

**Synthesis and conformational analysis of polypeptides
related to the inhibitor of the DNA binding
and cell differentiation Id2**

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat)
der Fakultät für Chemie und Pharmazie
der Universität Regensburg



vorgelegt von
Noemi Colombo
aus Correzzana

Regensburg 2006

Die Arbeit wurde angeleitet von:

Prof. Dr. A. Buschauer

Promotiongesucht eingereicht am:

14. Dezember 2006

Promotionkolloquium am

16. Januar 2007

Prüfungsausschuß:

Vorsitzender: Prof. Dr. R. M. Gschwind

1. Gutachter: Prof. Dr. A. Buschauer

2. Gutachter: Prof. Dr. M. L. Gelmi

3. Prüfer: Prof. Dr. B. König

Die vorliegende Arbeit wurde in der Zeit von Oktober 2003 bis Oktober 2006 an der Fakultät für Chemie und Pharmazie der Universität Regensburg in der Arbeitsgruppe von Dr. C. Cabrele unter der Leitung von Prof. Dr. A. Buschauer angefertigt.

Dr. C. Cabrele danke ich herzlich für die Überlassung des interessanten Themas, und ihre stetige Unterstützung.

Prof. Dr. A. Buschauer danke ich für die Möglichkeit, diese Arbeit an seinem Lehrstuhl durchführen zu dürfen.

Ai miei genitori

(To my parents)

Table of Contents

I.	The role of the Id2 protein in cell cycle, cancer and neurobiology	1
I.1	Introduction	1
I.2	The Id Proteins	2
I.3	The Id2 Protein	3
I.3.1	The Id2 protein and cancer	6
I.3.2	The Id2 protein in the nervous system	6
I.3.3	The Id2 protein and the lymphatic system	9
I.3.4	The Id2 protein can drive to apoptosis	10
I.3.5	Id2 protein activity regulation	11
I.3.6	Nucleo-cytoplasmic shuttling of the Id2 protein	12
I.3.7	The Id2 protein promotes axonal growth	13
I.4	Conclusions and perspectives	15
I.5	Literature	16
II.	A chemical approach to the synthesis of large Id2 protein fragments	20
II.1	Introduction	20
II.2	Chemical synthesis of Id2 protein fragments	21
II.2.1	The C-terminal fragments	21
II.2.2	The HLH motif	26
II.3	CD spectroscopy	28
II.3.1	Peptides containing the Id2 C-terminal region	28
II.3.2	Peptides related to the Id2 HLH motif	30
II.3.3	Noncovalent interactions between the Id2 N-terminus and the HLH containing peptides	35
II.4	Native chemical ligation approach for the synthesis of Id2 large fragments	36
II.5	Conclusions	39
II.6	Literature	40
III.	A short Id2 protein fragment containing the nuclear export signal forms amyloid-like fibrils	43
III.1	Introduction	43
III.2	Synthesis and conformational characterization of Id2 analogues containing the C-terminal NES sequence	44

III.3	Synthesis and conformational characterization of Id2 analogues containing the NES sequence in the helix-2	51
III.4	Characterization of the insoluble form of the C-terminal NES region of Id2	54
III.5	Conclusions	65
III.6	Literature	66
IV.	Toward peptidomimetics as modulators of Id protein-protein interactions	68
IV.1	Introduction	68
IV.2	3-Carboxy-cyclopentylglycine (Cpg) as a tool for N-N linkage of peptides	71
IV.2.1	Synthesis of Id peptide dimers containing (<i>SRS</i>)-Cpg or (<i>RSR</i>)-Cpg	75
IV.2.2	Conformational characterization of the Cpg-containing Id peptides by CD spectroscopy	80
IV.2.3	Investigation of the interaction of the Cpg-containing peptides with the native HLH motif by CD spectroscopy	82
IV.2.4	Preliminary assays of the Cpg-containing peptides on a cellular model for atherosclerosis	84
IV.3	Small peptides containing the building block 3,4-(aminomethano)proline (Amp)	89
IV.3.1	Incorporation of Amp in small α/γ -peptides	90
IV.3.2	Structural investigations of the Amp-containing α/γ -peptides	93
IV.3.3	Incorporation of Amp in small α -peptides	106
IV.3.4	Structural investigations of the Amp-containing α -peptides	110
IV.4	Literature	113
V.	Summary	118
VI.	Experimental part	122
VI.1	Materials	122
VI.2	Methods	122
VI.2.1	Solid phase peptide synthesis (SPPS)	122
VI.2.2	Circular dichroism (CD) spectroscopy	125
VI.2.3	Nuclear magnetic resonance (NMR) spectroscopy	127
VI.2.4	Computational studies	130
VI.3	General procedures for peptide synthesis	132
VI.3.1	Peptide chain assembly by automated SPPS	132
VI.3.2	Peptide chain assembly by manual SPPS for Cpg-containing peptides (IV.10-13a/b)	133

VI.3.3	Peptide chain assembly by manual SPPS for Amp-containing peptides (IV.14-19a/b)	134
VI.4	General procedures for peptide purification and characterization	135
VI.5	Procedure for CD spectroscopy analysis	136
VI.6	Procedure for NMR spectroscopy analysis	136
VI.7	Procedure for conformational search analysis	137
VI.8	Literature	138
VII.	Appendix	140

Abbreviations

Amino acids

Residue	One-letter code	Three-letter code
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val
6-Aminohexanoic acid	--	Ahx
Norleucine	--	Nle

Other abbreviations

Ac	Acetyl	min.	Minutes
ACN	Acetonitrile	MS	Mass spectrometry
Boc	<i>tert</i> -Butyloxycarbonyl	M.W.	Molecular Weight
CD	Circular Dichroism	NMR	Nuclear Magnetic Resonance
COSY	Correlation Spectroscopy	NMP	<i>N</i> -Methylpyrrolidinone
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene	NOE	Nuclear Overhauser Effect
DIC	N,N'-Diisopropylcarbodiimide	PG	Protecting Group
DIPEA	Diisopropylethylamine	ppb	Part per billion
DMF	N,N-Dimethylformamide	RMS	Root Mean Square
EDT	1,2-Ethanedithiol	RP-	Reverse Phase High Pressure
		HPLC	Liquid Chromatography
equiv.	Equivalents	SPPS	Solid Phase Peptide Synthesis
ESI	Electrospray ionization	TFA	Trifluoroacetic acid
Fmoc	9-Fluorenylmethoxycarbonyl	TFE	2,2,2-Trifluoroethanol
h	Hours	TIS	Triisopropylsilane
HBTU	O-benzotriazole-N,N,N',N'-Tetramethyluronium hexafluorophosphate	TOCSY	Total Correlation Spectroscopy
HLH	Helix-Loop-Helix	t_R	Retention time
HOBt	Hydroxybenzotriazole	UV	Ultraviolet
LC-MS	Liquid Chromatography-Mass Spectrometry		
MALDI-ToF	Matrix-Assisted-Laser-Desorption-Ionization Time of Flight		
MBHA	Methylbenzhydrylamine		

I. The role of the Id2 protein in cell cycle, cancer and neurobiology

I.1 Introduction

Transcription factors are proteins that bind directly to a specific DNA sequence or to other DNA-bound proteins, in order to promote or block the expression of a target gene. The most common DNA-binding motifs found in proteins are the helix-turn-helix (HTH), the zinc-fingers, the leucine-zipper (LZ) and the helix-loop-helix (HLH). The family of the HLH transcription factors includes more than two hundred members which have been identified in a variety of organisms from yeast to mammals and can be classified in four different groups, depending on the presence or absence of additional functional domains along with the HLH motif (Figure 1) [1].

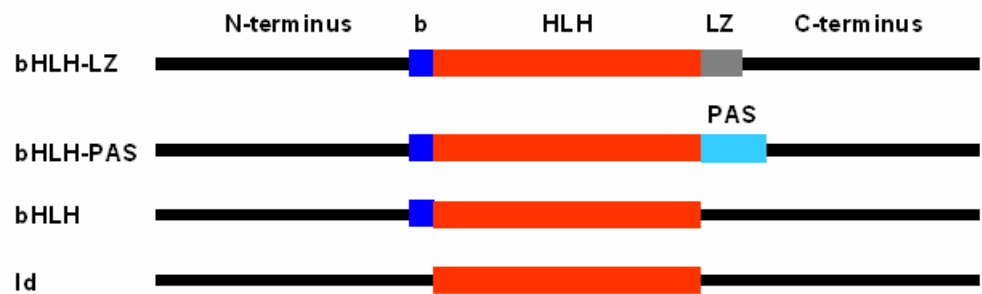


Figure 1: Classification of the HLH transcription factors based on the presence or absence of additional functional domains along with the HLH motif (b = basic region, DNA-binding site).

Another type of classification of the HLH proteins (classes I-VII) is based on their tissue distribution, dimerization capabilities and DNA-binding specificities. For example, class I [2] includes bHLH proteins which are ubiquitously expressed, like the factors E12 and E47, whereas class II includes mostly tissue-specific bHLH proteins, like NeuroD [3] and the myogenin-regulating factors MyoD and Myf5/6 [4]. Classes III (e.g. TFE3) and IV (e.g. Mad) contain HLH factors with an additional leucine-zipper C-terminal to the HLH region. Class V (emc and Ids) consists of proteins lacking the basic region N-terminal to the HLH motif which is essential for DNA binding, whereas class VI (e.g. HES-1) includes bHLH factors

displaying a conserved proline residue within the basic region. Finally, class VII (e.g. AHR) contains bHLH proteins with a conserved PAS domain C-terminal to the HLH motif.

The HLH proteins are highly conserved in the HLH region that consists of two amphipathic α -helices connected by a loop and is required for homo- and heterodimerization. Only after dimerization the bHLH proteins are able to recognize and bind a consensus DNA sequence, so called E-box (CANNTG), through the two correctly oriented basic regions (Figure 2). Therefore, protein dimerization by the formation of a parallel four-helix bundle is used to convert inactive monomeric molecules into transcriptionally active dimeric complexes at a specific time during cellular development [5]. Indeed, this mechanism underlies the bHLH-mediated regulation of the expression of tissue-specific- and cell-cycle-related genes [6-8].

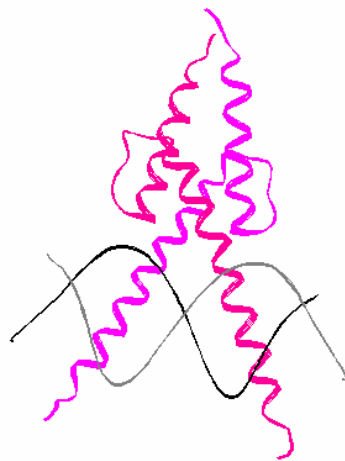


Figure 2: Crystal structure of the bHLH dimer of MyoD bound to DNA (PDB ID: 1MDY) [9].

1.2 The Id Proteins

The Id proteins (Id1 to Id4) belong to class V of the HLH transcription factors and are characterized by the absence of the DNA-binding region. They are involved in the control of cell-cycle progression and cell differentiation during development of many different cell types including nerves, muscles, cartilage and

bones. The name Id is related to the most important role of these proteins, which is that of being *in*hibitors of *DNA* binding and cell *d*ifferentiation.

The four Id proteins are similar in size (13-20 kDa) and share a highly conserved HLH region, whereas the rest of the sequence is unique for each protein. The Id HLH motif allows for dimerization with the related bHLH proteins of classes I and II, supposedly by forming a four- α -helix bundle stabilized by a combination of hydrophobic and electrostatic interactions. Due to the lack of the DNA-binding site, the Id proteins act as dominant-negative regulators of DNA transcription by sequestering ubiquitous (e.g. E47) and cell-type-restricted (e.g. MyoD) bHLH transcription factors, thus preventing the formation of the corresponding active dimers (Figure 3).

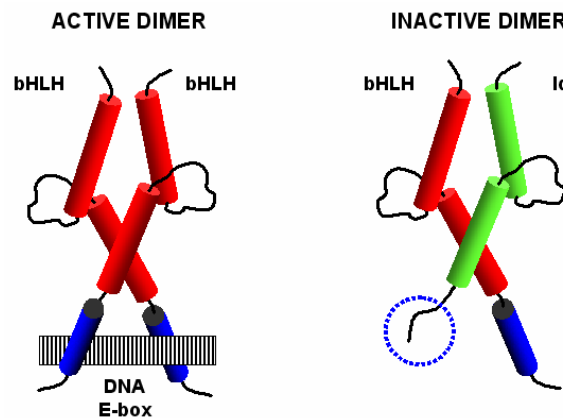


Figure 3: Mechanism of action of the Id proteins. The formation of Id-bHLH dimers prevents bHLH transcription factors from forming transcriptionally active DNA-binding complexes.

Whereas the Id proteins are widely expressed during development, they are low-expressed or absent in healthy adult cells. However, they are found to be overexpressed in many cancer types (Table 1) [10, 11].

Table 1: Examples of tumors in which the Id protein expression is dysregulated [11].

Tumor type	Dysregulated Id protein
Squamous cell carcinoma	Id1, Id2, Id3
Melanoma	Id1
Hepatocellular carcinoma	Id1
Colorectal adenocarcinoma	Id1, Id2, Id3
Pancreatic cancer	Id1, Id2
Thyroid cancer	Id1
Astrocytic tumor	Id1, Id2, Id3
Neuroblastoma	Id2
Ewing's sarcoma	Id2
Ovarian tumor	Id1, Id3
Cervical cancer	Id1
Endometrial carcinoma	Id1
Breast cancer	Id1, Id2, Id3
Prostate cancer	Id1, Id2
Malignant seminoma	Id1, Id2, Id3, Id4

The Id proteins can contribute to tumorigenesis by inhibiting cell differentiation, while stimulating proliferation and favoring tumor angiogenesis. Although there is thus far no evidence that the Id proteins act as oncogenes, however, their overexpression might mimic the activity of oncogenes or the loss of tumor suppressor mechanisms.

The role of the Id proteins in tumorigenesis, together with their low postnatal expression, makes them an attractive target for anti-cancer therapy [12].

I.3 The Id2 protein

Among the Id proteins, Id2 has attracted much interest because it is the only one to be expressed in the adult cerebellum upon physiological conditions. Besides heterodimerization with the parent HLH factors, Id2 can bind the retinoblastoma protein (pRb) [13] and the related pocket proteins p107 and p130. Moreover, Id2 is likely to be a target of the proto-oncogene *N-myc*, whose amplification and overexpression are hallmark of neuroblastoma [14].

I.3.1 The Id2 protein and cancer

The Id protein research is mainly focused on the role of these proteins in cancer. Id2 is a potent effector in tumorigenesis. Aberrant overexpression of Id2 has been associated, for example, with squamous cell carcinoma, colorectal adenocarcinoma, pancreatic cancer, astrocytic tumor, neuroblastoma, Ewing's sarcoma, breast cancer, prostate cancer and malignant seminoma (Table 1). However, it has been also reported that mice lacking Id2 develop intestinal tumors, indicating a tumor-inhibitory function of Id2 in the intestinal epithelium [12, 20-23].

I.3.2 The Id2 protein in the nervous system

In the peripheral nervous system Id2 has a pivotal role in cell-fate determination and oncogenesis. Recent data show that Id2 is highly expressed in two pediatric neuroectodermal tumors, neuroblastoma and Ewing's sarcoma (EWS). In these tumors, Id2 is constitutively activated by overexpression of *N-Myc* and EWS-Ets chimera proteins, respectively [24-26].

There is clear evidence that the Id proteins are essential proliferative factors for a variety of cell types. Several studies suggest that the progression in the G₁ cell-cycle phase requires the cooperative action of Ids on their natural targets, like bHLH, Ets and Pax. In mammalian cells, a crucial checkpoint control for proliferation is provided by the pocket proteins of the Rb family (pRb, p107, p130). Id2 is the only member of the Id family that is able to bind to each of the three pocket proteins in a cell-cycle regulated fashion [27]. Apparently, the very few non-conserved amino acid residues in the HLH region of Id2 compared with the corresponding domains of Id1, Id3 and Id4 are likely to confer the binding specificity of Id2 to the Rb family members.

Undifferentiated cells must overcome growth-inhibitory signals to maintain their proliferative potential, which requires the inactivation of negative regulators of cell-cycle progression, such as the pRb and the related pocket proteins. These share a stretch homologous sequence, so called pocket, which is necessary to bind other cellular proteins, including E2F, cyclins and MyoD. The interaction between

the Rb family members and these proteins results in the timely transcription of genes encoding factors essential for cell-cycle progression (E2F), or induction of the differentiated state (MyoD). Several studies indicate that pRb (or a related pocket protein) is required for differentiation of a variety of cell types. In normal cells, pRb activity is regulated by alternating changes in pRb phosphorylation. G₁ cyclins functioning as regulatory subunits of their partners cyclin-dependent kinases (Cdks) phosphorylate pRb, thus switching off its growth-inhibitory function.

Id2 physically associates with the hypophosphorylated active form of pRb, and this interaction negatively regulates Id2. Conversely, high levels of Id2 negatively regulate pRb-mediated growth suppression, probably in a direct manner by Id2-pRb association and, also, in an indirect manner by reducing the activity of the cyclin kinase inhibitors (CKIs) and altering the G₁ cyclin-Cdk complexes [28].

Disruption of the pRb pathway is a hallmark of cancer and it is widely accepted that normal pRb function must be removed in most of the human tumors. Extensive *in vivo* studies have been done to analyze the specific role of Id2 in driving the cell cycle. Comparison of genetic data from *pRb-null* and *Id2-pRb-double-null* embryos indicates that the ectopic proliferation and differentiation both in the nervous system and in the haematopoietic compartment of the *pRb-null* embryos are efficiently rescued by loss of Id2 [26]. Completion of normal development requires the inhibition of Id2 by pRb, which is essential for activation of the natural targets repressed by Id2 (bHLH and Ets-family members) [12].

Neuroblastoma, a tumor of the sympathetic nervous system, is the most common extracranial solid cancer in infants and children. Genomic amplification of the gene *N-Myc* is found in about 25% of primary neuroblastomas, which is one of the most reliable independent prognostic factors for this disease. *N-Myc* amplification clearly correlates with advanced stages of illness and is associated with poor outcome regardless of stage or age. Tumors holding this amplification are often refractory to treatment [29, 30]. *Myc* proto-oncogenes directly bind to the Id2 promoter activating it (Figure 5). By rising Id2 levels, *Myc* proteins avoid the blockage of cell-cycle progression imposed by the pRb pathway.

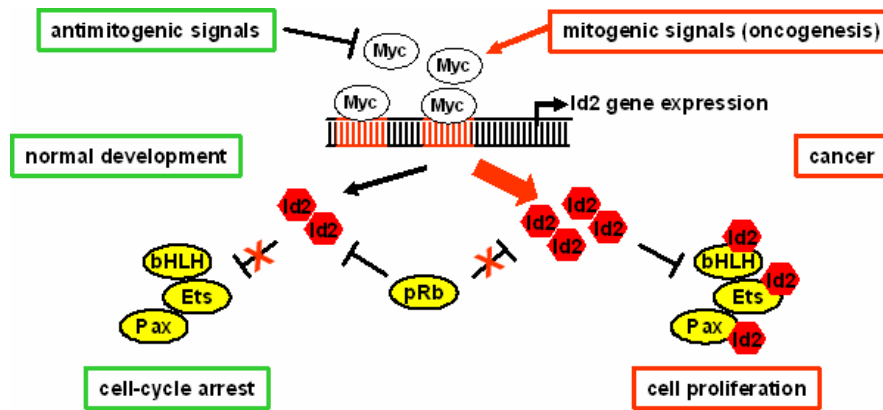


Figure 5: Role of pRb and Id2 in neuroblastoma.

This holds also for neuroblastoma cell lines carrying *N-Myc* gene amplification. Id2 expression determines the rate of proliferation of primary, immortalized and tumor cell lines. An immunohistochemical analysis of primary neuroblastoma shows that the expression of Id2 is strongly predictive of poor outcome. Overexpression of Id2 mediates cellular transformation and is required to maintain the malignant behavior of neuroblastoma cells [14]. Id2 is also a target for *c-Myc* and there are studies suggesting a direct correlation between *c-Myc* and Id2-mRNA in primary neuroblastomas [31].

Ewing's sarcomas (EWS) are highly malignant tumors arising in adolescents and young adults. They are characterized by a chromosomal translocation that results in the fusion of the 5' sequence of EWS with the 3' portion of various genes of the Ets oncoprotein family. The EWS-Ets chimera proteins act as aberrant transcription factors that can lead to oncogenic transformation. Id2 is a direct downstream target and potential mediator of the oncogenic EWS-Ets proteins. In addition, Id2 is highly expressed in EWS primary tumors, which may explain the high malignancy and poor prognosis of this disease. Thus, the oncogenic process of EWS may be attributed to the Id2-mediated inhibition of pRb that has a tumor suppressive activity in a series of cancers. Again, the *Myc*/Id2 pathway plays a crucial role in the tumorigenic process, as the induction of the Id2 gene is increased by *Myc* oncoproteins that are up-regulated by the EWS-Ets chimera proteins [24, 28, 32].

The cellular localization of Id2 has been recently proposed to be critical for the regulation of its function. In spite of the fact that other biological conditions

may regulate subcellular compartmentalization of Id2, the process of differentiation requires nuclear exclusion of the small protein. Cytoplasmic sequestration of Id2 has been described in two models of neuroectodermal and haematopoietic differentiation [18, 33]. Recently, the actin-associated protein enigma homolog (ENH) has been reported to act as a cytoplasm retention factor for Id2 [34]. ENH, whose expression increases during neuronal differentiation, sequesters Id2 in the cytoplasm and prevents cell-cycle progression and inhibition of bHLH driven by Id2. ENH belongs to a family of adaptor proteins that are anchored to the actin cytoskeleton. It possesses LIM domains, cysteine-rich double zinc fingers motifs known to mediate protein-protein interactions. Id2 nuclear shuttling is inhibited by the formation of a complex between the LIM region and the HLH motif. Therefore, ENH contributes to the differentiation of the nervous system and is a restraining factor of the oncogenic activity of Id2 in neuronal tumors [34]

I.3.3 The Id2 protein and the lymphatic system

Recent studies have established a role of Id2 in lymphopoiesis. Surviving *Id2*^{-/-} mice display a cell intrinsic defect in production of natural killer cells involved in immune function and also lack secondary lymphoid follicles of Peyer's patches in the intestine [35].

Hodgkin's lymphoma is one of the most common cancer forms in the lymphatic system. Pathologically, it is characterized by the presence of Reed-Sternberg cells deriving from a B lymphocyte (white blood cell) that has become cancerous. The Id proteins block B lymphocytes maturation at an early differentiation step, as demonstrated by gain-of-function studies [36]. Upon B cell activation, Id2 acts as a negative regulator to prevent potentially harmful effects caused by excessive immunological reactions. One of its special roles is to maintain low serum concentrations of immunoglobulin E by antagonizing E2A and Pax5 activities, which are both required for proper B cell activation. Primary Hodgkin-Reed/Sternberg (HRS) cells are characterized by the loss of B cell specific gene expression. Id2 is not detectable in normal B cells, whereas it is strongly and uniformly expressed in HRS cells in all cases of classical Hodgkin's lymphomas.

Thus, aberrant Id2 expression contributes significantly to the loss of B cell specific gene expression in HRS cells by inhibiting E2A and, probably, Pax5 [36-38].

I.3.4 The Id2 protein can drive to apoptosis

Apoptosis is one of the main mechanisms of programmed cell death (PCD), a process of deliberate life abandon by a cell in a multicellular organism. In contrast to necrosis, a form of cell death that results from acute cellular injury, apoptosis is carried out in an ordered process that generally confers advantages during an organism's life cycle. Defective apoptotic processes have been implicated in an extensive variety of diseases. Too much apoptosis causes cell-loss disorders, whereas too little results in uncontrolled cell proliferation, namely cancerous tumors.

The tendency of cells to undergo apoptosis when receiving divergent growth signals may result from a deregulated expression of cell-cycle modulators and the activation of their target genes at the wrong time. Apoptosis as well as all PCD processes are highly-controlled events. A feature shared by some positive regulators of mammalian cell-cycle progression is their ability to drive apoptosis when ectopically overexpressed [39, 40]. Several groups have reported on the pro-apoptotic properties of Ids in cell lines and in primary cell models. PCD induced by Id2 is recognizable as apoptosis by several criteria, including morphology, analysis of DNA fragmentation and of sub-diploid DNA content. Overexpression of Id2 increases apoptosis of 32D.3 cells, an interleukin-3 (IL-3)-dependent myeloid progenitor cell line, and of U2OS cells, an osteogenic sarcoma-derived cell line [17]. Id2 might drive apoptosis through a pathway that is different from the one utilized by the other Ids: in fact, Id2-induced apoptosis is independent of HLH-mediated dimerization and is not based on the interaction with the bHLH proteins [39, 40]. Florio *et al.* [17] have shown that the N-terminal region of Id2 and not the HLH motif is necessary for the enhancement of apoptosis: indeed, Id2 expression led to high mRNA levels of the pro-apoptotic molecule Bax in 32D.3 cells, and deletion of the Id2 HLH domain even incremented the PCD rate. Two models are suggested to explain such Id2 action. In the first model, the Id2 N-

terminal region is proposed to directly interact with the transcription factor complex mediating the expression of cell death genes within the nucleus. In the second model, the cytoplasmic Id2 pool might regulate cell death by binding molecules mediating or regulating apoptosis. In both models, these interactions would occur independently of dimerization via the HLH domain. However, Id2 enhancement of PCD may still be antagonized by proteins that normally bind to the Id2 HLH domain and may be physiologically regulated by the availability of HLH binding partners.

In a recent study the molecular regulation of apoptosis in fast plantaris muscles has been investigated [41]. Plantaris muscle mass has been found to be 22% lower in aged rather than in young adult animals, with the apoptotic index being 600% higher in the aged animals. This study also showed that changes in pro-apoptotic mRNA for apoptotic protease activating factor-1 (Apaf-1), Bax and Id2 are increased with aging. Bax and Bcl-2 protein levels were also altered differently in aged muscles when compared to young muscles. Significant positive correlations were observed between the changes in Id2 and Bax mRNAs, and Id2 and caspase-9 mRNAs. These data suggest that a pro-apoptotic environment may contribute to aging-associated atrophy in fast skeletal muscles, but apoptotic signaling differs by age [41]. Therefore, Id2 is a multifunctional protein that works as a gatekeeper of the G₁/S transition of the cell cycle [17].

1.3.5 Id2 protein activity regulation

Regulation of the level and of the activity of the Id proteins provides the cell with an important mechanism to control the balance between the active and inactive form of many bHLH transcription factors. Id proteins can be regulated at different levels, for example by transcriptional regulation, post-translational modifications and protein stability. Hara *et al.* [15] have reported that Id2 can undergo cyclin A(E)-Cdk2 mediated phosphorylation of a specific serine residue (Ser-5) in the N-terminus, where there is a consensus sequence for Cdks (Ser-Pro-Val-Arg), which is also present in Id2, Id3 and Id4. Phosphorylation prevents Id2 from interfering with the formation of DNA-binding complexes *in vitro*. Moreover, in serum-stimulated human diploid fibroblasts, phosphorylation occurs

in the late G₁, which correlates with the activation of cyclin E-Cdk2 and with the formation of a specific DNA-binding complex that would otherwise be prevented by excess of Id2. Therefore, Id2 must be phosphorylated to allow cell-cycle progression, and this event provides a link between the Cdk and bHLH activities, which may be crucial for the regulation of cell proliferation and differentiation.

Besides phosphorylation, Id2 activity in the cell is regulated also via the ubiquitin-proteasome system (UPS) that is responsible for specific degradation of numerous cellular regulatory proteins. The UPS activity involves the generation of a substrate-anchored polyubiquitin degradation signal and the destruction of the tagged protein by the 26S proteasome, followed by the release of free and reusable ubiquitin. The formation of ubiquitin conjugates requires the sequential action of three enzymes: E1, an ubiquitin-activating enzyme, E2, an ubiquitin carrier protein, and E3, a member of the ubiquitin-protein ligase family. In most cases, the first ubiquitin molecule is transferred to an ϵ -NH₂ group of an internal Lys residue. However, Id2 follows an N-terminal ubiquitination pathway, in which modification occurs at the N-terminal residue, whereas recognition of E3 probably involves a downstream motif.

Fajerman *et al.* [16] have reported that Id2 is ubiquitinated in an ATP-dependent manner in a cell-free reconstituted system. In addition, in this *in vitro* system the Ser-5/Ala Id2 mutant is equally sensitive to the UPS activity, thus excluding that phosphorylation of Id2 is a necessary signal for its ubiquitination and degradation. Lys-less (LL) Id2 analogue is degraded efficiently by the proteasome following ubiquitination, demonstrating that the first ubiquitin moiety is fused to the free α -NH₂ group of the protein. Fusion of a *Myc* tag to the N-terminal, but not to the C-terminal residue of Id2 stabilizes the protein, an effect reached also by deletion of the first 15 N-terminal amino acids, which proves that this domain serves as a recognition element for the ubiquitin ligase E3. The mechanisms and structural motifs that govern Id2 stability may have important implications in the regulation of the protein during normal differentiation and malignant transformation.

I.3.6 Nucleo-cytoplasmic shuttling of the Id2 protein

Appropriate subcellular localization is crucial for the proper function of numerous proteins. Some are constitutively nuclear, while others are actively transported into or out of the nucleus in a signal-dependent or -independent manner. Large proteins can shuttle between the nucleus and the cytoplasm through pore complexes due to their intrinsic nuclear localization signals (NLSs) and nuclear export signals (NESs). These domains are recognized by nuclear import and export receptors, respectively.

Due to their small size (13-20 kDa), passive diffusion is the main mechanism dictating the subcellular localization of the Id proteins, but it is likely that other regulatory pathways also exist [42, 43]. Id2 can move from the nucleus to the cytoplasm during neuronal differentiation into oligodendrocytes [44], and it is in the cytoplasm also during myeloid precursors differentiation [33].

Kurooka and Yokota [18] have shown that Id2 has the ability to shuttle between the nucleus and the cytoplasm. However, when passive diffusion is prevented by fusion with green fluorescent protein (GFP), Id2 is predominantly localized in the cytoplasm, and the C-terminus seems to be responsible for that. Id2 can be actively exported from the nucleus to the cytoplasm by a mechanism related to the chromosome region maintenance protein 1. In fact, Id2 contains two putative leucine-rich NESs, NES1 in the helix-2, and NES2 in the C-terminal region. Whereas the former is conserved among the Id members, the latter is characteristic only of Id2 and drives the nuclear export of the protein. This has an inhibitory effect on the Id2-mediated transcriptional repression, suggesting that the nuclear-cytoplasmic shuttling might be another type of post-translational regulation of the Id2 function [18].

I.3.7 The Id2 protein promotes axonal growth

The anaphase promoting complex/cyclosome and its activator Cdh1 (APC/C^{Cdh1}) restrain axonal growth, but their targets in neurons are unknown. In a recent work, Iavarone and co-workers [19] have reported that Id2 interacts with subunits of the APC/C complex in primary neurons. Expression of Cdh1 causes a

marked decrease in Id2 levels, an effect that can be prevented by proteasomal inhibition. Interestingly, activation of APC/C^{Cdh1} can prevent the sequestering of E47 by Id2 as well as by an Id2 mutant lacking the 15 N-terminal residues necessary for Id2 degradation via ubiquitination. APC/C^{Cdh1} triggers Id2 degradation through a destruction motif (D-box, residues 100-107) that is located at the C-terminal domain and is also present in Id1 and Id4. APC/C^{Cdh1} directs cell cycle-independent functions in postmitotic neurons, including the negative control of axonal growth. Activity of the Id proteins, which are depleted in quiescent cells, corresponds to the effect of inactivation of APC/C^{Cdh1} in postmitotic cells. As a matter of fact, overexpression of Id2 prevents cell-cycle arrest by a wide range of antiproliferative signals. Upon certain experimental conditions, ectopic Id2 is able to override the quiescent state and drive terminally differentiated cells back into the cell cycle. Id2 analogues containing mutations in the D-box enhance axonal growth in cerebellar granule neurons *in vitro* and overcome the myelin inhibitory signals for growth. In contrast, activation of bHLH transcription factors induces a cluster of genes with potent axonal inhibitory functions, including the gene coding for the Nogo receptor, a key transducer of myelin inhibition. Degradation of Id2 in neurons leads to accumulation of Nogo, linking the APC/C^{Cdh1} activity with bHLH target genes for the inhibition of axonal growth (Figure 6). Therefore, up-regulation of Id2 activity might be useful to reprogram quiescent neurons into the axonal growth mode [25].

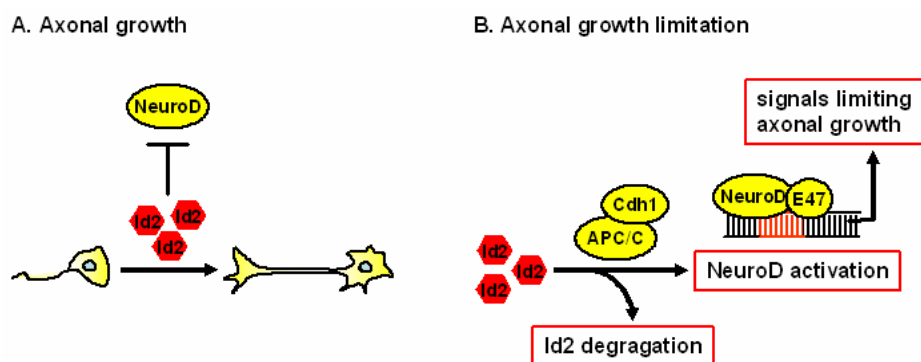


Figure 6: The role of Id2 during neuronal maturation.

I.4 Conclusions and perspectives

Therapeutic strategies based on Id protein targeting, in particular Id2, might be important for cancer patients, given the clear relationship of Id protein expression with aggressive cancer cell behavior. Drugs neutralizing the action of Id2 in tumors like neuroblastoma might reconstitute the integrity of the pRb pathway, the most critical antiproliferative defense available in mammalian cells. On the other hand, thinking of the ability of Id2 to promote axonal growth, specific molecules preventing Id2 degradation, thus allowing immature neurons to grow, would represent a hope for patients with spinal cord injuries.

While a lot of efforts have been made to study the biological and pathological aspects of the Id proteins, not much is known on their biophysical and structural features, although their mode of action is strongly dependent on the ability to fold into stable HLH domains which allow for homo- and heterodimerization with related bHLH proteins. Moreover, the N- and C-termini of the Id proteins have been shown to contain key motifs that exert important regulatory functions. The presented Ph. D. thesis has focused on a chemical approach to prepare large polypeptides based on Id2 and related analogues as tools for conformational investigations. Such synthetic and spectroscopic studies should be complementary to the biochemical ones and contribute to the further understanding of the complex behavior of the Id2 protein as well as to the future development of artificial Id2 protein modulators.

1.5 Literature

1. Massari, M. E., Murre, C. (2000) Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms, *Mol. Cell Biol.*, 20, 429-440.
2. Ik Tsen Heng, J., Tan, S. S. (2003) The role of class I HLH genes in neural development: have they been overlooked?, *Bioessays*, 25, 709-716.
3. Mutoh, H., Fung, B. P., Naya, F. J., Tsai, M. J., Nishitani, J., Leiter, A. B. (1997) The basic helix-loop-helix transcription factor BETA2/NeuroD is expressed in mammalian enteroendocrine cells and activates secretin gene expression, *Proc. Natl. Acad. Sci. USA*, 94, 3560-3564.
4. Perry, R. L., Rudnick, M. A. (2000) Molecular mechanisms regulating myogenic determination and differentiation, *Front. Biosci*, 5, 750-767.
5. Langlands, K., Yin, X., Anand, G., Prochownik, E. V. (1997) Differential interactions of Id proteins with basic-helix-loop-helix transcription factors, *J. Biol. Chem.*, 272, 19785-19793.
6. Lucas, M. E., Muller, F., Rudiger, R., Henion, P. D., Rohrer, H. (2006) The bHLH transcription factor hand2 is essential for noradrenergic differentiation of sympathetic neurons, *Development*, 133, 4015-4024.
7. Hershey, C. L., Fisher, D. E. (2004) Mitf and Tfe3: members of a b-HLH-ZIP transcription factor family essential for osteoclast development and function, *Bone*, 34, 689-696.
8. Reig, G., Cabrejos, M. E., Concha, M. L. (2006) Functions of BarH transcription factors during embryonic development, *Dev. Biol.*, in press.
9. Ma, P. C., Rould, M. A., Weintraub, H., Pabo, C. O. (1994) Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation, *Cell*, 77, 451-459.
10. Iavarone, A., Lasorella, A. (2004) Id proteins in neural cancer, *Cancer Lett.*, 204, 189-196.
11. Fong, S., Debs, R. J., Desprez, P. Y. (2004) Id genes and proteins as promising targets in cancer therapy, *Trends Mol. Med.*, 10, 387-392.
12. Perk, J., Iavarone, A., Benezra, R. (2005) Id family of helix-loop-helix proteins in cancer, *Nat. Rev. Cancer.*, 5, 603-614.

13. Iavarone, A., Garg, P., Lasorella, A., Hsu, J., Israel, M. A. (1994) The helix-loop-helix protein Id-2 enhances cell proliferation and binds to the retinoblastoma protein, *Genes Dev.*, **8**, 1270-1284.
14. Lasorella, A., Boldrini, R., Dominici, C., Donfrancesco, A., Yokota, Y., Inserra, A., Iavarone, A. (2002) Id2 is critical for cellular proliferation and is the oncogenic effector of N-myc in human neuroblastoma, *Cancer Res.*, **62**, 301-306.
15. Hara, E., Hall, M., Peters, G. (1997) Cdk2-dependent phosphorylation of Id2 modulates activity of E2A-related transcription factors, *Embo J.*, **16**, 332-342.
16. Fajerman, I., Schwartz, A. L., Ciechanover, A. (2004) Degradation of the Id2 developmental regulator: targeting via N-terminal ubiquitination, *Biochem. Biophys. Res. Commun.*, **314**, 505-512.
17. Florio, M., Hernandez, M. C., Yang, H., Shu, H. K., Cleveland, J. L., Israel, M. A. (1998) Id2 promotes apoptosis by a novel mechanism independent of dimerization to basic helix-loop-helix factors, *Mol. Cell Biol.*, **18**, 5435-5444.
18. Kurooka, H., Yokota, Y. (2005) Nucleo-cytoplasmic shuttling of Id2, a negative regulator of basic helix-loop-helix transcription factors, *J. Biol. Chem.*, **280**, 4313-4320.
19. Lasorella, A., Stegmuller, J., Guardavaccaro, D., Liu, G., Carro, M. S., Rothschild, G., de la Torre-Ubieta, L., Pagano, M., Bonni, A., Iavarone, A. (2006) Degradation of Id2 by the anaphase-promoting complex couples cell cycle exit and axonal growth, *Nature*, **442**, 471-474.
20. Mori, S., Nishikawa, S. I., Yokota, Y. (2000) Lactation defect in mice lacking the helix-loop-helix inhibitor Id2, *Embo J.*, **19**, 5772-5781.
21. Russell, R. G., Lasorella, A., Dettin, L. E. & Iavarone, A. (2004) Id2 drives differentiation and suppresses tumor formation in the intestinal epithelium, *Cancer Res.*, **64**, 7220-7225.
22. Nilsson, J. A., Nilsson, L. M., Keller, U., Yokota, Y., Boyd, K., Cleveland, J. L. (2004) Id2 is dispensable for myc-induced lymphomagenesis, *Cancer Res.*, **64**, 7296-7301.
23. Murphy, D. J., Swigart, L. B., Israel, M. A., Evan, G. I. (2004) Id2 is dispensable for Myc-induced epidermal neoplasia, *Mol. Cell. Biol.*, **24**, 2083-2090.
24. Fukuma, M., Okita, H., Hata, J., Umezawa, A. (2003) Upregulation of Id2, an oncogenic helix-loop-helix protein, is mediated by the chimeric EWS/ets protein in Ewing sarcoma, *Oncogene*, **22**, 1-9.

25. Iavarone, A., Lasorella, A. (2006) ID proteins as targets in cancer and tools in neurobiology, *Trends Mol. Med.*, 12, 588-94
26. Lasorella, A., Nosedà, M., Beyna, M., Yokota, Y., Iavarone, A. (2000) Id2 is a retinoblastoma protein target and mediates signalling by Myc oncoproteins, *Nature*, 407, 592-598.
27. Lasorella, A., Uo, T., Iavarone, A. (2001) Id proteins at the cross-road of development and cancer, *Oncogene*, 20, 8326-8333.
28. Lasorella, A., Iavarone, A., Israel, M. A. (1996) Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins, *Mol. Cell Biol.* 16, 2570-2578.
29. Lofstedt, T., Jogi, A., Sigvardsson, M., Gradin, K., Poellinger, L., Pahlman, S., Axelson, H. (2004) Induction of ID2 expression by hypoxia-inducible factor-1: a role in dedifferentiation of hypoxic neuroblastoma cells, *J. Biol. Chem.*, 279, 39223-39231.
30. Alaminos, M., Gerald, W. L., Cheung, N. K. (2005) Prognostic value of MYCN and ID2 overexpression in neuroblastoma, *Pediatr. Blood Cancer*, 45, 909-915.
31. Vandesompele, J., Edsjo, A., De Preter, K., Axelson, H., Speleman, F., Pahlman, S. (2003) ID2 expression in neuroblastoma does not correlate to MYCN levels and lacks prognostic value, *Oncogene*, 22, 456-460.
32. Nishimori, H., Sasaki, Y., Yoshida, K., Irifune, H., Zembutsu, H., Tanaka, T., Aoyama, T., Hosaka, T., Kawaguchi, S., Wada, T., Hata, J., Toguchida, J., Nakamura, Y., Tokino, T. (2002) The Id2 gene is a novel target of transcriptional activation by EWS-ETS fusion proteins in Ewing family tumors, *Oncogene*, 21, 8302-8309.
33. Tu, X., Baffa, R., Luke, S., Prisco, M., Baserga, R. (2003) Intracellular redistribution of nuclear and nucleolar proteins during differentiation of 32D murine hemopoietic cells, *Exp. Cell Res.*, 288, 119-130.
34. Lasorella, A., Iavarone, A. (2006) The protein ENH is a cytoplasmic sequestration factor for Id2 in normal and tumor cells from the nervous system, *Proc. Natl. Acad. Sci. USA.*, 103, 4976-4981.
35. Yokota, Y., Mansouri, A., Mori, S., Sugawara, S., Adachi, S., Nishikawa, S., Gruss, P. (1999) Development of peripheral lymphoid organs and natural killer cells depends on the helix-loop-helix inhibitor Id2, *Nature*, 397, 702-706.

36. Sugai, M., Gonda, H., Nambu, Y., Yokota, Y., Shimizu, A. (2004) Role of Id proteins in B lymphocyte activation: new insights from knockout mouse studies, *J. Mol. Med.*, 82, 592-599.
37. Renne, C., Martin-Subero, J. I., Eickernjager, M., Hansmann, M. L., Kuppers, R., Siebert, R., Brauninger, A. (2006) Aberrant expression of ID2, a suppressor of B-cell-specific gene expression, in Hodgkin's lymphoma, *Am. J. Pathol.*, 169, 655-664.
38. Mathas, S., Janz, M., Hummel, F., Hummel, M., Wollert-Wulf, B., Lusatis, S., Anagnostopoulos, I., Lietz, A., Sigvardsson, M., Jundt, F., Johrens, K., Bommert, K., Stein, H., Dorken, B. (2006) Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma, *Nat. Immunol.*, 7, 207-215.
39. Norton, J. D. (2000) ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis, *J. Cell. Sci.*, 113, 3897-3905.
40. Norton, J. D., Deed, R. W., Craggs, G., Sablitzky, F. (1998) Id helix-loop-helix proteins in cell growth and differentiation, *Trends Cell. Biol.*, 8, 58-65.
41. Pistilli, E. E., Siu, P. M., Alway, S. E. (2006) Molecular regulation of apoptosis in fast plantaris muscles of aged rats, *J. Gerontol. A Biol. Sci. Med. Sci.*, 61, 245-255.
42. O'Toole, P. J., Inoue, T., Emerson, L., Morrison, I. E., Mackie, A. R., Cherry, R. J., Norton, J. D. (2003) Id proteins negatively regulate basic helix-loop-helix transcription factor function by disrupting subnuclear compartmentalization, *J. Biol. Chem.*, 278, 45770-45776.
43. Samanta, J., Kessler, J. A. (2004) Interactions between ID and OLIG proteins mediate the inhibitory effects of BMP4 on oligodendroglial differentiation, *Development*, 131, 4131-4142.
44. Wang, S., Sdrulla, A., Johnson, J. E., Yokota, Y., Barres, B. A. (2001) A role for the helix-loop-helix protein Id2 in the control of oligodendrocyte development, *Neuron*, 29, 603-614.

II. A chemical approach to the synthesis of large Id2 protein fragments [1]

II.1 Introduction

Due to its implication in tumorigenesis and tumor growth, the Id family is a potential target for the diagnostics and therapy of tumor diseases, in which an increased Id protein activity has been detected [2]. Therefore, it is important to characterize these proteins from a biochemical and structural point of view, with the long-term aim of developing artificial molecules capable of modulating the Id function in pathological sceneries in a specific and efficient manner.

As described previously in chapter I, among the Id proteins, Id2 has attracted much interest because of its ability to bind pRb and the related pocket proteins, and because of the fact that it might be a target of the proto-oncogene N-Myc, whose amplification and overexpression are hallmark of neuroblastoma [3]. Our interest for this protein is related to its conformational properties and to the structural prerequisites dictating its unique protein–protein interaction profile. For this purpose, we decided to use a synthetic approach that would allow the preparation of a variety of polypeptides related to the native as well as chemically modified Id2 protein for structural studies. Therefore, we investigated the accessibility of large Id2 fragments by standard solid-phase peptide synthesis (SPPS). The sequence 36-110 was the largest one that could be assembled by stepwise SPPS in combination with the 9-fluorenylmethoxycarbonyl (Fmoc) strategy. Unfortunately, fragments containing the complete C-terminal region (residues 77-134) turned out to be poorly accessible with this chemical approach, which should be attributable to strong peptide chain aggregation during elongation. Peptide aggregation also prevented the possibility to obtain the full length C-terminal domain by native chemical ligation [4].

The secondary structure of the synthetic Id2 analogues was analyzed by circular dichroism (CD) spectroscopy. Noteworthy, all fragments containing the C-terminal sequence 77-110 showed low solubility and/or high tendency to aggregate. Even the usually highly stable Id2

HLH motif was negatively affected by the C-terminal elongation of helix-2, suggesting that this flanking region could modulate the folding and stability of the adjacent domain [1].

II.2 Chemical synthesis of Id2 protein fragments

II.2.1 The C-terminal fragments

The C-terminal domain of Id2 (product **II.1**, Table 1) consists of 58 residues, including a considerable number of the β -branched Leu (8) and Ile (4), as well as of Ser (9), Thr (5), Asx (6) and Glx (6). Peptide chain assembly starting from the C-end, Gly-134, using Fmoc chemistry and standard solid-phase procedures led to a crude peptide with very low homogeneity (Figure 1A), which prevented any attempt of purification. Instead, the shorter fragment 99-134 (**II.2**) was obtained as the major product (Figure 1B) and was then efficiently purified by preparative RP-HPLC. This suggests that the synthetic problems encountered in the preparation of the fragment 77-134 occurred during elongation from residue 98 to residue 77. However, the fact that the sequence 77-124 (**II.3**) was obtained with good homogeneity (Figure 1C) indicated that the difficulty of peptide chain assembly was dependent on the starting position. Indeed, when Glu-119 was chosen as the C-terminal residue, even after only 16 couplings a complex mixture of products was cleaved from the resin (product **II.4**, Table 1 and Figure 1D), which could not be further identified by mass spectrometry. The different accessibility of the last two Id2 fragments presumably reflected different aggregation propensity of the peptide chains; in fact, starting from Leu-124 rather than from Glu-119 was probably advantageous because the proline residue located four positions apart from the C-end, Pro-121, might have induced a backbone bending, thus reducing the aggregation of the growing sequence that was predicted to form preferentially β -sheets in the region 103-115 on the base of the Chou and Fasman analysis [5].

Table 1: Amino acid sequence of the human Id2 protein and analytical data of the synthetic Id2 fragments.

Human Id2 ^a				
M ¹ KAFSPVRSVRKNSLSDHSLGISRSKTPVDDPMSL ³⁵ <u>LYNMNDCYSKLKELVP</u> ⁵¹ <u>SIPQNK</u> <u>KVS</u> ⁶⁰ <u>KMEILQHVIDYILDLO</u> ⁷⁶ IALDSHPTIVSLHHQRPQGQNQASRTPLTTLNTDISILSLQA SEFPSELMSNDSKALCG ¹³⁴				
Synthetic Id2 polypeptides				
No	Chain length	MW _{calc.} (Da)	MW _{found} (Da)	t _R (min)
II.1	77-134 ^b	6285.09	n. d.	n. d.
II.2	99-134	3912.44	3915	29.0
II.3	77-124 ^{b, c}	5276.94	5278	28.0
II.4	103-119 ^{b, c}	1860.11	n. d.	n. d.
II.5	61-110 ^{b, c}	5727.59	5729	21.5
II.6	51-110 ^{b, c}	6806.87	6807.8 ^d	19.0
II.7	[S-42]-(36-110) ^{b, c}	8605.96	8605.5 ^d	19.5
II.8	60-76 ^{b, c}	2099.50	2100	21.5
II.9	36-52 ^{b, c}	2058.42	2060	18.7
II.10	36-76 ^{b, c}	4906.85	4907.7 ^d	20.5
II.11	[Q-38, -40, E-41]-(36-76) ^{b, c}	4948.93	4949.7 ^d	21.5
II.12	[Nle-39, -62, S-42]-(36-76) ^{b, c}	4854.71	4857	19.0
II.13	[F-37, -43]-(36-76) ^{b, c}	4874,85	4877	20.99
II.14	[F-37, -71]-(36-76) ^{b, c}	4874,85	4876	21.1
II.15	[F-43, -71]-(36-76) ^{b, c}	4874,85	4875	21.4
II.16	1-35 ^c	3843.48	3845	10.8

^a Underlined sequences: helix-1 (36-51) and helix-2 (61-76); double underlined sequence: loop (52-60). ^b The fragment is N-terminally acetylated. ^c The fragment is C-terminally amidated. ^d Determined by LC-ESI-MS (all others were determined by MALDI-ToF-MS).

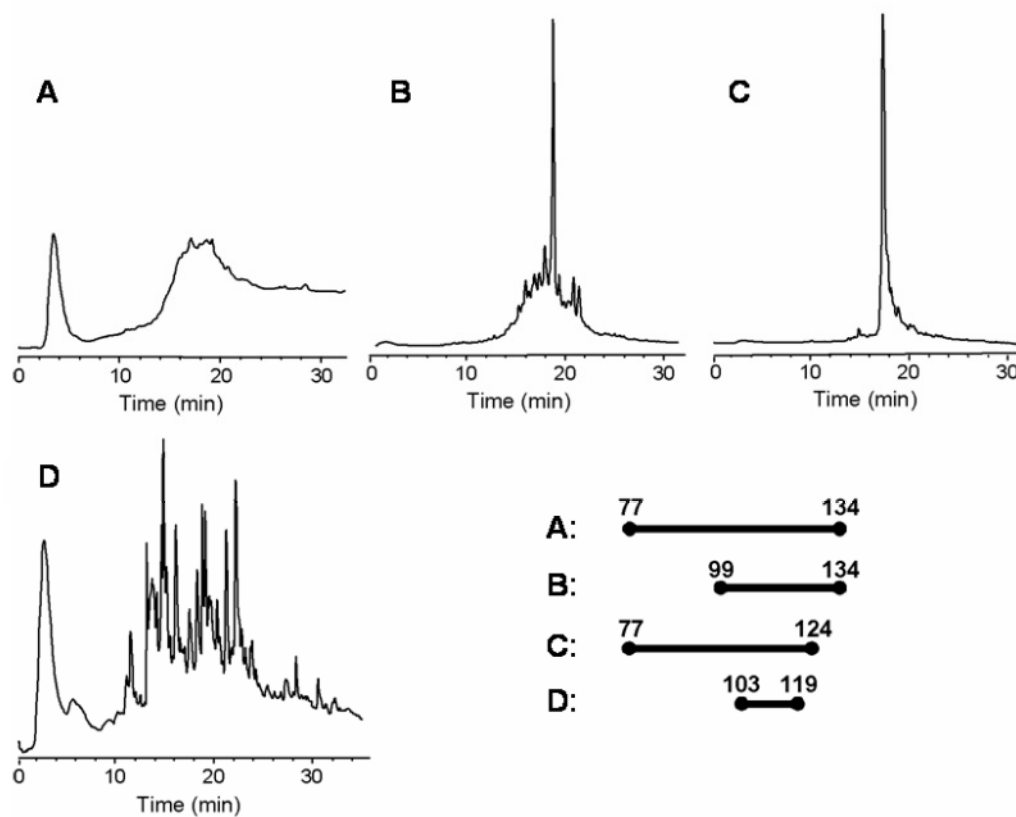


Figure 1: HPLC profiles of crude peptides **II.1** (A), **II.2** (B), **II.3** (C), and **II.4** (D).

It has been reported that solid-phase synthesis of peptides forming aggregates [6] could be improved by using the pseudoproline dipeptides developed by Mutter and co-workers (Figure 2) [7].

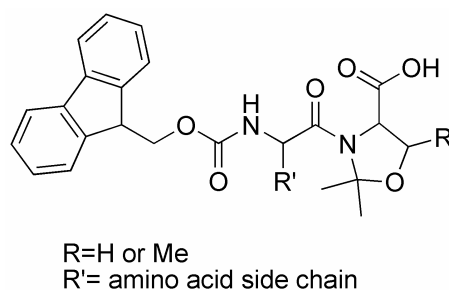


Figure 2: The pseudoproline dipeptide developed by Mutter and co-workers [7].

Pseudoproline dipeptides are masked serine or threonine residues that present some interesting features. First of all they allow the introduction of two residues in one step, and the side chain of serine or threonine can be regenerated by general cleavage conditions with TFA. Moreover, the introduction of a moiety that has a similar structure to proline helps disrupting secondary structure formation as well as aggregation. Additionally, peptides containing C-terminal pseudoproline dipeptide residues can be coupled without any risk of stereomutation. To obtain more effective results by using pseudoproline dipeptides, they must be introduced every four-six residues in the sequence, as it has been successfully done in the case of the human islet amyloid peptide (IAPP) [6].

As segment **II.4** was a suitable candidate for the pseudoproline chemistry due to the presence of several Thr and Ser residues, the synthesis was repeated by introducing the pseudoproline at positions 111 and 114. However, also this approach led to a poorly homogeneous product.

Another strategy that can be applied in the case of difficult sequences containing native serine residues is the depsipeptide strategy (Figure 3) [8, 9]. Insertion of a depsipeptide unit in a sequence interrupts the regular pattern of amide bonds at the site of a Ser/Thr unit, as from that point the peptide chain is extended via the β -hydroxyl function and not via the α -amino function. Each depsipeptide unit provides an additional ionizable moiety, thereby increasing solubility and facilitating purification, as reported for the case of the amyloid peptide A β (1-42) [10]. Conversion of the ester bonds to the target amide bonds is smoothly achieved through an O,N-acyl shift, which occurs quantitatively under mildly basic conditions over a short period of time.

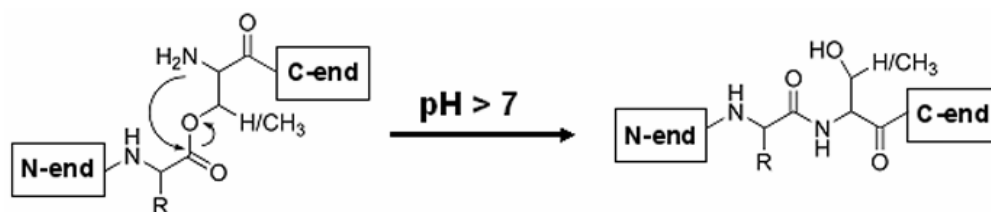
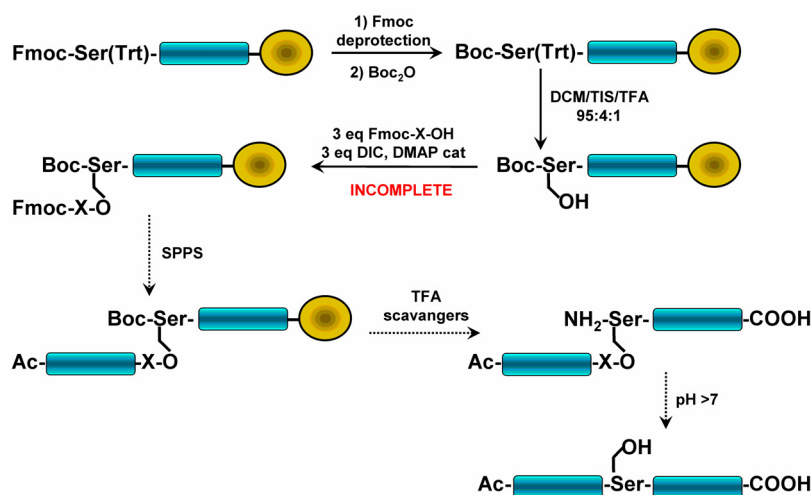


Figure 3: Conversion of a depsipeptide to the amide form through a base-promoted O,N-acyl shift.

This strategy was applied to assemble the Id2 peptide chain starting from residue 134, with the purpose of forming an ester bond between Ser-118 and Ala-117 or between Ser-114 and Leu-113 (Scheme 1). Unfortunately, in both cases the esterification reaction did not run to completion by using the conditions reported in the literature [8]. Therefore, further optimization of the reaction procedure is necessary to reach satisfactory results.

Scheme 1: Synthesis of Id2 C-terminal peptides by the depsipeptide strategy (the yellow sphere represents the Wang resin; the blue rectangles are the growing peptide chains; the dotted arrows represent the reactions that would have brought to the desired product after a successful esterification step).



In contrast to the synthetic difficulty of the Id2 sequence starting from Gly-134 (**II.1**), the Id2 sequence starting from Leu-124 (**II.3**) was chemically accessible by standard solid-phase methodology. However, this peptide was found to be only moderately soluble both in water and phosphate buffer, whereas it could be dissolved in methanol and in mixtures of methanol/water.

Another good starting point for the synthesis of C-terminally truncated Id2 analogs was Ile-110, as shown by the HPLC profiles of the crude products of peptides **II.5-7** (Figure 4). For the synthesis of **II.7** that contained the complete HLH motif with one mutation at position 42 (Cys was replaced with Ser), the pseudoproline dipeptide Asp(OtBu)-Ser($\Psi^{\text{Me,Me}}_{\text{pro}}$) was

coupled instead of the single amino acids Asp-41 and Ser-42, to reduce the risk of chain aggregation during elongation [7].

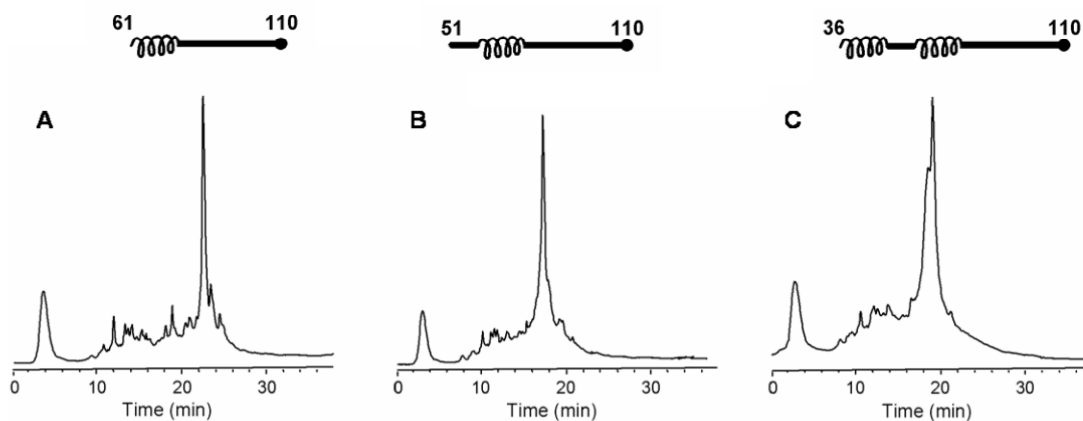


Figure 4: HPLC profiles of crude peptides **II.5** (A), **II.6** (B), and **II.7** (C).

II.2.2 The HLH motif

To investigate the intrinsic conformation of the two helices forming the HLH structural motif of Id2, the sequences 36-52 and 60-76 reproducing helix-1 and helix-2, respectively, were synthesized. Helix-1 (**II.9**) showed a major degree of difficulty in the chain assembly relative to helix-2 (**II.8**), as deduced by the different homogeneity of the corresponding crude products (Figure 5A-B). Also during the synthesis of the entire HLH motif (**II.10**) a consistent amount of byproducts was produced, which accumulated mostly in the region of helix-1 (Figure 5C). LC-ESI-MS analysis indicated the presence of the truncated sequences 43-76 and 42-76, and of the sequence lacking Cys-42. When the last couplings were carried out manually and monitored by the ninhydrin test [11], it was observed that especially Asp-41, Met-39 and Tyr-37 were difficult to be acylated. Longer coupling reactions and introduction of a capping step, however, did not lead to a significant improvement of the quality of the crude product.

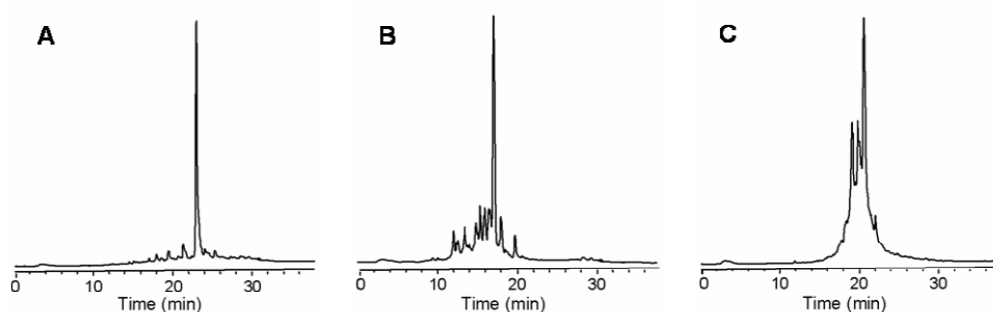


Figure 5: HPLC profiles of peptides **II.8** (A), **II.9** (B), and **II.10** (C)

One side reaction that could have occurred during the synthesis of **II.10** is the formation of aspartimide [12, 13]. In the Fmoc chemistry, this intramolecular cyclization can be catalyzed by the repetitive piperidine treatments needed for the Fmoc removal. Additional byproducts result from racemization of the imide moiety, ring opening by nucleophiles and even breakage of the backbone. Therefore, to exclude the formation of aspartimide, Asn-38, Asn-40 and Asp-41 were substituted by Gln and Glu (**II.11**). However, as the HPLC profiles of the crude peptides **II.10** and **II.11** were almost superimposable, it was concluded that the aspartimide formation, if ever occurred, was a minor event.

A consistent byproduct of the synthesis of the HLH motif was its Met-oxidized species (+16 Da), which resulted to be resistant against reduction with TMSBr/EDT/TFA. Peptide **II.12** containing the replacements Nle→Met-39/-62 and Ser→Cys-42 was prepared to have an HLH analog insensitive to oxidative processes and thus easier to handle, but it was obviously still contaminated by the same truncated and deleted sequences found for the native HLH peptide.

II.3 CD spectroscopy

II.3.1 Peptides containing the Id2 C-terminal region

The CD spectrum of the Id2 C-terminal segment 101-134 in phosphate buffer (100 mM, pH 7.3) was reported [14] and resembled that of a random coil. Unfortunately, in the present work the synthesis of the full-length C-terminus 77-134 by stepwise solid-phase methodology and Fmoc chemistry was not successful, but it was possible to obtain the analogue spanning residues 77-124. Because of the very low solubility both in water and in phosphate buffer, the CD spectra of **II.3** were recorded in methanol and methanol/water mixtures at a peptide concentration of 50 μ M (Figure 6). A low-intensity, α -helix-like curve was recorded in methanol, which was characterized by two negative bands at 223 nm (amide $n-\pi^*$ transition) and 209 nm (amide $\pi-\pi^*$ transition), and by a positive band near 190 nm with a crossover at 200 nm. The ratio R between the ellipticity values at 223 and 209 nm was 0.75. Despite the helix-like shape, this spectrum is remarkably less intense than it is expected for an α -helix, which is probably due to peptide aggregation. Upon water addition, the CD intensity further decreased, and at water percentage $\geq 60\%$ the shape changed from helix-like to random-like, showing a negative Cotton effect close to 200 nm. The absence of an isodichroic point suggests that it was not a two-state transition.

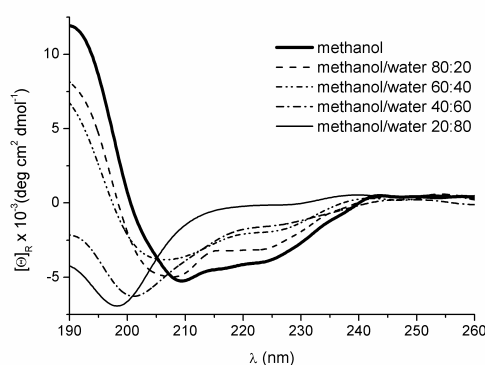


Figure 6: CD spectra of peptide **II.3** at the concentration of 50 μ M in methanol and methanol/water mixtures.

Unlike peptide **II.3**, peptides **II.5-7** were soluble in phosphate buffer (100 mM, pH 7.3), in which they were dissolved at the concentration of 30 μ M for CD analysis. The spectra showed two negative bands at 222 nm (amide $n-\pi^*$ transition) and 204 nm (amide $\pi-\pi^*$ transition) and a positive one near 190 nm, with a crossover at 195-197 nm (Figure 7A). Secondary structure element composition was estimated by using the algorithm Contin [15] (Table 2), which indicated the presence of similar amounts of α -helix and β -sheet structures (the α/β ratios were ~ 1 for **II.5** and **II.7**, and 0.6 for **II.6**). This suggests that these peptides were prone to aggregate, which was also supported by the observation that the CD spectra recorded at the concentrations of 30 and 100 μ M were characterized by different intensities (the more concentrated sample gave a less intense CD signal). The conformation of peptides **II.5-7** was not only concentration-dependent, but also time dependent; indeed, aging of **II.5** and **II.6** favored the α -helix at the expense of the β -sheet structure, whereas aging of **II.7** increased the β -sheet structure at the expense of the α -helix (Figure 7B).

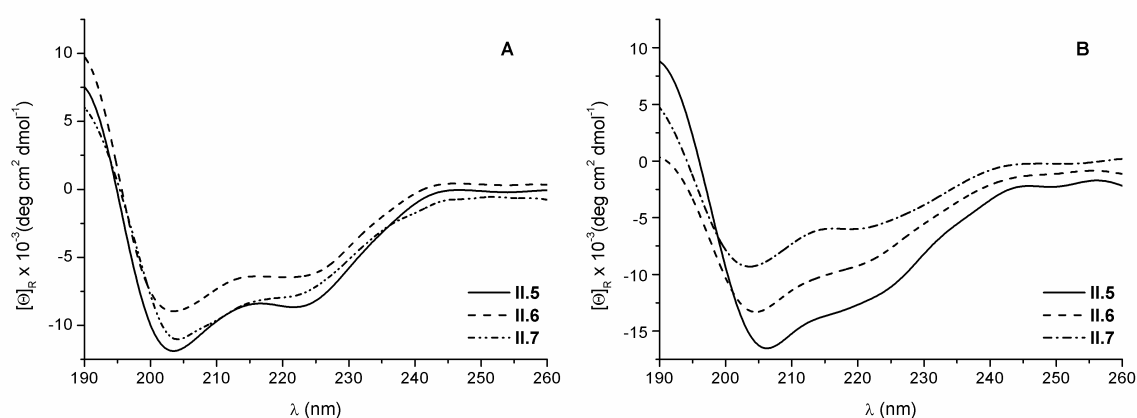


Figure 7: CD spectra of peptides **II.5-7** at the concentration of 30 μ M in phosphate buffer (100 mM, pH 7.3): (A) fresh and (B) one-day-old samples.

The weak helix character of the Id2 fragments **II.5-6** is not very surprising, as they contain only part of the HLH fold (helix-2 with or without the loop); instead, it was surprising to see that peptide **II.7**, although spanning the complete HLH motif, was not able to adopt a stable helix conformation. Of course, modification of helix-1 by replacing Cys-42 with Ser might have affected the stability of the HLH fold; however, this cannot explain the aggregation

propensity that was observed not only for peptide **II.7**, but also for peptides **II.5-6** that lack helix-1. Probably, as all three peptides shared the flanking residues C-terminal to helix-2, which were also present in the poorly soluble and aggregation-prone peptide **II.3**, it is likely that the common tendency of these four Id2 fragments to aggregate arise from such region.

Table 2: Conformational properties of the synthetic Id2 C-terminal fragments

No	Number of residues	<i>R</i> value ^a	Secondary structure element by Contin (%) ^b			
			Helix (Number of residues)	β - sheet	Turns	Un- ordered
II.5	50	0.70	23 (12)	20	23	34
II.6	60	0.72	16 (10)	27	21	36
II.7	75	0.71	21(16)	23	24	32

^a The *R* value is defined as the ratio between the CD intensities of the amide bands $n-\pi^*$ and $\pi-\pi^*$. ^b

The CD spectra of the 30 μ M samples were analyzed by using the online server Dichroweb [16]

II.3.2 Peptides related to the Id2 HLH motif

The Id2 HLH motif 36-76 has been previously investigated by CD spectroscopy in phosphate buffer (100 mM, pH 7.3), and it has been shown to adopt a helical conformation that is stable also in the presence of high concentrations of guanidine hydrochloride (up to 4 M) [14]. Moreover, the CD spectrum was characterized by an *R* value >1 . In this work the Id2 HLH analogue **II.12** was prepared by replacing the sulfur-containing residues Met and Cys with Nle and Ser, respectively. These substitutions have the advantage to avoid problems related with peptide oxidation; moreover, as they are highly conservative, no significant conformational changes were expected. Instead, the CD spectrum of the three-point mutated HLH sequence was found to be less intense than that of the native motif, and the amide $\pi-\pi^*$ band was slightly blue-shifted. In addition, the *R* value became <1 . Based

on the CD data reported in Figure 8, helical contents of 86% for the native and of 48% for the mutated HLH peptide were estimated with the algorithm Contin [15] (Table 3). These results suggest that the Cys and/or Met side chains are important for the HLH fold, and even subtle substitutions are badly tolerated. Thus, peptide **II.12** is not a suitable candidate to be used in place of the native sequence for further conformational analyses.

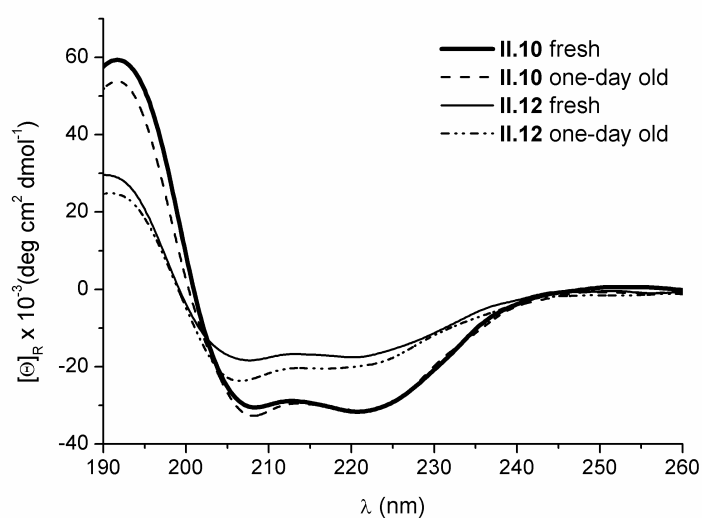


Figure 8: CD spectra of peptides **II.10** and **II.12** at the concentration of 30 μM in phosphate buffer (100 mM, pH 7.3), as fresh and one-day old samples.

Table 3: Conformational properties of the synthetic peptides related to the Id2 HLH region

No	Number of residues	<i>R</i> value ^a	Secondary structure element by Contin (%) ^b			
			Helix (Number of residues)	β - sheet	Turns	Un- ordered
II.10	41	1.3	86 (35)	0	8	6
II.12	41	0.96	48 (20)	9	21	22

^a The *R* value is defined as the ratio between the CD intensities of the amide bands $n-\pi^*$ and $\pi-\pi^*$.^b

The CD spectra of the 30 μ M samples were analyzed by using the online server Dichroweb [16].

In contrast to the one-point mutated HLH sequences conjugated to the Id2 C-terminal region **II.7**, the isolated HLH motifs of peptides **II.10** and **II.12** were found to be stable upon aging, with only a moderate increase in the intensity of the CD band corresponding to the amide $\pi-\pi^*$ transition (Figure 8). This observation again underlines the negative impact of the C-terminal extension on the folding and stability of the adjacent HLH region.

The Id2 HLH motif contains three tyrosine residues, two of them, Tyr-43 and Tyr-71, are conserved within the Id family, whereas the third one, Tyr-37, is found only in Id1 and Id2. In order to investigate the importance of these positions for the HLH fold, analogues of the native sequence **II.10** were synthesized (**II.13-15**), each containing two Phe→Tyr mutations. A significant reduction of the intensity and of the *R* value as well as a slight blue shift could be observed for the CD curves of the mutated peptides in comparison to the one of the native sequence **II.10** (Figure 9). In particular, Tyr-43 and Tyr-71 were found to be structurally more important than Tyr-37, as indicated by the higher loss of CD intensity for peptides **II.14-15** than for peptide **II.13** (Figure 9). These data reveal that the tyrosine side chains are likely to be involved not only in hydrophobic contacts through the phenyl ring but also in hydrogen bonds through the OH group.

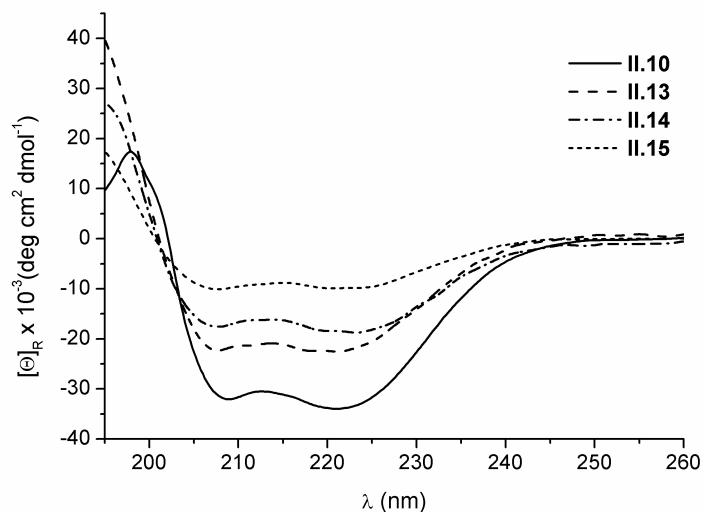


Figure 9: CD spectra of peptides **II.10**, **II.13**, **II.14** and **II.15** at the concentration of 100 μ M in phosphate buffer (100 mM, pH 7.3)

Further, to investigate whether both the N- and C-terminal helices of the HLH motif displayed similar intrinsic helix propensity, the CD spectra of the corresponding peptides **II.9** (helix-1) and **II.8** (helix-2) were recorded in phosphate buffer. The presence of an ordered conformation was found for helix-2 (**II.8**), whereas a disordered structure was found for helix-1 (**II.9**) (Figure 10 and Table 4). This suggests that the intrinsic helix propensity of helix-2 is much higher than that of helix-1. However, the conformational stability of helix-2 was concentration dependent, suggesting that a favorable intermolecular helix-packing, presumably through their hydrophobic faces, occurs by increasing the peptide concentration.

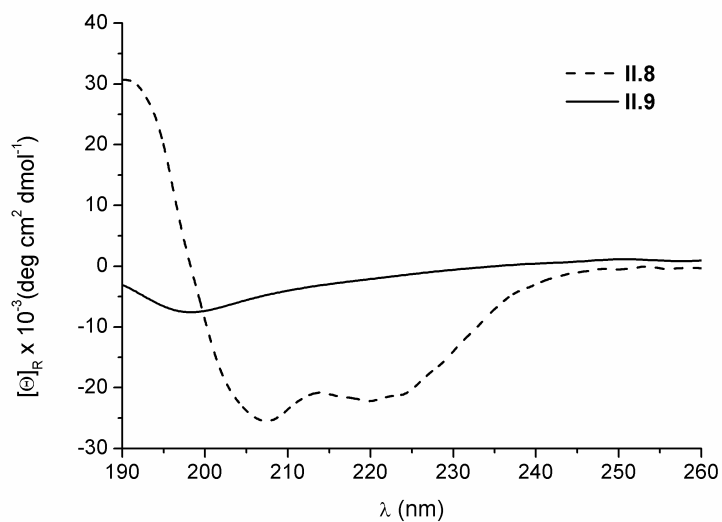


Figure 10: CD spectra of peptides **II.8** (220 μ M) and of peptide **II.9** (180 μ M) in phosphate buffer (100 mM, pH 7.3).

Table 4: Conformational properties of the synthetic Id2 helix-1 (**II.9**) and helix-2 (**II.8**).

No	Number of residues	<i>R</i> value ^a	Secondary structure element by Contin (%) ^b			
			Helix (Number of residues)	β - sheet	Turns	Un- ordered
II.8	17	0.87	61 (10)	6	1	32
II.9	17	0.3	0	6	10	84

^a The *R* value is defined as the ratio between the CD intensities of the amide bands $n-\pi^*$ and $\pi-\pi^*$. ^b

The CD spectra were analyzed by using the online server Dichroweb [16].

II.3.3 Noncovalent interactions between the Id2 N-terminus and the HLH containing peptides.

The Id2 N-terminal fragment 1-35 (**II.16**) has been shown to adopt an unordered conformation in phosphate buffer [14]. However, when it was mixed with the native HLH motif (**II.10**) in a 1:1 ratio, the CD spectrum was characteristic of an α -helix and the intensity of the two minima at 222 and 206 nm was higher than the arithmetic sum of the intensities of the two separated fragments (Figure 11A). This indicates that an interaction of the N-terminus with the HLH fold could be detected, which apparently led to helix stabilization in the complex and was maintained also upon aging.

When the experiment described above was carried out using the segment 36-110 (**II.7**) instead of the HLH motif (**II.10**), the CD spectrum of the mixture was less intense than the sum of the spectra of the two separated fragments (Figure 11B), suggesting an increase in β -sheet structure at the expense of α -helix upon mixing. However, such changes seemed to be reversible, as the CD spectrum of the one-day-old mixture gained intensity and was close to the sum of the CD spectra of the two individual components (Figure 11C). Presumably, the increased tendency of peptide **II.7** to self-aggregate became predominant, thus reducing the possibility of interaction with the N-terminus. The different effect of the N-terminus on the HLH motif alone and conjugated to the C-terminal tail 77-110 might reflect the different stability of the HLH fold; indeed, only in the former case the HLH domain was well-structured, whereas it was not well-defined in the latter case, as it was discussed above about the conformational properties of peptides **II.10** and **II.7**. This would imply a specific interaction pattern of the N-terminus with the folded HLH region, whereas the interaction would be mostly unspecific in the case of a partially unfolded HLH motif.

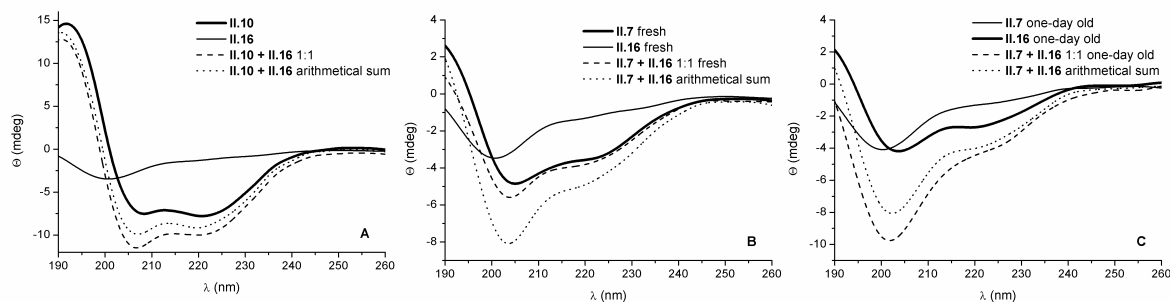


Figure 11: CD spectra of equimolar mixtures of the N-terminal sequence (II.16) with (A) the native HLH motif (II.10) and (B-C) the long fragment II.7. Each component in the mixture was at the concentration of 30 μ M. The CD spectra of the individual peptides (30 μ M) and their arithmetical sum are also shown.

II.4 Native chemical ligation approach for the synthesis of Id2 large fragments

Native chemical ligation (NCL) [4] is a powerful tool in protein chemistry to prepare proteins of small to medium size by coupling unprotected peptides through an amide bond, usually under aqueous buffer conditions. The amide bond can be formed when the C-terminal fragment possesses an N-terminal cysteine, and the N-terminal fragment is synthesized as C-terminal thioester (Figure 12). The ligation consists of a chemoselective capture between the thioester and the thiol function of the cysteine, followed by a spontaneous intramolecular rearrangement to give the desired amide bond.

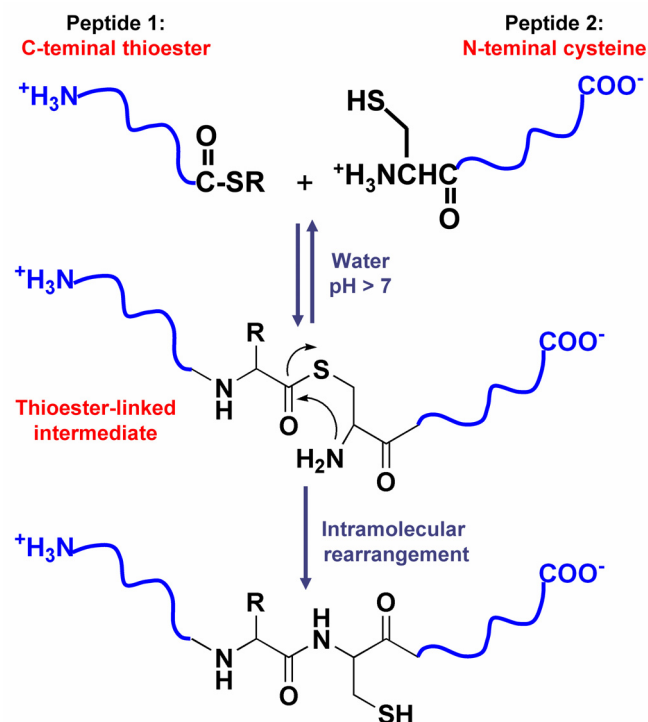


Figure 12: A schematic representation of native chemical ligation [4] between two unprotected peptides in aqueous buffer solution.

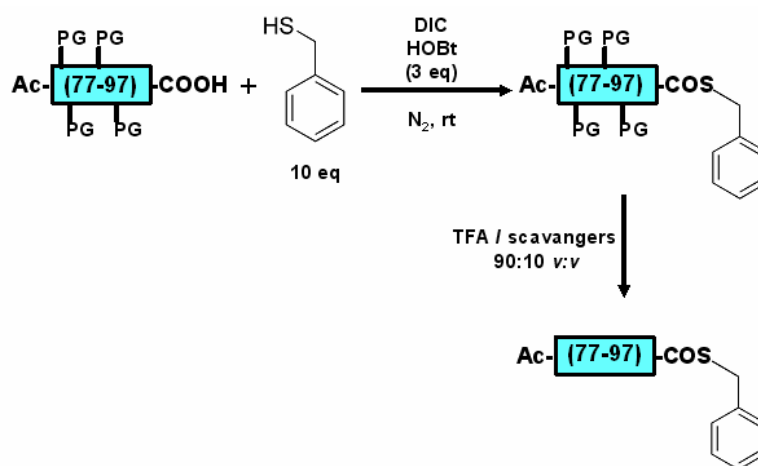
Because of the difficulties encountered during the syntheses of large Id2 analogues discussed above, the possibility to use the native chemical ligation was considered. A first trial was focused on the full-length C-terminus 77-134. As the only natural cysteine residues in this region is Cys-133 preceding the C-end, the introduction of an unnatural cysteine was necessary. Thus, the N-terminal cysteine-containing peptide component for NCL was [Cys-98]-(98-134)-Id2 (**II.17**), in which the native Ala-98 was substituted with Cys. The synthesis was performed by stepwise solid-phase methodology starting from the Wang resin, and the obtained crude product was purified by preparative HPLC (Table 5).

Table 5: Analytical data of the Id2 peptide analogues for NCL.

No	Chain length	MW _{calc} (Da)	MW _{found} (Da)	t _R (min)
II.17	[C-98]-(98-134)	3944.5	3947.8	19.6
II.18	Ac-(77-97)-SBn	2492	2495	11.1

Instead, the peptide thioester containing the Id2 sequence 77-97 was obtained in three steps (Scheme 2): first, the fully-protected peptide acid was synthesized by stepwise solid-phase methodology starting from 2-chloro-tritylchloride resin, and then it was converted in solution into a C-terminal thioester that was finally fully-deprotected to give **II.18** [17].

Scheme 2: Synthesis of the peptide thioester **II.18** (PG: protecting group).



Due to the high homogeneity of the product after the last step (Figure 13 and Table 5), it could be directly used for the NCL experiment, without any further purification.

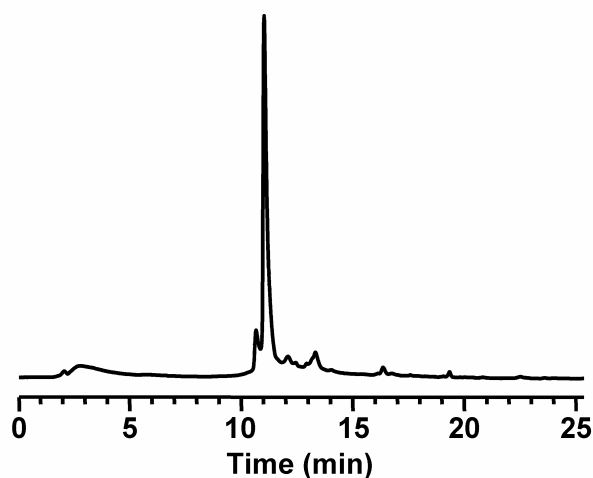


Figure 13: HPLC profile of the crude peptide thioester **II.18**.

Peptides **II.17-18** were dissolved in phosphate buffer (100 mM, pH 7.3) containing TFE to increase the peptide solubility and guanidinium chloride (6 M) as denaturing agent. Sodium thiophenolate was used in catalytic amount to promote the transthioesterification. The reaction mixture was stirred at room temperature and monitored by HPLC. Unfortunately, the formation of the ligation product was found to run slowly and with unsatisfactory yields. Presumably, the aggregation propensity generally observed for sequences related to the Id2 C-terminal domain compromised the efficiency of the ligation.

II.5 Conclusions

The Id2 protein is believed to be a promising target for tumor therapy, especially for neuroblastoma, where it has been shown to inhibit the function of pRb [18]. Therefore, both biochemical and conformational studies of Id2 are important to understand the mechanism of action and to develop chemical tools to block the Id2-pRb interaction. Synthetic peptides derived from amino acid replacement and N-/C-end truncation of Id2 are useful to study the importance of different parts of the sequence on the conformation. Here, we have reported on a series of Id2 analogues with variable length, which were prepared by stepwise SPPS

using Fmoc chemistry. The synthesis of the C-terminal domain was found to be difficult and to strongly vary with the truncation point. Moreover, a general poor solubility under physiological conditions as well as a propensity to self-aggregate was observed for Id2 fragments containing part of the C-terminus.

The Id2 HLH region was found to be very sensitive to amino acid substitutions: even conservative mutations like Nle/Met, Ser/Cys or Phe/Tyr were not-well tolerated. Also the presence of the domain immediately following the C-terminal helix-2 seemed to strongly alter the conformational properties of the HLH fold.

All together, these results suggest that the three main subdomains of Id2, the N-terminus, the HLH motif and the C-terminus, play a role during protein folding and modulate the dimerization profile of Id2 with the related bHLH factors and the members of the Rb family.

II.6 Literature

1. Colombo, N., Cabrele, C. (2006) Synthesis and conformational analysis of Id2 protein fragments: impact of chain length and point mutations on the structural HLH motif, *J Pept Sci.*, 12, 550-558.
2. Fong, S., Debs, R. J., Desprez, P. Y. (2004) Id genes and proteins as promising targets in cancer therapy, *Trends. Mol. Med.*, 10, 387-392.
3. Lasorella, A., Boldrini, R., Dominici, C., Donfrancesco, A., Yokota, Y., Inserra, A., Iavarone, A. (2002) Id2 is critical for cellular proliferation and is the oncogenic effector of N-myc in human neuroblastoma, *Cancer Res.*, 62, 301-306.
4. Dawson, P. E., Muir, T. W., Clark-Lewis, I., Kent, S. B. (1994) Synthesis of proteins by native chemical ligation, *Science*, 266, 776-779.
5. Chou, P. Y., Fasman, G. D. (1974) Prediction of protein conformation, *Biochemistry*, 13, 222-245.
6. Abedini, A., Raleigh, D. P. (2005) Incorporation of pseudoproline derivatives allows the facile synthesis of human IAPP, a highly amyloidogenic and aggregation-prone polypeptide, *Org Lett.*, 7, 693-696.
7. Mutter, M., Nefzi, A., Sato, T., Sun, X., Wahl, F., Wohr, T. (1995) Pseudo-prolines (Ψ Pro) for accessing "inaccessible" peptides, *Pept. Res.*, 8, 145-153.
8. Mutter, M., Chandravarkar, A., Boyat, C., Lopez, J., Dos Santos, S., Mandal, B., Mimna, R., Murat, K., Patiny, L., Saucedo, L., Tuchscherer, G. (2004) Switch peptides in statu nascendi: induction of conformational transitions relevant to degenerative diseases, *Angew. Chem. Int. Ed. Engl.*, 43, 4172-4178.
9. Coin, I., Dolling, R., Krause, E., Bienert, M., Beyermann, M., Sferdean, C. D., Carpino, L. A. (2006) Depsipeptide methodology for solid-phase peptide synthesis: circumventing side reactions and development of an automated technique via depsidipeptide units, *J. Org. Chem.*, 71, 6171-6177.
10. Taniguchi, A., Sohma, Y., Kimura, M., Okada, T., Ikeda, K., Hayashi, Y., Kimura, T., Hirota, S., Matsuzaki, K., Kiso, Y. (2006) "Click peptide" based on the "o-acyl isopeptide method": control of A beta1-42 production from a photo-triggered A beta1-42 analogue, *J. Am. Chem. Soc.*, 128, 696-697.

11. Kaiser, E., Colescott, R. L., Bossinger, C. D., Cook, P. I. (1970) Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides, *Anal. Biochem.*, *34*, 595-598.
12. Mergler, M., Dick, F., Sax, B., Weiler, P., Vorherr, T. (2003) The aspartimide problem in Fmoc-based SPPS. Part I, *J. Pept. Sci.*, *9*, 36-46.
13. Mergler, M., Dick, F., Sax, B., Stahelin, C., Vorherr, T. (2003) The aspartimide problem in Fmoc-based SPPS. Part II, *J. Pept. Sci.*, *9*, 518-526.
14. Kiewitz, S. D., Cabrele, C. (2005) Synthesis and conformational properties of protein fragments based on the Id family of DNA-binding and cell-differentiation inhibitors, *Biopolymer.*, *80*, 762-774.
15. Provencher, S. W., Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism, *Biochemistry*, *20*, 33-37.
16. Whitmore, L., Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res.*, *32*, 668-673.
17. von Eggelkraut-Gottanka, R., Klose, A., Beck-Sickinger, A. G., Beyermann, M. (2003) Peptide thioester formation using standard Fmoc-chemistry, *Tetrahedron Letter.*, *44*, 3551-3554.
18. Lasorella, A., Nosedà, M., Beyna, M., Yokota, Y., Iavarone, A. (2000) Id2 is a retinoblastoma protein target and mediates signalling by Myc oncoproteins, *Nature*, *407*, 592-598.

III. A short Id2 protein fragment containing the nuclear export signal forms amyloid-like fibrils [1]

III.1 Introduction

Whereas the HLH motif in all four Id proteins seems to have predominantly a structural function, the role of the N- and C-terminal regions is not yet completely understood: it is known that phosphorylation of Ser-5 modulates the inhibitory activity of Id2 and Id3 [2, 3], and that the N-terminus of Id2 is necessary for ubiquitination [4]. Moreover, the N-terminus of Id2 can induce apoptosis independently of the HLH dimerization motif [5]. Recently, Kurooka and Yokota have shown that the HLH region of Id2 is required for the protein nuclear localization, whereas residues 103-119 in the central part of the C-terminus are necessary for the localization of the protein in the cytoplasm [6]. They also demonstrated that the isolated 17-residue long Id2 fragment displays nuclear export activity when conjugated to other proteins. Indeed, the sequence 106-115 has been identified as nuclear export signal (NES), in which Leu-106, Ile-110, Leu-113, and Leu-115 are the key residues. An additional NES sequence is present in the helix-2 (residues 65-75): this is conserved among the Id family members but is not functionally active in Id2 [6].

While investigating synthetic peptides related to the C-terminal domain of the Id2 protein, we found that the fragment 99-134 was soluble and unstructured in phosphate buffer, whereas the fragment 77-124 was insoluble under the same conditions (Figure 1) [7, 8]. Apparently, the different behavior of these two peptides, which share the same NES segment, is dictated by the neighboring regions. However, as the isolated Id2 NES is functional, we were wondering whether it adopts a preferred conformation or is rather flexible. Thus, we have undertaken the following study to elucidate the structural features of short synthetic peptides based on the Id2 NES sequence.

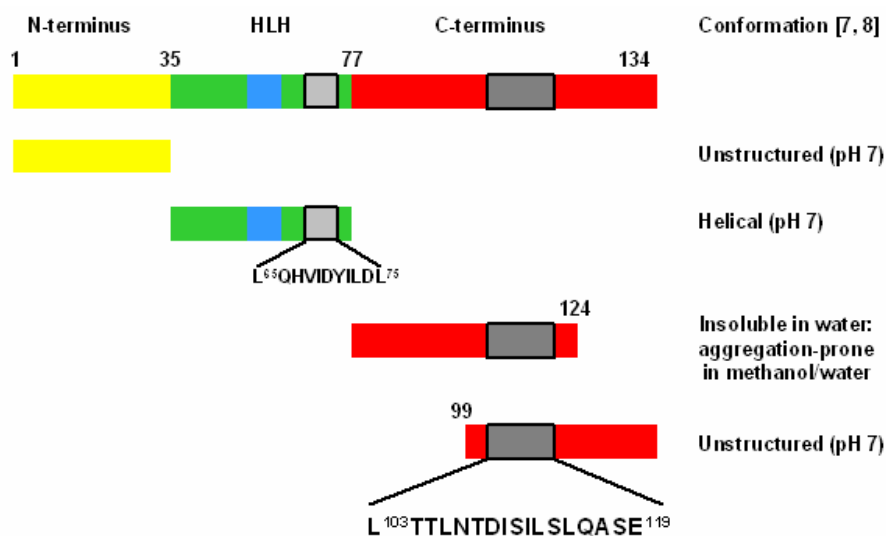


Figure1: Schematic representation of the full-length Id2 protein and of the synthetic fragments reproducing the N-terminus, the HLH motif, and parts of the C-terminus. The gray bars represent the two NES sequences; the motif 103-119 is required for the Id2 nucleo-cytoplasmic transport, whereas the motif 65-75 is inactive. The key residues of the C-terminal NES are underlined.

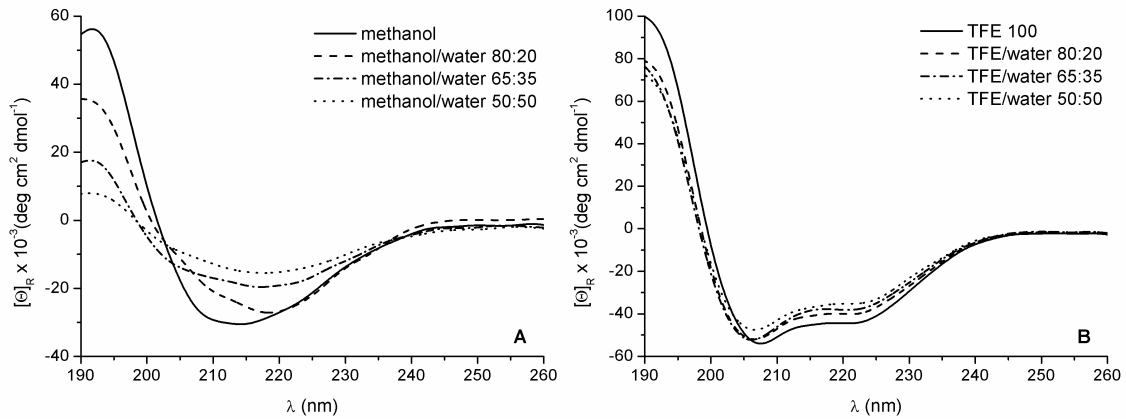
III.2 Synthesis and conformational characterization of Id2 analogues containing the C-terminal NES sequence

Kurooka and Yokota have identified the active NES of Id2 in the fragment 103-119 [6]. However, Glu-119 is not a good starting point for the synthesis of Id2 analogues, as already discussed in chapter II [7], and it was thus chosen to study the longer fragment 103-124 (peptide **III.1**, Table 1). This was found to be easily accessible by SPPS, but displayed the disadvantage of being insoluble in water. For this reason, the CD spectra of **III.1** were recorded in methanol or TFE and upon addition of up to 50% water (at higher water percentages there was instantaneous peptide precipitation). In methanol, the NES sequence adopted a β -sheet structure, as indicated by the negative CD band at 214 nm and the positive one at 193 nm, with a crossover at 202 nm (Figure 2A). Upon water addition, there was a significant decrease in the CD intensity, which should be attributable to the formation of oligomers stabilized by hydrophobic interactions. In contrast, a helix conformation was induced in TFE, which remained stable upon addition of up to 50% water (Figure 2B).

Table 1: Synthetic peptides containing the functionally active NES region of Id2

No.	Sequence ^a	MW _{calc} (Da)	MW _{found} (Da) ^b	t _R (min) ^c
III.1	L ¹⁰³ TTLNTDISILSLQASEFPSEL ¹²⁴	2432	2343	24.0
III.2	KKKX-L ¹⁰³ TTLNTDISILSLQASEFPSEL ¹²⁴	2930	2931	21.5
III.3	L ¹⁰³ TTLNTDISILSLQASEFPSEL ¹²⁴ -XKKK	2930	2932	21.1
III.4	XKKK-L ¹⁰³ TTLNTDISILSLQASEFPSEL ¹²⁴	2930	2935	17.2
III.5	L ¹⁰³ TTLNTDISILSLQASEFPSEL ¹²⁴ -KKKX	2930	2933	18.8

^a The sequences are N-terminally acetylated and C-terminally amidated (X = Ahx). ^b The masses [M+H]⁺ were measured by MALDI-ToF-MS. ^c The retention times t_R refer to the HPLC runs on a reverse-phase C18 column, with the binary system (A) 0.012% TFA in water and (B) 0.01% TFA in acetonitrile (gradient: 10-70% B over 40 min).

**Figure 2:** CD spectra of peptide **III.1** at the concentration of 50 μM (A) in methanol/water or (B) in TFE/water.

In order to improve the water solubility of peptide **III.1** or at least to slow down its precipitation from alcohol/water, we decided to modify the sequence by increasing the number of positively charged residues, an approach that has been successful to stabilize amyloidogenic peptides in solution like PrP 174-195 [9]. As such modification should obviously not affect the conformation significantly, we chose the motif (Lys)₃-Ahx to be coupled to the N-terminal residue 103 (peptide **III.2**, Table 1), or to the C-terminal residue 124 (peptide **III.3**, Table 1). Besides three ammonium ions, this motif displays the 6-

aminohexanoyl unit (Ahx) as a spacer between the basic moiety and the NES region, which should also reduce the risk of influencing the conformation of the latter. Although both analogues were still insoluble in water, they could be dissolved in methanol/water or TFE/water mixtures containing up to 80% water. Moreover, the results of the turbidity assays showed that the solution form of peptide **III.3** was much more stable than that of peptide **III.2** (Figure 3).

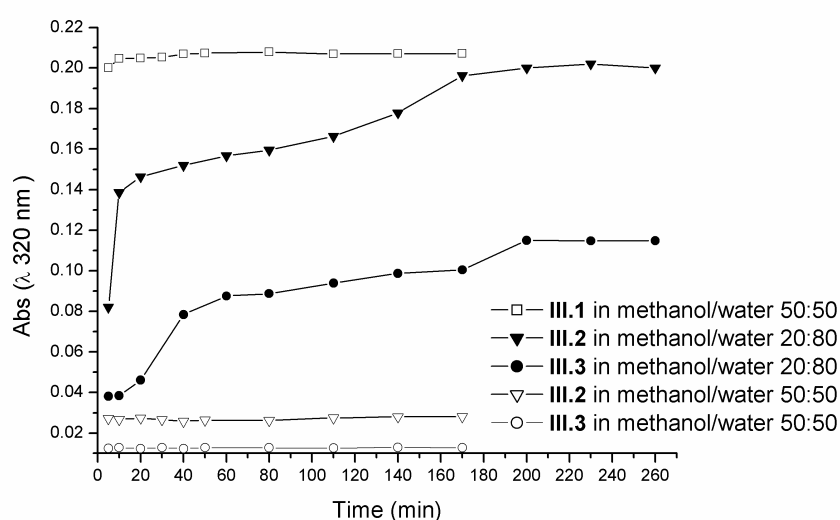


Figure 3: Time-dependent appearance of turbidity in solutions of peptides **III.1-3** (50 μ M in methanol/water mixtures).

The effect of N- or C-terminal modification of peptide **III.1** on its conformation was investigated by CD spectroscopy. In methanol peptide **III.2** was characterized by a very intense β -sheet-like CD curve (Figure 4A). It is plausible to assume that the β -strands were in an antiparallel arrangement, which would allow favorable electrostatic interactions involving the basic N-terminal lysine residues and the C-terminal glutamic acid residues (Glu-119/-123) (Figure 5). Additionally, aromatic-cation interactions between Phe-120 and the lysine side chains might also have a stabilizing effect [10]. As observed above for

peptide **III.1**, the intensity of the CD spectra of **III.2** decreased in methanol/water mixtures; however, unlike **III.1**, **III.2** was probably displaying a helical intermediate at the methanol/water ratios 80:20 and 60:40, as indicated by the appearance of a negative band close to 207 nm. This band disappeared at lower methanol/water ratios, suggesting that the helix conformer has low stability, probably because of unfavorable contacts between the poles of the helix dipole and the positively or negatively charged residues adjacent to the N- and C-ends, respectively. In contrast, increasing the water content in the solvent mixture favored the formation of β -sheet aggregates stabilized by hydrophobic interactions. Importantly, the CD spectrum of peptide **III.2** in methanol/water 40:60 or 20:80 was similar in shape and intensity to that of peptide **III.1** in methanol, indicating that the analogue **III.2** was structurally close to the native fragment 103-124, but it had the advantage of requiring smaller amounts of organic solvent to be dissolved and remain in solution for a long period. Also in TFE peptide **III.2** behaved like peptide **III.1**, building an α -helix that was stable upon addition of up to 60% water (Figure 4B). At higher water percentage, peptide **III.2** underwent a conformational change from α -helix to β -sheet, which was not observed for peptide **III.1** (Figure 2B), thus showing its strong tendency to adopt a β -sheet structure even in the presence of a well-known α -helix-stabilizing solvent like TFE.

Surprisingly, when the motif (Lys)₃-Ahx was conjugated to the C-terminus of the sequence 103-124, the obtained peptide **III.3** (Table 1) adopted a helical conformation both in methanol and in TFE, which was partially destabilized upon water addition (Figure 4C-D). The helix preference of **III.3** might reflect a favorable interaction between the negative pole of the helix dipole and the positively charged residues at the C-end. Thus, modifying the NES region at the C-terminus was structurally not equivalent to modifying it at the N-terminus, as only in the latter case the peptide was allowed to form a β -sheet structure. One reason might be the generation of unfavorable electrostatic interactions upon peptide-chain alignment, thus preventing the formation of β -sheets.

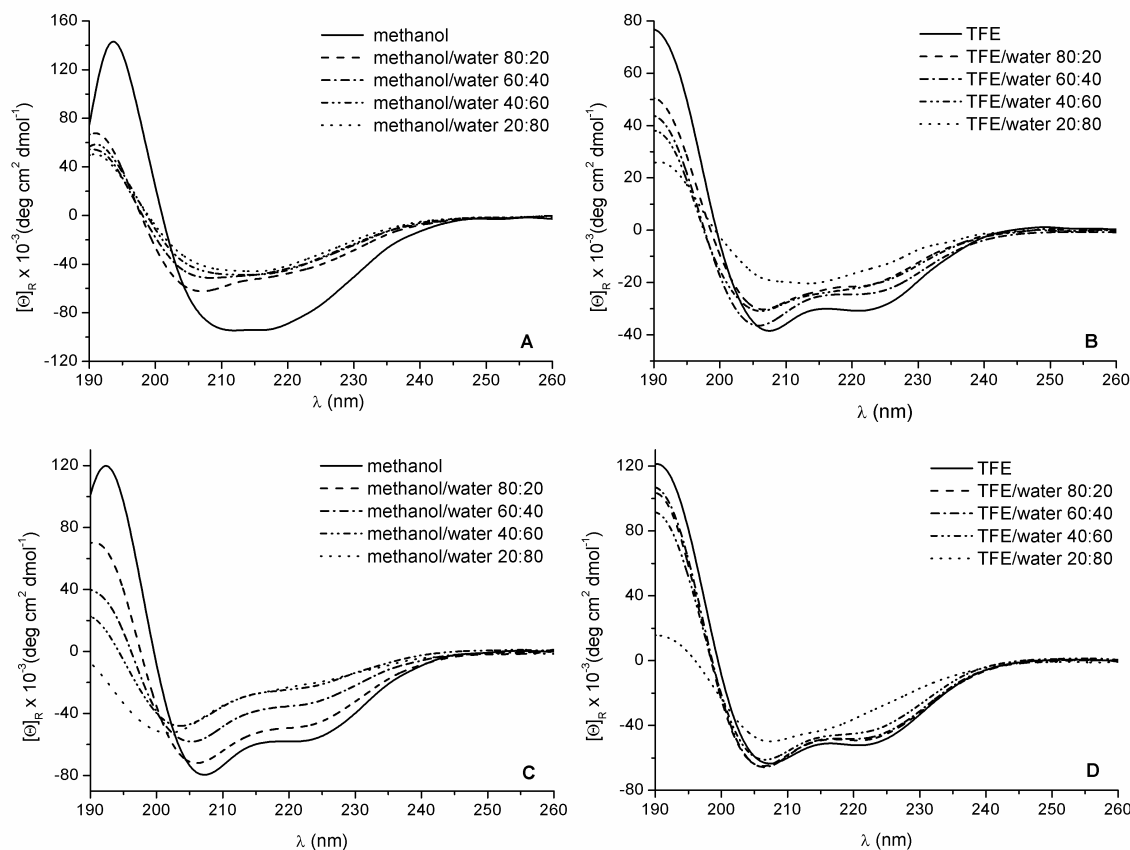


Figure 4: CD spectra of peptide **III.2** at the concentration of 50 μM (A) in methanol/water and (B) in TFE/water, and of peptide **III.3** at the concentration of 50 μM (C) in methanol/water and (D) in TFE/water.

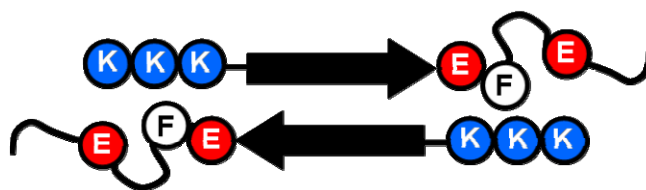


Figure 5: Proposed antiparallel alignment of β -strands of **III.2**.

The possibility that peptides **III.2** and **III.3** could interact in solution was investigated by CD spectroscopy, in methanol and in water with 20% methanol (Figure 6A-B). As already described above (Figure 4A), in both cases **III.2** alone shows a β -sheet structure, whereas **III.3** displays a stable α -helix in methanol which undergoes a conformational transition to 3_{10} -helix [11] upon water addition (Figure 4C). The equimolar mixture of the Short Id2 C-terminal peptide forms amyloid-like fibrils

two peptides in both conditions was found to behave more like **III.3** alone than **III.2** alone, showing a well-defined helix character, especially in methanol. Moreover, the CD curve of the mixture was not only characteristic of a helix but was also more intense than the curve expected in the case of an equimolar mixture of non-interacting species in the region between 202 and 230 nm.

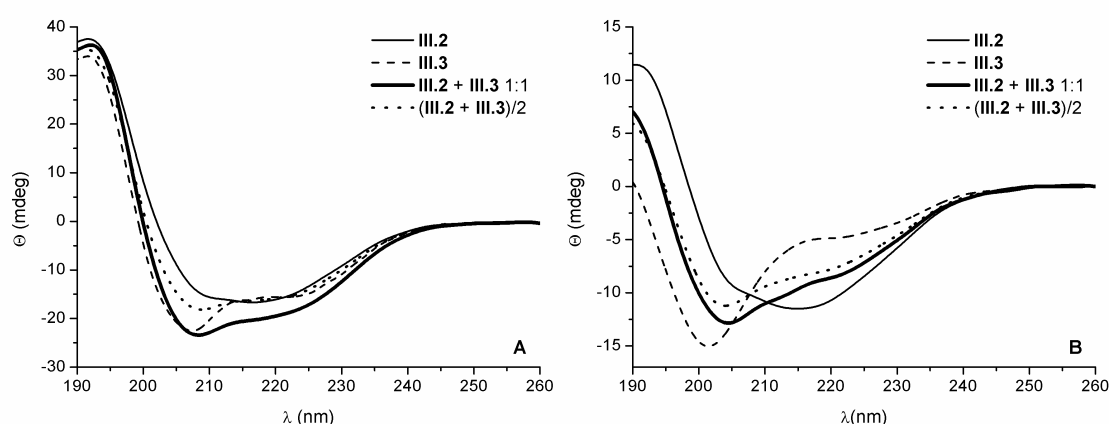
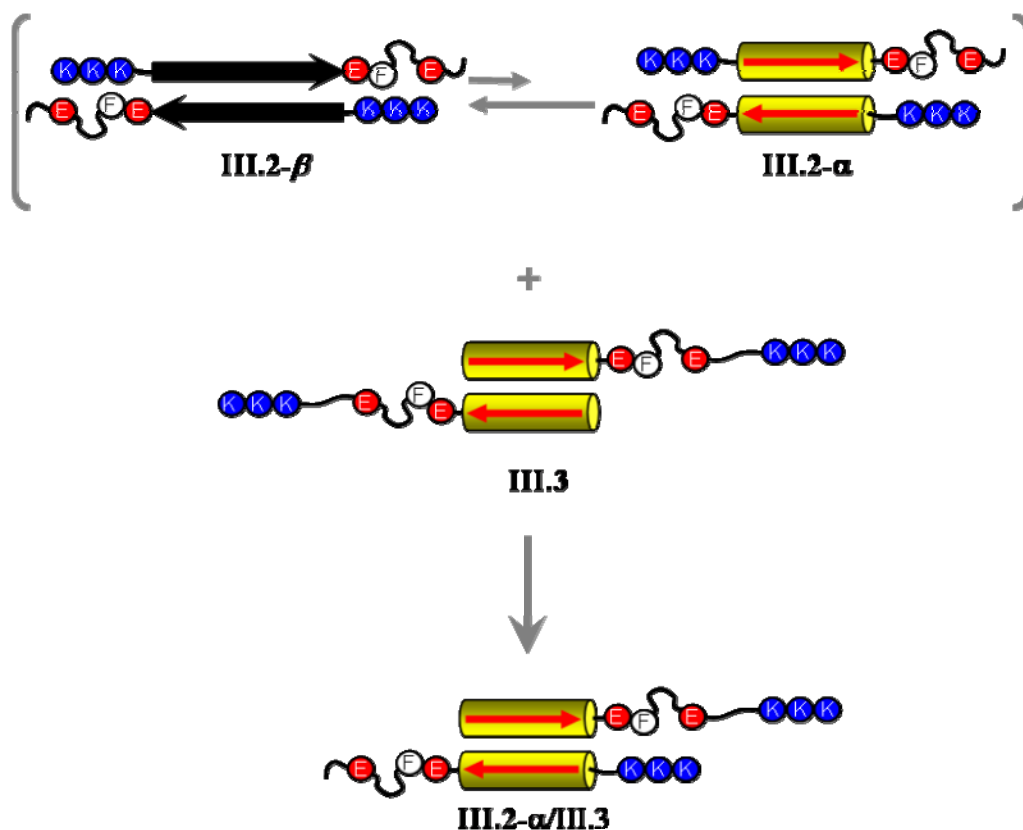


Figure 6: CD spectra of peptides **III.2** and **III.3**, and of their equimolar mixture (A) in methanol and (B) in methanol/water 20:80. In each sample the overall peptide concentration was 50 μ M.

The results of this experiment suggest that an interaction between the α - 3_{10} -helical peptide **III.3** and the β -sheet peptide **III.2** can occur, and that the hetero-association not only induces a helical conformation at the expense of the β -sheet one, but also improves the helicity in comparison to the self-association (Scheme 1).

Scheme 1: Suggested mechanism of interaction between the β -sheet peptide **III.2** and the α -helix peptide **III.3** (the black thick arrow represents a β -strand, whereas the yellow cylinder represents a α -helix. The red arrow represents the helix dipole. Blue, red and white spheres represent lysine, glutamic acid and phenylalanine, respectively). Although an antiparallel helix packing is usually stabilized by a favorable interaction of the helix dipoles, a parallel helix packing can not be excluded.



To investigate the role of the spacer moiety connecting the three lysine residues with the target peptide, we also synthesized two analogues of **III.2** and **III.3**, in which the Ahx unit was shifted before or after the lysine residues, respectively (**III.4** and **III.5**, Table 1). Both peptides showed helix propensity that was higher for peptide **III.5** than for peptide **III.4**, thus indicating that positively charged residues improved the helix stability more efficiently while located at the C-end rather than at the N-end (Figure 7). However, these results also show that the presence of a spacer between the lysine residues and the target

peptide chain can reduce the impact of the lysine tag on the structural properties of the target peptide chain.

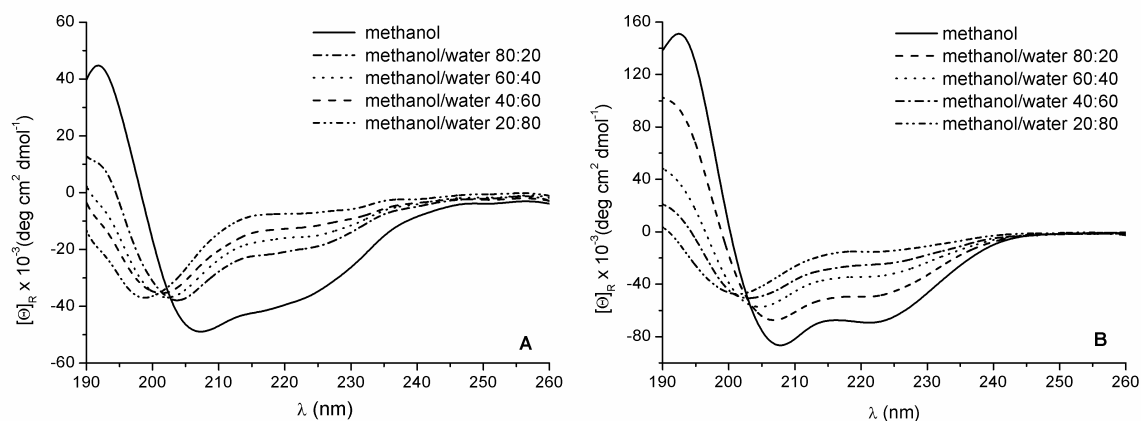


Figure 7: CD spectra of (A) **III.4** and (B) **III.5** at the concentration of 50 μ M in methanol/water.

Based on the CD study of peptides **III.1-5**, we conclude that the isolated Id2 fragment containing the NES domain can form stable β -sheets or α -helices depending on the solvent. Among the lysine-tagged analogues, only the N-terminally modified one containing a spacer between the lysine-tag and the peptide chain (**III.2**) presented conformational properties very similar to those of the unmodified peptide, whereas the N-terminally modified one lacking the spacer (**III.4**) as well as the C-terminally modified ones (**III.3** and **III.5**) lost the ability to form a β -sheet structure.

III.3 Synthesis and conformational characterization of Id2 analogues containing the NES sequence in the helix-2

A putative NES sequence spans a leucine-rich region in the helix-2. The helix-2 of Id2 (peptide **II.8**) has been shown to be easy to synthesize and to adopt an α -helical conformation depending on the peptide concentration. [7]. We have synthesized helix-2

peptide analogues containing the motif (Lys)₃-Ahx coupled to the N-terminal residue 60 (peptide **III.6**, Table 2), or to the C-terminal residue 76 (peptide **III.7**, Table 2) to investigate whether this motif can affect the structural features of a helix preferring sequence.

Table 2: Synthetic peptides containing the putative NES sequence in the helix-2 of Id2.

No.	Sequence ^a	MW _{calc} (Da)	MW _{found} (Da) ^b	t _R (min) ^c
II.8	S ⁶⁰ KMEILQHVIDYILDLQ ⁷⁶	2099.5	2100	21.5
III.6	KKKX-S ⁶⁰ KMEILQHVIDYILDLQ ⁷⁶	2597.2	2599.5	17.5
III.7	S ⁶⁰ KMEILQHVIDYILDLQ ⁷⁶ -XKKK	2597.2	2600.7	18.8

^a The sequences are N-terminally acetylated and C-terminally amidated (X = Ahx). ^b The masses [M+H]⁺ were measured by MALDI-ToF-MS. ^c The retention times t_R refer to the HPLC runs on a reverse-phase C18 column, with the binary system (A) 0.012% TFA in water and (B) 0.01% TFA in acetonitrile (gradient: 10-70% B over 40 min).

The three compounds were dissolved at the concentration of 50 μM in phosphate buffer (100 mM, pH 7.2) and analyzed by CD spectroscopy (Figure 8). As expected, peptide **II.8** assumed a flexible conformation at this concentration, giving a CD spectrum characterized by a negative band at 201 nm followed by a shoulder at 225 nm.

When the motif (Lys)₃-Ahx was coupled to the N-terminal residue of **II.8**, the resulting peptide **III.6** was helical, with an intense CD spectrum characterized by two negative bands at 206 nm and 221 nm (*R* = 0.85), and an intense positive band below 195 nm with a crossover at 198 nm. When the motif (Lys)₃-Ahx was coupled to the C-terminal residue of **II.8**, the resulting peptide **III.7** displayed a CD curve with a blue-shifted minimum at 204 nm, a shoulder near 222 nm (*R* = 0.74), but no maximum below 200 nm. This shape reflects the presence of the both α- and the 3₁₀-helix types. Therefore, upon these conditions peptide **III.7** probably represents a conformational intermediate state between the mostly flexible state of peptide **II.8** and the α-helix folded state of peptide **III.6**, which is likely to be characterized by the presence of a nascent helix, like the 3₁₀-helix [11].

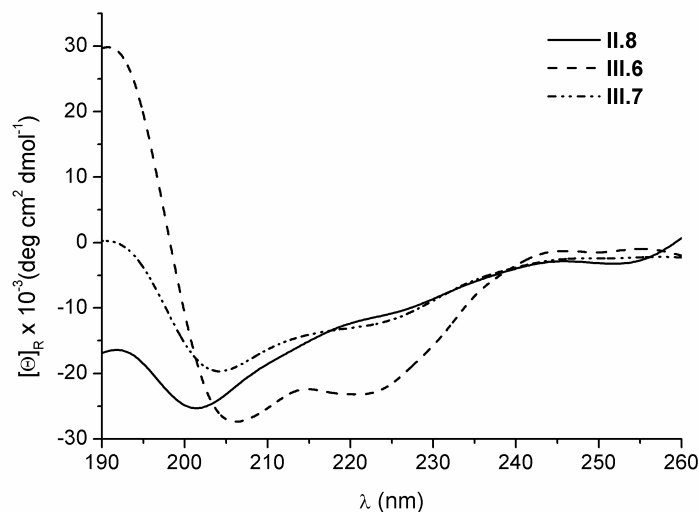


Figure 8: CD spectra of peptides **II.8**, **III.6** and **III.7** (50 μ M in 100 mM phosphate buffer, pH 7.2).

The α -helix propensity of peptides **II.8** and **III.7** could be enhanced by addition of methanol to the phosphate buffer (Figure 9, panels A and C). Indeed, increasing the amount of organic solvent was accompanied by an increase in the intensity of a α -helix-like CD curve, thus indicating an increase in the helix content. In contrast, addition of up to 40% methanol to the buffer solution of peptide **III.6** caused a destabilization of the α -helical structure, which, however, was overcome at higher methanol contents (Figure 9B). Interestingly, the intensity of the CD spectrum of **III.6** at the highest methanol content investigated (80%) was inferior to that obtained for peptides **II.8** and **III.7**, which might be attributable to the presence of a fraction of β -sheet aggregates.

All together, these results suggest that the N-terminal Lys-tag stabilized the α -helix structure of the target peptide **II.8** under physiologic conditions, whereas the C-terminal Lys-tag stabilized the 3_{10} -helix type. However, in the presence of high contents of methanol, the C-terminal Lys-tag was found to be superior in stabilizing the α -helix, as indicated by the high intensity of the CD band below 200 nm for peptide **III.7** (Figure 8C).

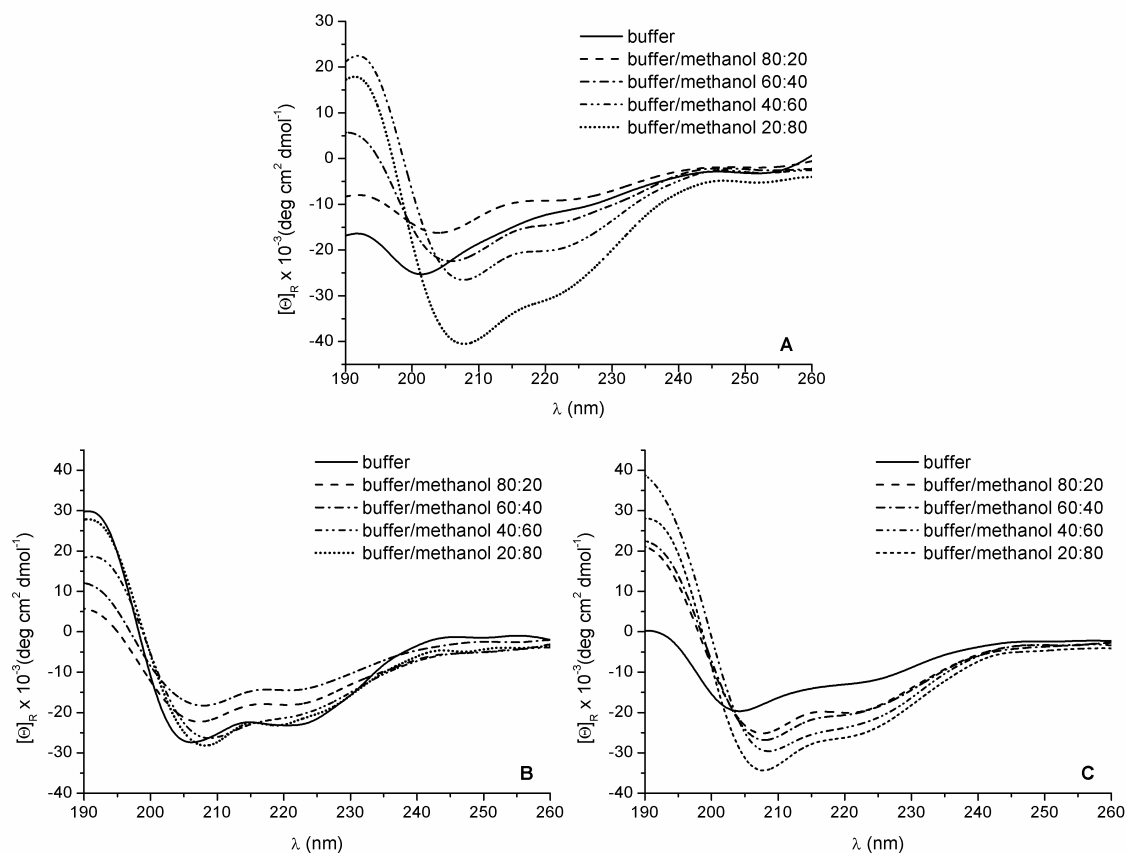


Figure 9: CD spectra of helix-2 peptide analogues (50 μ M) in methanol/phosphate buffer mixtures. (A) **II.8**, (B) **III.6** and (C) **III.7**.

III.4 Characterization of the insoluble form of the C-terminal NES region of Id2

Beside the conformation of the Id2 NES region in solution, we also wanted to characterize its insoluble form. To do that, we analyzed at first the morphology of the aggregates formed by the NES peptide **III.1** and by the variant **III.2** by transmission electron microscopy (TEM). Two samples of **III.1** were prepared in methanol and in methanol/water 80:20 (the high alcohol content was necessary to slow down the precipitation), while **III.2** was dissolved in methanol/water 20:80 because of its improved hydrophilic character. After incubation for ten days or one month, the obtained suspensions

were stained with a 2% aqueous solution of phosphotungstic acid (pH 7.2) and analyzed by TEM. After ten days, peptide **III.2** had formed fibrils with widths of 6-8 nm and lengths ranging from 60 to 200 nm (Figure 10A). Fibrils characterized by a banding pattern with alternating light and dark zones of 8 and 3 nm heights, respectively, were also visible (Figure 10B). Instead, isolated or assembled spherical particles (~10 nm in diameter) as well as amorphous material were present in the sample of peptide **III.1** containing 80% methanol (Figure 10C). Similar isolated spheres were detected for peptide **III.1** in methanol, although in this case the absence of water had strongly lowered the formation of deposits (Figure 10D).

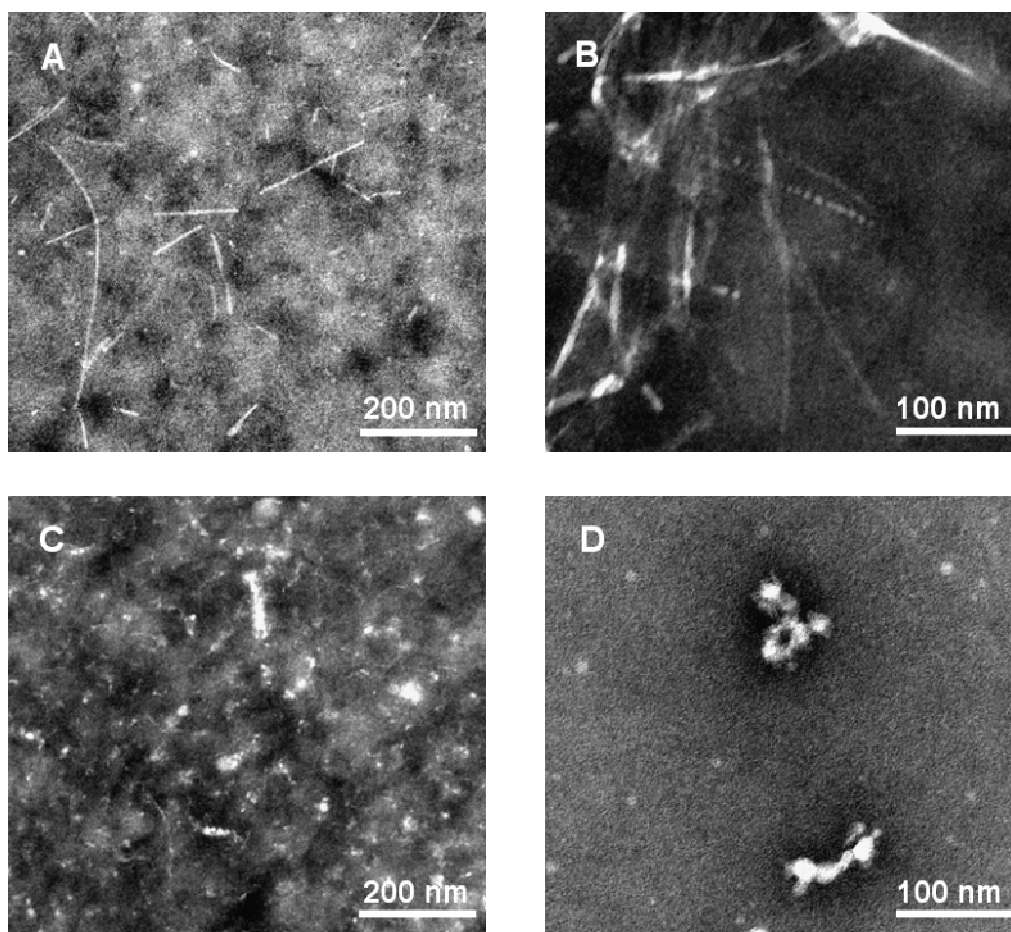


Figure 10: TEM images of the insoluble aggregates formed by peptides **III.1** and **III.2** after ten days. (A) Fibrils obtained by a 50 μM solution of peptide **III.2** in methanol/water 20:80; magnification grade of 20000. (B) Fibrils obtained by a 50 μM solution of peptide **III.2** in methanol/water 20:80; magnification grade of 40000 (C) Spherical particles and amorphous material obtained by a 50 μM solution of peptide **III.1** in methanol/water 80:20; magnification grade of 20000. (D) Spherical particles and amorphous material obtained by a 50 μM solution of peptide **III.1** in methanol; magnification grade of 40000. The staining was performed by using 2% aqueous solution of phosphotungstic acid (pH 7.2).

As it is known that β -sheet fibrils like those formed by amyloid peptides and proteins bind Thioflavin T (ThT) [12], we investigated whether the Id2 NES aggregates were also displaying this property. Indeed, by analyzing the ten-day-old samples of peptides **III.1** and **III.2**, we found that only peptide **III.2** increased the fluorescence emission of ThT at 484 nm after excitation at 446 nm (Figure 11).

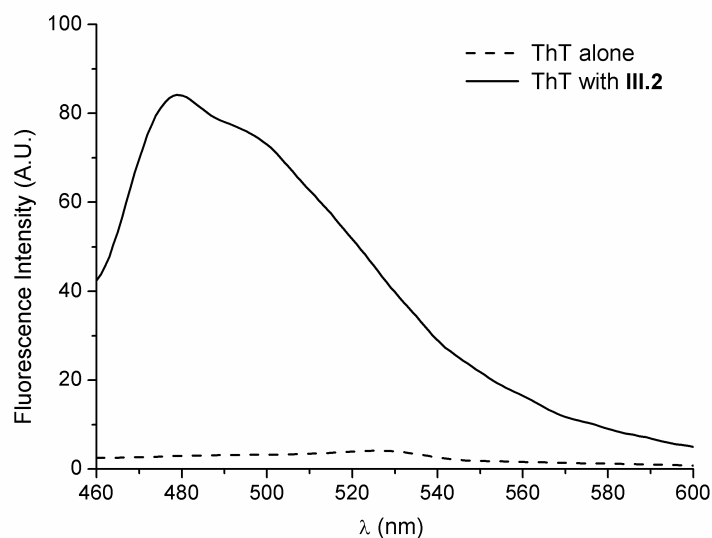


Figure 11: ThT assay. Fluorescence of ThT alone (5 μM in methanol/water 20:80) or mixed with a fibrillar sample of peptide **III.2** (50 μM in methanol/water 20:80). In the latter case, 25 μL of a ten-day-old peptide suspension were added to the ThT solution, which was then excited at 446 nm.

After one month, the peptide **III.2** fibrils reached up to 10 nm thickness and up to 1 μm length (Figure 12A-B), while peptide **III.1** precipitated abundantly from methanol with and without 20% methanol, forming highly dense deposits which appeared to be mostly amorphous (Figure 12C-D).

To control if the negative staining procedure could affect the morphology of the solid aggregates, the same one-month-old samples described above were stained with a 2% aqueous solution of uranyl acetate (pH 4.1). In the case of peptide **III.2**, besides the already observed fibrils, well-defined spherical particles with diameters in the range of 12-30 nm were observed (Figure 13A). Such granules were also characteristic of peptide **III.1** in 80% methanol (Figure 13B), whereas the sample in methanol presented compact deposits of both fibrillar and amorphous material (Figure 13C-D).

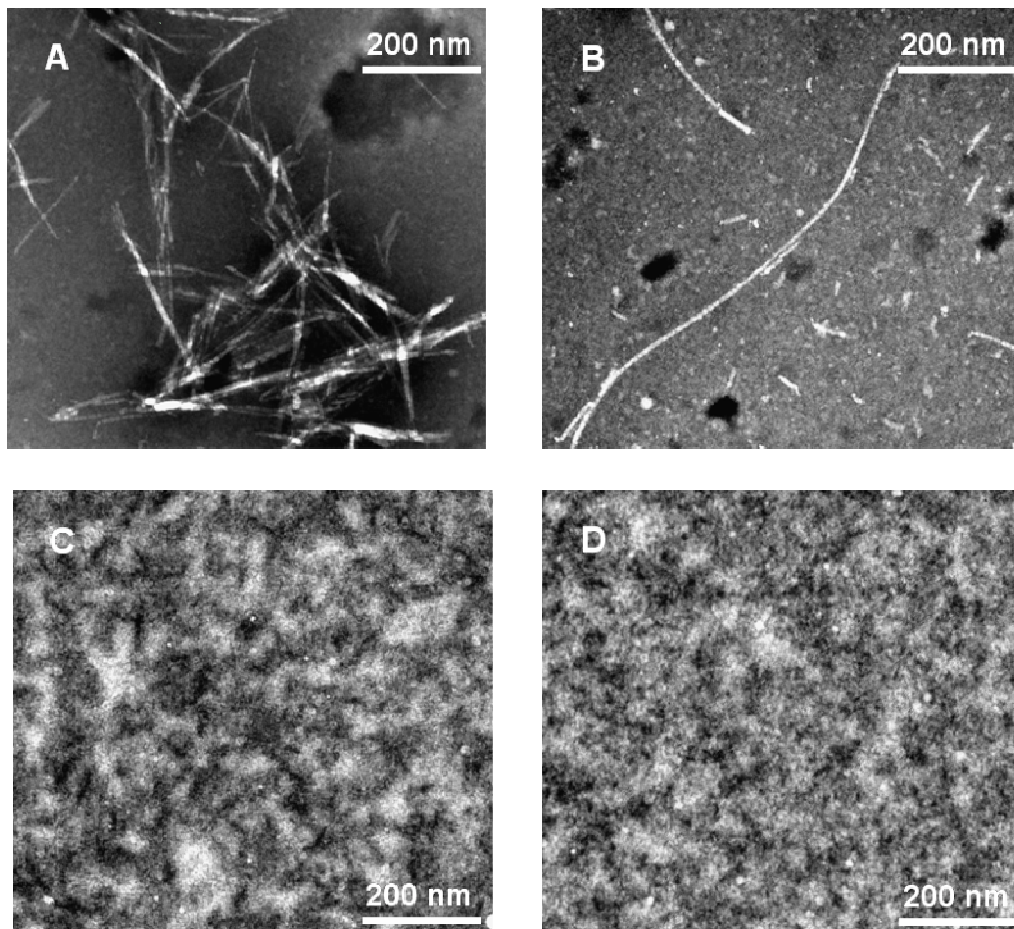


Figure 12: TEM images of the insoluble aggregates formed by peptides **III.1** and **III.2** after one month. (A-B) Fibrils obtained by a 50 μ M solution of peptide **III.2** in methanol/water 20:80. (C) Spherical particles and amorphous material obtained by a 50 μ M solution of peptide **III.1** in methanol/water 80:20. (D) Spherical particles and amorphous material obtained by a 50 μ M solution of peptide **III.1** in methanol. The staining was performed by using 2% aqueous solution of phosphotungstic acid (pH 7.2); the magnification grade is 20000 for all four pictures.

These results suggest that the aggregates of peptides **III.1** and **III.2** are pH sensitive: in particular, the fibrils of peptide **III.2** seem to be unstable at low pH, where the formation of spherical aggregates is favored. This would be consistent with the antiparallel β -sheet model described above (Figure 5), stabilized by electrostatic interactions between the positively charged lysine side chains and the negatively charged side chains of the C-terminal glutamic acid residues. Such favorable interactions would be partially lost at pH values near 4, thus destabilizing the β -sheet aggregates (Figure 14). In the case of peptide **III.1**, the morphologies of the aggregates shown after acidic staining appeared more ordered than those shown after neutral staining: indeed, well-defined granular deposits and clumped fibrils were found for the sample in 20% water and in methanol, respectively. Thus, it is likely that the staining procedure performed at neutral pH partly converted the ordered aggregates to amorphous material, due to the increase in the peptide net charge after deprotonation of the aspartic and glutamic acid residues.

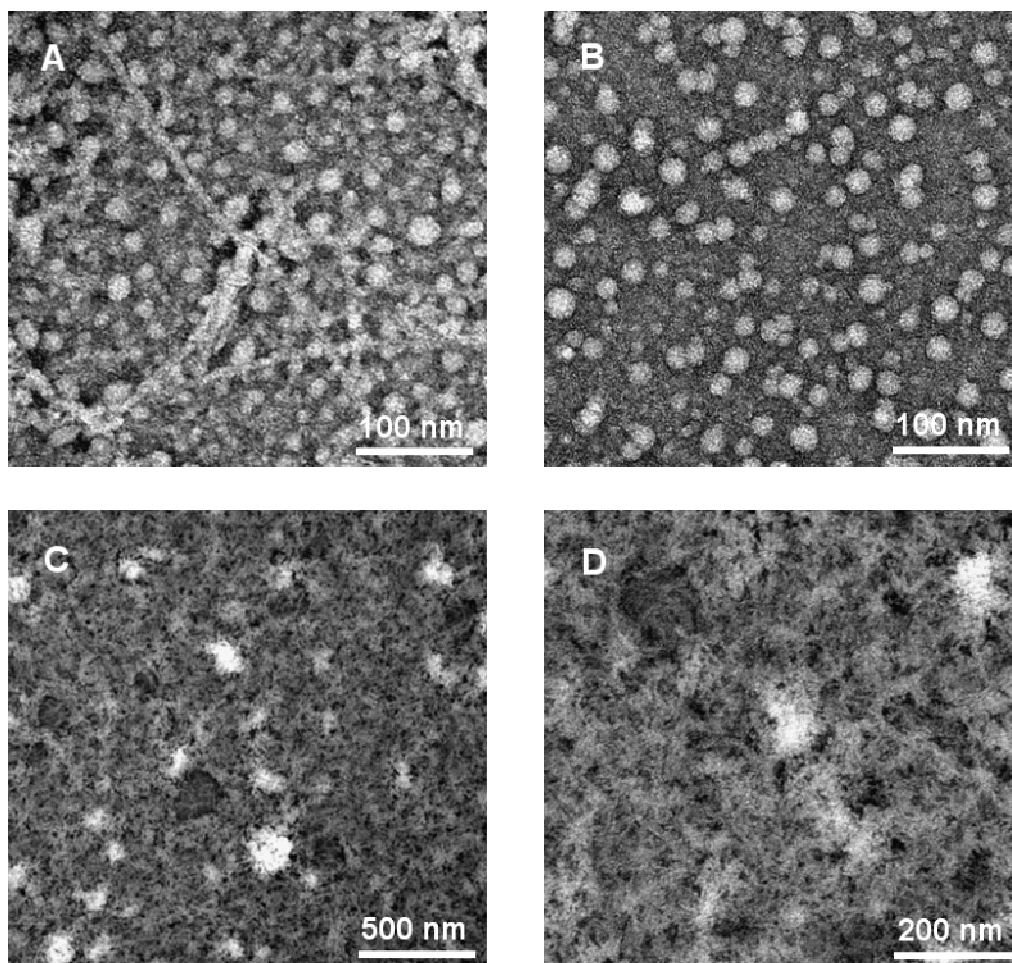


Figure 13: TEM images of the insoluble aggregates formed by peptides **III.1** and **III.2** after one month. (A) Fibrils and spherical particles obtained by a 50 μ M solution of **III.2** in methanol/water 20:80; magnification grade of 40000. (B) Spherical particles obtained by a 50 μ M solution of **III.1** in methanol/water 80:20; magnification grade of 40000. (C-D) Clumped fibrils and amorphous material obtained by a 50 μ M solution of **III.1** in methanol; magnification grade of 10000 and 20000, respectively. The staining was performed by using 2% aqueous solution of uranyl acetate (pH 4.1).

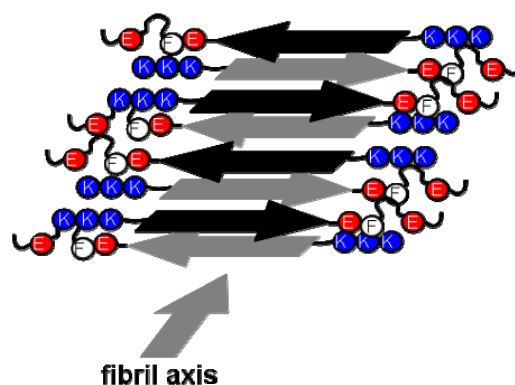


Figure 14: Proposed antiparallel alignment of **III.2** β -strands in the fibril.

For comparison, also peptide **III.3** was analyzed by electron microscopy. A sample of **III.3** in methanol/water 20:80 was incubated for 14 days or one month, and the obtained suspensions were stained with a 2% aqueous solution of phosphotungstic acid (pH 7.2) and analyzed by TEM (Figure 15, panels A and C). Interestingly, already after 14 days incubation peptide **III.3** was able to form mature long fibrils, characterized by a banding pattern with alternating light and dark zones of 6-9 nm and 5-6 nm width, respectively (Figure 15A). After one month, the fibrils grew further and formed compact assemblies (Figure 15C). Unlike the morphology of peptides **III.1** and **III.2**, the one shown by the aggregates formed by peptide **III.3** was not sensitive to the pH value of the staining procedure. As a matter of fact, the same type of fibrils was detected by staining with a 2% aqueous solution of uranyl acetate (pH 4.1) and with a 2% aqueous solution of phosphotungstic acid (pH 7.2) (Figure 15, panels B and D).

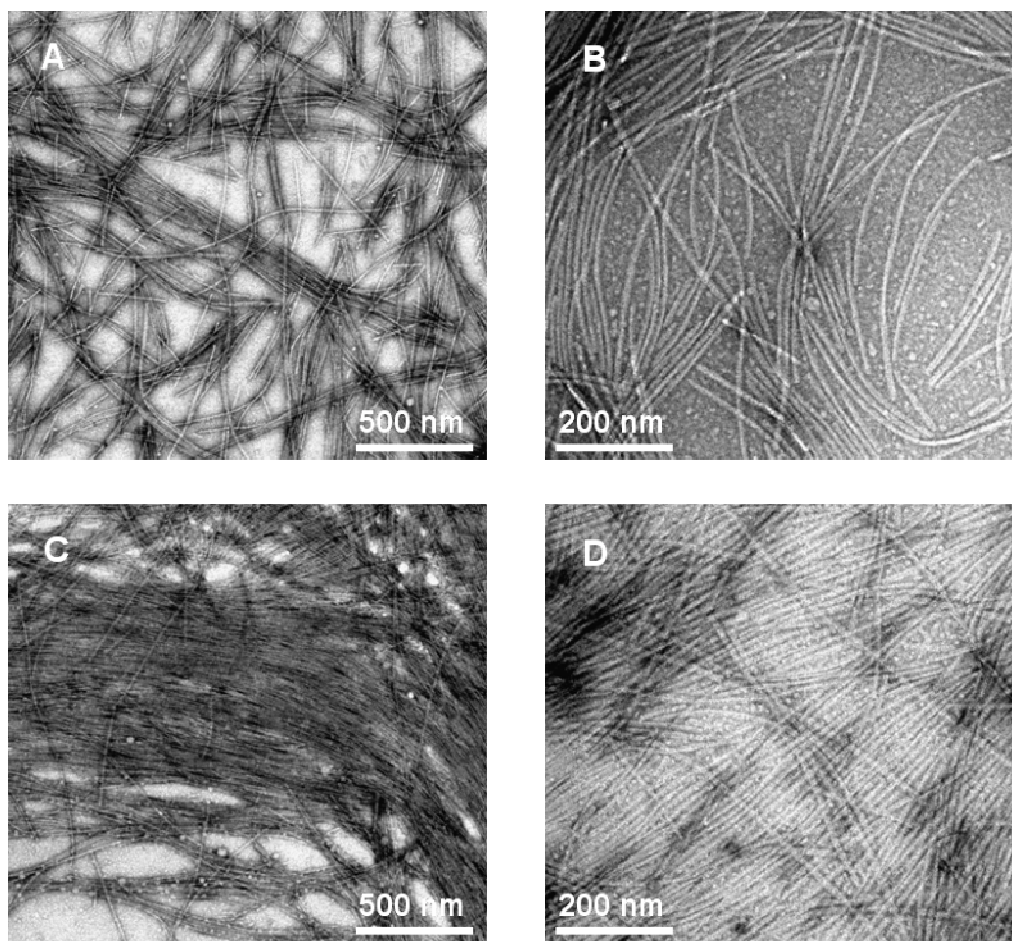


Figure 15: TEM images of the aggregates formed by **III.3** from a 50 μ M peptide solution in methanol/water 20:80. (A) Fibrils obtained after 14 days incubation, staining with 2% aqueous solution of phosphotungstic acid (pH 7.2); magnification grade of 10000. (B) Fibrils obtained after 14 days incubation, staining with 2% aqueous solution of uranyl acetate (pH 4.1); magnification grade of 20000. (C) Fibrils obtained after one month incubation, staining with 2% aqueous solution of phosphotungstic acid (pH 7.2); magnification grade of 10000. (D) Fibrils obtained after one month incubation, staining with 2% aqueous solution of uranyl acetate (pH 4.1); magnification grade of 20000.

A proposed mechanism of fibril formation for the peptide **III.3** is shown in Figure 16. An antiparallel organization of the helix monomers in a bilayer with a hydrophobic core and two polar surfaces might be stabilized by favorable interactions between the helix dipoles. In this model the flexible lysine-tags would lie at opposite sites of the bilayer.

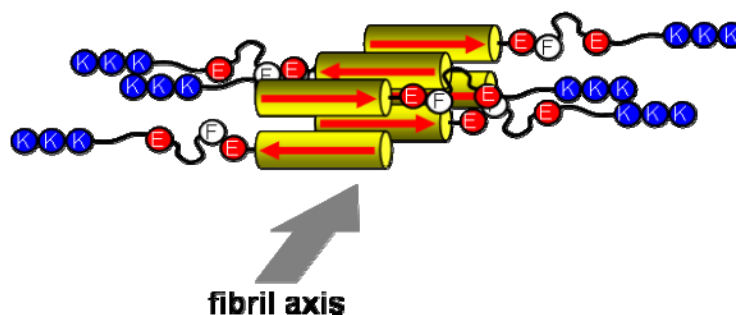


Figure 16: Proposed antiparallel alignment of **III.3** helices in the fibril.

Also an equimolar mixture of the peptides **III.2** and **III.3** in methanol/water 20:80 was studied by electron microscopy. After incubation for 14 days or one month, the obtained suspensions were stained with a 2% aqueous solution of phosphotungstic acid (pH 7.2) (Figure 17, panels A and C) or with a 2% aqueous solution of uranyl acetate (pH 4.1) (Figure 17, panels B and D) and analyzed. Fibrils with morphology similar to that of the aggregates formed by **III.3** were observed. This is in agreement with the CD data obtained for the same mixture, which have been described above (Figure 6). Accordingly, hetero-association of **III.2** and **III.3** is likely to improve helicity and helix stability and might lead to self-assembly with fibril formation, as proposed in Figure 18.

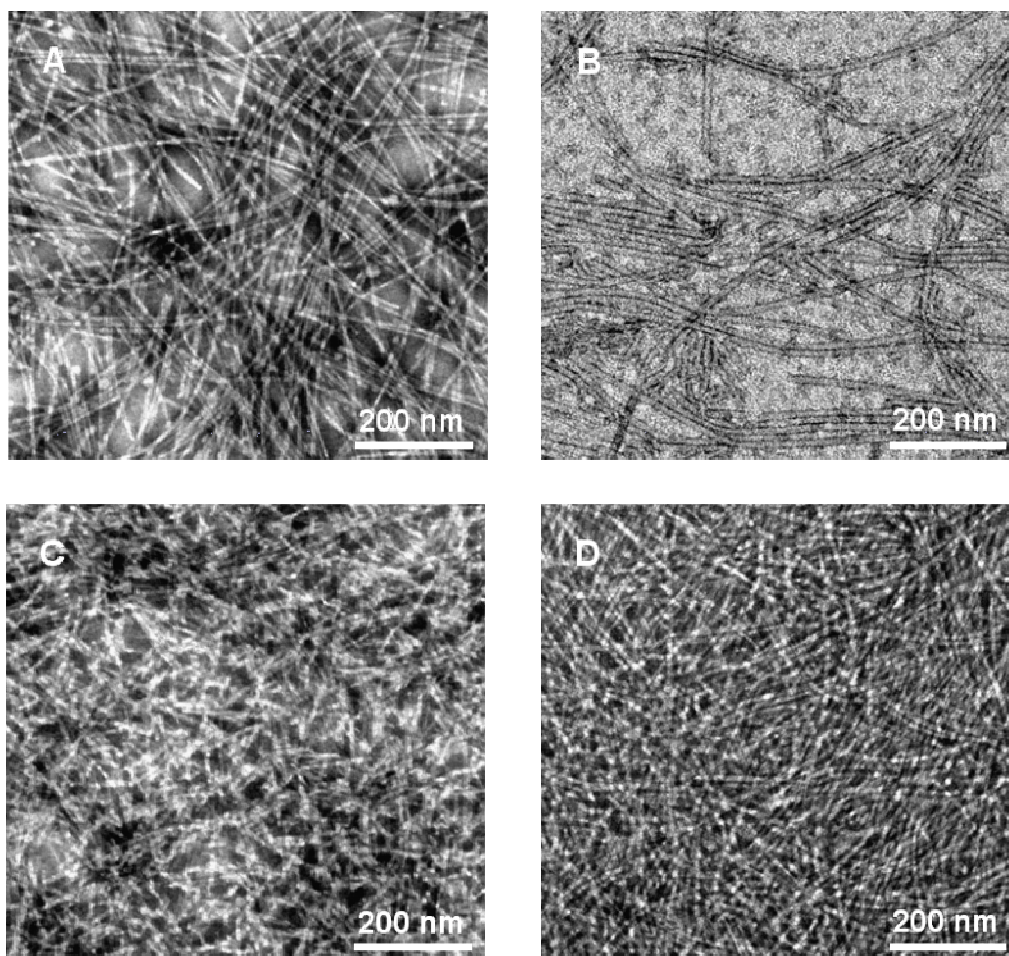


Figure 17: TEM images of the aggregates formed by an equimolar mixture of peptides **III.2** and **III.3** (50 μ M in methanol/water 20:80). (A) Fibrils obtained after 14 days incubation, staining with 2% aqueous solution of phosphotungstic acid (pH 7.2). (B) Fibrils obtained after 14 days incubation, staining with 2% aqueous solution of uranyl acetate (pH 4.1). (C) Fibrils obtained after one month incubation, staining with 2% aqueous solution of phosphotungstic acid (pH 7.2). (D) Fibrils obtained after one month incubation, staining with 2% aqueous solution of uranyl acetate (pH 4.1). The magnification grade was of 20000 for all pictures.

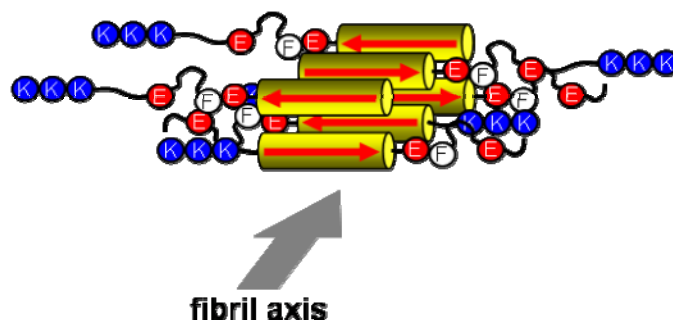


Figure 18: Proposed antiparallel alignment between the peptides **III.2** and **III.3** in the fibril.

III.5 Conclusions

We have shown that the Id2 fragment 103-124 containing the leucine-rich active NES motif adopts a β -sheet structure in methanol/water, from which it aggressively precipitates in amorphous, spherical or fibril-like forms depending on the environment. We could also show that the conjugation of target peptides to lysine-tags at the N- or C-ends can affect the conformation, and that, consequently, a careful analysis on the structure of lysine-tagged peptides is required.

Thus far, there is no evidence that the Id2 protein is subjected to oligomerization/aggregation *in vivo*; however, it is known that Id2 can form homodimers stabilized by an intermolecular disulfide bond involving the HLH residue Cys-42, and that the Cys-42/Ala mutation is deleterious for the Id2 activity [13]. Therefore, Id2 can undergo self-association, which might be a potential mechanism for auto-regulation or modulation of the inhibitory activity. Further studies are necessary to clarify whether the intrinsic propensity of the Id2 NES subdomain to aggregate has any significance for the folding and biology of this protein.

III.6 Literature

1. Colombo, N., Schroeder, J., Cabrele, C. (2006) A short Id2 protein fragment containing the nuclear export signal forms amyloid-like fibrils, *Biochem. Biophys. Res. Commun.*, *346*, 182-187.
2. Hara, E., Hall, M., Peters, G. (1997) Cdk2-dependent phosphorylation of Id2 modulates activity of E2A-related transcription factors, *Embo J.*, *16*, 332-342.
3. Deed, R.W., Hara, E., Atherton, G. T., Peters, G., Norton, J. D. (1997) Regulation of Id3 cell cycle function by Cdk-2-dependent phosphorylation, *Mol. Cell Biol.*, *17*, 6815-6821.
4. Fajerman, I., Schwartz, A.L., Ciechanover, A. (2004) Degradation of the Id2 developmental regulator: targeting via N-terminal ubiquitination, *Biochem. Biophys. Res. Commun.*, *314*, 505-512.
5. Florio, M., Hernandez, M. C., Yang, H., Shu, H. K., Cleveland, J. L., Israel, M. A. (1998) Id2 promotes apoptosis by a novel mechanism independent of dimerization to basic helix-loop-helix factors, *Mol. Cell Biol.*, *18*, 5435-5444.
6. Kurooka, H., Yokota, Y. (2005) Nucleo-cytoplasmic shuttling of Id2, a negative regulator of basic helix-loop-helix transcription factors, *J. Biol. Chem.*, *280*, 4313-4320.
7. Colombo, N., Cabrele C. (2006) Synthesis and conformational analysis of Id2 protein fragments: impact of chain length and point mutations on the structural HLH motif, *J. Pept. Sci.*, *12*, 550-558.
8. Kiewitz, S.D., Cabrele, C. (2005) Synthesis and conformational properties of protein fragments based on the Id family of DNA-binding and cell-differentiation inhibitors, *Biopolymers*, *80*, 762-774.
9. Bosques, C.J., Imperiali, B. (2003) Photolytic control of peptide self-assembly. *J. Am. Chem. Soc.*, *125*, 7530-1.
10. Ma, J.C., Dougherty, D.A. (1997) The Cation minus sign pi Interaction, *Chem. Rev.*, *97*, 1303-1324.
11. Crisma, M., Formaggio, F., Moretto, A., Toniolo, C. (2006) Peptide helices based on alpha-amino acids, *Biopolymers*, *84*, 3-12.

12. Le Vine 3rd, H. (1999) Quantification of β -sheet amyloid fibril structures with thioflavin T, *Methods Enzymol.*, 309, 274-284.
13. Liu, J., Shi, W., Warburton, D. (2000) A cysteine residue in the helix-loop-helix domain of Id2 is critical for homodimerization and function, *Biochem. Biophys. Res. Commun.*, 273, 1042-1047.

IV. Toward peptidomimetics as modulators of Id protein-protein interactions

IV.1 Introduction

One of the major challenges in medical research is to understand how a protein correlates to a specific biological or pathological effect. Many cellular processes are regulated through protein-protein interactions: as a result of the central role of this mechanism, the aberrant or inappropriate formation of protein complexes can lead to pathological events. One example for that is provided by the Id proteins, as already reported in Chapter I. Indeed, these proteins act as dominant negative regulators of DNA transcription by forming heterodimers with ubiquitously as well as tissue-specifically expressed bHLH factors [1]. As increased Id protein activity has been observed in several cancer types, perturbation of the interaction between the Id proteins and their natural dimerization partners by using synthetic molecules might represent an interesting approach for anti-cancer therapy. Moreover, the fact that the Id proteins are present at high levels in tumor cells, whereas they are generally not found in healthy adult cells, offers the advantage of selectively distinguishing between pathological and physiological systems.

There are different methods to understand the mechanisms of protein complex formation and to design new compounds that could interfere with it [2]. Computational approaches are a useful tool for the identification of small-molecule modulators of protein-protein interactions. They have been successfully used, for example, to discover inhibitors of the formation of a complex between Bcl-2 and Bak, two proteins involved in the apoptotic process. A second possibility consists of the screening of chemical libraries. This approach led to the identification of inhibitors of Myc/Max dimerization [2, 3]. Myc and Max belong to the HLH family of transcription factors and have a leucine-zipper domain at their C-termini. Based on the crystal structure of the related Max/Max homodimer, it has been assumed that the Myc/Max heterodimer forms a parallel four-helix bundle by interaction of the respective HLH domains. The protein-protein interface is formed mostly by hydrophobic amino acid residues. Vogt and co-workers [3] performed a screening of

several inhibitor candidates using fluorescence resonance energy transfer (FRET), and identified two compounds that were able to disrupt the formation of the complex (Figure 1). This was further confirmed in biological tests performed on chicken embryo fibroblasts, in which oncogenic c-Myc induced transformation was inhibited.

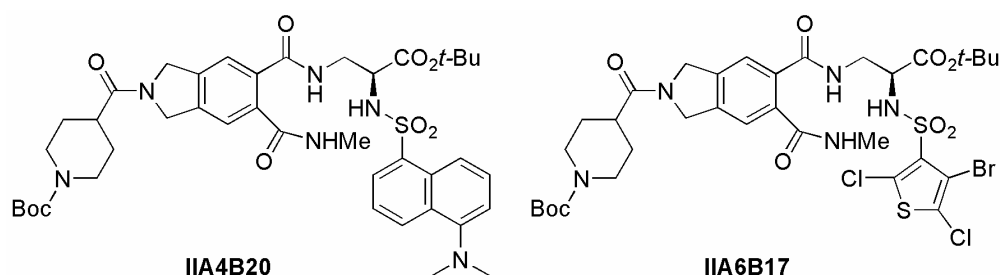


Figure1: Small-molecule inhibitors of the Myc/Max dimerization [3].

Another approach to target protein assembly consists of the identification of peptides derived from the interface between interacting proteins, which are able to bind to one of the complex subunits in a competitive fashion. These peptides are subsequently “converted” into peptidomimetics by the incorporation of unnatural amino acids or other chemical modifications to improve their specificity and proteolytic stability.

The field of the peptidomimetics is rapidly evolving because of their potential as precursors of efficient compounds that could lead to new therapeutically useful drugs. Peptidomimetics have some advantages as therapeutics in terms of pharmacokinetic properties, due to their major stability towards the hydrolytic activity of proteases. Moreover, the introduction of conformational restrictions can minimize the binding to non-target proteins and enhance the selectivity for the desired one, while the increase in hydrophobicity can result in better transport properties through cellular membranes. Accordingly, methylation of the amide nitrogen enhances the hydrophobic character of the peptide backbone, but also changes the network of intra-molecular interactions. The role of the peptide backbone can be investigated by the introduction of amide bond isosteres, like the $-\text{CH}_2\text{NH}-$ or $-\text{CH}_2\text{S}-$ groups, while the insertion of a *trans*-olefin is the best way to maintain the planarity and rigidity of the amide bond (Figure 2) [4, 5].

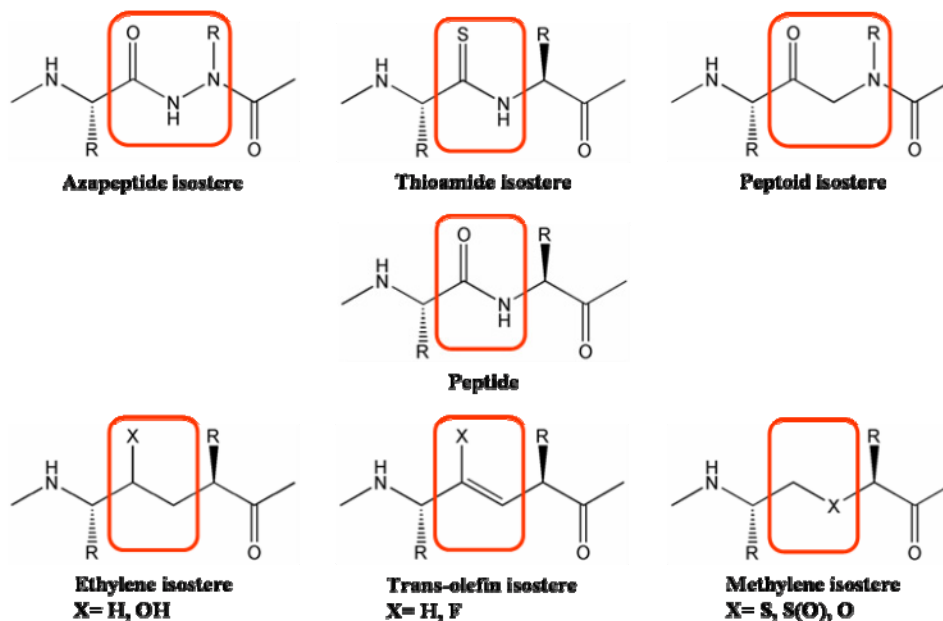


Figure 2: Examples of possible backbone modifications in the design of peptidomimetics.

The most common strategy for the design of peptidomimetics is based on the substitution of native amino acid residues with unnatural ones. Also the introduction of D-amino acids in bioactive sequences can positively affect the proteolytic stability and provide some oral bioavailability [6]; additionally, it can lead to a stabilization of secondary structure motifs, in particular of turns [7-9]. As a matter of fact, the ability of a peptidomimetic to fold into an ordered structure is, indeed, an important prerequisite. Polymers of α -amino acids present a well-known set of secondary structures characterized by precise hydrogen bond patterns and torsional angles along their backbones (Figure 3). The α -helix is the most abundant helix conformation found in globular proteins. It consists of a spiral arrangement of the peptide backbone with 3.6 amino acids per wheel stabilized by hydrogen bonds between residues i and $i+4$. The 3_{10} -helix differs from the α -helix in the number of residues per wheel, which is 3.2, and in the hydrogen bonding pattern that involves residues i and $i+3$. Another common conformation in nature is the β -sheet that is formed by two neighboring polypeptide chains aligned in a parallel or antiparallel manner and stabilized by intermolecular hydrogen bonds [10].

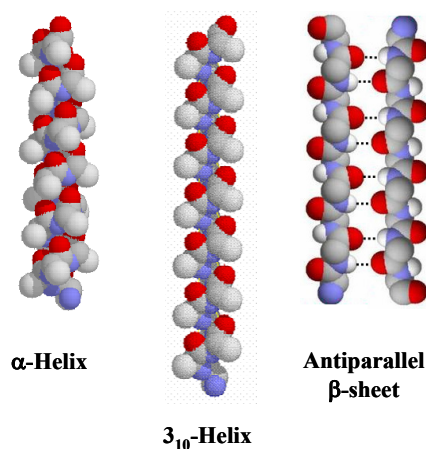


Figure 3: The most common secondary structures found in peptides and proteins.

The above mentioned secondary structure elements are connected in proteins by loops. They are important to change the direction of the backbone and are often involved in ligand-receptor recognition. Turns are classified according to the number of amino acids involved: for example, γ -turns are made of three residues, whereas β -turns are formed by four residues [11].

Many examples of synthetic amino acids as building blocks in bioactive sequences are known from literature [12-14]. A large number contain constraints that are important to rigidify the peptide structure. The application of synthetic building blocks to prepare analogues of the Id HLH motif and their conformational and pharmacological implications will be described in the following sections.

IV.2 3-Carboxy-cyclopentylglycine (Cpg) as a tool for N-N linkage of peptides

The non-proteinogenic amino acid 3-carboxy-cyclopentylglycine (Cpg) can be considered a conformationally restricted analogue of 2-amino-adipic and 2-amino-pimelic acid, two amino acids found in bioactive molecules, as well as a superior homologue of the natural α -amino acid Glu. Additionally, it also has the potential to be used as δ -amino acid,

while the presence of the two carboxylic groups connected by a chain of four carbon atoms allows this building block to be a tool for intra- or intermolecular crosslinking (Figure 4).

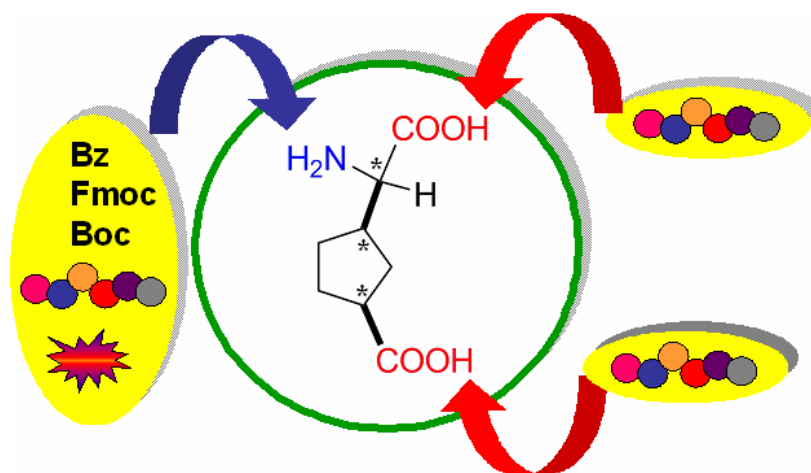
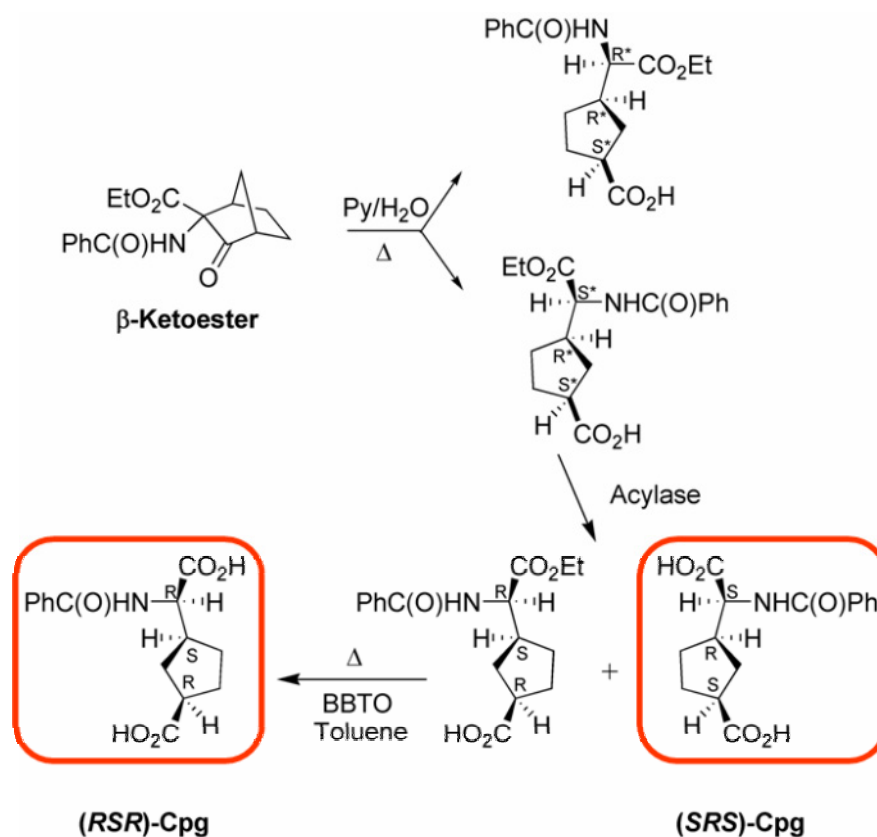


Figure 4: The Cpg scaffold: the carboxylic groups are suitable for conjugation to the N-terminus of peptides, whereas the amino group can be masked with different protecting groups, but also conjugated to peptides or labeled with fluorophores.

The synthesis of Cpg has been developed by Gelmi and co-workers [15]. Enantiopure **(SRS)-Cpg** and **(RSR)-Cpg** derive from a base-catalyzed retro-Claisen reaction on the β -ketoester shown in Scheme 1, followed by chemo-enzymatic resolution performed on a semi-preparative scale using the acylase from *Aspergillus melleus*.

Scheme 1: Synthesis of the Cpg enantiomers [15].

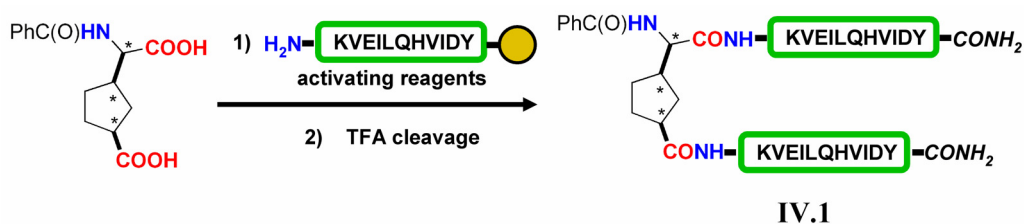


The main structural and functional domain of the Id proteins consists of two amphipathic α -helices connected by a loop. As mentioned above, these connecting regions are important to determine the folding of a biomolecule and they can play a role in protein-ligand and protein-protein interactions. In comparison to the helical regions, the loop of the Id proteins is much less conserved, although the ones of Id1 and Id2 present many similarities. The importance of the loop in the Id HLH folding has been investigated by its substitution with a C5-aliphatic chain [16]. CD spectroscopy analysis of these analogues has shown that they all dramatically lost helix content, with the Id3 HLH analogue being the only one to still contain some helical elements. This suggests that the Id loop plays a role in determining the correct orientation of the two flanking helices, thus favoring the folding and stability of this domain.

Starting from this observation, we were interested in the use of unnatural amino acids as loop surrogates, which should preferably stabilize a parallel arrangement of the two flanking regions, as it is suggested that the HLH motif of the Id proteins also displays a parallel helix packing, at least as heterodimer subunit. Loop-modified Id HLH peptides might be interesting models for the development of highly structured protein domains as well as of molecules interacting with the native HLH region of the Id proteins.

The Cpg scaffold presents two carboxylic groups separated by four carbon atoms, three of them are constrained in a cyclopentyl ring (Figure 4). Hence, Cpg can be used as a linker between two peptide portions and tested as potential loop surrogate. However, a limitation of the presented Cpg scaffold is the lack of an orthogonal protection for the two carboxylic acids, which reduces its application to the synthesis of covalent homodimers. Nevertheless, we decided to explore this application and to prepare Cpg-linked dimers of peptides corresponding to the helix-2 of the Id HLH motif that is more conserved and has higher intrinsic helix propensity than helix-1. Preliminary studies to evaluate the suitability of Cpg as a building block for solid phase peptide synthesis have been done on its racemic mixture. This was coupled manually to a resin-immobilized peptide related to the helix-2 of Id1 (residues 91-101 or 91-106) (Scheme 2). Due to the tendency of Cpg to epimerize in the presence of bases, the coupling reaction was accomplished using only DIC and HOBT [17]. These base-free acylation conditions have been used for the syntheses of all Cpg-containing peptides presented in this work (Figure 5). Unfortunately, the two diastereomers obtained for peptide **IV.1** could not be separated by HPLC, which prevents the use of Cpg as a racemate for our scope.

Scheme 2: Synthesis of a covalent dimer of the Id1 sequence 91-101. The yellow sphere represents the Rink amide resin (the Cpg building block was used as a racemate)



IV.2.1 Synthesis of Id peptide dimers containing (*SRS*)-Cpg or (*RSR*)-Cpg

A series of peptides analogous to **IV.1** have been prepared using the enantiopure Cpg derivatives, as shown in Scheme 3. The Id fragments were first synthesized by automated stepwise solid phase synthesis, and their amino acid sequences are listed in Table 1. The quality of the assembled peptide chain was controlled by performing a small-scale TFA cleavage of the peptide from the solid support: the crude peptides were found to be highly homogeneous, with the exception of **IV.4**, derived from helix-2 of Id3, and of **IV.5**, derived from helix-1 of Id1. In the first case, a truncated peptide lacking a glutamine was present as side product, while in the second one there was a more complex mixture of undesired products. Therefore, **IV.4-5** were not further used in this study.

Scheme 3: Synthesis of Id peptide dimers containing Cpg in the *SRS* or *RSR* configuration. The peptide sequence (green colored) was anchored to the Rink amide resin (yellow sphere).

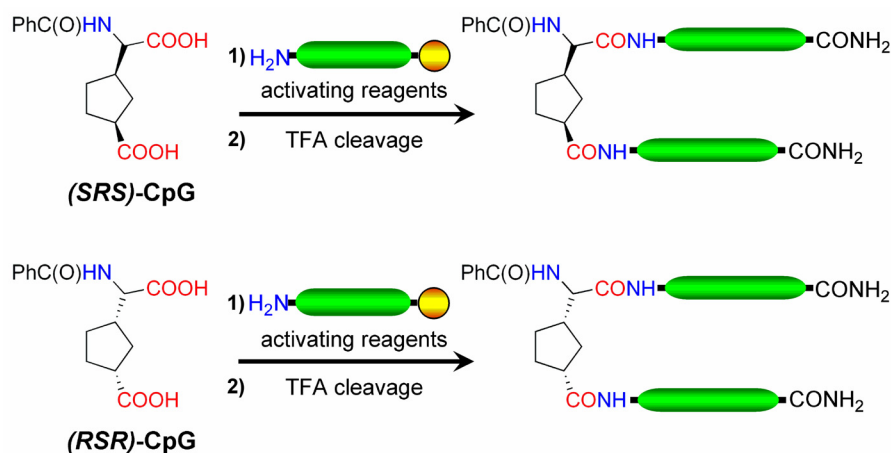


Table 1: Id sequences C-terminally anchored on Rink amide resin for coupling with the Cpg building block. They reproduce part of the helix-1 (**IV.5-7**) and of the helix-2 (**IV.2-4** and **IV.8**) of the Id HLH motifs.

Peptide no.	Amino acid sequence	Id fragment
IV.2	KVEILQHVIDY	Id1 91-101
IV.3	KMEILQHVIDY	Id2 61-71
IV.4	QVEILQRVIDY	Id3 66-76
IV.5	NGCYSRLKELV	Id1 70-80
IV.6	NHCYSRLRELV	Id3 45-55
IV.7	NDCYSRLRRLV	Id4 69-79
IV.8	KVEIL	Id1 91-95
IV.9	VKEIQLVHIYD	Id1 91-101 scrambled

Subsequently, the enantiopure Cpg units were coupled manually to the Id fragments to connect two equal sequences through their free N-termini (Figure 5). Although working in heterogeneous conditions, the Cpg dicarboxylate was not added in excess but rather stoichiometrically, in order to minimize the formation of mono-substituted Cpg derivatives. Therefore, only 0.5 equiv. Cpg were dissolved in DMF together with DIC and HOBt as activating reagents, whereas no base was added to avoid epimerization. The coupling mixture was added to the previously swelled peptidyl-resin and shaken for 24 hours at room temperature. Successively, further DIC and HOBt were added and the reaction was left to run for another day. The reaction completion was controlled by cleaving a small amount of peptide from the resin and analyzing it by HPLC and MALDI-ToF-MS. Surprisingly, the two Cpg enantiomers were successfully coupled to the Id helix-2 related sequences, but not to the Id helix-1 related sequences. Indeed, **IV.2**, **IV.3** and **IV.9** were converted into the corresponding covalent dimers **IV.10a/b**, **IV.11a/b** and **IV.12a/b** (Figures 5 and 6), whereas **IV.6** and **IV.7** were only partially converted into the two possible monosubstituted Cpg derivatives.

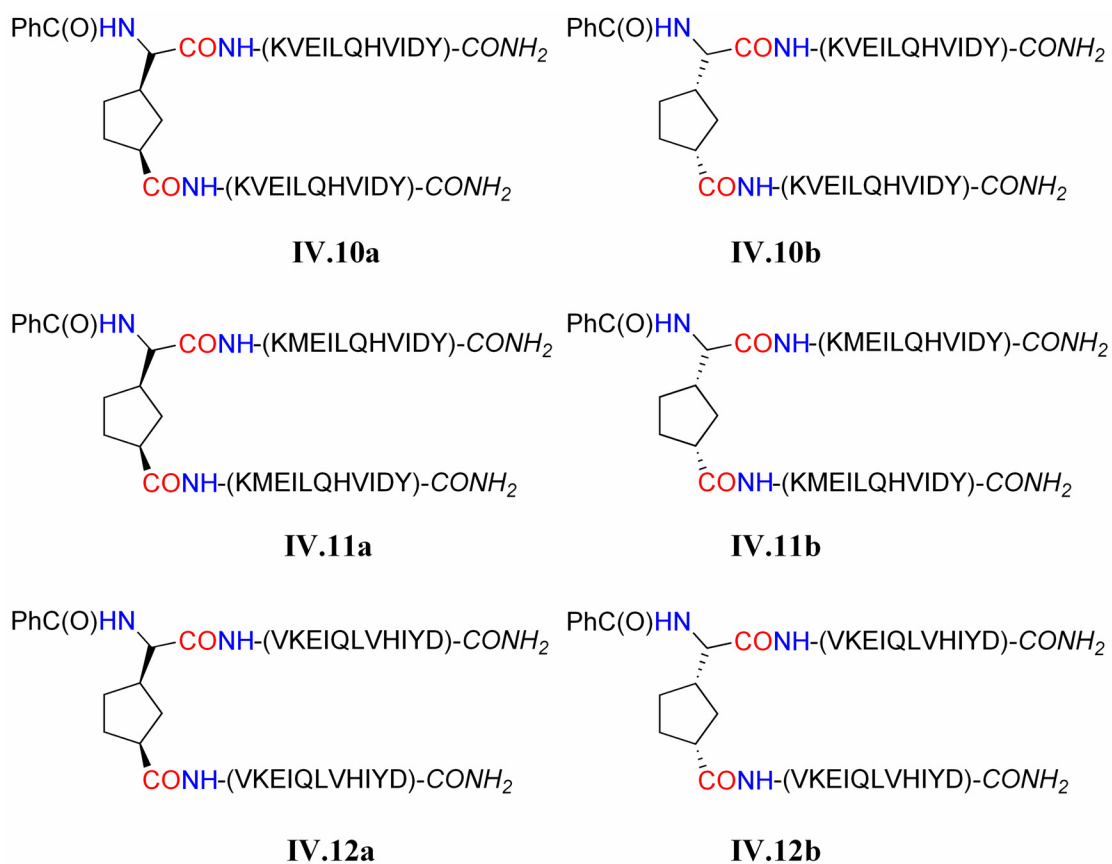


Figure 5: Id peptide dimers containing Cpg in the *SRS* or *RSR* configuration.

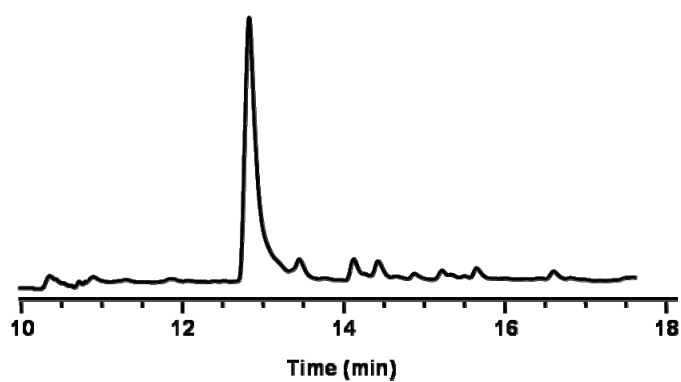


Figure 6: HPLC profile of the crude product of **IV.10a**.

Also the coupling of Cpg to peptide **IV.8** was not successful: (*SRS*)-Cpg incompletely reacted with the pentapeptide, whereas its enantiomer could only be monosubstituted. This suggests that the stereochemistry of Cpg may also play a role in the coupling reaction. Nevertheless, the low yield of crosslinking for **IV.8** was a bit surprising, but it is likely that the pseudo-dilution effect exerted by the solid support on the short-length peptide backbone disfavored intermolecular reactions [18].

Further, the Cpg building block has been used to anchor three peptide chains, as shown in Figure 7. In this case, the α -amino group of Cpg was Fmoc-protected, which allowed the assembly of the third chain after the coupling of Cpg on the peptidyl-resin. More precisely, the Fmoc-protected Cpg was first coupled to peptide **IV.2** under the same reaction conditions described above for the benzoyl-protected-Cpg. Acetylation with ten equiv. acetic anhydride followed to cap the eventually unreacted NH₂ groups of the resin-bound peptide chains. Due to the epimerization tendency of Cpg in the presence of bases, the commonly used mixture of piperidine in DMF was replaced with a mixture of DBU and HOBt in DMF to perform the Fmoc cleavage; indeed, although DBU is a stronger base than piperidine, it is less nucleophilic because of its sterically hindered nature. Moreover, highly diluted solutions of DBU and short reaction times (3-5 min) are already able to completely cleave the Fmoc group. The addition of HOBt is made to control the overall basicity of the mixture. The assembly of the Id1 fragment 70-80 was done manually using a double coupling procedure, with a high molar excess (8 equiv.) of the commercially available amino acids activated with DIC/HOBt (8 equiv. each). The Fmoc cleavage was performed with the DBU solution described above. Finally, the last amino acid was acetylated. With this protocol, peptides **IV.13a/b** were obtained with high homogeneity.

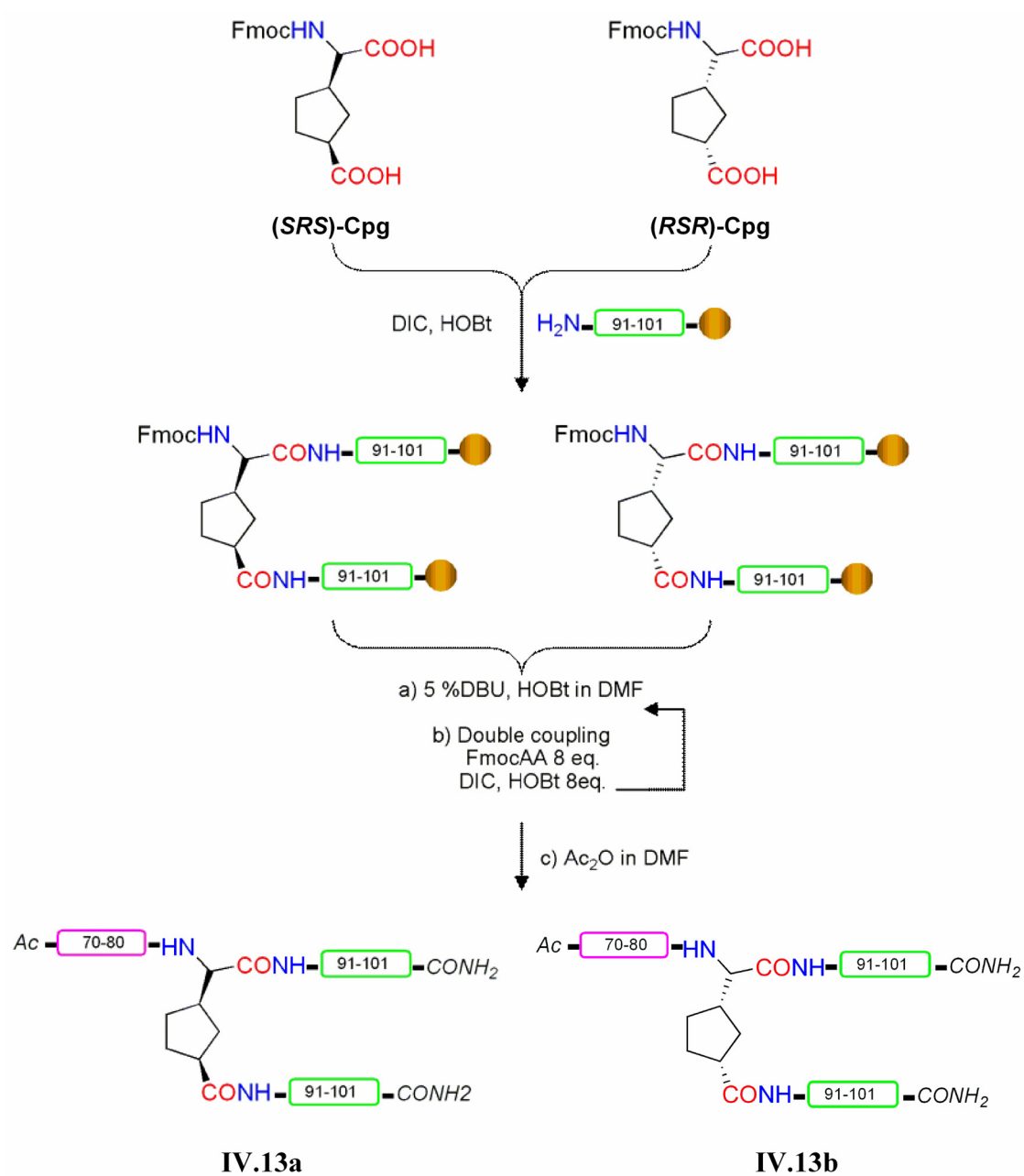


Figure 7: Synthesis of peptides IV.13a/b. For the preparation of Fmoc-protected Cpg see reference [15].

IV.2.2 Conformational characterization of the Cpg-containing Id peptides by CD spectroscopy

Peptides **IV.10a/b** and **IV.11a/b** differ only for the presence of a valine in place of a methionine, respectively, and this is not expected to strongly influence the conformation. Therefore, we chose to study the valine-containing peptides **IV.10a/b** because of the absence of oxidizing groups. The two diastereomers **IV.10a/b** were dissolved at the concentration of 90 μ M in phosphate buffer (100 mM, pH 7.3) at room temperature, and were measured immediately after preparation as well as after one day. The CD spectra are reported in Figure 8. Compound **IV.10a** showed a broad negative band below 210 nm, followed by a negative maximum at 219 nm and a weak negative minimum at 225 nm (Figure 8A). In contrast, peptide **IV.10b** showed a negative band below 200 nm, followed by a weak positive maximum at 207 nm and a negative minimum at 222 nm. These dichroic data suggest that **IV.10a** was mostly disordered (local polyproline-II contributions are probably present, due to the presence of the negative maximum at 219 nm), whereas **IV.10b** was partially ordered, as suggested by the presence of the well-defined minimum at 222 nm and the positive contribution at 207 nm. This is supported also from the CD spectra of the one-day old samples, as the one of **IV.10a** became clearly random-shaped, while the one of **IV.10b** underwent only a slight red shift (Figure 8B).

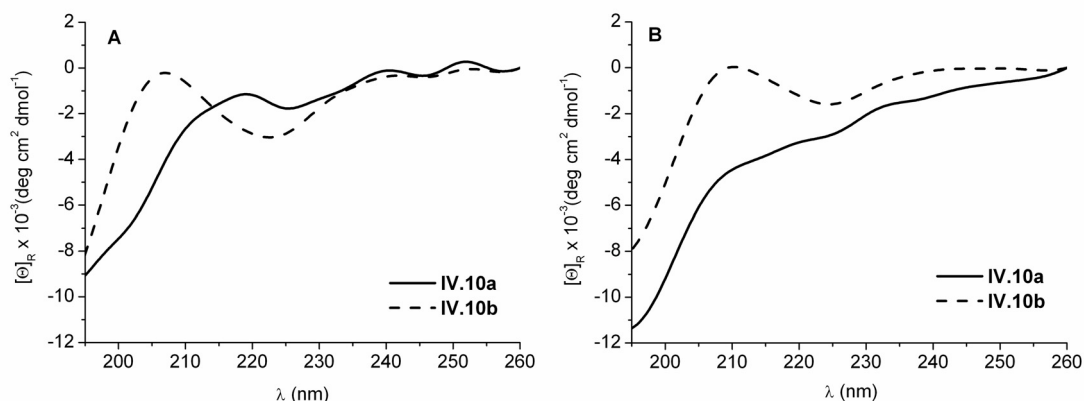


Figure 8: CD spectra of peptides **IV.10a/b** at the concentration of 90 μM in 100 mM phosphate buffer, pH 7.3. (A) Freshly prepared and (B) one-day old samples.

Further, comparison of the CD spectrum of the Id1 fragment alone with those of peptides **IV.10a/b** clearly showed that the Cpg unit with the configuration *RSR* induced a conformational transition from the disordered state of the single peptide chain to a partly ordered state of the corresponding covalent dimer (Figure 9A). Indeed, by subtracting the random component represented by the Id1 fragment alone from the spectrum of **IV.10b**, the resulting curve displayed a negative band at 223 nm and a positive one at 203 nm, which is indicative of the presence of a β -turn-like motif (Figure 9B). In contrast, the Cpg unit with the *SRS* configuration seemed to have almost no effect on the random conformation of the Id1 fragment, as suggested by the CD spectrum difference which is not significant over the region 205-260 nm

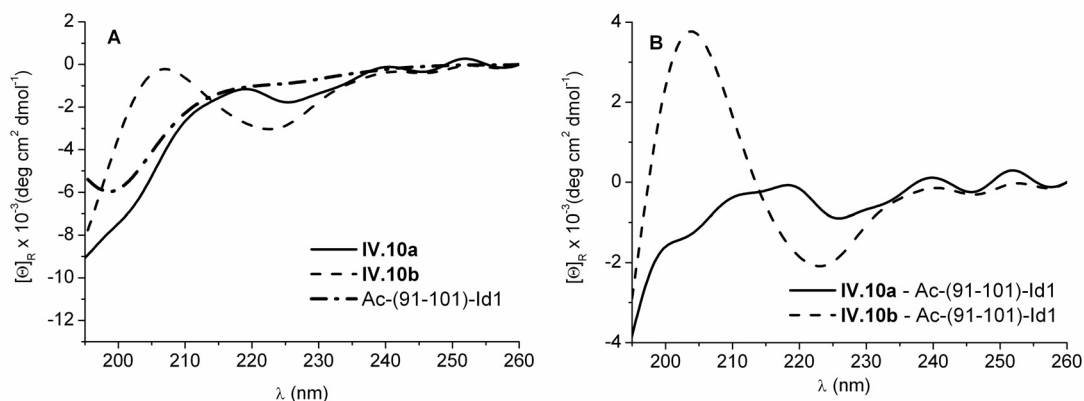


Figure 9: A) Comparison of the CD spectrum of Ac-(91-101)-Id1 with those of the peptides **IV.10a/b**. B) CD spectrum differences showing the CD intensity contribution of the Cpg units to the CD spectra of **IV.10a/b**.

Taking these observations together, it can be suggested that (*RSR*)-Cpg favored a β -turn or a β -bend conformation in **IV.10b**, which should allow the two flanking peptide chains to point into the same direction, whereas its enantiomer was probably unable to induce a preferred conformation.

IV.2.3 Investigation of the interaction of the Cpg-containing Id peptides with the native Id1 HLH motif by CD spectroscopy

As peptides **IV.10a/b** represent a dimeric form of part of the Id1 helix-2, we investigated their potential to interact with the native HLH motif of Id1 by CD spectroscopy. Each Cpg-containing peptide was mixed with the HLH motif in the ratio 3:1, in order to favor partner- rather than self-interactions of the latter. The peptide mixtures were measured immediately after preparation as well as the day after (Figure 10). Very weak peptide interactions were detected in the fresh mixture containing **IV.10a**, as indicated by the fact that the CD spectrum of the mixture was almost superimposable to the arithmetic sum of the CD spectra of the two separated peptides (Figure 10A). However, the one-day old mixture showed a well-defined helix-like CD curve that was characterized by a

significantly more intense maximum at 197 nm than the one obtained by the sum of the CD spectra of the two separated peptides (Figure 10C). This suggests that an interaction of **IV.10a** with the Id1 HLH motif occurred, which resulted in increased helicity. In the case of peptide **IV.10b**, an interaction with the Id1 HLH motif could be detected already in the fresh mixture. Indeed, the CD spectrum of the mixture was completely different from the one corresponding to the arithmetical sum of the two separated peptides (Figure 10B), displaying a negative band at 220 nm. Also the CD spectrum of the one-day old mixture showed the same shape but the minimum at 220 nm became more intense and a weak positive band at 201 nm appeared (Figure 10D). Thus, the peptide **IV.10b** was strongly interacting with the Id1 HLH motif, leading to a conformational change characterized by the formation of a β -sheet structure which was stabilized with the time. The presence of an isodichroic point at 200 nm suggests that a transition from an α -helix conformation to a β -sheet one took place in the mixture.

Taken together, these data suggest that both the Cpg-containing peptides were interacting with the Id1 HLH motif, but in a different manner: the one containing **RSR-Cpg** induced a rapid β -sheet conformation, whereas its diastereomer weakly stabilized the α -helix conformation.

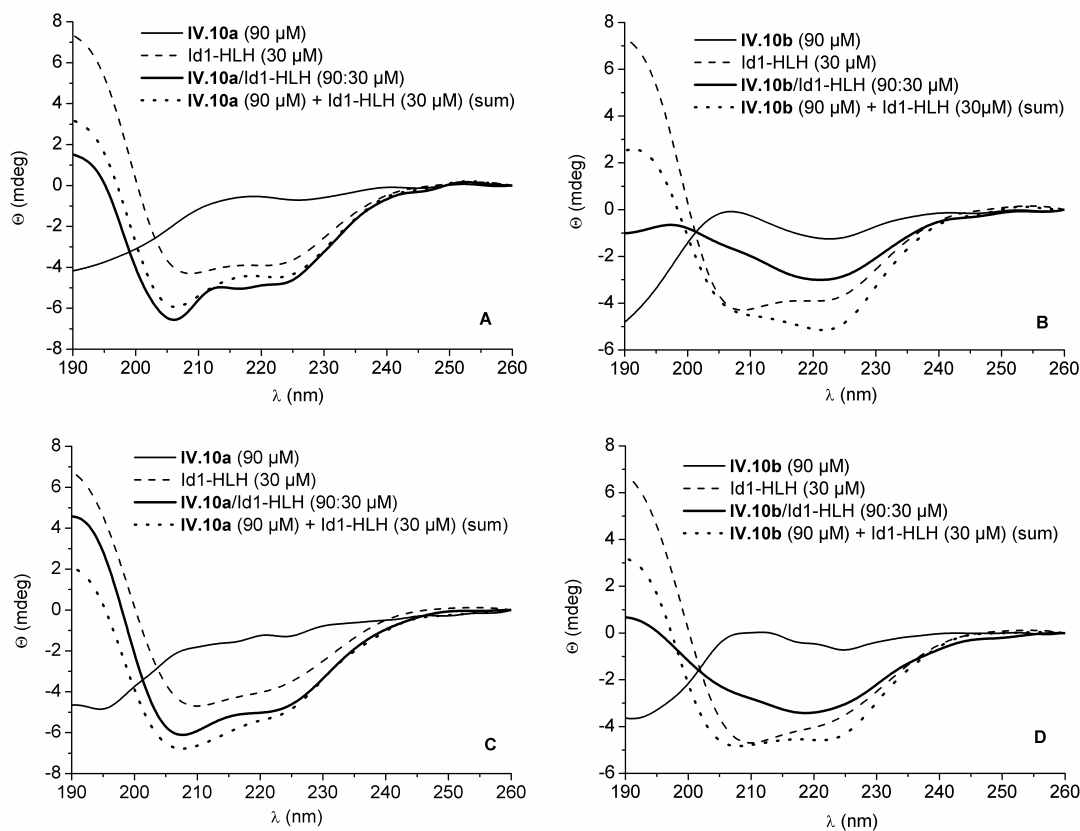


Figure 10: CD spectra of the peptides **IV.10a** and Id1 HLH motif, and of their 3:1 mixture. The three samples were measured immediately after preparation (A) or the day after (C). CD spectra of the peptides **IV.10b** and Id1 HLH motif, and of their 3:1 mixture. The three samples were measured immediately after preparation (B) or the day after (D). All samples were prepared in 100 mM phosphate buffer, pH 7.3. The dotted CD curves correspond to the arithmetic sum of the CD spectra of the two components of the mixture.

IV.2.4 Preliminary assays of the Cpg-containing Id peptides on a cellular model for atherosclerosis

There is clear evidence that the Id proteins regulate vascular smooth muscle cell (VSMC) growth and differentiation, and, consequently, their dysregulation contributes to vasculoproliferative disorders.[19]. Several data suggest that the Id proteins are mediators of various mitogenic factors acting generally through receptor tyrosine kinases: for

example, receptor tyrosine kinase stimulation by the platelet derived growth factor (PDGF) or serum increases both Id protein expression and VSMC growth. Another early response gene is c-Myc that activates Id2 gene transcription and is upregulated after vascular injury. In addition to their growth promoting properties, the Id proteins act as negative regulators of cellular differentiation. Modulation of the VSMC phenotype consisting in an alteration of their differentiated state is a hallmark of vascular lesion formation. As the primary function of VSMC is to contract, the expression of smooth muscle (SM) specific contractile proteins, like SM α -actin, is commonly used as an indicator of VSMC differentiation. *In vivo* expression of VSMC differentiation markers, including α -actin, decreases after vascular injury at time points coincident with the upregulation of Id2 and Id3. Therefore, the Id proteins are likely to be important mediators of the dedifferentiated state of VSMC during lesion formation.

Based on the CD spectroscopy data reported above, the Cpg-containing Id peptides **IV.10a/b** are apparently able to interact with the native Id HLH motif. Therefore, we were interested in the evaluation of these compounds in a cellular system. The synthetic phenotype of the human VSMC is characterized by decreased α -actin expression and increased proliferation and migration, thus offering a suitable cellular system to study atherosclerosis. To investigate the effect of the compounds on cell proliferation, VSMC cultures synchronized at the G₀ phase of the cell cycle were stimulated for three days with 10% FCS medium in the presence and in the absence of growing concentrations of peptides. As shown in Figure 11, compounds **IV.10a/b** showed a similar antiproliferative effect in the low micromolar range by reducing the number of cells of about 40%. As expected, also the mixture of the two diastereomers represented by peptide **IV.1** induced a similar decrease in cell proliferation. In contrast, the peptide corresponding to the single helix-2 of Id2 had no antiproliferative effect. These preliminary results suggest that the dimeric form of the Id helix-2 fragment connected by the (SRS)- or (RSR)-Cpg units were able to negatively affect cell proliferation, whereas the monomeric form of the Id helix-2 was not. Although the potency and efficacy of the tested Cpg-containing Id peptides is not satisfactory, nevertheless they may provide a starting point for the development of peptide-based compounds with improved pharmacological properties. One of the problems that must be overcome is the difficult cell uptake: one possible solution might be the

substitution of the benzoyl group of **IV.10a/b** with an arginine-rich peptide chain, which is known to favor cell membrane penetration and act as a molecular transporter for drug delivery [20].

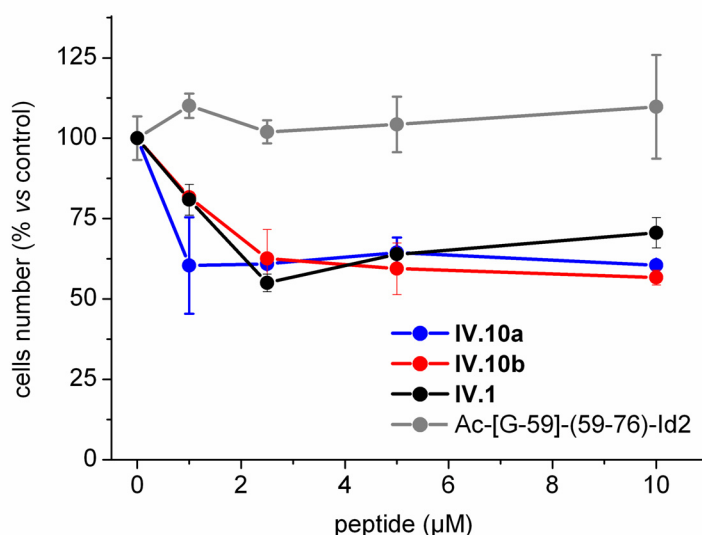


Figure 11: Effect of the Cpg-containing Id peptides **IV.10a/b** and of their diastereomer mixture (**IV.1**) on the proliferation of the synthetic phenotype of VSMC, and comparison with the effect of the Id2 helix-2.

As mentioned above, α -actin is a marker for differentiation of SM cells. The ability of the Cpg-containing peptides to induce differentiation of the synthetic phenotype of VSMC was evaluated by detecting the α -actin expression in the presence of growing peptide concentration. The effect of **IV.1** and **IV.10a** on the expression levels of α -actin is shown in Figure 12: both peptides positively affected the expression of the protein, whereas the peptide **IV.11a** containing the scrambled peptide sequence had not effect. This suggests that compounds **IV.1** and **IV.10a** not only were able to reduce cell proliferation, but also to promote cell differentiation. Moreover, this effect was specific for the compounds containing the correct Id helix-2 sequence, as the compound displaying the corresponding scrambled sequence was inactive.

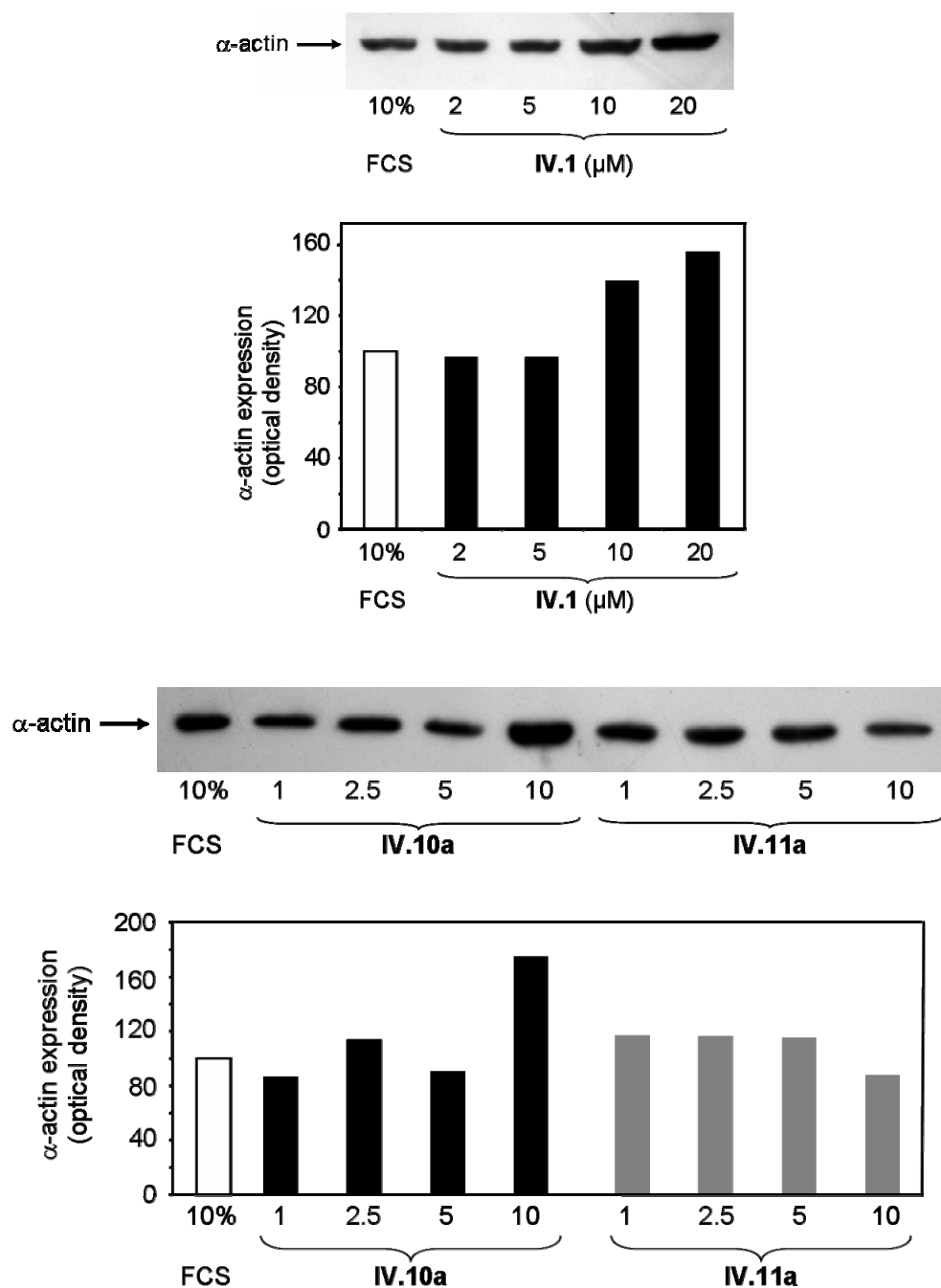


Figure 12: Effect of the Cpg-containing Id peptides **IV.1** and **IV.10a** on the expression levels of α -actin in the synthetic phenotype of VSMC, and comparison with the effect of **IV.11a** containing the scrambled sequence of the Id1 fragment 91-101.

Finally, the effect of **IV.1** on cell migration was tested by using a modified Boyden chamber assay. In the lower chamber, the medium contained 20 ng/ml PDGF as the chemotactic agent and the peptide (two different concentrations were tested: 1 μ M and 10 μ M). Migration of the cells from the upper chamber to the lower chamber was assessed after eight hours incubation. A moderate but still significant reduction of cell migration was observed, as shown in Figure 13.

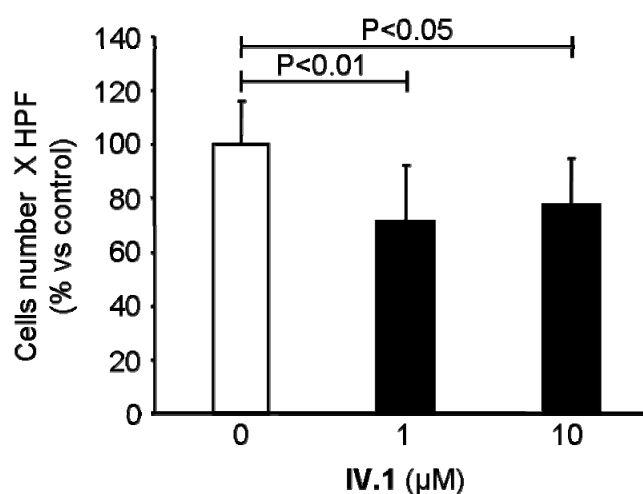


Figure 13: Effect of the Cpg-containing Id peptide **IV.1** on cell migration of the synthetic phenotype of VSMC (HPF, high power field).

All together, these preliminary data suggest that the Cpg-containing Id peptides investigated in this work might influence the inhibitory activity of the Id proteins, as they were able to inhibit VSMC proliferation and migration, while inducing differentiation. However, their potency is only in the micromolar range and needs to be improved, together with their cellular uptake and, of course, proteolytic stability.

IV.3 Small peptides containing the building block 3,4-(aminomethano)proline (Amp)

In the last decade, many efforts have been made to design and develop new oligopeptides, so called foldamers, which exist in a well-defined conformation in solution. These molecules aim to mimic or even to expand the functions of nature. The pioneering work of Gellman [21-23] and Seebach [24-26] on β -peptides designing oligomers of cyclic and acyclic β -amino acids has shown the power of these chemical tools in a number of biological applications. β -Peptides can in fact adopt ordered structures in solution, like helices defined by 8- up to 14-membered hydrogen-bonded rings [26]. Besides β -peptides, also α/β -peptides [12, 27], γ -peptides [28] and δ -peptides [29] have been described as promising biologically active compounds characterized by ordered solution structures. Therefore, the synthesis of new unnatural, structurally constrained amino acids and their application in the development of new oligomers with interesting architectures are a worthwhile research field in chemistry.

Recently, the synthesis of the unnatural amino acid 3,4-(aminomethano)proline (Amp) has been reported by de Meijere and co-workers [30]. Amp not only represents a derivative of proline, but it can be also seen as a conformationally constrained analogue of lysine, ornithine and, upon guanidinylation of the γ -amino group, of arginine. Moreover, it can be used both as cyclic α - and bicyclic γ -amino acid, as well as for the synthesis of branched peptides (Figure 14).

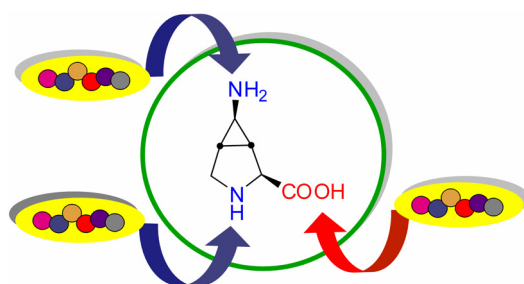
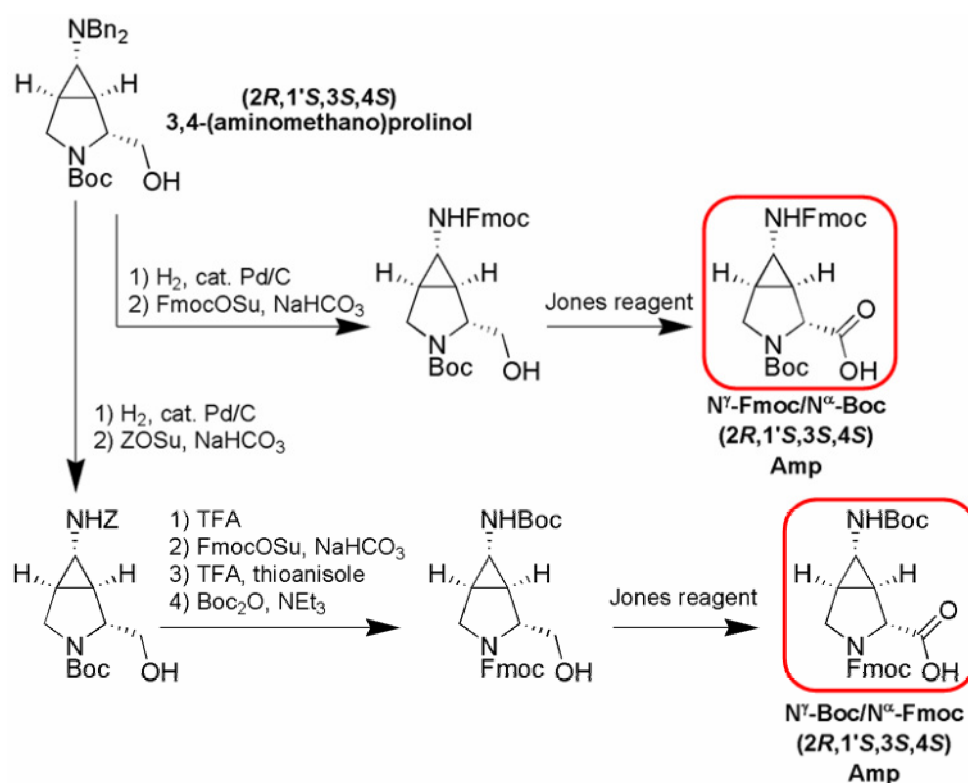


Figure 14: The Amp building block contains a secondary and a primary amine that can be used for the synthesis of α - and γ -peptides, respectively, or can be connected to the C-end of peptides to obtain branched peptides.

To investigate the potential of Amp as building block for peptide and foldamer synthesis, de Meijere and co-workers developed a synthetic route to obtain both the Boc/Fmoc-protected Amp enantiomers (Scheme 4) [31].

Scheme 4: Synthesis of **(2*R*, 1'*S*, 3*S*, 4*S*)-Amp** from the enantiopure 3,4-(aminomethano)-prolinol (for its synthesis see reference [31]). The enantiomer **(2*S*, 1'*R*, 3*R*, 4*R*)-Amp** comes from the corresponding enantiopure 3,4-(aminomethano)-prolinol with the same procedure (for details see reference [32]).



IV.3.1 Incorporation of Amp in small α/γ peptides [31]

An extensive research of the β -peptide class of foldamers has provided several new secondary structures. Besides β -peptides, also γ -peptides can fold in regular structures both in solution and on the solid state. Theoretical predictions by Hofmann and co-workers about γ -peptides revealed that structures with 14- and 9-membered helices are the most

stable in solution [33]. Hanessian *et al.* showed that γ -amino acids can favor a 14-helix [34] or reverse turns, whose screw sense depends on the stereochemistry of the γ -residue [35]. These findings are supported also by the studies of Seebach and co-workers, who observed the presence of 9-membered hydrogen-bonded rings in γ -peptides [36, 37]. Other examples by Ferrera-Sinfreu *et al.* [28] describe 9-membered sheets in γ -peptides containing cyclic γ -amino acids [28], while the acyclic γ -amino acids of Balaram and co-workers [38] made the peptides adopt 9-membered helices and ribbons. Kunwar and co-workers synthesized mixed γ -peptides [39] which can fold into a left-handed 9-helix in chloroform.

It has been shown that the structural features of all β - or mixed α/β -peptides are compatible with the world of the natural α -peptides, and that these unnatural foldamers can be potent tools to target protein-protein interactions [40, 41]. As the mode of action of the Id proteins is dictated by intermolecular contacts between helices, developing and using foldamers might be a powerful approach to target the Id proteins. Due to the ability of peptides containing γ -amino acids to adopt helical structures, we decided to undergo a synthetic and structural study of alternating α/γ -peptides containing the Amp unit [31].

Both enantiomers of N' -Fmoc/ N^α -Boc-protected Amp shown in Figure 15 were used for solid phase synthesis of alternating α/γ -oligomers consisting of the dipeptide repeat α -Ala- γ -Amp (Figure 15). The chain assembly was performed manually on Rink amide resin with an initial loading of 0.7 mmol/g. The dipeptide Gly-Tyr was chosen as general C-terminal motif for two reasons: (i) Gly can function as short spacer between the C-end and the core sequence, whereas (ii) Tyr provides the oligomer with an internal chromophore, which is advantageous for the determination of the concentration by UV spectroscopy. As the oligomers presented here were cleaved-off from the resin before removing the last Fmoc group, their concentration was determined from the UV absorption of the fluorene moiety at 301 nm.

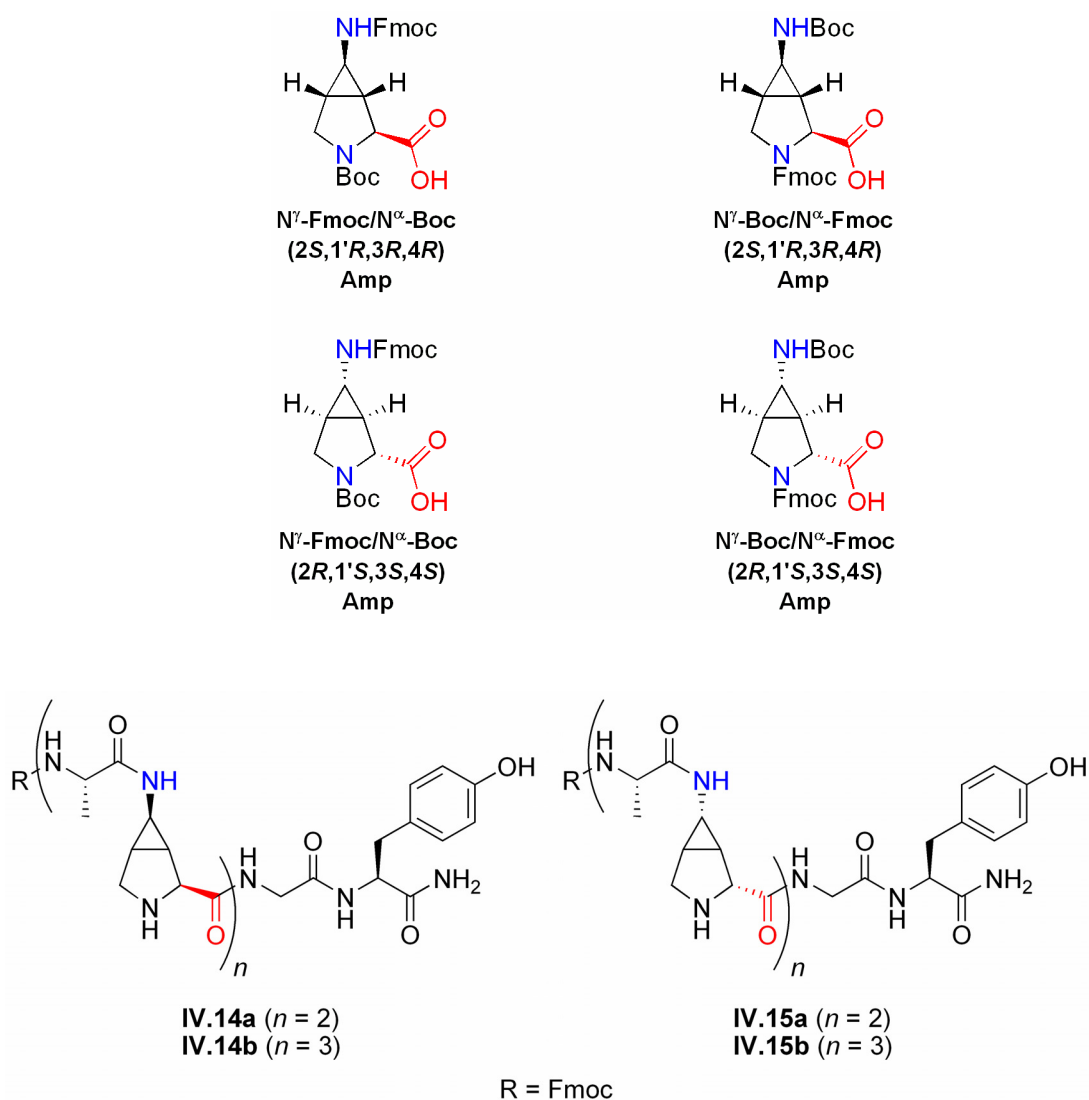


Figure 15: Orthogonally protected Amp derivatives (top) used for the solid phase synthesis of the oligomers **IV.14a/b** and **IV.15a/b** (bottom).

A single-coupling procedure was applied for the building block as well as for the natural amino acids in the presence of DIC and HOBt; however, whereas the acylation steps involving Amp were carried out using a moderate molar excess (2.5 equiv.) and a longer reaction time (16 h), those involving the commercially available amino acids were carried out using higher molar excess (4 equiv.) and a shorter reaction time (4 h). To control the quality of the growing peptide chain, a small-scale TFA cleavage was performed after each

γ -Amp coupling, and the cleaved-off product was characterized by analytical HPLC and mass spectrometry. The Fmoc cleavage steps were accomplished by using a mixture of DBU and HOBT in DMF (3-5 min). With this synthetic procedure, hexamers **IV.14a** and **IV.15a** and octamers **IV.14b** and **IV.15b** could be obtained with high homogeneity already after TFA cleavage from the resin and precipitation from ice-cold diethyl ether (Figure 16).

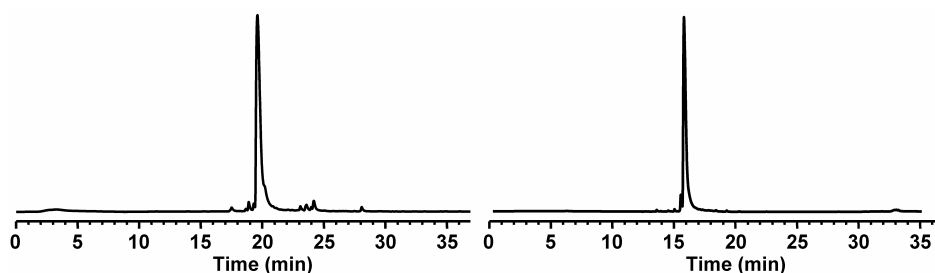


Figure 16: HPLC profiles of peptides **IV.14a** (left) and **IV.14b** (right)

IV.3.2 Structural investigation of the Amp-containing α/γ peptides

The synthesized α/γ -peptides were studied by CD and NMR spectroscopy. The CD spectra were analyzed in the far-UV region of the amide bond transitions (195-260 nm). Both the hexamer **IV.14a** and octamer **IV.14b** containing the γ -Amp unit with the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration were characterized in water by a positive signal in the range 200-240 nm, with a maximum at 208 nm for the hexamer or at 212 nm for the octamer, followed by a shoulder at 225 nm. The CD signal became negative below 198 nm (Figure 17A-B). The same CD shape was found in methanol, in which, however, the positive maximum for both oligomers was slightly blue-shifted to 205 nm; moreover, in the organic solvent peptide **IV.14b** showed a more intense spectrum with respect to peptide **IV.14a**, indicating a superior conformational stability for the longer peptide chain. Interestingly, these CD spectra are similar in the shape to those recently reported by Kunwar and co-workers [42] for solutions in methanol of α/γ -peptides containing acyclic γ -amino acids: NMR studies of

the latter had shown the presence of a helical conformation with high and moderate stability in chloroform and methanol, respectively [42] .

In contrast to **IV.14a** and **IV.14b**, oligomers **IV.15a** and **IV.15b** containing the γ -Amp unit with the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration presented completely different Cotton effects, as they were characterized by a negative CD signal in the range 195-225 nm. In water, **IV.15a** and **IV.15b** displayed an intense minimum near 207 nm, which was slightly red-shifted in methanol (to 211 nm for **IV.15a** and to 209 nm for **IV.15b**). Additionally, a weak positive band was detected near 230 nm for the two peptides in both solvents, with a crossover close to 225 nm (Figure 17C-D).

The dichroic properties of the four oligomers suggest that the enantiomers of γ -Amp induced two different ordered conformations that were described by opposite Cotton effects in the region 200-225 nm, positive in the case of the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration and negative in the case of the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration. Moreover, the induced conformations appeared to be not significantly affected by the change of the solvent from water to methanol.

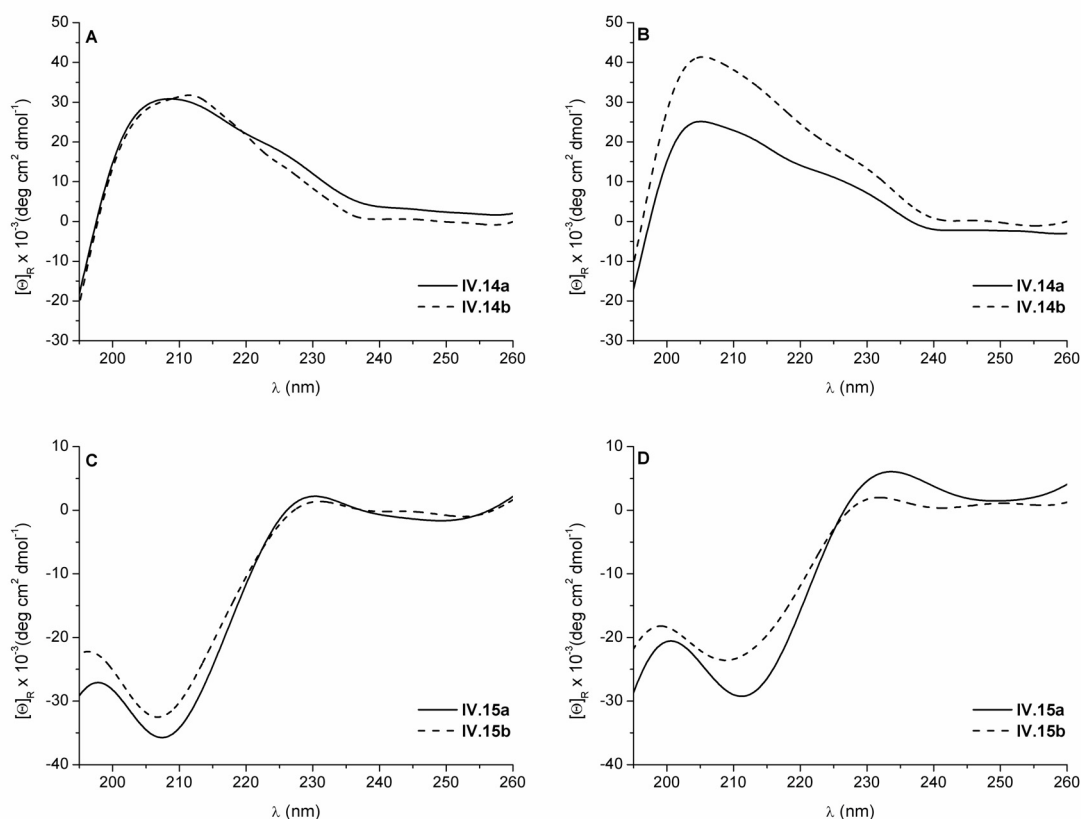


Figure 17: CD spectra of **IV.14a/b** at the concentration of 100 μ M in water (A) and in methanol (B). CD spectra of **IV.15a/b** at the concentration of 50 μ M in water (C) and in methanol (D).

Further, the four α/γ -peptides were dissolved in methanol- d_3 and studied by NMR spectroscopy. 1D experiments were first acquired at four different temperatures (278, 283, 293 and 303 K), and the one that gave the better resolved spectrum was chosen to record the 2D experiments. The complete ^1H NMR resonance and sequence assignments were done by using the COSY, TOCSY and NOESY spectra. Sequential and interresidue NOE crosspeak intensities were classified as strong (2-2.5 Å), medium (2-4 Å) and weak (2-5 Å) based on the number of contours in the contour plot of the NOESY spectrum. The values of the 3J coupling constants ($^3J_{\alpha(\gamma)\text{N}}$) between the α -NH and the α -CH, or the γ -NH and the γ -CH in the case of the Amp unit, were estimated by the using the amide region in the 1D-NMR spectrum.

The NOE pattern, the $^3J_{\alpha(\gamma)\text{N}}$ values and the temperature coefficients for the α - and γ -NH obtained for the hexamer **IV.14a** containing Amp in the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration are reported in Table 2 and Figure 18. With the exception of the C-terminal Tyr residue, the other α -residues all presented a vicinal coupling constant value of 6, whereas the two γ -Amp units showed the very low value of 2.4. The temperature coefficients were in the range of 3.7-7.6 ppb/K, with the Ala-3 and Gly-5 amide protons showing the lower values (4.2 and 3.7, respectively). The strong $d_{\alpha\text{N}}$ NOE and the weak d_{NN} and $d_{\beta\text{N}}$ NOEs between Ala-1 and γ -Amp-2 might indicate the presence of a turn. Also the medium $d_{\beta\text{N}}$ NOEs between γ -Amp-2 and Ala-3, Ala-3 and γ -Amp-4, and γ -Amp-4 and Gly-5 might reflect the presence of turn elements. Unfortunately, the number of NMR constraints observed under these conditions was not sufficient to carry out a structure calculation.

Table 2: NOE pattern, $^3J_{\alpha(\gamma)\text{N}}$ values (all at 278 K) and temperature coefficients for the α - and γ -NH observed for **IV.14a** in methanol. Bar thickness reflects the NOE intensities as strong, medium and weak.

Residue	Ala-1	γ -Amp-2	Ala-3	γ -Amp-4	Gly-5	Tyr-6
$^3J_{\alpha(\gamma)\text{N}}$	6	2.4	6	2.4	6	7.8
$-\Delta\delta/\Delta T$ (ppb/K)	7.6	6.1	4.2	5.4	3.7	7.0
$d_{\text{NN}}(i; i+1)$	—				—	
$d_{\alpha\text{N}}(i; i+1)$	—	—	—	—	—	
$d_{\beta\text{N}}(i; i+1)$	—	—	—	—	—	

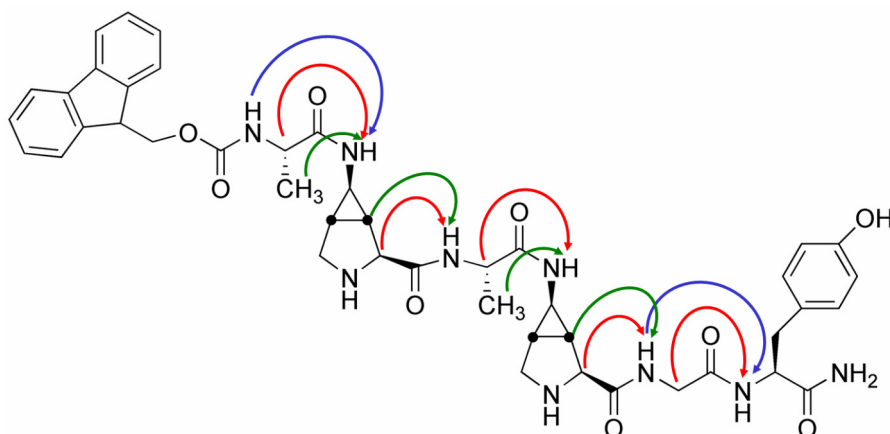


Figure 18: Summary of the interresidue NOE constraints observed for the peptide **IV.14a** in methanol at 278 K (red arrows: $d_{\alpha N}$ NOEs; green arrows: $d_{\beta N}$ NOEs; blue arrows: d_{NN} NOEs).

The NOE pattern, the $^3J_{\alpha(\gamma)N}$ values and the temperature coefficients for the α - and γ -NH obtained for the octamer **IV.14b** containing Amp in the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration are reported in Table 3. As observed for peptide **IV.14a**, the alanine residues presented the same value of 6.3 for the vicinal coupling constant, which is only slightly higher than the $^3J_{\alpha N}$ value of 6 found for the alanine residues in the hexamer. In contrast, the glycine coupling constant decreased from 6 to 5.3, and that of the three γ -Amp units could not be estimated from the 1D NMR spectrum. However, it should be noted that the NMR experiments of peptide **IV.14b** were recorded at 283 K and not at 278 K, as it was done for the hexamer **IV.14a**. The different temperature might explain the change in the values of the coupling constants between the hexamer and the octamer. The temperature coefficients were in the range of 3.7-8.0 ppb/K, with the Ala-3 and Gly-7 amide protons showing the lower values (4.0 and 3.7, respectively), while the N-terminal Ala-1 and the C-terminal Tyr-8 showing the higher values (8.0 and 7.0 ppb/K, respectively). Strong and medium $d_{\alpha N}$ (i ; $i+1$) NOEs were found in the region from Ala-1 to Gly-7, and weak d_{NN} (i ; $i+1$) NOEs were observed between Ala-3 and γ -Amp-4 and in the region from Ala-5 to Tyr-8. Medium $d_{\beta N}$ NOEs were found between Ala-5 and γ -Amp-6, and between Gly-7 and Tyr-8, whereas a weak one was found between Ala-1 and γ -Amp-2. Besides these sequential NOEs, also medium-range NOEs were detected: in particular, two medium $d_{\beta N}$ (i ; $i+3$) NOEs were

detected between γ -Amp-2 and Ala-5, and between γ -Amp-4 and Gly-7, together with two weak $d_{\text{NN}}(i; i+3)$ NOEs between γ -Amp-2 and Ala-5, and between Ala-5 and Tyr-8 (Figure 19). This $(i; i+3)$ periodicity might indicate the presence of an ordered structure, thus supporting the CD data.

Table 3: NOE pattern, $^3J_{\alpha(\gamma)\text{N}}$ values (all at 283 K) and temperature coefficients for the α - and γ -NH observed for **IV.14b** in methanol. Bar thickness reflects the NOE intensities as strong, medium and weak.

Residue	Ala-1	γ -Amp-2	Ala-3	γ -Amp-4	Ala-5	γ -Amp-6	Gly-7	Tyr-8
$^3J_{\alpha(\gamma)\text{N}}$	6.30	n.d.	6.30	n.d.	6.30	n.d.	5.30	7.60
$-\Delta\delta/\Delta T$ (ppb/K)	8.0	6.5	4.8	5.7	4.0	5.7	3.7	7.0
$d_{\text{NN}}(i; i+1)$			—		—			
$d_{\text{NN}}(i; i+3)$		—		—	—		—	
$d_{\alpha\text{N}}(i; i+1)$	—	—	—	—	—	—	—	
$d_{\beta\text{N}}(i; i+1)$	—				—		—	
$d_{\text{N}\gamma}(i; i+1)$			—		—			
$d_{\beta\text{N}}(i; i+3)$		—	—	—	—	—		

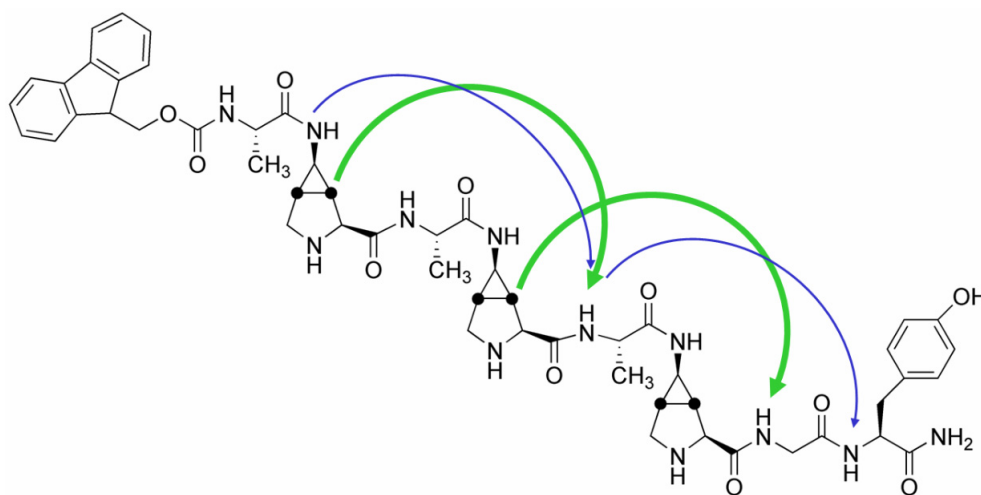


Figure 19: Summary of the medium-range interresidue NOE constraints observed for the peptide **IV.14b** in methanol at 283 K (green arrows: $d_{\beta\text{N}}(i; i+3)$ NOEs; blue arrows: $d_{\text{NN}}(i; i+3)$ NOEs).

Theoretical studies performed by Hofmann and co-workers have shown that structural similarities exist between acyclic α/γ -peptides and β -peptides [43]; moreover, these studies have also indicated that the most stable helix type for α/γ -alternating peptides in a polar medium is the helix-12, stabilized by hydrogen bonds between the carbonyl group of residue i and the amide proton of residue $i+3$, whereas the mixed 12/10-helix is less favored in a polar solvent [43]. In order to investigate which kind of energetically favored conformations could be assumed by the α/γ -peptide **IV.14b**, a systematic search was performed on a simplified analogue of this molecule, in which the N-terminal Fmoc-alanine was substituted by an acetyl group, whereas the C-terminal glycine and tyrosine were omitted (Figure 20). For practical reasons, however, the same residue numbers have been maintained as in peptide **IV.14b**. The conformational search was performed assuming that (i) bond lengths and bond angles are fixed, (ii) the amide bonds are in the *trans* configuration and can not rotate, (iii) the distance constraints corresponding to the two medium $d_{\beta\text{N}}$ ($i; i+3$) NOEs (2-4 Å) must be satisfied. All rotatable bonds in the molecule (Figure 20), with the exception of ϕ 1, were systematically rotated through 360° using a fixed increment of 30° (for Ψ 2, ϕ 5, Ψ 6, ϕ 9 and Ψ 10) and of 60° (for ϕ 3, Ψ 4, ϕ 7, and Ψ 8). 225 possible conformations were found and each of them was minimized by using the dielectric constant of methanol (30), in order to reproduce the same polar environment of the NMR experiments.

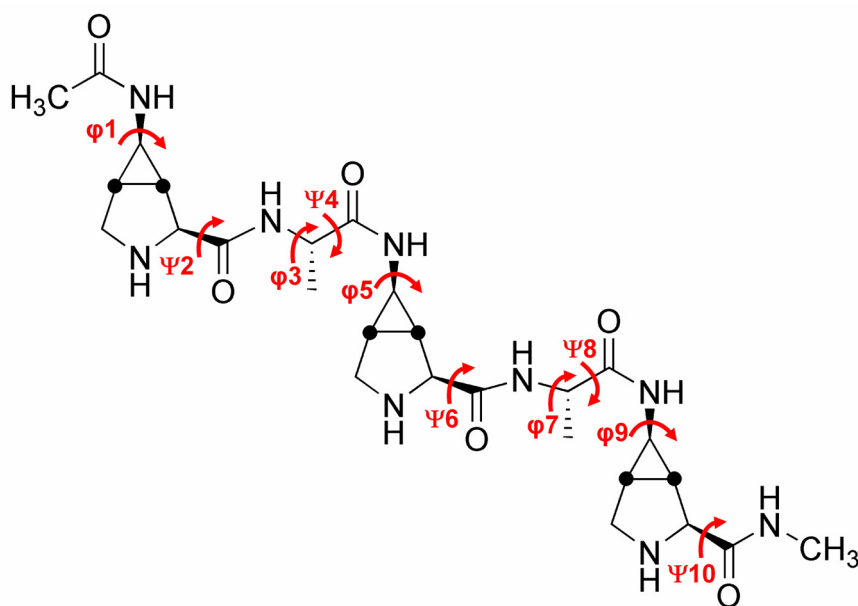


Figure 20: Structure of the **IV.14b** analogue used for the systematic search. Red arrows indicate all the backbone torsion angles in the structure; $\phi 1$ was not considered in the study.

Among all the minimized conformations, special attention was given to the conformer families presenting a periodicity in torsion angle values, which could indicate a helical fold. In particular, there was a group of structures stabilized by the presence of two hydrogen bonds involving 9 or 7 atoms, with a periodicity of $i, i+2$. In these conformers the values of torsion angles $\Psi 2$ were similar to those of $\Psi 6$ and $\Psi 10$, and at the same time $\phi 3$ was similar to $\phi 7$, $\Psi 4$ to $\Psi 8$, and $\phi 5$ to $\phi 9$ (structures A and B in Table 4).

Table 4: Backbone torsion angles ^a of the conformations A-D presented in Figure 21.

	γ -Amp-2		Ala-3		γ -Amp-4		Ala-5		γ -Amp-6	
	ϕ 1	Ψ 2	ϕ 3	Ψ 4	ϕ 5	Ψ 6	ϕ 7	Ψ 8	ϕ 9	Ψ 10
A	61	179	-83	160	56	176	-76	148	55	-175
B	60	-173	-72	83	74	-170	-69	81	98	-169
C	59	-175	-78	153	54	177	-60	-47	-149	-55
D	60	174	-78	161	53	176	-64	-40	-117	-171

^a Torsion angles are defined for Ala as follows : $\phi \rightarrow \text{C-N}^\alpha\text{-C}^\alpha\text{-C}$; $\Psi \rightarrow \text{N-C}^\alpha\text{-C-N}^\gamma$

Torsion angles are defines for Amp as follows: $\phi \rightarrow \text{C-N}^\gamma\text{-C}^\gamma\text{-C}^\beta$; $\Psi \rightarrow \text{N}^\text{R}\text{-C}^\alpha\text{-C-N}^\alpha$

One example is shown in Figure 21A, where the folded molecule presents 9-membered-hydrogen-bonded rings, in which the first bond takes place between the CO of Ala-3 and the NH of Ala-5, and the second one between the CO of Ala-5 and the NH at the C-terminus that corresponds to the amide of Gly-7 in **IV.14b**. Furthermore, the C-terminal NH group can interact also with the CO of γ -Amp-2, as the distance between the proton and oxygen atoms is of 2.13 Å.

The conformation represented in Figure 21B is characterized by 7-membered-hydrogen-bonded rings, with a hydrogen bond formed between the CO of γ -Amp-2 and the NH of γ -Amp-4, followed by a second one between the CO of γ -Amp-4 and the NH of γ -Amp-6.

The conformations with the lowest values of minimum energy were characterized by a periodicity in torsion angle, in which $\Psi_2 \sim \Psi_6$ and $\phi_3 \sim \phi_7$ (structures C and D in Table 4). They can also fold into helical structures, but no hydrogen bonds can be detected. Among these conformers, two of them resulted particularly interesting and are reported in panels C and D of Figure 21. In the first one (Figure 21C) the distance between the amide proton of Ala-5 and the oxygen of the CO group of Ala-3 is of ~ 3 Å, as well as the distance between the NH of Ala-3 and the CO that would correspond to the one of Ala-1 in **IV.14b**. In this case, it can be said that two long hydrogen bonds could be formed with a periodicity of $i, i+2$. The structure reported in Figure 21D is also characterized by two long hydrogen bonds, involving the CO of Ala-3 that can interact either with the NH of γ -Amp-6 or with the one of Ala-5. Interestingly, when these two conformers are minimized using a dielectric constant of 1 to simulate an apolar environment, the distances between the cited

atoms decrease, reaching values of ~ 1.8 Å. This means that in these conditions hydrogen bonds could be formed, and that the conformations are more stable in an apolar solvent than in methanol.

These data show that peptide **IV.14b** can adopt an ordered helical conformation. All possible conformations are in agreement with the NMR data and with the values of temperature coefficients of amide protons of the core structure of peptide **IV.14b**. As a matter of fact, these values indicate that these amide protons could be either involved in hydrogen bonds or solvent shielded, and both cases are possible in the conformations found by the systematic search. Moreover, the observation that some of the structures were stabilized in apolar solvents is in agreement with the data reported in literature for acyclic α/γ -peptides [42, 43], with the difference that the periodicity of hydrogen bond pattern we found is $i, i+2$ and not $i, i+3$ or the mixed $i, i+1/i, i+3$. Anyway, this is not surprising, as we are dealing with a by-cyclic γ -amino acid that is, of course, more constrained than an acyclic one.

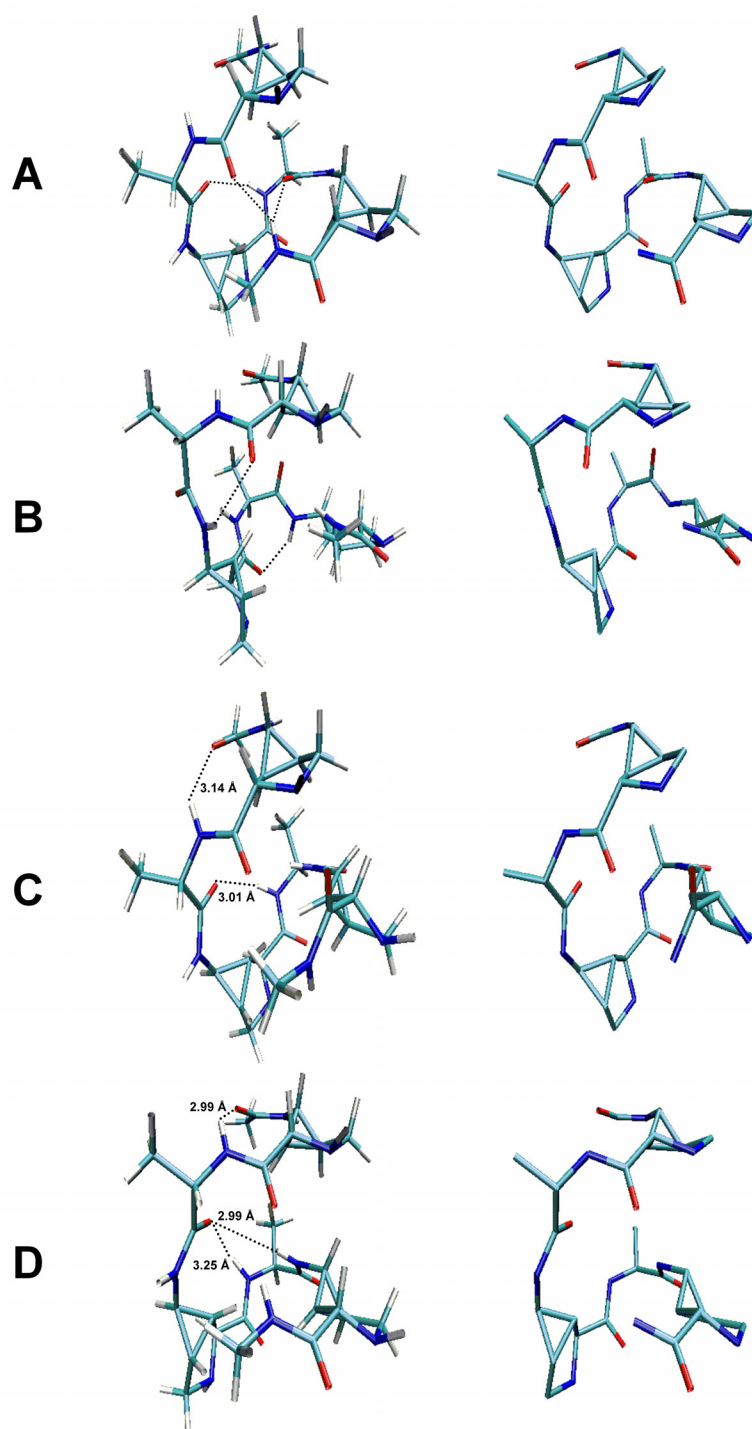


Figure 21: Examples of possible conformations adopted by the simplified analogue of peptide **IV.14b** in a polar environment, as found by the systematic search. Dotted black lines represent hydrogen bonds; in the case of long hydrogen bonds, the distances are also shown.

Summarizing, the hexamer **IV.14a** and the octamer **IV.14b** have displayed CD spectra similar to those reported for acyclic α/γ -peptides with moderate to high helix propensity [42]. Moreover, the NMR data observed for the octamer, together with the computational analysis, support the presence of an ordered structure with a well-defined periodicity that may be further stabilized by increasing the apolar character of the solvent.

The NMR data collected for peptide **IV.15a**, the hexamer containing the γ -Amp unit in the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration, are reported in Table 5 and Figure 22. The $^3J_{\alpha(\gamma)\text{N}}$ values from the 1D NMR spectrum at 283 K could be detected only for Ala-1 (6.5), Ala-3 (6.1) and Tyr-7 (7.7). The temperature coefficients were in the range of 4.0-7.5 ppb/K, with the Ala-3 and Gly-5 amide protons showing the lower values (4.2 and 4.0, respectively). The similar pattern of sequential NOEs in the region from γ -Amp-2 to γ -Amp-4 (strong to medium $d_{\alpha\text{N}}$ NOEs, medium to weak $d_{\beta\text{N}}$ NOEs and weak d_{NN} NOEs) suggest the formation of a secondary structure, probably two sequential bends. Also the weak d_{NN} NOE and the medium $d_{\beta\text{N}}$ NOE between γ -Amp-4 and Gly-5 might reflect the presence of a turn.

Table 5: NOE pattern, $^3J_{\alpha(\gamma)\text{N}}$ values (all at 283 K) and temperature coefficients for the α - and γ -NH observed for **IV.15a** in methanol. Bar thickness reflects the NOE intensities as strong, medium and weak.

Residue	Ala-1	γ -Amp-2	Ala-3	γ -Amp-4	Gly-5	Tyr-6
$^3J_{\alpha(\gamma)\text{N}}$	6.5	n.d.	6.1	n.d.	n.d.	7.7
$-\Delta\delta/\Delta T$ (ppb/K)	7.5	6.3	4.2	5.1	4.0	6.9
$d_{\text{NN}}(i; i+1)$		<div></div>				
$d_{\alpha\text{N}}(i; i+1)$	<div></div>	<div></div>	<div></div>	<div></div>	<div></div>	<div></div>
$d_{\beta\text{N}}(i; i+1)$	<div></div>	<div></div>	<div></div>	<div></div>	<div></div>	

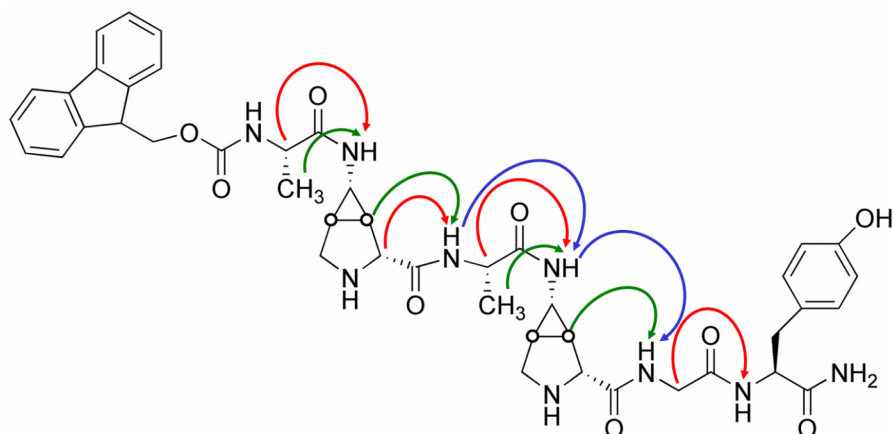


Figure 22: Summary of the interresidue NOE constraints observed for the peptide **IV.15a** in methanol at 283 K (red arrows: $d_{\alpha N}$ NOEs; green arrows: $d_{\beta N}$ NOEs; blue arrows: d_{NN} NOEs).

The presence of a structured peptide is supported from the splitting ($\Delta\delta$ of 0.3 ppm) of the diastereotopic α -protons of Gly-5 (Figure 23). This splitting has been usually observed in peptides, in which the glycine residue was locked in a turn edge, and the magnitude of the splitting has been also directly correlated with the turn stability in a number of structural studies [44].

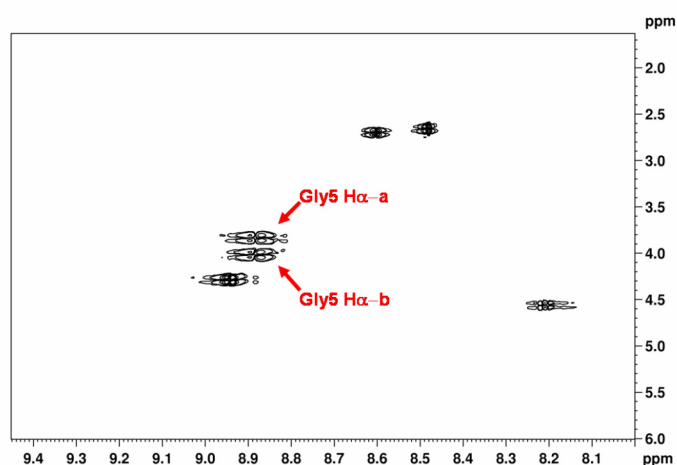


Figure 23: Fingerprint region of the COSY spectrum of peptide **IV.15a** (283 K) showing the diastereotopic α -protons of Gly-5.

Moreover, Gly-5, together with Ala-3, possesses a low temperature coefficient (4.0 ppb/K), suggesting that the amide proton might be involved in a hydrogen bond that would stabilize the turn conformation. Unfortunately, the number of NMR constraints observed under these conditions was not sufficient to carry out a structure calculation.

The NMR spectra of the octamer **IV.15b** containing the γ -Amp units with the same configuration like in the hexamer **IV.15a** were characterized by strong overlap under the conditions used for the other three peptides described above. Consequently, no further conformational data could be obtained for **IV.15b**.

In conclusion, the CD and NMR investigations presented here suggest that alternating α/γ -peptides containing the γ -Amp building block can adopt ordered secondary structures which depend on the configuration of the γ -amino acid. In fact, the γ -Amp in the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration might favor a helical backbone, whereas the γ -Amp in the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration might induce a bend-like structure.

IV.3.3 Incorporation of Amp in small α -peptides

Besides its application as a γ -amino acid, we also tested the Amp unit as proline analogue by using the enantiomers of N ^{γ} -Boc/N ^{α} -Fmoc-protected Amp for solid phase peptide synthesis (Figure 15). Depending on their stereochemistry, the Amp amino acids can be seen as analogues of ^LPro in the case of the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration, and of ^DPro in the case of the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration.

Proline is the amino acid with the highest propensity to occur into turn structures in proteins [45], due to the constraints of the pyrrolidine ring. Proline-containing turns stabilized by a 4 \rightarrow 1 hydrogen bond have been suggested to be natural templates for nucleating both β -hairpins and helical structures [46]. There different types of β -turns which vary in the dihedral angles of the residues *i*+1 and *i*+2: the most common types of β -turns are the type I and the type II, together with their mirror images I' and II'. The segment ^DPro-^LPro is known to stabilize the type II' turn and can favor helix propagation, but it can

also be used as an effective hairpin inducer in both cyclic and acyclic peptides [47]. Also ^DPro-Xxx has demonstrated to be an effective hairpin nucleator [48]. Instead, a consecution of II'-I turns can nucleate a right-handed helix, whereas the enantiomeric structure (II-I') can facilitate the formation of a left-handed helix [46].

The N^γ-Boc/N^α-Fmoc-protected Amp with the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration was used for preliminary studies on small peptides consisting of the repeats Gly- α -Amp. The chain assembly was performed manually on Rink amide resin with an initial loading of 0.7 mmol/g. The dipeptide Gly-Tyr was chosen as general C-terminal motif for the reasons discussed above for the α/γ -peptides. A single-coupling procedure was applied for the building block as well as for the natural amino acids in the presence of HBTU, HOBt and DIPEA; however, whereas the acylation steps involving α -Amp were carried out using a moderate molar excess (2.2 equiv.) and a longer reaction time (90 min), those involving the commercially available amino acids were carried out using a higher molar excess (4 equiv.). The Fmoc cleavage steps were accomplished with a mixture of DBU and HOBt in DMF, by treating the peptidyl-resin for short times (3-5 min). To monitor the peptide chain growth, a small amount of peptidyl-resin was subjected to TFA cleavage after the first α -Amp coupling, and the cleaved-off product was characterized by mass spectrometry, which gave the expected mass for the tripeptide. Therefore, after glycine coupling, it was decided to elongate the peptide chain with the second Gly- α -Amp repeat without further monitoring of the acylation reactions. Unfortunately, when a small-scale TFA cleavage was performed at this point, only the product corresponding to the peptide lacking the second α -Amp unit was detected by mass spectrometry. Thus, the pentapeptide Fmoc-Gly-Gly- α -Amp-Gly-Tyr-amide (**IV.16**) was totally cleaved from the resin and further characterized. Its CD spectrum in methanol was characterized by a negative band at 208 nm, followed by an almost equally intense positive signal at 229 nm, with a crossover at 219 nm (Figure 24A). This shape might reflect an exciton splitting [49], which takes place when two chromophores are close enough in the space to allow their dipoles interact. In the case of peptide **IV.16**, an exciton splitting can occur if the C-terminal tyrosine side chain and the N-terminal Fmoc group are spatially close, which would imply the presence of a central turn in the peptide.

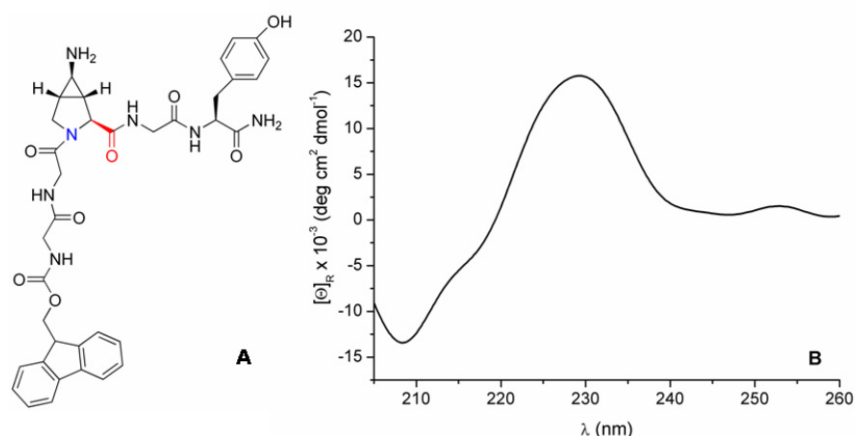


Figure 24: (A) Structure of peptide **IV.16** and (B) its CD spectrum (100 μ M in methanol).

In order to distinguish the contribution of α -Amp from that of additional favorable interactions between aromatic residues to induce and stabilize turns and hairpin-like structures, two pentapeptides differing only in the N-terminal residue were synthesized: peptides **IV.17a** contained Ala-1, whereas peptide **IV.17b** contained Phe-1 (Figure 25). The chain assembly was performed manually on Rink amide resin using a single-coupling procedure for the building block (2.5 equiv.) as well as for the natural amino acids (4 equiv.) in the presence of DIC and HOBt. The Fmoc cleavage steps were accomplished with a mixture of DBU and HOBt in DMF, by treating the peptidyl-resin for short times (3–5 min). Acetylation of the N-terminal residue was performed with acetic anhydride (10 equiv.) in DMF for 30 min. After final TFA cleavage, both crude peptides showed a purity near 90% and, thus, were not further purified. For a comparative study, the corresponding pentapeptides **IV.18a/b** containing L Pro instead of α -Amp were synthesized, as well as the pentapeptides **IV.19a/b** containing the α -Amp with the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration (Figure 23).

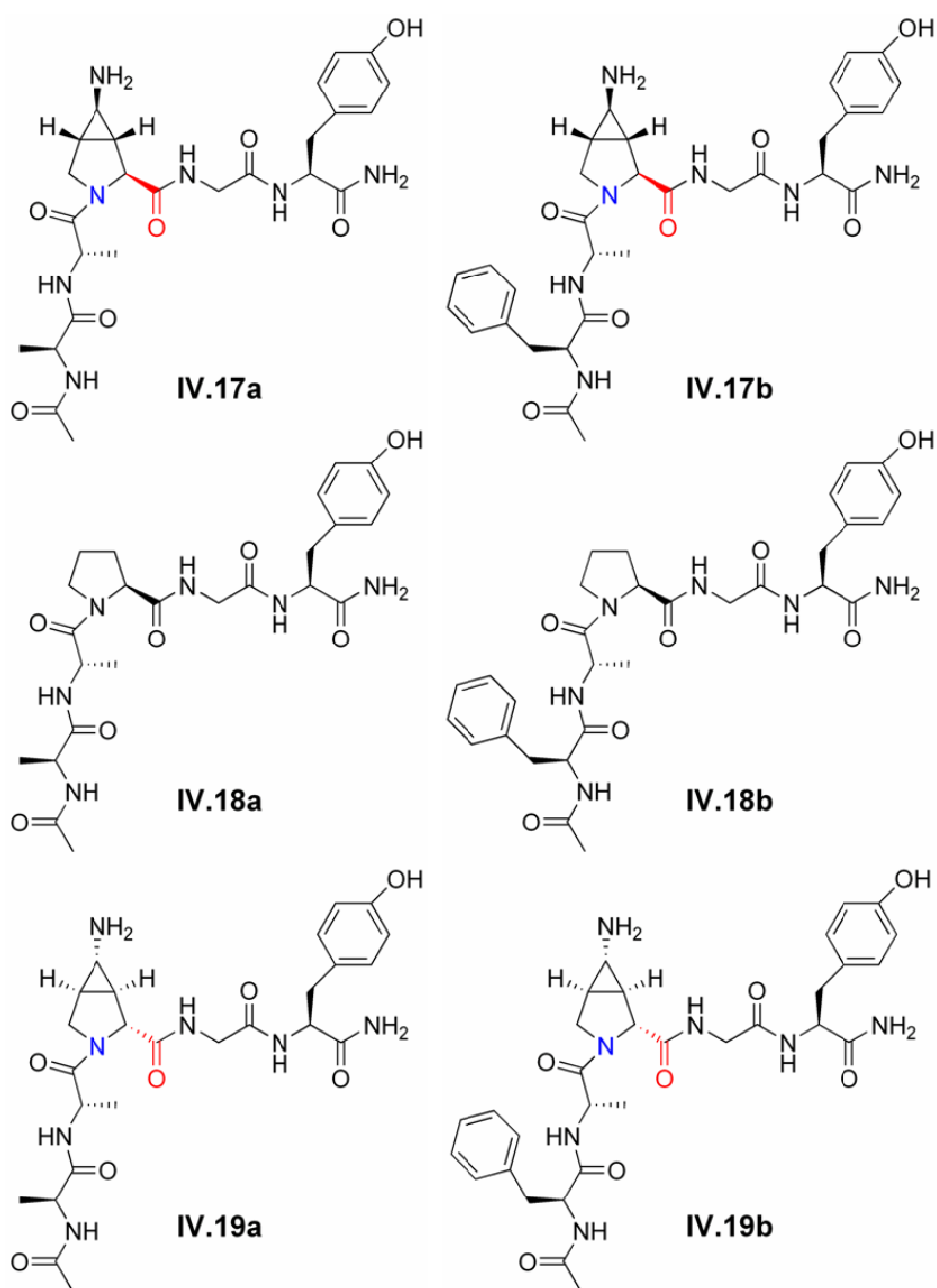


Figure 25: Structures of peptides IV.17a/b, IV.18a/b and IV.19a/b.

IV.3.4 Structural investigation of peptides containing α -Amp

The peptides were characterized by CD spectroscopy in methanol. The proline-containing sequences **IV.18a/b** were both characterized by negative ellipticity over the amide region, indicating that these two compounds adopt a disordered conformation (Figure 26A). In contrast, peptides **IV.17a/b** containing the α -Amp with the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration showed a different CD behavior that was also depending on the N-terminal residue. Indeed, with Ala-1 (**IV.17a**) the CD spectrum was positive between 215 and 240 nm, with a weak maximum at 228 nm, and negative below 215 nm, whereas with Phe-1 (**IV.17b**) the CD curve was characterized by two bands of opposite intensity at 222 nm (positive) and 203 nm (negative), with a crossover at 213 nm (Figure 26B). The spectrum of **IV.17a** is similar to the one obtained for polyproline II peptides, whose hallmark is a positive CD signal between 220 and 230 nm and a strong negative band near 200 nm. Instead, in the spectrum of peptide **IV.17b** the contribution of an exciton splitting could be detected, thus suggesting an interaction between the two aromatic residues Phe-1 and Tyr-5. These observations would support a stabilization of a turn element in the peptides **IV.17b** and **IV.16** by π - π contacts between aromatic side chains.

A different CD shape has been found by the peptides **IV.19a/b** containing the α -Amp with the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration (Figure 26C), which was characterized by a negative band below 195 nm, a maximum at 201-203 nm followed by a shoulder at 216 nm, and by a weak minimum near 240 nm. This CD shape can be attributed to the presence of turns. Interestingly, there is no evidence for an exciton splitting in the peptide containing Phe-1 (**IV.19b**). This suggests that the α -Amp with the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration favors a turn type, in which the side chains of Phe-1 and Tyr-5 are not able to couple.

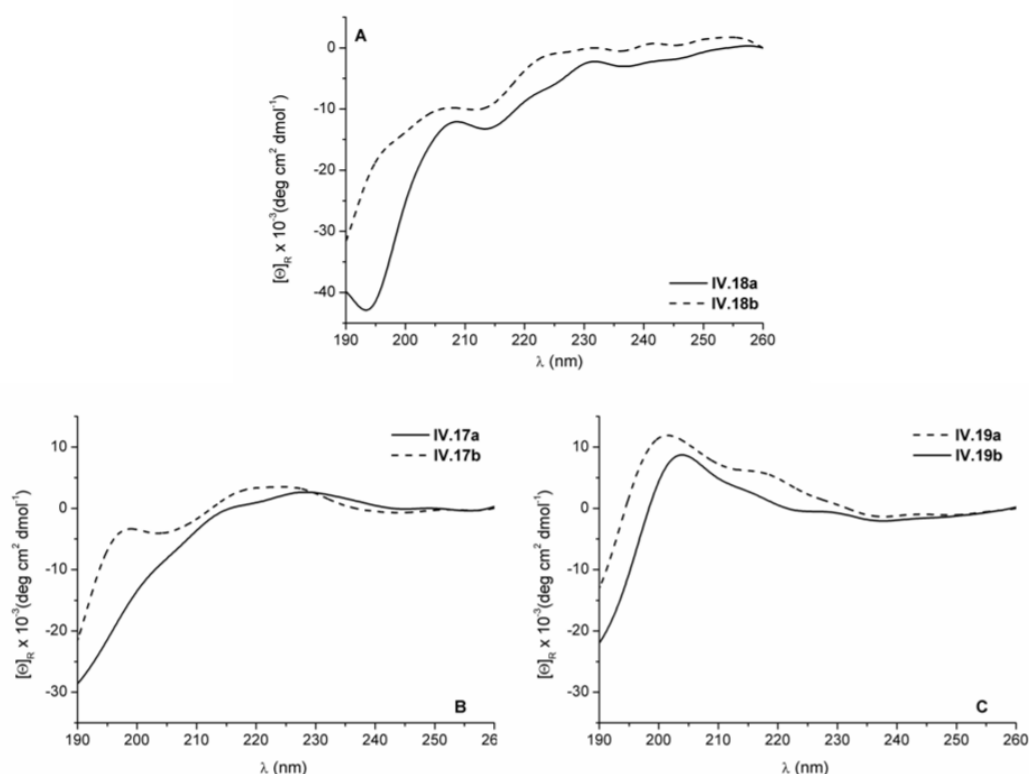


Figure 26: (A) CD spectra of peptides **IV.18a/b** (100 μ M in methanol). (B) CD spectra of peptides **IV.17a/b** (200 μ M in methanol). (C) CD spectra of peptides **IV.19a/b** (200 μ M in methanol).

In order to see if the conformational properties of the α -Amp-containing peptides were dependent on the solvent, the CD spectra of the peptides dissolved in water were recorded as well (Figure 27). Peptide **IV.17a** showed a blue shift of the positive band from 228 to 222 nm, but the shape remained the same as in methanol. In contrast, the exciton splitting observed for peptide **IV.17b** in methanol disappeared in water, suggesting that probably a more open conformation was present in water. For peptides **IV.19a/b**, CD spectra similar to those found in methanol were observed, although the CD intensity was reduced, probably because of a partial destabilization.

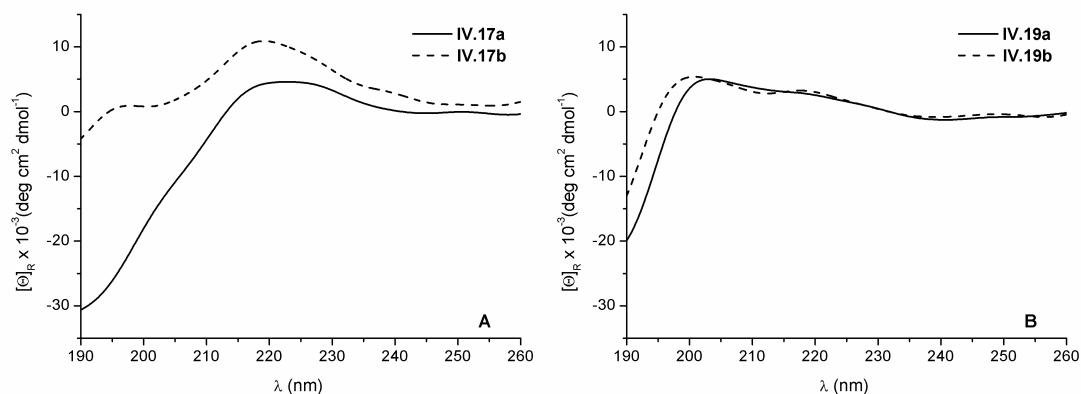


Figure 27: (A) CD spectra of peptides **IV.17a/b** (200 μ M in water). (B) CD spectra of peptides **IV.19a/b** (200 μ M in water).

In summary, these data suggest that the substitution of ^LPro by the α -Amp unit with the *S*-configuration at the α -carbon may lead to a stabilization of a polyproline II conformation at the expense of a flexible structure (Figure 26A-B). However, a turn structure could also be induced in conjunction with additional stabilizing effects like π - π interactions between aromatic residues (Figure 26B). On the other hand, the α -Amp unit with the *R*-configuration at the α -carbon seems to be superior in the stabilization of a turn motif: indeed, the induced conformation does not significantly depend on additional π - π interactions and remains moderately stable in water (Figures 26C and 27B). Moreover, the CD shape of the peptides containing this residue (**IV.19a/b**) corresponds to the one described for the class C' of linear α -peptides displaying Xxx-^DYyy or Xxx-Gly type II β -turns [50].

IV.4 Literature

1. Perk, J., Iavarone, A., Benezra, R. (2005) Id family of helix-loop-helix proteins in cancer, *Nat. Rev. Cancer*, 5, 603-614.
2. Berg, T. (2003) Modulation of protein-protein interactions with small organic molecules, *Angew. Chem. Int. Ed. Engl.*, 42, 2462-2481.
3. Berg, T., Cohen, S. B., Desharnais, J., Sonderegger, C., Maslyar, D. J., Goldberg, J., Boger, D. L., Vogt, P. K. (2002) Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts, *Proc. Natl. Acad. Sci. U S A*, 99, 3830-3835.
4. Gante, J. (1994) Peptidomimetics-Tailored Enzyme Inhibitors, *Angew. Chem. Int. Ed. Engl.*, 33, 1699-1720.
5. Giannis, A., Kolter, T. (1993) Peptidomimetics for Receptor Ligands-Discovery, Development, and Medical Perspectives, *Angew. Chem. Int. Ed. Engl.*, 32, 1244-1267.
6. Garber, D. W., Handattu, S. P., Datta, G., Mishra, V. K., Gupta, H., White, C. R., Anantharamaiah, G. M. (2006) Atherosclerosis and vascular disease: effects of peptide mimetics of apolipoproteins, *Curr. Pharm. Biotechnol.*, 7, 235-240.
7. Kovacs, J. M., Mant, C. T., Kwok, S. C., Osguthorpe, D. J., Hodges, R. S. (2006) Quantitation of the nearest-neighbour effects of amino acid side-chains that restrict conformational freedom of the polypeptide chain using reversed-phase liquid chromatography of synthetic model peptides with L- and D-amino acid substitutions, *J. Chromatogr. A*, 1123, 212-224.
8. Bang, D., Gribenko, A. V., Tereshko, V., Kossiakoff, A. A., Kent, S. B., Makhatadze, G. I. (2006) Dissecting the energetics of protein alpha-helix C-cap termination through chemical protein synthesis, *Nat. Chem. Biol.*, 2, 139-143.
9. Thundimadathil, J., Roeske, R. W., Guo, L. (2006) Conversion of a porin-like peptide channel into a gramicidin-like channel by glycine to D-alanine substitutions, *Biophys. J.*, 90, 947-955.
10. Sewald, N., Jakubke, H. D. in *Peptides: Chemistry and Biology*, Wiley-New York

11. Rotondi, K. S., Gierasch, L. M. (2006) Natural polypeptide scaffolds: beta-sheets, beta-turns, and beta-hairpins, *Biopolymers*, *84*, 13-22.
12. De Pol, S., Zorn, C., Klein, C. D., Zerbe, O., Reiser, O. (2004) Surprisingly stable helical conformations in alpha/beta-peptides by incorporation of cis-beta-aminocyclopropane carboxylic acids, *Angew. Chem. Int. Ed. Engl.*, *43*, 511-514.
13. Gracia, C., Isidro-Llobet, A., Cruz, L. J., Acosta, G. A., Alvarez, M., Cuevas, C., Giralt, E., Albericio, F. (2006) Convergent approaches for the synthesis of the antitumoral peptide, Kahalalide F. Study of orthogonal protecting groups, *J. Org. Chem.*, *71*, 7196-7204.
14. Martinell, M., Salvatella, X., Fernandez-Carneado, J., Gordo, S., Feliz, M., Menendez, M., Giralt, E. (2006) Synthetic ligands able to interact with the p53 tetramerization domain. Towards understanding a protein surface recognition event, *Chembiochem.*, *7*, 1105-1113.
15. Cabrele, C., Clerici, F., Gandolfi, R., Gelmi, M. L., Molinari, F., Pellegrino, S. (2006) Chemoenzymatic resolution of epimeric cis 3-carboxycyclopentylglycine derivatives, *Tetrahedron*, *62*, 3502-3508.
16. Kiewitz, S. D., Cabrele, C. (2005) Synthesis and conformational properties of protein fragments based on the Id family of DNA-binding and cell-differentiation inhibitors, *Biopolymer*, *80*, 762-774.
17. Pellegrino, S. Heterosubstituted carbocyclic amino acids: synthesis and stereocontrolled transformations, *Phd Thesis*.
18. Gonthier, E., Breinbauer, R. (2005) Solid-supported reagents and catalysts for the preparation of large ring compounds, *Mol. Divers.*, *9*, 51-62.
19. Forrest, S., McNamara, C. (2004) Id family of transcription factors and vascular lesion formation, *Arterioscler. Thromb. Vasc. Biol.*, *24*, 2014-2020.
20. Rothbard, J. B., Kreider, E., VanDeusen, C. L., Wright, L., Wylie, B. L., Wender, P. A. (2002) Arginine-rich molecular transporters for drug delivery: role of backbone spacing in cellular uptake, *J. Med. Chem.*, *45*, 3612-3618.
21. Appella, D. H., Christianson, L. A., Klein, D. A., Powell, D. R., Huang, X., Barchi, J. J., Jr., Gellman, S. H. (1997) Residue-based control of helix shape in beta-peptide oligomers, *Nature*, *387*, 381-384.

22. Langenhan, J. M., Gellman, S. H. (2004) Effects of alternative side chain pairings and reverse turn sequences on antiparallel sheet structure in beta-peptide hairpins, *Org. Lett.*, *6*, 937-940.
23. Potocky, T. B., Menon, A. K., Gellman, S. H. (2005) Effects of conformational stability and geometry of guanidinium display on cell entry by beta-peptides, *J. Am. Chem. Soc.*, *127*, 3686-3687.
24. Arvidsson, P. I., Ryder, N. S., Weiss, H. M., Gross, G., Kretz, O., Woessner, R., Seebach, D. (2003) Antibiotic and hemolytic activity of a beta2/beta3 peptide capable of folding into a 12/10-helical secondary structure, *Chembiochem.*, *4*, 1345-1347.
25. Gademann, K., Hane, A., Rueping, M., Jaun, B., Seebach, D. (2003) The fourth helical secondary structure of beta-peptides: the (P)-28-helix of a beta-hexapeptide consisting of (2R,3S)-3-amino-2-hydroxy acid residues, *Angew. Chem. Int. Ed. Engl.*, *42*, 1534-1537.
26. Seebach, D., Hook, D. F., Glattli, A. (2006) Helices and other secondary structures of beta- and gamma-peptides, *Biopolymers*, *84*, 23-37.
27. Schmitt, M. A., Weisblum, B., Gellman, S. H. (2004) Unexpected relationships between structure and function in alpha,beta-peptides: antimicrobial foldamers with heterogeneous backbones, *J. Am. Chem. Soc.*, *126*, 6848-6849.
28. Farrera-Sinfreu, J., Zaccaro, L., Vidal, D., Salvatella, X., Giralt, E., Pons, M., Albericio, F., Royo, M. (2004) A new class of foldamers based on cis-gamma-amino-L-proline, *J. Am. Chem. Soc.* *126*, 6048-6057.
29. Zhao, X., Jia, M. X., Jiang, X. K., Wu, L. Z., Li, Z. T., Chen, G. J. (2004) Zipper-featured delta-peptide foldamers driven by donor-acceptor interaction. Design, synthesis, and characterization, *J. Org. Chem.*, *69*, 270-279.
30. Brackmann, F., Schill, H., de Meijere, A. (2005) An access to 3,4-(aminomethano)proline in racemic and enantiomerically pure form, *Chemistry*, *11*, 6593-6600.
31. Brackmann, F., Colombo, N., Cabrele, C. & De Meijere, A. (2006) An Improved Synthesis of 3,4-(Aminomethano)proline and Its Incorporation into Small Oligopeptides, *Eur. J. Org. Chem.* *2006*, 4440-4450.

32. <http://classes.aces.uiuc.edu/AnSci312/Muscle/hlh%20structure%208bit.jpg>.
33. Baldauf, C., Gunther, R., Hofmann, H. J. (2003) Helix Formation and Folding in γ -Peptides and Their Vinylogues, *Helv. Chim. Acta.* 86, 2573-2588.
34. Hanessian, S., Luo, X., Schaum, R., Michnick, S. (1998) Design of Secondary Structures in Unnatural Peptides: Stable Helical γ -Tetra-, Hexa-, and Octapeptides and consequences of α -Substitution, *J. Am. Chem. Soc.*, 120, 8569-8570.
35. Hanessian, S., Luo, X., Schaum, R. (1999) Synthesis and folding preferences of γ -amino acid oligopeptides: stereochemical control in the formation of a reverse turn and a helix, *Tetrahedron Lett.*, 40, 4925-4929.
36. Hintermann, T., Gademann, K., Jaun, B., Seebach, D. (1998) γ -Peptides Forming Synthesis and Helical More Stable Secondary Structures than α -Peptides: NMR-Solution Structure of the γ -Hexapeptide Analog of H-(Val-Ala-Leu)-OH, *Helv. Chim. Acta.*, 81, 983-1002.
37. Seebach, D., Brenner, M., Rueping, M., Jaun, B. (2002) Gamma2-, gamma3-, and gamma(2,3,4)-amino acids, coupling to gamma-hexapeptides: CD spectra, NMR solution and X-ray crystal structures of gamma-peptides, *Chemistry*, 8, 573-584.
38. Vasudev, P. G., Shamala, N., Ananda, K., Balaram, P. (2005) C9 helices and ribbons in gamma-peptides: crystal structures of gabapentin oligomers, *Angew. Chem. Int. Ed. Engl.*, 44, 4972-4975.
39. Sharma, G. V., Jayaprakash, P., Narsimulu, K., Ravi Sankar, A., Ravinder Reddy, K., Radha Krishna, P., Kunwar, A. C. (2006) A left-handed 9-helix in gamma-peptides: synthesis and conformational studies of oligomers with dipeptide repeats of C-linked carbo-gamma4-amino acids and gamma-aminobutyric acid, *Angew. Chem. Int. Ed. Engl.*, 45, 2944-2947.
40. Sadowsky, J. D., Schmitt, M. A., Lee, H. S., Umezawa, N., Wang, S., Tomita, Y., Gellman, S. H. (2005) Chimeric (alpha/beta + alpha)-peptide ligands for the BH3-recognition cleft of Bcl-XL: critical role of the molecular scaffold in protein surface recognition, *J. Am. Chem. Soc.*, 127, 11966-11968.
41. Kritzer, J. A., Hodsdon, M. E., Schepartz, A. (2005) Solution structure of a beta-peptide ligand for hDM2, *J. Am. Chem. Soc.*, 127, 4118-4119.

42. Sharma, G. V., Jadhav, V. B., Ramakrishna, K. V., Jayaprakash, P., Narsimulu, K., Subash, V., Kunwar, A. C. (2006) 12/10- and 11/13-Mixed Helices in alpha/gamma- and beta/gamma-Hybrid Peptides Containing C-Linked Carbo-gamma-amino Acids with Alternating alpha- and beta-Amino Acids, *J. Am. Chem. Soc.*, *128*, 14657-14668.
43. Baldauf, C., Guenther, R., Hofmann, H. J. (2006) Helix Formation in alpha-gamma- and beta-gamma-Hybrid Peptides: Theoretical Insights into Mimicry of α - and β -Peptides, *J. Org. Chem.*, *71*, 1200-1208.
44. Butterfield, S. M., Waters, M. L. (2003) A designed beta-hairpin peptide for molecular recognition of ATP in water, *J. Am. Chem. Soc.*, *125*, 9580-9581.
45. Chou, P. Y., Fasman, G. D. (1977) Beta-turns in proteins, *J. Mol. Biol.*, *115*, 135-175.
46. Rai, R., Aravinda, S., Kanagarajadurai, K., Raghothama, S., Shamala, N., Balaram, P. (2006) Diproline templates as folding nuclei in designed peptides. Conformational analysis of synthetic peptide helices containing amino terminal Pro-Pro segments, *J. Am. Chem. Soc.*, *128*, 7916-7928.
47. Shankaramma, S. C., Athanassiou, Z., Zerbe, O., Moehle, K., Mouton, C., Bernardini, F., Vrijbloed, J. W., Obrecht, D., Robinson, J. A. (2002) Macrocyclic hairpin mimetics of the cationic antimicrobial peptide protegrin I: a new family of broad-spectrum antibiotics, *Chembiochem.*, *3*, 1126-1133.
48. Rai, R., Raghothama, S., Balaram, P. (2006) Design of a peptide hairpin containing a central three-residue loop, *J. Am. Chem. Soc.*, *128*, 2675-2681.
49. MacColl, R., Lam, I., Choi, C. Y., Kim, J. (1994) Exciton splitting in phycoerythrin 545, *J. Biol. Chem.*, *269*, 25465-25469.
50. Perczel, A. & Holl si, M. (1996) Circular dichroism and the conformational analysis of biomolecules, *Plenum Press, New York*.

V. Summary

The Id proteins (Id1-4) are negative regulators of several bHLH transcription factors, including the ubiquitous E factors and the tissue-specific myogenin-regulating factors. The Id HLH domains consist of two helices (a N-terminal helix-1 and a C-terminal helix-2, 16 residues each) connected by a loop of nine residues. The Id HLH sequences are highly conserved, whereas their N- and C-terminal extensions differ both in length and amino acid composition. Beside heterodimerization with the parent bHLH factors, the 134-residue long Id2 protein additionally interacts with the retinoblastoma protein (pRb) and, when overexpressed, Id2 inhibits the pRb-mediated cell-cycle arrest. As pRb is one of the most important tumor suppressors in the cells and Id2 is upregulated in several cancer types (e. g. neuroblastoma), Id2 represents an interesting target for cancer therapy based on the inhibition of protein-protein interactions.

The HLH region is essential for Id protein activity; however, the N- and C-termini are also important for the regulation of the inhibitory effects of the Id proteins on DNA transcription. Here we have presented the synthesis and conformational characterization of peptides derived from point mutations and N-/C-terminal truncations of Id2. The helix character of the HLH domain (residues 36-76) could be reduced upon substitution of Met-39/-62 and Cys-42 with Nle and Ser, respectively, suggesting that these side chains play a role in the HLH fold. The largest sequence that could be obtained by stepwise solid-phase peptide synthesis (SPPS) with Fmoc strategy spanned the entire HLH motif (with Cys-42 replaced by Ser) and part of the C-terminus (residues 77-110). This 75-residue long fragment was less helical than the isolated HLH domain and had propensity to aggregate, which was correlated with the presence of the flanking residues C-terminal to helix-2. By CD analysis of an equimolar mixture of the sequence 36-110 with the N-terminus 1-35, noncovalent interactions between the two peptides were detected, which, however, changed upon aging. In contrast, the mixture of the HLH sequence 36-76 with the N-terminus was characterized by a stabilized helix structure that was maintained also upon aging. Presumably, the N-terminal region interacted with the folded HLH motif in a specific

manner, whereas only unspecific, weak contacts occurred with the partly unfolded HLH domain and/or the immediate flanking residues 77-110.

Id2 possesses a C-terminal nuclear export signal (NES, residues 103-119) that is responsible for the transport of Id2 from the nucleus to the cytoplasm. C-terminal Id2 fragments containing the NES sequence are either disordered or aggregation-prone. To study the conformational properties of the isolated NES region, we have synthesized the Id2 segment 103-124. The latter was insoluble in water and only temporarily soluble in water/alcohol mixtures, in which it displayed a β -sheet conformation that quickly precipitated in amorphous aggregates. Introduction of a positively charged N-terminal tail bearing three lysines prevented aggressive precipitation and led to aggregates consisting of long β -sheet fibrils that bound thioflavin T. On the other hand, when the positively charged tail was introduced at the C-terminus, the peptide was able to stay in solution for a longer time and to adopt a helical conformation. Transmission electron micrographs of this C-terminally Lys-tagged NES analogue showed a highly ordered morphology consisting of aligned fibrils. These presumably originated from an antiparallel organization of helix monomers in a bilayer stabilized by favorable helix-dipole as well as hydrophobic side-chain interactions. It was also shown that the two Lys-tagged NES peptides could associate in helical structures and form aggregates with a morphology similar to that observed for the C-terminally Lys-tagged peptide alone.

The last part of this PhD work has been focused on the synthesis of peptidomimetics as potential modulators of protein-protein interactions. To this purpose, the unnatural amino acid 3-carboxy-cyclopentylglycine (**Cpg**) has been used. This scaffold presents two carboxylic groups separated by four carbon atoms, three of them are constrained in a cyclopentyl ring. Hence, one application of **Cpg** is the N,N-linkage of two peptide fragments to form covalent homodimers. This has been done for Id helix-2 fragments, and the effect of both enantiomers of Cpg on the conformation has been investigated by CD spectroscopy. It was found that the Cpg-linked dimer of the Id1 helix-2 fragment 91-101 could interact with the native Id1 HLH motif, leading to helix stabilization or destabilization of the latter depending on the Cpg configuration. The Cpg-containing peptide was also tested on human vascular smooth muscle cells (VSMC) presenting a synthetic phenotype: preliminary data suggest that the peptide might influence the

inhibitory activity of the Id proteins, as it was able to reduce VSMC proliferation and migration, while inducing differentiation. However, its potency was only in the micromolar range and thus needs to be improved, together with the cellular uptake and, of course, proteolytic stability.

A second unnatural amino acid, 3,4-(aminomethano)proline (**Amp**), has been used for the synthesis of peptidomimetics. **Amp** is not only a derivative of proline, but it is also a bicyclic γ -amino acid. In this work α/γ -alternating peptides (hexamers and octamers) containing both enantiomers of **Amp** have been synthesized by manual SPPS and characterized by CD and NMR spectroscopy. The results suggest that α/γ -peptides containing the γ -Amp building block adopt ordered secondary structures which depend on the configuration of the γ -amino acid. Indeed, the γ -Amp in the (2*S*, 1'*R*, 3*R*, 4*R*)-configuration favors a helical backbone, as supported also by computational analysis, whereas the γ -Amp in the (2*R*, 1'*S*, 3*S*, 4*S*)-configuration might induce a bend-like structure.

The Amp amino acid has been used also as a proline analogue, and in order to distinguish the contribution of α -**Amp** from that of additional favorable interactions between aromatic residues to induce and stabilize turns and hairpin-like structures, two pentapeptides have been prepared, both containing a C-terminal Tyr residue while differing in the N-terminal residue, which was either an alanine or a phenylalanine. CD spectroscopy data suggest that the substitution of ^LPro by the α -**Amp** unit with the *S*-configuration at the α -carbon may lead to a stabilization of a polyproline II conformation at the expense of a flexible structure. However, a turn structure could also be induced in conjunction with additional stabilizing effects like π - π interactions between the aromatic residues at the N- and C-ends. On the other hand, the α -**Amp** unit with the *R*-configuration at the α -carbon seems to be superior in the stabilization of a turn motif: indeed, the induced conformation does not significantly depend on additional π - π interactions and remains moderately stable in water.

In conclusion, the synthetic and conformational studies on Id2 protein fragments and on oligomers containing new structurally constrained amino acid building blocks have provided interesting data concerning (i) the usefulness of the solid phase methodology for the preparation of large polypeptides related to the Id2 protein, (ii) the secondary structure propensity of the Id HLH domain and the influence of the flexible flanking regions, (iii) the

ability of a short Id2 fragment to form highly ordered fibrillar structures, and *(iii)* the development of novel backbone-modified peptides with ordered conformations, which might become valuable tools in the design of peptidomimetics and foldamers for biological applications. All together, these data will help planning the future work aiming to the total synthesis and biophysical characterization of the Id2 protein and to the development of artificial modulators of the Id2 protein inhibitory activity.

VI. Experimental part

VI.1 Materials

The N- α -Fmoc protected amino acids were purchased by MultiSynTech (Witten, Germany) and Novabiochem (Merck Biosciences GmbH, Schwalbach/Ts., Germany). Rink amide 4-methylbenzhydrylamine (MBHA) resin and HBTU were from MultiSynTech. Wang resin and 2-chloro-tritylchloride resin were from Novabiochem.

HOBt, DIPEA, DBU, TFA, triisopropylsilane (TIS), trimethylsilylbromide (TMSBr), α -cyano-4-hydroxycinnamic acid and sinapinic acid were from Fluka (Taufkirchen, Germany).

The following peptide-synthesis-grade reagents were purchased by Biosolve (Valkenswaard, The Netherlands): piperidine, 1-methyl-2-pyrrolidinone (NMP), DMF and diethylether. HPLC-grade acetonitrile (ACN) and TFA for UV spectroscopy were from Biosolve. Methanol- d_3 and D_2O for NMR measurements were from Deutero GmbH (Kastellaum, Germany). N,N-diisopropylcarbodiimide (DIC) was from Acros (Geel, Belgium). Sodium hydrogenphosphate and dihydrogenphosphate, acetic anhydride, acetic acid and 1,2-ethanedithiol (EDT) were obtained from Merck (Darmstadt, Germany).

VI.2 Methods

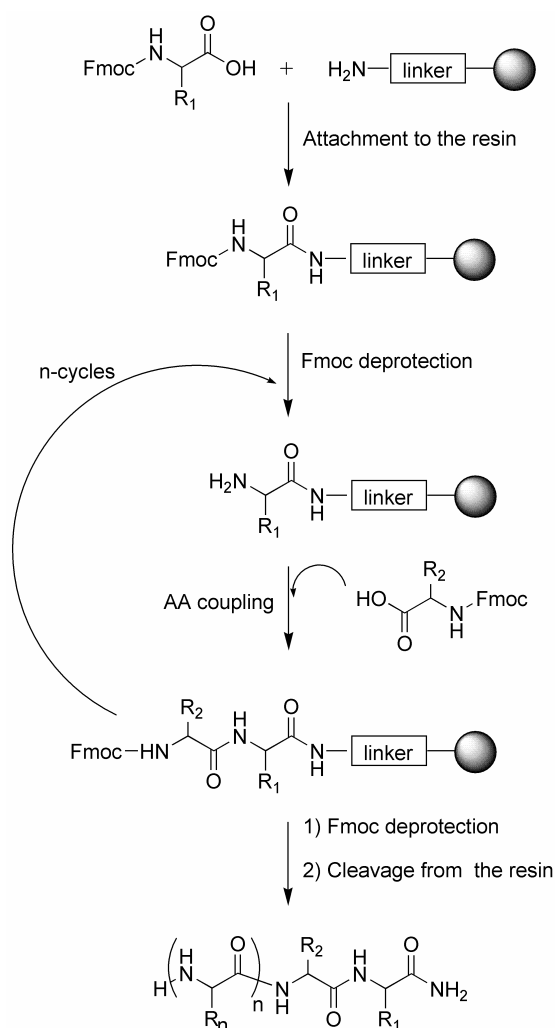
VI.2.1 Solid phase peptide synthesis (SPPS)

The SPPS methodology was introduced by Bruce R. Merrifield in the early sixties [1], and since then it has become the method of choice for the preparation of native and chemically modified peptides containing up to 50 residues. Also larger polypeptides and small proteins with more than 100 residues have been synthesized by stepwise SPPS [2].

The concept of SPPS is based on the build-up of the peptide backbone from the C-terminal to the N-terminal end, while the C-terminus is anchored to a solid support through a linker

(Scheme 1). Each amino acid is orthogonally protected at the α -amino group and, when required, at the side chain, in order to suppress polymerization. Two temporary protecting groups are mostly used for the N-protection: the 9-fluorenylmethyloxycarbonyl (Fmoc) and the *tert*-butyloxycarbonyl (Boc). The Fmoc group is base labile and usually removed with 20-40% piperidine in DMF, whereas the Boc group is acid labile and can be released with 50% TFA. After attachment of the N- α -protected C-terminal amino acid to the resin, the temporary N-protection is cleaved and the next N- α -protected and activated amino acid is coupled to the first one. The deprotection/coupling cycles are repeated till the sequence is completed. Finally, the crude product is cleaved from the solid support and simultaneously side-chain deprotected.

Scheme 1: SPPS of peptide amides using the Fmoc chemistry.



The first solid support used by Merrifield was based on polystyrene crosslinked with 1% m-divinylbenzene and on a chlorobenzyl linker. This resin is known as the Merrifield resin and has been in the past the standard resin for the synthesis of peptide acids by Boc chemistry. However, strong acids like HF are required to remove the peptide chain from the solid support. As in the presented PhD work the SPPS has been performed in combination with the Fmoc chemistry, resin linkers which are labile under milder acidic conditions have been chosen (Figure 1). For the preparation of peptide amides, a polystyrene/1%-divinylbenzene copolymer bearing the Rink amide MBHA linker has been used. As the MBHA linker is easily acylated, no preloading is necessary and the resin can be directly

used for the automatic synthesis. By contrast, the two resins applied in the synthesis of peptide acids, the 4-hydroxybenzyl (Wang) and 2-chloro-tritylchloride (trityl) resins, must be manually preloaded and capped before starting the automatic peptide chain assembly. The Wang linker is labile in concentrated TFA (>80%), whereas the trityl linker is labile already in the presence of 1% TFA. Therefore, the trityl resin can be used to prepare fully-protected peptide acids which may be further modified at the C-terminus in solution. [3].

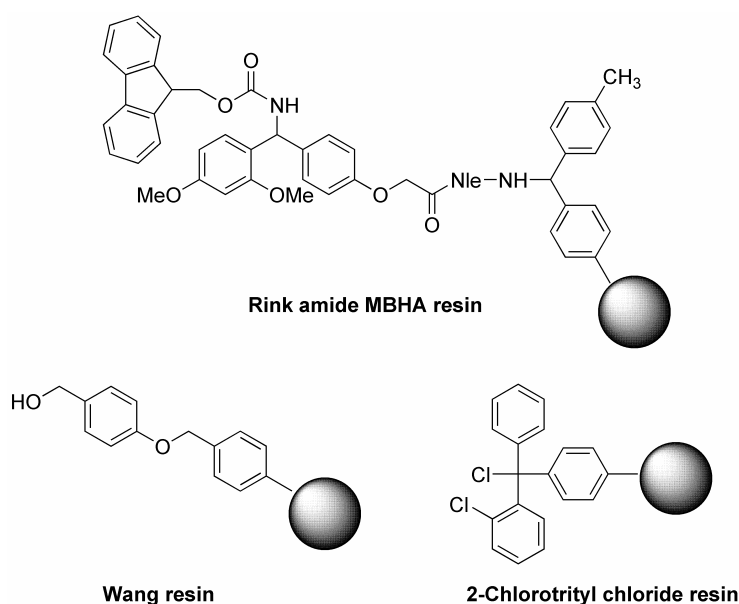


Figure 1: Chemical structures of the three different resin linkers used in this work for SPPS (the Rink amide MBHA resin is purchased in the Fmoc-protected form).

VI.2.2 Circular dichroism (CD) spectroscopy

CD spectroscopy is a very convenient method to study the secondary structure of peptides and proteins in solution [4]. The CD signal is the consequence of the absorption of polarized light by a chiral medium. A monochromatic plane-polarized light beam can be described as the sum of two vectors representing two in-phase circularly polarized beams of equal amplitude ($E_L - E_R$) but spinning in opposite directions with a frequency identical to the one of the light. When linearly polarized light passes through a chiral sample or through

a symmetrical chromophore placed in an asymmetrical environment, a different absorption for the left and the right circularly polarized light occurs. The left and right components indeed possess different extinction coefficients (ϵ_L and ϵ_R) and also different reflection indexes (n_L and n_R). Dichroism is the different absorption of left (A_L) and right (A_R) circularly polarized light as a function of the frequency, and can be described by the following equation 1:

$$\Delta\epsilon = (\epsilon_L - \epsilon_R) = \frac{A_L - A_R}{l \cdot c}$$

(equation 1)

where $\Delta\epsilon$ is the differential molar extinction coefficient in $M^{-1} \text{ cm}^{-1}$ or $10^3 \text{ cm}^2 \text{ mol}^{-1}$, l the path length, c the concentration in g ml^{-1} and A the absorbance.

The unit of CD is the ellipticity Θ , which is related to the absorbance by the following equation 2:

$$\Theta = 32,98 \cdot (A_L - A_R)$$

(equation 2)

The ellipticity Θ can be converted into the mean residue molar ellipticity $[\Theta]_R$, which is given in $\text{deg cm}^2 \text{ dmol}^{-1}$, as follows (equation 3):

$$[\Theta]_R = \frac{100 \cdot \Theta}{l \cdot c \cdot n}$$

(equation 3)

where c is expressed in M , l in cm , and n is the number of residues.

CD spectra show the ellipticity as a function of the wavelength and are usually recorded in the far UV range from 190 to 260 nm, where the peptide bond transitions can be

detected. Absorptions are observed around 190 nm as a result of the peptide bond $\pi-\pi^*$ transitions and at 210-220 nm due to the $n-\pi^*$ transitions. The most common secondary structure elements found in peptides and proteins, α -helix, β -turn and random coil, display characteristic patterns for these transitions which are reflected in characteristic CD spectra, as shown in Figure 2. The latter can be used as references for the conformational analysis of peptides and proteins, for which the structure is not yet known.

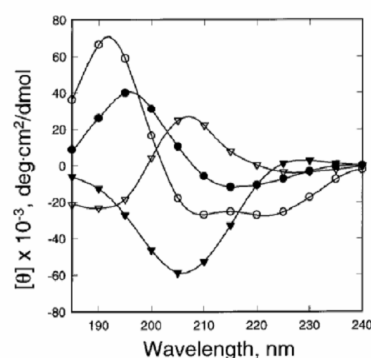


Figure 2: Characteristic patterns for the CD spectra of the most common secondary structures of peptides and proteins: \triangleright β -turn, \blacktriangle random coil, \bullet β -sheet, \circ α -helix [4].

VI.2.3 Nuclear magnetic resonance (NMR) spectroscopy

One of the most common and useful tools to investigate the secondary and tertiary structure of peptides and proteins in solution is the NMR spectroscopy [5]. The NMR technique is unique among the ones available for three-dimensional structure determination of peptides and proteins at atomic resolution, since the sample can be investigated in solution. Moreover, solution parameters such as temperature, pH and salt concentration can be adjusted to mimic physiological conditions, a fact that, especially for bioactive molecules, is particularly important. NMR studies can also highlight the presence of conformational equilibria and the dynamics of the folding.

Important structural insights can be obtained from both mono- and two-dimensional NMR spectra. However, in the case of biopolymers, 2D NMR techniques are required for the complete identification and assignment of the amino acid sequence.

A dispersion of the chemical shift values is the first indication of the presence of spatial folding in the polymer chain. When a polymer has a defined conformation in solution, a proton in the sequence will have a different microenvironment compared to those in a random coil, and it will be also different from the same proton contained in an identical residue type but at a different point of the sequence. This fact leads to different specific chemical shift values compared to those for an unfolded structure and to a good dispersion of the signals.

The vicinal scalar coupling constants 3J are one of the most important NMR parameters, because their magnitude depends on the dihedral angle between vicinal hydrogens. The secondary structure of peptides is characterized by distinct torsion angles along the backbone. The well known Karplus relationship [6] between the vicinal spin-spin coupling constants 3J of two protons and the corresponding torsion angle Θ is as follows (equation 4):

$$^3J = A \cos^2 \Theta + B \cos \Theta + C$$

(equation 4)

This equation has been empirically defined for α -polypeptides on the base of crystallographic data for the dihedral angles φ and of NMR data for the $^3J_{\alpha N}$ coupling constants [7], leading to the empirical coefficients A, B and C reported in equation 5, where $\Theta = |\varphi - 60^\circ|$:

$$^3J_{\alpha N} = 6.4 \cos^2 \Theta - 1.4 \cos \Theta + 1.9$$

(equation 5)

Generally, values of $^3J_{\alpha\text{N}} < 6$ Hz indicate helical conformations, whereas bigger values are referred to unordered structures and extended conformations. Values of $^3J_{\alpha\text{N}}$ of about 9 Hz are indicative of β -sheets.

The presence of hydrogen bonds can be detected from mono-dimensional NMR spectra, as the hydrogen-bonded NHs are less sensitive to perturbations such as changing of temperature, concentration or solvents, than the NHs that are exposed to the solvent. The determination of the temperature coefficients ($\Delta\delta/\Delta T$) for the amide protons is based on the observations that, upon warming, NHs involved in hydrogen bonds display shifts at a smaller extent than the NHs that are exposed to the solvent [8]. Generally, values of $-\Delta\delta/\Delta T > 4.5$ ppb K⁻¹ indicate solvent accessible amide protons, while values of $-\Delta\delta/\Delta T < 4.5$ ppb K⁻¹ are considered an indication of hydrogen bonds or buried protons.

2D NMR experiments for peptide and protein analysis consist usually of correlation spectroscopy (COSY) experiments, together with total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). A COSY spectrum represents a map of the complete scalar coupling network in a macromolecular structure, where a cross peak indicates the correlation between two nuclei A and B that are separated by not more than three covalent bonds. More information can be obtained by the TOCSY experiment, in which also scalar correlations between nuclei with shared coupling partners are visible. Cross signals are formed between all nuclei in a spin system, and in the case of peptides each amino acid has a particular spin-system pattern that allows the identification of every residue type in the sequence. With the help of these two experiments, it is possible to assign completely all the chemical shifts of the protons of a peptide. Further, NOESY experiments allow the sequential assignment of the residues in the peptide chain. This type of experiment is based on the dipolar spin-spin coupling between two spins, also known as nuclear Overhauser effect (NOE), which takes place between two nuclei that are close in the space up to 5 Å. A cross peak indicates in this case, that two non-covalently bonded nuclei A and B are close enough to interact or to exchange. NOE cross peaks are usually defined as sequential distances when they are between backbone protons, or between a backbone proton and a β proton of residues that are nearest neighbors in the sequence. Medium range distances are all non-sequential signals between residues within a segment

of five consecutive residues. Long-range distances are between the backbone protons of the residues that are at least six positions away in the sequence.

Medium and long range distances can be connected to secondary structure elements: for example, medium range contacts in helices are usually $i; i+2$, $i; i+3$, and $i; i+4$. Tight turns show $i; i+3$ distances, whereas β -sheets do not present medium range distances, since the polypeptide segments are almost fully extended. Short sequential and medium range distances that are found in known α -peptide conformations are listed in Table 1.

Table 1: Short sequential and medium range distances in some common secondary structures of α -peptides.

Distance	α -helix	3_{10} -helix	β -sheet
$d_{\alpha N}$	3.5	3.4	2.2
$d_{\alpha N}(i; i+2)$	4.4	3.8	
$d_{\alpha N}(i; i+3)$	3.4	3.3	
$d_{\alpha N}(i; i+4)$	4.2		
d_{NN}	2.8	2.6	4.2-4.3
$d_{NN}(i; i+2)$	4.2	4.1	
$d_{\beta N}$	2.5-4.1	2.9-4.4	3.2-4.7
$d_{\alpha N}(i; i+3)$	2.5-4.4	3.1-5.1	

VI.2.4 Computational studies

An important tool for the conformational analysis by a computational approach is the conformational search, which means to identify the “preferred” conformations of a molecule in solution [9]. Once a range of conformations is generated, an energy function is used to discriminate between them. As only those conformations that are located in

minimum points on the energy surface are considered with this approach, energy minimization methods play a crucial role in conformational analysis.

It is important to recognize the difference between a conformational search and a molecular dynamics simulation. The conformational search is concerned exclusively with locating minimum energy structures, whereas the simulation generates an ensemble of states that includes structures which are not necessarily at energy minima.

There are different conformational search methods. One of these is the systematic search that explores the conformational space by making regular and predictable changes to the conformation. In peptides and proteins, it is assumed that bond lengths and bond angles are fixed, and that the amide bond is adopting a *trans* configuration, therefore, only two torsion angles (ϕ and ψ , Figure 3) can vary. In the systematic search each of these rotatable bonds is systematically changed through 360° using a fixed increment. Every conformation so generated is then minimized to derive the associated minimum energy conformation. The search stops when all possible combinations of torsion angles have been generated and minimized.

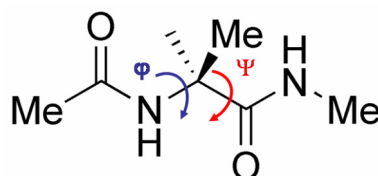


Figure 3: Representation of torsion angles for an alanine residue.

The energy is then function of the two variable considered, in this case the ϕ and ψ angles, and it can be represented in contour diagrams known as Ramachandran plots [10] (Figure 4), which show that amino acids are restricted to a limited range of conformations. In these plots two regions are particularly important, corresponding to α -helix and β -sheet structures.

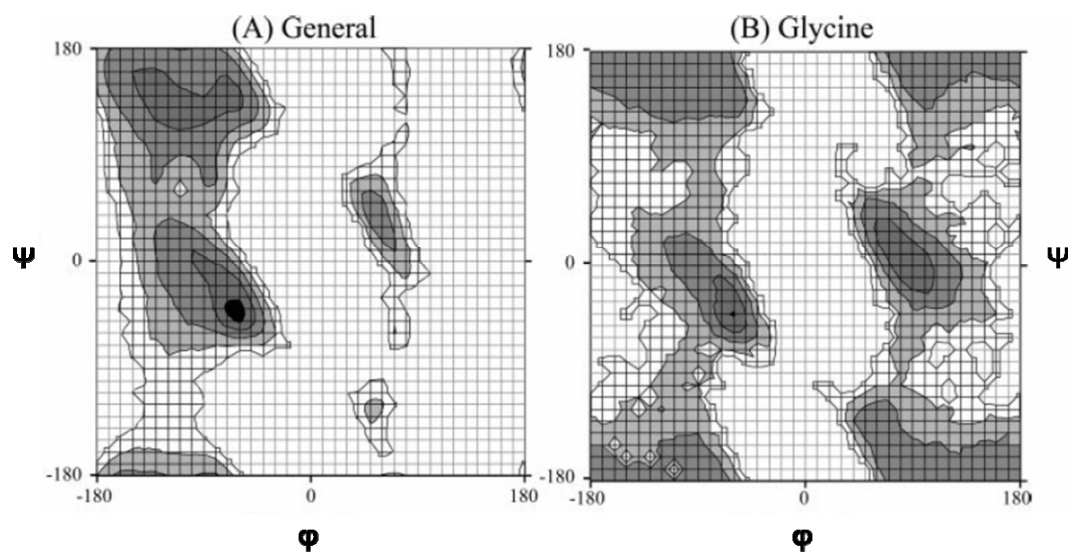


Figure 4: Ramachandran plots of L-amino acids in *trans* configuration. (A) Non-glycine and non-proline residues. (B) Glycine [11].

Systematic search can be done also taking into account particular constraints, like, for example, distance constraints coming from NMR analysis, to better understand if a peptide can fold or not into an ordered conformation. In this case the result of the search will consider only the conformations that reach a minimum energy and that, at the same time, do not violate the introduced distance constraints. However, even if important information on the secondary structure of a protein can be provided by a systematic search, this method cannot be used to characterize the dynamics of the folding process.

VI.3 General procedures for peptide synthesis

VI.3.1 Peptide chain assembly by automated SPPS

All peptides presented in chapter II, III, and the Id helix analogues to be coupled to the Cpg amino acid (chapter IV), were prepared on a 0.018 mmol scale by solid-phase methodology using the automatic synthesizer *Syro-I* (*MultiSynTech*). Double couplings (2×40 min) were performed with N- α -Fmoc-amino acids (5 equiv.) activated *in situ* with

HBTU (5 equiv.), HOBt (5 equiv.) and DIPEA (10 equiv.). Fmoc cleavage was accomplished by treating the peptidyl-resin with 40% piperidine in DMF/NMP (80:20 v/v) for 3 min, followed by a second treatment with 20% piperidine for 10 min.

N-terminally truncated Id2 analogues were acetylated at the N-terminus by acetic anhydride (10 equiv.) and DIPEA (10 equiv.) in DMF for 20 min. C-terminally truncated Id2 analogues were synthesized as C-terminally amidated peptides using the Rink amide MBHA resin (loading: 0.6 mmol g⁻¹). The Id2 segments containing the entire C-terminus were prepared starting from the Wang resin (loading: 1.0 mmol g⁻¹), which was first manually preloaded with Fmoc-Gly-OH (5 equiv.), HOBt (5 equiv.), DIC (5 equiv.) and DIPEA (5 equiv.). The new loading was 0.7 mmol g⁻¹, as determined by the UV absorption of the fluorenylmethyl-piperidine adduct at 300 nm ($\epsilon = 7800 \text{ M}^{-1} \text{ cm}^{-1}$). The remaining free hydroxybenzyl groups were then acetylated.

Peptide **II.18** was synthesized starting from the trityl resin (loading: 1.6 mmol g⁻¹), which was first manually preloaded with Fmoc-Gln(Trt)-OH (1 equiv.) and DIPEA (2 equiv.) in DCM. The new loading was 0.59 mmol g⁻¹, as determined spectrophotometrically. The resin capping was then performed by washings with methanol/DIPEA/DCM.

The peptides were cleaved from the resin and simultaneously deprotected with the mixture TFA/water/TIS (90:5:5 v/v) or, in the case of thiol-containing sequences, with the mixture TFA/water/TIS/EDT (90:3:4:3 v/v) for 2.5 h. The crude peptides were precipitated from ice-cold diethyl ether and recovered by centrifugation at 3 °C for 8 min. Several ether washes/centrifugation cycles were carried out to efficiently remove the scavengers.

VI.3.2 Peptide chain assembly by manual SPPS for Cpg-containing peptides (IV.10-13a/b)

The enantiopure N- α -benzoyl-Cpg units (0.5 equiv.) were dissolved in DMF and coupled manually to the resin-bound Id fragments **IV.2-9**, which were first synthesized by automated stepwise solid phase synthesis, as described above, using DIC (0.5 equiv.) and HOBt (0.5 equiv.). After 24 hours reaction time at room temperature, further 0.5 equiv. of the activating reagents were added and the reaction was left to run for another day. The

reaction completion was controlled by cleaving a small amount of peptide from the resin and analyzing it by analytical HPLC and MALDI-ToF-MS. The peptides were finally cleaved from the resin and simultaneously deprotected with the mixtures TFA/water/TIS (90:5:5 v/v) or TFA/water/TIS/EDT (90:3:4:3 v/v) for 2.5 h. The crude peptides were precipitated from ice-cold diethyl ether and recovered by centrifugation at 3 °C for 8 min. Several ether washes/centrifugation cycles were carried out to efficiently remove the scavengers.

The coupling of N- α -Fmoc-Cpg to the peptidyl-resin carrying the same sequence of peptide **IV.2** was performed by using the procedure described above for N- α -benzoyl-Cpg. After an acetylation step with acetic anhydride (10 equiv.) in DMF for 30 min, the Fmoc cleavage was done by treating the peptidyl-resin with 5% DBU in a 60 mM HOBt solution in DMF for 3 min, for two times. The chain assembly of the sequence corresponding to (70-80)-Id1 was carried out by performing double couplings (2×1 h) with N- α -Fmoc-amino acids (8 equiv.) activated with HOBt (8 equiv.) and DIC (8 equiv.). For the Fmoc cleavage the same DBU/HOBt solution described above was used. The N-terminus was then treated with acetic anhydride (10 equiv.) in DMF for 30 min. The peptides were finally cleaved from the resin and simultaneously deprotected with TFA/water/TIS (90:5:5 v/v) for 2.5 h. The crude peptides were precipitated from ice-cold diethyl ether and recovered by centrifugation at 3 °C for 8 min. Several ether washes/centrifugation cycles were carried out to efficiently remove the scavengers. Analysis by HPCL and MALDI-ToF-MS on the crude peptides confirmed the success of the synthesis.

VI.3.3 Peptide chain assembly by manual SPPS for Amp-containing peptides (IV.14-19a/b)

Each oligomer was synthesized manually starting from 30 mg of Rink amide MBHA resin (loading: 0.7 mmol g⁻¹). Single coupling of the N- α -Fmoc amino acids was performed using a mixture of amino acid/HOBt/DIC (each 4 equiv.) in DMF for 4 h. Single coupling of the N- γ -Fmoc Amp was carried out with 2.5 equiv. of the amino acid in the presence of equimolar HOBt and DIC in DMF for 16 h. The Fmoc group was removed with

5% DBU in a solution of 60 mM HOBt in DMF (2×3 min). Each coupling and Fmoc cleavage step was followed by washes with DMF ($3 \times$). After peptide chain completion, the peptidyl-resin was treated with TFA/TIS/water (90:5:5 v/v) for 2.5 h, the resin was filtered off, ice-cold diethyl ether was added to the filtrate to induce peptide precipitation. The precipitate was then recovered by centrifugation at 3 °C for 8 min, washed several times with ice-cold ether and finally dried in vacuo. The peptides were characterized by analytical HPLC and MALDI-ToF-MS.

VI.4 General procedure for peptide purification and characterization

The crude products were dissolved in 30% acetic acid and purified by preparative RP-HPLC using an Agilent equipment and a Luna C-18(2) column from Phenomenex (10 μ m, 250×21.2 mm). The binary elution system was (A) 0.004% (v/v) TFA in water and (B) ACN. The gradient was 10-70% B over 40 min, with a flow rate of 21 ml min⁻¹. UV detection was made at 220 nm. Analytical RP-HPLC was performed on the following equipment: L-6200A Intelligent pump from Merck, HP detector series 1050 from Agilent, and Luna C-18(2) column from Phenomenex (90 Å, 3 μ m, 150×4.60 mm). The binary elution system was (A) 0.012% (v/v) TFA in water and (B) 0.01% (v/v) TFA in ACN. The gradients used were: (i) 15-70% B over 35 min for the fully protected thioester peptide **II.18**; (ii) 5-60% B over 40 min for all Amp-containing peptides in chapter **IV**; (iii) 10-70% B over 40 min for all the other peptides (the flow rate was always of 1 ml min⁻¹). UV detection was made at 220 nm. The mass spectra were recorded on a Future GSG spectrometer (Bruchsal, Germany) for MALDI-ToF analysis, and on a ThermoQuest spectrometer (Finnigan) for LC-ESI analysis.

VI.5 Procedure for CD spectroscopy analysis

The CD spectra were recorded on a JASCO J710 spectropolarimeter, using a quartz cell with a path length of 0.02 cm. Peptide stock solutions were prepared in phosphate buffer (100 mM, pH 7.3), water or methanol, and their concentration was determined by measuring the UV absorbance of the Tyr residues at 280 nm ($\varepsilon = 1480 \text{ M}^{-1} \text{ cm}^{-1}$) [12], or of the fluorene group at 301 nm ($\varepsilon = 7800 \text{ M}^{-1} \text{ cm}^{-1}$) for peptides **IV.14a/b** and **IV.15a/b**. In the case of peptides lacking any chromophore, a stock solution was prepared assuming a peptide content of 70%.

For each CD spectrum ten scans were accumulated using the following parameters: 1 nm step resolution and band width, 2 s response time, 20 nm/min scan speed and 20 mdeg sensitivity. The CD spectrum of the solvent was subtracted from that of the peptide to eliminate interferences from the cell, solvent and optical equipment. Noise reduction was obtained by a Fourier transform filter with the program Origin (OriginLab Corporation, Northampton, MA, USA). The secondary structure element composition of the peptides reported in chapter II was extrapolated by submitting the experimental CD spectra to the Dichroweb online server [13].

VI.6 Procedure for NMR spectroscopy analysis

Peptides **IV.14a/b** and **IV.15a/b** were dissolved in CD_3OH at the concentration of 2 up to 4 mg ml^{-1} . The experiments were recorded on Bruker Avance 600 (600.13 MHz). The 1D-NMR spectra were collected between 278 K and 303 K by applying either the presaturation or the Watergate method to suppress the water resonance. 2D-NMR experiments (COSY, TOCSY, NOESY) were performed at 278 K for peptide **IV.14a** and at 283 K for all others by using the presaturation method. The mixing times were 80 ms for TOCSY and 200 ms for NOESY. All spectra were processed with the program TOPSPIN of Bruker (version 1.2).

Proton chemical shifts and TOCSY spectra are reported in the Appendix.

VI.7 Procedure for conformational search analysis

A systematic search on a simplified analogue of peptide **IV.14b** was performed by using the software Sybyl 7.2, considering the distance constraints found by NMR analysis. Moreover, the amide bond was fixed in the *trans* configuration. Torsion angles Ψ_2 , ϕ_5 , Ψ_6 , ϕ_9 and Ψ_{10} were incremented of 30° from 0° to 360° , whereas ϕ_3 , Ψ_4 , ϕ_7 and Ψ_8 were incremented of 60° . All conformations with bumps between any two atoms were not considered. The 225 conformations found were minimized by the Tripos force field with the Gasteiger-Huckel method for charges, using the dielectric constant of methanol (30) to reproduce the conditions used for the NMR experiments. After minimization, all conformations were checked against all preceding conformations by a match of all heavy atoms (in this case, heavy atoms are all atoms except hydrogens). All conformations with a root mean square (RMS) value lower than 0.3 \AA were rejected.

VI.8 Literature

1. Fields, G. B., Noble, R. L. (1990) Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids, *Int. J. Pept. Protein Res.*, **35**, 161-214.
2. Svobodova, J., Cabrele, C. (2006) Stepwise solid-phase synthesis and spontaneous homodimerization of the helix-loop-helix protein Id3, *Chembiochem*, **7**, 1164-1168.
3. von Eggelkraut-Gottanka, R., Klose, A., Beck-Sickinger, A. G. & Beyermann, M. (2003) Peptide thioester formation using standard Fmoc-chemistry, *Tetrahedron Letters*, **44**, 3551-3554.
4. Greenfield, N. J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data, *Anal. Biochem.*, **235**, 1-10.
5. Wüthrich, K. (1986) NMR of proteins and nucleic acids. *Wiley, New York*.
6. Karplus, M. (1959) Contact electron-spin interactions of nuclear magnetic moments, *J. Chem. Phys.*, **30**, 11-15.
7. Pardi, A., Billeter, M., Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C-alpha proton coupling constants, $^3J_{\text{HNH}}$, in globular proteins: use of $^3J_{\text{aNH}}$ for identification of helical secondary structure, *J. Mol. Biol.*, **180**, 741-751.
8. Cierpicki, T., Zhukov, I., Byrd, R. A., Otlewski, J. (2002) Hydrogen bonds in human ubiquitin reflected in temperature coefficients of amide protons, *J. Magn. Reson.*, **157**, 178-180.
9. Leach, A. R. (2001) Molecular modelling, principles and application. *Pearson, Prentice Hall, London*
10. Ramachandran, G. N., Ramakrishnan, C., Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations, *J. Mol. Biol.*, **7**, 95-99.
11. Anderson, R. J., Weng, Z., Campbell, R. K., Jiang, X. (2005) Main-chain conformational tendencies of amino acids, *Proteins*, **60**, 679-689.
12. Mach, H., Middaugh, C. R., Lewis, R. V. (1992) Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins, *Anal. Biochem.*, **200**, 74-80.

13. Whitmore, L., Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res.*, 32,668-673.

VII. Appendix: NMR data

Table 1: Summary of chemical shifts (ppm) of peptide **IV.14a** at 278K

Residue	NH	α -CH	β -CH	γ -CH	Others
Ala-1	7.58	4.03	1.30		
γ -Amp-2	8.53 (γ -NH)	4.44	2.10	2.59	3.51 (CH ₂ -pyrrolidine)
Ala-3	8.9	4.24	1.33		
γ -Amp-4	8.61 (γ -NH)	4.44	2.06	2.59	3.61 (CH ₂ -pyrrolidine)
Gly-5	8.79	3.89			
Tyr-6	8.25	4.51	3.04; 2.77		7.05 (ArHQ); 6.66 (ArHE)

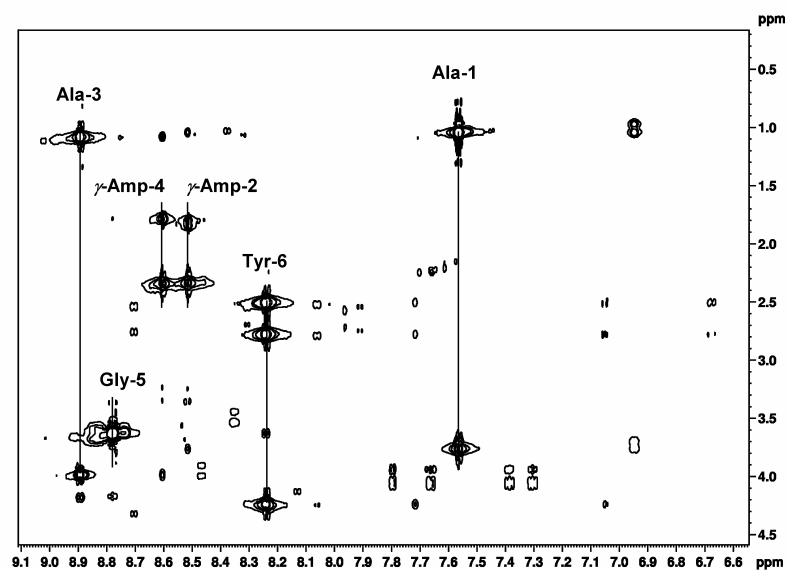


Figure 1: NH-aliphatic region (fingerprint) of the TOCSY spectrum of peptide **IV.14a**, recorded in CD₃OH at 278K

Table 2: Summary of chemical shifts (ppm) of peptide **IV.14b** at 283K

Residue	NH	α -CH	β -CH	γ -CH	Others
Ala-1	7.56	4.06	1.33		
γ -Amp-2	8.55 (γ -NH)	4.46	2.17	2.64	4.52-3.5 (CH ₂ -pyrrolidine)
Ala-3	8.87	4.27	1.35		
γ -Amp-4	8.63 (γ -NH)	4.48	2.11	2.62	4.52-3.5 (CH ₂ -pyrrolidine)
Ala-5	8.90	4.29	1.37		
γ -Amp-6	8.63 (γ -NH)	4.48	2.11	2.62	4.52-3.51 (CH ₂ -pyrrolidine)
Gly-7	8.80	3.92			
Tyr-8	8.26	4.52	3.07; 2.80		7.07 (ArHQ); 6.70 (ArHE)

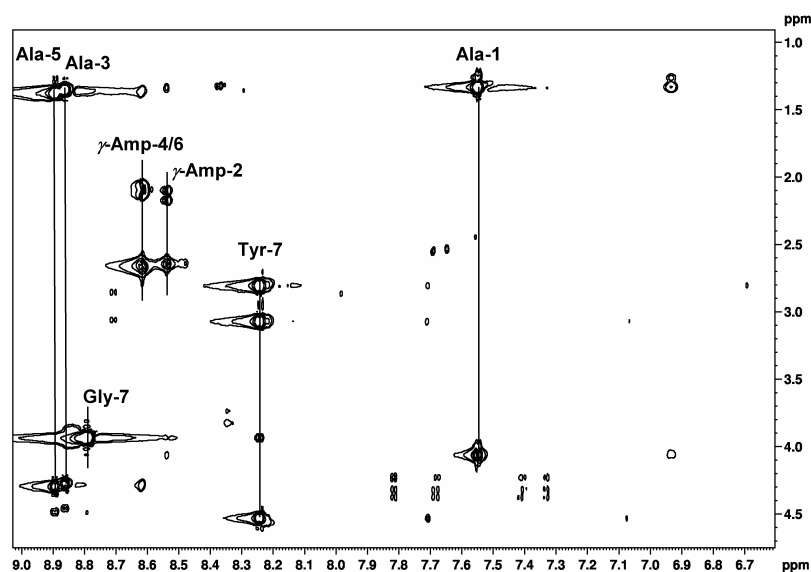


Figure 2: NH-aliphatic region (fingerprint) of the TOCSY spectrum of peptide **IV.14b**, recorded in CD₃OH at 283 K

Table 3: Summary of chemical shifts (ppm) of peptide **IV.15a** at 283K

Residue	NH	α -CH	β -CH	γ -CH	Others
Ala-1	7.54	4.02	1.31		
γ -Amp-2	8.48 (γ -NH)	4.47	2.04	2.66	3.53 (CH ₂ -pyrrolidine)
Ala-3	8.94	4.28	1.36		
γ -Amp-4	8.6 (γ -NH)	4.47	2.09	2.69	3.65 (CH ₂ -pyrrolidine)
Gly-5	8.88	4.02; 3.82			
Tyr-6	8.20	4.56	3.08; 2.79		7.06 (ArHQ); 6.69 (ArHE)

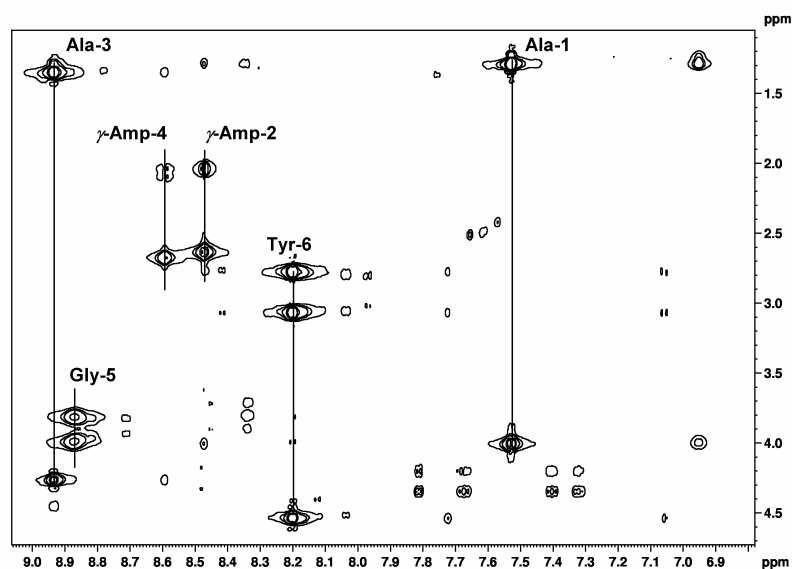


Figure 3: NH-aliphatic region (fingerprint) of the TOCSY spectrum of peptide **IV.15a**, recorded in CD₃OH at 283 K

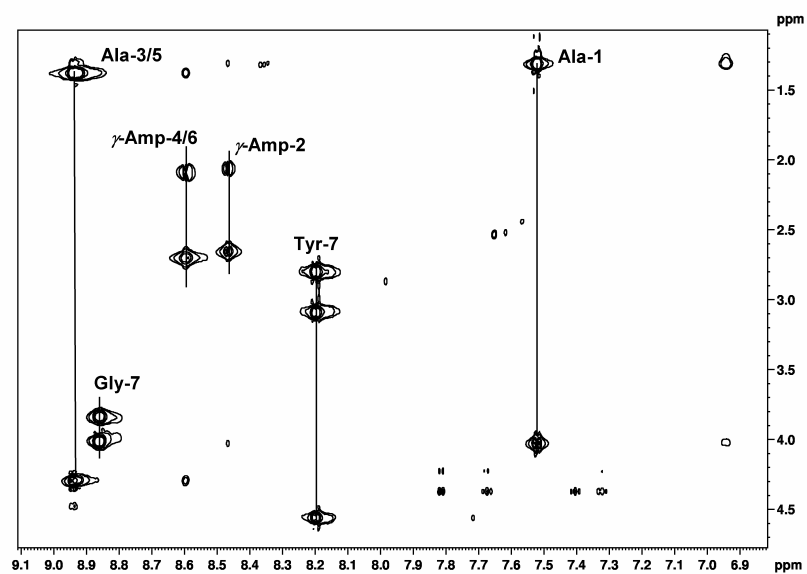


Figure 4: NH-aliphatic region (fingerprint) of the TOCSY spectrum of peptide **IV.15b**, recorded in CD₃OH at 283 K. The overlapping between the signals relative to Ala-3 and Ala-5 and between Amp-4 and Amp-6 can be observed.

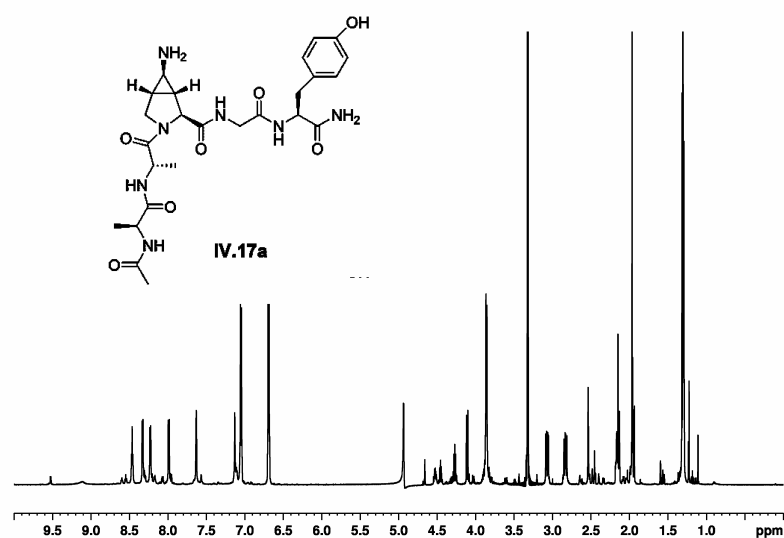


Figure 5: ^1H -NMR spectrum of peptide **IV.17a**, in CD_3OH at 293 K

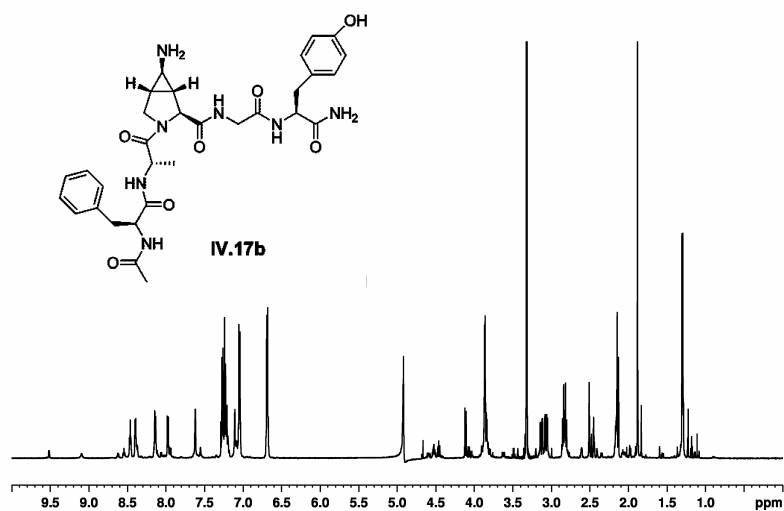


Figure 6: ^1H -NMR spectrum of peptide **IV.17b**, in CD_3OH at 293 K

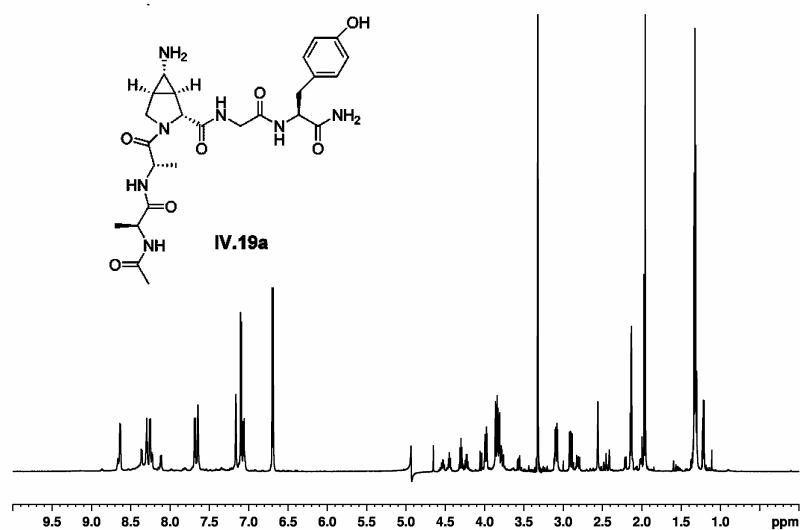


Figure 7: ^1H -NMR spectrum of peptide **IV.19a**, in CD_3OH at 293 K

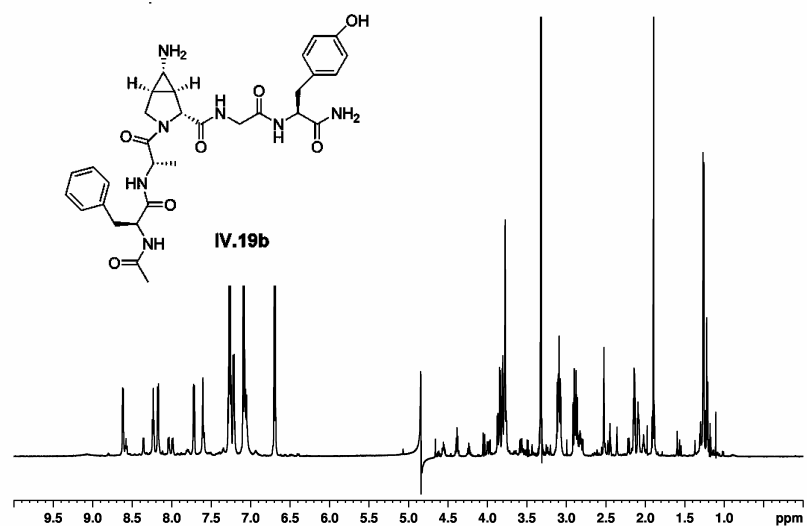


Figure 8: ^1H -NMR spectrum of peptide **IV.19b**, in CD_3OH at 303 K

Publications

Colombo N., Cabrele C. (2006), Synthesis and conformational analysis of Id2 protein fragments: impact of chain length and point mutations on the structural HLH motif., *J.Pept Sci.*, 12: 550-558

Colombo N., Schroeder J., Cabrele C. (2006), A short fragment of the Id2 protein containing the nuclear export signal forms amyloid-like fibrils. *Biochem. Biophys. Res. Commun.*, 346: 182-187

Brackmann F., **Colombo N.**, Cabrele C., de Meijere A. (2006), An improved synthesis of 3,4-(Aminomethano)proline and its incorporation in small oligopeptides. *Eur. J. Org. Chem.*, 4440-4450

Poster presentations

Colombo N., Cabrele C., Synthesis and conformational analysis of peptide fragments from the DNA-binding inhibitor Id-2. 7th German peptide symposium, 2005, Braunschweig (Germany)

Colombo N., Havrdová J. and Cabrele C., Id Proteins. 2nd Summer school of Medicinal Chemistry, 2004, Regensburg (Germany)

Curriculum Vitae

Noemi Colombo

Place of birth:

Monza (Italy), June 26th 1976

Education and studies:

July 2002:

Laurea degree in “Chemistry and Pharmaceutical Technologies” at the University of Milano (Italy).

Title of the thesis:

Asymmetric synthesis of 2-amino-3-nitro-norbornan-2-carboxylic acids.

Supervisor: Prof. Dr. M. L. Gelmi

Co-supervisor: Prof. Dr. D. Pocar

Final grade:

107/110

June 1995:

High School degree at “Liceo Scientifico A. Banfi”, Vimercate (Italy).

Scientific activity:

Since September 2003:

PhD student in the group of Prof. A. Buschauer under the supervision of Dr. C. Cabrele, Department of Chemistry and Pharmacy of the University of Regensburg (Germany)

September 2004:

Structural investigations of peptides by mono- and bi-dimensional NMR techniques, under the supervision of Dr. O. Zerbe, at the University of Zürich (Switzerland)

July 2002-July 2003:

Proteomics techniques; laboratories of Mass Spectrometry, at DIMESAB department of the University of Milano Bicocca (Italy)

Acknowledgements

I would like to thank *Prof. Dr. Armin Buschauer*, who gave me the opportunity to perform my PhD in Germany.

I am specially thankful to *Dr. Chiara Cabrele*, who offered me the chance to work in her group, giving me an interesting research project in the world of peptides, and supported its development at any time.

I am grateful to *Prof. Dr. Josef Schröder*, *Beate Voll* and *Heiko Siegmund*, for the TEM measurements; to *Prof. Dr. Maria Luisa Gelmi*, *Dr. Sara Pellegrino* and *Dr. Nicola Ferri* concerning the Cpg project and to *Prof. Dr. Armin de Meijere* and *Dr. Farina Brackmann* concerning the Amp project.

A special thanks to *Prof. Dr. Stefan Dove* and to *Dott.ssa Maria Elena Silva*, for their precious help with the computational techniques.

I am thankful to *Dr. Oliver Zerbe*, who taught me patiently the basics of 2D-NMR analysis interpretation.

I thank *Dr. Burgermeister*, *Mr. Kastner*, *Ms. Schramm* and *Ms. Stülher* for recording NMR spectra; *Dr. Vasold* and *Mr. Lautenschlager* for their assistance during analytical and preparative HPLC analysis; *Dr. Mayer*, *Mr. Kiermeier* and *Mr. Söllner* for their assistance during MALDI-ToF-MS measurements and for the LC-ESI analysis.

Additionally, I thank the working group of *Prof. Dr. O. Wolfbeis* for the use of the circular dichroism spectrometer.

For the financial support, I am grateful to the Deutsche Forschungsgemeinschaft (DFG), and to the DAAD (International Quality Network Medicinal Chemistry).

A special thanks to my PhD colleagues *Jaroslava Svobodová* and *Sebastian Kiewitz*, for the nice time we spent together in the lab, their constant support, and for their friendship.

Moreover, I would like to thank all the people who gave, inside and outside the lab, an important contribution to my experience in Regensburg:

Elena, for her constant presence, her kindly patience and helpfulness, and, moreover, for all the crazy weekends in Regensburg; *Eva*, for her precious help whenever I needed it and for the nice Bavarian evenings; *Michi*, for the funny discussions about differences between Italian and German people; *Francesco (Lo Zio Franz)*, for his sympathy and kindness; *Giovanni, Silvia* and *Elena Sancassiani*, the first friends I found here; *Valerio*, for his sympathy and all the music and movies he shared; *Edith* and *Ralf* for the beer weekends of last winter; all the people of the König's group, especially *Stefan (Milli)*, *Stefan (Maschio)*, *Daniel, Jiri*, and *Harald* for the nice time spent together during parties and dinners.

At the end I would like to thank all the people who supported me from a distance:

Nadia and *Marta*, whose friendship never failed over years and distance; *Lella, Lucia*, and *Sonia*, for their sympathy and precious help when I decided to come here; *Sara*, for her friendship and interesting discussions also about the world of peptides; *Laura, Gottardo, il Massa, Giovanna, Stefano* and, last but not least, *Andrea (il Ciccio)*, for all e-mails, SMS and other little things that helped me a lot during these years.

I am especially grateful to *Serafina* and *Nino*, for their constant support and farsighted advices

My parents, *Ileana* and *Mario*, together with my brother *Giorgio* and his wife *Federica*, receive my greatest wholeheartedly thank for their unconditioned understanding, their encouragement, and for their presence at any time.