sekundärstruktur zu erkennen. Der $D_{1/2}$ -Wert liegt bei 2,7 M Harnstoff und ist damit sehr nahe am $D_{1/2}$ -Wert der HisF_N Hälfte, der 2,8 M Harnstoff beträgt (vgl. Abbildung 28A).

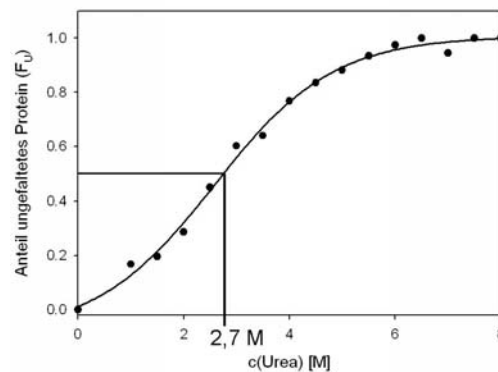


Abbildung 37: Entfaltung der Sekundärstruktur von HisF_N1 durch chemische Denaturierung mit Harnstoff.

Auffaltung der Sekundärstruktur von HisF_N1 in 50 mM KP, pH 7,5 bei 25 °C, verfolgt am CD-Spektrometer bei 220 nm. Die Proteinkonzentration betrug 0,2 mg/ml, die Schichtdicke der Küvette 1 mm. Der $D_{1/2}$ -Wert ist durch Hilfslinien eingezeichnet und liegt bei 2,7 M Harnstoff. Die Sekundärstruktur von HisF_N1 entfaltet wenig kooperativ.

Die sigmoidale Ausgleichsline soll das Auge des Betrachters leiten.

4.5 Stabilisierung von HisF_N1 mittels rationalem Design

Aufgrund der Ergebnisse aus Abschnitt 4.4 wurde die Struktur eines theoretischen (HisF_N1)₄-Moleküls mit dem SwissPDB-Viewer modelliert (Abbildung 38). Vor allem in *layer* 4 des hypothetischen ($\beta\alpha$)₈-barrels, der in wildtypischem HisF aus einem Salzbrückencluster besteht (vgl. Abbildung 8), konnten mögliche Probleme, die während der Oligomerisierung der vier ($\beta\alpha$)₂-Module auftreten, aufgedeckt werden. Durch die Anwesenheit von insgesamt vier Argininseitenketten (R5) aus β -Strang 1 und vier Glutamatresten (E46) aus β -Strang 2 könnte es zu sterischen und elektrostatischen Hinderungen bei der Ausbildung des Tetramers kommen.

Frühere Arbeiten zeigten, dass das theoretische ($\beta\alpha$)₈-barrel, bestehend aus der duplizierten C-terminalen Hälfte von HisF (HisF_CC) durch die Rekonstruktion des Salzbrückencluster in *layer* 4 (HisF_C*C) stabilisiert und kompaktiert wurde (Höcker *et al.*, 2004)

Analog sollte in HisF_N1 ein Wasserstoffbrückennetzwerk durch die Einführung von vier Glutaminseitenketten konstruiert werden. Dazu wurde der Rest Arginin 5 gegen Alanin und der Rest Glutamat 46 gegen Glutamin ausgetauscht. Alanin wurde für Position 5 gewählt, da es aufgrund seiner relativ kleinen aliphatischen Seitenkette,

die in das Zentrum des zentralen β -barrel orientiert ist, die Wechselwirkungen der eingeführten Glutaminreste möglichst wenig stört.

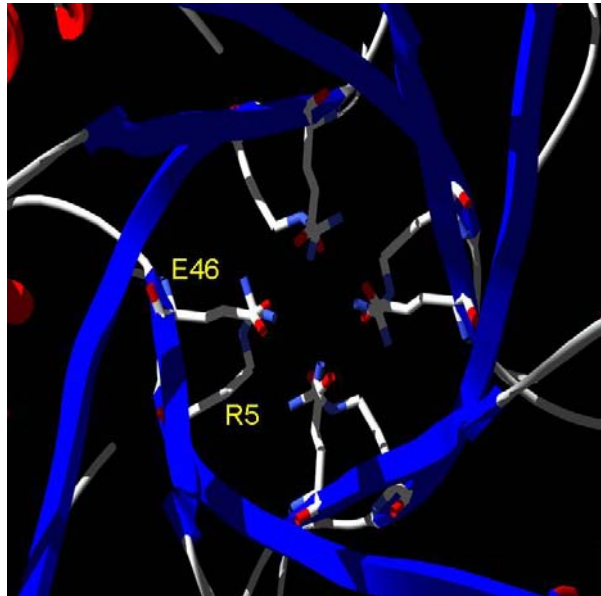


Abbildung 38: Bänderdiagramm von layer 4 des modellierten HisF_N1-Tetramers.

Die β -Stränge des zentralen β -Fasses sind in blau gezeichnet, die Aminosäureseitenketten von R5 und E46 sind als *sticks* dargestellt. In layer 4 des zentralen β -barrels kommt es durch die Seitenketten von R5 und E46 zu elektrostatischen und sterischen Hinderungen, da jedes der vier identischen $(\beta\alpha)_2$ -Module beide Reste enthält.

4.5.1 Klonierung der HisF_N1-Varianten HisF_N1_E46Q und HisF_N1_R5A_E46Q

Nukleotidaustausche, die zu den Aminosäureaustauschen E46Q und R5A führen, wurden aufeinanderfolgend mittels PCR (3.2.3) in *hisF_N1* eingeführt. Dazu wurde für den ersten Austausch E46Q eine SOE-PCR (3.2.5) mit den Austausch-Primern (5'N1_E46Q; 3'N1_E46Q) und dem T7-Primerpaar (2.8.6) durchgeführt. Als Matrize diente *hisF_N1*, das in dem Vektor pET21a vorlag (4.4.1). Nach Klonierung von *hisF_N1_E46Q* über *NdeI* und *BamHI* in den Vektor pET21a (2.7.1.1) und Verifizierung der Nukleotidabfolge durch DNA-Sequenzierung (3.2.10), wurde dieses Konstrukt zur Erzeugung der zweiten Mutante *hisF_N1_R5A_E46Q* in einer weiteren PCR als Matrize verwendet. Mit dem Austausch-Primer 5'N1_R5A und dem T7-Terminatorprimer wurde der Nukleotidaustausch, der zum Aminosäureaustausch R5A führt, in *hisF_N1_E46Q* eingeführt. Nach Klonierung über *NdeI* und *BamHI* in

den Vektor pET21a und Verifizierung der Nukleotidabfolge mittels DNA-Sequenzierung (3.2.10) wurden beide Konstrukte einer Probeexpression unterzogen.

4.5.2 Reinigung der HisF_N1-Varianten

Zunächst wurde die Expression der beiden Varianten in BL21(DE3)-Zellen im analytischen Maßstab (3.3.1) überprüft. Die Abbildungen 39A und 39B zeigen, dass beide Proteine nach Induktion mit 1 mM IPTG größtenteils in unlöslicher Form hergestellt wurden. Von HisF_N1_E46Q fanden sich nur ca. 1 %, von HisF_N1_R5A_E46Q lediglich 10 % im löslichen Teil des Zelllysates (Spur 1 in Abbildung 39A bzw. Spur 4 in Abbildung 39B). Um an den Hauptanteil der Proteinmenge, die im Pellet (Spur 2 in Abbildung 39A bzw. Spur 2 in Abbildung 39B) vorlag, zu gelangen, wurden bei der präparativen Expression (3.3.2) die Proteine mittels schrittweiser Rückfaltung (3.3.5) gewonnen. Der Überstand nach der Dialyse gegen nativen Puffer (50 mM KP, pH 7,5) wurde mit Hilfe eines *Amicon Centrifugal Device* (MWCO: 5 kDa) einkonzentriert (3.3.9) und in vier aufeinander folgenden Läufen über präparative Gelfiltration weiter aufgereinigt (3.3.4). Saubere Proteinfractionen, identifiziert über die Analyse mittels SDS-PAGE (3.4.1.2), wurden vereinigt, über Nacht bei 4 °C gegen 50 mM KP, pH 7,5 dialysiert, einkonzentriert und bei -80 °C gelagert (3.3.9). Beide HisF_N1 Varianten bilden wie das wildtypische Protein ein Homodimer über eine Disulfidbrücke zwischen zwei Cysteinseitenketten an Position 9 (Spur 4 in Abbildung 39A bzw. Spur 6 in Abbildung 39B). Die beiden Absorptionsspektren zeigen kein wesentliches Meßsignal oberhalb von 300 nm, was auf eine gute Löslichkeit und die Abwesenheit von Aggregaten hindeutet (Abbildung 39C). Aus einem Liter Kulturmedium konnten 68 mg HisF_N1_E46Q und 56 mg HisF_N1_R5A_E46Q gereinigt werden.

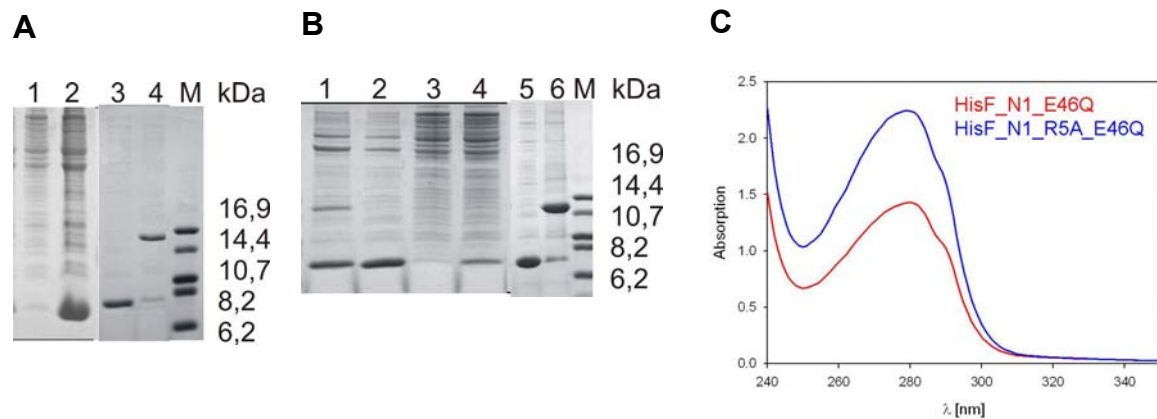


Abbildung 39: Reinigung von HisF_N1_E46Q und HisF_N1_R5A_E46Q, verfolgt über die Analyse wichtiger Fraktionen mittels SDS-PAGE auf Tris-Tricin-Gelen (20% Polyacrylamid).

- (A) Tris-Tricin SDS-PAGE (20 % Polyacrylamid) relevanter Fraktionen der Reinigung von HisF_N1_E46Q. Das Protein wird nach Induktion mit 1 mM IPTG größtenteils unlöslich (Spur 2) exprimiert. Etwa 1 % der Proteinmenge findet sich im löslichen Zellextrakt (Spur 1). Nach Rückfaltung und präparativer Gelfiltration liegt das Protein zu 98 % sauber vor (Spur 3) und bildet ein Homodimer über eine Disulfidbrücke (Spur 4) zwischen den Cysteinresten zweier Monomere.
(1: löslicher Teil des Zellextrakts nach Induktion; 2: unlöslicher Teil des Zellextrakts nach Induktion; 3: mit Mercaptoethanol versetzt; 4: ohne Mercaptoethanol; M: *small fragment marker*)
- (B) Tris-Tricin SDS-PAGE (20 % Polyacrylamid) relevanter Fraktionen der Reinigung von HisF_N1_R5A_E46Q. Das Protein wird nach Induktion mit 1 mM IPTG größtenteils unlöslich (Spur 2) exprimiert. Etwa 10 % der Proteinmenge finden sich im löslichen Zellextrakt (Spur 4). Nach Rückfaltung und präparativer Gelfiltration liegt das Protein zu 95 % sauber vor (Spur 5) und bildet ein Homodimer über eine Disulfidbrücke (Spur 6) zwischen den Cysteinresten zweier Monomere.
(1: unlöslicher Teil des Zellextrakts; 2: unlöslicher Teil nach Induktion; 3: löslicher Teil des Zellextrakts; 4: löslicher Teil nach Induktion; 5: mit Mercaptoethanol versetzt; 6: ohne Mercaptoethanol; M: *small fragment marker*)
- (C) Absorptionsspektren von gereinigtem HisF_N1_E46Q und HisF_N1_R5A_E46Q in 50 mM KP, pH 7,5 bei RT.
HisF_N1_E46Q: $c = 1,7 \text{ mg/ml}$; $A_{280/250} = 2,14$
HisF_N1_R5A_E46Q: $c = 2,8 \text{ mg/ml}$; $A_{280/250} = 2,69$

Die Parameter zur Berechnung der Proteinkonzentration der HisF_N1 Varianten sind der Tabelle 23 im Anhang zu entnehmen.

4.5.3 Bestimmung der Homogenität und des Assoziationszustandes von HisF_N1_E46Q und HisF_N1_R5A_E46Q

4.5.3.1 Analytische Gelfiltration

Die Homogenität und der Oligomerisierungsgrad der beiden gereinigten HisF_N1 Varianten wurde mithilfe von analytischer Gelfiltration überprüft (3.4.5). Von den beiden Proteinlösungen (4.5.2) wurden jeweils 100 µl auf eine Superdex 75-Säule aufgetragen und mit 50 mM KP, pH 7,5 eluiert. Die Elution wurde bei 280 nm verfolgt und ist in Abbildung 40 im Vergleich zu HisF_N1 (4.4.2.1) gezeigt.

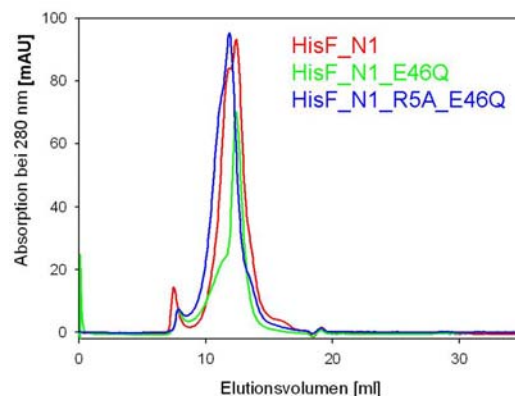


Abbildung 40: Analytische Gelfiltration von HisF_N1_E46Q und HisF_N1_R5A_E46Q im Vergleich zu HisF_N1 (Superdex 75-Säule).

Die Elution von HisF_N1_E46Q und HisF_N1_R5A_E46Q wurde bei 280 nm verfolgt. Es wurden 100 µl HisF_N1_E46Q ($c = 1,7$ mg/ml) und 100 µl HisF_N1_R5A_E46Q ($c = 2,8$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, 300 mM KCl, pH 7,5. Die Daten von HisF_N1, das unter identischen Bedingungen eluiert wurde stammen aus Abbildung 30.

Das Elutionsprofil von HisF_N1_E46Q zeigt einen Hauptpeak, der laut Kalibrierungsgerade (siehe Anhang) einem Molekulargewicht von 43 kDa und damit wie bei HisF_N1 dem Pentamer entspricht. Die zusätzliche Schulter bei einem niedrigeren Elutionsvolumen deutet auf zusätzliche Spezies mit einem größeren Molekulargewicht hin. HisF_N1_R5A_E46Q zeigt nur einen Peak, der einem Molekulargewicht von 56 kDa, entsprechend dem Heptamer entspricht. Durch die beiden Austausche geht der Anteil höherer Aggregate (Elutionsvolumen: 8 ml) zurück.

Tabelle 11 fasst die Ergebnisse der analytischen Gelfiltration zusammen.

Tabelle 11: Aus analytischer Gelfiltration ermittelte apparente und für das Monomer berechnete Molekulargewichte, sowie daraus abgeleitete Assoziationszustände von HisF_N1 und seiner Varianten.

	MW _{app} [kDa]	MW _{kalk} [kDa]	Assoziations- zustand
HisF_N1	43	8,4	Pentamer
HisF_N1_E46Q	43	8,4	Pentamer
HisF_N1_R5A_E46Q	56	8,3	Heptamer

4.5.3.2 Chemisches *cross-linking*

Um die Ergebnisse der analytischen Gelfiltration zu überprüfen und genauere Informationen über den Assoziationszustand zu erhalten, wurde chemisches *cross-linking* mit Glutardialdehyd als Quervernetzungsreagenz durchgeführt. Die Durchführung und die Analyse der Experimente erfolgten wie in Abschnitt 4.4.2.2 für HisF_N1 beschrieben. Bei beiden Varianten waren nach dem *cross-linking* Mono-, Di-, Tetramere, sowie höhere Oligomere vorhanden (Spuren 2 und 3 in Abbildung 41).

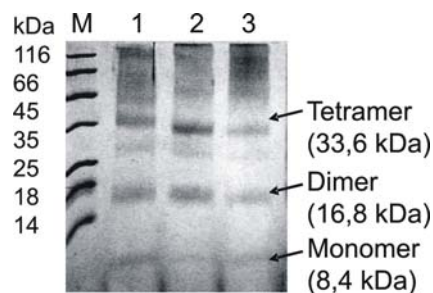


Abbildung 41: Chemisches *cross-linking* von HisF_N1 und der HisF_N1-Varianten, analysiert über SDS-PAGE auf einem Tris-Tricin-Gel (20% Polyacrylamid).

Die HisF_N1-Varianten werden wie HisF_N1 zu einem Tetramer verlinkt. Durch eine unvollständige Reaktion sind aber auch noch Mono- und Dimere zu sehen.

Die Reaktion lief in einem Volumen von 200 µl mit einer Proteinkonzentration von 0,2 mg/ml und 0,2% (v/v) Glutardialdehyd für 2 min ab. Anschließend wurde mit NaBH₄ in NaOH (Endkonz.: 0,01 M) für 20 min bei RT gestoppt.

(M = low molecular weight marker, 1 = HisF_N1; 2 = HisF_N1_E46Q; 3 = HisF_N1_R5A_E46Q)

Mono- und Dimere lassen sich zum Einen durch den unvollständigen Ablauf der chemischen Reaktion an einigen HisF_N1-Oligomermolekülen erklären. Zum Anderen sind sie ein Hinweis darauf, dass HisF_N1 nicht ausschließlich als Tetramer vorliegt. Die höheren Aggregate entstehen durch unterschiedlich lange Glutardialdehyd-Polymere. Dadurch wird auch eine Vernetzung über eine größere Distanz möglich (Peters & Richards, 1977). Wie bei HisF_N1 (Spur 1 in Abbildung 41) ist auch bei HisF_N1_E46Q und HisF_N1_R5A_E46Q eine Bande auf der Höhe des Tetramers zu sehen. Bei HisF_N1_E46Q dominiert diese deutlich gegenüber den Banden des Dimers und des Monomers. Bei HisF_N1_R5A_E46Q ist die Tetramer-Bande nicht mehr dominant, aber noch deutlich sichtbar. Die Ergebnisse des chemischen *cross-linking* liefern somit keinen Hinweis darauf, dass sich der Assoziationszustand von HisF_N1 durch die beiden Austausche E46Q und R5A wesentlich verändert.

4.5.3.3 Analytische Ultrazentrifugation von HisF_N1_E46Q und HisF_N1_R5A_E46Q

Analog wie in Abschnitt 4.4.2.3 für HisF_N1 beschrieben wurden mit HisF_N1_E46Q und HisF_N1_R5A_E46Q in 50 mM KP, pH 7,5 Gleichgewichtsläufe in der analytischen Ultrazentrifuge (3.4.6.2) durchgeführt. Im Falle der Doppelmutante wurde zusätzlich ein Sedimentationslauf (3.4.6.1) durchgeführt. Die Ergebnisse sind in Abbildung 42 gezeigt.

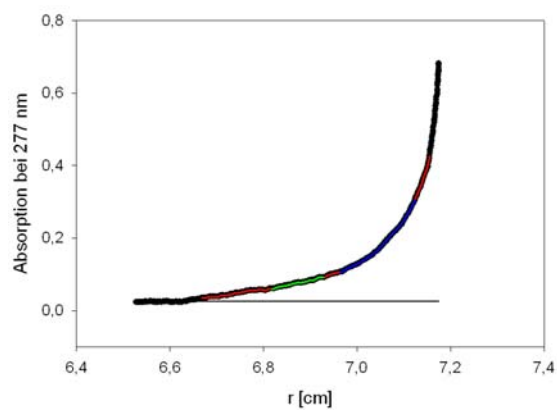
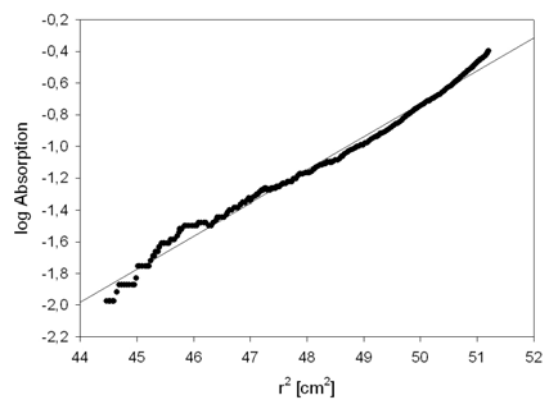
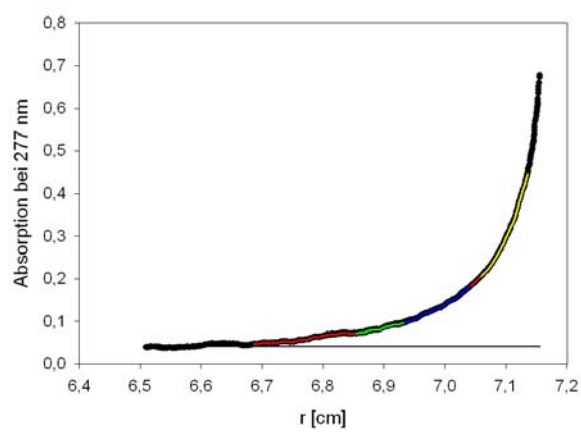
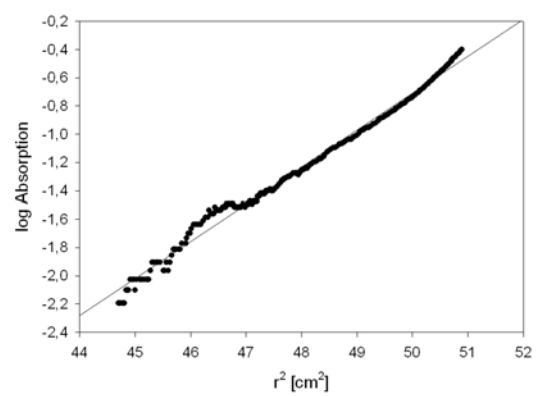
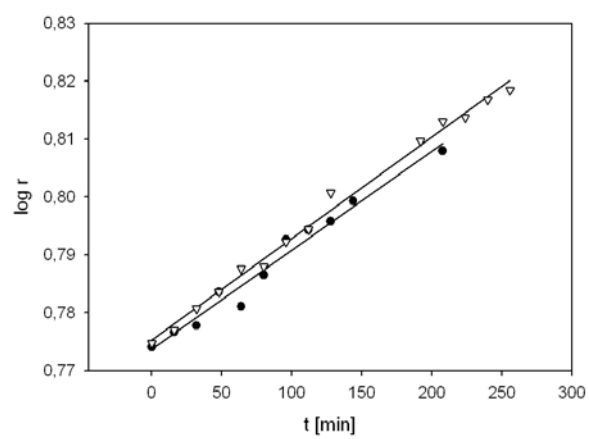
A**B****C****D****E**

Abbildung 42: Auswertung der analytischen Ultrazentrifugation von HisF_N1_E46Q und HisF_N1_R5A_E46Q.

- (A) Absorption in der Messzelle im Gleichgewichtslauf von HisF_N1_E46Q (0,2 mg/ml in 50 mM KP, pH 7,5) bei 24 °C, verfolgt bei 277 nm nach 48 h. Die durchgezogene Linie stellt die Basislinie dar, um die alle Absorptionswerte korrigiert wurden. Der rot markierte Bereich (Gesamtbereich) diente zur Erstellung der Geraden in (B). Es wurde für zwei Teilbereiche das zugehörige Molekulargewicht bestimmt: grün: 29,5 kDa; blau: 39,2 kDa.
- (B) Halblogarithmische Auftragung des Gesamtbereichs in (A). Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.2 ein Molekulargewicht von 34,6 kDa bestimmt. Dieses entspricht annähernd dem berechneten Molekulargewicht des Tetramers von 33,6 kDa.
- (C) Absorption in der Messzelle im Gleichgewichtslauf von HisF_N1_R5A_E46Q (0,2 mg/ml in 50 mM KP, pH 7,5) bei 24 °C, verfolgt bei 277 nm nach 48 h. Die durchgezogene Linie stellt die Basislinie dar, um die alle Absorptionswerte korrigiert wurden. Der rot markierte Bereich (Gesamtbereich) diente zur Erstellung der Geraden in (D). Es wurde für drei Teilbereiche das zugehörige Molekulargewicht bestimmt: grün: 41,7 kDa; blau: 42,2 kDa; gelb: 61,2 kDa.
- (D) Halblogarithmische Auftragung des Gesamtbereichs in (C). Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.2 ein Molekulargewicht von 43,5 kDa bestimmt. Dieses stellt das Molekulargewichtsmittel aller in Lösung vorhandenen Spezies dar (persönliche Mitteilung Dr. H. Durchschlag).
- (E) Auswertung des Sedimentationsgeschwindigkeitslaufes von HisF_N1_R5A_E46Q (offene Dreiecke) bei 21 °C im Vergleich zu HisF_N1 (Kreise). Die Proteinkonzentration betrug 0,2 mg/ml in 50 mM KP, pH 7,5. Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.1 ein Sedimentationskoeffizient von 2,21 S berechnet werden, der exakt dem von HisF_N1 entspricht (vgl. Abbildung 33).

Nach halblogarithmischer Auftragung der Absorption in der Messzelle von HisF_N1_E46Q (roter Bereich in Abbildung 42A = Gesamtbereich; $r = 6,65$ bis $7,15$ cm) erkennt man, dass die Messpunkte annäherungsweise auf einer Geraden zu liegen kommen (Abbildung 42B). Aus der Steigung wurde ein Molekulargewicht von 34,6 kDa berechnet, das ziemlich exakt dem Tetramer mit einem berechneten Molekulargewicht von 33,6 kDa entspricht. Es wurden zwei Teilbereiche mit unterschiedlicher Steigung und somit verschiedenen Molekulargewichten detektiert. Zum Einen wurde der in Abbildung 42A grün markierte Bereich ($r = 6,8$ bis $6,95$ cm), zum Anderen der blau markierte Bereich ($r = 7,0$ bis $7,15$ cm) ausgewertet. Die berechneten Molekulargewichte von 29,5 (grün) und 39,2 (blau) kDa liegen leicht unterhalb bzw. oberhalb des Molekulargewichts des Tetramers. In diesen Bereichen erhält man das Molekulargewichtsmittel, da die Proteinlösung nicht monodispers ist. Die Ergebnisse sprechen eindeutig dafür, dass in der Proteinlösung die tetramere Form von HisF_N1_E46Q deutlich überwiegt (persönliche Mitteilung Dr. H. Durchschlag).

Nach halblogarithmischer Auftragung der Absorption in der Messzelle von HisF_N1_R5A_E46Q (roter Bereich in Abbildung 42C = Gesamtbereich; $r = 6,7$ bis $7,13$ cm) erkennt man, dass die Messpunkte ebenfalls annäherungsweise auf einer Geraden zu liegen kommen (Abbildung 42D). Aus der Steigung wurde ein Molekulargewicht von 43,5 kDa berechnet. Es wurden drei Teilbereiche mit

unterschiedlicher Steigung und somit verschiedenen Molekulargewichten detektiert. Als erstes wurde der in Abbildung 42C grün markierte Bereich ($r = 6,85$ bis $6,95$ cm) ausgewertet und ein Molekulargewicht von $41,7$ kDa berechnet. Für den zweiten Bereich (blau in 41C, $r = 6,95$ bis $7,05$ cm) konnte ein Molekulargewicht von $42,2$ kDa ermittelt werden. Die Messpunkte des gelb markierten Bereichs ($r = 7,08$ bis $7,13$ cm) in Abbildung 42C lieferten ein Molekulargewicht von $61,2$ kDa. Die ermittelten Molekulargewichte liegen alle oberhalb des Molekulargewichts des Tetramers von $33,6$ kDa. Vermutlich ist die Proteinlösung auch in diesem Fall nicht monodispers und HisF_N1_R5A_E46Q bildet Aggregate, weshalb sich nur noch das Molekulargewichtsmittel detektieren lässt. Es lässt sich trotzdem annehmen, dass auch HisF_N1_R5A_E46Q überwiegend als Tetramer in Lösung vorliegt, da der für den Gesamtbereich gefundene Mittelwert noch nahe dem Molekulargewicht des Tetramers liegt (persönliche Mitteilung Dr. H. Durchschlag). Zur Verifizierung dieser Annahme, wurde für HisF_N1_R5A_E46Q wie in Abschnitt 4.4.2.3 beschrieben der Sedimentationskoeffizient ermittelt, um daraus das Molekulargewicht abschätzen zu können. Nach Auswertung der Messdaten des Sedimentationsgeschwindigkeitslaufes (3.4.6.1) konnte ein Sedimentationskoeffizient von $2,21$ S bestimmt werden, der exakt dem von HisF_N1 entspricht (Abbildung 42E).

Die analytische Ultrazentrifugation liefert somit einen weiteren Hinweis für das Vorhandensein des Tetramers in Lösung sowohl von HisF_N1_E46Q, als auch von HisF_N1_R5A_E46Q. In Übereinstimmung mit den Ergebnissen der chemischen Quervernetzung (vgl. Abbildung 41) scheint jedoch der R5A-Austausch die Aggregationstendenz von HisF_N1 zu erhöhen.

4.5.4 Spektroskopische Charakterisierung von HisF_N1_E46Q und HisF_N1_R5A_E46Q

4.5.4.1 Fern-UV CD-Spektroskopie

Die Sekundärstruktur der HisF_N1-Varianten wurde mit Hilfe von fern-UV Circular Dichroismus spektroskopischen Messungen untersucht. Die auf molare Elliptizität pro Aminosäure (Θ_{MRW}) normierten fern-UV CD-Spektren der beiden HisF_N1-Varianten sind in Abbildung 43 im Vergleich zu dem von HisF_N1 gezeigt. Aus der Identität der

Spektren lässt sich schließen, dass sich die Sekundärstrukturanteile durch die eingeführten Aminosäureaustausche nicht wesentlich verändern.

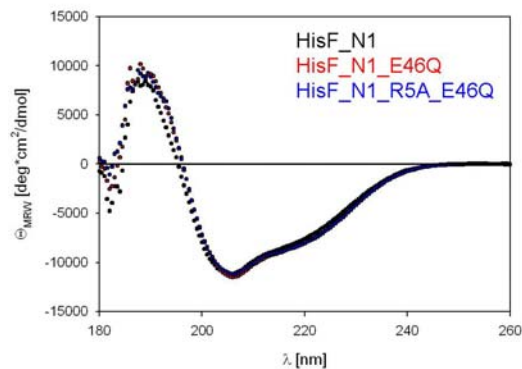


Abbildung 43: Fern-UV CD-Spektren von HisF_N1_E46Q und HisF_N1_R5A_E46Q im Vergleich zu HisF_N1.

Die Spektren von HisF_N1_E46Q und HisF_N1_R5A_E46Q wurden in einer 1 mm Quartzküvette mit 0,2 mg/ml Protein in 50 mM KP, pH 7,5 aufgenommen. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C. Die Daten von HisF_N1, die unter identischen Bedingungen aufgenommen wurden, stammen aus Abbildung 34.

Wie für wildtypisches HisF_N1 wurde auch für die beiden Varianten eine Titration mit Trifluorethanol (TFE) durchgeführt, um den Einfluss des Kosolvens auf die Sekundärstrukturausbildung zu untersuchen (4.4.3.1). Die Ergebnisse sind in Abbildung 44 gezeigt.

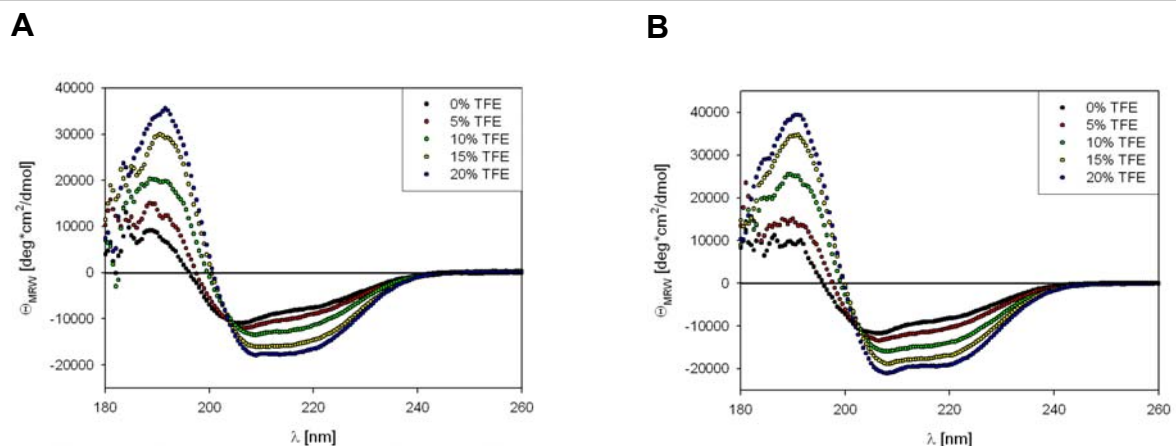


Abbildung 44: TFE-Titration der HisF_N1-Varianten, verfolgt mittels CD-Spektroskopie im fern-UV Bereich.

(A) TFE-Titration von HisF_N1_E46Q.

(B) TFE-Titration von HisF_N1_R5A_E46Q.

Die fern-UV CD-Spektren wurden mit einer Proteinkonzentration von jeweils 0,2 mg/ml in 50 mM KP, pH 7,5 und steigenden TFE-Konzentrationen [%] (v/v) aufgenommen. Die Schichtdicke der Küvette betrug 1 mm. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C.

Basierend auf den Spektren aus Abbildung 44 wurde mit dem Programm ContinLL auf dem Dichroweb-Server (4.2.3.1) der Sekundärstrukturgehalt der Varianten bei verschiedenen TFE-Konzentrationen berechnet. Die ermittelten Werte sind in Tabelle 12 gezeigt.

Tabelle 12: Vorhergesagte Sekundärstrukturelementanteile von HisF_N1_E46Q und HisF_N1_R5A_E46Q in Abhängigkeit von der TFE-Konzentration. Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

	<div>Konz. TFE (v/v) [%] Sekundär- strukturelement</div>	0	5	10	15	20	kalk. aus 1thf
		Vorhersage aus CD					
HisF_N1_ E46Q	α-Helix	22	29	33	43	44	31
	β-Strang	20	16	19	18	22	22
	random coil	58	55	48	39	34	47
HisF_N1_ R5A_ E46Q	α-Helix	25	32	40	50	55	31
	β-Strang	18	14	14	10	7	22
	random coil	57	54	46	40	38	47

Durch die Anwesenheit des Kosolvens TFE wird die Sekundärstruktur wie folgt verändert: Bei einer Konzentration von ca. 10% (v/v) TFE im Falle von HisF_N1_E46Q und bereits bei einer Konzentration von ca. 5% TFE (v/v) im Falle von HisF_N1_R5A_E46Q wird diejenige Sekundärstrukturverteilung erreicht, die aus der Kristallstruktur von HisF für das Viertel berechnet wurde. Beide Varianten reagieren somit genau wie der Wildtyp auf die Anwesenheit von TFE. Die Polypeptidkette mit dem Doppelaustausch scheint ein besseres Grundgerüst darzustellen, um die erwartete Sekundärstruktur einzunehmen.

4.5.4.2 Fluoreszenzemissionsspektroskopie

Um Einblicke in die Tertiärstruktur von HisF_N1_E46Q und HisF_N1_R5A_E46Q zu erhalten, wurden Fluoreszenzemissionsspektren (3.4.8) aufgenommen. Die Fluoreszenzemissionsmaxima von HisF_N1_E46Q und HisF_N1_R5A_E46Q nach Anregung bei 280 nm liegen bei 343 nm bzw. 345 nm. Nach vollständiger Denaturierung (8 M Harnstoff) kommt es zu einer Rotverschiebung des Emissionsmaximums um 11 nm bzw. 9 nm (Abbildung 45). Diese Ergebnisse entsprechen wildtypischen HisF_N1, was darauf hindeutet, dass sich durch die beiden Austausche die Lösungsmittelexposition des Tryptophan 23 nicht wesentlich verändert.

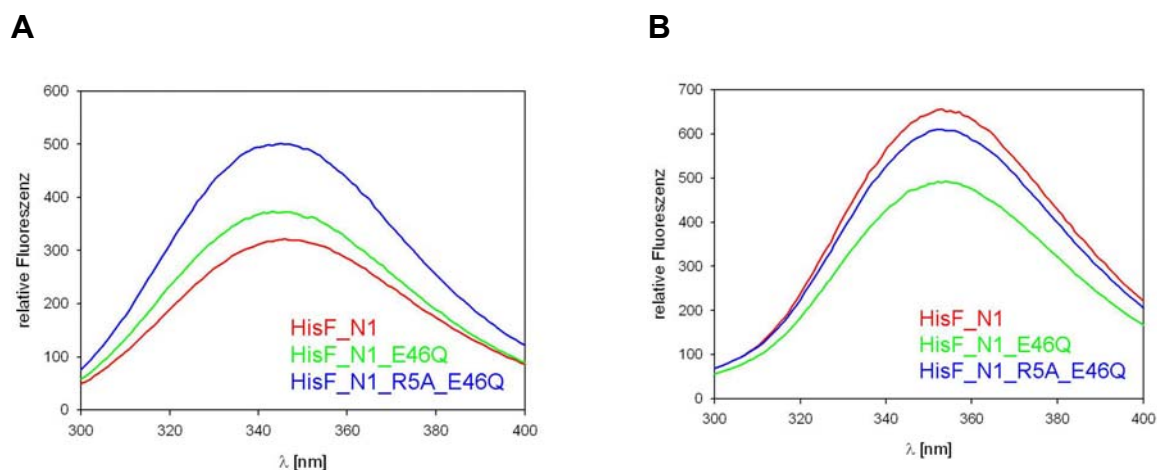


Abbildung 45: Fluoreszenzemissionsspektren von HisF_N1, HisF_N1_E46Q und HisF_N1_R5A_E46Q im nativen und denaturierten Zustand.

(A) Die Emissionsspektren von HisF_N1, HisF_N1_E46Q und HisF_N1_R5A_E46Q unter nativen Pufferbedingungen (50 mM KP pH 7,5) nach Anregung bei 280 nm. Die Emissionsmaxima liegen bei 345, 343 und 345 nm.

(B) Die Emissionsspektren von HisF_N1, HisF_N1_E46Q und HisF_N1_R5A_E46Q unter denaturierenden Pufferbedingungen (8 M Harnstoff, 50 mM KP pH 7,5) nach Anregung bei 280 nm. Die Emissionsmaxima liegen bei 356, 354 und 353 nm.

Die Proteinkonzentrationen betrugen 5 μ M, die Schichtdicke der Küvette 1 cm. Die gezeigten Spektren sind das Mittel aus zehn Einzelmessungen bei 25 °C.

4.5.5 Stabilitätsuntersuchung von HisF_N1_E46Q und HisF_N1_R5A_E46Q

Die Stabilität der HisF_N1-Varianten HisF_N1_E46Q und HisF_N1_R5A_E46Q wurde durch Denaturierung mittels Harnstoff über fern-UV CD bei 220 nm untersucht

(3.4.10). Wie bei wildtypischem HisF_N1 erfolgt die Entfaltung der Sekundärstruktur mit niedriger Kooperativität (Abbildung 46A). Auch die $D_{1/2}$ -Werte liegen mit Konzentrationen von 2,7 M (HisF_N1), 2,4 M (HisF_N1_E46Q) und 2,7 M (HisF_N1_R5A_E46Q) Harnstoff im vergleichbaren Bereich.

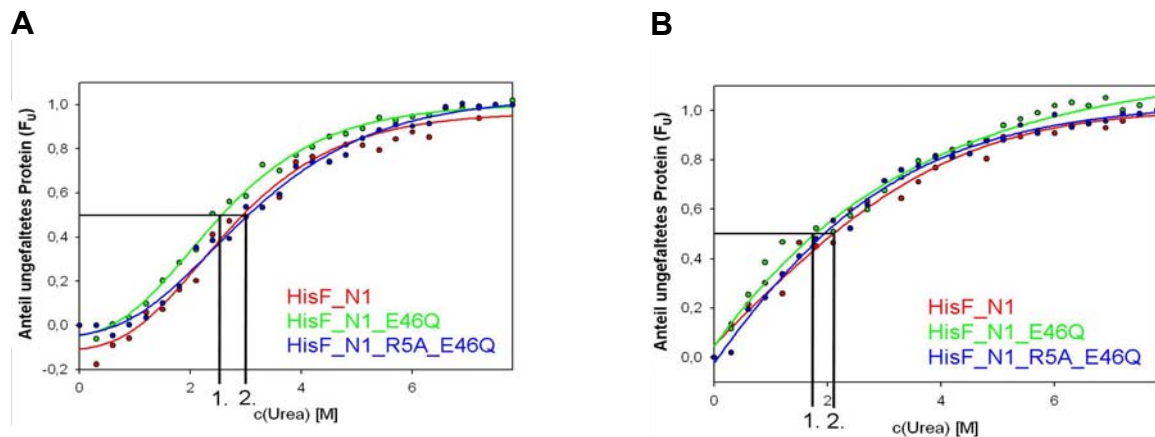


Abbildung 46: Entfaltung der Sekundärstruktur von HisF_N1, HisF_N1_E46Q und HisF_N1_R5A_E46Q durch chemische Denaturierung mit Harnstoff unter oxidierenden (A) und reduzierenden (B) Bedingungen.

- (A) Auffaltung der Sekundärstruktur von HisF_N1, HisF_N1_E46Q und HisF_N1_R5A_E46Q unter oxidierenden Bedingungen (50 mM KP, pH 7,5) bei 25 °C, verfolgt am CD-Spektrometer bei 220 nm. Die Proteinkonzentrationen betrugen jeweils 0,2 mg/ml, die Schichtdicke der Küvette 1 mm. Die $D_{1/2}$ -Werte sind durch Hilfslinien eingezeichnet und liegen bei 2,4 M (HisF_N1_E46Q) bzw. 2,7 M (HisF_N1 und HisF_N1_R5A_E46Q) Harnstoff. Die Sekundärstruktur von HisF_N1 und seinen Varianten entfaltet wenig kooperativ.
- (B) Auffaltung der Sekundärstruktur von HisF_N1, HisF_N1_E46Q und HisF_N1_R5A_E46Q unter reduzierenden Bedingungen (2 mM TCEP, 50 mM KP, pH 7,5) bei 25 °C, verfolgt am CD-Spektrometer bei 220 nm. Die Proteinkonzentrationen betrugen jeweils 0,2 mg/ml, die Schichtdicke der Küvette 1 mm. Die $D_{1/2}$ -Werte sind durch Hilfslinien eingezeichnet und liegen bei 1,6 M (HisF_N1_E46Q) bzw. 2,1 M (HisF_N1 und HisF_N1_R5A_E46Q) Harnstoff. Die Kooperativität der Entfaltung nimmt in Anwesenheit von TCEP weiter ab.

Die Ausgleichskurven sollen das Auge des Betrachters leiten.

Durch die Austausche kommt es nicht zu einer Erhöhung der Stabilität von HisF_N1, wie dies in HisF_CC der Fall war, als in diesem Kontext in *layer 4* das Salzbrückencluster nachgebaut wurde (Höcker *et. al*; 2004). Interessanter Weise entsprechen die $D_{1/2}$ -Werte der HisF_N1-Viertel ziemlich genau dem $D_{1/2}$ -Wert von 2,8 M Harnstoff, der bei der Auffaltung der HisF_N-Hälfte ermittelt wurde (vgl. Abbildung 28A).

Da zwei HisF_N1-Monomere über ihre Cysteinreste an Position 9 kovalent über eine Disulfidbrücke zu einem Homodimer verbunden sind, sollte die Entfaltung der HisF_N1-Varianten unter reduzierenden Bedingungen untersucht werden. Dies sollte Aufschluss darüber geben, inwieweit die Disulfidbrücke zur Stabilität von HisF_N1

und der beiden Varianten beiträgt. Durch eine Reduktion der Disulfidbrücke mit Tris(2-carboxyethyl)phosphin Hydrochlorid (TCEP) kommt es zu einer drastischen Erniedrigung des $D_{1/2}$ -Wertes auf 1,6 bis 2,1 M Harnstoff (Abbildung 46B). Auch die Kooperativität der Entfaltung nimmt in Anwesenheit von TCEP ab. Die Ausbildung einer Disulfidbrücke stellt somit ein wichtiges stabilisierendes Element von HisF_N1 dar.

4.6 Resümee zum Viertel HisF_N1 und den Varianten HisF_N1_E46Q und HisF_N1_R5A_E46Q

Verschiedene Methoden (analytische Gelfiltration, chemisches *cross-linking* und analytische Ultrazentrifugation) zeigen, dass sich das Viertel HisF_N1, bestehend aus den Aminosäuren 1 – 74 von HisF, in Lösung überwiegend zum Homotetramer zusammenlagert. Über das wildtypische Cystein 9 werden zwei HisF_N1-Monomere über eine Disulfidbrücke kovalent zum Homodimer verbunden, welche dann wiederum dimerisieren. In Anwesenheit des Kosolvens Trifluorethanol (TFE) in einer Endkonzentration von 10 % (v/v) entspricht die Sekundärstrukturelementverteilung von HisF_N1 im Homotetramer der aus der Kristallstruktur von HisF ermittelten. Die Entfaltung der Sekundärstruktur in Anwesenheit von Harnstoff verläuft wenig kooperativ und setzt bereits bei niedrigen Harnstoffkonzentrationen ein. Durch Reduktion der Disulfidbrücke zwischen zwei HisF_N1-Monomeren nehmen sowohl die Kooperativität als auch die Stabilität der Faltung weiter ab.

Durch rationales Design wurden in HisF_N1 die Reste Arginin 5 und Glutamat 46, die im $(\beta\alpha)_8$ -barrel von HisF in *layer* 4 zu liegen kommen, gegen Alanin bzw. Glutamin ausgetauscht. Ziel war die Stabilisierung des aus vier $(\beta\alpha)_2$ -Einheiten bestehenden $(\beta\alpha)_8$ -barrels von HisF_N1. Weder die HisF_N1_E46Q- noch die HisF_N1_R5A_E46Q-Variante unterscheiden sich jedoch in ihren Eigenschaften wesentlich von HisF_N1.

4.7 Das Viertel HisF_C1

4.7.1 Expression und Anreicherung von HisF_C1

Das *hisF_C1* Gen (Basenpaare 366 – 588 in *hisF*) lag bereits über *NdeI* und *BamHI* kloniert im Vektor pET21a vor (Bosnali, 2003). Da HisF_C1 nach Expression in BL21(DE3)-Zellen und Induktion mit IPTG (Endkonz.: 1 mM) vollständig im unlöslichen Teil des Zellextrakts vorlag, wurde es wie HisF_N1 über schrittweise Rückfaltung (3.3.3) mit sich anschließender präparativer Gelfiltration (3.3.4) gereinigt. Über SDS-PAGE (3.4.1.2) wurden diejenigen Fraktionen ermittelt in denen HisF_C1 sauber vorlag, diese wurden vereinigt, über Nacht bei 4 °C gegen 50 mM KP, pH 7,5 dialysiert (3.3.8), einkonzentriert und bei -80 °C gelagert (3.3.9). Der Reinigungserfolg ist in Abbildung 47A gezeigt. Die Ausbeute betrug 3,25 mg aus einem Kulturvolumen von einem Liter. Das Absorptionsspektrum von HisF_C1 (Abbildung 47B) zeigt ein schwaches Meßsignal oberhalb von 300 nm, was auf eine – wenn auch schwach ausgeprägte – Aggregationstendenz hindeutet.

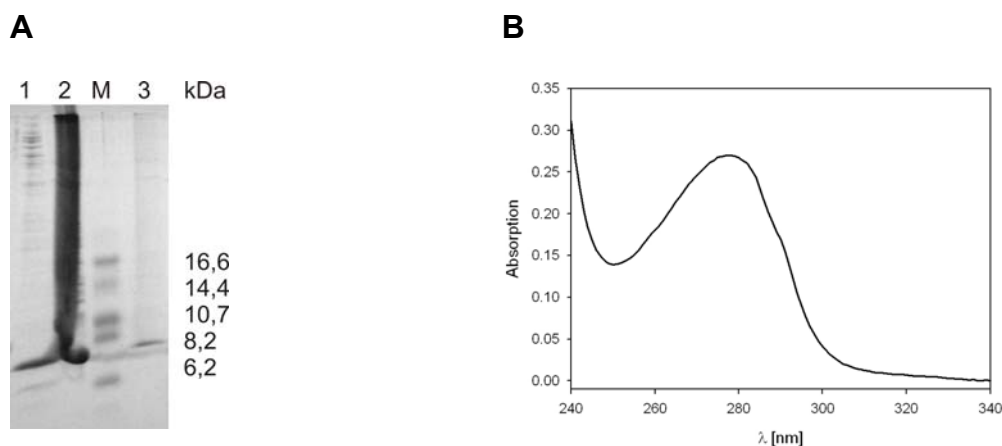


Abbildung 47: Reinigung von HisF_C1, verfolgt über die Analyse wichtiger Fraktionen mittels SDS-PAGE auf einem Tris-Tricin-Gel (20% Polyacrylamid).

- (A) Tris-Tricin SDS-PAGE (20% Polyacrylamid) relevanter Fraktionen der Reinigung von HisF_C1. Nach der Rückfaltung findet sich nur ein geringer Anteil von HisF_C1 im löslichen Teil der Dialyse (Spur 1). Ein Großteil des Proteins fällt während der Rückfaltung wieder aus (Spur 2). Durch präparative Gelfiltration können höhermolekulare Proteine entfernt werden, anschließend liegt das Protein mit einer Homogenität von ca. 90 % vor (Spur 3). (1: löslicher Teil nach der Dialyse; 2: unlöslicher Teil nach der Dialyse; 3: gereinigtes HisF_C1; M: *small fragment marker*)
- (B) Absorptionsspektrum von gereinigtem HisF_C1 in 50 mM KP, pH 7,5 bei RT.
HisF_C1: c = 0,26 mg/ml; $A_{280/250}$: 1,92

4.7.2 Untersuchung der Homogenität und des Assoziationszustandes von HisF_C1 mittels analytischer Gelfiltration

Die Homogenität von HisF_C1 wurde mithilfe von analytischer Gelfiltration auf einer Superdex 75-Säule überprüft (3.4.5). Es wurden 200 µl der Lösung mit gereinigtem HisF_C1 (4.7.1) auf die Säule aufgetragen und mit 50 mM KP, 300 mM KCl, pH 7,5 eluiert. Das Elutionsprofil von HisF_C1 zeigt drei Peaks (Abbildung 48). Peak 1 geht auf Proteinaggregate zurück, die im Ausschlußvolumen von der Säule eluieren. Die Peaks 2 und 3 entsprechen nach Vergleich mit der Kalibrierungsgeraden (im Anhang) Proteinen mit Molekulargewichten von 31,7 bzw. 8,5 kDa, entsprechend dem Tetramer (Peak 2) und dem Monomer (Peak 3). Die Ergebnisse der analytischen Gelfiltration sind in Tabelle 13 zusammengefasst.

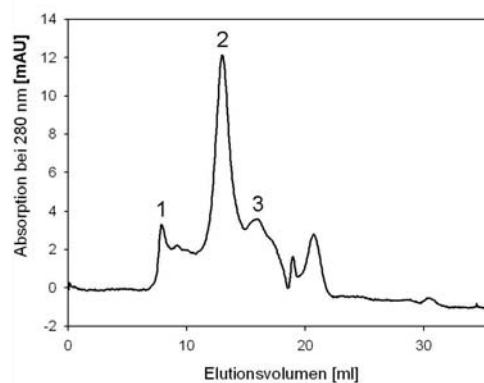


Abbildung 48: Analytische Gelfiltration von HisF_C1 (Superdex 75-Säule).

Die Elution von HisF_C1 wurde bei 280 nm verfolgt. Es wurden 200 µl HisF_C1 ($c = 0,26$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, 300 mM KCl, pH 7,5. Das Elutionsprofil von HisF_C1 zeigt drei Peaks, die höheren Aggregaten (Peak 1), dem Tetramer (Peak 2) und dem Monomer (Peak 3) entsprechen.

Tabelle 13: Aus analytischer Gelfiltration ermittelte apparente und für das Monomer berechnete Molekulargewichte von HisF_C1, sowie daraus abgeleitete Assoziationszustände.

	MW _{app} [kDa]	MW _{kalk} [kDa]	Assoziationszustand
Peak 1	> 100	8,3	Aggregate
Peak 2	31,7	8,3	Tetramer
Peak 3	8,5	8,3	Monomer

Wie man aus dem Elutionsprofil von HisF_C1 und den daraus ermittelten Assoziationszuständen erkennen kann, handelt es sich bei der gereinigten Proteinlösung um keine homogene Lösung. Die tetramere Form von HisF_C1 scheint in Lösung zu dominieren, das Protein besitzt aber auch eine gewisse Tendenz zu aggregieren bzw. liegt als Monomer vor.

4.7.3 Spektroskopische Charakterisierung von HisF_C1

4.7.3.1 Fern-UV CD-Spektroskopie

Zur Untersuchung der Sekundärstruktur von HisF_C1 wurden CD-Spektren im fern-UV Bereich aufgenommen (3.4.9). Die Messung wurde fünfmal wiederholt und die Ergebnisse gemittelt. Das auf die molare Elliptizität pro Aminosäure (Θ_{MRW}) normierte Spektrum ist zusammen mit den Spektren von HisF_C und HisF in Abbildung 49 dargestellt.

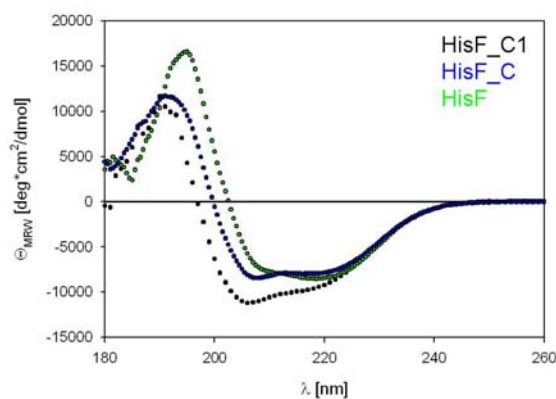


Abbildung 49: Fern-UV CD-Spektrum von HisF_C1 im Vergleich zu HisF_C und HisF.

Das Spektrum von HisF_C1 wurde in einer 1 mm Quartzküvette mit 0,2 mg/ml Protein in 50 mM KP, pH 7,5 aufgenommen. Das gezeigte Spektrum ist das Mittel aus fünf Einzelmessungen bei 25 °C. Die Daten von HisF_C und HisF, die unter identischen Bedingungen ermittelt wurden, stammen aus Abbildung 26.

Die Auswertung der Sekundärstrukturelementanteile erfolgte wiederum mit dem Programm ContinLL auf dem Dichroweb-Server (4.2.3.1) und ist in Tabelle 14 zusammengefasst.

Tabelle 14: Vorhergesagte Sekundärstrukturelementanteile von HisF_C1 (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

Sekundärstrukturelement	HisF_C1	
	Vorhersage aus CD	kalk. aus 1thf
α -Helix	28	24
β -Strang	15	37
<i>random coil</i>	57	39

Die α -helikalen Anteile von HisF_C1 sind nach der Berechnung voll ausgebildet. Die Anteile an β -Faltblatt sind deutlich weniger stark ausgeprägt als im Kontext des wildtypischen HisF, weshalb insgesamt fast 60 % der Polypeptidkette ungeordnet vorliegen.

4.7.3.2 Fluoreszenzemissionsspektroskopie

Um Einsichten in die Lösungsmittelexposition des einzigen Tryptophanrestes 156 in α -Helix 5 zu erhalten, wurde HisF_C1 mittels Fluoreszenzemissionsspektroskopie untersucht. Das Emissionsmaximum von HisF_C1 nach Anregung bei 280 nm liegt bei 343 nm und verschiebt sich nach Denaturierung mit 8 M Harnstoff um 11 nm auf 354 nm (Abbildung 50). Dies deutet darauf hin, dass HisF_C1 Tertiärstruktur besitzt, Tryptophan 156 aber deutlich mehr lösungsmittel exponiert ist als in HisF_C, das ein Emissionsmaximum bei 336 nm aufweist (vgl. Abbildung 27B).

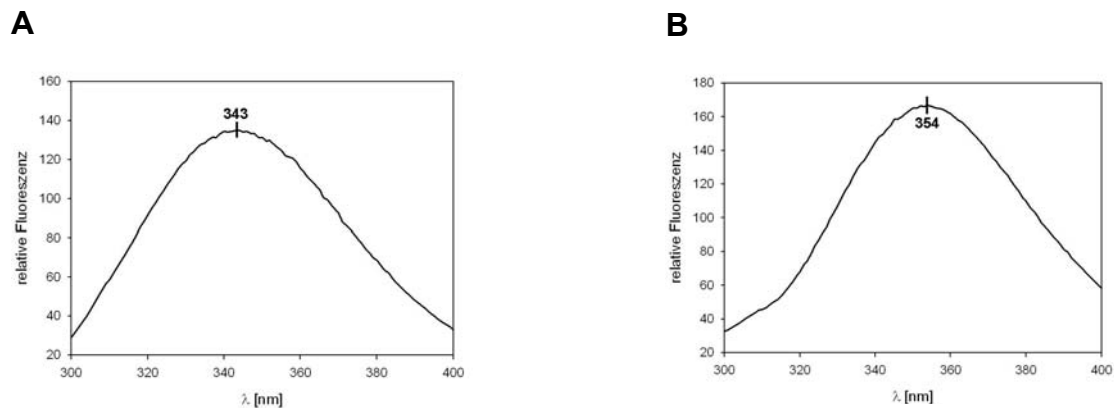


Abbildung 50: Fluoreszenzemissionsspektren von HisF_C1 im nativen und denaturierten Zustand.

- (A) Das Emissionsspektrum von HisF_C1 unter nativen Pufferbedingungen (50 mM KP, pH 7,5) nach Anregung bei 280 nm. Das Emissionsmaximum liegt bei 343 nm.
- (B) Das Emissionsspektrum von HisF_C1 unter denaturierenden Pufferbedingungen (8 M Harnstoff, 50 mM KP, pH 7,5) nach Anregung bei 280 nm. Das Emissionsmaximum liegt bei 354 nm.

Die Proteinkonzentrationen betrugen 5 μ M, die Schichtdicke der Küvette 1 cm. Die gezeigten Spektren sind das Mittel aus zehn Einzelmessungen bei 25 °C. Das Emissionsspektrum von HisF_C1 deutet auf ein im Vergleich zu HisF_C weniger stark vom Lösungsmittel abgeschirmtes Tryptophan hin.

4.7.4 Stabilitätsuntersuchung von HisF_C1

Um die Stabilität von HisF_C1 zu untersuchen, wurden Auffaltungskurven von HisF_C1 mit Harnstoff als Denaturierungsmittel sowohl mittels fern-UV CD (Sekundärstruktur) als auch fluorimetrisch (Tertiärstruktur) vermessen (Abbildung 51).

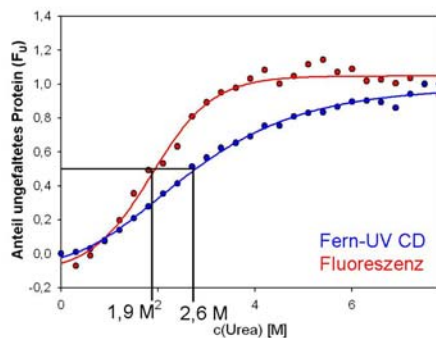


Abbildung 51: Entfaltung der Sekundär- und Tertiärstruktur von HisF_C1 durch chemische Denaturierung mit Harnstoff.

Auffaltung der Sekundärstruktur von HisF_C1 in 50 mM KP, pH 7,5 bei 25 °C, verfolgt am CD-Spektrometer bei 220 nm. Die Proteinkonzentration betrug 5 μ M, die Schichtdicke der Küvette 0,5 cm. Der $D_{1/2}$ -Wert ist durch Hilfslinien eingezeichnet und liegt bei 2,6 M Harnstoff. Die Sekundärstruktur von HisF_C1 entfaltet wenig kooperativ. Auffaltung der Tertiärstruktur von HisF_C1 in 50 mM KP, pH 7,5 bei 25 °C, verfolgt am Fluoreszenzspektrometer bei 320 nm nach Anregung bei 280 nm. Die Proteinkonzentration betrug 5 μ M, die Schichtdicke der

Küvette 1 cm. Der $D_{1/2}$ -Wert ist durch die Hilfslinien eingezeichnet und liegt bei 1,9 M Harnstoff. Die Tertiärstruktur von HisF_C1 entfaltet etwas kooperativer.

Die sigmoidalen Ausgleichslinien sollen das Auge des Betrachters leiten.

Die Entfaltung der Sekundärstruktur verläuft wie bei HisF_N1 wenig kooperativ und ergibt mit einem $D_{1/2}$ -Wert von 2,6 M Harnstoff praktisch die identische Stabilität (vgl. Abbildung 37). Die Entfaltung der Tertiärstruktur erfolgt wesentlich kooperativer und mit einem $D_{1/2}$ -Wert von 1,9 M Harnstoff. Somit ist sowohl die Sekundär- als auch die Tertiärstruktur von HisF_C1 weniger stabil als im Falle von HisF_C, dessen entsprechende $D_{1/2}$ -Werte bei 3,5 M bzw. 2,8 M Harnstoff liegen (Abbildung 28).

4.8 Stabilisierung von HisF_C1 mittels rationalem Design

4.8.1 Planung der Aminosäureaustausche in HisF_C1

Wie die vorhergehenden Messungen gezeigt haben, besitzt HisF_C1 zwar eine einigermaßen ausgeprägte Sekundär- und Tertiärstruktur, nimmt aber keinen definierten Assoziationszustand an und ist auch nicht besonders stabil. Außerdem hat HisF_C1 eine höhere Aggregationstendenz als HisF_N1 und fällt bei der Rückfaltung größtenteils wieder aus, weshalb bei der Reinigung eine deutlich geringere Ausbeute erzielt werden konnte (4.7.1). Deshalb sollten durch rationalen Aminosäureaustausch die Charakteristika von HisF_C1 verbessert werden.

Durch eine Disulfidbrücke zwischen zwei Cysteinseitenketten bildet HisF_N1 unter oxidierenden Bedingungen ein Homodimer, das sich wiederum zu einem Dimer zusammenlagert und somit ein $(\text{HisF_N1})_4$ Tetramer bildet. Der $D_{1/2}$ -Wert und die Kooperativität der Auffaltung nehmen in Anwesenheit eines reduzierenden Agens ab (vgl. Abbildung 46), wodurch die stabilisierende Wirkung der Disulfidbrücke belegt wird. Um diesen Effekt für HisF_C1 zu erreichen, sollte auch in dieses HisF-Viertel ein Cysteinrest eingeführt werden. Durch Superpositionierung der beiden Viertel mit dem Programm STAMP wurde Alanin 128 in HisF_C1 als korrespondierender Rest zu C9 in HisF_N1 identifiziert. Die Orientierung der beiden Seitenketten stimmt überein und beide liegen etwa in der Mitte des ersten β -Stranges des jeweiligen Moduls (vgl. Abbildung 23). Des Weiteren konnte in unserer Arbeitsgruppe gezeigt

werden, dass der Aminosäureaustausch Tyrosin 143 gegen Histidin eine Erhöhung der Löslichkeit und Stabilität von HisF_C**C* zur Folge hat (Seitz *et al.*, 2007). Deshalb sollte auch die zu diesem Austausch führende Mutation in *hisF_C1* eingeführt werden. Durch einen dritten Aminosäureaustausch sollte in *layer 4* ein Wasserstoffbrückennetzwerk hergestellt werden, um eine weitere Stabilisierung des potentiellen HisF_C1-($\beta\alpha$)₈-*barrels* zu erreichen. Dazu wurde Glutamat 167 gegen Glutamin ausgetauscht. In Abbildung 52 ist das Bänderdiagramm von HisF_C1 mit den auszutauschenden Aminosäureresten gezeigt.



Abbildung 52: Bänderdiagramm von HisF_C1 (Strukturausschnitt aus 1thf) mit den als sticks dargestellten auszutauschenden Aminosäureresten.

Im Bänderdiagramm sind die β -Stränge in blau und die α -Helices in rot dargestellt. A128 (gelb) ist der korrespondierende Rest zu C9 aus HisF_N1, der dort zu Bildung einer Disulfidbrücke zwischen zwei Monomeren führt (vgl. Abbildung 29). Y143 (türkis) soll nach Austausch gegen Histidin die Löslichkeit und Stabilität erhöhen (Seitz *et al.*, 2007). E167 ist der Rest, der in *layer 4* an der Bildung des Salzbrückenclusters beteiligt ist. Er soll gegen Glutamin mit dem Ziel der Bildung eines Wasserstoffbrückennetzwerkes ausgetauscht werden (vgl. 4.5, Abbildung 38).

4.8.2 Klonierung der Varianten von HisF_C1

Nukleotidaustausche, die zu den Aminosäureaustauschen A128C, Y143H und E167Q führen, wurden über SOE-PCR (3.2.5) in *hisF_C1* eingeführt. Dazu wurden die unter 2.8.3 aufgeführten Mutationsprimer in Kombination mit vektorspezifischen Primern (2.8.6) verwendet. Tabelle 15 gibt eine Übersicht über die Primerkombinationen und die verwendeten Matrizen für die PCR.

Tabelle 15: Zur Herstellung der HisF_C1 Varianten verwendete Primer und Matrizen.

HisF_C1-Variante	5'-Primer	3'-Primer	Matrize (in pET21a kloniert)
C1_A128C	C1_A128C_Ndel	T7-Terminator	<i>hisF_C1</i>
C1_Y143H	C1_Y143H	T7-Terminator	<i>hisF_C1</i>
	T7-Promotor	C1_Y143H	
C1_E167Q	C1_E167Q	T7-Terminator	<i>hisF_C1</i>
	T7-Promotor	C1_E167Q	
C1_A128C_Y143H	C1_Y143H	T7-Terminator	<i>hisF_C1_A128C</i>
	T7-Promotor	C1_Y143H	
C1_A128C_E167Q	C1_E167Q	T7-Terminator	<i>hisF_C1_A128C</i>
	T7-Promotor	C1_E167Q	
C1_Y143H_E167Q	C1_E167Q	T7-Terminator	<i>hisF_C1_Y143H</i>
	T7-Promotor	C1_E167Q	
C1_A128C_Y143H_E167Q	C1_A128C_Ndel	T7-Terminator	<i>hisF_C1_Y143H_E167Q</i>

Alle Mutanten, deren Nukleotidaustausche sukzessive von den Einzel- bis zur Dreifachvariante eingeführt wurden, wurden über *NdeI* und *BamHI* in den Vektor pET21a (2.7.1.1) kloniert. Die beabsichtigten Nukleotidabfolgen der Konstrukte wurden durch Sequenzierung verifiziert (3.2.10).

4.8.3 Expression und Anreicherung der HisF_C1-Varianten

Expressionen im analytischen Maßstab (3.3.1) zeigten, dass alle HisF_C1-Varianten in BL21(DE3)-Zellen nach Induktion mit IPTG (Endkonz.: 1 mM) im unlöslichen Teil des Zellextrakts vorlagen. Nach Expression im präparativen Maßstab (3.3.2) – die Dreifachvariante wurde in LB*-Medium versetzt mit Betain/Sorbitol (Vollmedium) angezogen – wurden die Proteine aus dem Pellet durch Behandlung mit Guanidiniumhydrochlorid (GdmCl) zuerst denaturiert und anschließend schrittweise durch Dialyse renaturiert (3.3.3). Der lösliche Anteil nach Zentrifugation der Dialyse (ca. 110 ml) wurde mit Hilfe eines *Amicon Centrifugal Device* (MWCO: 5 kDa) auf ein Volumen von ca. 40 ml einkonzentriert (3.3.9). Das Konzentrat wurde nach Filtration über einen Spritzenvorsatzfilter in vier Läufen über die präparative Gelfiltrationssäule

Superdex 75 bei 4 °C weitergereinigt, wobei 50 mM KP, 300 mM KCl, pH 7,5 als Elutionspuffer diente (3.3.4). Mit Hilfe von SDS-PAGE (3.4.1.2) wurden die Fraktionen mit sauberem HisF_C1 identifiziert, vereinigt, über Nacht bei 4 °C gegen 50 mM KP, pH 7,5 dialysiert (3.3.8) und einkonzentriert. Die Lagerung der gereinigten Proteine erfolgte bei -80 °C (3.3.9).

Die Reinigung der Varianten HisF_C1_A128C, HisF_C1_E167Q, HisF_C1_A128C_E167Q und HisF_C1_Y143H_E167Q war nicht möglich, da der Hauptteil der Proteine nach der Rückfaltung in unlöslicher Form vorlag. In Tabelle 16 sind die Reinigungserfolge und Ausbeuten aus 1 l Kulturmedium für die einzelnen Varianten zusammengefasst.

Tabelle 16: Reinigungserfolg und Ausbeuten aus 1 l Kulturmedium der einzelnen HisF_C1-Varianten.

HisF_C1-Variante	Reinigung	Ausbeute [mg/l]
C1_A128C	nicht möglich	-----
C1_Y143H	möglich	10,1
C1_E167Q	nicht möglich	-----
C1_A128C_Y143H	möglich	8,4
C1_A128C_E167Q	nicht möglich	-----
C1_Y143H_E167Q	nicht möglich	-----
C1_A128C_Y143H_E167Q	möglich	3,4

In Abbildung 53A sind die Tris-Tricin-Gelspuren der gereinigten HisF_C1-Varianten HisF_C1_Y143H, HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q gezeigt. Demnach bilden beide Varianten (HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q) mit dem Austausch A128C unter oxidierenden Bedingungen das erhoffte Homodimer über eine Disulfidbrücke zwischen den eingeführten Cysteinen. Die Absorptionsspektren der drei Proteine zur Ermittlung der Proteinkonzentration zeigen kein wesentliches Meßsignal oberhalb von 300 nm, was auf eine gute Löslichkeit der HisF_C1 Varianten hinweist (Abbildung 53B).

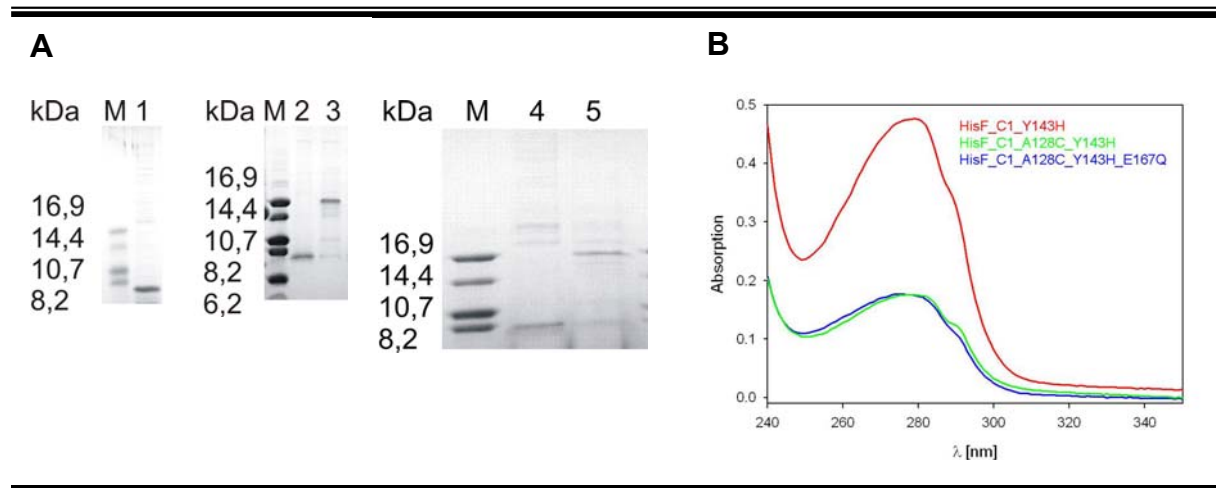


Abbildung 53: Die gereinigten HisF_C1-Varianten, analysiert mittels SDS-PAGE auf einem Tris-Tricin-Gel (20% Polyacrylamid).

- (A) Die drei Varianten (HisF_C1_Y143H (Spur 1), HisF_C1_A128C_Y143H (Spur 2) und HisF_C1_A128C_Y143H_E167Q (Spur 4)) konnten durch Rückfaltung und sich anschließende präparative Gelfiltration gereinigt werden. Die zwei Varianten mit dem künstlichen eingeführten Cystein an Position 128 bildeten unter oxidierenden Bedingungen ein Homodimer über eine Disulfidbrücke (HisF_C1_A128C_Y143H (Spur 3) und HisF_C1_A128C_Y143H_E167Q (Spur 5)).
 (1: HisF_C1_Y143H; 2: HisF_C1_A128C_Y143H mit Mercaptoethanol versetzt, 3: HisF_C1_A128C_Y143H ohne Mercaptoethanol; 4: HisF_C1_A128C_Y143H_E167Q mit Mercaptoethanol versetzt; 5: HisF_C1_A128C_Y143H_E167Q ohne Mercaptoethanol; M: *small fragment marker*)
- (B) Absorptionsspektren von gereinigtem HisF_C1_Y143H, HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q in 50 mM KP, pH 7,5 bei RT.
 HisF_C1_Y143H: $c = 0,56 \text{ mg/ml}$; $A_{280/250}$: 2,02
 HisF_C1_A128C_Y143H: $c = 0,21 \text{ mg/ml}$; $A_{280/250}$: 1,7
 HisF_C1_A128C_Y143H_E167Q: $c = 0,24 \text{ mg/ml}$; $A_{280/250}$: 1,85

Die Parameter zur Berechnung der Proteinkonzentration der einzelnen HisF_C1-Varianten sind der Tabelle 23 im Anhang zu entnehmen.

4.8.4 Untersuchung der Homogenität und des Assoziationszustandes der HisF_C1 Varianten

4.8.4.1 Analytische Gelfiltration

Die Homogenität der drei gereinigten HisF_C1-Varianten wurde mit Hilfe von analytischer Gelfiltration auf einer Superdex 75-Säule überprüft (3.4.5). Dazu wurden

200 µl der Lösung mit gereinigtem Protein (4.8.3) auf die Säule aufgetragen und mit 50 mM KP, 300 mM KCl, pH 7,5 eluiert (Abbildung 54).

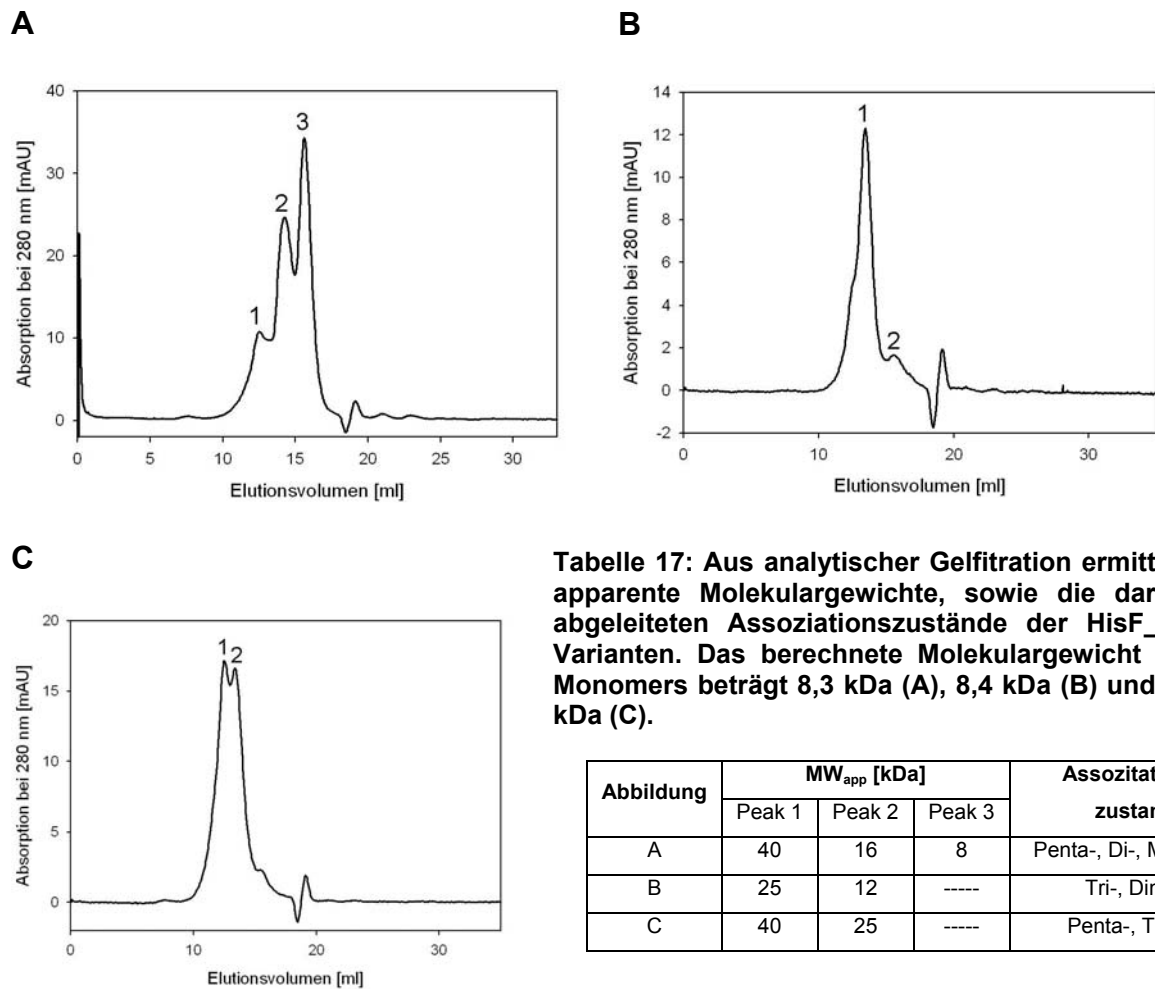


Abbildung 54: Analytische Gelfiltration der drei HisF_C1-Varianten (Superdex 75-Säule).

- (A) Die Elution von HisF_C1_Y143H wurde bei 280 nm verfolgt. Es wurden 200 µl HisF_C1_Y143H ($c = 0,56$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, pH 7,5.
- (B) Die Elution von HisF_C1_A128C_Y143H wurde bei 280 nm verfolgt. Es wurden 200 µl HisF_C1_A128C_Y143H ($c = 0,21$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, pH 7,5.
- (C) Die Elution von HisF_C1_A128C_Y143H_E167Q wurde bei 280 nm verfolgt. Es wurden 200 µl HisF_C1_A128C_Y143H_E167Q ($c = 0,24$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, pH 7,5.

Die Ergebnisse der analytischen Gelfiltration sind in Tabelle 17 zusammengefasst.

Das Elutionsprofil von HisF_C1_Y143H (Abbildung 54A) besitzt drei Peaks, die Molekulargewichten von 40 (Peak 1), 16 (Peak 2) und 8 (Peak 3) kDa entsprechen und somit auf die Ausbildung eines Penta-, Di- und Monomers hinweisen. Das Monomer ist die prominenteste Spezies in der HisF_C1_Y143H-Lösung.

Im Elutionsprofil von HisF_C1_A128C_Y143H (Abbildung 54B) sind zwei Peaks zu erkennen. Peak 1 ist der prominentere und entspricht einem Molekulargewicht von 25 kDa, was auf die Zusammenlagerung zu einem Trimer hindeutet. Es ist aber auch eine leichte Schulter bei einem geringeren Elutionsvolumen und somit Spezies mit höherem Molekulargewicht zu sehen. Es scheinen also auch höhere Oligomere als das Trimer in der Lösung enthalten zu sein. Peak 2 eluiert bei ca. 16 ml und ist nur gering ausgeprägt. Das resultierende Molekulargewicht (12 kDa) und der daraus abgeleitete Assoziationszustand deuten auf das Dimer hin.

Während der Elution von HisF_C1_A128C_Y143H_E167Q (Abbildung 54C) treten zwei Peaks mit gleich starker Absorption auf, die ineinander übergehen. Die Molekulargewichte (40 kDa bei Peak 1 und 25 kDa bei Peak 2) entsprechen denen eines Penta- und Trimers. Im Mittel würde man die Zusammenlagerung zu einem Tetramer annehmen, da durch eine nicht ganz globuläre Proteinform das Elutionsverhalten von der tatsächlichen Größe abweichen kann.

4.8.4.2 Chemisches *cross-linking*

Um die Ergebnisse der analytischen Gelfiltration zu überprüfen und genauere Informationen über den Assoziationszustand zu erhalten, wurde mit HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q chemisches *cross-linking* mit Glutardialdehyd durchgeführt. Die HisF_C1-Variante HisF_C1_Y143H wurde nicht weiter untersucht, da das Protein keinen homogenen Assoziationszustand annimmt (4.8.4.1). Die Durchführung und Analyse der Reaktionen erfolgte wie in Abschnitt 4.4.2.2 beschrieben. Das Ergebnis für HisF_C1_A128C_Y143H ist in Abbildung 55 gezeigt. Für HisF_C1_A128C_Y143H_E167Q erhielt man ein ähnliches Ergebnis. Wie für HisF_N1 ist das Mono-, Di- und Tetramer zu sehen (vgl. Abbildung 31 und 41). Das *cross-linking* ist aber nicht so eindeutig, dies liegt vermutlich an den weniger Lysinresten an der Oberfläche von HisF_C1 im Vergleich zu HisF_N1. Kleinere Oligomere kommen auch hier durch den unvollständigen Ablauf der Reaktion, höhere Oligomerisierungsgrade durch eine zu starke Vernetzung mit längeren Glutardialdehyd-Polymeren zustande.

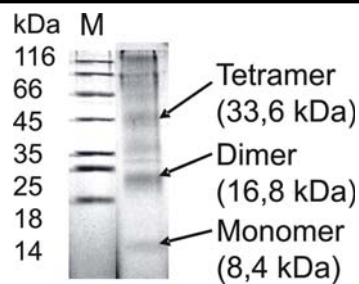


Abbildung 55: Chemisches *cross-linking* der HisF_C1-Variante HisF_C1_A128C_Y143H, analysiert mittels SDS-PAGE auf einem Tris-Tricin-Gel (20% Polyacrylamid).

Bei der HisF_C1-Variante HisF_C1_A128C_Y143H ist nach dem *cross-linking* neben dem Mono- und Dimer das Tetramer zu sehen (für die Dreifachvariante war das Ergebnis ähnlich).

Die Reaktion lief in einem Volumen von 200 µl mit einer Proteinkonzentration von 0,2 mg/ml und 0,2% (v/v) Glutardialdehyd für 2 min ab. Anschließend wurde mit NaBH₄ in NaOH (Endkonz.: 0,01 M) für 20 min bei RT gestoppt. (M: *low molecular weight marker*)

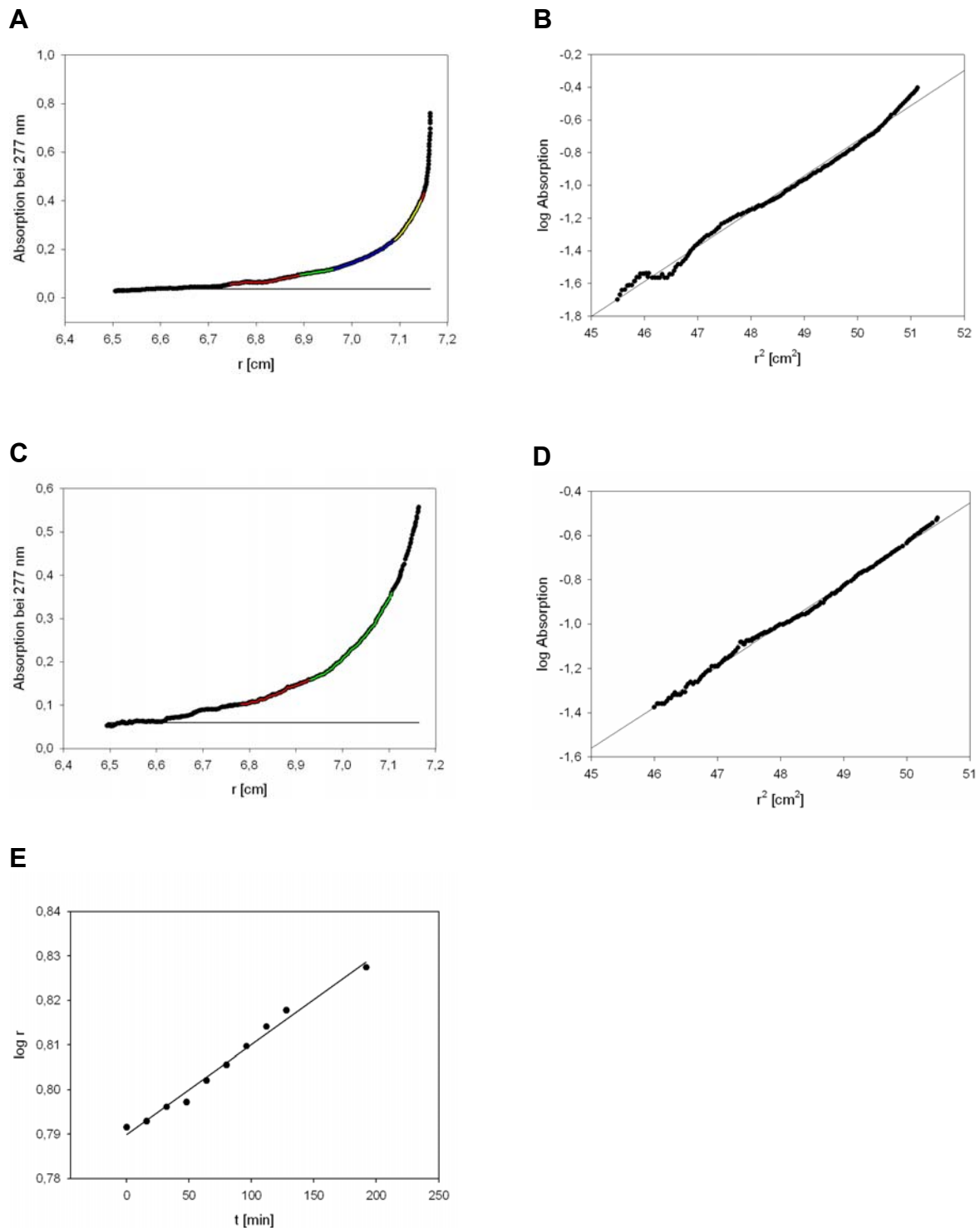
Das chemische *cross-linking* liefert somit einen ersten Hinweis für die Zusammenlagerung von HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q zu Homotetrameren.

4.8.4.3 Analytische Ultrazentrifugation

Das chemische *cross-linking* von HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q lieferte einen ersten Hinweis auf die Tetramerisierung der beiden HisF_C1-Varianten (4.8.4.2). Um dieses Ergebnis weiter zu verifizieren, wurde analytische Ultrazentrifugation durchgeführt. Dazu wurden wie in Abschnitt 4.4.2.3 beschrieben zunächst Gleichgewichtsläufe (3.4.6.2) durchgeführt (Abbildung 56A bis D). HisF_C1_A128C_Y143H_E167Q wurde zusätzlich mittels eines Sedimentationsgeschwindigkeitslaufes untersucht (Abbildung 56E).

Abbildung 56: Auswertung der analytische Ultrazentrifugation von HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q (nächste Seite).

- (A) Absorption in der Messzelle im Gleichgewichtslauf von HisF_C1_A128C_Y143H (0,2 mg/ml in 50 mM KP, pH 7,5) bei 24 °C, verfolgt bei 277 nm nach 48 h. Die durchgezogene Linie stellt die Basislinie dar, um die alle Absorptionswerte korrigiert wurden. Der rot markierte Bereich (Gesamtbereich) diente zur Erstellung der Geraden in (B). Es wurde für drei Teilbereiche das zugehörige Molekulargewicht bestimmt: grün: 24,8 kDa; blau: 34,2 kDa; gelb: 53,8 kDa.
- (B) Halblogarithmische Auftragung des Gesamtbereichs in (A). Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.2 ein Molekulargewicht von 34,6 kDa bestimmt. Dieses entspricht annähernd dem berechneten Molekulargewicht des Tetramers von 33,6 kDa.
- (C) Absorption in der Messzelle im Gleichgewichtslauf von HisF_C1_A128C_Y143H_E167Q (0,2 mg/ml in 50 mM KP, pH 7,5) bei 24 °C, verfolgt bei 277 nm nach 48 h. Die durchgezogene Linie stellt die Basislinie dar, um die alle Absorptionswerte korrigiert wurden. Der rot markierte Bereich (Gesamtbereich) diente zur Erstellung der Geraden in (D). Es wurde für einen Teilbereich das zugehörige Molekulargewicht bestimmt: grün: 32,3 kDa.
- (D) Halblogarithmische Auftragung des Gesamtbereichs in (C). Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.2 ein Molekulargewicht von 30,2 kDa bestimmt. Dieses entspricht annähernd dem berechneten Molekulargewicht des Tetramers von 33,6 kDa.
- (E) Auswertung des Sedimentationsgeschwindigkeitslaufes von HisF_C1_A128C_Y143H_E167Q (0,2 mg/ml in 50 mM KP, pH 7,5) bei 21 °C. Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.1 ein Sedimentationskoeffizient von 2,54 S berechnet. Ein Vergleich mit globulären Proteinen, die den gleichen S-Wert besitzen, deutet auf ein Molekulargewicht von 20 bis 40 kDa hin.



Nach halblogarithmischer Auftragung der Absorption in der Messzelle von HisF_C1_A128C_Y143H (roter Bereich in Abbildung 56A = Gesamtbereich; $r = 6,75$ bis $7,18$ cm) erkennt man, dass die Messpunkte annäherungsweise auf einer Geraden zu liegen kommen (Abbildung 56B). Aus deren Steigung wurde ein

Molekulargewicht von 34,6 kDa berechnet, das dem berechneten Molekulargewicht des Tetramers von 33,6 kDa entspricht. Es wurden drei Teilbereiche mit unterschiedlicher Steigung und somit verschiedenen Molekulargewichten detektiert. Zum Einen wurde der in Abbildung 56A grün markierte Bereich ($r = 6,9$ bis $6,98$ cm), zum Anderen der blau markierte Bereich ($r = 6,98$ bis $7,09$ cm) und schließlich der gelb markierte Bereich ($r = 7,09$ bis $7,15$ cm) ausgewertet. Die berechneten Molekulargewichte von 24,8 (grün), 34,2 (blau) und 53,8 (gelb) kDa liegen unterhalb (grün) bzw. oberhalb (blau, gelb) des Molekulargewichts des Tetramers (33,6 kDa), da die Proteinlösung nicht monodispers ist. Da das berechnete Molekulargewicht vom Gesamtbereich nahe am Molekulargewicht des Tetramers liegt, befindet sich in der Proteinlösung überwiegend die tetramere Form von HisF_C1_A128C_Y143H (persönliche Mitteilung Dr. H. Durchschlag). In Bodennähe (gelb markierter Bereich) ist das Molekulargewicht der Spezies um 20 kDa höher als für das Tetramer erwartet, dies deutet auf teilweise Aggregation von HisF_C1_A128C_Y143H nach 48 h bei 24 °C hin.

Nach halblogarithmischer Auftragung der Absorption in der Messzelle von HisF_C1_A128C_Y143H_E167Q (roter Bereich in Abbildung 56C = Gesamtbereich; $r = 6,79$ bis $7,10$ cm) erkennt man, dass die Messpunkte ebenfalls annäherungsweise auf einer Geraden zu liegen kommen (Abbildung 56D). Aus deren Steigung wurde ein Molekulargewicht von 30,2 kDa berechnet, das am ehesten dem Tetramer entspricht. Es gibt einen Teilbereich mit unterschiedlicher Steigung und somit abweichendem Molekulargewicht. Es wurde der in Abbildung 56C grün markierte Bereich ($r = 6,92$ bis $7,10$ cm) ausgewertet und ein Molekulargewicht von 32,3 kDa berechnet. Vermutlich ist auch hier die Proteinlösung nicht monodispers, dadurch erhält man nur noch das Molekulargewichtsmittel. Es lässt sich trotzdem schließen, dass HisF_C1_A128C_Y143H_E167Q überwiegend als Tetramer in Lösung vorliegt (persönliche Mitteilung Dr. H. Durchschlag). Zur Verifizierung dieser Annahme, wurde für HisF_C1_A128C_Y143H_E167Q wie in Abschnitt 4.4.2.3 beschrieben der Sedimentationskoeffizient ermittelt (3.4.6.1). Nach Auswertung der Messdaten des Sedimentationsgeschwindigkeitslaufes konnte ein Sedimentationskoeffizient von 2,54 S bestimmt werden (Abbildung 56E). Dieser liegt nahe am Sedimentationskoeffizienten von HisF_N1, der 2,16 S betrug. Die analytische Ultrazentrifugation liefert somit einen weiteren Hinweis für das Vorhandensein des

Tetramers in Lösung sowohl von HisF_C1_A128C_Y143H, als auch von HisF_C1_A128C_Y143H_E167Q.

4.8.5 Spektroskopische Charakterisierung

4.8.5.1 Fern-UV CD-Spektroskopie

Mit Hilfe von CD-Spektroskopie im fern-UV Bereich (3.4.9) wurde die Sekundärstruktur der drei HisF_C1-Varianten untersucht. Die auf molare Elliptizität pro Aminosäure (Θ_{MRW}) normierten Spektren sind in Abbildung 57 im Vergleich zu HisF_C und HisF dargestellt.

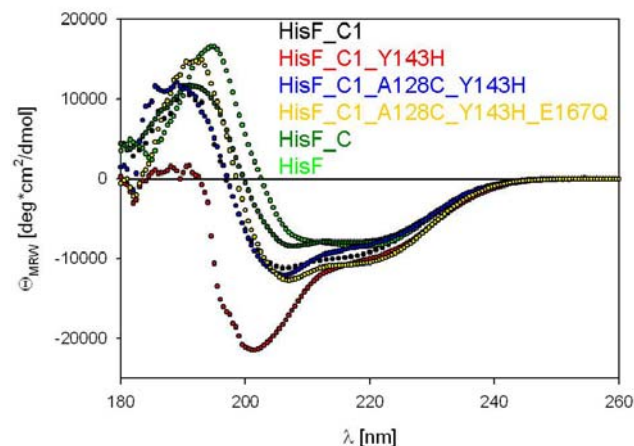


Abbildung 57: Fern-UV CD-Spektren der HisF_C1 Varianten im Vergleich zu HisF_C und HisF.

Die Spektren von HisF_C1_Y143H, HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q wurden in einer 1 mm Quarzküvette mit 0,2 mg/ml Protein in 50 mM KP pH 7,5 aufgenommen. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C. Die Daten von HisF_C1, HisF_C und HisF, die unter identischen Bedingungen aufgenommen wurden, stammen aus den Abbildungen 26 und 49.

Nach der Normierung der Messdaten wurden wie im Abschnitt 4.2.3.1 beschrieben auf dem Dichroweb-Server mit dem Programm ContinLL die Sekundärstrukturelementanteile berechnet. Die Ergebnisse der Vorhersage sind in Tabelle 18 im Vergleich zu den, aus der Struktur von HisF, berechneten Sekundärstrukturelementanteilen zusammengefasst.

Tabelle 18: Vorhergesagte Sekundärstrukturelementanteile der HisF_C1-Varianten (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

Sekundärstruktur- element	HisF_C1				
	wt	Y143H	A128C_Y143H	A128C_Y143H_E167Q	kalk. aus 1thf
	Vorhersage aus CD				
α -Helix	28	28	25	33	24
β -Strang	15	7	19	14	37
<i>random coil</i>	57	65	56	53	39

Die drei HisF_C1-Varianten bilden lediglich die α -Helices in vollem Umfang aus. Die drei Polypeptidketten besitzen keine besondere Tendenz β -Faltblatt Strukturen auszuprägen, weshalb mehr als 50 % der Aminosäurekette in Lösung ungeordnet vorliegen.

Wie für HisF_N1 und seine Varianten (4.4.3.1; 4.5.4.2) wurde auch für HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q eine Titration mit TFE durchgeführt, um den Einfluss des Lösungsmittels auf die Sekundärstrukturausbildung zu untersuchen. Die Ergebnisse sind in Abbildung 58 dargestellt.

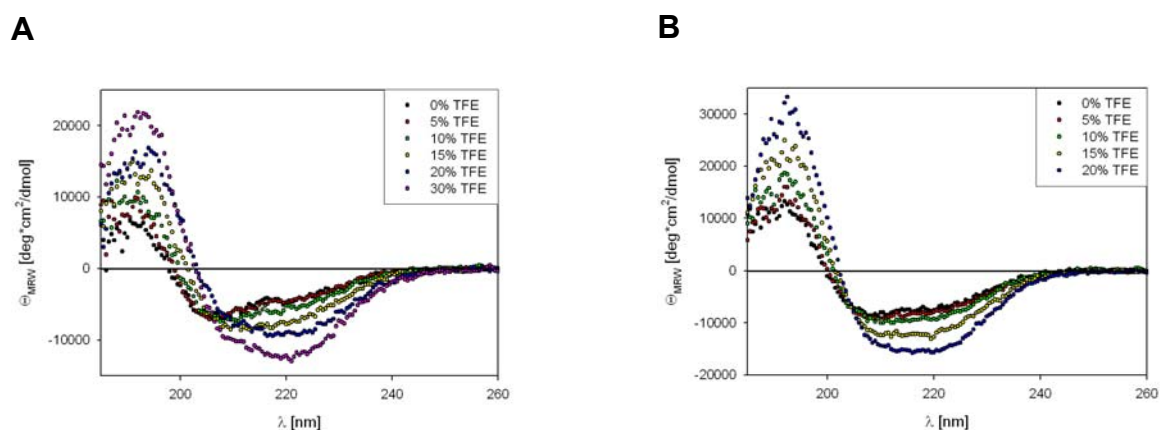


Abbildung 58: TFE-Titration der HisF_C1-Varianten, verfolgt mittels CD-Spektroskopie im fern-UV Bereich.

(A) TFE-Titration von HisF C1 A128C Y143H.

(B) TFE-Titration von HisF C1 A128C Y143H E167Q.

Die fern-UV CD-Spektren wurden mit einer Proteinkonzentration von jeweils 0,1 mg/ml in 50 mM KP, pH 7,5 und steigenden TFE-Konzentrationen [%] (v/v) aufgenommen. Die Schichtdicke der Küvette betrug 1 mm. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C.

Basierend auf den Spektren aus Abbildung 58 wurde der Sekundärstrukturgehalt der Varianten auf dem Dichroweb-Server mit dem Programm ContinLL berechnet (4.2.3.1). Die ermittelten Werte sind in Tabelle 19 zusammengefasst.

Tabelle 19: Vorhergesagte Sekundärstrukturelementanteile von HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q in Abhängigkeit von der TFE-Konzentration. Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

	<div>Konz. TFE (v/v) [%]</div> <div>Sekundär- strukturelement</div>	0	5	10	15	20	30	kalk. aus 1thf
		Vorhersage aus CD						
HisF_C1_A128C_Y143H	α -Helix	17	17	18	21	23	31	24
	β -Strang	24	27	28	29	29	23	37
	<i>random coil</i>	59	56	54	50	48	46	39
HisF_C1_A128C_Y143H_E167Q	α -Helix	22	23	25	32	40	---	24
	β -Strang	23	25	26	23	22	---	37
	<i>random coil</i>	55	52	49	45	38	---	39

Der berechnete Sekundärstrukturgehalt der beiden Varianten ohne TFE weicht hier etwas von den in Tabelle 18 berechneten Werten ab. Da kein frisch gereinigtes Protein verwendet wurde kommt es vermutlich während des Auftauens zu Umlagerungen in der Sekundärstruktur. Die Variante HisF_C1_A128C_Y143H reagiert nicht so stark auf die Anwesenheit des Kosolvens TFE wie z.B. HisF_N1_R5A_E46Q, da sich hier die Sekundärstrukturverteilung nur langsam ändert und ca. 20 % TFE benötigt werden, um vollständig die α -helikalen Anteile auszuprägen. Bei der Variante HisF_C1_A128C_Y143H_E167Q werden nur ca. 15 % TFE benötigt, um die berechnete Sekundärstrukturelementverteilung zu erreichen. Das Viertel reagiert somit viel sensitiver auf die Anwesenheit des Kosolvens.

4.8.5.2 Fluoreszenzemissionsspektroskopie

Die Tertiärstruktur der Varianten wurde mit Hilfe von Fluoreszenzemissionsspektroskopie (3.4.8) untersucht. Die Ergebnisse sind in Abbildung 59 gezeigt.

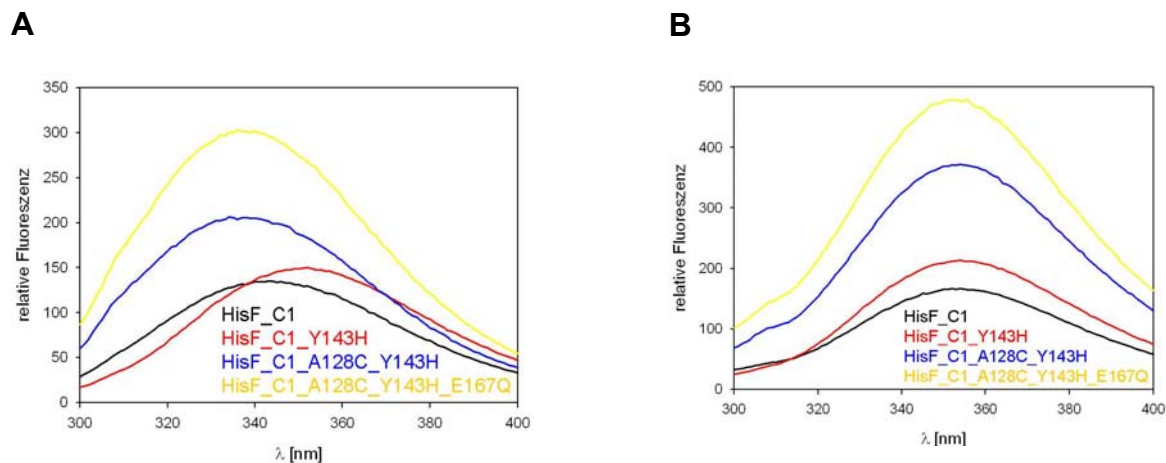


Abbildung 59: Fluoreszenzemissionsspektren von HisF_C1 und seinen Varianten unter nativen und denaturierenden Bedingungen.

(A) Emissionsspektren von HisF_C1 und seinen Varianten unter nativen Pufferbedingungen (50 mM KP, pH 7,5) nach Anregung bei 280 nm. Die Emissionsmaxima liegen bei 343 nm (HisF_C1), 352 nm (HisF_C1_Y143H), 334 nm (HisF_C1_A128C_Y143H) und 336 nm (HisF_C1_A128C_Y143H_E167Q).

(B) Emissionsspektren von HisF_C1 und seinen Varianten unter denaturierenden Bedingungen (8 M Harnstoff, 50 mM KP, pH 7,5) nach Anregung bei 280 nm. Die Emissionsmaxima verschieben sich nach Denaturierung um 2 nm (HisF_C1_Y143H) – 20 nm (HisF_C1_A128C_Y143H) auf ca. 354 nm.

Die Proteinkonzentrationen betrugen 5 μ M, die Schichtdicke der Küvette 1 cm. Die gezeigten Spektren sind das Mittel aus zehn Einzelmessungen bei 25 °C.

Obwohl durch den Austausch Y143H die Löslichkeit des Viertels erhöht wird, deutet das Emissionsmaximum von HisF_C1_Y143H bei 352 nm (zum Vgl.: HisF_C1: 343 nm) auf ein komplett lösungsmittel-exponiertes Tryptophan 156 hin. Unter denaturierenden Bedingungen verschiebt sich das Emissionsmaximum lediglich um zwei nm auf 354 nm. Aus diesem Ergebnis und dem in Abbildung 57 gezeigten fern-UV CD-Spektrum lässt sich schließen, dass diese Variante keine nennenswerte Sekundär- und Tertiärstruktur besitzt und der Austausch Y143H entgegen den Erwartungen eher perturbierend als stabilisierend wirkt (vgl. die fern-UV CD- und Fluoreszenzspektren von wildtypischem HisF_C1: Abbildung 49 und 50). Das Emissionsmaximum von HisF_C1_A128C_Y143H liegt bei 334 nm. Unter denaturierenden Pufferbedingungen kommt es zu einer Rotverschiebung des Maximums um 20 nm auf 354 nm. Auch HisF_C1_A128C_Y143H_E167Q zeigt ein Emissionsmaximum bei 336 nm das sich nach Denaturierung um 16 nm auf 352 nm verschiebt. Die Doppel- und Dreifachvariante schirmen das Tryptophan 156 somit effizienter vom Lösungsmittel ab als das wildtypische Viertel HisF_C1

(Emissionsmaximum: 343 nm) und sind auf dieser Ebene vergleichbar mit HisF_C (Emissionsmaximum: 336 nm; Abbildung 27B).

4.8.6 Stabilitätsuntersuchungen der HisF_C1-Varianten

Die Stabilität der HisF_C1-Varianten HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167 wurde über chemische Auffaltung analysiert (3.4.10). Durch das hohe Emissionsmaximum von HisF_C1_Y143H und die kaum ausgeprägte Sekundärstruktur war für diese Variante eine Messung der Auffaltung nicht möglich. Für die Doppel- und Dreifach-Variante wurde wie für das wildtypische HisF_C1 sowohl die Entfaltung der Sekundärstruktur am CD-Spektrometer, als auch die Entfaltung der Tertiärstruktur am Fluoreszenzspektrometer verfolgt (4.7.4). In Abbildung 60 sind die Ergebnisse gezeigt.

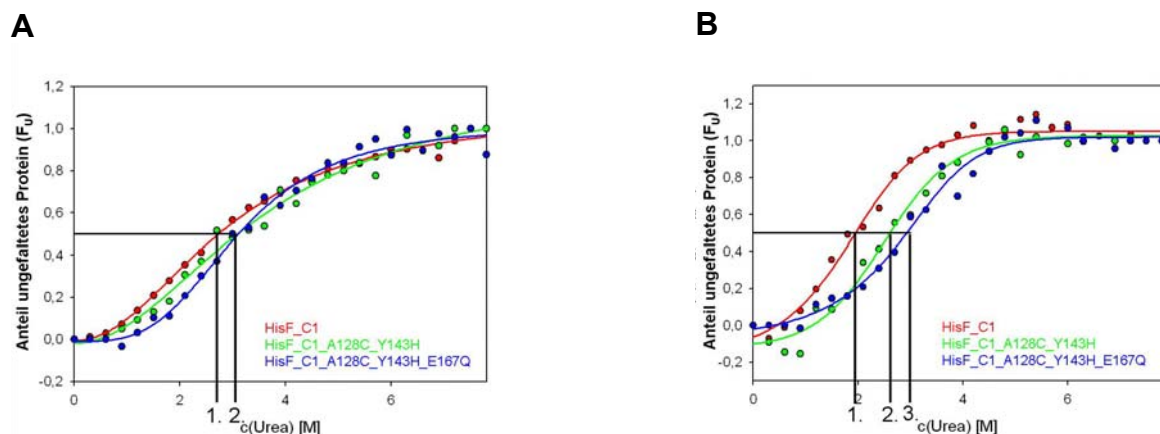


Abbildung 60: Entfaltung der Sekundär- (A) und Tertiärstruktur (B) von HisF_C1 und den Varianten HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q durch chemische Denaturierung mit Harnstoff.

- (A) Auffaltung von HisF_C1, HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q bei 25 °C, verfolgt am CD-Spektrometer bei 220 nm. Die Schichtdicke der Küvette betrug 0,5 cm. Die $D_{1/2}$ -Werte sind durch die Hilfslinien gekennzeichnet und liegen bei 2,6 M (1.; HisF_C1) und 3,1 M (2.; HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q) Harnstoff. Die Entfaltung der Sekundärstruktur verläuft wie bei HisF_N1 wenig kooperativ.
- (B) Auffaltung von HisF_C1, HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q bei 25 °C, verfolgt am Fluoreszenzspektrometer bei 320 nm nach Anregung bei 280 nm. Die Schichtdicke der Küvette betrug 1 cm. Die $D_{1/2}$ -Werte sind durch die Hilfslinien gekennzeichnet und liegen bei 1,9 M (1.; HisF_C1), 2,6 M (2.; HisF_C1_A128C_Y143H) und 2,8 M (3.; HisF_C1_A128C_Y143H_E167Q) Harnstoff. Die Tertiärstruktur der HisF_C1-Varianten entfaltet etwas kooperativer.

Die Proteinkonzentrationen betragen jeweils 5 μ M in 50 mM KP, pH 7,5.

Die sigmoidalen Ausgleichslinien sollen das Auge des Betrachters leiten.

Die Entfaltung der Sekundärstruktur aller HisF_C1-Varianten verläuft wie bei HisF_N1 eher unkooperativ und setzt schon bei niedrigen Harnstoffkonzentrationen ein (Abbildung 60A). Die $D_{1/2}$ -Werte der beiden Varianten sind mit jeweils 3,1 M Harnstoff gegenüber HisF_C1 um 0,5 M Harnstoff nach rechts verschoben und somit ähnlich denen der HisF_N1-Varianten unter oxidierenden Bedingungen (vgl. Abbildung 46A) und etwas niedriger als der Wert von HisF_C (3,5 M; Abbildung 28A). Die Stabilität der Tertiärstruktur nimmt in der Reihenfolge HisF_C1 < C1_A128C_Y143H < C1_A128C_Y143H_E167Q zu (Abbildung 60B). Der $D_{1/2}$ -Wert der Dreifach-Variante ist mit 2,8 M Harnstoff identisch mit dem Wert von HisF_C (vgl. Abbildung 28B). Dies bedeutet zum Einen einen Stabilitätsgewinn durch die Einführung des Cystein, welches zu einer Dimerisierung über eine Disulfidbrücke führt. Entsprechend sinkt nach Reduktion der Disulfidbrücke mit TCEP die Stabilität und Kooperativität der beiden Varianten HisF_C1_A128C_Y143H und HisF_C1_A128C_Y143H_E167Q (Daten nicht gezeigt). Zum Anderen scheint bei dem Viertel HisF_C1 die geplante Etablierung eines Wasserstoffbrückennetzwerks in *layer 4* größere Auswirkung zu zeigen (Vgl.: Doppel- und Dreifach-Variante), als dies in HisF_N1 der Fall war.

4.9 **Resümee zum Viertel HisF_C1 und seinen Varianten**

Die Reinigungsprozedur und die analytische Gelfiltration haben gezeigt, dass das Viertel HisF_C1, bestehend aus den Aminosäuren 123 – 196 von HisF eine hohe Aggregationstendenz besitzt und im Gegensatz zu HisF_N1 keinen definierten Assoziationszustand einnimmt.

Über ein künstlich eingeführtes Cystein an Position 128 kommt es zur Dimerisierung zweier HisF_C1-Monomere über eine Disulfidbrücke. Das kovalent verknüpfte Dimer assoziiert dann weiter zum Homotetramer, das durch chemisches *cross-linking* und analytische Ultrazentrifugation eindeutig nachgewiesen wurde. Durch rationales Design wurde der Rest Glutamat 167 in *layer 4* des putativen $(\beta\alpha)_8$ -barrels gegen Glutamin ausgetauscht. Dies führte zu einem weiteren leichten Stabilitätsgewinn, eventuell durch die Ausbildung eines Wasserstoffbrückennetzes. Sowohl HisF_C1_A128C als auch HisF_C1_A128C_E167Q enthielten zusätzlich den

Austausch Y143H, der im HisF_C**C* Kontext zu einer Erhöhung der Löslichkeit führt (Seitz *et al.*, 2007). Obwohl die Variante HisF_C1_Y143H eine weniger ausgeprägte Sekundär- und Tertiärstruktur aufweist als wildtypisches HisF_C1, erwies sich der Einbau des Histidinrestes im Kontext der beiden anderen Mutationen als vorteilhaft, da in seiner Abwesenheit sowohl HisF_C1_A128C als auch HisF_C1_A128C_E167Q bei der Rückfaltung vollständig ausfielen (vgl. Tabelle 16).

Die Sekundärstruktur ist in den HisF_C1-Varianten verschieden stark ausgeprägt, konnte aber durch das Kosolvens TFE soweit verändert werden, dass in Anwesenheit von ca. 10 – 20 % (v/v) TFE die Sekundärstrukturverteilung angenommen wird, die aus der Kristallstruktur von HisF berechnet werden kann.

4.10 Die Viertel HisF_N2 und HisF_C2

4.10.1 Klonierung von HisF_N2 und HisF_C2 mit Intein als Fusionprotein und deren Reinigung

Da HisF_N2 (Aminosäure 75 -122 in HisF) und HisF_C2 (Aminosäure 197 – 254 in HisF) allein nicht gereinigt werden können und die Expression und Reinigung mit Glutathion-S-Transferase (GST) als Fusionsprotein ebenfalls nicht zum Erfolg führte (Bosnali, 2003), wurde nunmehr Intein als Fusionspartner verwendet (3.3.6). Das *hisF_N2*-Gen wurde mit dem Primerpaar *hisF_N2_5'NdeI* und *hisF_N2_3'intein* in einer PCR amplifiziert. Als Matrize diente *hisF_N*, in das über eine SOE-PCR (3.2.5) zuvor die Mutation für den Aminosäureaustausch F86W im Bereich von HisF_N2 eingeführt worden war. *hisF_C2* wurde mit dem Primerpaar *hisF_C2_5'NdeI* und *hisF_C2_3'intein* amplifiziert. Als Matrize diente *hisF_C*, in das ebenfalls über eine SOE-PCR (3.2.5) zuvor die Mutation für den Aminosäureaustausch F227W im Bereich von HisF_C2 eingeführt worden war. Durch den Einbau der Tryptophanreste war in beiden Vierteln ein spektroskopischer Marker vorhanden. Über *NdeI* und *XhoI* wurden die Fragmente in den Vektor pTYB1 (2.7.2.1) kloniert, wodurch an beide Proteine N-terminal bei der Expression das *Saccharomyces cerevisiae* Intein angehängt wird. Nach Verifizierung der gewünschten Nukleotidabfolge durch Sequenzierung (3.2.10) wurde zur Expression der Fusionsproteine der Stamm *E. coli*

ER2566 mit den beiden Plasmiden transformiert (3.1.3.2). Durch Probeexpression (3.3.1) konnten die Fusionsproteine HisF_N2/Intein/CBD und HisF_C2/Intein/CBD im löslichen Teil der Zellpräparation nachgewiesen werden. Da die Menge an rekombinanten Fusionsproteinen selbst nach Induktion mit 1 mM IPTG sehr gering war erfolgte dies mit Hilfe eines *Western Blots* (3.4.3) unter Verwendung eines anti-CBD-Antikörpers (im IMPACT-Kit enthalten). Nach Präparation der Rohextrakte (3.3.2) wurden diese auf die vorbereitete Chitin-Säule geladen und die Affinitätsreinigung durchgeführt (3.3.6). Nach der Elution wurden die Fraktionen mittels SDS-PAGE (3.4.1.2) analysiert. In Abbildung 61 ist das Ergebnis von HisF_N2 gezeigt, das dem von HisF_C2 entsprach.

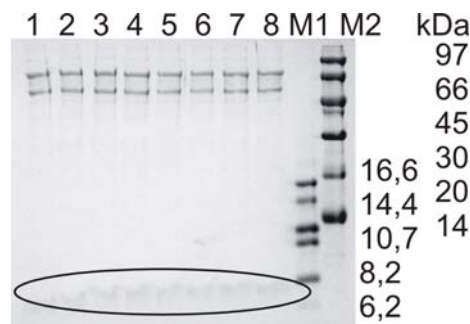


Abbildung 61: Reinigungsversuch von HisF_N2 mit Intein als Fusionsprotein, analysiert mittels SDS-PAGE auf einem Tris-Tricin-Gel (20 % Polyacrylamid).

Die auffälligsten Banden bei der Elution von der Chitin-Affinitätssäule (Spuren 1 – 8) entsprechen Proteinen mit 67 kDa (= Fusionsprotein) und mit 55 kDa (= Intein). Das Zielprotein HisF_N2 ist nur als ganz schwache Bande bei ca. 6 kDa zu sehen. Die Reinigung von HisF_C2 erbrachte ein ähnliches Ergebnis.

(1 bis 8: Fraktionen der Elution von der Chitin-Säule; M1: *small fragment marker*, M2: *low molecular weight marker*)

Offenbar konnte hauptsächlich nur die unerwünschte Elution als Fusionsprotein bzw. des Intein von der Säule erreicht werden. Dagegen konnte nur eine geringe Menge HisF_N2 und HisF_C2 identifiziert werden, vermutlich aufgrund hoher Aggregationstendenz. Auch die Variation der Pufferbedingungen (Salzgehalt, pH-Wert) und der Inkubationsbedingungen während des Selbstspaltungsprozesses (Temperatur, Inkubationsdauer) erbrachten nicht den gewünschten Erfolg. Die Funktionstüchtigkeit des Systems wurde mit Hilfe der im IMPACT-Kit enthaltenen Testvektoren bestätigt (Daten nicht gezeigt).

4.10.2 Klonierung einer Faktor Xa-Schnittstelle zwischen die Viertel HisF_N1 und HisF_N2 bzw. HisF_C1 und HisF_C2

Die Hälften HisF_N und HisF_C lassen sich gut reinigen und falten sich annähernd nativ (vgl. Abbildung 24 und 26), weshalb zwischen die Viertel HisF_N1, HisF_N2 und HisF_C1, HisF_C2 in diese Hälften die Nukleotidsequenz einer Faktor Xa-Schnittstelle kloniert werden sollte. Dadurch sollte nach der Reinigung der Hälften und proteolytischer Spaltung, eine Trennung von HisF_N2 und HisF_C2 vom jeweiligen Partnerviertel möglich werden. Durch den N-terminalen His₆-Tag an HisF_N2 und HisF_C2 (nach Klonierung über die Schnittstelle für *Xho*I in den Vektor pET21a (2.7.1.1)) sollten die Viertel anschließend über Nickel-Chelataffinitätschromatographie angereichert werden.

In das *hisF_N2*-Gen wurde in einer ersten SOE-PCR (3.2.5) mit den Primerpaaren 5'N2_F86W und T7-Terminator sowie 3'N2_F86W und T7-Promotor ein Nukleotidaustausch eingeführt, welcher zu dem Aminosäureaustausch F86W (spektroskopischer Marker in HisF_N2) führt. In einer zweiten SOE-PCR (3.2.5) diente dieses Produkt als Matrize für die Primerpaare *hisF_N1_5'Nde*I und *hisF_N1XaN2.rev*, sowie *hisF_N1XaN2.for* und *hisF_N2_5'intein*, um die Nukleotidsequenz, die für die Faktor Xa-Erkennungssequenz kodiert, einzufügen.

In das *hisF_C2*-Gen wurde analog in einer ersten SOE-PCR (3.2.5) mit den Primerpaaren 5'C2_F227W und T7-Terminator sowie 3'C2_F227W und T7-Promotor ein Nukleotidaustausch eingeführt, welcher zu dem Aminosäureaustausch F227W (spektroskopischer Marker in HisF_C2) führt. In einer zweiten SOE-PCR (3.2.5) diente dieses Produkt als Matrize für die Primerpaare *hisF_C1_5'Nde*I und *hisF_C1XaC2.rev*, sowie *hisF_C1XaC2.for* und *hisF_C2_5'intein*, um die Nukleotidsequenz, die für die Faktor Xa-Erkennungssequenz kodiert, einzufügen. Die *hisF_N*- und *hisF_C*-Gene mit den eingefügten Faktor Xa-Schnittstellen wurden über *Nde*I und *Xho*I in den Vektor pET21a kloniert, überprüft (3.2.10) und BL21(DE3)-Zellen mit den resultierenden Plasmiden transformiert (3.1.3.2). Wie HisF_N und HisF_C finden sich HisF_N1XaN2 und HisF_C1XaC2 nach Induktion mit 1 mM IPTG im unlöslichen Teil des Zellextrakts. Nach präparativer Expression (3.3.2) wurden die Proteine durch Rückfaltung auf einer HisTrapTM-Säule gereinigt (3.3.5). Das Ergebnis der Reinigung von HisF_N1XaN2 ist in Abbildung 62A gezeigt und entspricht dem für HisF_C1XaC2. Die sauberen Proteinfractionen wurden vereinigt, über Nacht bei 4 °C

gegen 50 mM KP, pH 7,5 dialysiert und bei -80 °C gelagert. Die Ausbeuten aus 1 l Kulturmedium betrugen 16,6 mg HisF_N1XaN2 und 4,3 mg HisF_C1XaC2.



Abbildung 62: Analyse der Reinigung von HisF_N1XaN2 mittels Rückfaltung auf einer HisTrap™-Säule und der anschließenden proteolytischen Spaltung. Gezeigt ist jeweils eine SDS-PAGE auf einem Tris-Glycin-Gel (12,5% Polyacrylamid) (A) bzw. Tris-Tricin-Gel (20% Polyacrylamid) (B).

- (A) Analyse ausgewählter Fraktionen während der Elution von HisF_N1XaN2 von einer HisTrap™-Säule nach Rückfaltung. HisF_N1XaN2 liegt nach Elution in reiner Form vor (Spuren 2 - 10). Die Spuren 2 bis 9 wurden vereinigt und gegen 50 mM KP, pH 7,5 dialysiert. (1 - 12: Fraktionen der Elution von der HiTrap-Säule; M: *low molecular weight marker*)
- (B) Analyse der proteolytischen Spaltung von HisF_N1XaN2 mit Faktor Xa. Die Abbaubanden in Spur 2 resultieren aus der Fragmentierung des N-terminalen Viertels HisF_N1. Es ist keine definierte Bande bei ca. 6 kDa (HisF_N2) zu sehen. Ohne Faktor Xa findet kein merklicher Abbau statt (Spur 1). (1: Kontrolle ohne Faktor Xa; 2: inkubiert mit Faktor Xa für 10 h; M: *small fragment marker*)

Der Verdau der gereinigten Proteine mit der Protease Faktor Xa wurde wie in 3.3.7 beschrieben durchgeführt und mittels SDS-PAGE auf einem Tris-Tricin-Gel (20 % Polyacrylamid) analysiert (3.4.1.2). In Abbildung 62B ist exemplarisch das Ergebnis für HisF_N1XaN2 gezeigt. Das Konstrukt wird zwar verkürzt, die Banden auf dem Gel entsprechen aber alle Fragmenten des N-terminalen Teil des Proteins (also HisF_N1), wie durch N-terminale Sequenzierung (geführt von der Arbeitsgruppe von Prof. Deutzmann an der Universität Regensburg) gezeigt werden konnte. Nach dem Verdau von HisF_C1XaC2 zeigte sich ein ähnliches Bild. Eine mögliche Erklärung für dieses Ergebnis ist eine Maskierung der Faktor Xa-Schnittstelle, bei welcher der $\alpha\beta$ -loop zwischen HisF_N1 und HisF_N2 bzw. HisF_C1 und HisF_C2 von der umliegenden Struktur verdeckt wird. Dadurch wird ein vollständiger proteolytischer Verdau des Fusionsproteins unmöglich.

4.10.3 Festphasenpeptidsynthese von HisF_N2

Da alle Reinigungsversuche von HisF_N2 zu keinem Ergebnis geführt hatten, wurde das Protein in der Arbeitsgruppe von Frau Dr. Cabrele (Institut für Organische Chemie der Universität Regensburg) mit Hilfe der Fmoc-Methode synthetisiert (3.3.10). Dabei wurden drei Aminosäureaustausche vorgenommen (Abbildung 63). Phenylalanin 120 wurde gegen Tryptophan, zur Einführung eines spektroskopischen Markers, Serin 101 gegen Cystein, zur Bildung einer Disulfidbrücke zwischen zwei HisF_N2-Monomeren, und Lysin 99 gegen Glutamin, zur Ausbildung eines Wasserstoffbrückennetzwerks in *layer* 4, ausgetauscht (vgl. 4.5 und 4.8). Aus einer Superpositionierung der vier HisF-Viertel ging eigentlich Glycin 80 als korrespondierender Rest zu C9 in HisF_N1 hervor. Eine Untersuchung der theoretischen Tetramerstruktur von HisF_N2 mit dem Programm Moloc ermittelte jedoch S101 als geeignete Position zur Einführung des Cysteinrestes (Marco Bocola Lehrstuhl Biochemie II der Universität Regensburg, unveröffentlichte Daten).

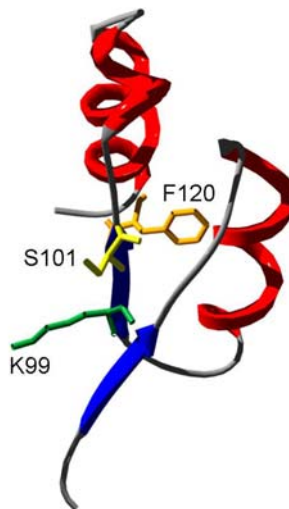


Abbildung 63: Bänderdiagramm von HisF_N2 (Strukturausschnitt aus 1thf) mit den als sticks dargestellten auszutauschenden Aminosäureresten.

Im Bänderdiagramm sind die β -Stränge in blau und die α -Helices in rot dargestellt. S101 (gelb) soll zur Ausbildung einer intermolekularen Disulfidbrücke gegen Cystein ausgetauscht werden. F120W (orange) dient der Einführung eines spektroskopischen Markers in HisF_N2. Glutamin soll anstelle von K99 (grün) in *layer* 4 bei der Ausbildung eines Wasserstoffbrückennetzwerkes helfen (vgl. 4.5, Abbildung 38).

HisF_N2 wurde vom C-Terminus (Serin 122) aus synthetisiert, wobei in regelmäßigen Abständen Zwischenprodukte mittels RP-HPLC gereinigt und angereichert wurden.

In einer ersten Phase wurde bis zu Glycin 81 synthetisiert, wobei die Austausch F120W und S101C eingeführt wurden. Dieses Zwischenprodukt zeigt im Elutionsprofil der RP-HPLC nur einen symmetrischen Peak (Abbildung 64A), was auf keinerlei Aggregation hindeutet. Durch Anheften der letzten sechs Aminosäuren bis hin zu Isoleucin 75, die β -Strang 3 in HisF entsprechen (vgl. Abbildung 23), kommt es zu einer starken Aggregation des Endproduktes (Abbildung 64B). Das Volllängenfragment HisF_N2_S101C_F120W konnte dem Peak bei einer Elutionszeit von ca. 22 min (mit einem Pfeil in Abbildung 64B markiert) zugeordnet werden. Mit Hilfe einer massenspektrometrischen Analyse konnte die Reinheit des Produktes nachgewiesen werden (Daten nicht gezeigt). Es konnten ca. 0,5 mg HisF_N2_S101C_F120W in lyophilisierter Form dargestellt werden. Dieses Lyophilisat wurde in 1 ml 50 mM KP, pH 7,5 gelöst und anschließend mit Hilfe eines *Midi D-Tube Dialyzer* (Calbiochem) vier Tage gegen 50 mM KP, pH 7,5 dialysiert (3.3.8).

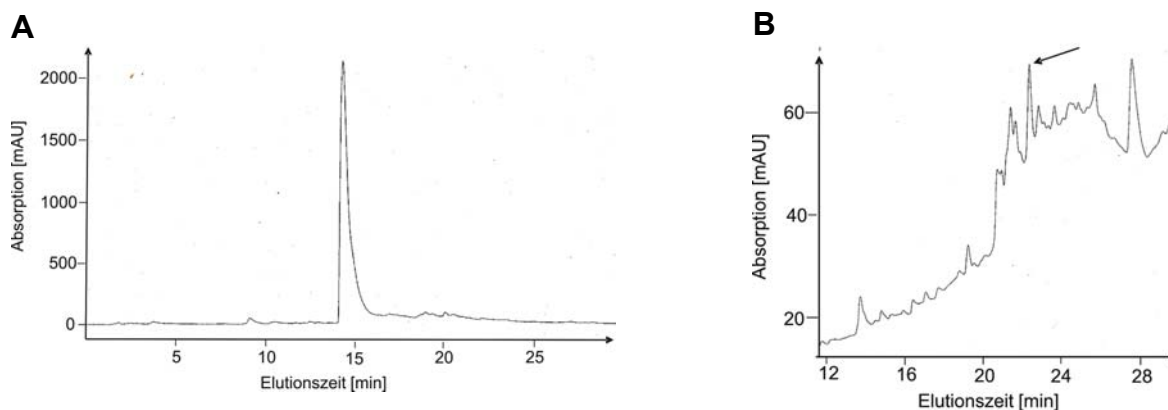


Abbildung 64: Überprüfung des Fortschrittes der Festphasenpeptidsynthese von HisF_N2 mittels RP-HPLC.

- (A) Elutionsprofil der RP-HPLC-Analyse des Peptidfragments der 42 C-terminalen Aminosäurereste von HisF_N2 (G81 – S122). Es fehlt der vollständige β -Strang 3. Dieses Zwischenprodukt zeigt nur einen homogenen Peak und keine Aggregation.
- (B) Elutionsprofil der RP-HPLC-Analyse des Endlängenfragmentes HisF_N2. Durch das Anheften des β -Stranges 3 kommt es zu einer starken Aggregation. Unter dem Peak bei einer Elutionszeit von ca. 22 min (mit Pfeil markiert) fand sich das Endprodukt. Die Ausbeute betrug 0,5 mg HisF_N2_S101C_F120W.

Die Analysen wurden von der Arbeitsgruppe Dr. Cabrele am Institut für Organische Chemie der Universität Regensburg durchgeführt.

4.10.4 Strukturuntersuchungen an HisF_N2

Das Absorptionsspektrum zur Konzentrationsbestimmung der in 4.10.3 präparierten Proteinlösung deutet auf hochmolekulare Aggregate hin (Abbildung 65A). Nach Korrektur der Streuung durch lineare Extrapolation des Bereiches von 350 bis 300 nm konnte eine Proteinkonzentration von ca. 0,18 mg/ml abgeschätzt werden.

Um Einblicke in die Sekundärstruktur von HisF_N2_S101C_F120W zu erhalten, wurde mit dieser Proteinlösung eine Circular Dichroismus spektroskopische Analyse im fern-UV Bereich (3.4.9) durchgeführt. Das auf molare Elliptizität pro Aminosäure (Θ_{MRW}) normierte Spektrum von HisF_N2_S101C_F120W deutet auf eine eher ungeordnete Sekundärstruktur hin, die vor allem von β -Strängen dominiert wird (Abbildung 65B).

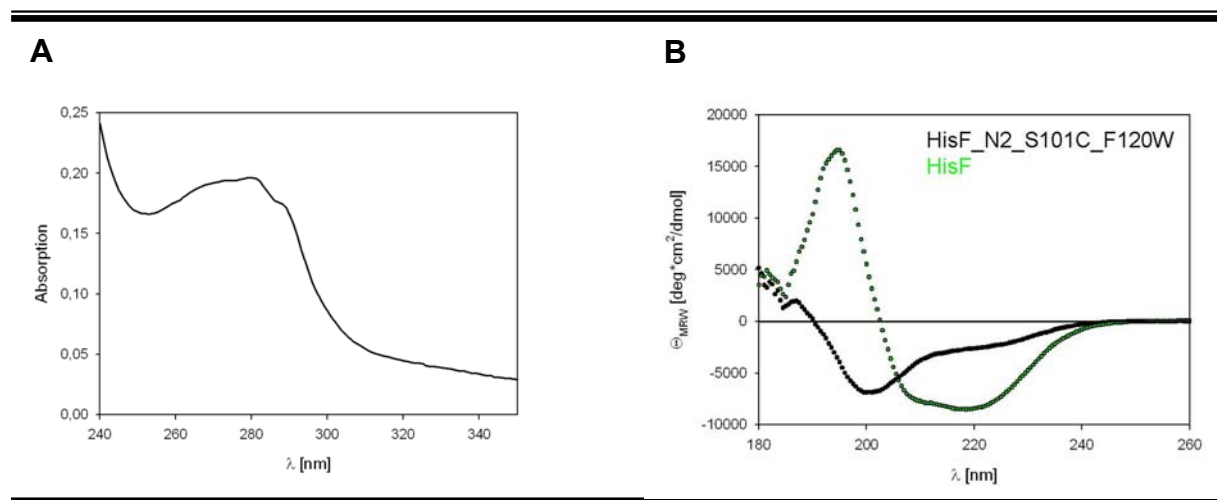


Abbildung 65: Absorptions- und fern-UV CD-Spektrum von chemisch synthetisiertem HisF_N2_S101C_F120W.

- (A) Das Absorptionsspektrum von HisF_N2 zeigt nach Lösen von 0,5 mg Protein in 50 mM KP, pH 7,5 und Dialyse gegen diesen Puffer ein starkes Meßsignal oberhalb von 300 nm, was ein eindeutiges Indiz für Aggregation ist. Die Konzentration wurde nach Korrektur der Streuung (lineare Extrapolation des Bereiches von 350 bis 300 nm) zu ca. 0,18 mg/ml abgeschätzt.
- (B) Das fern-UV CD-Spektrum von HisF_N2 (Lösung aus (A)) im Vergleich zu HisF. Das gezeigte Spektrum ist das Mittel aus fünf Einzelmessungen bei 25 °C. Die Schichtdicke der Küvette betrug 1 mm. Es sind kaum α -helikale Strukturen vorhanden, was die Abwesenheit der Bande bei 208 nm deutlich macht. Das Spektrum deutet auf viel ungeordnete Struktur mit einem geringen Anteil an β -Faltblatt hin.

Die aus der visuellen Inspektion des fern-UV CD-Spektrums gewonnenen Einblicke über die Sekundärstrukturelementverteilung wurden durch eine Analyse mit dem Programm ContinLL auf dem Dichroweb-Server (4.2.3.1) bestätigt (Tabelle 20). Demnach liegen über 60 % der Peptidkette in Lösung unstrukturiert vor. Das

Programm bestätigt den hohen Anteil an β -Strang-Konformation. Dieser ist doppelt so groß wie der aus der Struktur von HisF für das Viertel HisF_N2 abgeleitete Wert.

Tabelle 20: Vorhergesagte Sekundärstrukturelementanteile von HisF_N2_S101C_F120W (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

Sekundärstrukturelement	HisF_N2_S101C_F120W synthetisiert	
	Vorhersage aus CD	kalk. aus 1thf
α -Helix	11	46
β -Strang	25	12
<i>random coil</i>	64	42

Die Analyse der Tertiärstruktur erfolgte mittels Fluoreszenzemissionsspektroskopie (3.4.8). In Abbildung 66 ist das Fluoreszenzemissionsspektrum von HisF_N2_S101C_F120W gezeigt. Das Emissionsmaximum bei 355 nm weist auf ein vollständig lösungsmittel-exponiertes Tryptophan 120 hin. HisF_N2_S101C_F120W hat somit vermutlich in Lösung keine definierte Tertiärstruktur.

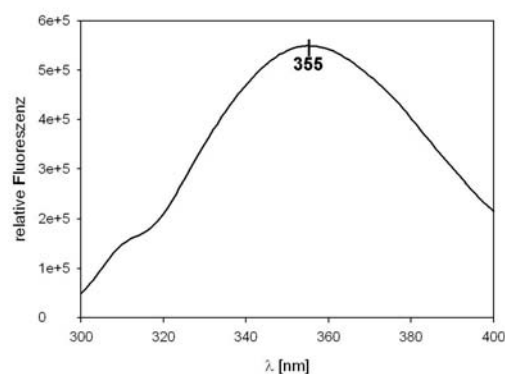


Abbildung 66: Fluoreszenzemissionsspektrum von HisF_N2_S101C_F120W unter nativen Bedingungen.

Das Fluoreszenzemissionsspektrum von HisF_N2_S101C_F120W unter nativen Bedingungen (50 mM KP, pH 7,5) nach Anregung bei 280 nm. Das Emissionsmaximum liegt bei 355 nm. Die Proteinkonzentration betrug 0,18 mg/ml. Das gezeigte Spektrum ist das Mittel aus zehn Einzelmessungen.

4.10.5 Festphasenpeptidsynthese von HisF_C2

HisF_C2 wurde ebenfalls in der Arbeitsgruppe von Frau Dr. Cabrele mit Hilfe der Fmoc-Methode synthetisiert (3.3.10). Auch hier wurden drei Aminosäureaustausche vorgenommen. Phenylalanin 210 wurde gegen Tryptophan, zur Einführung eines spektroskopischen Markers, Serin 201 gegen Cystein, zur Bildung einer intermolekularen Disulfidbrücke, und Alanin 220 gegen Glutamin, zur Ausbildung eines Wasserstoffbrückennetzwerks in *layer 4*, ausgetauscht (Abbildung 67A).

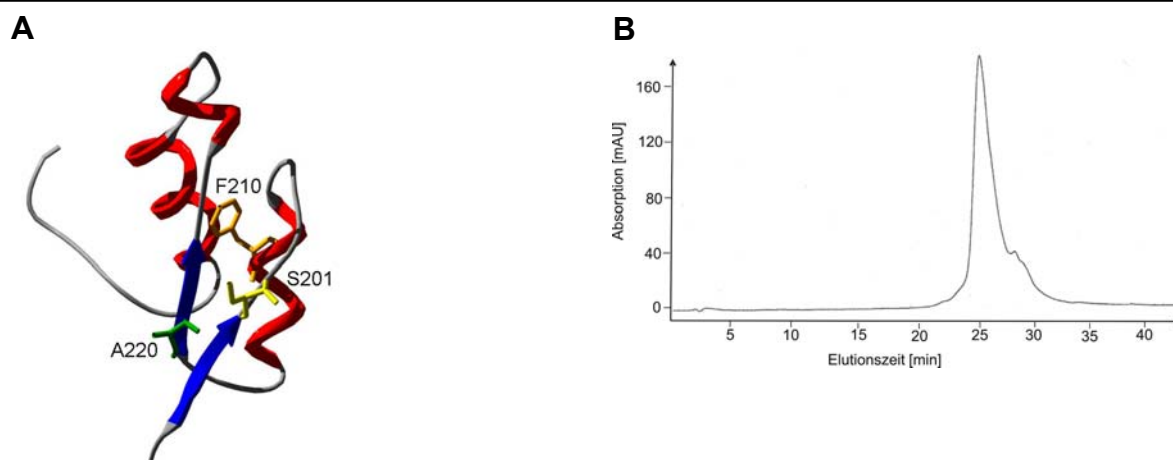


Abbildung 67: Bänderdiagramm von HisF_C2 (Strukturausschnitt aus 1thf) mit den als *sticks* dargestellten auszutauschenden Aminosäureresten (A). RP-HPLC-Elutionsprofil des synthetisierten Peptids (B).

- (A) Im Bänderdiagramm sind die β -Stränge in blau und die α -Helices in rot dargestellt. S201 (gelb) soll zur Ausbildung einer intermolekularen Disulfidbrücke gegen Cystein ausgetauscht werden. F210W (orange) dient der Einführung eines spektroskopischen Markers in HisF_C2. Glutamin soll anstelle von A220 (grün) in *layer 4* bei der Ausbildung eines Wasserstoffbrückennetzwerkes helfen (vgl. 4.5, Abbildung 38).
- (B) Elutionsprofil der RP-HPLC-Analyse des Endlängenfragmentes HisF_C2. Unter dem Peak bei einer Elutionszeit von ca. 25 min fand sich das Endprodukt. Die Ausbeute betrug 0,58 mg HisF_C2_S201C_F210W_A220Q. Die Analyse wurde von der Arbeitsgruppe Dr. Cabrele am Institut für Organische Chemie der Universität Regensburg durchgeführt.

Die Festphasenpeptidsynthese von HisF_C2 erfolgte ebenfalls vom C-Terminus (Leucin 253) aus, wobei in regelmäßigen Abständen Zwischenprodukte mittels RP-HPLC gereinigt und angereichert wurden.

Wie bei HisF_N2 setzt nach Anheften der letzten sechs Aminosäuren bis hin zu Prolin 197, die in diesem Fall β -Strang 7 von HisF entsprechen (vgl. Abbildung 23), Aggregation des Endproduktes ein. Diese ist allerdings nicht so stark wie bei HisF_N2 (Abbildung 66B). Ebenfalls konnten ca. 0,5 mg

HisF_C2_S201C_F210W_A220Q in lyophyllisierter Form dargestellt werden. Von diesem Lyophyllisat wurden jeweils 100 µg in 1 ml 50 mM KP, pH 7,5 gelöst, das entweder 5 % oder 60 % (v/v) TFE enthielt. Das Absorptionsspektrum zur Konzentrationsbestimmung deutet im Fall der 5 %igen TFE-Lösung auf starke Aggregation hin (Abbildung 68A). Im Fall der 60 %igen TFE-Lösung wird ein proteintypisches Absorptionsspektrum ohne nennenswerte Absorption oberhalb von 300 nm erhalten (Abbildung 68B). Es zeigt sich, dass das organische Lösungsmittel TFE, das sich an das Protein anlagert, in hohen Konzentrationen das synthetische HisF_C2_S201C_F210W_A220Q in Lösung halten kann.

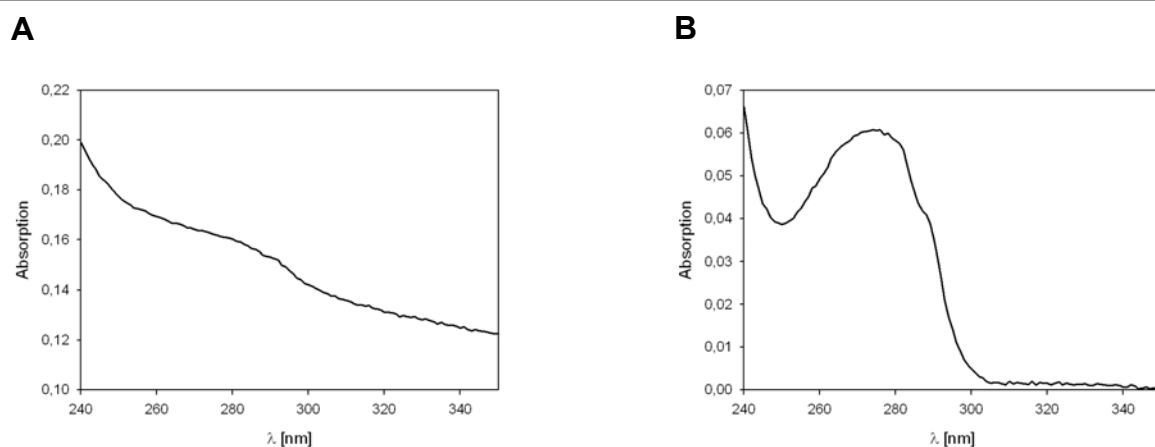


Abbildung 68: Absorptionsspektren von HisF_C2_S210C_F210W_A220Q in Anwesenheit von (A) 5 % (v/v) und (B) 60 % (v/v) TFE.

- (A) Das Absorptionsspektrum von 100 µg HisF_C2_S210C_F210W_A220Q, gelöst in 5% (v/v) TFE, 50 mM KP, pH 7,5, weist starke Lichtstreuung auf. Dies deutet auf Aggregation des Peptids hin.
- (B) Das Absorptionsspektrum von 100 µg HisF_C2_S210C_F210W_A220Q, gelöst in 60 % (v/v) TFE, 50 mM KP, pH 7,5, zeigt keine Absorption bei Wellenlängen > 300 nm. TFE lagert sich an das synthetisierte Peptid und hält es in Lösung.

4.10.6 Strukturuntersuchungen an HisF_C2_S210C_F210W_A220Q

Um Einblicke in die Sekundärstruktur von HisF_C2_S201C_F210W_A220Q zu erhalten, wurden mit den in 4.10.5 beschriebenen präparierten Proteinlösungen Circular Dichroismus spektroskopische Untersuchungen im fern-UV Bereich (3.4.9) durchgeführt. Durch unterschiedliche Mischungsverhältnisse von zwei Stammlösungen ohne bzw. mit 60 % (v/v) TFE konnte, analog zu den entsprechenden Messungen der Viertel HisF_N1 (vgl. Abbildung 44) und HisF_C1

(vgl. Abbildung 58), die Konzentration an TFE variiert werden, ohne die Proteinkonzentration zu verändern. Somit war es möglich die Veränderung der Sekundärstruktur in Abhängigkeit von der TFE-Konzentration zu verfolgen. Die auf molare Elliptizität pro Aminosäure (Θ_{MRW}) normierten Spektren (Abbildung 69) von HisF_C2_S201C_F210W_A220Q zeigen, dass TFE die Bildung von Sekundärstruktur induziert.

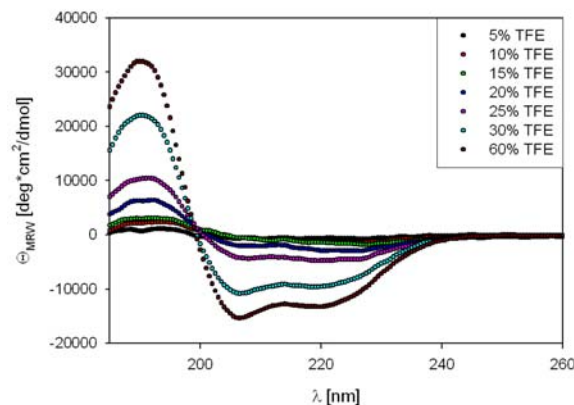


Abbildung 69: Titration von HisF_C2_S201C_F210W_A220Q mit TFE, verfolgt über CD-Spektroskopie im fern-UV Bereich.

Die fern-UV CD-Spektren wurden mit einer Proteinkonzentration von jeweils 0,1 mg/ml in 50 mM KP, pH 7,5 und steigenden TFE-Konzentrationen [%] (v/v) aufgenommen. Die Schichtdicke der Küvette betrug 1 mm. Die gezeigten Spektren sind das Mittel aus fünf Einzelmessungen bei 25 °C.

Die Auswertung der Spektren mit dem Programm ContinLL auf dem Dichroweb-Server zeigt, dass es sich dabei vor allem um α -Helices handelt (Tabelle 23).

Tabelle 21: Vorhergesagte Sekundärstrukturelementanteile von HisF_C2_S201C_F210W_A220Q in Abhängigkeit von der TFE-Konzentration. Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen in Prozent dar. Im Vergleich dazu sind die aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

<div>Konz. TFE (v/v) [%]</div> <div>Sekundär- strukturelement</div>	5	10	15	20	25	30	60	kalk. aus 1thf
	Vorhersage aus CD							
α -Helix	6	8	8	12	17	31	43	38
β -Strang	37	35	35	32	27	20	12	12
random coil	57	57	57	56	54	49	45	50

In Abwesenheit von TFE liegt der Großteil der Peptidkette der Variante in Lösung unstrukturiert vor, der Rest fast ausschließlich in Form von β -Strängen. Mit Erhöhung der TFE-Konzentration geht der Anteil der β -Stränge zurück und es bilden sich immer mehr α -Helices aus. Erst bei einer Konzentration von ca. 60 % TFE (v/v) wird die Sekundärstrukturelementverteilung angenommen, die man aus der Kristallstruktur von HisF für das Viertel berechnen konnte.

4.11 *Erstellung eines Modells zur Evolution der $(\beta\alpha)_8$ -barrel Faltung aus $(\beta\alpha)_2$ -Modulen*

Aufgrund der Beobachtung, dass im Viertel HisF_N1 ein konserviertes Cystein für die Stabilität und homogene Assoziation verantwortlich ist (4.4) und der Tatsache, dass das in HisF_C1 eingeführte Cystein zu verbesserten biophysikalischen Eigenschaften führt (4.8), wurde ein Evolutionsmodell zur Entstehung der $(\beta\alpha)_8$ -Faltung ausgehend von einem $(\beta\alpha)_2$ -Modul erstellt. Im strukturbasierten Sequenzalignment der vier HisF-Viertel (vgl. Abbildung 23) kommen am C-terminalen Ende des jeweils ersten β -Stranges eines $(\beta\alpha)_2$ -Moduls konservierte Cysteine bzw. im Fall von HisF_C2 ein Serin zu liegen, das jedoch durch nur einen Nukleotidaustausch aus Cystein entstehen kann.

Entsprechend geht das in Abbildung 70 skizzierte Modell von einem $(\beta\alpha)_2$ -Modul aus, das mithilfe einer Disulfidbrücke mit einem zweiten Monomer zu einem Homodimer assoziieren kann und anschließend über nicht-kovalente Wechselwirkungen mit einem weiteren Homodimer ein Tetramer bildet (1). Durch Duplikation des Gens, das für das ursprüngliche $(\beta\alpha)_2$ -Modul kodiert, kann nach Fusion der Tochtergene eine $(\beta\alpha)_4$ -Hälfte aus zwei identischen $(\beta\alpha)_2$ -Modulen exprimiert werden. Diese Hälfte kann über maximal zwei intermolekulare Disulfidbrücken mit einem zweiten $(\beta\alpha)_4$ -Modul zu einem Dimer verknüpft werden (2). Da, wie am Beispiel von HisF_N gezeigt, bereits eine Disulfidbrücke zur Dimerisierung von zwei Hälften genügt, könnte auf dieser Stufe bereits die Mutation eines der vier Cysteine erfolgen ohne die Struktur zu pertubieren. Durch Duplikation des für das $(\beta\alpha)_4$ -Modul kodierenden Gens und anschließende Fusion wäre dann das $(\beta\alpha)_8$ -barrel entstanden (3). Hier ist eine Verknüpfung der Fragmente über eine Disulfidbrücke nicht mehr notwendig,

sodass daran beteiligte Cysteinreste mutieren können. Dies wird im Fall von HisF aus *Thermotoga maritima* deutlich, bei welchem nun nur noch ein vollständig konserviertes Cystein zu finden ist, nämlich C9 in HisF_N1 (vgl. Abbildung 23).

Um dieses Modell weiter zu untermauern, wurden $(\beta\alpha)_4$ -Halbbarrels, bestehend aus zwei identischen $(\beta\alpha)_2$ -Modulen, hergestellt und untersucht.

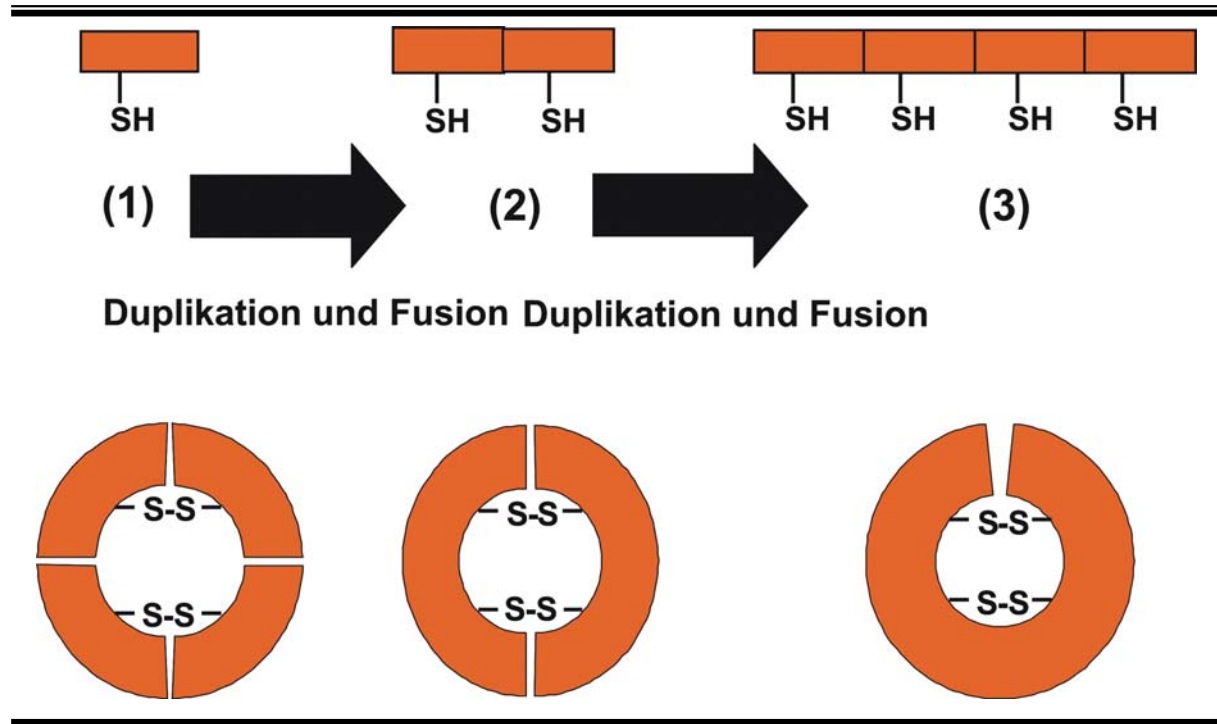


Abbildung 70: Evolutionsmodell zur Entstehung der $(\beta\alpha)_8$ -barrel Faltung aus $(\beta\alpha)_2$ -Modulen.

Im oberen Teil der Abbildung sind schematisch die Evolutionseignisse auf DNA-Ebene, im unteren Teil die resultierenden Genprodukte dargestellt. Man geht von einem Gen aus, das für ein $(\beta\alpha)_2$ -Modul kodiert. Dieses $(\beta\alpha)_2$ -Modul verknüpft sich über eine Disulfidbrücke mit einem zweiten Monomer zu einem Homodimer. Zwei Homodimere lagern sich dann zum Tetramer zusammen (1). Durch Duplikation und Fusion des Ursprungs-Gens entsteht eine $(\beta\alpha)_4$ -Hälfte, die ebenfalls über maximal zwei Disulfidbrücken zum Dimer verknüpft wird (2). Durch erneute Duplikation und Fusion des duplizierten Gens entsteht das heutige $(\beta\alpha)_8$ -barrel (3). Auf dieser Stufe sind die Disulfidbrücken nicht mehr notwendig, sodass die daran beteiligten Cysteine ihren konservierten Status ablegen können.

4.12 Die duplizierten und fusionierten Viertel HisF_N1N1 und HisF_C1C1

4.12.1 Klonierung, Expression und Anreicherung von HisF_N1N1

Mit *hisF_N1_R5A_E46Q* (4.5) als Matrize wurden mit den Primern T7-Promotor, *hisF_N1-1_3'_NheI* und *5'_N1_R5A_NheI*, *hisF_N1_3'_BamHI* zwei *hisF_N1*-Gene

mittels PCR (3.2.3) mit unterschiedlichen Restriktionsschnittstellen amplifiziert. Das erste Fragment erhielt am 5'-Ende eine *NdeI*- und am 3'-Ende eine *NheI*-Schnittstelle, das zweite Fragment 5'-seitig eine *NheI*- und 3'-seitig eine *BamHI*-Schnittstelle. Zunächst wurden beide Fragmente nach Verdau mit *NheI* ligiert. Das resultierende ca. 440 bp lange Fragment wurde mit *NdeI* und *BamHI* verdaut und in den Vektor pET21a (2.7.1.1) kloniert. Die beabsichtigte Nukleotidabfolge wurde mittels DNA-Sequenzierung verifiziert (3.2.10). Dieses Protein trägt in beiden HisF_N1-Vierteln die Austausche R5A, F23W und E46Q, wird im Folgenden aber nur noch als HisF_N1N1 bezeichnet.

Eine analytische Expression (3.3.1) von *hisF_N1N1* in BL21(DE3)-Zellen zeigte, dass das rekombinante Protein nach Induktion mit 1 mM IPTG zwar zum Großteil im Pellet zu finden ist, der lösliche Anteil des Proteins im Vergleich zu dem Viertel HisF_N1_R5A_E46Q jedoch zugenommen hat (vgl. Abbildung 39). Die Reinigung der ($\beta\alpha$)₄-Einheit erfolgte wie bereits für die anderen Fragmente beschrieben über Rückfaltung mit anschließender Gelfiltration (4.4; 4.5; 4.7; 4.8). Das Reinigungsergebnis und das Absorptionsspektrum von HisF_N1N1 sind in Abbildung 71 gezeigt. Die Ausbeute aus 1 l Kulturmedium betrug 42,5 mg HisF_N1N1.

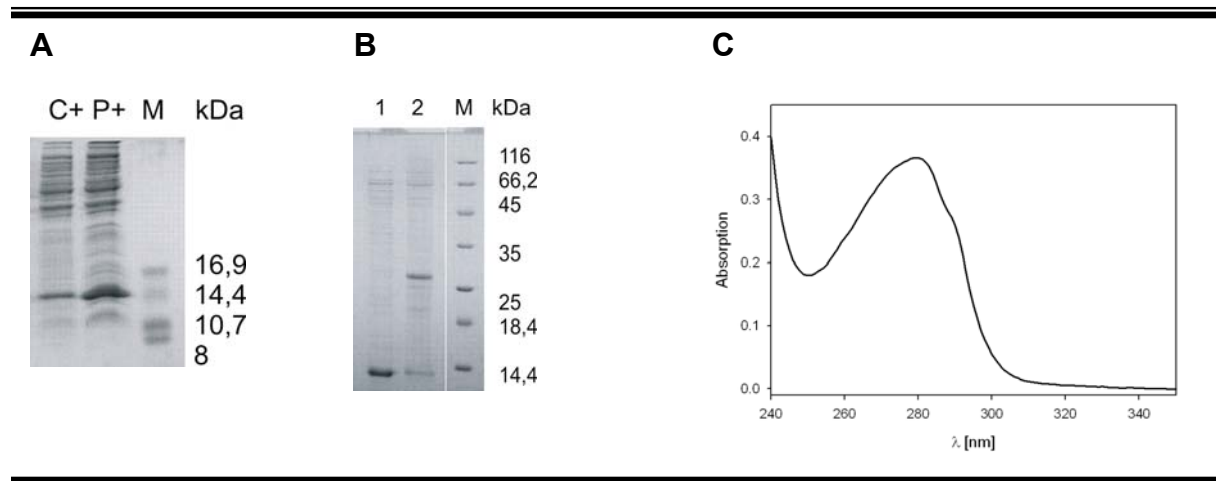


Abbildung 71: Probeexpression und Reinigung von HisF_N1N1, verfolgt mittels SDS-PAGE.

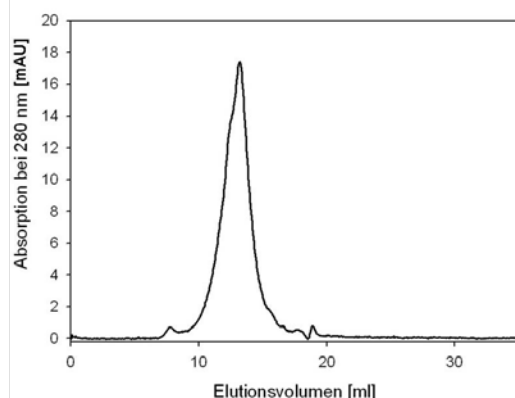
- (A) Probeexpression von HisF_N1N1 in BL21(DE3)-Zellen, verfolgt mittels SDS-PAGE auf einem Tris-Tricin-Gel (20% Polyacrylamid). Der Großteil des Proteins wird nach Induktion mit IPTG (1mM) im unlöslichen Teil exprimiert (P+). Es finden sich aber mehr als 10 % des Proteins im löslichen Teil des Zelllysates (C+). (C+: lösliche Zellfraktion nach Induktion; P+: unlösliche Zellfraktion nach Induktion; M: *small fragment marker*)
- (B) Reinigung von HisF_N1N1 mittels Rückfaltung und sich anschließender präparativer Gelfiltration, verfolgt mittels SDS-PAGE auf einem Tris-Glycin-Gel (15% Polyacrylamid). Unter reduzierenden Bedingungen ist nur das Monomer bei ca. 14 kDa (Spur 1) zu sehen, unter oxidierenden Bedingungen überwiegt das Dimer bei ca. 30 kDa (Spur 2). (1: gereinigtes Protein mit Mercaptoethanol versetzt; 2: gereinigtes Protein ohne Mercaptoethanol; M: *low molecular weight marker*)
- (C) Absorptionsspektrum von HisF_N1N1 in 50 mM KP, pH 7,5 bei RT.
HisF_N1N1: c = 1,7 mg/ml; $A_{280/250}$: 2,03

Durch die vorhandenen Cysteine kann sich prinzipiell sowohl eine inter- als auch intramolekulare Disulfidbrücke bilden. Wie man in Abbildung 71B erkennen kann, liegt ein Großteil des Proteins als Homodimer vor, dessen Untereinheiten über die intermolekulare Disulfidbrücke verbrückt sind (vgl. (2) in Abbildung 70). Das Absorptionsspektrum von HisF_N1N1 zeigt kein wesentliches Meßsignal oberhalb von 300 nm, was ein Indiz für gute Löslichkeit und wenig Aggregation ist.

4.12.2 Bestimmung der Homogenität und des Assoziationszustandes von HisF_N1N1

4.12.2.1 Analytische Gelfiltration

Die Homogenität der Proteinlösung wurde durch analytische Gelfiltration mit einer Superdex 75-Säule untersucht (3.4.5). Es wurden 200 µl HisF_N1N1 (4.12.1) auf die Säule aufgetragen und mit 50 mM KP, 300 mM KCl, pH 7,5 eluiert. Im Elutionsprofil (Abbildung 72) ist nur ein symmetrischer Peak zu erkennen, der einem Molekulargewicht von 28,4 kDa entspricht. Bei einem berechneten Molekulargewicht von 16,1 kDa für das Monomer handelt es sich hier um das Dimer. Es sind keine höhermolekularen Aggregate vorhanden.



	MW _{app} [kDa]	MW _{kalk} [kDa]	Assoziations- zustand
HisF_N1N1	28,4	16,1	Dimer

Abbildung 72: Analytische Gelfiltration von HisF_N1N1 (Superdex 75-Säule).

Die Elution von HisF_N1N1 wurde bei 280 nm verfolgt. Es wurden 200 µl HisF_N1N1 ($c = 0,14$ mg/ml) aufgetragen. Die Elution erfolgte bei 24 °C mit einer Flussrate von 0,5 ml/min in 50 mM KP, 300 mM KCl, pH 7,5. Rechts ist das mit Hilfe der Kalibrierungsgeraden bestimmte apparente Molekulargewicht, das berechnete Molekulargewicht für das Monomer und der abgeleitete Assoziationszustand des Proteins angegeben.

4.12.2.2 Analytische Ultrazentrifugation

Um das Ergebnis der analytischen Gelfiltration von HisF_N1N1 zu verifizieren wurde wie für HisF_N1 beschrieben (4.4.2.3) über einen Gleichgewichtslauf in der analytischen Ultrazentrifuge das Molekulargewicht bestimmt (3.4.6.2). Die Auswertung des Gleichgewichtslaufes ist in Abbildung 73 dargestellt.

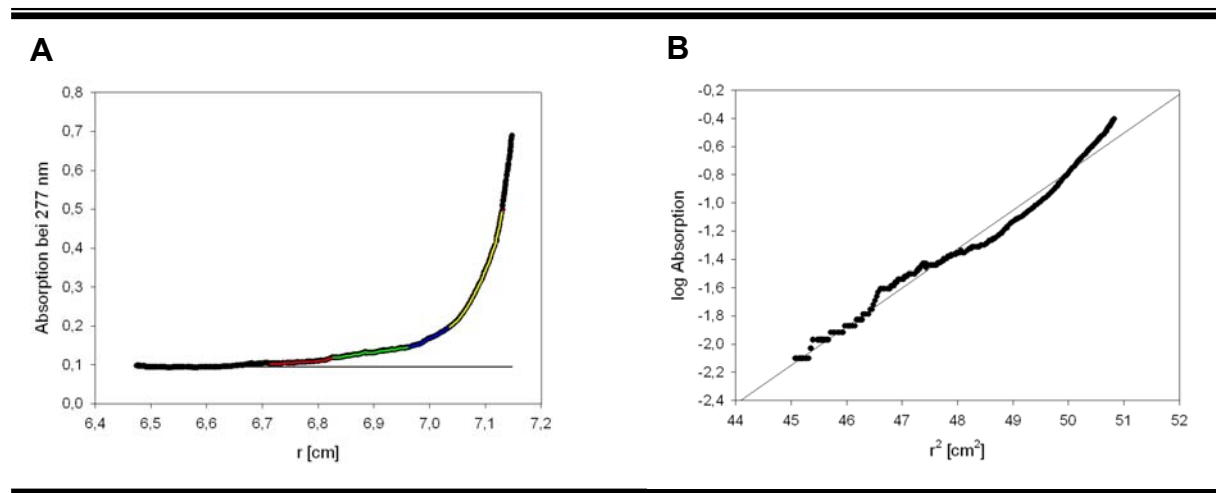


Abbildung 73: Auswertung der analytischen Ultrazentrifugation von HisF_N1N1.

- (A) Absorption in der Messzelle im Gleichgewichtslauf von HisF_N1N1 (0,2 mg/ml in 50 mM KP, pH 7,5) bei 24 °C, verfolgt bei 277 nm nach 48 h. Die durchgezogene Linie stellt die Baseline dar, um die alle Absorptionswerte korrigiert wurden. Der rot markierte Bereich (Gesamtbereich) diente zur Erstellung der Geraden in (B). Es wurde für drei Teilbereiche das zugehörige Molekulargewicht bestimmt: grün: 28,7 kDa; blau: 50,8 kDa; gelb: 73,4 kDa.
- (B) Halblogarithmische Auftragung des Gesamtbereichs in (A). Aus der Geradensteigung wurde mit den Formeln aus 3.4.6.2 ein Molekulargewicht von 45,7 kDa bestimmt. Dieses entspricht dem Molekulargewichtsmittel mehrere Spezies.

Nach halblogarithmischer Auftragung der Absorption in der Messzelle (roter Bereich in Abbildung 73A = Gesamtbereich; $r = 6,73$ bis $7,15$ cm) erkennt man, dass die Messpunkte annäherungsweise auf einer Geraden zu liegen kommen (Abbildung 73B), aus deren Steigung ein Molekulargewicht von 45,7 kDa berechnet wurde. Es wurden drei Teilbereiche mit unterschiedlicher Steigung und somit verschiedenen Molekulargewichten detektiert. Zum Einen wurde der in Abbildung 73A grün markiert Bereich ($r = 6,83$ bis $6,98$ cm), zum Anderen der blau markierte Bereich ($r = 6,98$ bis $7,05$ cm) und schließlich der gelb markierte Bereich ($r = 7,05$ bis $7,15$ cm) ausgewertet. Die berechneten Molekulargewichte mit 28,7 (grün), 50,8 (blau) und 73,4 (gelb) kDa liegen unterhalb (grün) bzw. oberhalb (blau und gelb) des Molekulargewichts des Dimers (32,2 kDa). Im grün markierten Bereich der Messzelle

erhält man das Molekulargewichtsmittel, da die Proteinlösung nicht monodispers ist. Da das berechnete Molekulargewicht in diesem Bereich nahe am Molekulargewicht des Dimers liegt, befindet sich in der Proteinlösung überwiegend die dimere Form von HisF_N1N1 (persönliche Mitteilung Dr. H. Durchschlag). Im blau markierten Bereich und in Bodennähe (gelb markierter Bereich) sind die Molekulargewichte der Spezies um 18 bis 40 kDa höher als für das Dimer erwartet, was auf einen beträchtlichen Anteil Aggregate von HisF_N1N1 bei höheren Proteinkonzentrationen hindeutet.

4.12.3 Spektroskopische Charakterisierung von HisF_N1N1

4.12.3.1 Fern-UV CD-Spektroskopie

Die Sekundärstruktur des Proteins wurde mit Hilfe von CD-Spektroskopie im fern-UV Bereich analysiert (3.4.9). Das nach Normierung auf die molare Elliptizität pro Aminosäure (Θ_{MRW}) erhaltene Spektrum ist in Abbildung 74 zusammen mit den Spektren von HisF_N1 und HisF gezeigt.

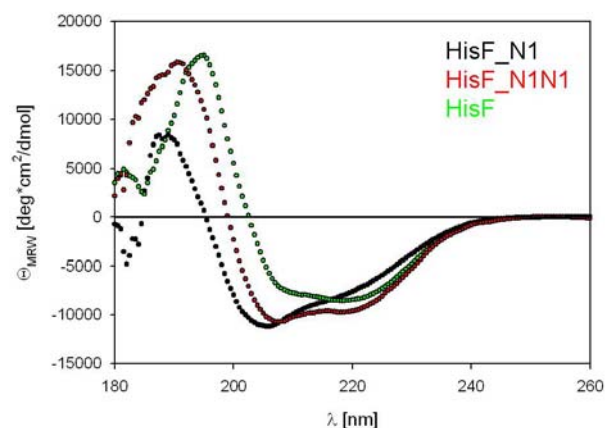


Abbildung 74: Fern-UV CD-Spektrum von HisF_N1N1 im Vergleich zu HisF_N1 und HisF.

Das Spektrum von HisF_N1N1 wurde in einer 1 mm Küvette mit 0,14 mg/ml Protein in 50 mM KP, pH 7,5 aufgenommen. Das gezeigte Spektrum ist das Mittel aus fünf Einzelmessungen bei 25 °C. Die Daten von HisF_N1 und HisF, die unter identischen Bedingungen aufgenommen wurden, stammen aus den Abbildungen 26 und 34.

Die Auswertung der Sekundärstrukturelementverteilung erfolgte wiederum auf dem Dichroweb-Server mit dem Programm ContinLL (4.2.3.1). Die gemittelten Ergebnisse der Strukturvorhersage sind in Tabelle 22 zusammengefasst.

Tabelle 22: Vorhergesagte Sekundärstrukturelementanteile von HisF_N1N1 (Programm ContinLL, Dichroweb-Server). Die angegebenen Werte stellen den Mittelwert der vorhergesagten Strukturanteile auf der Basis von drei unabhängigen experimentellen Datensätzen des Dichroweb-Servers in Prozent dar. Im Vergleich dazu sind die von HisF_N1 bestimmten (vgl. Tabelle 9) und aus der Kristallstruktur (1thf) berechneten Sekundärstrukturelementanteile gezeigt.

	HisF_N1N1	HisF_N1	HisF_N1
Sekundärstruktur- element	Vorhersage aus CD		kalk. aus 1thf
α -Helix	29	19	31
β -Strang	17	24	22
<i>random coil</i>	54	57	47

HisF_N1N1 ist durch die gegenseitige Stabilisierung der kovalent verknüpften Einzelviertel in der Lage, der für HisF_N1 erwarteten Sekundärstrukturelementverteilung näher zu kommen als das isolierte HisF_N1 Viertel. Die Duplikation und Fusion von HisF_N1 hat somit einen positiven Effekt auf die Sekundärstruktur des resultierenden $(\beta\alpha)_8$ -barrels.

4.12.3.2 Fluoreszenzemissionsspektroskopie

Die Tertiärstruktur von HisF_N1N1 wurde über die Aufnahme von Fluoreszenzemissionsspektren untersucht (3.4.8). Die Messungen zeigen, dass HisF_N1N1 unter nativen Pufferbedingungen ein Emissionsmaximum von 341 nm besitzt. Dieses ist um 4 nm niedriger im Vergleich zu HisF_N1 (vgl. Abbildung 36A). D.h. das Tryptophan 23 ist in den fusionierten Vierteln etwas mehr vom Lösungsmittel abgeschirmt. Nach Denaturierung in 8 M Harnstoff kommt es wie bei HisF_N1 zu einer Rotverschiebung des Maximums um 12 nm (Abbildung 75).

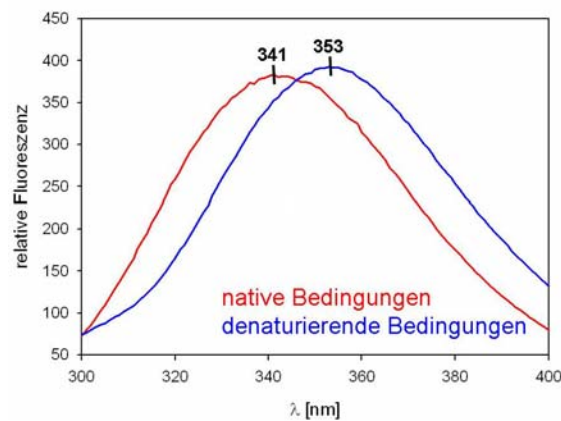


Abbildung 75: Fluoreszenzmissionsspektren von HisF_N1N1 im nativen und denaturierten Zustand.

Unter nativen Pufferbedingungen (50 mM KP, pH 7,5) besitzt HisF_N1N1 ein Emissionsmaximum von 341 nm. Unter denaturierenden Pufferbedingungen (8 M Harnstoff, 50 mM KP, pH 7,5) kommt es zu einer Rotverschiebung des Maximums um 12 nm auf 353 nm.

Die Proteinkonzentration betrug 5 μ M, die Schichtdicke der Küvette 1 cm. Die Emission wurde von 300 – 400 nm nach Anregung bei 280 nm aufgenommen. Die gezeigten Spektren sind das Mittel aus zehn Einzelmessungen bei 25 °C.

4.12.4 Stabilitätsuntersuchung von HisF_N1N1

Um Einblicke in die Stabilität des Proteins zu erhalten, wurde die Entfaltung von HisF_N1N1 am CD- und Fluoreszenzspektrometer verfolgt (3.4.10). Die Ergebnisse sind in Abbildung 76 gezeigt.

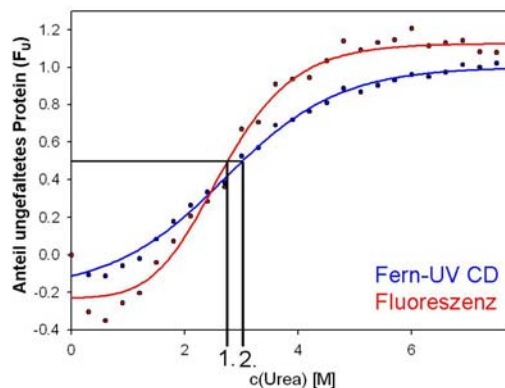


Abbildung 76: Chemische Entfaltung der Sekundär- und Tertiärstruktur von HisF_N1N1.

Die Auffaltung der Sekundärstruktur verfolgt am CD-Spektrometer (220 nm; Schichtdicke der Küvette 1 mm) zeigt eine niedrige Kooperativität. Die Auffaltung der Tertiärstruktur verfolgt am Fluoreszenzspektrometer (Anregung 280 nm; Emission 320 nm; Schichtdicke der Küvette 1 cm) zeigt eine höhere Kooperativität. Die $D_{1/2}$ -Werte (durch Hilfslinien eingezeichnet) liegen bei Konzentrationen von 2,8 M (1.) bzw. 2,9 M (2.) Harnstoff. Die Proteinkonzentration betrug jeweils 0,2 mg/ml in 50 mM KP, pH 7,5. Die sigmoidalen Ausgleichslinien sollen das Auge des Betrachters leiten.

Die Sekundärstruktur entfaltet ähnlich wie HisF_N1 eher unkooperativ (Abbildung 76) und der $D_{1/2}$ -Wert liegt mit 2,9 M Harnstoff nur um 0,2 M Harnstoff höher als beim isolierten Viertel (vgl. Abbildung 37).

Die Tertiärstrukturentfaltung, die bei den Varianten von HisF_N1 nicht messbar war, zeigt im Fall von HisF_N1N1 Ansätze von Kooperativität (Abbildung 76). Man kann also schlussfolgern, dass die Verdoppelung und kovalente Verknüpfung von HisF_N1 einen großen Einfluss auf die Stabilität der Tertiärstruktur hat.

4.12.5 Resümee zur Charakterisierung von HisF_N1N1

Durch die Fusion von zwei *hisF_N1*-Genen erhöht sich der Anteil an löslich exprimiertem Protein. Dies ist ein erster Hinweis auf die Verbesserung des HisF_N1N1-Konstruktes im Vergleich zu HisF_N1 und seinen Varianten. Die Homogenität der Proteinlösung nimmt ebenfalls zu und bei der analytischen Gelfiltration findet man nur eine Spezies, das Dimer. Dieser Befund konnte durch analytische Ultrazentrifugation gestützt werden. Durch die Verknüpfung zweier HisF_N1-Viertel kommt es weiterhin zu einer wechselseitigen Unterstützung in der Faltung und es wird nahezu die Sekundärstrukturelementverteilung erreicht, die man aus der Kristallstruktur von HisF für das Viertel HisF_N1 berechnet hat. Ein Stabilitätsgewinn findet sich nur in Hinblick auf die Tertiärstruktur, deren Auffaltung auch einen Ansatz von Kooperativität zeigt.

4.12.6 Klonierung und Reinigung von HisF_C1C1

In einem zweiten Ansatz zur Überprüfung des aufgestellten Evolutionsmodells wurde das Gen des *hisF_C1*-Viertels verdoppelt und die Tochtergene fusioniert. Zur Klonierung von *hisF_C1C1* wurden zwei Fragmente mit den Primern 5'C1_A128C_NdeI, 3'C1a_NheI und 5'C1b_NheI, *hisF_C1_3'BamHI* in einer PCR (3.2.3) amplifiziert. Als Matrize diente *hisF_C1_A128C_Y143H_E167Q*. Zunächst wurden beide Fragmente mit *NheI* verdaut und ligiert. Das ca. 440 bp lange Produkt wurde mit *NdeI* und *BamHI* verdaut und in den Vektor pET28a (2.7.1.3) kloniert.

Dadurch trägt das rekombinante Protein nach Expression N-terminal einen His₆-Tag, der mit Hilfe der Protease Thrombin abgetrennt werden kann. Die beabsichtigte Nukleotidabfolge wurde mittels DNA-Sequenzierung verifiziert (3.2.10). Die Expression des *hisF_C1C1* Gens erfolgte in *E. coli* Rosetta Zellen üN bei 37 °C. Das Protein wurde selbst nach Induktion mit IPTG (1mM) nur in geringen Mengen produziert und musste mittels *Western Blot* (3.4.3) mit dem Anti-His₆-Antikörper nachgewiesen werden. HisF_C1C1 fand sich vollständig im unlöslichen Teil des Zelllysates und wurde deshalb mittels Rückfaltung auf einer HisTrap™-Säule gereinigt (3.3.5). Elutionsfraktionen mit sauberem Protein wurden mittels SDS-PAGE (3.4.1.1) identifiziert (Abbildung 77), vereinigt und üN bei 4 °C gegen 50 mM KP, pH 7,5 dialysiert. Während der Dialyse kam es zur Trübung der Proteinlösung. Zusammen mit dem Ergebnis des Absorptionsspektrums, das auf erhebliche Streuung hindeutet (nicht gezeigt), lässt dies auf Aggregation des gereinigten Proteins schließen. Nach Zentrifugation der Lösung konnte sowohl über die Messung der Absorption als auch über SDS-PAGE (3.4.1.1) kein Protein mehr nachgewiesen werden. Die Duplikation des stabilsten HisF_C1-Viertels HisF_C1_A128C_Y143H_E167Q (4.8) hatte keine Verbesserung der biophysikalischen Eigenschaften zur Folge (Daten nicht gezeigt).

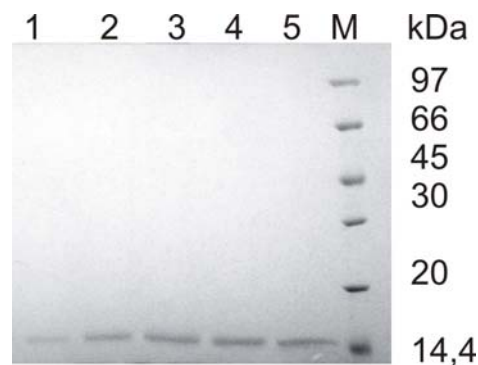


Abbildung 77: Analyse der Reinigung von HisF_C1C1 mittels Rückfaltung auf einer HisTrap™-Säule, verfolgt über SDS-PAGE auf einem Tris-Glycin-Gel (15 % Polyacrylamid).

HisF_C1C1 kann in reiner Form nach Rückfaltung aus 6 M GdmCl dargestellt werden. Nach Vereinigung und Dialyse der Spuren 1 - 5 fällt das Protein jedoch aus (Daten nicht gezeigt). (Spur 1 – 5: relevante Fraktionen nach Elution des Proteins von der HisTrap™-Säule; M: *low molecular weight marker*)

5 Diskussion

Proteine sind aus strukturell abgeschlossenen Domänen mit eigenständiger Faltung und Funktion aufgebaut. Diese Domänen stellen die evolutiven Einheiten von Proteinfaltungstypen dar (Jaenicke, 1999; Vogel *et al.*, 2004; Orengo & Thornton, 2005). Duplikation von Domänen mit sich anschließender Rekombination und Fusion dieser Einheiten sind die zentralen Mechanismen, die die Proteinevolution und Komplexität der Proteome vorantreiben (Vogel *et al.*, 2005). Deshalb ist die Identifizierung einzelner Domänen innerhalb komplexer Strukturen ein wichtiger Schritt, um die Entstehung der heute existierenden Proteinstrukturen verstehen zu können. Weiterhin können innerhalb der definierten Domänen kleine, eigenständige Subdomänen mithilfe von Sequenz- und Strukturanalysen identifiziert werden.

Thoma *et al.* (1999) und Lang *et al.* (2000) postulierten auf der Basis der Ergebnisse von Sequenz- und Strukturanalysen der beiden $(\beta\alpha)_8$ -barrel Enzyme HisA und HisF aus *Thermotoga maritima* (*tm*), dass die auf den ersten Blick „Eindomänen-Faltung“ der beiden Proteine tatsächlich aus zwei Modulen besteht, wobei das eine der N- und das andere der C-terminalen Hälfte entspricht. Den biochemischen Beweis für die Existenz zweier unabhängiger Domänen innerhalb des $(\beta\alpha)_8$ -barrels lieferte der Befund, dass sich die isolierten Hälften des $(\beta\alpha)_8$ -barrel Enzyms HisF aus *T. maritima*, HisF_N und HisF_C, zu stabilen Proteinen falten können (Höcker *et al.*, 2001). Diese Ergebnisse stützen das Modell der Evolution von HisA, HisF und anderer $(\beta\alpha)_8$ -barrel Enzyme aus einem gemeinsamen Vorläufer halber Länge (Abbildung 7).

Im Rahmen dieser Arbeit wurden die beiden Hälften, HisF_N und HisF_C aus *tm*HisF, mit Hilfe eines im Vergleich zu Höcker *et al.* (2001) optimierten Protokolls aufgereinigt und charakterisiert. Sowohl HisF_N als auch HisF_C liegen nach Rückfaltung aus *Inclusion bodies* und sich anschließender präparativer Gelfiltration als Dimere in Lösung vor (4.2.2). Eine im Rahmen dieser Arbeit gewonnene Erkenntnis ist, dass es unter oxidierenden Bedingungen zwischen zwei HisF_N-Monomeren über die Cysteinseitenketten an Position 9 zur Ausbildung einer intermolekularen Disulfidbrücke kommt (Abbildung 24A). Höcker *et al.* (2001) fanden neben Dimeren auch höhere Assoziationszustände (Tetramere, Dekamere) von HisF_N und HisF_C. Durch die sich an die Rückfaltung anschließende Gelfiltration

der Proteine, die von Höcker *et al.* (2001) nicht durchgeführt wurde, konnten diese höheren Oligomerisierungszustände entfernt werden.

Anhand von CD-spektroskopischen Messungen im fern-UV Bereich konnte für beide Proteine eine native Sekundärstruktur nachgewiesen werden (Abbildung 26). Mithilfe einer Analyse der CD-Spektren durch das Sekundärstruktur-Vorhersageprogramm ContinLL wurde für HisF_N ein α -helikaler Anteil von 34 % bestimmt. Dieser ist identisch mit dem aus der Kristallstruktur von *tmHisF* (PDB-Eintrag: 1thf) berechneten Wert. Da der Anteil an β -Strängen im Vergleich zur Kristallstruktur erniedrigt ist, muss davon ausgegangen werden, dass ein Teil des β -Faltblattes von HisF_N als *random coil* vorliegt. HisF_C weist ein deutlich erhöhtes Verhältnis von α -Helices zu β -Strängen im Vergleich zum berechneten Verhältnis aus *tmHisF* auf. Der unstrukturierte Anteil des Proteins hingegen entspricht dem berechneten Wert. Es ist deshalb zu vermuten, dass es in HisF_C vor allem zu Umlagerungen von β -Strängen zu α -Helices kommt (Tabelle 8).

Die Ergebnisse der Fluoreszenzemissionsspektroskopie deuten auf eine ausgeprägte Tertiärstruktur innerhalb von (HisF_N)₂ und (HisF_C)₂ hin (Abbildung 27). Das künstlich eingeführte Tryptophan an Position 23 (*loop* β 1/ α 1) in HisF_N ist dabei nur geringfügig vom Lösungsmittel abgeschirmt, wie das Emissionsmaximum bei 341 nm zeigt. Durch die Dimerisierung von zwei Hälften kommt es dennoch zu einer teilweisen Abschirmung des Restes vom polaren Lösungsmittel, da andernfalls ein Emissionsmaximum bei einer Wellenlänge > 350 nm zu erwarten wäre. Das einzige wildtypische Tryptophan 156 in *tmHisF* (α -Helix 5) ist im Gesamtprotein gut vom Lösungsmittel abgeschirmt, wie durch das Emissionsmaximum bei 323 nm deutlich wird (Höcker *et al.*, 2001). Dieser Rest zeigt in (HisF_C)₂ ein nur leicht erhöhtes Emissionsmaximum bei 336 nm und ist damit fast ebenso gut vom Lösungsmittel abgeschirmt wie im wildtypischen Protein.

Die Stabilität beider Hälften ist bezogen auf die Entfaltung ihrer Sekundärstrukturelemente annähernd identisch und verläuft mit einer geringen Kooperativität (Abbildung 28A). Die Entfaltung der Tertiärstruktur von HisF_C zeigt dagegen eine etwas erhöhte Kooperativität (Abbildung 28B), was auf ein kompakteres Protein schließen lässt. Diese Ergebnisse, die sich im Wesentlichen mit denen von Höcker *et al.* (2001) decken, gaben den Anstoß, Fusionsproteine aus zwei identischen HisF_C Einheiten bzw. eine HisAF Chimäre herzustellen und zu optimieren (Höcker *et al.*, 2004; Seitz *et al.*, 2007; J. Claren, unveröffentlichte Daten).

Die Ergebnisse, insbesondere die Tendenz der beiden Hälften zu vollständigen $(\beta\alpha)_8$ -Einheiten zu assoziieren, stützen die Annahme, dass es sich bei den $(\beta\alpha)_4$ -Modulen um eigenständige Domänen handelt.

Im Hauptteil der vorliegenden Arbeit sollten biochemische Evidenzen dafür erbracht werden, dass die beiden Proteinhälften HisF_N und HisF_C wiederum aus jeweils zwei kleineren Subdomänen zusammengesetzt sind, die jeweils einem $(\beta\alpha)_2$ -Modul entsprechen. Auf diese Weise wäre es möglich, das bisherige „Zweistufen“-Modell der Evolution der $(\beta\alpha)_8$ -barrel Faltung (Abbildung 7) um eine evolutive Stufe zu erweitern. Im resultierenden „Dreistufen“-Modell geht man von einem Gen aus, das für ein $(\beta\alpha)_2$ -Modul kodiert. Das DNA-Fragment bildet durch Duplikation und anschließende Fusion die genetische Grundlage für die Expression eines $(\beta\alpha)_4$ -Moduls, das der Hälfte HisF_N bzw. HisF_C entspricht. Nach erneuter Duplikation und Fusion der Gene ist die Grundlage für die heute existierende $(\beta\alpha)_8$ -barrel Faltung geschaffen.

Erste Hinweise auf den vierfachen modularen Aufbau der $(\beta\alpha)_8$ -barrel Faltung (insbesondere von HisF aus *Thermotoga maritima*) ergaben sich aus der Topologie des zentralen β -Fasses (Abbildung 8) sowie auf der Basis von Sequenzanalysen (Abbildung 9 und 10). *tmHisF* wurde basierend auf diesen Hinweisen in die $(\beta\alpha)_2$ -Module HisF_N1, HisF_N2, HisF_C1 und HisF_C2 unterteilt. Ein Sequenzvergleich zeigte, dass HisF_N1 mit durchschnittlich 20 % die höchste Ähnlichkeit zu den drei anderen Vierteln besitzt und es sich bei ihm vermutlich um das ursprünglichste $(\beta\alpha)_2$ -Modul handelt, aus dem sich sukzessive HisF_C1, HisF_C2 und HisF_N2 entwickelt haben (Abbildung 22; Tabelle 6).

Von den vier Modulen konnten nur das erste N-terminale (HisF_N1) und das erste C-terminale (HisF_C1), jeweils aus 74 Aminosäureresten bestehend, in Lösung gebracht werden (4.4.1; 4.7.1). Die beiden kürzeren Viertel HisF_N2 (48 Aminosäurereste) und HisF_C2 (58 Aminosäurereste) konnten nur als Fusionsproteine löslich hergestellt werden, ließen sich aber in isolierter Form nicht anreichern. Auch die Festphasenpeptidsynthese beider Fragmente hatte keinen Erfolg. Die Endlängenfragmente zeigten eine erhöhte Tendenz zur Aggregation und ließen sich nur in Anwesenheit des Kosolvens TFE, das Sekundärstrukturen stabilisiert (Buck, 1998), in Lösung bringen und halten (4.10). Daher ist anzunehmen, dass weder HisF_N2 noch HisF_C2 die Mindestlänge besitzen, einen ausreichend großen hydrophoben Kern im Inneren des Proteins auszubilden, um den durch die

Ordnung des Peptidrückgrates und der Seitenketten auftretenden Entropieverlust auszugleichen (Zitzewitz & Matthews, 1999). Diese Tatsache wurde bereits bei Studien an Fragmenten der Barnase (Kippen *et al.*, 1994), des Chymotrypsin-Inhibitors 2 (De Prat Gay *et al.*, 1995) und der Dihydrofolat-Reduktase (Gegg *et al.*, 1997) beobachtet. Bei den Fragment-Experimenten mit diesen „Eindomänen-Enzymen“ konnten nur Polypeptidketten mit einer Mindestlänge von 100 Aminosäureresten signifikante Sekundärstruktur annehmen. Auch die Analysen von Fragmenten des Tachylectin 2, einem Protein mit β -Propeller Faltung (Klasse der WD40-*repeats*; 1.3.2) (Yadid & Tawfik, 2007) und des Tumorsuppressors p16^{INK4} (Klasse der Ankyrin-*repeats*; 1.3.3) (Zhang & Peng, 2000) deuten darauf hin, dass nur Fragmente bestehend aus 100 (Tachylectin 2) bzw. 66 (p16^{INK4}) Aminosäureresten stabil sind. Die Untersuchungen nach Zerlegung in jeweils einzelne *repeats*, bestehend aus 50 bzw. 33 Aminosäuren, zeigten keine Aktivität der Oligomere (Tachylectin 2) bzw. die Module lagen unstrukturiert in Lösung vor (p16^{INK4}). Betrachtet man die $(\beta\alpha)_8$ -*barrel* Faltung ebenfalls als repetitive Abfolge von gemischten $\beta\alpha$ -Einheiten lässt sich erklären, warum die $(\beta\alpha)_2$ -Module HisF_N2 und HisF_C2 nur schlecht Sekundärstruktur annehmen und somit nicht in Lösung bleiben. Die beiden Module benötigen anscheinend ihren Partner aus der jeweiligen Hälfte (HisF_N1, HisF_C1) um, nachdem dieser eine Teilfaltung angenommen hat, sich an dessen vordefinierten Strukturen zu orientieren bzw. zu stabilisieren.

HisF_N1 lässt sich gut in reiner Form darstellen. Es besitzt wie HisF_N im β -Strang 1 einen Cysteinrest an Position 9, über den durch eine Disulfidbrücke zwei Monomere dimerisieren können (Abbildung 29A). Da das intrazelluläre Milieu reduzierend ist, sind nur wenige cytosolische Proteine mit Cystinen bekannt. Diese existieren normalerweise nur transient und haben eher funktionelle als stabilisierende Bedeutung (Raines, 1997; Linke & Jakob, 2003). Die Röntgenstrukturanalyse von Proteinen aus hyperthermophilen Archaeen legt jedoch nahe, dass Disulfidbrücken zur extremen Thermostabilität dieser Proteine beitragen könnten. Die bioinformatische Analyse mikrobieller Genome stützt diese Annahme (Mallick *et al.*, 2002; Beeby *et al.*, 2005). Auch im Fall von HisF_N1 aus dem hyperthermophilen Bakterium *T. maritima* hat die Bildung des Cystins eine stabilisierende Wirkung, wie durch den Vergleich der Auffaltungskurven unter oxidierenden und reduzierenden Bedingungen gezeigt werden konnte (Abbildung 46). Die Entropie des entfalteten Zustands sinkt durch die Ausbildung von Disulfidbrücken und somit wird die Differenz

zur freien Energie des gefalteten Zustands erhöht, wodurch die Proteinstabilität steigt (Betz, 1993). Einen weiteren Grund für die relativ hohe Stabilität von HisF_N1 stellt sein Oligomerisierungsgrad dar. Mit Hilfe von chemischem *cross-linking* und analytischer Ultrazentrifugation konnten deutliche Hinweise für das Vorhandensein eines Tetramers gewonnen werden (4.4.2.2; 4.4.2.3). Zusammenfassend kann gesagt werden, dass der Assoziationszustand von HisF_N1 durch die Ausbildung eines intermolekularen Cystins vom Monomer auf das Dimer und über nicht kovalente Wechselwirkungen zweier Homodimere vom Dimer auf das Tetramer erhöht werden kann. Die Erhöhung des Oligomerisierungsgrads ist ein zentraler Mechanismus, der zur Erhöhung der Proteinstabilität beiträgt (Vieille & Zeikus, 2001). Ein Teil der Thermostabilität der dimeren Phosphoribosylanthranilat (PRA) Isomerase (TrpF) aus *Thermotoga maritima* (Thoma *et al.*, 2000) bzw. der tetrameren TIM aus *Pyrococcus wosei* (Walden *et al.*, 2001) beruht auf ihrem erhöhten Oligomerisierungsgrad im Vergleich zu den monomeren bzw. dimeren Enzymen aus mesophilen Organismen. Die Sekundärstruktur von HisF_N1 entspricht bezogen auf das β -Faltblatt weitgehend der aus der Röntgenstruktur von HisF abgeleiteten, wohingegen sich ein Teil der α -Helices nicht auszubilden scheint (Tabelle 9). Durch die Anwesenheit relativ geringer Mengen (10 % v/v) des organischen Kosolvens Trifluorethanol (TFE), das sich an die Polypeptidkette anlagert und dabei vor allem die Bildung von α -Helices fördert (Buck, 1998), konnte die native Sekundärstrukturelementverteilung induziert werden (Tabelle 10). Die Lösungsmittelexposition des Tryptophan 23 ist in (HisF_N1)₄ ähnlich wie in (HisF_N)₂. Dies wird durch das Fluoreszenzemissionsmaximum unter nativen Bedingungen bei 345 nm für (HisF_N1)₄ (Abbildung 36) im Vergleich zu 341 nm für (HisF_N)₂ (Abbildung 27) deutlich. Auch die chemische Entfaltung der Sekundärstruktur des tetrameren Viertels und der dimeren N-terminalen Hälfte sind identisch (Abbildung 28A und 37). Auf der Grundlage dieser Ergebnisse kann somit spekuliert werden, dass die Sekundär- und Tertiärstruktur, sowie die Stabilität von HisF_N vornehmlich auf den gefalteten Zustand von HisF_N1 in der Hälfte zurückgehen. HisF_N1 ist somit der Faltungskeim der N-terminalen Hälfte, an dem sich das zweite Viertel HisF_N2 orientiert.

Um die tetramere Form von HisF_N1 weiter zu stabilisieren, wurde versucht, in *layer* 4 von (HisF_N1)₄ ein Wasserstoffbrückennetzwerk zu etablieren. Durch rationales Design in diesem *layer*, in dem sich im wildtypischen HisF ein Salzbrückencluster

bildet, konnte das artifizielle Protein HisF_CC, bestehend aus zwei kovalent verknüpften HisF_C-Hälften kompaktiert werden (Höcker *et al.*, 2004). Durch die beiden Austausche E46Q und R5A in HisF_N1 wurde ein größerer Anteil der resultierenden Varianten HisF_N1_E46Q und HisF_N1_R5A_E46Q in der löslichen Fraktion exprimiert (4.5.2). In den Varianten bildete sich ebenfalls über die Cysteine an Position 9 zweier Monomere eine Disulfidbrücke (Abbildung 39). Wie beim wildtypischen HisF_N1 assoziieren zwei Homodimere weiter zum Tetramer, wie mit Hilfe von chemischem *cross-linking* und analytischer Ultrazentrifugation gezeigt werden konnte (4.5.3.2; 4.5.3.3). Die Sekundärstrukturen der Varianten und des Wildtyps von HisF_N1 sind identisch (Abbildung 43). Auf die Anwesenheit von TFE reagieren HisF_N1_E46Q und HisF_N1_R5A_E46Q ähnlich wie der Wildtyp, bei der Doppelvariante reichen allerdings schon 5 % (v/v) TFE aus, um die Sekundärstrukturelementverteilung zu erreichen, die aus der Kristallstruktur berechnet worden war (Tabelle 12). Die Tertiärstruktur des resultierenden ($\beta\alpha$)₈-barrels dürfte der des wildtypischen HisF_N1 Proteins ähnlich sein, da die Emissionsmaxima der Varianten mit 343 nm (HisF_N1_E46Q) und 345 nm (HisF_N1_R5A_E46Q) dem des Wildtyp-Viertels von 341 nm ähneln (Abbildung 45). Die beiden Austausche führen nicht zur erhofften Stabilitätserhöhung von HisF_N1 (Abbildung 46A).

Das Viertel HisF_C1 kann nicht in so großen Mengen rein dargestellt werden wie HisF_N1 (4.7.1). Auch lag das Protein nicht in homogener Form vor, da mittels analytischer Gelfiltration von dem Protein der monomere, der tetramere und auch höhermolekulare Zustände nachgewiesen werden konnten (4.7.2). Die Sekundär- und Tertiärstruktur ist nicht so stark ausgebildet wie in HisF_N1, HisF_N und HisF_C. Ein erheblicher Anteil der vorhergesagten α -Helices wird nicht ausgebildet (Tabelle 14) und das Emissionsmaximum von 343 nm unter nativen Bedingungen (Abbildung 50) deutet darauf hin, dass das Tryptophan 156 deutlich mehr Kontakt mit dem Lösungsmittel hat als in (HisF_C)₂, dessen Emissionsmaximum bei 336 nm liegt (Abbildung 27B).

Auch HisF_C1 wurde versucht weiter zu stabilisieren. In einer ersten Variante wurde der Austausch Tyrosin 143 gegen Histidin vorgenommen. Dieser Austausch konnte durch gerichtete Evolution der N-terminalen Hälfte von HisF_C*C identifiziert werden (Seitz *et al.*, 2007) und führte in diesem Kontext zu einer Erhöhung der Löslichkeit der resultierenden HisF_C*C-Variante. Basierend auf Moleküldynamik- (MD) Simulationen *in silico* wurde postuliert, dass das eingeführte Histidin entweder über

seinen Imidazolring ein Wasserstoffbrückennetzwerk über die Kontaktfläche der beiden Hälften hinweg bildet oder aufgrund seiner größeren Polarität im Vergleich zum Imidazolring des Tyrosins und der daraus resultierenden höheren Solvatisierungsenergie den nativen Zustand des Proteins stabilisiert (Seitz *et al.*, 2007). In HisF_C1_Y143H führt der Austausch zu einer erhöhten Löslichkeit der Variante nach Rückfaltung (Abbildung 53). Analytische Gelfiltration zeigt, dass HisF_C1_Y143H als Mono-, Di- und Tetramer vorliegt, die Proteinlösung ist also wie bei HisF_C1 nicht homogen. Höhermolekulare Aggregate sind jedoch nicht mehr vorhanden (Abbildung 54A), der Austausch scheint damit die Aggregationstendenz zu verringern. Die Sekundär- und Tertiärstruktur der Variante ist im Vergleich zum Wildtyp deutlich pertubiert, mehr als 60 % der Polypeptidkette liegen ungeordnet in Lösung vor (Abbildung 57) und das Fluoreszenzemissionsmaximum bei 353 nm deutet auf ein vollkommen lösungsmittel-exponiertes Tryptophan 156 hin (Abbildung 59A). Der Austausch Y143H hat somit zwar einen positiven Effekt auf die Löslichkeit des Proteins, verringert aber gleichzeitig die in HisF_C1 nur schwach vorhandene Tendenz native Sekundär- und Tertiärstruktur auszubilden.

Zwei HisF_N1-Moleküle bilden über ihre Cysteine an Position 9 eine Disulfidbrücke und dimerisieren dadurch, was zu einer Stabilisierung des daraus resultierenden $(\beta\alpha)_4$ -Moduls führt. Nach Superpositionierung von HisF_N1 mit HisF_C1 wurde Alanin 128 in HisF_C1 als korrespondierender Rest zu C9 in HisF_N1 identifiziert und durch Cystein ersetzt. Dieser Austausch führt dazu, dass die resultierende Doppelvariante HisF_C1_A128C_Y143H nun überwiegend als Tetramer vorliegt (Abbildung 54B, 56) und eine höhere Stabilität aufweist als HisF_C1 (Abbildung 60). Tryptophan 156 ist in $(\text{HisF_C1_A128C_Y143H})_4$ mit einem Fluoreszenz-emissionsmaximum von 334 nm (Abbildung 59A) vergleichbar gut vom Lösungsmittel abgeschirmt wie in $(\text{HisF_C})_2$, das ein Maximum bei 336 nm aufweist (Abbildung 27). Dies spricht dafür, dass $(\text{HisF_C1_A128C_Y143H})_4$ und $(\text{HisF_C})_2$ ähnliche Quartär- und Tertiärstrukturen aufweisen. Mit zwei Aminosäureaustauschen konnte somit das *molten globule*-ähnliche HisF_C1 – mit nativer Sekundär-, aber ohne definierte Tertiärstruktur – stabilisiert werden. Damit wurden die Kriterien für eine eigenständige Domäne erstmals für das erste C-terminale $(\beta\alpha)_2$ -Modul erfüllt.

In *layer 4* von $(\text{HisF_C1_A128C_Y143H})_4$ kommen die Glutamatreste (E167) der vier Polypeptide zu liegen. Diese wurden gegen Glutamin ausgetauscht, um die Bildung eines Wasserstoffbrückennetzwerkes zu ermöglichen. Auch die Variante

HisF_C1_A128C_Y143H_E167Q lagert sich wiederum überwiegend zum Tetramer zusammen (Abbildung 54C, 56), zeigt aber im Vergleich zu HisF_A128C_Y143H eine nochmals höhere Stabilität gegenüber Harnstoff (Abbildung 60B). Sekundär- und Tertiärstruktur sind vergleichbar mit (HisF_C1_A128C_Y143H)₄ und (HisF_C)₂ (Tabelle 18; Abbildung 59). Auf die Anwesenheit von TFE reagiert die Dreifach-Variante allerdings empfindlicher als die Doppel-Variante (Tabelle 19), weshalb man annehmen kann, dass HisF_C1_A128C_Y143H_E167Q ein geeigneteres Polypeptidgerüst zur Ausbildung nativer Sekundärstruktur besitzt. Die Optimierung der Kontaktfläche am Stabilitätspol der ($\beta\alpha$)₈-Einheit von (HisF_C1_A128C_Y143H)₄ durch E167Q führte also zu einem Stabilitätsgewinn und einer Kompaktierung des Proteins im Vergleich zu HisF_C1. Dagegen erzielten die durchgeführten Aminosäureaustausche an der Kontaktfläche von HisF_N1 nicht diesen Effekt. Nach dem „Rosetta Stone Modell“ zur Evolution von Protein-Protein-Interaktionen (Marcotte *et al.*, 1999) kommt es nach der Fusion zweier Subdomänen zu einer Optimierung der unspezifischen Wechselwirkungen an der Kontaktfläche, wodurch ein stabileres Protein entstehen kann. Daher kann man vermuten, dass HisF_C1 eine unspezifischere Kontaktfläche zur Interaktion mit weiteren HisF_C1-Molekülen oder anderen Subdomänen besitzt, als HisF_N1.

Eine andere Möglichkeit zur weiteren Stabilisierung der beiden Viertel HisF_N1 und HisF_C1 könnte die gerichtete Evolution bieten. Die gängigen Selektionssysteme, wie die Selektion auf Löslichkeit eines Protein-Repertoires mit fusionierter Chloramphenicolacetyltransferase (CAT) (Maxwell *et al.*, 1999) oder die Selektion basierend auf der erhöhten Fluoreszenz des fusionierten *Green-Fluorescent-Protein* (GFP) (Waldo *et al.*, 1999), erscheinen jedoch eher ungeeignet. Durch die Fusion relativ großer Selektionsproteine (CAT, 25 kDa; GFP, 27 kDa) würden die kleinen Viertel HisF_N1 und HisF_C1 (jeweils ca. 8,3 kDa) vermutlich automatisch in Lösung gehalten werden, was die Wahrscheinlichkeit zur Isolierung verbesserter Varianten stark verringert.

Die Ergebnisse der Untersuchungen an HisF_N1 und HisF_C1 haben Auswirkung auf die Vorstellungen über den Faltungsmechanismus von ($\beta\alpha$)₈-*barrel* Proteinen, der bisher anhand von Fragmentstudien untersucht wurde. So zeigten Zitzewitz *et al.* (1999) anhand von aus 105 bis 214 Aminosäuren bestehenden Fragmenten der α -Untereinheit der Tryptophansynthase (TrpA), die sie durch sukzessive Verkürzung des C-Terminus erhalten hatten, dass lokale $\beta\alpha\beta$ -Module ein Grundgerüst bilden,

das die Faltung des gesamten Proteins unterstützt. In diesem Zusammenhang könnten HisF_N1 und HisF_C1 zwei solche Initiationseinheiten darstellen, die bei der Faltung von *tmHisF* als erstes ihre Sekundär- und Tertiärstrukturen ausbilden. In einer weiteren Studie wurde das N-terminale und das C-terminale Fragment der Phosphoribosylanthranilat (PRA) Isomerase aus *Saccharomyces cerevisiae*, bestehend aus den $\beta\alpha$ -Einheiten 1 bis 6 und 7 bis 8 untersucht (Eder & Kirschner, 1992). Dabei nahm lediglich das N-terminale Fragment eine definierte Struktur an, während sich das C-terminale Fragment nur nach gemeinsamer Rückfaltung mit der $(\beta\alpha)_{1-6}$ -Einheit zu einem stöchiometrischen und katalytisch aktiven Komplex mit dieser zusammengelagert hat. Daher wurde für dieses Enzym ein [6+2]-Mechanismus für die Faltung postuliert. Die nativen Strukturen von HisF_N und HisF_C, sowie ihre Fähigkeit einen katalytisch aktiven HisF_NC- Komplex zu bilden, legen dagegen einen [4+4]-Mechanismus nahe (Höcker *et al.*, 2001). Die Ergebnisse der vorliegenden Arbeit lassen zu, diesen Mechanismus in einen [(2+2)+(2+2)]-Mechanismus zu zerlegen, wobei die Initiation der Faltung von HisF_N1 und HisF_C1 ausgeht.

Die Tatsache, dass HisF_N1 und HisF_C1 durch die Bildung einer intermolekularen Disulfidbrücke zwischen zwei Cysteinen (HisF_N1 nativ; HisF_C1 künstlich eingeführt) native Struktur annehmen, homogene Assoziationszustände bilden und stabil sind, untermauert die Annahme, dass Proteine und Peptide durch die Anwesenheit eines Cystins stabilisiert werden können (Matthews *et al.*, 1987; Eder & Wilmanns, 1992; Clarke & Fersht, 1993; Wakarchuk *et al.*, 1994; Mansfeld *et al.*, 1997; Van den Burg *et al.*, 1998; Ivens *et al.*, 2002; Fenel *et al.*, 2004). Sie stützen weiterhin die aufgestellte Hypothese, dass die Evolution der $(\beta\alpha)_8$ -barrel Faltung mit einem Gen begonnen hat, das für ein Viertel *tmHisF* kodiert. In diesem Rahmen sind HisF_N1 und HisF_C1 aufgrund bisheriger Untersuchungen die nächsten „Nachkommen“ dieses ursprünglichen $(\beta\alpha)_2$ -Moduls.

Die im Lichte der biochemischen Ergebnisse durchgeführte Inspektion bekannter HisF-Sequenzen (Abbildung 23) führte zu einer Modifizierung des „Dreistufen“-Modells der Evolution von $(\beta\alpha)_8$ -barrel Enzymen. Der Sequenzvergleich zeigte, dass jedes $(\beta\alpha)_2$ -Modul im C-terminalen Bereich des ersten β -Strangs ein konserviertes Cystein bzw. Serin (in HisF_C2) besitzt (Abbildung 23). Es erscheint somit plausibel, dass sowohl die $(\beta\alpha)_2$ -Einheiten, als auch die $(\beta\alpha)_4$ -Einheiten nach Duplikation und Fusion über intermolekulare Cystine stabilisierte Homotetramere bzw. Homodimere

bilden (Abbildung 70). Um die erste Stufe des Modells, also die Duplikation und anschließende Fusion von Vierteln experimentell nachzuahmen, wurden zwei Exemplare *hisF_N1* in tandem kloniert, was zu dem artifiziellen Protein *HisF_N1N1* führte. Diese $(\beta\alpha)_4$ -Einheit wird ebenfalls über die beiden Cysteine an Position 9 zum Homodimer verknüpft (Abbildung 71), wie durch analytische Gelfiltration und Ultrazentrifugation bestätigt werden konnte (Abbildung 72, 73). Die Sekundärstrukturelemente, vor allem die α -Helices, sind ausgeprägter als in *HisF_N1* (Tabelle 22). Die Auffaltung der Sekundärstrukturelemente von $(\text{HisF_N1N1})_2$ erfolgt bei ähnlichen Harnstoffkonzentrationen wie bei $(\text{HisF_N1})_4$ und verläuft mit vergleichbar geringer Kooperativität (Abbildung 76). Die Tertiärstruktur von *HisF_N1N1* muss aufgrund des Emissionsmaximums bei 341 nm (Abbildung 75) nahezu identisch zu der Tertiärstruktur von $(\text{HisF_N})_2$ sein. Ihre Entfaltung ist mit der von $(\text{HisF_C})_2$ vergleichbar und verläuft ähnlich kooperativ (Abbildung 76). Da *HisF_N1N1* ein stabiles Protein mit nativer Sekundärstruktur ist und ähnliche biophysikalische und proteinchemische Eigenschaften besitzt wie *HisF_N1*, stabilisiertes *HisF_C1* und die Hälften *HisF_N*, *HisF_C*, handelt es sich auch bei der duplizierten und fusionierten $(\beta\alpha)_2$ -Einheit um eine Proteindomäne. Dieses Ergebnis stützt das in dieser Arbeit aufgestellte Evolutionsmodell, wonach die Bildung eines intermolekularen Cystins zwischen zwei identischen, evolutiven Einheiten eine Stabilisierung und die Etablierung der $(\beta\alpha)_8$ -Faltung unterstützt.

Zusammenfassend lässt sich feststellen, dass die beiden Viertel *HisF_N1* und *HisF_C1* als Subdomänen der beiden Hälften *HisF_N* und *HisF_C* anzusehen sind. Beide Viertel bilden stabile und nativ gefaltete Proteine, die sich zu einem Tetramer zusammenlagern. *HisF_C1* benötigt allerdings mehrere Aminosäureaustausche, um die biophysikalischen und proteinchemischen Eigenschaften des ersten N-terminalen Viertels zu erreichen. In beiden $(\beta\alpha)_2$ -Modulen wirkt sich die Bildung eines intermolekularen Cystins positiv auf die Stabilität und den Assoziationszustand des Proteins aus. Gestützt auf diese Ergebnisse und die Analyse eines Sequenzalignments konnte ein neues Evolutionsmodell, in dem die Stabilisierung durch ein Cystin eine zentrale Rolle spielt, aufgestellt werden. Das artifizielle Protein *HisF_N1N1*, das die zweite Stufe im Evolutionsmodell repräsentiert, besitzt identische proteinchemische Eigenschaften wie die Hälften *HisF_N* und *HisF_C*. Weitere Experimente könnten darin bestehen zu versuchen, sowohl die Viertel *HisF_N1* und *HisF_C1*, als auch *HisF_N1N1* weiter zu stabilisieren. Weiterhin wäre

es interessant, ein Protein aus vier identischen, fusionierten $(\beta\alpha)_2$ -Modulen herzustellen, um die letzte Stufe des neuen Evolutionsmodells (Abbildung 70) experimentell nachzuahmen.

6 Zusammenfassung

Im Rahmen dieser Arbeit wurde die Evolution der $(\beta\alpha)_8$ -barrel Faltung aus kleineren Struktureinheiten am Beispiel der Imidazolglycerinphosphat Synthase (HisF) aus *Thermotoga maritima* untersucht. Die $(\beta\alpha)_8$ -barrel Faltung besteht auf den ersten Blick aus einer einzigen Domäne. Es gibt in der Literatur jedoch zahlreiche Hinweise dafür, dass dieser Faltungstyp aus zwei $(\beta\alpha)_4$ -halbbarrel entstanden ist. Die bereits früher charakterisierten N- und C-terminalen Hälften von HisF, namentlich HisF_N $[(\beta\alpha)_{1-4}]$ und HisF_C $[(\beta\alpha)_{5-8}]$, wurden erneut heterolog in *Escherichia coli* exprimiert und mit Hilfe eines optimierten Protokolls nochmals aufgereinigt. Analytische Gelfiltration zeigte, dass sich die Hälften zum jeweiligen Homodimer zusammenlagern, welches im Falle von HisF_N über eine bisher nicht identifizierte Disulfidbrücke stabilisiert wird. Beide Proteine besitzen native Sekundär- und Tertiärstrukturen, die in chemischen Auffaltungssexperimenten kooperativ denaturieren. Diese Ergebnisse stützen die in früheren Studien getroffenen Aussagen, dass es sich bei den $(\beta\alpha)_4$ -Einheiten um eigenständige Domänen handelt und diese eine Vorläufereinheit der $(\beta\alpha)_8$ -barrel Faltung darstellen.

Sequenz- und Strukturuntersuchungen deuteten neben der zweifachen Symmetrie auch auf eine vierfache Symmetrie innerhalb der $(\beta\alpha)_8$ -barrel Faltung hin. Diese findet sich auch in HisF, dessen vier $(\beta\alpha)_2$ -Einheiten HisF_N1 $[(\beta\alpha)_{1-2}]$, HisF_N2 $[(\beta\alpha)_{3-4}]$, HisF_C1 $[(\beta\alpha)_{5-6}]$ und HisF_C2 $[(\beta\alpha)_{7-8}]$ zusätzlich Sequenz- und Strukturähnlichkeiten aufweisen. Diese Befunde deuteten darauf hin, dass HisF aus einem $(\beta\alpha)_2$ -Vorläuferbarrel entstanden sein könnte. Diese Hypothese sollte durch die Herstellung und Charakterisierung der HisF-Viertel überprüft werden. Dazu wurden die Viertel heterolog in *E. coli* exprimiert.

Bei der anschließend versuchten Aufreinigung zeigte sich jedoch, dass HisF_N2 und HisF_C2 in isolierter Form nicht stabil waren. Nach Fusion mit der Glutathion S-Transferase (GST) waren beide Proteine zwar nachweisbar, konnten aber nach Abtrennung der GST nicht in Lösung gehalten werden. Aus diesen Gründen wurden HisF_N2 und HisF_C2 mithilfe der Festphasenpeptidsynthese dargestellt. Jedoch zeigte sich auch hier, dass die Proteine keine native Faltung annehmen und erst mit dem Kosolvens TFE in Lösung zu bringen sind. Ein Grund dürfte in der unzureichenden Länge der Polypeptidketten liegen, die mit 48 (HisF_N2) bzw. 58

Aminosäureresten (HisF_C2) wohl zu kurz sind, um einen hydrophoben Kern auszubilden.

Das aus 74 Aminosäuren bestehende Viertel HisF_N1 bildete wie die Hälfte HisF_N ein über eine Disulfidbrücke verknüpftes Homodimer. Analytische Gelfiltration, chemisches *cross-linking* und analytische Ultrazentrifugation ergaben, dass sich zwei Homodimere zum Tetramer zusammenlagern und somit vermutlich ein vollständiges $(\beta\alpha)_8$ -*barrel* ausbilden. Mittels fern-UV Circular dichroismus- und Fluoreszenz-emissionsspektroskopie konnte gezeigt werden, dass das (HisF_N1)₄-Molekül in Lösung native Sekundär- und Tertiärstruktur aufweist. Die schwach kooperative Entfaltung der Sekundärstruktur mittels Harnstoff belegte, dass es sich um ein stabiles und kompaktes Protein handelt. Versuche zur weiteren Stabilisierung von HisF_N1 mittels rationalem Design schlugen fehl.

Das ebenfalls aus 74 Aminosäuren bestehende Viertel HisF_C1 erwies sich als weniger stabil als HisF_N1, besitzt aber Anteile an nativer Sekundär- und Tertiärstruktur. In Lösung liegt HisF_C1 als ein Gemisch von Mono-, Tetra- und höheren Oligomeren vor. Nach Einführung von zwei Aminosäureaustauschen (A128C und Y143H) bildete jedoch auch HisF_C1 ein Tetramer aus, wie durch analytische Gelfiltration, chemisches *cross-linking* und analytische Ultrazentrifugation gezeigt werden konnte. Während der Y143H Austausch die Löslichkeit förderte, wurde durch den Austausch A128C die Ausbildung einer Disulfidbrücke induziert, worauf sich – analog zu HisF_N1 und HisF_N – zwei kovalent verknüpfte Homodimere zum Tetramer zusammenlagerten. Auffaltungskurven zeigten, dass die beiden Aminosäureaustausche auch zu einer Erhöhung der Stabilität im Vergleich zum wildtypischen HisF_C1 führten. Eine weitere leichte Stabilisierung gelang durch den Austausch E167Q, der mit dem Ziel der Rekonstruktion eines nativen Wasserstoffbrückennetzes im zentralen β -*barrel* eingeführt wurde.

Basierend auf diesen Ergebnissen wurde ein Modell zur Evolution von $(\beta\alpha)_8$ -*barrels* aus $(\beta\alpha)_2$ -Einheiten erstellt. Demnach lagerten sich zunächst zwei identische $(\beta\alpha)_2$ -Einheiten zum über eine Disulfidbrücke stabilisierten Homodimer zusammen. Durch die Assoziation zweier Homodimere zum Tetramer bildete sich bereits eine erste $(\beta\alpha)_8$ -*barrel* Struktur aus. Im nächsten Schritt kam es zur Duplikation des $(\beta\alpha)_2$ -Gens, gefolgt von der in tandem Fusion der beiden Tochtergene. Das Produkt des resultierenden $(\beta\alpha)_4$ -Gens lagerte sich über eine oder mehrere Disulfidbrücken zum Homodimer zusammen. Um diese Hypothese weiter zu überprüfen, wurden zwei

hisF_N1 Gene in tandem fusioniert und das resultierende *hisF_N1N1*-Gen in *E. coli* exprimiert. Das erzeugte rekombinante Protein HisF_N1N1 besitzt native Sekundär- und Tertiärstruktur, ist vergleichbar stabil wie das $(\beta\alpha)_4$ -Halbbarrel HisF_N und bildet ebenso wie dieses in Lösung ein Homodimer.

Die Ergebnisse dieser Arbeit liefern Hinweise darauf, dass es sich bei HisF_N1 und HisF_C1 um eigenständige Domänen handelt, die Eigenschaften einer ursprünglichen $(\beta\alpha)_2$ -Einheit aufweisen. HisF_N1 ist dabei wohl die ursprünglichere $(\beta\alpha)_2$ -Einheit, da es von vornherein ein Tetramer bildete und nicht erst wie HisF_C1 durch gezielte Aminosäureaustausche stabilisiert werden musste.

7 Literaturverzeichnis

- Amersham Biosciences** (1999). Rapid and efficient purification and refolding of a (His)₆-tagged recombinant protein produced in *E. coli* as inclusion bodies. Application Note HiTrap Chelating: 1-4.
- Andrade, M. A. und P. Bork** (1995). HEAT repeats in the Huntington's disease protein. *Nat Genet* **11**(2): 115-6.
- Andrade, M. A., C. Petosa, S. I. O'Donoghue, C. W. Muller und P. Bork** (2001). Comparison of ARM and HEAT protein repeats. *J Mol Biol* **309**(1): 1-18.
- Andrade, M. A., C. P. Ponting, T. J. Gibson und P. Bork** (2000). Homology-based method for identification of protein repeats using statistical significance estimates. *J Mol Biol* **298**(3): 521-37.
- Banner, D. W., A. C. Bloomer, G. A. Petsko, D. C. Phillips, C. I. Pogson, I. A. Wilson, P. H. Corran, A. J. Furth, J. D. Milman, R. E. Offord, J. D. Priddle und S. G. Waley** (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature* **255**(5510): 609-14.
- Batchelor, A. H., D. E. Piper, F. C. de la Brousse, S. L. McKnight und C. Wolberger** (1998). The structure of GABP α /beta: an ETS domain- ankyrin repeat heterodimer bound to DNA. *Science* **279**(5353): 1037-41.
- Beeby, M., B. D. O'Connor, C. Ryttersgaard, D. R. Boutz, L. J. Perry und T. O. Yeates** (2005). The genomics of disulfide bonding and protein stabilization in thermophiles. *PLoS Biol* **3**(9): e309.
- Bertolaet, B. L. und J. R. Knowles** (1995). Complementation of fragments of triosephosphate isomerase defined by exon boundaries. *Biochemistry* **34**(17): 5736-43.
- Betz, S. F.** (1993). Disulfide bonds and the stability of globular proteins. *Protein Sci* **2**(10): 1551-8.
- Bloemendal, H., W. de Jong, R. Jaenicke, N. H. Lubsen, C. Slingsby und A. Tardieu** (2004). Ageing and vision: structure, stability and function of lens crystallins. *Prog Biophys Mol Biol* **86**(3): 407-85.
- Bolton, D., P. A. Evans, K. Stott und R. W. Broadhurst** (2001). Structure and properties of a dimeric N-terminal fragment of human ubiquitin. *J Mol Biol* **314**(4): 773-87.
- Bork, P. und E. V. Koonin** (1998). Predicting functions from protein sequences-- where are the bottlenecks? *Nat Genet* **18**(4): 313-8.

- Bosnali, M.** (2003). Herstellung und Charakterisierung von $(\beta\alpha)_2$ -Modulen eines $(\beta\alpha)_8$ -Barrel Enzyms: Implikationen für Evolution und Faltung. Fachbereich Biologie der Mathematisch-Naturwissenschaftlichen Fakultät, Universität zu Köln. **Diplomarbeit**.
- 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**: 248-54.
- Buck, M.** (1998). Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q Rev Biophys* **31**(3): 297-355.
- Caetano-Anolles, G., H. S. Kim und J. E. Mittenthal** (2007). The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc Natl Acad Sci U S A* **104**(22): 9358-63.
- Chervenka, C. H.** (1973). A manual of methods for the analytical ultracentrifuge. Palo Alto, Spinicio Divisions of Beckman Instruments.
- Chothia, C., J. Gough, C. Vogel und S. A. Teichmann** (2003). Evolution of the protein repertoire. *Science* **300**(5626): 1701-3.
- Chothia, C. und A. M. Lesk** (1986). The relation between the divergence of sequence and structure in proteins. *Embo J* **5**(4): 823-6.
- Clarke, J. und A. R. Fersht** (1993). Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. *Biochemistry* **32**(16): 4322-9.
- Cohn, E. und J. Edsall** (1943). Proteins, amino acids and peptides as ions and dipolar ions. New York, Reinhold: 370 - 381, 428 - 431.
- Coulson, A. F. und J. Moult** (2002). A unfold, mesofold, and superfold model of protein fold use. *Proteins* **46**(1): 61-71.
- Das, A. K., P. W. Cohen und D. Barford** (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *Embo J* **17**(5): 1192-9.
- De Prat Gay, G., J. Ruiz-Sanz, J. L. Neira, L. S. Itzhaki und A. R. Fersht** (1995). Folding of a nascent polypeptide chain in vitro: cooperative formation of structure in a protein module. *Proc Natl Acad Sci U S A* **92**(9): 3683-6.
- Doolittle, R. F.** (1995). The multiplicity of domains in proteins. *Annu Rev Biochem* **64**: 287-314.
- Doolittle, W. F.** (1999). Phylogenetic classification and the universal tree. *Science* **284**(5423): 2124-9.

- Douangamath, A., M. Walker, S. Beismann-Driemeyer, M. C. Vega-Fernandez, R. Sterner und M. Wilmanns** (2002). Structural Evidence for Ammonia Tunneling across the $(\beta\alpha)_8$ Barrel of the Imidazole Glycerol Phosphate Synthase Bienenzyme Complex. *Structure (Camb)* **10**(2): 185-93.
- Durchschlag, H.** (1986). *Specific Volumes of Biological Macromolecules*. Berlin, Springer-Verlag.
- Eder, J. und K. Kirschner** (1992). Stable substructures of eightfold beta alpha-barrel proteins: fragment complementation of phosphoribosylanthranilate isomerase. *Biochemistry* **31**(14): 3617-25.
- Eder, J. und M. Wilmanns** (1992). Protein engineering of a disulfide bond in a beta/alpha-barrel protein. *Biochemistry* **31**(18): 4437-44.
- Eriksson, A. E., L. S. Cousens und B. W. Matthews** (1993). Refinement of the structure of human basic fibroblast growth factor at 1.6 Å resolution and analysis of presumed heparin binding sites by selenate substitution. *Protein Sci* **2**(8): 1274-84.
- Fani, R., M. Brilli, M. Fondi und P. Lio** (2007). The role of gene fusions in the evolution of metabolic pathways: the histidine biosynthesis case. *BMC Evol Biol* **7 Suppl 2**: S4.
- Fani, R., P. Lio, I. Chiarelli und M. Bazzicalupo** (1994). The evolution of the histidine biosynthetic genes in prokaryotes: a common ancestor for the *hisA* and *hisF* genes. *J Mol Evol* **38**(5): 489-95.
- Farber, G. K. und G. A. Petsko** (1990). The evolution of alpha/beta barrel enzymes. *Trends Biochem Sci* **15**(6): 228-34.
- Fenel, F., M. Leisola, J. Janis und O. Turunen** (2004). A de novo designed N-terminal disulphide bridge stabilizes the *Trichoderma reesei* endo-1,4-beta-xylanase II. *J Biotechnol* **108**(2): 137-43.
- Fukuyama, K., Y. Nagahara, T. Tsukihara, Y. Katsube, T. Hase und H. Matsubara** (1988). Tertiary structure of *Bacillus thermoproteolyticus* [4Fe-4S] ferredoxin. Evolutionary implications for bacterial ferredoxins. *J Mol Biol* **199**(1): 183-93.
- Gaspar, J. A., C. Liu, K. A. Vassall, G. Meglei, R. Stephen, P. B. Stathopoulos, A. Pineda-Lucena, B. Wu, A. Yee, C. H. Arrowsmith und E. M. Meiering** (2005). A novel member of the YchN-like fold: solution structure of the hypothetical protein Tm0979 from *Thermotoga maritima*. *Protein Sci* **14**(1): 216-23.
- Gegg, C. V., K. E. Bowers und C. R. Matthews** (1997). Probing minimal independent folding units in dihydrofolate reductase by molecular dissection. *Protein Sci* **6**(9): 1885-92.
- Gindhart, J. G., Jr. und L. S. Goldstein** (1996). Tetrapeptide repeats are present in the kinesin light chain. *Trends Biochem Sci* **21**(2): 52-3.

- Gorina, S. und N. P. Pavletich** (1996). Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science* **274**(5289): 1001-5.
- Gouet, P., E. Courcelle, D. I. Stuart und F. Metoz** (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**(4): 305-8.
- Govindarajan, S. und R. A. Goldstein** (1996). Why are some proteins structures so common? *Proc Natl Acad Sci U S A* **93**(8): 3341-5.
- Greenfield, N. J.** (1996). Methods to Estimate the Conformation of Proteins and Polypeptides from Circular Dichroism Data. *Anal Biochem* **235**(1): 1-10.
- Groves, M. R., N. Hanlon, P. Turowski, B. A. Hemmings und D. Barford** (1999). The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* **96**(1): 99-110.
- Hanahan, D.** (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**(4): 557-80.
- Heinz, D. W., L. O. Essen und R. L. Williams** (1998). Structural and mechanistic comparison of prokaryotic and eukaryotic phosphoinositide-specific phospholipases C. *J Mol Biol* **275**(4): 635-50.
- Henn-Sax, M., B. Hocker, M. Wilmanns und R. Sterner** (2001). Divergent evolution of (betaalpha)₈-barrel enzymes. *Biol Chem* **382**(9): 1315-20.
- Heukeshoven, J. und R. Dernick** (1988). Improved silver staining procedure for fast staining in PhastSystem Development Unit. I. Staining of sodium dodecyl sulfate gels. *Electrophoresis* **9**(1): 28-32.
- Hildebrand, A., M. Romaris, L. M. Rasmussen, D. Heinegard, D. R. Twardzik, W. A. Border und E. Ruoslahti** (1994). Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor beta. *Biochem J* **302** (Pt 2): 527-34.
- Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen und L. R. Pease** (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**(1): 51-9.
- Höcker, B., S. Beismann-Driemeyer, S. Hettwer, A. Lustig und R. Sterner** (2001). Dissection of a (β α)₈-barrel enzyme into two folded halves. *Nat Struct Biol* **8**(1): 32-6.
- Höcker, B., J. Claren und R. Sterner** (2004). Mimicking enzyme evolution by generating new (β α)₈-barrels from (β α)₄-half-barrels. *Proc Natl Acad Sci U S A* **101**(47): 16448-53.
- Höcker, B., S. Schmidt und R. Sterner** (2002). A common evolutionary origin of two elementary enzyme folds. *FEBS Lett* **510**(3): 133-5.

- Holm, L. und C. Sander** (1998). Dictionary of recurrent domains in protein structures. *Proteins* **33**(1): 88-96.
- Holtzhauer, M.** (1996). Methoden in der Proteinanalytik. Berlin, Springer-Verlag.
- Inoue, H., H. Nojima und H. Okayama** (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**(1): 23-8.
- Iozzo, R. V.** (1998). Matrix proteoglycans: from molecular design to cellular function. *Annu Rev Biochem* **67**: 609-52.
- Ito, N., S. E. Phillips, C. Stevens, Z. B. Ogel, M. J. McPherson, J. N. Keen, K. D. Yadav und P. F. Knowles** (1991). Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. *Nature* **350**(6313): 87-90.
- Ivens, A., O. Mayans, H. Szadkowski, C. Jurgens, M. Wilmanns und K. Kirschner** (2002). Stabilization of a (betaalpha)₈-barrel protein by an engineered disulfide bridge. *Eur J Biochem* **269**(4): 1145-53.
- Jaenicke, R.** (1999). Stability and folding of domain proteins. *Prog Biophys Mol Biol* **71**(2): 155-241.
- Jansen, R. und M. Gerstein** (2000). Analysis of the yeast transcriptome with structural and functional categories: characterizing highly expressed proteins. *Nucleic Acids Res* **28**(6): 1481-8.
- Johnson, W. C., Jr.** (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins* **7**(3): 205-14.
- Jung, J. und B. Lee** (2001). Circularly permuted proteins in the protein structure database. *Protein Sci* **10**(9): 1881-6.
- Kajava, A. V.** (2001). Review: proteins with repeated sequence--structural prediction and modeling. *J Struct Biol* **134**(2-3): 132-44.
- Katoh, K., K. Kuma, H. Toh und T. Miyata** (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**(2): 511-8.
- Katoh, K., K. Misawa, K. Kuma und T. Miyata** (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**(14): 3059-66.
- Kinch, L. N. und N. V. Grishin** (2002). Evolution of protein structures and functions. *Curr Opin Struct Biol* **12**(3): 400-8.
- Kippen, A. D., J. Sancho und A. R. Fersht** (1994). Folding of barnase in parts. *Biochemistry* **33**(12): 3778-86.
- Knowles, J. R.** (1991). To build an enzyme. *Philos Trans R Soc Lond B Biol Sci* **332**(1263): 115-21.

- Kobe, B. und J. Deisenhofer** (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci* **19**(10): 415-21.
- Kobe, B. und J. Deisenhofer** (1995). A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374**(6518): 183-6.
- Laemmli, U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259): 680-5.
- Lamb, J. R., S. Tugendreich und P. Hieter** (1995). Tetrapeptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* **20**(7): 257-9.
- Lang, D., R. Thoma, M. Henn-Sax, R. Sterner und M. Wilmanns** (2000). Structural evidence for evolution of the beta/alpha barrel scaffold by gene duplication and fusion. *Science* **289**(5484): 1546-50.
- Lees, J. G., A. J. Miles, F. Wien und B. A. Wallace** (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* **22**(16): 1955-62.
- Lindqvist, Y. und G. Schneider** (1997). Circular permutations of natural protein sequences: structural evidence. *Curr Opin Struct Biol* **7**(3): 422-7.
- Linke, K. und U. Jakob** (2003). Not every disulfide lasts forever: disulfide bond formation as a redox switch. *Antioxid Redox Signal* **5**(4): 425-34.
- Lochner, A.** (2008). Stabilisierung und Aktivierung eines künstlichen $(\beta\alpha)_8$ Fass Proteins. Lehrstuhl für physikalische Biochemie. Regensburg, Universität Regensburg. **Diplomarbeit.**
- Lupas, A. N., C. P. Ponting und R. B. Russell** (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**(2-3): 191-203.
- Lux, S. E., K. M. John und V. Bennett** (1990). Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**(6261): 36-42.
- Lynch, M. und J. S. Conery** (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**(5494): 1151-5.
- Mallick, P., D. R. Boutz, D. Eisenberg und T. O. Yeates** (2002). Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. *Proc Natl Acad Sci U S A* **99**(15): 9679-84.
- Mansfeld, J., G. Vriend, B. W. Dijkstra, O. R. Veltman, B. Van den Burg, G. Venema, R. Ulbrich-Hofmann und V. G. Eijssink** (1997). Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J Biol Chem* **272**(17): 11152-6.

- Marcotte, E. M., M. Pellegrini, H. L. Ng, D. W. Rice, T. O. Yeates und D. Eisenberg** (1999). Detecting protein function and protein-protein interactions from genome sequences. *Science* **285**(5428): 751-3.
- Marcotte, E. M., M. Pellegrini, T. O. Yeates und D. Eisenberg** (1999). A census of protein repeats. *J Mol Biol* **293**(1): 151-60.
- Matthews, B. W., H. Nicholson und 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): 6663-7.
- Maxwell, K. L., A. K. Mittermaier, J. D. Forman-Kay und A. R. Davidson** (1999). A simple in vivo assay for increased protein solubility. *Protein Sci* **8**(9): 1908-11.
- McDonald, N. Q. und G. Peters** (1998). Ankyrin for clues about the function of p16INK4a. *Nat Struct Biol* **5**(2): 85-8.
- McLachlan, A. D.** (1979). Three-fold structural pattern in the soybean trypsin inhibitor (Kunitz). *J Mol Biol* **133**(4): 557-63.
- McLachlan, A. D.** (1987). Gene duplication and the origin of repetitive protein structures. *Cold Spring Harb Symp Quant Biol* **52**: 411-20.
- Merril, C. R., D. Goldman und M. L. Van Keuren** (1984). Gel protein stains: silver stain. *Methods Enzymol* **104**: 441-7.
- Mullis, K. B. und F. A. Faloona** (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* **155**: 335-50.
- Murzin, A. G., S. E. Brenner, T. Hubbard und C. Chothia** (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**(4): 536-40.
- Murzin, A. G., A. M. Lesk und C. Chothia** (1992). beta-Trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors interleukins-1 beta and 1 alpha and fibroblast growth factors. *J Mol Biol* **223**(2): 531-43.
- Nagano, N., C. A. Orengo und J. M. Thornton** (2002). One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions. *J Mol Biol* **321**(5): 741-65.
- Neer, E. J., C. J. Schmidt, R. Nambudripad und T. F. Smith** (1994). The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371**(6495): 297-300.
- Neer, E. J. und T. F. Smith** (1996). G protein heterodimers: new structures propel new questions. *Cell* **84**(2): 175-8.
- NewEnglandBiolabs** Instruction manual IMPACT Kit.

- Nicolet, Y. und C. L. Drennan** (2004). AdoMet radical proteins--from structure to evolution--alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Res* **32**(13): 4015-25.
- Notredame, C., D. G. Higgins und J. Heringa** (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**(1): 205-17.
- Ohno, S.** (1970). *Evolution by Gene Duplication*. Heidelberg, Germany, Springer-Verlag.
- Orengo, C. A.** (1999). CORA--topological fingerprints for protein structural families. *Protein Sci* **8**(4): 699-715.
- Orengo, C. A., A. D. Michie, S. Jones, D. T. Jones, M. B. Swindells und J. M. Thornton** (1997). CATH--a hierarchic classification of protein domain structures. *Structure* **5**(8): 1093-108.
- Orengo, C. A., I. Sillitoe, G. Reeves und F. M. Pearl** (2001). Review: what can structural classifications reveal about protein evolution? *J Struct Biol* **134**(2-3): 145-65.
- Orengo, C. A. und J. M. Thornton** (2005). Protein families and their evolution-a structural perspective. *Annu Rev Biochem* **74**: 867-900.
- Ortiz, A. R., C. E. Strauss und O. Olmea** (2002). MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison. *Protein Sci* **11**(11): 2606-21.
- Pace, C. N. und J. M. Scholtz** (1997). Measuring the conformational stability of a protein. Proofs for Protein Structure: A practical approach. T. E. Creighton. Oxford, IRL Press: 299-321.
- Pace, C. N., F. Vajdos, L. Fee, G. Grimsley und T. Gray** (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* **4**(11): 2411-23.
- Pan, H., A. S. Raza und D. L. Smith** (2004). Equilibrium and kinetic folding of rabbit muscle triosephosphate isomerase by hydrogen exchange mass spectrometry. *J Mol Biol* **336**(5): 1251-63.
- Peifer, M., S. Berg und A. B. Reynolds** (1994). A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell* **76**(5): 789-91.
- Peters, K. und F. M. Richards** (1977). Chemical cross-linking: reagents and problems in studies of membrane structure. *Annu Rev Biochem* **46**: 523-51.
- Peti, W., T. Herrmann, O. Zagnitko, S. K. Grzechnik und K. Wuthrich** (2005). NMR structure of the conserved hypothetical protein TM0979 from *Thermotoga maritima*. *Proteins* **59**(2): 387-90.

- Poirot, O., E. O'Toole und C. Notredame** (2003). Tcoffee@igs: A web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res* **31**(13): 3503-6.
- Ponting, C. P., L. Aravind, J. Schultz, P. Bork und E. V. Koonin** (1999). Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. *J Mol Biol* **289**(4): 729-45.
- Ponting, C. P. und R. B. Russell** (1995). Swaposins: circular permutations within genes encoding saposin homologues. *Trends Biochem Sci* **20**(5): 179-80.
- Ponting, C. P. und R. R. Russell** (2002). The natural history of protein domains. *Annu Rev Biophys Biomol Struct* **31**: 45-71.
- Price, S. R., P. R. Evans und K. Nagai** (1998). Crystal structure of the spliceosomal U2B''-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* **394**(6694): 645-50.
- Priestle, J. P., H. P. Schar und M. G. Grutter** (1988). Crystal structure of the cytokine interleukin-1 beta. *Embo J* **7**(2): 339-43.
- Provencher, S. W. und J. Glockner** (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**(1): 33-7.
- Raines, R. T.** (1997). Nature's transitory covalent bond. *Nat Struct Biol* **4**(6): 424-7.
- Redfern, O., A. Grant, M. Maibaum und C. Orengo** (2005). Survey of current protein family databases and their application in comparative, structural and functional genomics. *J Chromatogr B Analyt Technol Biomed Life Sci* **815**(1-2): 97-107.
- Rossmann, M. G., D. Moras und K. W. Olsen** (1974). Chemical and biological evolution of nucleotide-binding protein. *Nature* **250**(463): 194-9.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis und H. A. Erlich** (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**(4839): 487-91.
- Salem, G. M., E. G. Hutchinson, C. A. Orengo und J. M. Thornton** (1999). Correlation of observed fold frequency with the occurrence of local structural motifs. *J Mol Biol* **287**(5): 969-81.
- Sambrook, J., E. F. Fritsch und Maniatis, T.** (1989). *Molecular cloning: a laboratory manual*. New York, Cold Spring Harbour.
- Sanger, F., S. Nicklen und A. R. Coulson** (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**(12): 5463-7.
- Schägger, H. und G. von Jagow** (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**(2): 368-79.

- Scheufler, C., A. Brinker, G. Bourenkov, S. Pegoraro, L. Moroder, H. Bartunik, F. U. Hartl und I. Moarefi** (2000). Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* **101**(2): 199-210.
- Schmid, F. X.** (1997). Optical spectroscopy to characterize protein conformation and conformational changes. Protein Structure: a practical approach. T. E. Creighton. New York, Oxford University Press: 261-297.
- Sears, L. E., L. S. Moran, C. Kissinger, T. Creasey, H. Perry-O'Keefe, M. Roskey, E. Sutherland und 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): 626-33.
- Sedgwick, S. G. und S. J. Smerdon** (1999). The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci* **24**(8): 311-6.
- Seitz, T., M. Bocola, J. Claren und R. Sterner** (2007). Stabilisation of a (beta/alpha)₈-barrel protein designed from identical half barrels. *J Mol Biol* **372**(1): 114-29.
- Sharma, S., Y. Zhou und B. R. Singh** (2006). Cloning, expression, and purification of C-terminal quarter of the heavy chain of botulinum neurotoxin type A. *Protein Expr Purif* **45**(2): 288-95.
- Sharp, P. A., B. Sugden und J. Sambrook** (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* **12**(16): 3055-63.
- Sikorski, R. S., M. S. Boguski, M. Goebel und P. Hieter** (1990). A repeating amino acid motif in CDC23 defines a family of proteins and a new relationship among genes required for mitosis and RNA synthesis. *Cell* **60**(2): 307-17.
- Slatko, B. E.** (1994). Thermal cycle dideoxy DNA sequencing. *Methods Mol Biol* **31**: 35-45.
- Smith, M. H.** (1970). CRC Handbook of Biochemistry. H. A. Sober. Cleveland, Chemical Rubber: C3 - C35.
- Soberon, X., P. Fuentes-Gallego und G. Saab-Rincon** (2004). In vivo fragment complementation of a (beta/alpha)₈ barrel protein: generation of variability by recombination. *FEBS Lett* **560**(1-3): 167-72.
- Söding, J. und A. N. Lupas** (2003). More than the sum of their parts: on the evolution of proteins from peptides. *Bioessays* **25**(9): 837-46.
- Söding, J., M. Remmert und A. Biegert** (2006). HHrep: de novo protein repeat detection and the origin of TIM barrels. *Nucleic Acids Res* **34**(Web Server issue): W137-42.

- Stark, G. R.** (1965). Reactions of cyanate with functional groups of proteins. 3. Reactions with amino and carboxyl groups. *Biochemistry* **4**(6): 1030-6.
- Sterner, R. und B. Höcker** (2005). Catalytic versatility, stability, and evolution of the (betaalpha)8-barrel enzyme fold. *Chem Rev* **105**(11): 4038-55.
- Studier, F. W. und B. A. Moffatt** (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**(1): 113-30.
- Svensson, L., I. Narlid und A. Oldberg** (2000). Fibromodulin and lumican bind to the same region on collagen type I fibrils. *FEBS Lett* **470**(2): 178-82.
- Terwilliger, T. C.** (1995). Engineering the stability and function of gene V protein. *Adv Protein Chem* **46**: 177-215.
- Thoma, R., M. Hennig, R. Sterner und K. Kirschner** (2000). Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. *Structure* **8**(3): 265-76.
- Thoma, R., G. Obmolova, D. A. Lang, M. Schwander, P. Jenö, R. Sterner und M. Wilmanns** (1999). Efficient expression, purification and crystallisation of two hyperthermostable enzymes of histidine biosynthesis. *FEBS Lett* **454**(1-2): 1-6.
- Thoma, R., M. Schwander, W. Liebl, K. Kirschner und R. Sterner** (1998). A histidine gene cluster of the hyperthermophile *Thermotoga maritima*: sequence analysis and evolutionary significance. *Extremophiles* **2**(4): 379-89.
- Tiefenbach, K.-J.** (1997). Hydrodynamische und Thermodynamische Untersuchungen an Protein-SDS-Komplexen. Institut für Biophysik und physikalische Biochemie, Universität Regensburg. **Diplomarbeit**.
- Towbin, H., T. Staehelin und J. Gordon** (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**(9): 4350-4.
- Urfer, R. und K. Kirschner** (1992). The importance of surface loops for stabilizing an eightfold beta alpha barrel protein. *Protein Sci* **1**(1): 31-45.
- Van den Burg, B., G. Vriend, O. R. Veltman, G. Venema und V. G. Eijssink** (1998). Engineering an enzyme to resist boiling. *Proc Natl Acad Sci U S A* **95**(5): 2056-60.
- van Stokkum, I. H., H. J. Spoelder, M. Bloemendal, R. van Grondelle und F. C. Groen** (1990). Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal Biochem* **191**(1): 110-8.
- Vieille, C. und G. J. Zeikus** (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**(1): 1-43.

- Vision, T. J., D. G. Brown und S. D. Tanksley** (2000). The origins of genomic duplications in Arabidopsis. *Science* **290**(5499): 2114-7.
- Vogel, C., M. Bashton, N. D. Kerrison, C. Chothia und S. A. Teichmann** (2004). Structure, function and evolution of multidomain proteins. *Curr Opin Struct Biol* **14**(2): 208-16.
- Vogel, C., S. A. Teichmann und J. Pereira-Leal** (2005). The relationship between domain duplication and recombination. *J Mol Biol* **346**(1): 355-65.
- Wakarchuk, W. W., W. L. Sung, R. L. Campbell, A. Cunningham, D. C. Watson und M. Yaguchi** (1994). Thermostabilization of the *Bacillus circulans* xylanase by the introduction of disulfide bonds. *Protein Eng* **7**(11): 1379-86.
- Walden, H., G. S. Bell, R. J. Russell, B. Siebers, R. Hensel und G. L. Taylor** (2001). Tiny TIM: a small, tetrameric, hyperthermostable triosephosphate isomerase. *J Mol Biol* **306**(4): 745-57.
- Waldo, G. S., B. M. Standish, J. Berendzen und T. C. Terwilliger** (1999). Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* **17**(7): 691-5.
- Whitmore, L. und B. A. Wallace** (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* **32**(Web Server issue): W668-73.
- Whitmore, L. und B. A. Wallace** (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89**(5): 392-400.
- Wierenga, R. K.** (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett* **492**(3): 193-8.
- Wilson, G. G. und N. E. Murray** (1991). Restriction and modification systems. *Annu Rev Genet* **25**: 585-627.
- Wolfe, K. H. und D. C. Shields** (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**(6634): 708-13.
- Woolley, D. W. und R. B. Merrifield** (1963). Anomalies of the structural specificity of peptides. *Ann N Y Acad Sci* **104**: 161-71.
- Yadid, I. und D. S. Tawfik** (2007). Reconstruction of functional beta-propeller lectins via homo-oligomeric assembly of shorter fragments. *J Mol Biol* **365**(1): 10-7.
- Yphantis, D. A.** (1964). Equilibrium Ultracentrifugation of Dilute Solutions. *Biochemistry* **3**: 297-317.
- Zhang, B. und Z. Peng** (2000). A minimum folding unit in the ankyrin repeat protein p16(INK4). *J Mol Biol* **299**(4): 1121-32.

- Zitzewitz, J. A. und C. R. Matthews** (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**(31): 10205-14.

8 Abkürzungsverzeichnis

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

A	Absorption; Adenin; Ampère
Å	Ångström ($1 \text{ Å} = 10^{-10} \text{ m}$)
app	apparent
APS	Ammoniumperoxodisulfat
A_x	Absorption bei einer Wellenlänge x nm
<i>bla</i>	β -Lactamase-Gen
bp	Basenpaare
BSA	Rinderserumalbumin (<i>bovine serum albumine</i>)
bzw.	beziehungsweise
C	lösliche Fraktion des Zellextraktes (<i>crude extract</i>); Cytosin
c	Konzentration
ca.	circa
CBD	Chitinbindedomäne
CD	Circular-Dichroismus
cm	Zentimeter ($1 \cdot 10^{-2} \text{ m}$)
C-terminal	Carboxyterminales Ende einer Polypeptidkette
d	Schichtdicke der Küvette [cm]
d.h.	das heißt
Da	Dalton [g/mol]
DNA	Desoxyribonukleinsäure
dNTP	Didesoxynukleotidtriphosphat (N = A, C, G oder T)
dsDNA	doppelsträngige Desoxyribonukleinsäure
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraessigsäure
<i>et al.</i>	und andere (Autoren) (<i>et alii</i>)
EtBr	Ethidiumbromid
f	Verdünnungsfaktor

FPLC	<i>fast performance liquid chromatography</i>
F _U	Anteil ungefaltetes Protein
g	Gramm; Zentrifugalbeschleunigung
G	Guanosin
GdmCl	Guanidiniumhydrochlorid
h	Stunde
<i>his</i>	Gene, die für Enzyme des Histidinbiosyntheseweges kodieren, z.B. kodiert <i>hisF</i> für das Enzym HisF
HisA	5'-ProFAR Isomerase
HisF	Imidazolglycerinphosphat Synthase
HisF_C	[(β α) ₅₋₈]-Einheit von HisF
HisF_C1	[(β α) ₅₋₆]-Einheit von HisF
HisF_C2	[(β α) ₇₋₈]-Einheit von HisF
HisF_N	[(β α) ₁₋₄]-Einheit von HisF
HisF_N1	[(β α) ₁₋₂]-Einheit von HisF
HisF_N2	[(β α) ₃₋₄]-Einheit von HisF
His-Tag	Aminosäureschwanz bestehend aus sechs Histidinen
HPLC	<i>high performance liquid chromatography</i>
IPTG	Isopropyl-β-D-thiogalaktopyranosid
K	Kelvin
kalk	kalkuliert
<i>kan</i>	Kanamycinnukleotidtransferase-Gen
kb	Kilobasenpaare
kDa	kiloDalton (1*10 ³ g/mol)
KP	Kaliumphosphat
l	Liter
LB	Luria-Bertani (-Medium)
ln	natürlicher Logarithmus
log	dekadischer Logarithmus
M	Molar (mol/l); Längenstandard (Marker)
m	Steigung; Milli (1*10 ⁻³); Meter
max.	maximal
MCS	<i>multiple cloning site</i>
min	Minute

MPa	Megapascal ($1 \cdot 10^3$ Pa)
MRW	mittleres Molekulargewicht der Aminosäuren [g/mol]
MW	Molekulargewicht (<i>molecular weight</i>)
MWCO	<i>molecular weight cut off</i>
n	Nano ($1 \cdot 10^{-9}$)
N _A	Anzahl der Aminosäuren
N-terminal	Aminoterminales Ende einer Polypeptidkette
OD _x	optische Dichte bei der Wellenlänge x nm
<i>ori</i>	Replikationsstartpunkt eines Plasmids (<i>origin of replication</i>)
P	unlösliche Fraktion des Zellextraktes (Pellet)
PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerase-Kettenreaktion
PD	Pellet nach der Dialyse
PDB	Protein Datenbank
pH	negativer dekadischer Logarithmus der Protonenkonzentration
pmol	pico mol ($1 \cdot 10^{-12}$ mol)
<i>rmsd</i>	<i>root mean square deviation</i>
rpm	Umdrehungen pro Minute (<i>rotations per minute</i>)
RT	Raumtemperatur
s	Sekunde; Sedimentationskoeffizient
S	Svedberg [$1 \cdot 10^{-13}$ s]
SCOP	<i>Structural Classification of Proteins</i>
SDS	Natriumdodecylsulfat (<i>sodium dodecyl sulfate</i>)
SOE	<i>overlap extension</i>
SV	Säulenvolumen
T	Temperatur; Thymidin
<i>T. maritima / tm</i>	<i>Thermotoga maritima</i>
T _A	Annealing-Temperatur [°C]
TBE	Tris-Borat-EDTA-Puffer
TCEP	Tris(2-carboxyethyl)phosphin Hydrochlorid
TEMED	N, N, N', N'-Tetramethylethyldiamin
TIM	Triosephosphat Isomerase
T _A	Annealing-Temperatur [°C]
T _M	Schmelztemperatur [°C]

T_E	Transformationseffizienz
Tris	Tris(hydroxymethyl)aminomethan
<i>trp</i>	Gene, die für Enzyme des Tryptophanbiosynthesweges kodieren
U	Unit, 1 U entspricht der Menge an Enzym, die unter Standardbedingungen 1 μ mol Substrat pro Minute umsetzt; Uracil
Upm	Umdrehungen pro Minute
UV	Ultraviolettes Licht
ÜD	Überstand nach der Dialyse
V	Volt
v/v	Volumen pro Volumen (<i>volume per volume</i>)
vgl.	vergleiche
w/v	Masse pro Volumen (<i>weight per volume</i>)
wt / WT	Wildtyp
z.B.	zum Beispiel
%GC	relativer molarer GC-Gehalt
Ø	Durchmesser
°C	Grad Celsius
μ	Mikro ($1 \cdot 10^{-6}$)
$0,1\% A_x$	spezifischer Extinktionskoeffizient bei x nm [$\text{cm}^2 \cdot \text{mg}^{-1}$]
λ	Wellenlänge [nm]
Θ_{MRW}	durchschnittliche Elliptizität pro Aminosäure [$\text{grad} \cdot \text{cm}^2 / \text{dmol}$]
Θ_{obs}	gemessene Elliptizität [mgrad]
ϵ_x	molarer Extinktionskoeffizient bei einer Wellenlänge x [$\text{M}^{-1} \cdot \text{cm}^{-1}$]

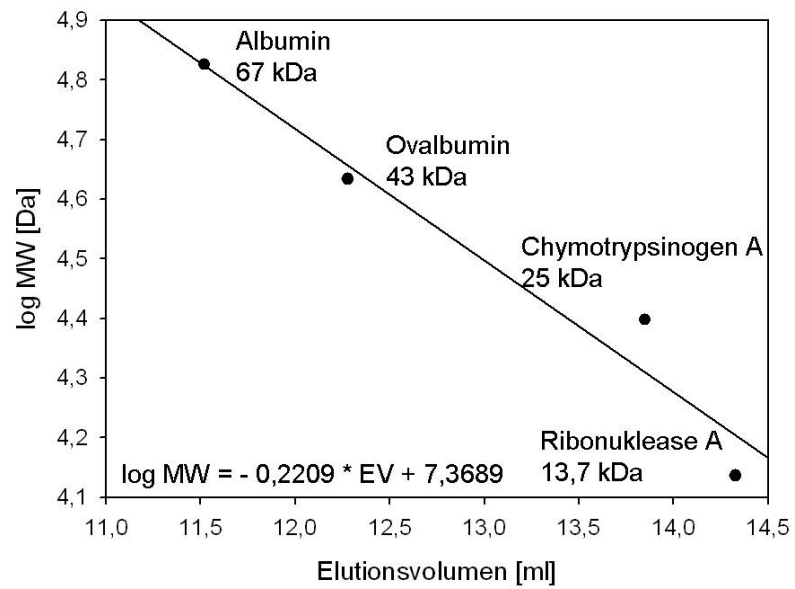
9 Anhang

Tabelle 23: Charakteristische Größen der vermessenen Proteine

Protein	Anzahl der Aminosäuren	M _{kalk} [Da] [#]	ε ₂₈₀ [*]	0,1% A ₂₈₀ [*]
HisF	253	27722	11460	0,41
HisF_N	122	13285	6990	0,53
HisF_C	131	14455	9970	0,69
HisF_N1	74	8352	6990	0,84
HisF_N1_E46Q	74	8353	6990	0,84
HisF_N1_R5A_E46Q	74	8268	6990	0,84
HisF_N2	48	4988	---	---
HisF_N2_S101C_F120W	48	4988	5500	1,1
HisF_N1XaN2	134	14884	12490	0,84
HisF_C1	74	8297	8480	1,02
HisF_C1_Y143H	74	8271	6990	0,85
HisF_C1_A128C_Y143H	74	8388	6990	0,83
HisF_C1_A128C_Y143H_E167Q	74	8315	6990	0,84
HisF_C2	58	6439	1490	0,23
HisF_C2_S201C_F210W_A220Q	58	6439	6990	1,1
HisF_C1XaC2	136	15141	15470	1,02
HisF_N1N1	148	16074	13980	0,87
HisF_C1C1	148	16597	13980	0,84

[#] berechnet mit dem Programm GeneRunner

^{*} berechnet nach Pace *et al.*, 1995

Kalibrierungsgerade für die analytische Gelfiltration (Superdex 75):

10 Danksagung

An dieser Stelle möchte ich mich bei all denen bedanken, die direkt bzw. indirekt zum Gelingen dieser Arbeit beigetragen haben.

Ein besonderer Dank gilt meinem Doktorvater Prof. Dr. Reinhard Sterner für das spannende, aber oft auch knifflige Thema und die Möglichkeit dieses sehr selbstständig bearbeiten zu dürfen. Stets war er am Fortschritt der Arbeit interessiert und hatte immer ein offenes Ohr für Probleme. Jeder Zeit stand seine Tür für Diskussionen und Anregungen offen.

Bei PD Dr. Wolfram Gronwald möchte ich mich ganz herzlich für die Übernahme des Zweitgutachtens bedanken.

Bei den Professoren Dr. Günther Hauska und Dr. Dr. Hans Robert Kalbitzer bedanke ich mich für die Übernahme des Vorsitzes bzw. des Drittprüfers bei der Doktorprüfung.

Frau Dr. Chiara Cabrele und Herrn Eduard Hochmuth danke ich für die Durchführung der Peptidsynthese bzw. der Aufnahme der Massenspektren.

Jeannette Ueckert gilt ein ganz großes Dankeschön für die Hilfe bei der täglichen Laborarbeit. Sie hat mir bei den ständigen Reinigungen und Gelfiltrationen sehr unter die Arme gegriffen. Herrn Dr. Helmut Durchschlag sei für seine Hilfe bei den UZ-Läufen, deren Auswertung und Interpretation gedankt. Christiane Endres danke ich herzlich für ihre ständige Bereitschaft Bestellungen aufzunehmen. Außerdem hat sie mir immer den Vormittag im Labor versüßt. Bei Klaus-Jürgen Tiefenbach möchte ich mich für anregende Diskussionen über die CD-Messungen und andere biophysikalische Methoden bedanken.

Sonja Fuchs, Hermine Reisner und Barbara Kellerer möchte ich für die Unterstützung bei der Bewältigung des Laboralltags danken. Matthias Zwick und vor allem Felix List danke ich für die zahlreichen angeregten Gespräche und Diskussionen im Büro, Labor und in der Freizeit. Obwohl Felix einige Zeit ausgezogen war, hat sein trockener Humor und Zynismus die Arbeit im Labor sehr aufgelockert. Matthias danke ich zudem für die Hilfe bei der Erstellung des strukturbasierten Sequenzalignments.

Bei Susanne Dietrich, Sandra Schlee und Alexander Ehrmann möchte ich mich für die unzähligen unterhaltsamen Mittagspausen bedanken. Für willkommene

Zigarettenpausen und angeregte Diskussionen über den FC Bayern München war Daniel Schneider stets zu haben. André Fischer danke ich für sein ständiges Interesse an meiner Arbeit und manch skurrile Idee, sowie für die Lauftreffs im ersten Jahr. Harald Guldán, Thomas Schwab, Jörg Claren, Tobias Seitz, Dr. Patrick Babinger, PD Dr. Rainer Merkl, Claudia Pauer, Hermann Zellner und allen ehemaligen Kollegen möchte ich für die freundschaftliche Arbeitsatmosphäre am Lehrstuhl danken.

Bei meinen Freunden Egi, Gerald, Obbe und Mollix möchte ich mich für die unterhaltsame Zeit während des Studiums und der Doktorarbeit bedanken. Egi sei besonders für die spannenden Fußballabende und zusammen mit Gerald für jeden Dienstagabend im Dombrowski gedankt.

Bei meiner Familie, vor allem bei meinen Eltern möchte ich mich für die psychologische und finanzielle Unterstützung während des Studiums und der Doktorarbeit bedanken. Meinem Bruder Michael danke ich für manch entspanntes Wochenende in Regensburg oder Kelheim.

Susanne möchte ich ganz ganz herzlich für die andauernde Unterstützung und die Aufmunterungen während der letzten vier Jahre danken. Obwohl sie lange Zeit über 600 km entfernt war, hat sie ständig zu mir gehalten und an mich geglaubt. Ohne sie hätte ich es nicht geschafft.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Regensburg, August 2008

Markus Richter