Molecular Recognition of Integrin $\alpha_3\beta_1$ and Inorganic Compounds by Tailor-made Peptides

Yi Yang

Dissertation submitted to
the Department of Chemistry, Bielefeld University
for partial fulfillment of the requirements for the degree
of *Doctor rerum naturalium* (*Dr. rer. nat.*)

Bielefeld 2008

Acknowledgments

As I nearly completed to compile my doctoral thesis, I came to realize that I am finally approaching the end of my study in Bielefeld. I will probably have to say goodbye to my well-beloved supervisor and friendly colleagues soon. It is the time I am expecting and dreading.

Of all the persons that play important role in my life during these years, my supervisor Prof. Dr. Norbert Sewald stands out as the first one I am deeply appreciating. It was he who helped me see the light at the end of the tunnel in the darkest hour of my life up to now. It was he who led me into the wonderland of peptide of which I nearly knew nothing at the time. It was he who bestowed me the courage and wisdom to fight all kinds of difficulties in my research and life. I do believe he is the one of the persons who exerts maximal influences on my life.

I would like to thank Prof. Dr. Johannes Eble from Frankfurt Univesity Hospital with whom we established wonderful cooperation and achieved fruitful results concerning peptide inhibitors for the interactions between invasin and integrin $\alpha_3\beta_1$. His guidance in ELISA experiment and advice of the peptide inhibitor optimization are extremely crucial and valuable.

Prof. Dr. Dirk Volkmer from University of Ulm offered me the possibility to research on challenging but interesting project concerning biomineralization. I regard our cooperation as fruitful, successful and irradiative.

I would like to thank Dr. Zsuzsa Majer from Eötvös Loraánd University who kindly and patiently instructed me CD measurement during her stay in Bielefeld.

The atmosphere in our research group is so harmonious that I almost never feel that I am a foreigner here except during the World Cup 2006. Dr. Ulf Strijowski kindly instructed me the process of solid phase peptide synthesis, which led me on the right track of my research. I would like to thank Dr. Sven Weigelt as my labmate who helped me so much during my stay. His advices are extremely instructive on both chemistry and everyday life. I do not believe I could try to speak German without his courage and patience. I established fruitful collaboration with Dr. Sylwia Urman who contributed so much to the biological tests of β-Acc derivative of peptide inhibitors. I am grateful to her endeavors. Dr. Soledad Royo assisted me to tackle all kinds of problems originated from both chemistry and everyday life. She and Dr. Anna Norgren kindly offered me the help to review and correct my doctoral thesis, which saved me a lot of work. I would like to thank Katharina Gaus who taught me so much on both theory and practice of peptide NMR. Dr. Eckhart Guthöhrlein helped me design peptide inhibitors with Molecular Modeling. Anke Nieß and Marco Wißbrock are pretty helpful as they afforded me so much convenience in my research.

I would like to thank *Cell* and *Science* for their generosity to permit me to cite their corresponding academic figures in my doctoral thesis.

I am extremely grateful to the contributions that my parents and my wife Dan made. Their unselfish love, patience and support are the most important sources of my courage to face and fight all the difficulties during such a long time.

Finally, I would like to thank Arminia Bielefeld, which I support and love so much. This not-so-strong but so tough football team means so much for my immaterial world. I am always inspired by the spirit of this team as I got to know that "Leidenschaft ist unabsteigbar".

It might be the time to say goodbye to Bielefeld. However, all the beautiful scenes will be cherished in my heart for eternity.

Abbreviations

Abs	absorption
Ac	acetyl
β-Acc	β-aminocyclopropane carboxylic acid
ACN	acetonitrile
ADAM	a disintegrin and metalloproteinase
Aib	α-aminoisobutyric acid
Aloc	allyloxycarbonyl
Boc	*tert*-butoxycarbonyl
BOP	benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate
BroP	bromotris(dimethylamino)phosphonium hexafluorophosphate
CD	circular dichroism
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCC	*N*,*N*′-dicyclohexylcarbodiimide
DCM	dichloromethane
DHB	2,5-dihydroxybenzoic acid
DIPEA	diisopropylethylamine
DMF	dimethylformamide
DNA	deoxyribonucleic acid
DPPA	diphenyl phosphorazidate
ECM	extracellular matrix
EDT	ethane-1,2-dithiol
EDTA	ethylene diamine tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
Et	ethyl
Fmoc	9-fluorenylmethoxycarbonyl
HATU	*N*-[(dimethylamino)(1*H*-1,2,3-triazolo[4,5-*b*]pyridine-1-yl)methylene]-*N*-methylmethanaminium hexafluorophosphate-*N*-oxide

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFIP	hexafluoroisopropanol
Hmb	2-hydroxy-4-methoxybenzyl
HOBt	1,2,3-benzotriazol-1-ol
IC_{50}	half maximal inhibitory concentration
ICAMs	intercellular adhesion molecules
MALDI-ToF	matrix-assisted laser desorption/ionization time of flight
MS	mass spectrometry
NMR	nuclear magnetic resonance
NPY	neuropeptide Y
ORD	optical rotary dispersion
ORD	optical rotary dispersion
Pgm	platinum group metal
PyBOP	benzotriazo-1-yloxytripyrrolidinophosphonium hexafluorophosphate
PyBroP	bromotripyrrolidinophosphonium hexafluorophosphate
PyCloP	chlorotripyrrolidinophosphonium hexafluorophosphate
R.T.	room temperature
RNA	ribonucleic acid
RP-HPLC	reverse phase high performance liquid chromatography
SDS	sodium dodecyl sulfate
SPPS	solid-phase peptide synthesis
Su	succinimide
TBS	tris-buffered saline
TBTU	*N*-[(1*H*-benzotriazol-1-yl) -*N*-methylmethanaminium tetrafluoroborate *N*-oxide
tBu	*tert*-butyl
TCTU	1-[Bis(dimethylamino)methylen]-5-chlorobenzotriazolium-3-oxide tetrafluoroborate
TFA	trifluoroacetic acid

TFE	trifluoroethanol
TGF-β	transforming growth factor beta
TIS	triisopropylsilane
t_R	retention time
Trt	triphenylmethyl
Trt(2-Cl)	(2-chlorophenyl)diphenylmethyl
UNCAs	urethane-protected α-amino acid *N*-carboxyanhydride
Z	benzyloxycarbonyl

Contents

1 Supramolecular Chemistry of Peptides and Recognition by Receptors

1.1 Thermodynamics of the Association of Peptide Ligands and Protein Receptors

A new class of low-molecular compounds, including synthetic vaccines, synthetic diagnostics, peptide-originated drugs, are flourishing based on developments of peptide libraries, supramolecular chemistry of peptides in their recognition to corresponding protein receptors, and structure-based peptidyl ligand design.

Structure-based design is based on the observation that ligands bind to clearly defined molecular targets. Receptors are mostly membrane-bound proteins that selectively bind small molecules, referred to as ligands that elicit some physiological response. A strong and selective binding can be obtained from a high structural and chemical complementarity between the receptor and the ligand.[1,2] Structure-based ligand design may therefore be described as the search for small molecules that fit into the binding site of the target and can form favorable interactions.

The basic force to promote the ligand-receptor interaction could be regarded as the reach of low energy state of the ligand-receptor complex, which is described by the equilibrium shown below:

$$\text{ligand} + \text{receptor} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{ligand-receptor complex}$$

The biological activity of a ligand is referred to its binding affinity for the corresponding receptor, which could be reflected as the stability of the generated

ligand-receptor complex. This stability is commonly quantified with K_d, the dissociation constant for the ligand-receptor complex at equilibrium:

$$K_d = [\text{ligand}]\,[\text{receptor}]/\,[\text{ligand-receptor}]$$

The smaller the K_d, the larger the concentration of the ligand-receptor complex at its balanced equilibrium, the more stable is the generated complex, and hence the greater is the affinity of the ligand to its receptor.

From the thermodynamics point of view, the process of association of ligand and its receptor, namely, ligand-receptor complex formation, is entropically unfavorable, since it causes a loss in conformational degrees of freedom for both the receptor and ligand, as well as the loss of three rotational and three translational degrees of freedom.[3] That is to say, the strong association of ligand to its receptor greatly limits the degree of freedom of both ligand and receptor, which results in the loss of the net entropy of this literally close-system. Solely from the entropic viewpoint, this association is regarded as adverse. Therefore, highly favorable enthalpic gain, which is corresponded to the interactions between the receptor and the ligand, must compensate for the adverse entropic loss during the same association process.

The selective binding of a low-molecular-weight peptide ligand to a specific protein receptor is determined by structural recognition of peptide ligand to its protein receptor and their thermodynamics parameters in the association process. The binding affinity can be measured from the experimentally measured binding constant K_d:

$$\Delta G = -\,RT \ln K_d = \Delta H - T\Delta S$$

The experimentally determined binding constant K_d is typically in the range of 10^{-2} to 10^{-12} M, corresponding to a Gibbs free energy of binding ΔG between -10 and -70 *kJ*/mol in aqueous solutions.[4,5]

The enthalpic and entropic components of the binding affinity can be determined experimentally, e.g., by isothermal titration calorimetry (ITC). The available data indicate that there is always a substantial compensation between enthalpic and entropic contributions.[6-8] The available data also show that the binding may be enthalpy-driven (e.g., streptavidin-biotin, ΔG = -76.5 kJ/mol, ΔH = -134 kJ/mol) or entropy driven (e.g., streptavidin-HABA, ΔG = -22.0 kJ/mol, ΔH = 7.1 kJ/mol).[9] Data from protein mutants yield estimates of 5 ± 2.5 kJ/mol for the contribution from individual hydrogen bonds to the binding affinity.[10-12] Similar values have been obtained for the contribution of an intramolecular hydrogen bond to protein stability.[13-15] The consistency of values derived from different proteins suggests some degree of additives in the hydrogen bonding interactions.

Identifying a ligand to a given receptor is just the first step in the process of discovering a new drug that offers a therapeutic opportunity for the treatment of a disease.

1.2 Non-Covalent Interactions between Peptide Ligands and Receptors

1.2.1 General Introduction

Low-molecular peptide ligands are capable of interacting with their corresponding macromolecular protein receptors through both covalent and non-covalent interactions. Non-covalent and reversible covalent binding could be normally characterized by the equilibrium thermodynamics of their association process.

The majority of the currently available peptide ligands functions through non-covalent interactions with their corresponding protein receptors. Non-covalent interactions are therefore of particular interest in peptide ligand design and the deeper study of the mechanisms of non-covalent interactions should be hugely beneficial to optimize the

affinities of peptide ligands to their receptors. Important non-covalent interactions basically include hydrogen bonds, ionic interactions and hydrophobic interactions, as well as π-π interactions.

Generally speaking, direct interactions between the protein receptor and peptide ligand are very important for binding. The most important direct non-covalent interactions are highlighted in *Figure 1.1*. Structural data on unfavorable protein-ligand interactions are sparser, despite that they are sometimes also pretty instructive to the design of appropriate peptide ligand candidates to some certain receptors. Understandably, structures of weakly binding ligands are more difficult to obtain and are usually considered less interesting by many structural biologists. However, these data are vital for the development of scoring functions. Some conclusion can be drawn from the available data: unpaired buried polar groups at the protein-ligand interfaces are strongly adverse to binding. Few buried CO and NH groups in folded protein fail to form hydrogen bonds.[16] Therefore, in the ligand design process one has to ensure that polar functional groups, either of the protein or the ligand, will find suitable counterparts if they become buried upon ligand binding. Another situation that leads to a decreased binding affinity is imperfect steric fit, leading to holes at the hydrophobic part of the protein-ligand interface.

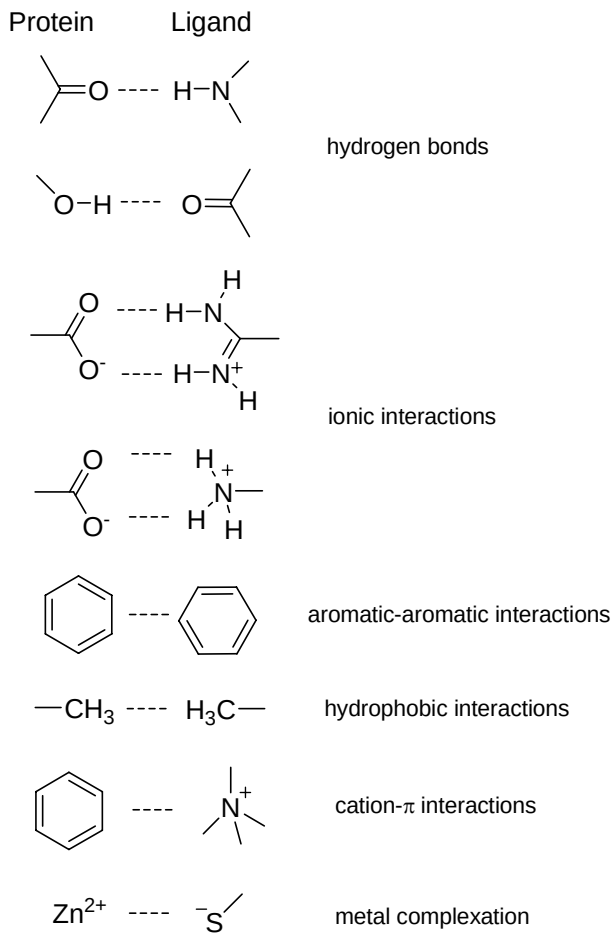


Figure 1.1 Typical non-covalent interactions found in protein-ligand complexes

1.2.2 Category of Non-covalent Interactions in the Association of Peptide Ligands and Protein Receptors

1. Ionic (Electrostatic) Interactions

For the protein receptor at physiological pH (generally taken as 7.4) basic amino acid residues such as Lys, Arg, and to a less extent, His are normally protonated and, thus, provide a cationic microenvironment. Acidic residues, such as Asp and Glu, are deprotonated and bear anionic groups under this circumstance. Ligand and receptor will be mutually attracted if they bear opposite charges. This electrostatic interaction can function across larger distances than those required for other types of interactions,

and it lasts longer. This electrostatic interaction could thus possibly associate the ligands and receptors bearing opposite charges provided a steric complementarity exists between them.

2. Ion-Dipole and Dipole-Dipole Interactions

The greater electronegative properties of atoms such as F, Cl, Br, I, S, N, and O, compared to C, will result in an asymmetric distribution of electron density and lead to electronic dipole in C-X bonds, where X is an electronegative atom shown above. These dipoles in ligand/receptor can interact with ions (dipole-ion interactions) or with other dipoles (dipole-dipole interactions) in the opposite counterpart of ligand/receptor, which contributes to the binding of ligands to their receptors.

3. Hydrogen Bonds

Hydrogen bonds could be categorized to the dipole-dipole interaction formed between the protons of a group X-H, where X represents an electronegative atom, while other electronegative atom (Y) contains a pair of non-bonded electrons. Hydrogen bonds are essential to maintain the structural integrity of helix and β-sheet conformation of peptides and proteins. Hydrogen bonds play crucially important roles in the recognition of ligands to their receptors. Binding affinities increase by about one order of magnitude per hydrogen bond. Some ligands form many hydrogen bonds with the protein but exhibit only very weak binding.[17-19] Interestingly, these ligands are all characterized by the absence of significant hydrophobic groups. Despite that the affinities they exert on binding of ligands to their receptors are not as strong as those from hydrophobic interaction, the existence of hydrogen bonds can increase the specificity of the bindings between ligands and their receptors.

4. Hydrophobic Interactions

In the presence of a nonpolar molecule or region of a molecule, the surrounding water molecules orient themselves and, therefore, are in a higher energy state than in the bulk water. When two hydrophobic molecules, such as peptide ligand and its protein receptor, surrounded by ordered water molecules, approach each other, these water molecules commence to disassociate from their bound ligand/receptor and associate with other water molecules. This water-liberation process increases the systematic entropy as the water molecules are no more limited to the certain domains of the ligand or receptor, thus resulting in a decrease of the net Gibbs free energy, which stabilizes the ligand-receptor complex and in turn reinforces the binding of ligands to their receptors. This interaction is popularly described as hydrophobic interaction. In principle, this is not the type of the force that fuse the two hydrophobic molecules/domains of molecules, but rather an inclination to decrease the free energy of nonpolar molecules/groups because of the increased entropy of the surrounding water molecules. Hydrophobic interactions might be the most important single factor responsible for non-covalent interactions between the molecules in aqueous solution.

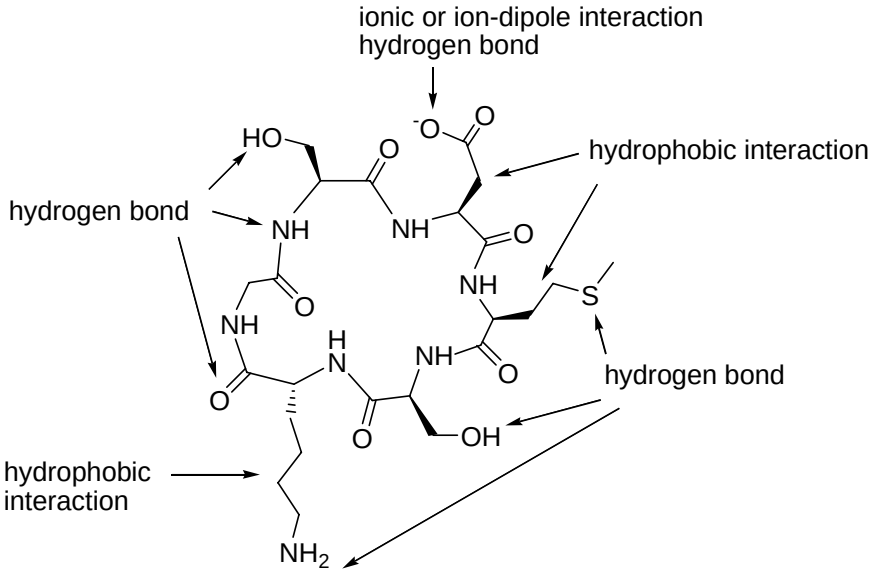


Figure 1. 2 Example of potential multiple peptide-receptor interactions.

In *Figure 1. 2* the cyclic peptide cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) was used as an example of peptide ligand to show the variety of non-covalent interactions that are possible. This peptide was designed in this thesis as the ligand of the membrane-protein receptor integrin $\alpha_3\beta_1$. The binding of this cyclic peptide to integrin $\alpha_3\beta_1$ could block the association of invasin with its membrane-protein receptor, thus inhibiting the internalization of invasin into mammal cell.

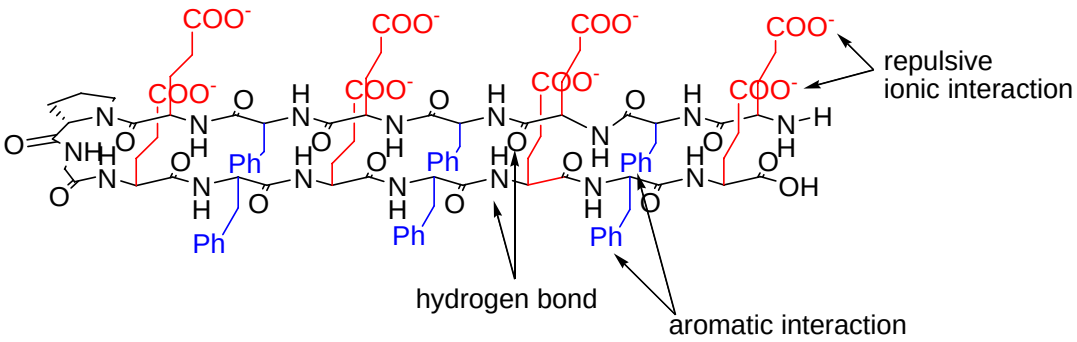


Figure 1.3 Non-covalent intra-molecular interactions in a β-hairpin peptide that sustain its conformation

Non-covalent interactions are crucial not only to the recognition of ligand to its receptor, but also decisive in determining or sustaining bioconformations of peptides under certain circumstances. *Figure 1.3* exhibits the multiple non-covalent intra-molecular interactions in an amphiphilic peptide which adopts β-hairpin conformation in halogenated alcohol solvent such as HFIP. Hydrogen bond, ionic interactions, and aromatic interactions, which might be either counteracting or stimulative, determine the conformation of this peptide jointly. Factors that jeopardize or promote some of these non-covalent forces would probably lead to destabilization or reinforcement of its original β-hairpin conformation.

1.2.3 Evaluation of the Non-covalent Interactions in the Recognition of Peptide Ligands to Protein Receptors

1.2.3.1 Water Molecules in Peptide Ligand/Protein Receptor Recognition

The biggest challenge in the quantitative treatment of protein-ligand interactions is still an accurate description of the role of water molecules. In particular, the contribution of hydrogen bonds to the binding affinity strongly depends on solvation and desolvation effects (see *Figure 1.4*). It has been shown by comparing the binding affinities of ligand pairs differing by just one hydrogen bond that the contribution of an individual hydrogen bond to the binding affinity can sometimes be very small or even adverse to binding.[20] Charge-assisted hydrogen bonds are stronger than neutral ones, but this is paid for by higher desolvation penalties. The electrostatic interaction of an exposed salt bridge is worth as much as a neutral hydrogen bond (5±1 kJ/mol),[21] while the same interaction in the interior of a protein can be significantly larger.[22]

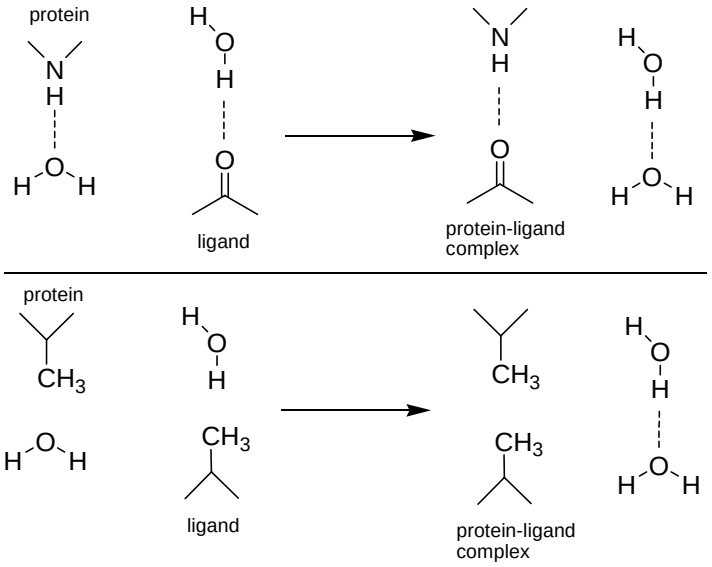


Figure 1.4 Role of water molecules in hydrogen bonds (upper part) and hydrophobic interactions (lower part). In the unbound state (left side), the polar groups of the ligand and the protein form hydrogen bonds to water molecules. These water molecules are replaced upon complex formation. Total number of hydrogen bonds does not change. In contrast, the formation of hydrophobic contact increases the total number of hydrogen bonds due to the release of water molecules from the unfavorable hydrophobic environment.

1.2.3.2 Electrostatic Interactions in Peptide Ligand/Protein Receptor Recognition

Our current knowledge indicates that unpaired buried polar groups at the protein-ligand interface are strongly adverse to binding. A statistical analysis of high-resolution protein structures showed that in protein less than 2% of the polar atoms are buried without forming a hydrogen bond.[23]

1.2.3.3 Hydrophobic Interactions in Peptide Ligand/Protein Receptor Recognition

Hydrophobic interactions are essentially contacts between apolar parts of the protein and the ligand. The generally accepted view is that hydrophobic interactions are mainly due to the replacement and release of ordered water molecules and are therefore entropy-driven.[24,25] The entropy gain is obtained when the water molecules are no longer positionally confined. There are also enthalpic contributions to hydrophobic interactions. Water molecules occupying hydrophobic binding sites are unable to form hydrogen bond. If they are released, they can form strong hydrogen bonds with bulk water. It has been shown in many cases that the contribution to the binding affinity is proportional to the hydrophobic surface area buried from solvent with values in the range of 80-200 J/(mol $Å^2$).[26,27]

Many protein-ligand complexes are characterized by the presence of both polar and hydrophobic interactions. The bound conformation of the ligand is determined by the relative importance of these contributions. The process of forming a complex between a small molecule ligand and a protein is a complex equilibrium process. A solvated ligand likely exists as an equilibrium mixture of several conformers; likewise the solvated protein also exists as several conformers in equilibrium. To form a complex the solvent molecules that occupy the binding sites are displaced by the ligand to produce a solvated complex. Tight complexes will result when the protein-ligand

interactions are significantly stronger than the interactions of the protein alone and ligand alone with solvent. Small variations in complex stability measured in kJ/mol translate into significant differences in ligand affinity for a protein.

1.2.4 Empirical Approaches in the Design of Peptide Ligands

1.2.4.1 General Experiences of the Tight Associations of Peptide Ligands and Protein Receptors

What are the most important properties that allow a peptide ligand to bind tightly and selectively to a protein receptor? Our current understanding of protein-ligand interactions is still far from being sufficient to answer this question fully. The most important prerequisite appears to be a good steric and electronic complementarity between protein and ligand. However, due to desolvation effects this criterion alone is not sufficient to fully describe tight binding of ligands.

Basically, there are several features found in all complexes of tightly binding ligand:

1. There is a high level of steric complementarity between the peptide ligand and protein receptor. This observation is also described as the lock-and-key paradigm.
2. There is usually high complementarity of the surface properties between the protein receptor and the ligand. Many tight binding ligands form significant hydrophobic interactions with the protein receptor. If the hydrophobic contact surface can be enlarged by an additional hydrophobic substituent, enhanced binding affinity is frequently observed. Therefore, the search for unoccupied hydrophobic pockets should always be one of the first steps in the design of new ligands. Hydrophobic parts of the ligands are most frequently found to be in contact with hydrophobic parts of the protein receptor. Polar groups are usually paired with suitable polar protein groups to form hydrogen bonds or ionic

interactions. A burial of unpaired polar groups upon ligand binding leads to a loss of binding affinity and should be avoided. With very few exceptions, there are no repulsive interactions between the ligand and the protein.

3. Additional hydrogen bonds do not always lead to improved binding affinities but may be nevertheless able to improve the selectivity and to make the compound more water soluble.
4. The binding of a ligand to a protein receptor always leads to the displacement of water molecules. If the ligand can form more hydrogen bonds than the water molecules that are released, then a very tight binding can be achieved.
5. The ligand usually binds in an energetically favorable conformation. Rigid ligands can bind more strongly than flexible ligands because the entropy loss due to the freezing of internal degrees of freedom is smaller.
6. Water can form strong hydrogen bonds but is not particularly well suited as a transition metal ligand. For transition metal-containing enzymes such as metalloproteases, it is therefore a good idea to incorporate functional groups into the ligands that are known to bind well to metal ions (e.g. thiols, hydroxamic acid).

1.2.4.2 Approaches to Structure-Based Ligand Design

Structure-based ligand design basically originated from the thorough research of certain target receptor molecule. In most of these cases this knowledge is resided in the natural substrate, a cofactor or the extracellular protein receptor. In some cases, such prototype molecules could possibly lead directly to derivative compounds with drug properties. In other cases, structural resemblance to a known ligand is too elusive to be carried out (structurally too complex, too difficult to synthesize, unwanted physicochemical properties, ect.) and alternative design approaches are favored.

Ligands Derived from Substrate or Natural Ligand

For most of the targets, the natural substrate or ligand is luckily known. Structural analogs of the substrate can serve as inhibitors if they are no longer processed or transformed by the enzyme. In other cases, modified mimetics of natural ligands could lead to antagonists or agonists of their corresponding protein receptors. Typically, natural ligands of protein receptor can be modified to improve its affinity to the receptor, increase the selectivity, or optimize the physicochemical properties.

Structures Derived from 3D Database Searchers

Since the pharmacological action is mediated through the 3D shape of the ligand molecule in the receptor-bound conformation, efforts have been made to construct low-energy 3D structures from 2D databases. A pharmacophore hypothesis can be derived from a series of known inhibitors and their consensus 3D features. From this technology the molecules with lead structure could be designed and derived. A lot of computer-assisted program such as *Catalyst* have been invented and applied to generate a pharmacophore hypothesis using a number of reference compounds as input.

De-novo Design of Ligands

The sequence and structure information in biostructure research is growing exponentially. Increasingly, we find ourselves in a situation where the structure of an enzyme or a receptor protein is known and suggestions for new inhibitors are desirable. This is where so-called *de-novo* design tools are applied.[28] In this approach to structure-based design, it is attempted to construct new molecules completely from scratch. These tools either build up candidate ligands from atoms or fragments, or they search databases of existing structures for complementary molecules.

2 Peptides as Integrin Ligands

2.1 Integrins

Integrins form a protein family present in many animal cells. They are permanently attached to the enclosing or separating tissue called the plasma membrane, which acts as a barrier around a cell. Proteins such as integrin, that are permanently attached to a biological membrane such as the plasma membrane, are called integral membrane proteins. The plasma membrane is composed of lipid molecules, such as fats, and is semipermeable.

Integrins play a role in the attachment of cells to other cells, as well as in the attachment of a cell to the extracellular matrix. Besides the attachment role, integrins also mediate signal transduction, a process by which a cell tranforms one kind of signal or stimulus into another. The signal that the integrin converts comes from the extracellular matrix to the cell.

There are many types of integrin, and many cells have multiple types on their surface. Integrins are of vital importance to all animals and have been found in all animals tested, from sponges to mammals. Integrins have been extensively studied in humans. Other types of protein that play a role in cell-cell and cell-matrix interaction and communication are cadherins, CAMs and selectins.

Integrins are approximately 280 Å long heterodimeric membrane glycoproteins, composed of an α- (150 to 180 kD) and a β- (~90 kD) subunit, both of which are type I membrane proteins.[29] 19 α and 8 β mammalian subunits are known, which assemble noncovalently to give 24 different heterodimers.[29] Although these subunits could in theory associate to give more than 100 integrin heterodimers, the actual diversity appears to be much more restricted. Contacts between the α and β subunits

primarily involve their *N*-terminal halves, which together form a globular head, the remaining portions form two rod-shaped tails that also span the plasma membrane. Each integrin subunit has a large extracellular domain, a single membrane spanning domain and usually a short cytoplasmic domain (40-60 amino acids).[29] These short cytoplasmic domains of the α and β integrin subunits do not have intrinsic enzymatic activities, but can interact with a variety of cytoplasmic proteins, including cytoskeletal and signaling molecules. The α-cytoplasmic domains are highly diverse, whereas the β-cytoplasmic domains are somewhat conserved but they are necessary and sufficient for integrin-dependent signaling.[29] Association of α- and β-subunits defines distinct, although largely overlapping ligand specificity. Integrin binding to extracellular matrices can be classified as either RGD-dependent (binding e.g. fibronectin, vitronectin and fibrinogen) or RGD-independent (binding e.g. collagen and invasin). In addition, some integrins can bind to counterreceptors (such as intercellular adhesion molecules ICAMs) on adjacent cells leadings to homotypic and heterotypic cell-cell interaction. Like other receptors, integrins transmit signals to the cell interior ("outside-in" signaling), which regulates organization of the cytoskeleton, activates kinase-signaling cascades, and modulates the cell cycle and gene expression.[30] Unlike other receptors, ligand binding to integrins is not generally constitutive but is regulated to reflect the activation state of the cell. This "inside-out" regulation of integrin protects the host from pathological integrin–mediated adhesion.[31] It is known that inside-out and outside-in signaling are associated with distinct conformational changes in the integrin extracellular segment.

Table 2.1 compiles the members of the β_1-integrin family together with their putative or confirmed extracellular ligands.

Table 2.1 Extracellular ligands for β_1-integrins

Integrin	Ligand
$\alpha_3\beta_1$	Epiligrin, Fibronectin, Reelin, **Invasin**, Thrombospondin, Laminin
$\alpha_4\beta_1$	Fibronectin, Osteopontin, **Invasin**, VCAM-1, Prepro von Willebrandt factor, Coagulation factor XIII, Angiostatin, Tissue transglutaminase, Rotavirus
$\alpha_5\beta_1$	Fibronectin, Fibrinogen, **Invasin**, Tissue transglutaminase, ADAM-15, -17
$\alpha_6\beta_1$	Laminin, Sperm fertilin, Cystein-rich angiogenic protein 61, Fisp12/mCTGF, Papilloma virus, **Invasin**
$\alpha_v\beta_1$	Fibronectin, **Invasin**, Vitronectin, TGFβ latency-associated peptide, Parechovirus

2.2 State-of-the-art: Cyclic Peptides

2.2.1 General Introduction

Homodetic cyclic peptides are polypeptide chains whose amino and carboxyl termini are themselves linked together with a peptide bond, forming a circular chain. They create a large family of naturally occurring or synthetic compounds which hold a variety of unique biological properties compared to their linear counterpart.[32] One interesting property of cyclic peptides is that they tend to be resistant to degradation by proteases. This property makes cyclic peptides attractive in drug design as scaffolds.

Furthermore, the cyclization of the linear precursor peptide could possibly lock the peptide inhibitor in an advantageous conformation, thus decreasing the relevant entropy loss upon its binding to the receptor, enhancing its binding affinity. The constrained geometry of cyclic peptides is understandably capable of locking the referred peptides in a favorable and stable conformation as the bioactivities are concerned. The cyclic peptide scaffolds have gained extreme importance as templates for turn-forming peptide structures. Unusual amino acids residues, such as β-amino acids, γ-amino acids, D-amino acids, α,β-didehydro amino acids, *N*-alkyl amino acids, proline and analogues, and $C^{\alpha,\alpha}$-disubstituted amino acids, are frequently incorporated into the natural or synthetic cyclic peptides, resulting in peptides with certain geometry. These special properties of cyclic peptides afford the drug designer a superb idea reservoir to construct peptide ligands with biologically active conformations. Structure-activity relationship studies on bioactive cyclic peptides are one of the main interests concerning these special peptides, which include conformational explorations with respect to peptide ligand/receptor interaction, since the ring closure of linear precursor peptides could possibly bring forward some special conformational properties and improved bioactivities. Cyclic peptides serve as models of protein-recognition motifs, and are used to mimic β-sheets, β-hairpins, β-turns, or γ-turns. These properties of cyclic peptides have already been fused into the study of structure-based ligand design.

2.2.2 Examples of Cyclic Peptide Ligands with -Arg-Gly-Asp- Recognition Motif

2.2.2.1 Cilengitide: cyclo-(-Arg-Gly-Asp-D-Phe-*N*-Me-Val-)

Among the cyclic peptides bearing biological activities, those with -Arg-Gly-Asp- motif are most intensively researched. Cilengitide (see *Figure 2.1*) is one of the most known examples of this family. It is a cyclic pentapeptide containing -Arg-Gly-Asp- motif, designed and synthesized at Technical University Munich in collaboration with Merck KGaA in Darmstadt.[33]

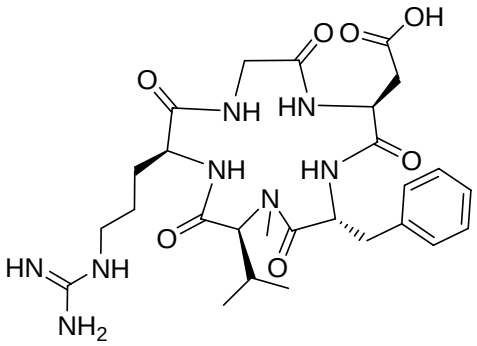


Figure 2.1 Cyclic pentapeptide Cilengitide: cyclo-(-Arg-Gly-Arg-D-Phe-*N*-Me-Val-)

Inhibitors with the -RGD- sequence locked in a favorable conformation have been extensively studied as potential competitors against some natural ligands of the integrin family, since integrin receptors frequently recognize a primary peptide sequence, -Arg-Gly-Asp-, with certain conformation in their target ligands. Cilengitide, with the sequence cyclo-(-Arg-Gly-Asp-D-Phe-*N*-Me-Val-), was found to be a highly active ligand addressing integrin $\alpha_v\beta_3$, which is expressed by endothelial cells and plays a role in the attachment of endothelial cells to the extracellular matrix. Cilengitide was proved to exhibit high affinity to integrin $\alpha_v\beta_3$ and integrin $\alpha_5\beta_1$ *in vitro*.[34] Inhibition of integrin $\alpha_v\beta_3$ activity by cyclic RGD peptides has been shown to induce endothelial apoptosis,[35] inhibit angiogenesis,[36,37] and increase endothelial monolayer permeability.[38] The inhibition of integrin $\alpha_v\beta_3$ activity has been associated with decreased tumor growth in breast cancer xenografts and melanoma xenografts.[39,40]

2.2.2.2 Cilengitide Derivatives: cyclo-(-Arg-Gly-Asp-(±)-β-Acc-Val-)

On the basis of the finding of cilengitide, the research group of Prof. Sewald decided to optimize the acitivity of the peptide inhibitor with -RGD- primary sequence. For the rational design of a peptide with improved biological activity, the elucidation of the three-dimentional structure is highly necessary, in order to establish a structure-activity relationship. The idea of spatial screenings of peptides and peptidomimetics is an important concept,[41] which applied the concerned peptides with diverse conformations in order to search the bioactive conformation of peptide ligand.

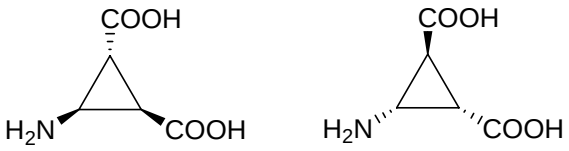
It is known that the incorporation of a β-amino acid residue in a cyclopeptide may stabilize its conformation, leading to a possible improved activity of concerned peptide ligand through the entropy effect, provided that the receptor-bound conformation of the peptide ligand remained obtainable. If a β-amino acid residue is incorporated into a cyclopentapeptide, it would preferentially occupy the central position of a γ-turn, which is elongated by a methylene group compared with a regular γ-turn, named derivatively as pseudo-γ-turn (Ψγ).[42] A binding motif, such as -RGD- sequence could be combined with a secondary structure inducer like D-amino acid, *N*-alkyl amino acid, or β-amino acid in order to engender a more favorable conformation as bioactivity is concerned.

Inspired by this methodology, derivatives of *cis*-β-aminocyclopropane carboxylic acid (β-Acc)[43-46] were introduced into the target peptides (see *Figure 2.2*). Two cyclic pentapeptides containing (+)-β-Acc or (-)-β-Acc, as well as -RGD- motif as derivatives of cilengitide were synthesized, and tested in cell adhesion assay. These two peptides, with the sequence cyclo-(-Arg-Gly-Asp-(+)-β-Acc-Val-) **1** and cyclo-(-Arg-Gly-Asp-(-)-β-Acc-Val-) **2**, respectively, were evaluated in cell adhesion assays with two cancer cell lines WM115 and K562. Adhesion of K562 and WM115 cell lines to their ligands fibronectin and vitronectin is predominantly mediated by integrin $\alpha_5\beta_1$ and integrin $\alpha_v\beta_3$, respectively.[47] The ability of the RGD peptides **1** and

2 to inhibit the adhesion of the K562 and WM115 cells to their ligands was compared to the previously described peptides cyclo-(-Arg-Gly-Asp-D-Phe-Val-) **3** and cyclo-(-Arg-Gly-Asp- D-Phe-β-Ala-) **4** as references.

It turned out that peptide **1** showed 10 fold higher affinity biological activity[48] than the reference peptide **3**,[41] as inhibition of adhesion of WM 115 cells to vitronectin through integrin $\alpha_v\beta_3$ is concerned. The inhibition capacity of **2** in this regard is nearly comparable to **3**. Peptide **4** displayed rather moderate influence on WM115 cell adhesion to vitronectin. Integrin $\alpha_5\beta_1$ mediated cell adhesion of K562 cells to fibronectin was nearly equally inhibited by peptides **1** and **2**.[48] In this assay both peptides **1** and **2** were approximately five- to sixfold more active than reference peptide **4**, which in turn was about twice as active as peptide **3**.[48]

The significantly increased affinity of the RGD peptide **1** to integrin $\alpha_v\beta_3$ is interpreted as a consequence of the introduction of the rigid β-Acc derivative, with **1** as the most active ligand of the integrin $\alpha_v\beta_3$ investigated so far.[48]

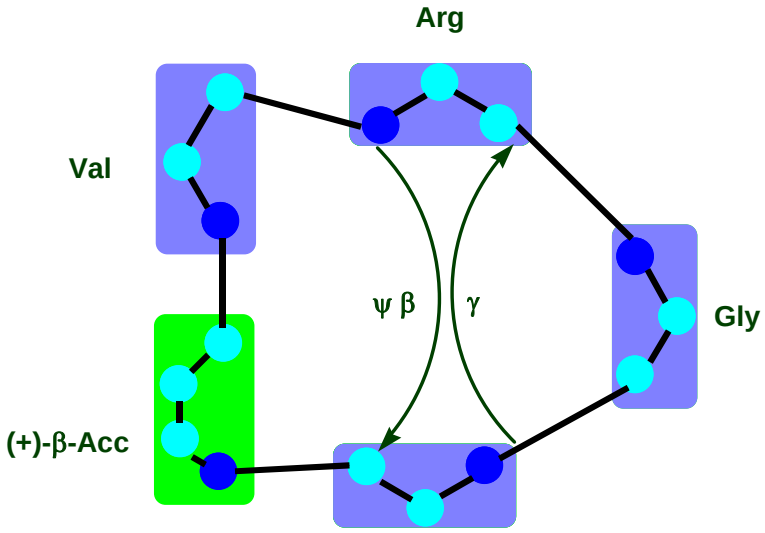


(+)-β-Acc (-)-β-Acc

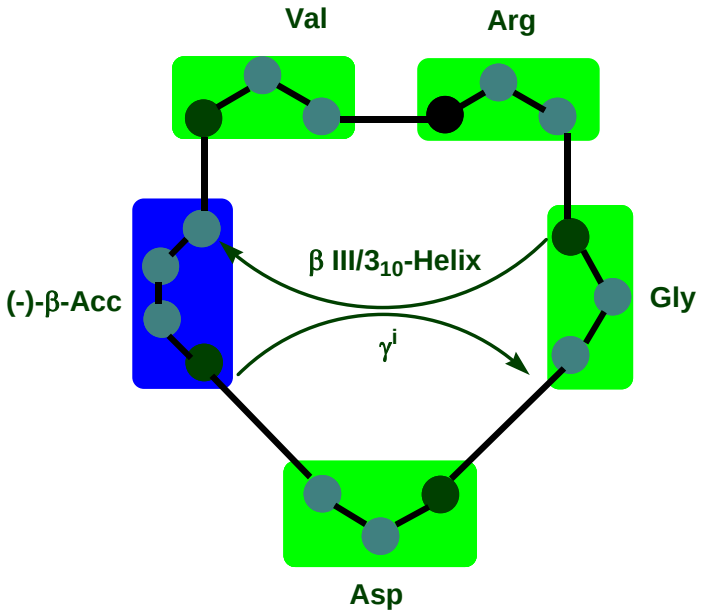
Figure 2.2 Secondary structure inducer: (+)-β-Acc and (-)-β-Acc

Through the introduction of cyclic β-amino acid (+/-)-β-Acc, the conformation of their parent cyclopentapeptide were enormously restricted. In peptide **1** Gly was locked in the central position of a γ-turn, while (+)-β-Acc was resided in *i*+1 position of a pseudo-β-turn. In peptide **2** Asp was locked in the central position of a γ^i-turn. This structure has the torsion angle of a 3_{10}-helix between the residue (-)-β-Acc and Gly, which was originally characterized as βIII-turn, in which the Val and Arg resided in *i*+1 and *i*+2 positions, respectively.[48] (see *Figure 2.3*)

This project shows the important roles cyclic peptide play as conformational restricted ligands to certain receptors. The underlying structure-activity relationship and diverse conformational propensity of amino acid residues afford abundant valuable informations for the structure-based design of peptide drugs.



(1)



(2)

Figure 2.3 Schematic structures of peptide **1** and **2** obtained by NMR and molecular dynamic calculations

2.3 Internalization of *Yersinia Sp*. through Integrins

2.3.1 General Introduction of *Yersinia*

Yersinia is a genus of Gram-negative rod shaped bacteria in the family Enterobacteriaceae. It is a few micrometers long and fractions of a micrometer in diameter, and is facultative anaerobe.[49] Some members of *Yersinia* are pathogenic to humans. Rodents are the natural reservoirs of *Yersinia*; less frequently other mammals serve as the host. Infection may occur either through blood (in the case of *Yersinia pestis*) or in an alimentary fashion, through occasionally via consumption of food products (especially vegetables, milk-derived products and meat) contaminated with infected urine or feces.[49]

There are three human-pathogenic *Yersinia* species: *Yersinia pestis, Yersinia enterocolitica and Yersinia pseudotuberculosis.*

2.3.1.1 *Yersinia pseudotuberculosis*

Yersinia pseudotuberculosis is a Gram-negative bacterium which primarily causes disease in animals; humans occasionally get infected zoonotically, most often through the food-borne route.[49] *Yersinia pseudotuberculosis* is the least common of the 3 main *Yersinia* species to cause infections in humans. It is primarily a zoonotic infection with variable hosts, including domestic and sylvatic animals and birds. The condition has been associated with food-borne infection, including a few outbreaks. The organism primarily leads to a gastroenteritis (diarrheal component uncharacteristic) characterized by a self-limited mesenteric lymphadenitis that mimics appendicitis.[50] The organism invades mammalian cells and survives intracellularly; the primary virulence factor is a plasmid-encoded protein that causes increased invasiveness. Postinfectious complications include erythema nodosum and reactive arthritis.[51] Thus, a major triad for infection with this organism include fever,

abdominal pain, and rash. Rarely, it has been associated with septic complications (often in patients who are immunocompromised with chronic liver diseases).

In animals, *Y. pseudotuberculosis* can cause tuberculosis-like symptoms, including localized tissue necrosis and granulomas in the spleen, liver, and lymph node.[51] In humans, symptoms are similar to those of infection with *Y. enterocolitica* (fever and right-sided abdominal pain), except that the diarrheal component is often absent, which sometimes makes the resulting condition difficult to diagnose. *Y. pseudotuberculosis* infections can mimic appendicitis, especially in children and younger adults, and, in rare cases the disease may cause skin complaints (erythema nodosum), joint stiffness and pain (reactive arthritis), or spread of bacteria to the blood (bacteremia).[52]

Infection usually becomes apparent 5–10 days after exposure and typically lasts 1–3 weeks without treatment.[53] In complex cases or those involving immunocompromised patients, antibiotics[53] may be necessary for resolution; ampicillin, aminoglycosides, tetracycline, chloramphenicol, or a cephalosporin may all be effective.

The recently described syndrome Izumi-fever has been linked to infection with *Y .pseudotuberculosis*.[54]

2.3.1.2 *Yersinia enterocolitica*

Yersinia enterocolitica is a species of gram-negative coccobacillus-shaped bacterium, belonging to the family Enterobacteriaceae. It causes primarily a zoonotic disease (cattle, deer, pigs, and birds); animals which recover frequently become asymptomatic carriers of the disease.[55]

Acute *Y. enterocolitica* infections produce severe diarrhea in humans, along with Peyer's patch necrosis, chronic lymphadenopathy, and hepatic or splenic abscesses. Additional symptoms may include entero-colitis, fever, mesenteric adenitis, erythema

nodosum and acute terminal ileitis, which may be confused with appendicitis or Crohn's disease.[56]

Treatment of *Y. enterocolitica* infections requires aggressive antibiotic therapy, typically involving a combination of chloramphenicol, ampicillin, and polymyxin.[55]

Y. enterocolitica infections are sometimes followed by chronic inflammatory diseases such as arthritis.[49] *Y. enterocolitica* seems to be associated with autoimmune Graves-Basedow thyroiditis.[57] Whilst indirect evidence exists, direct causative evidence is limited,[58] and *Y. enterocolitica* is probably not a major cause of this disease, but may contribute to the development of thyroid autoimmunity arising for other reasons in genetically susceptible individuals.[59] It has also been suggested that *Y. enterocolitica* infection is not the cause of auto-immune thyroid disease, but rather is only an associated condition; with both having a shared inherited susceptibility.[60]

2.3.1.3 *Yersinia pestis*

Yersinia pestis is a Gram-negative facultative anaerobic bipolar-staining (giving it a safety pin appearance) bacillus bacterium belonging to the family Enterobacteriaceae.[55] The infectious agent of bubonic plague, *Y. pestis* infection can also cause pneumonic and septicemic plague.[49] All three forms have been responsible for high mortality rates in epidemics throughout human history, including the Great Plague and the Black Death.

Pathogenicity of *Y. pestis* is in part due to two anti-phagocytic antigens, named F1 (Fraction 1) and V, both important for virulence.[55] These antigens are produced by the bacterium at 37 °C. Furthermore, *Y. pestis* survives and produces F1 and V antigens within blood cells such as monocytes, but not in polymorphonuclear neutrophils. Natural or induced immunity is achieved by the production of specific opsonic antibodies against F1 and V antigens; antibodies against F1 and V induce phagocytosis by neutrophils.[61]

The traditional first line treatment for *Y. pestis* has been streptomycin,[62,63] chloramphenicol, tetracycline,[64] and fluoroquinolones. There is also good evidence to support the use of doxycycline or gentamicin.[65]

It should be noted that strains resistant to one or two agents specified above have been isolated: treatment should be guided by antibiotic sensitivities where available. Antibiotic treatment alone is insufficient for some patients, who may also require circulatory, ventilatory, or renal support.

2.3.2 Invasin

2.3.2.1 Entry into Mammalian Cells by *Yersinia pseudotuberculosis*

The interaction of bacterial pathogens with host cells is an important step in the establishment of a productive infection. Bacterial adhesion to the mammalian cell surface allows either extracellular colonization[66-68] or penetration within the host cell.[69-72] Once internalized, the pathogen may replicate within the protective niche of the cell or translocate into deeper tissues where multiplication occurs.[73,74]

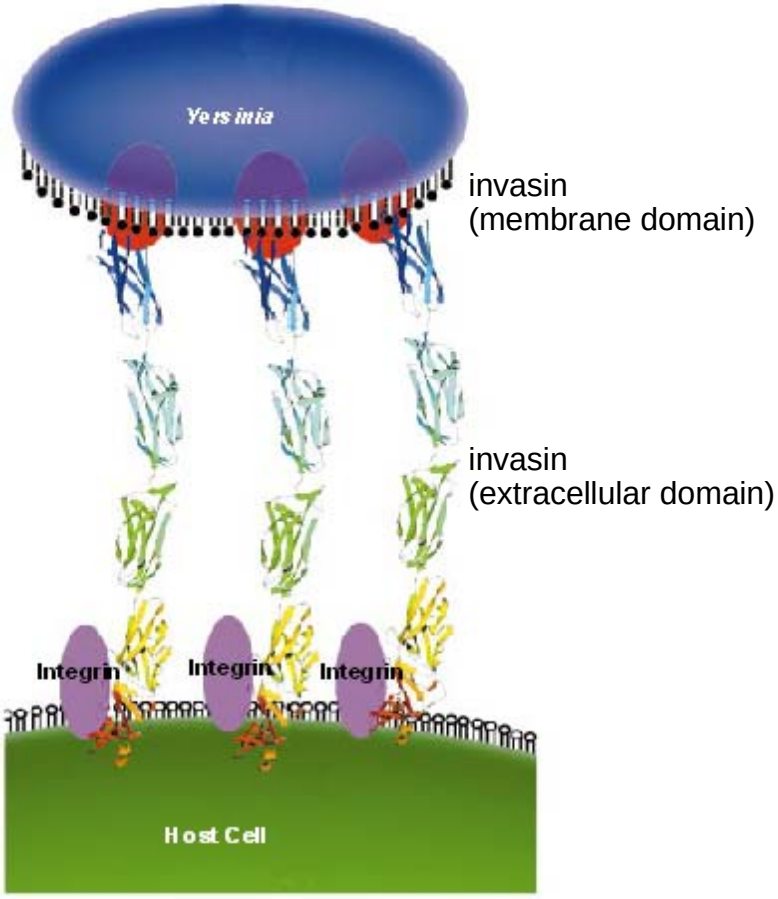


Figure 2.4 Invasin-mediated *Yersinia* Internalization.[75]

The latter strategy is used by the enteropathogenic bacterium *Yersinia pseudotuberculosis*.[76-78] *Yersinia pestis* and *Y. enterocolitica* are entero- pathogenic Gram-negative bacteria that cause gastroenteritis when they are translocated across

the intestinal epithelium at Peyer's patches by way of M cells.[79] Translocated bacteria enter the lymphatic system and colonize the liver and spleen, where they grow mainly extracellularly.[80] *Y. pseudotuberculosis* is internalized by normally nonphagocytic cultured mammalian cells via two pathways.[81] The best studied and most efficient of these mechanisms is mediated by the bacterial protein invasin.[82] Members of the invasin family of proteins are found in a variety of enteropathogens, each of which contributes to the interaction with host cells.[83,84]

Invasin mediates entry into eukaryotic cells by binding to members of the β_1 integrin family that lack I, or insertion, domains, such as $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_6\beta_1$, and $\alpha_v\beta_1$[85] (see *Figure 2.4*). Integrins are heterodimeric integral membrane proteins that mediate communication between the extracellular environment and the cytoskeleton by binding to cytoskeletal components and either extracellular matrix proteins or cell surface proteins.[86] Invasin binding to β_1 integrins is thought to activate a reorganization of the host cytoskeleton to form pseudopods that envelop the bacterium.[87] Another family of enteropathogenic bacterial proteins related to invasin, the intimins, does not appear to use integrins as its primary receptors for invasin.[88] Instead, intimins mediate attachment of the bacteria to host cells by binding to a bacterially secreted protein Tir, which upon secretion becomes inserted into the host membrane.[88]

2.3.2.2 Structure of Invasin

Yersinia pseudotuberculosis invasin is a 986-residue protein. The approximately 500 NH_2-terminal amino acids, which are thought to reside in the outer membrane,[89] are related (~36% sequence identity) to the analogous regions of intimins.[90] The COOH-terminal 497 residues of invasin, which make up the extracellular region, can be expressed as a soluble protein (Inv497) that binds integrins and promotes uptake when attached to bacteria or beads.[91] The shortest invasin fragment capable of binding integrins consists of the COOH-terminal 192 amino acids.[89] This fragment is not homologous to the integrin-binding domains of fibronectin [the fibronectin type

III repeats 9 and 10 (Fn-III 9–10)],[90] although mutagenesis studies and competition assays indicate that invasin and fibronectin bind to $\alpha_3\beta_1$ and $\alpha_5\beta_1$ integrins at the same or overlapping sites.[92] The integrin-binding region of invasin also lacks significant sequence identity with the corresponding regions of intimins (~20% identity).[90] To gain insight into enteric bacterial pathogenesis and to compare the structural basis of integrin binding by invasin and Fn-III domains, the crystal structure of Inv497 was solved by Isberg *et al*.[93]

Inv497 was expressed in *Escherichia coli* and purified.[91] The structure was solved to 2.3 Å by multiple isomorphous replacement with anomalous scattering.[94,95] Inv497 is a rodlike molecule with overall dimensions of ~180 Å × 30 Å × 30 Å (see *Figure 2.5 A*), consistent with analytical ultracentrifugation analyses suggesting an extended monomeric structure of the fragment in solution.[93] The Inv497 structure overall resembles that of another $\alpha_5\beta_1$-binding fragment, Fn-III repeats 7 through 10 (Fn-III 7–10),[96] as they are both elongated molecules composed of tandem domains. The first four Inv497 domains (D1, D2, D3, and D4) are composed mainly of β-structure, and the fifth domain (D5) includes α-helices and β-sheets. Despite only 20% sequence identity, [90] the D3 to D5 region of Inv497 is structurally similar to a 280-residue fragment of the extracellular portion of enteropathogenic *E. coli* intimin.[96]

The four NH_2-terminal domains of Inv497 adopt folds resembling eukaryotic members of the immunoglobulin superfamily (IgSF),[97] although the Inv497 domains do not share significant sequence identity with IgSF domains and lack the disulfide bond and core residues conserved in IgSF structures.[90,97] D1 belongs to the I2 set of the IgSF, and D2 and D3 belong to the I1 set.[97] D4 adopts the folding topology of the C1 set of IgSF domains, a fold seen in the constant domains of antibodies, T cell receptors, and major histocompatability complex (MHC) molecules.[97] Unlike these C1 domains, D4 of Inv497 includes a 15–amino acid insertion between strands A and B that forms two additional β-strands (see *Figure 2.5 B*). D1 and D2 of the intimin

fragment are also Ig-like, and the second domain includes an insertion similar to that found in Inv497 D4.[96]

Inv497 D5 is composed of two antiparallel β-sheets with interspersed α helical and loop regions and includes a disulfide bond linking helix-1 to β-strand 5 (see *Figure 2.5 B*).

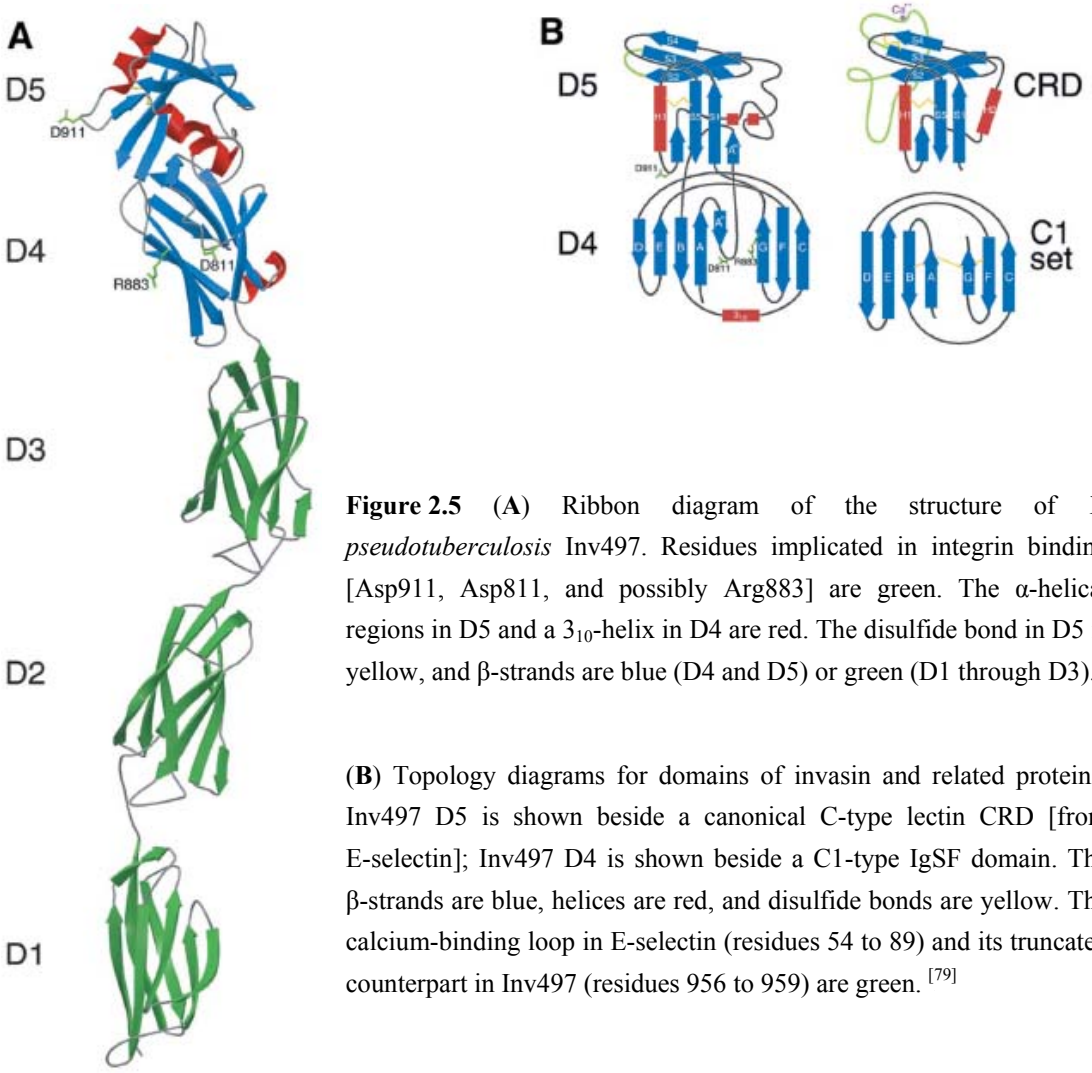


Figure 2.5 (**A**) Ribbon diagram of the structure of *Y. pseudotuberculosis* Inv497. Residues implicated in integrin binding [Asp911, Asp811, and possibly Arg883] are green. The α-helical regions in D5 and a 3_{10}-helix in D4 are red. The disulfide bond in D5 is yellow, and β-strands are blue (D4 and D5) or green (D1 through D3).

(**B**) Topology diagrams for domains of invasin and related proteins. Inv497 D5 is shown beside a canonical C-type lectin CRD [from E-selectin]; Inv497 D4 is shown beside a C1-type IgSF domain. The β-strands are blue, helices are red, and disulfide bonds are yellow. The calcium-binding loop in E-selectin (residues 54 to 89) and its truncated counterpart in Inv497 (residues 956 to 959) are green. [79]

Extensive interactions between Inv497 D4 and D5 create a superdomain that is composed of the 192 residues identified as necessary and sufficient for integrin binding.[89] The interface between D4 and D5 is significantly larger than the interfaces between tandem IgSF domains and between the Ig-like invasin domains (D4 to D5 buried surface area is 1925 $Å^2$ in comparison with ~500 $Å^2$ for IgSF interfaces).[79] The D4-D5 interface is predominantly hydrophobic, although a number of hydrogen bonds are also present (see *Figure 2.6*).

The large buried surface area at the D4-D5 interface (see *Figure 2.5 A*) and the consequent rigidity of this portion of invasin contrasts with the flexibility between the integrin-binding portions of fibronectin, inferred from interdomain buried surface areas that are lower than average at these interfaces (Fn-III 9–10 and Fn-III 12–13).[79,96] Interdomain flexibility in fibronectin was proposed to facilitate integrin binding[97] and is also observed in the structures of two other integrin-binding proteins, ICAM-1[96] and VCAM-1.[96] However, invasin, which shows little or no interdomain flexibility in its integrin-binding region, binds at least five different integrins and binds $\alpha_5\beta_1$ with an affinity that is ~100 times that of fibronectin.[87,92] High-affinity binding of invasin is necessary for bacterial internalization, as studies have shown that bacteria coated with lower affinity ligands for $\alpha_5\beta_1$ bind, but do not penetrate, mammalian cells.[87,92]

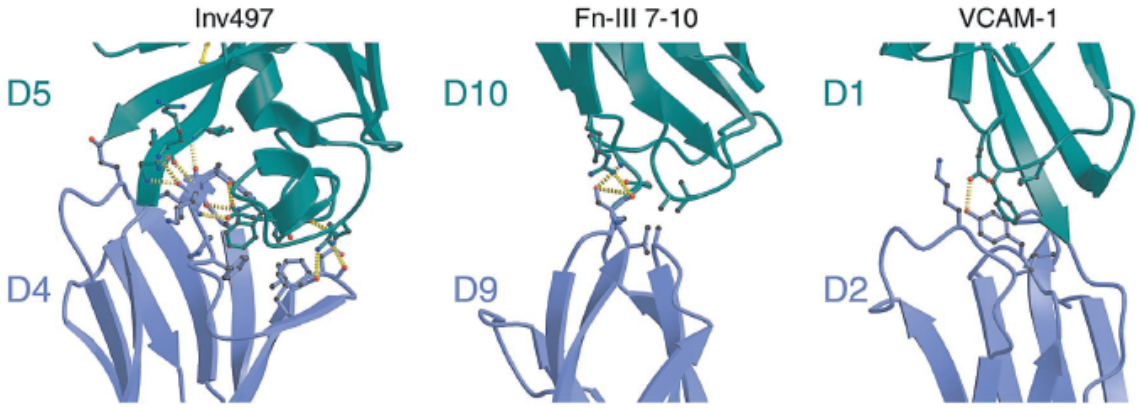


Figure 2.6 Comparison of interdomain interfaces in integrin-binding regions of Inv497 (D4-D5), fibronectin type III repeats 9 and 10 (D9-D10),[96] and VCAM-1 (D1-D2). Hydrogen bonds are shown as dashed yellow lines. Additional hydrogen bonds, van der Waals contacts, and a three- to fivefold larger interdomain surface area stabilize Inv497 D4-D5 and restrict interdomain flexibility, compared to the other interfaces. [79]

Invasin residues that are important for integrin binding include 903 to 913,[89,98] which form helix 1 and the loop after it in D5. Although invasin lacks an -Arg-Gly-Asp- (RGD) sequence, which is critical for the interaction of Fn-III 10 with integrins,[86] an aspartate in Inv497 D5 (Asp911) is required for integrin binding.[89,98] Like the aspartate in the Fn-III RGD sequence, Asp911 is located in a loop (see *Figure 2A* and *4B*). Other host proteins, such as VCAM-1 and MAdCAM-1, which bind integrins

that lack I domains, also contain a critical aspartate residue on a protruding loop.[97] A second region of invasin that is ~100 amino acids apart from Asp911 contains additional residues that are implicated in integrin binding, including Asp811 (see *Figure 2.5 A* and *2.7 B*).[98] This region of invasin is reminiscent of the fibronectin synergy region located in Fn-III 9, which is required for maximal $\alpha_5\beta_1$ integrin-dependent cell spreading.[99] Invasin Asp811 is located in D4 and lies on the same surface as Asp911, separated by approximately 32 Å (measured between carbon-α atoms). The distance between Fn-III 10 Asp1495 in the RGD sequence and Fn-III 9 Asp1373 in the synergy region is also 32 Å,[80] although the side-chain orientation of Asp1373 differs from that of Asp811 in invasin (see *Figure 2.7*). Within the Fn-III synergy region, a critical residue for integrin binding is Arg1379 [32 Å from Asp1495 (see *Figure 2.7 B*)].[99] The invasin synergy-like region also includes a nearby arginine, Arg883 [32 Å from Asp911 (see *Figure 2.7 B*)]. The overall similarity in the relative positions of these three residues suggests that invasin and host proteins share common integrin-binding features.

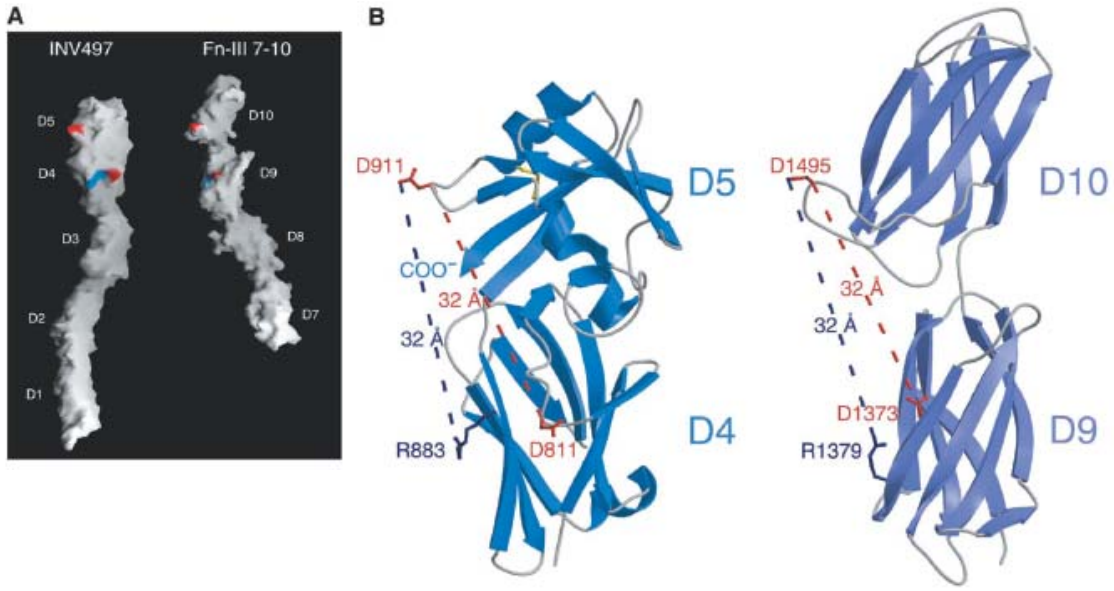


Figure 2.7 Comparison of integrin-binding regions of invasin and fibronectin. Despite different folding topologies and surface structures, the relative positions of several residues implicated in interactions with integrins are similar [Asp811, Asp911, and Arg883 in Inv497; Asp1373, Asp1495, and Arg1379 in Fn-III 9 and 10; (aspartates are red; arginines are blue)].
(**A**) Surface representations of the structures of Inv497 and Fn-III 7-10.[96]
(**B**) Ribbon representations of Inv497 D4-D5 and Fn-III 9-10.[79] Addition of one or more residues to the COOH-terminus of invasin (indicated as "COO^-") interferes with integrin binding,[79] suggesting that the rather flat region between Asp811 and Asp911 is at the integrin-binding interface. By contrast, the integrin-binding surface of fibronectin contains a cleft resulting from the narrow link between Fn-III 9 and 10.

The transmembrane regions of outer membrane proteins of known structure are β-barrels, as represented by the structures of porins.[89] Assuming that the membrane-associated region of invasin is also a β-barrel,[89] the structure of intact invasin may resemble the model shown in *Figure 2.7*, in which the cell-binding region projects ~180 Å away from the bacterial surface, ideally positioned to contact host cell integrins. Similarities between invasin and fibronectin demonstrate convergent evolution of common integrin-binding properties. However, the integrin-binding surface of invasin does not include a cleft, as found on the binding surface of fibronectin (see *Figure 2.6*); thus, invasin may bind integrins with a larger interface. Together with the restricted orientation of the invasin integrin-binding domains, a larger binding interface provides a plausible explanation for the increased

integrin-binding affinity of invasin as compared with fibronectin. Differences between the integrin-binding properties of invasion and fibronectin illustrate how a bacterial pathogen is able to efficiently compete with host proteins to establish contact and subsequent infection, thereby exploiting a host receptor for its own purposes.

2.4 Aims of the Cyclopeptide Invasin Inhibitor Research

Upon the thorough study of the crystal structure of invasin, especially the domain responsible for binding to integrin $\alpha_3\beta_1$, peptide ligands to this integrin that mimic the binding motif of invasin could be rationally designed, imitating both its primary sequence and spatial geometry. This structure-based design could generate the synthetic low-molecular ligands to integrin $\alpha_3\beta_1$ that inhibit the association of invasin to its receptor, which switches off the path of the internalization of invasin, thus in turn blocking the invasion of human pathogenic *Yersinia Sp*. into the cell. If the synthesized peptide/peptidomimetic inhibitors possess higher or matchable binding affinities relative to natural ligand such as invasin, they would be capable of blocking the harmful association of natural ligand to their membrane receptor under physiological conditions, which contains underlying applications in medicine to cure diseases such as gastroenteritis, erythema, reactive arthritis,[51] Graves-Basedow thyroiditis,[57] pneumonic and septicemic plague,[49] which are caused by the infection of pathogenic *Yersinia Sp*..

The designed peptide inhibitors are supposed to block the interactions between the invasin and integrin $\alpha_3\beta_1$. Ideally, the binding between the artificial peptides and invasin is supposed to be as strong as the interactions between the natural ligands and receptors. The most common peptide inhibitors are reversible, and mostly compete

with the substrate fro active site binding. These are known as competitive reversible inhibitors, compounds that have structures similar to those of the ligands (in this case invasin) of receptors (integrin $\alpha_3\beta_1$) and which bind at the ligand binding sites, thereby blocking ligand binding (see *Figure 2.8*). Actually, these peptide inhibitors establish their binding equilibria with the receptors rapidly, so that inhibition is observed as soon as the receptor is assayed for activity, although there are some cases in which inhibition can be relatively slow.[100] A common measurement for inhibition is IC_{50} value, the inhibitor concentration that produces 50 % receptor inhibition in the presence of ligands. The lower the IC_{50} value is, the stronger and more efficient the inhibitor is regarding its inhibitions property.

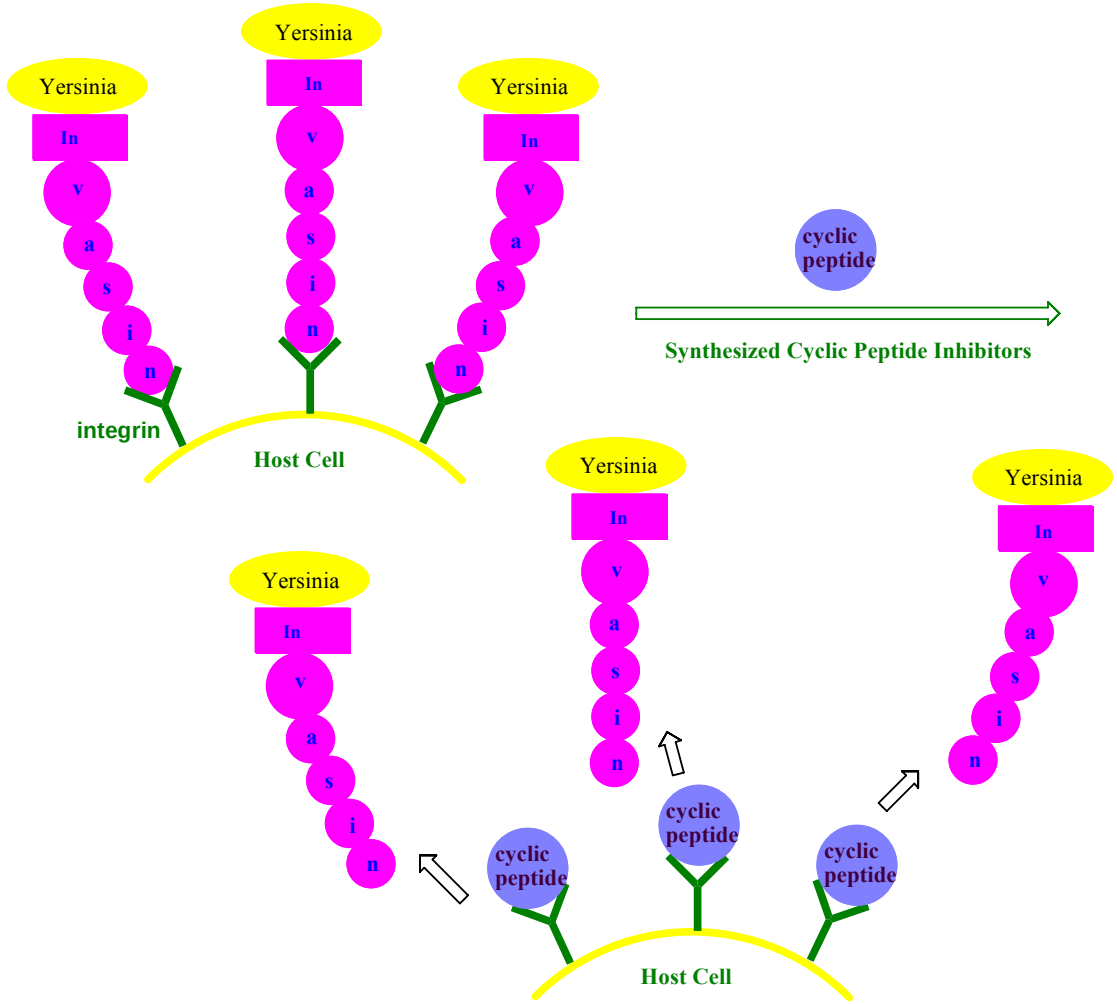


Figure 2.8 Synthesized cyclic peptides as inhibitors for the binding between integrin $\alpha_3\beta_1$ and invasin.

As is already explained, cyclization of peptide could lower the flexibility of the precursor linear peptide and reinforce some certain kind of secondary structure in cyclic peptide by introduction some special amino acid residues. This conformational restriction by cyclization could possibly result in the increase of biological activities of concerned peptides as ligands of some protein receptors.[48] One of the main purposes of this project is to rationally design the high-efficient low-molecular synthetic peptides as ligands to the membrane protein receptor integrin $\alpha_3\beta_1$, which contains underlying applications in medicine domain, such as the blocking of the infections of pathogenic *Yersinia Sp.*.[49]

Through the close study of the crystal structure of invasin, which is the natural ligand to integrin $\alpha_3\beta_1$ promoting the infection of *Yersinia* by internalization into the mammal cells,[82,85] as well as the similarity[92,96] of its binding domains with those of fibronectin, which is also the natural ligand of integrin $\alpha_3\beta_1$,[87,92,97] rational design of peptide inhibitors to integrin $\alpha_3\beta_1$ could be carried out. These structure studies crucially facilitate the structure-based ligand design. By mimicking the integrin binding regions of invasin, a library of peptide inhibitors could be rationally designed and synthesized, among which the lead is supposed to be screened out. Furthermore, through the residue substitution, the binding mechanism could be to some extent elucidated, which can in turn promote the optimization of ligand design.

2.5 Results and Discussion

2.5.1 Structure-based Peptide Ligand Design

2.5.1.1 General Introduction

2.5.1.1.1 State-of-the-art

The state-of-the-art design process relies, in large part, on a quantitative understanding of the molecular recognition of protein-ligand complexes based largely upon analogies to other systems. This structural information is usually obtained from X-ray crystal structures of the protein-ligand complex. Protein crystallography does not determine these structures to atomic resolution; therefore, there is some uncertainty in the exact position of each atom. Given a high resolution structure of a protein, new ligands can be designed completely *de novo*, in which the ligand is designed purely on considerations of protein structure, or more commonly, the design can be based on the modification of known ligands for the target protein.

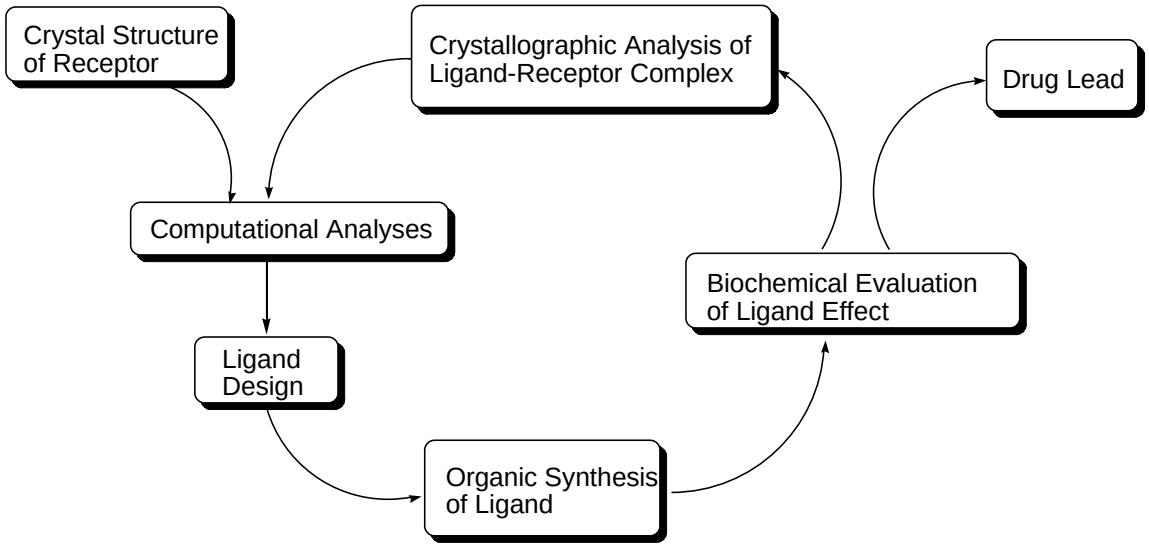


Figure 2.9 Schematic design of protein ligands as drug candidates

The utilization of iterative design processes has greatly advanced the use of protein structure for the design of drug candidates (see *Figure 2.9*).[101,102]

2.5.1.1.2 Cyclic Peptides

As is already elucidated, one interesting property of cyclic peptides is that they tend to be extremely resistant to degradation. This property makes cyclic peptides attractive in drug design for use as scaffolds. Furthermore, the cyclization of the linear precursor peptide could possibly lock the peptide inhibitor in an advantage conformation, thus decreasing the relevant entropy loss upon its binding to the receptor, promoting its binding affinity. In this project, incorporation of special amino acid residue into the peptide sequence and cyclization of peptides was performed in order to finetune the bioactive conformations for integrin $\alpha_3\beta_1$ binding.

2.5.1.1.3 Alanine Scan

A rational approach to investigate the role which the ligand plays is the alanine scan, a method frequently used to screen the amino acid sequence of peptide ligands for the contribution of each residue to the receptor-ligand interaction. In an alanine scan, every single amino acid of a peptide's natural sequence is replaced by L-alanine. Alanine residues in the natural sequence are usually replaced by glycine. For example, a complete alanine scan of neuropeptide Y revealed residues that are important for NPY's interactions with the NPY-Y_1 and NPY-Y_2 receptors.[103] This idea and methodology is especially efficient to elucidate the role of the side-chain residues in the interaction between the ligand and receptor, since the substitutions of these amino acid residues in the natural sequence by alanine eliminate the potential interactions between the ligand and receptor while holding the structure of the original peptide ligand basically untouched.

Based on the findings from an alanine scan, further peptides can be synthesized in which residues involved in the receptor recognition are substituted by homologous amino acids in order to improve the binding affinity between the ligand and receptor. Interesting in this context is also the use of conformationally constrained analogues in which amino acid residues such as the helix breaker proline or the turn-inducing motive Ala-Aib are introduced into the natural peptide sequence.

2.5.1.2 Design of Peptides as Inhibitors for the Binding between Invasin and Integrin $\alpha_3\beta_1$

2.5.1.2.1 Design of the Cyclic Peptide/Peptidomimetic Templates

The integrin binding motif consists of one loop region with conservative Asp911 residue and two synergistic regions. The critical Asp911 residue is found in *i*+1 position of the β turn in the sequence -Ser-Asp-Met-Ser- (see *Figure 2.10*)

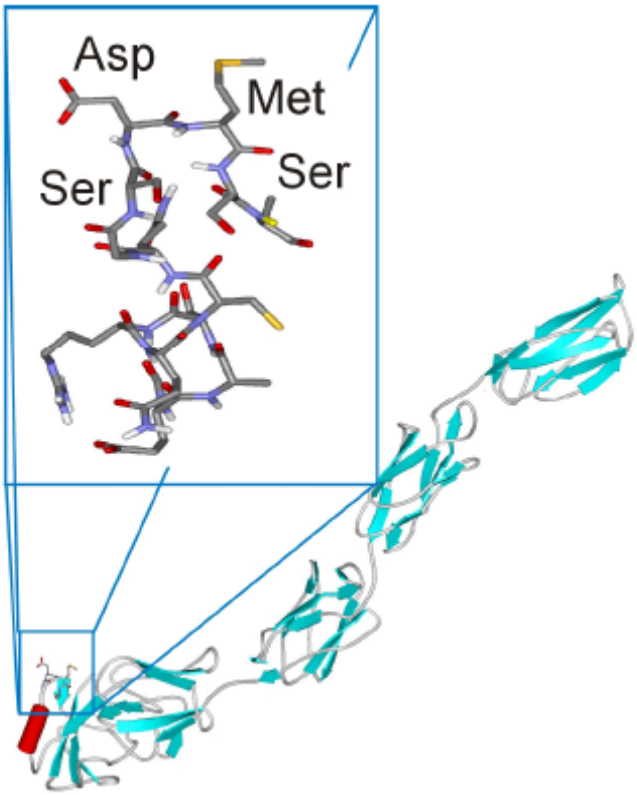


Figure 2.10 Binding Epitope of Invasin with Crucial Asp on *i*+1 Position in β turn

In this project, peptides and peptidomimetics would be designed to mimic the binding moieties in invasin which are responsible for the specific binding to its correlative receptor, namely integrin $\alpha_3\beta_1$, resided on cell surfaces. The crucial primary sequence -Ser-Asp-Met-Ser- found in invasin, which is assumably responsible for the specific binding to integrin $\alpha_3\beta_1$, is supposed to be included in the designed peptide or peptidomimetic inhibitors. Moreover, according to the crystal structure of the native invasin, the corresponding fragment of this sequence adopts a β-turn strcture in which

the crucial aspartate residue is locked in *i*+1 position. The design of peptide inhibitors should hence take this information into account. The synthesized peptide inhibitors should not only fuse the -Ser-Asp-Met-Ser- moiety into their primary sequence, but also guarantee that that this binding moiety be locked in a active conformation.

The cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) peptide was synthesized as template inhibitor of invasin binding to β_1 integrins such as integrin $\alpha_3\beta_1$ which is applied in this project. Peptide conformation of the recognition sequence -Ser-Asp-Met-Ser- is locked for binding to the integrin by incorporation of a single D-amino acid (see *Figure 2.11*). The D-amino acid induces usually a βII'-turn in a cyclic hexapeptide and leads to the formation of an additional complementary β-turn, in which Asp occupies *i*+1 position.

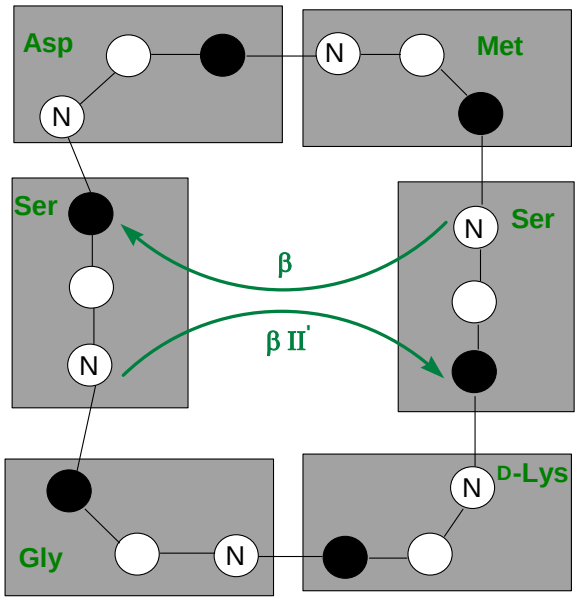


Figure 2.11 Proposed Conformation of cyclo-(-Ser-Asp-Met-D-Lys-Gly-)

Furthermore, a library of synthetic peptide inhibitors (Table 2.2) was synthesized with the aim to elucidate the interaction modes between the invasin/laminin-332 and integrin $\alpha_3\beta_1$ through mutations of component or conformational building blocks of the reference peptide inhibitor cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) (Table 2.3). By

screening the pepide library, the lead could be probably figured out, and the informations gathered in this process is beneficial to rationally optimize the design of the next generation of peptide/peptidomimetic inhibitors for integrin $\alpha_3\beta_1$.

Table 2.2 List of the Library of Synthetic Peptide Inhibitors without Spacers

No.	Sequence
PI 1	**H-Ser-Asp-Met-Ser-D-Lys-Gly-OH**
PI 2	**H-Ser-Asp-Met-Ser-Lys-Gly-OH**
PI 3	**cyclo-(-Ser-Asp-Met-Ser-Lys-Gly-)**
PI 4	**cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-)**
PI 5	**cyclo-(-Ser-Asp-Met(O)-Ser-Lys-Gly-)**
PI 6	**cyclo-(-Ser-Asp-Met(O)-Ser- D-Lys-Gly-)**
PI 7	**cyclo-(-Ser-D-Asp-Met-Ser-D-Lys-Gly-)**
PI 8	**cyclo-(-Ser-D-Asp-Met-Ser-Lys-Gly-)**
PI 9	**cyclo-(-Ser-Asp-Aib-Ser-D-Lys-Gly-)**
PI 10	**cyclo-(-Ser-Asp-Aib-Ser-Lys-Gly-)**
PI 11	**cyclo-(-Ser-Asp-Met-Ser-Lys-Sar-)**
PI 12	**cyclo-(-Ser-Asp-Met-Ser-D-Lys-Sar-)**
PI 13	**cyclo-(-Ser-Glu-Met-Ser-Lys-Gly-)**
PI 14	**cyclo-(-Ser-Glu-Met-Ser-D-Lys-Gly-)**
PI 15	**cyclo-(-Ser-Asp-Met-Ser-Ala-Gly-)**
PI 16	**cyclo-(-Ser-Ala-Met-Ser-D-Lys-Gly-)**
PI 17	**cyclo-(-Ser-Ala-Met-Ser-Lys-Gly-)**
PI 18	**cyclo-(-Ser-Asp-Ala-Ser-D-Lys-Gly-)**
PI 19	**cyclo-(-Ser-Asp-Met-Ser-Arg-Gly-)**
PI 20	**cyclo-(-Ser-Asp-Met-Ser-Gly-)**
PI 21	**cyclo-(-Ser-Asp-Met-Ser-Asp-Met-)**
PI 22	**cyclo-(-Ala-Asp-Met-Ser-D-Lys-Gly-)**
PI 23	**H-Ser-Asp-Met-Ser-OH**

Comparison of the binding affinity of the linear peptide **PI 1** with its cyclic counterpart **PI 4** was performed in order to find out whether the cyclization of **PI 1** could lock the inhibitor in a conformationally advantageous status. Furthermore, compare the binding affinity of **PI 1** with its diastereomer **PI 2** was carried out as well to identify the conformational influence of the incorporation of D-amino acid into cyclic hexapeptide and its effect on inhibitory capacities. Similarly, **PI 2** was synthesized in order to compare the its binding affinity with its cyclic counterpart **PI 3** to find out whether the cyclization of **PI 2** could lead to a bioactive conformation as inhibitory effect was concerned. Compare the binding affinity of **PI 2** with its diastereomer **PI 1** was another purpose to find out the role of configuration of lysine.

Comparison of the binding affinity of **PI 3** with its D-Lys-containing diastereomer counterpart **PI 4** was performed to find out what influence it could bring forward when a β-turn inducer such as D-amino acid is incorporated in the reference cyclic hexapeptide at *i*+1 position of a βII'-turn and Asp is locked in *i*+1 position of the complementary β-turn..

PI 5 was synthesized to explore the role of methionine in the reference cyclic hexapeptide **PI 3** when binding to integrin $\alpha_3\beta_1$. Oxidation of methionine to its sulfoxide derivative could lower the hydrophobicity of the reference peptide and drastically increase its ability as hydrogen bond acceptor. Similarly, **PI 6** was designed and synthesized to find out the role of methionine in the reference cyclic hexapeptide **PI 4** when binding to integrin $\alpha_3\beta_1$.

PI 7 was designed and synthesized with the aim to explore the role of the configuration of crucial aspartate in the sequence -Ser-Asp-Met-Ser- and to find out the advantageous conformation as inhibition is concerned. Comparison of its inhibitory effect with its diastereomer counterpart **PI 4** was performed. Similarly, **PI 8** was synthesized to find the role of the configuration of crucial aspartate in the sequence -Ser-Asp-Met-Ser-. Comparsion of its inhibitory capacity with its diastereomer counterpart **PI 3** was carried out.

The incorporation of Aib into the cyclic hexapeptide **PI 9** has two purposes: Evaluation of the role which the methionine residue plays for binding to integrin $\alpha_3\beta_1$; the occupation of *i*+2 position in the βII (II')-turn motif by a α,α-tetrasubstituted amino acid residue such as Aib[104] could reinforce the stability of β-turn in which the crucial aspartate occupies *i*+1 position as is required. Comparison of the inhibitory capacity of **PI 9** with **PI 4** was performed. Similarly, comparison of the inhibitory capacity of **PI 10** with **PI 3** was carried out with the aim to elucidate the function of methionine upon the association of peptide ligand to integrin $\alpha_3\beta_1$.

The substitution of glyine by its *N*-alkyl amino acid such as sarcosine in the reference cyclic hexapeptide **PI 11** could reinforce the stability of β-turn conformation at -Ser-Asp-Met-Ser- region, as an *N*-alkyl amino acid prefers to occupy the *i*+2 position in the β-turn, thus locking the crucial aspartate in *i*+1 position of hence generated complementary β-turn in the cyclic hexapeptide. Comparison of the inhibitory capacity of **PI 11** with **PI 3** was performed to prove this hypothesis. Similarly, **PI 12** was designed and synthesized in order to compare its inhibitory capacity with that of **PI 4**.

The substitution of Asp by Glu increases the freedom degree of the side chain by introduction a extra methylene group. **PI 13** was therefore designed under this guidance to explore the role the crucial Asp plays upon the association of the ligand with its receptor integrin $\alpha_3\beta_1$. Comparison of the inhibitory capacity of **PI 13** with its Asp-containing counterpart **PI 3** was performed. Similarly, comparison of the inhibitory capacity of **PI 14** with its Asp-containing counterpart **PI 4** was also carried out.

Substitution of Lys by Ala through the systematic alanine-scan is to explore the function of Lys upon the binding of peptide inhibitor to integrin $\alpha_3\beta_1$. **PI 15** was designed and synthesized to find out whether the cationic side chain of Lys under physiological conditions could establish favourable electrostatic interactions with the complementary binding domain on the receptor integrin $\alpha_3\beta_1$.

Substitution of Asp with Ala by systematic alanine-scan is to explore the function of Asp upon the binding of peptide inhibitor to integrin $\alpha_3\beta_1$. Comparison of the inhibitory capacity of **PI 16** with **PI 4** was performed. Similarly, **PI 17** was designed and synthesized to elucidate the function of Asp upon association to integrin $\alpha_3\beta_1$ through the comparison of its inhibitory capacity with that of **PI 3**.

Substitution of Met with Ala by systematic alanine-scan is to explore the function of Met upon the binding of peptide inhibitor to integrin $\alpha_3\beta_1$. **PI 18** was designed and synthesized with this purpose. Comparison of the inhibitory capacity of **PI 18** with **PI 4** was performed.

Substitution of basic residue Lys with basic residue Arg is to explore the function of Lys upon the binding of peptide inhibitor, finding out whether electrostatic interactions exist at this position upon the binding of peptide inhibitor to integrin $\alpha_3\beta_1$. **PI 19** was designed and synthesized with this purpose. Comparison of its inhibitory capacity with those of **PI 3** and **PI 4** was performed.

PI 20 was synthesized with the conversion of cyclic hexapeptide to pentapeptide by excluding Lys. This approach is to find out whether Lys is imperative in the association of peptide ligand to its receptor integrin $\alpha_3\beta_1$. The inhibitory capacity could be improved by shrinking of the peptide size and restricting the flexibility of the concerned inhibitor which results in the decrease of entropic loss upon binding. The alteration of the peptide conformation by excluding lysine residue could also lead to a change of inhibitory capacity.

Incorporation of the dual -Ser-Asp-Met-Ser- sequence in a single cyclic hexapeptide could possibly result in an improved inhibitory capacity, since this crucial sequence happens to start and end with Ser. **PI 21** was therefore designed under this guidance, with the purpose to find out whether the inhibitory capacity could be increased since this motif is crucially important for the binding of peptide inhibitors to integrin $\alpha_3\beta_1$.

Substitution of Ser with Ala through systematic alanine-scan is to explore the function of Ser upon the binding of peptide inhibitor to integrin $\alpha_3\beta_1$. Comparison of the inhibitory capacity of **PI 22** with **PI 4** was performed.

PI 23 was designed and synthesized to find out whether shortening of the inhibitor to linear truncated sequence SDMS could lead to an improved inhibitory capacity.

2.5.1.2.2 Design of the Peptides with Scaffold

As is already discussed in the foregoing chapter, Asp811 ad Arg883 in the native invasin protein could probably also participate the binding of invasin to integrin $\alpha_3\beta_1$. These synergy regions therefore bestow the researcher another domain which could possibly be beneficial to peptide/peptidomimetic inhibitor design. The fusion of these domains into peptide/peptidomimetic inhibitors that contain the -Ser-Asp-Met-Ser- sequence through the rational design and reasonable spatial incorporation, taking the native invasin as template, could assumably enhance the affinity strength of the synthesized peptide/peptidomimetic inhibitors to integrin $\alpha_3\beta_1$, provided that a accurate balance is found between the additions of binding moiety and the entropy effect. This optimization should be obtained by the introduction of a linker with the reasonable length bearing Asp/Arg on the flank to the cyclic peptide with SDMS binding motif (see *Figure 2.12*). According to the crystal structure of invasin, the distance between Asp911 C^{α} and Asp811 C^{α} is 31.54 Å, while that between Asp911 C^{α} and Arg883 C^{α} is 27.29 Å. Two categories of linkers were designed with the aid of Molecular Modelling in order to simulate these distances in the natural protein, as is shown in Table 2.3 and 2.4. In the synthesized peptide derivatives, the distance between C^{α} of aspartate in cyclic peptide and C^{α} of Asp/Arg on the top of the linker peptide is supposed to be kept similar to that in invasin through the incorporation of linker template with the reasonable length. Two categories of linker were designed based on oligo β-Ala and $(Pro\text{-}Pro\text{-}Ala)_n$ which adopt different conformation under physiological conditions.

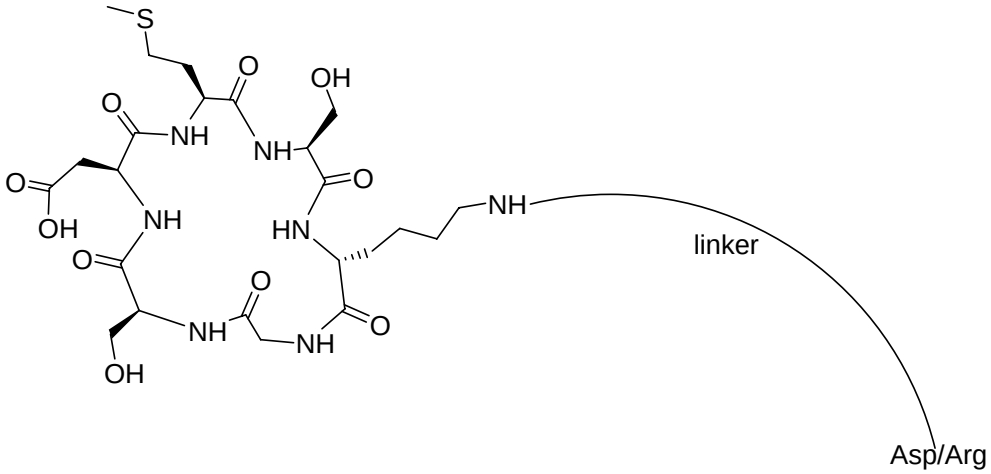


Figure 2.12 Introduction of Linker Bearing Asp/Arg to cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-)

The distance between the C^{α} of Asp911 and the C^{α} of Asp811 in native invasin is approximately 31.54Å, this length falls into the range of polymerization degree between 5 and 7 for linker 1, namely oligo β-alanine, provided that poly β-alanine peptide adopts an extended conformation according to the theory (see also Table 2.3).

Similarly, the distance between the C^{α} of Asp911 and the C^{α} of Arg883 C^{α} in native invasion is approximately 27.29 Å. This length falls into the range of polymerization degree between 4 and 6 for linker1, provided that poly β-alanine peptide adopts an extended conformation according to the theory (see also Table 2.3).

A: c-(-Ser-Asp-Met-Ser-D-Lys (H-Arg-Linker1)-Gly-)

B: c-(-Ser-Asp-Met-Ser-D-Lys (H-Asp-Linker1)-Gly-)

Linker1: **$(\beta\text{-Ala})_n$** extended conformation A (n=4-6), B (n=5-7)

Table 2.3 Correlation between the length of linker $(\beta\text{-Ala})_n$ and degree of polymerization

n	d_{max} (Å)
4	26.2
5	30.4
6	34.4
7	38.4

On the other hand, The distance between the C^{α} of Asp911 and the C^{α} of Asp811 in native invasin is approximately 31.54Å, this length is in accordance with the polymerization degree between 8 and 10 for linker 2, namely $(Pro\text{-}Pro\text{-}Ala)_n$, provided that $(Pro\text{-}Pro\text{-}Ala)_n$ adopts polyproline conformation according to the theory. In the same way, the distance between the C^{α} of Asp911 and the C^{α} of Arg883 C^{α} in native invasin is approximately 27.29 Å, which is in accordance with the polymerization degree between 7 and 9 for linker 2 (see also table 2.4).

A: c-(-Ser-Asp-Met-Ser-D-Lys (H-Arg-Linker2)-Gly-)

B: c-(-Ser-Asp-Met-Ser-D-Lys (H-Asp-Linker2)-Gly-)

Linker2: **$(Pro\text{-}Pro\text{-}Ala)_n$** polyproline helix φ = -79º, ψ = +149º, A (m=7-9), B (m=8-10) (m represents the number of amino acid residues in the linker peptide)

Table 2.4 Correlation between the length of linker $(Pro\text{-}Pro\text{-}Ala)_n$ and degree of polymerization

m	d_{max} (Å)
7	27.0
8	29.6
9	33.2
10	35.2

On these molecular modeling bases, a library of peptide inhibitors for integrin $\alpha_3\beta_1$ was designed and synthesized (Table 2.5).

Table 2.5 List of the Library of Synthetic Peptide Inhibitors with Spacers

No.	Sequence
PI 24	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Arg**-(**β-Ala**)$_4$-Gly-)
PI 25	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Arg**-(**β-Ala**)$_5$-Gly-)
PI 26	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Arg**-(**β-Ala**)$_6$-Gly-)
PI 27	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Asp**-(**β-Ala**)$_5$-Gly-)
PI 28	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Asp**-(**β-Ala**)$_6$-Gly-)
PI 29	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Asp**-(**β-Ala**)$_7$-Gly-)
PI 30	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Arg**-**Ala-Pro-Pro-Ala-Pro-Pro-Ala**-)-Gly-)
PI 31	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Arg**-**Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala**-)-Gly-)
PI 32	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Arg**-(**Pro-Pro-Ala**)$_3$-Gly-)
PI 33	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Asp**-**Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala**-)-Gly-)
PI 34	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Asp**-(**Pro-Pro-Ala**)$_3$-Gly-)
PI 35	Cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-**Asp**-**Ala**-(**Pro-Pro-Ala**)$_3$-Gly-)

2.5.2 Synthesis of Peptides/Peptidomimetics as Inhibitors of the Interaction between Integrin $\alpha_3\beta_1$ and Invasin

2.5.2.1 General Introduction

The peptide synthesis in this project was fulfilled by SPPS (solid-phase peptide synthesis) methodology according to Fmoc/tBu strategy (see *Figure 2.13*). The first amino acid was immobilized on Trityl(2-Cl) chloride resin across its carboxy group. For the synthesis of a cyclic peptide, the choice of the sequence of precursor linear peptide could be multiple and hence should be carefully balanced. In this project glycine was chosen as the *C*-terminus of the precursor linear peptide, since this residue is achiral and racemization upon the activation of the *C*-terminal component during the cyclization reaction could be avoided.

Fmoc protecting group on the *N*-terminus was quantitatively removed by a secondary amine such as piperidine. The second N^{α}-Fmoc-protected amino acid was coupled with activation of the carboxy group by a coupling reagent such as TBTU or HATU with DIPEA as base. The peptide chains bound to the resin were thus elongated until the full sequence was completed. The thus obtained peptidyl resin was treated with 1% TFA in DCM to liberate the fully side-chain protected peptide from the solid support. The synthesis of fully protected linear peptides in this project turned out to be smooth and problem free. Since there were almost no so-called "difficult coupling" in the SPPS of the concerned peptides in this project, most fully protected linear peptides are obtained with almost quantitative yields.

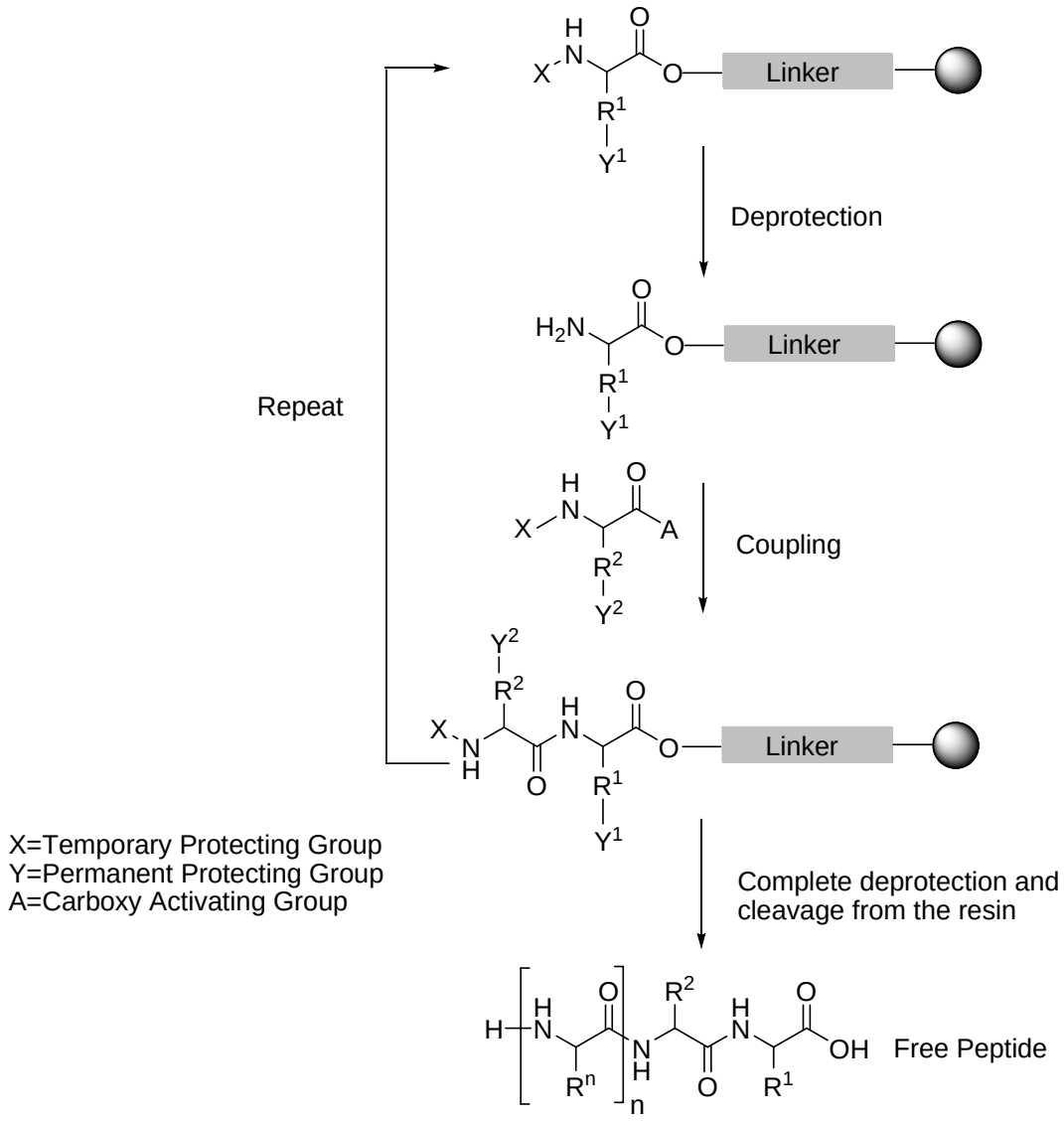


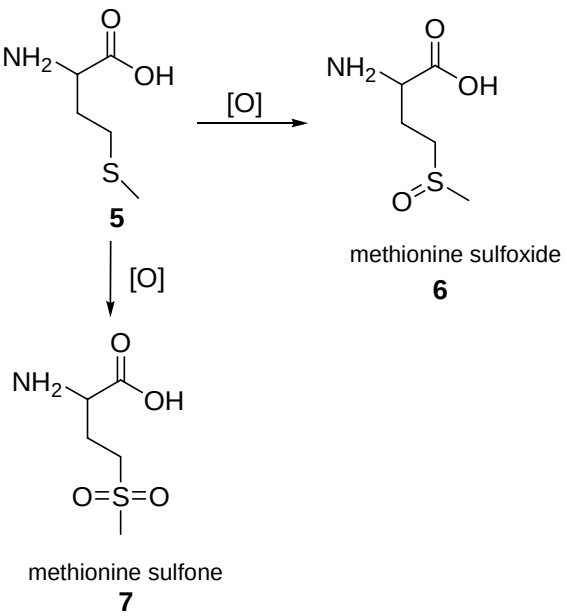
Figure 2.13 Strategy of SPPS

This process can be done manually or by an automated peptide synthesizer (see *Figure 2.14*). Additionally, peptide synthesis was performed under microwave irradiation as the input energy to drive the whole process of SPPS. The irradiation of microwave could possibly destroy the intra/inter-molecular hydrogen bonds which impede the smooth elongation of the peptide chains on the resin. The introduction of microwave to SPPS is therefore advantageous for some certain kinds of peptide synthesis. In this project, however, manual synthesis also achieved comparative and satisfactory yields.

Figure 2.14 CEM Microwave Peptide Synthesizer

2.5.2.2 Avoidance of the Alkylation and Oxidation of Methionine Residues

Two of the main concerns in the synthesis of methionine-containing peptides, no matter in solution or on solid support, are alkylation and oxidation at the sulfur atom on the thioether side chain of the methionine. Carbocations generated at the side-chain deprotection steps are capable of provoking thioether alkylation reactions to give sulfonium salts which might undergo subsequent degradation reactions.[105 - 107] Ox- idation of methionine **5** to the correspond- ing sulfoxide Met(O) **6** or even



Scheme 2.1 Oxidation of Methionine

sulfone **7** occurs in a multistep peptide synthesis (*Scheme 2.1*), especially under the acidic conditions which are applied in the final side-chain cleavage reactions, unless precautions are taken to operate in degassed, peroxide-free solvents under inert gas atmosphere. Acid-catalyzed oxidation of methionine is extremely prone to take place in peptide deprotection and resin-cleavage steps with strong acids. The generation of side-product Met(O) is usually detected with RP-HPLC, since the Met(O) derivatives, which are more polar with the introduction of the oxygen atom, exhibit shorter retention times than the non-oxidized peptides in RP-HPLC.

Since the methionine resiue in the reference peptide inhibitors in this project could play an extremely important role in the interactions between the peptide ligand and receptor integrin $\alpha_3\beta_1$, such as affording the necessary binding affinities and selectivities, it is therefore crucial to assure the thioether functional group in the side chain of methionine to be intact during the whole process of synthesis. Oxidation of thioether to sulfoxide or even sulfone could potentially reduce the binding capacities of the concerned peptide inhibitors, which should be definitely abstained in the synthetic process of this project. During the synthesis of SDMS derivative peptides, every single step is supposed to take place under the protection of inert gas atmosphere such as argon. When a potential peroxide-containing solvent such as ethyl ether was applied in the process of the peptide synthesis, it was distilled before usage in order to remove the peroxide derivatives which might lead to methionine oxidation. Extreme care should be taken for the acidic side-chain removal reactions. The choice of scavengers seems to play an important role in avoiding thioether oxidation. When the routine deprotecting cocktail such as 95%TFA/2.5%H_2O/ 2.5%triisopropylsilane was subjected to the final deprotection, SDMS derivative peptides were partially oxidized under the standard reaction condition. Reagent K, which is composed of TFA, phenol, H_2O, thioansiol, ethane-1,2-dithiol (82.5:5:5:5:2.5), is the preferable choice of deprotection cocktail for the peptides which contain methionine residues. No methionine sulfoxide or methionine sulfone derivatives were detected by MALDI-ToF-MS or RP-HPLC in this project with argon protection and the

application of Reagent K. Meanwhile, no alkylation side-reactions took place upon the acidic side-chain cleavage reactions. The carbocations generated in this reaction were efficiently trapped by the corresponding nucleophilic components (phenol, dithiol, thioether, water) in Reagent K. The thioether functionality in the concerned SDMS derivative peptides remained unoxidized and unalkylated in the whole process of synthesis in this project.

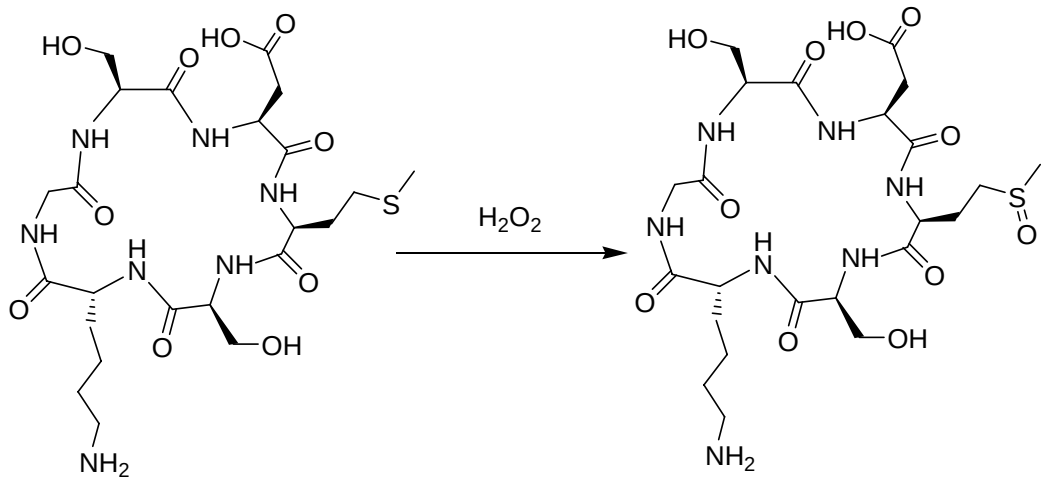
2.5.2.3 Oxidation of Methionine Residue to Sulfoxide Derivative in Cyclic SDMS Peptide

Contrary to the cautions to avoid methionine oxidation in the synthetic process in this project, deliberate oxidation of methionine to methionine sulfoxide in cyclic SDMS derivative peptides was carried out in the synthesis of **PI 5** and **PI 6**, in order to analyze the role which methionine residue plays in the interaction between SDMS binding domain in invasin and integrin $\alpha_3\beta_1$. The interaction mode of methionine to its corresponding binding domains on the receptor could therefore be explored by the comparison of reference peptides **PI 3** and **PI 4** with their oxidized derivatives **PI 5** and **PI 6**.

Generally, oxidation of L-methionine yields a mixture of *S*- and *R*-sulfoxides and, depending on the conditions used, even a mixture of the related sulfone (*Scheme 2.1*). At least by oxidation with hydrogen peroxide the *R*-sulfoxide is formed in a preferred manner,[108] isomeric pure Met(O) derivatives are obtained by isolation of the isomers from the *S,R*-sulfoxide mixture.

The cyclic peptides subjected to the oxidation reaction in this project were fully deprotected (*Scheme 2.2*), which assure them of the avoidance of the final deprotection reaction, as deprotecting cocktails contain substances with reducing abilities, such as phenol and EDT which could possibly reduce the methionine sulfoxide back to thioether derivatives. On the other hand, no other residues in the referred peptides could be influenced by the subjected oxidant, thus validating this

methodology.

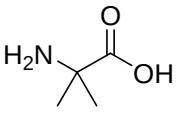


Scheme 2.2 Oxidation of Methionine-Containing Peptide to the Correspoding Methionine Sulfoxide Derivative

The oxidation of cyclic peptide cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) **PI 4** was carried out with utmost caution in order to avoid the further oxidation of sulfoxide functional group to corresponding sulfone derivative (*Scheme 2.1*). Single-Channel-Syringe-Pump was utilized to assure the slow injection of hydrogen peroxide into the cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) **PI 4** solution in order to avoid the real time excess of oxidant relative to unoxidized peptide, assuring the oxidation reaction to retain at the stage of methionine sulfoxide. Subsequent MALDI-ToF-MS and analytical RP-HPLC analysis testified the efficiency of this methodology since neither sulfone nor unoxidized thioether derivatives was detected. The oxidized species of peptides exhibited obvious shorter retention time on analytical RP-HPLC than their unoxidized counterparts.

2.5.2.4 Introduction of α-Aminoisobutyric Acid through Azidoacidchloride

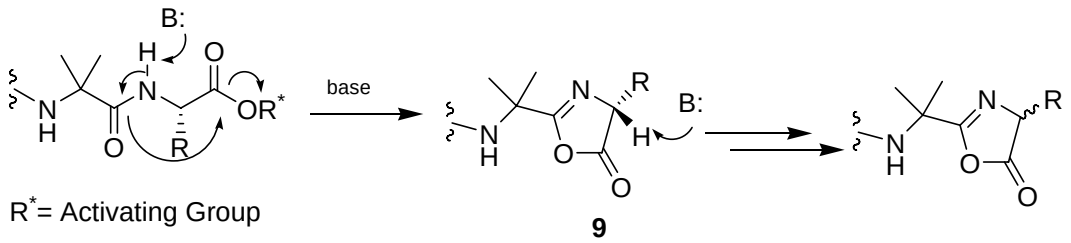
The incorporation of some constraints into the cyclic peptides will induce or reinforce certain kinds of secondary structures. α,α-disubstituted amino acids such as Aib **8** (α-aminoisobutyric acid) possess the inclination to induce the 3_{10}-turn (also known as type III β-turn) at *i*+2 position. To replace the methionine residue by Aib in the reference peptides **PI 3** and **PI 4** could thus potentially reinforce the β-turn in which crucial aspartic acid resides at *i*+1 position. On the other hand, the substitution of methionine by Aib could also testify the function of thioether side chain on methionine upon the association to corresponding binding domain on the integrin $\alpha_3\beta_1$.



8

The incorporation of sterically hindered Aib into the peptide chain proved to be difficult and inefficient with the routine coupling methodology both in solution and on solid support. Due to the sluggish coupling reactions concerning α,α-disubstituted amino acid residues into the peptide chain, series of problematic side reactions are to be aroused in this very process.

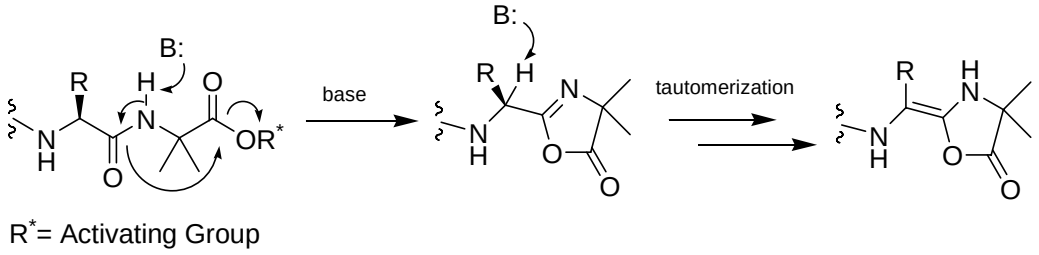
Diketopiperazine cyclization, attributed to the *gem* dialkyl effect, can occur readily with certain Aib-containing dipeptides and related fragment during *N*-deprotection. Another obvious problem associated with Aib incorporation is extensive racemization in R-Aib-Xaa-OH sequences during carbodiimide activation through oxazolone intermediate **9** (*Scheme 2.3*).



9

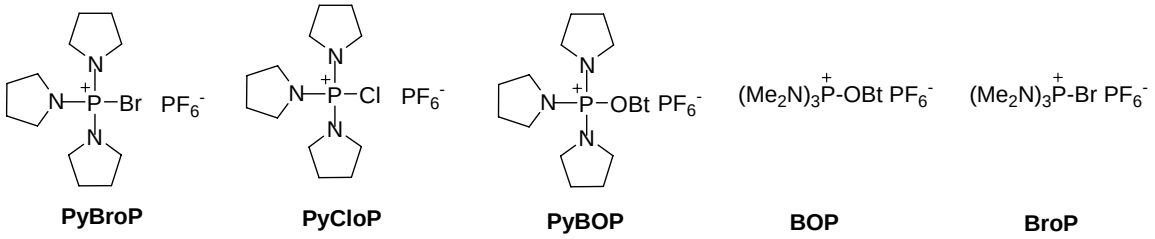
Scheme 2.3 Racemization of *C*-Terminal residue of Peptide R-Aib-Xaa-OH upon the Activation of *C*-Terminus

Racemization of the penultimate residue of peptides containing *C*-terminal Aib residues during activation has also been documented (*Scheme 2.4*) despite the fact that Aib is racemization-free.



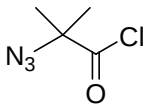
Scheme 2.4 Racemization of the Penultimate Residue of the Peptide with Aib as *C*-Terminus

Despite the advances of phosphonium salts-type coupling reagent with various levels of efficiency of the Aib coupling, such as BOP, PyBOP, BroP, PyBroP, and PyCloP (*Scheme 2.5*), the methodology for this kind of coupling reaction still remains to be improved, especially for those with extremely hindered amino acid derivatives.



Scheme 2.5 Phosphonium Salts-Type Coupling Reagent Applied in the Coupling of Aib Residues into the Peptides

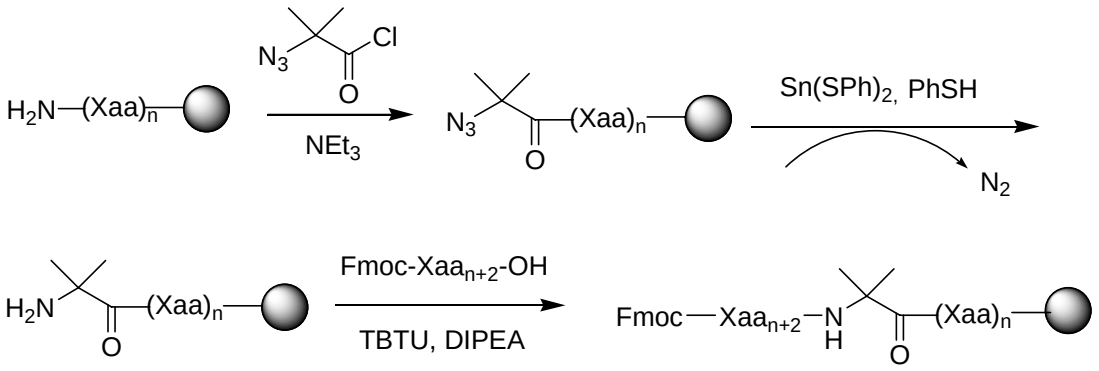
In order to bypass the problems caused by the sluggish coupling of α,α-disubstituted amino acid residues, α-azido acid chlorides, such as 2-azido-2-aminoisobutyric acid chloride **10**, which are readily reduced to α-amino derivative, are introduced as building blocks in the peptide synthesis.



10

As the N^{α}-azido protecting groups is much less hindered than the bulky routine carbamate protecting groups such as Fmoc, Boc, and Z, the incorporation of N_2=Aib-Cl **10** to the growing peptide chain will be much faster and efficient than the routine methods, which lead to the avoidance of the side reactions aforementioned. Furthermore, the acid chloride is more efficient as the property of electrophilicity is concerned.

The N^{α}-azido group is smoothly reduced with $Sn(SPh)_2$ and thiophenol under basic conditions to liberate the free amine. The aforementioned reduction process could be performed on solid support, making this methodology compatible with SPPS (*Scheme 2.6*).



Scheme 2.6 Incorporation of Aib Residue to the Peptide Chain on Solid Support through N_2=Aib-Cl, Subsequent Reduction of Azide to Amine and Chain Elongation

Upon the introduction of Aib residue to the growing peptide chain on solid support, the next incoming Fmoc-amino acid residue could be coupled to the peptide chain, whose *N*-terminal component is Aib, according to the routine methodology, namely, TBTU as coupling reagent and DIPEA as base. The efficiency of this coupling could be confirmed by the MALDI-ToF-MS. Inhibitor peptides **PI 9** and **PI 10** were synthesized successfully according to this method.

2.5.2.5 Cyclization of N^{α}-free side-chain Protected Linear Peptide

General Method of Peptide Cyclization in Solution

The fully protected linear peptide precursor is cyclized in DMF with 3 equiv. HATU as coupling reagent, and DIPEA as base.

Oligomerization Side Reaction in Peptide Cyclization

In the synthesis of cyclic peptides, the readiness of a linear precursor to cyclize depends on the size of the ring to be closed, and usually no difficulties arise with the cyclization of peptides containing six or more amino acid residues. However, the increase of the number of amino acid residues in the linear peptide to be cyclized could lead to side reacitions such as dimerization or oligomerization. Although ring closure with pentapeptides is more hampered, the cyclization is facilitated by the presence of turn structure-inducing amino acids such as glycine, proline or a D-amino acid.[109,110] The only pentapeptide synthesized in this project is **PI 20**, with the sequence cyclo-(-Ser-Asp-Met-Ser-Gly-). The occurence of glycine in the peptide guarantees the smooth ring closure reaction. No linear precursor was detected after the cyclization reaction was completed. For linear peptides which do not contain amino acid residues which stabilize turn structures, the cyclization reaction may be an inherently improbable or slow process, and side reactions, such as cyclodimerization, may dominate even at high dilutions (10^{-3}-10^{-4}M).[111]

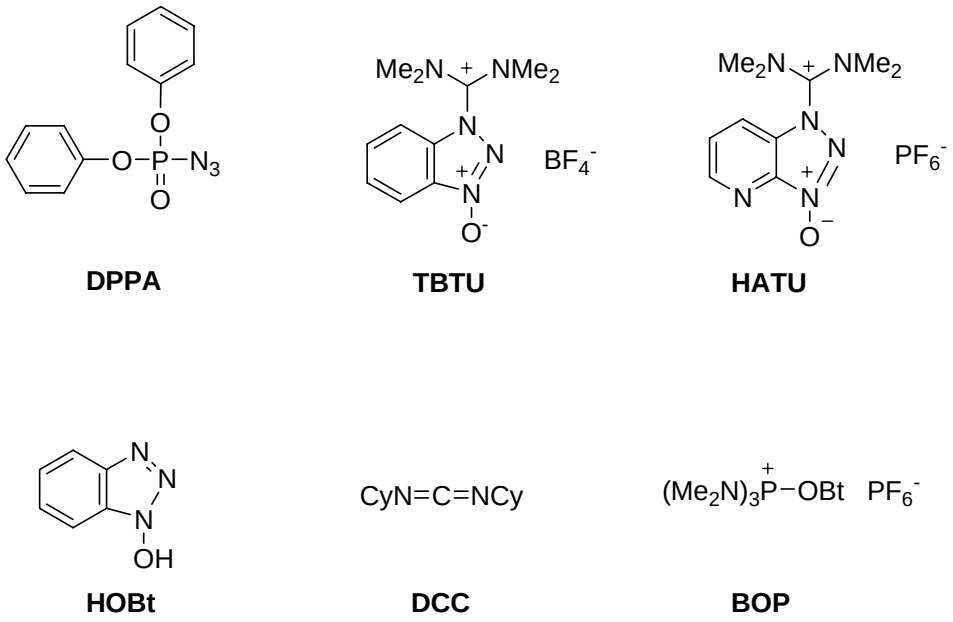
Formation of Guanidinium of *N*-terminus in Peptide Cyclization

The rate decrease of a cyclization reaction could also result in the formation of guanidinium adduct at the *N*-terminus of the linear peptide precursor, since coupling reagents such as HATU are capable of reacting with the amino moiety of the amino acid under basic conditions. In this project, however, no guanidinium adduct side products were found in the cyclization reactions. The occurrence of secondary inducing amino acid such as Gly, Aib, D-amino acid, and Sar favored the successful ring closure reactions, with no dectectable oligomer or guanidinium side products.

The unique cyclic peptide which does not contain such as secondary strucuture inducer is **PI 21**, with the sequence cyclo-(-Ser-Asp-Met-Ser-Asp-Met-). Despite that this cyclic peptide consists of exclusively L-amino acids which do not possess clear secondary structure inclinations, the cyclization reactions turned out to be problem free. The property of hexapeptide seems to be in favor of smooth cyclization in terms of the tension status of an 18-atom-member ring molecule.

Racemization of *C*-terminus in Peptide Cyclization

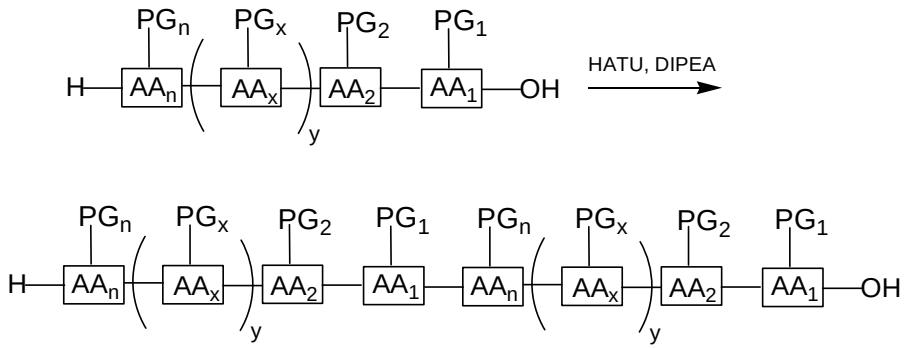
The prolonged existence of the activated carboxyl group resulted from the slow ring closure reaction could increase the likelihood of racemization of the *C*-terminal residue in the linear peptide through the oxazol-5-(4*H*)-one intermediate. The extent of racemization may be diminished by application of the azide method[112] or modified azide methods using DPPA (*Scheme 2.7*).[113] However, these cyclization methods are extremely slow, usually requiring many hours or even several days.[112,114] In comparison with DPPA, TBTU[115] and BOP[116] (*Scheme 2.7*) provide fast cyclization,[117,118] but may also lead to racemization levels which are comparable to those observed with DCC/HOBt.[119] In this project, the racemization of *C*-terminal residue is not the major concern, since the cyclization reactions process at relatively high rates which barely bestow chances to the racemization side reaction through oxazolone formation. More importantly, most linear peptide cyclized in this project possess glycine or sarcosine as their *C*-terminus which are racemization free because of the absence of a chiral center on their C^{α}, with the mere exceptions of **PI 21**. The rational choices for the *C*-terminus in the SPPS of linear peptide precursors avoid the problems of the *C*-terminus racemization upon ring closure reactions.



Scheme 2.7 Coupling reagent applied in peptide cyclization

Initiation of Pseudo High Dilution Condition in Peptide Cyclization

As the main concern in peptide ring closure reaction centers on the possibility of dimerization or oligomerization of the linear precursors to be cyclized under routine conditions, it is therefore necessary to apply a method which could to some extent abstain from the reaction conditions which favor such a side reaction. However, it is not feasible to cyclize a large amount of peptide under high dilution conditions from the aspect of economy. However, dimerization or even higher degree of oligo-merization occurs if the concentration of the peptide to be cyclized is not dilute enough to avoid the collision between the linear peptide molecules (*Scheme 2.8*), especially for the lengths and sequences which are not prone to cyclize under routine conditions, such as pentapeptide, and those which do not contain favorable residues like Gly, α,α-dialkyl amino acid, or D-amino acid. Cyclization on resin could abstain from the problems asscociated with high concentration. However, this methodology requests the *C*-terminal residue to be trifunctional and thus does not possess the universality.

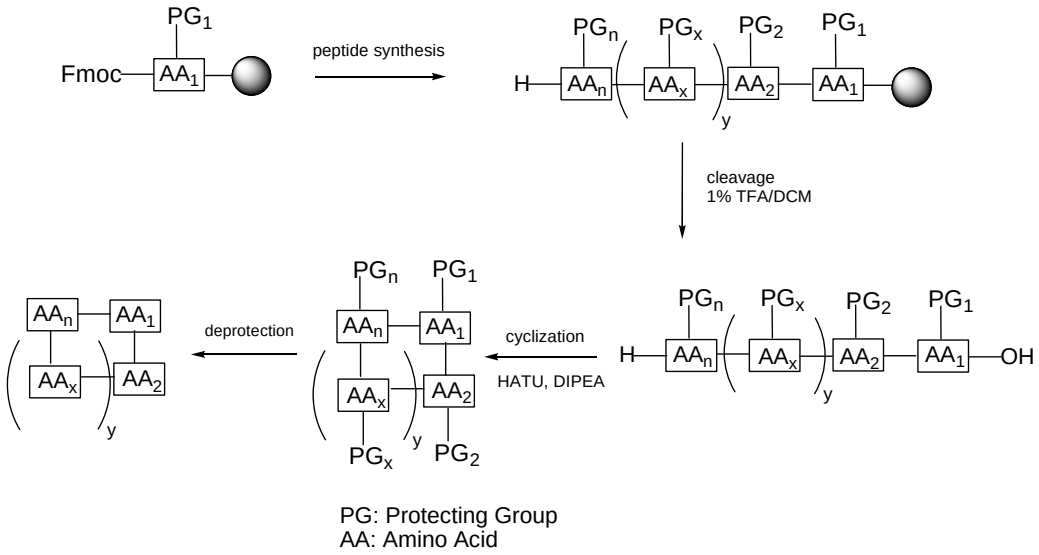


Scheme 2.8 Dimerization side reaction in peptide cyclization

In this project the linear precursor peptides which contain SDMS sequence were synthesized by SPPS (*Scheme 2.9*), the linear peptides were cleaved from the resin by 1%TFA with the side-chain protecting groups remaining intact and subsequently cyclized under the pseudo high dilution conditions which were developed by the Sewald group.[120] (see *Figure 2.15*)

Figure 2.15 Pseudo high dilution conditions generated by two-channel-syringe-pump

According to the grogram, both the peptide solution and the coupling reagent solution are added dropwise by the syringe to the three-neck flask containing solvent for cyclization. The added linear peptide molecules are consumed continually with the conversion to cyclic derivatives in the presence of HATU. The "real-time" dilution condition for linear peptides is thus created. This "pseudo high dilution" condition circumvents the dimerization or oligomerization side reactions which are promoted by the relatively high concentration of linear peptides. Furthermore, this reaction condition is favored from the point of economy consideration.



Scheme 2.9 Procedure of peptide cyclization in solution

No dimerized side products are detected by MALDI-ToF-MS or RP-HPLC through the aforementioned cyclization methodology in this project, which testifies its feasibility to some certain species such as hexapeptides. The herein obtained totally protected cyclic peptide is then subjected to the acid-catalyzed side-chain cleavage reactions. Reagent K is used on this occasion for the reason which has been clarified beforehand. The whole synthetic process proved to be smooth and problem-free.

2.5.2.6 Segment Condensation

2.5.2.6.1 Preconditions of Segment Condensation

As is introduced in chapter **2.5.1.2.2**, the remote synergic Asp/Arg binding moieties are supposed to be incorporated into the cyclic peptides containing -Ser-Asp-Met-Ser- sequence through molecular modelling programmed spacers with certain spatial conformation, in order to enhance the binding affinities through the additional association forces introduced by these remote binding moieties. The introduction of these spacers has to be achieved through a hinge which is capable of "anchoring" the linear spacers covalently to the parent cyclic peptides in which -Ser-Asp-Met-Ser- primary sequence is resided. From the viewpoint of chemical reaction, this operation

needs a reactive functional group on the cyclic peptide which could covalently bond to the corresponding functionality on the spacer. As the aspect of binding between ligand and receptor is taken into consideration, the "consumption" of this functional group on the cyclic peptide should not jeopardize the potential interactions between it and the opposite functionalities on the receptor. The choice of this anchor position is therefore supposed to be balanced between the chemical and biological standpoints. Ideally, the cyclic peptide side chain which does not participate in the association of peptide inhibitors to the integrin $\alpha_3\beta_1$ is chosen preferably. D-Lys residue in the reference peptide **PI 4** is regarded putatively as the starting point to introduce the spacer through its primary ω-amine functional group on the flank of its side chain, which is perfectly in accordance with the peptide synthetic methodologies. On this basis, peptide-based inhibitors **PI 24-35** (see also Table 2.5) were designed and synthesized.

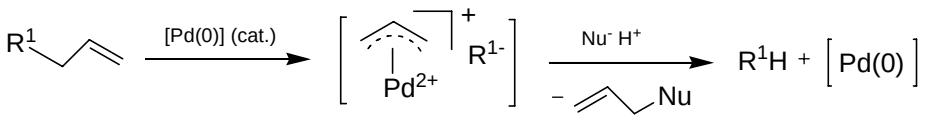
2.5.2.6.2 Aloc as ω-amino Protecting Groups for Orthogonal Deprotection Strategy

The introduction of peptide-based spacers to the parent cyclic peptide containing -Ser-Asp-Met-Ser- sequence at the side chain of D-Lys is performed on the premise that the protecting group for amine be cleaved. However, the acid labile tBu-based routine protecting groups such as Boc, which are compatible with Fmoc SPPS, are not appropriate for the selective removal, since the OtBu protecting group on aspartic acid could be removed as well under the same acidic deprotection conditions. The carboxy group on the aspartic acid side chain should be unconditionally protected during the segment condensation reaction since its reactivity is equivalent or even superior to that of the *C*-terminal carboxy group on the linear scaffold peptide. This brings forward the necessity of selective cleavage of the protecting group on the lysine side chain. Orthogonal protection of lysine is therefore imperative, which means the removal of the protecting group on lysine is supposed to be performed under the condition in which the other protecting groups, especially the one for aspartic acid, remain intact.

The allyl-type protecting groups are characterized by full stability towards moderate strong acids and towards basic conditions as used in Boc/Bzl and Fmoc/tBu strategies in solution and on solid supports, but are readily removed by palladium-catalyzed cleavage. Correspondingly, these carboxy, hydroxy, and amino protecting groups offer a new level of full orthogonality in side-chain protecting for regioselective deprotection of selected functionalities on resin or in solution for additional postsynthetic site-directed chemical reactions. The readily and quantitatively removed allyl-type protecting groups through the application of palladium-mediated cleavage reactions make this type of protecting groups ideal choice for the aforementioned site-selective modification reactions. The orthogonality of allyl-based protection towards Boc and Fmoc groups allows its use for the side-chain protection of linear and branched peptides as well as for glycopeptides. It also enables specific deprotection at preselected side-chain groups for macrocyclization of peptides.

Cleavage of the Aloc Protecting Group

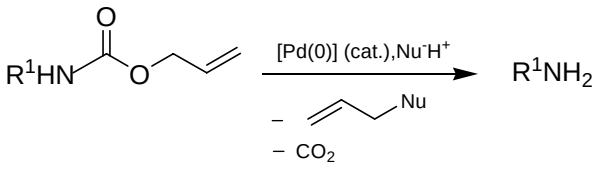
The deprotection of allyl-protectected functional groups is achieved through a palladium-catalyzed transfer of the allylic group to a nucleophilic species as represented in *Scheme 2.10*.



Scheme 2.10 Palladium-Catalyzed Transfer of an Allyl Group to a Nucleophile

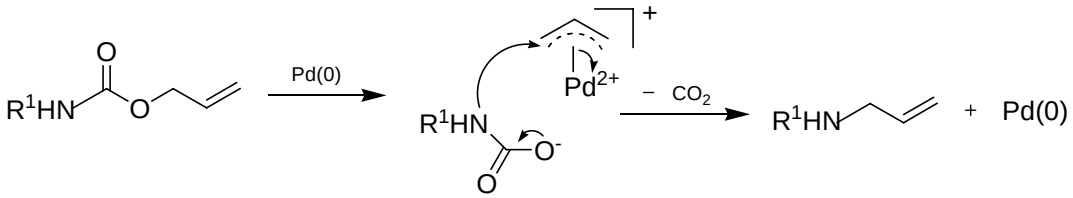
Many complexes of palladium in the zerovalent state, differing by the nature of ligand on the metal, can be used for allylic cleavage but in the majority of the cases it is the commercially available tetrakis(triphenylphosphine)palladium(0) [$Pd(PPh_3)_4$] that is used. $Pd(PPh_3)_4$ partially dissociates in solution into triphenylphosphine and $Pd(PPh_3)_2$, which is assumed to be the true catalytic species. Typically, 2 to 5 mol% of catalyst are used in deprotection reactions in solution but greater amounts are often required in solid-phase chemistry. The reaction is preferentially run under an inert

atmosphere (argon, nitrogen). A necessary condition for successful removal of allyl groups is that the protected function possesses sufficient leaving group ability to allow the formation of the π-allyl complex. As far as peptide synthesis is concerned, allyl esters of carboxylic, phosphonic and phosphoric acids, allyl phenoxides, and allyloxycarbonylderivatives of amines are the main protected functions, which fulfill this condition. On the contrary allyl ethers or sulfides and allylamines in their nonprotonated form are not cleaved by the palladium(0) catalyst. In most cases allylic carbamates undergo decarboxylation contamitantly with allylic cleavage thus leading directly to the amine (*Scheme 2.11*).



Scheme 2.11 Allyl Cleavage with Concomitant Decarboxylation of Allyl Carbamates

The deprotected amine may then compete with the nucleophilic scavenger in the trapping of the η^3-allylpalladium complex. To avoid the formation of allylamines it is therefore better to use deprotection systems that lead to the amine in a protonated or another masked non-nucleophilic form or to use an allyl group scavenger in large excess. In the presence of the palladium catalyst alone allyl carbamates may also form allylamines [121,122] through unimolecular decarboxylative rearrangement of the corresponding carbamato π-allyl complex [122] as presented in *Scheme 2.12*.



Scheme 2.12 Formation of Allylamines from Allyl Carbamates

Fortunately, this process, which is quite rapid in the case of carbamates derived from derived from secondary amines, is relatively slow with Aloc derivatives of primary

amines in general and N^{α}-Aloc derivatives of amino acids in particular.[123] It is, nevertheless, highly advisable not to expose N^{α}-Aloc derivatives to the catalyst without adding nucleophilic scavengers.

Deprotection with Tin or Silicon Hydrides or Amine-Based Complexes

A large variety of nucleophiles may be used as allyl group scavengers in the palladium-mediated deprotection of allylic derivatives, including oxygen, nitrogen, carbon, or sulfur nucleophiles, and hydride donors.[123]

Various hydride donors have been used as allyl group scavengers in palladium-catalyzed deallylation reactions leading to propene as the byproduct. These include formic acid[124] and several semimetal hydrides, including tributyltin hydride (Bu_3SnH),[125] phenylsilane ($PhSiH_3$),[126] lithium or sodium borohydrides,[127,128] and amine-borane complexes such as borane-ammonia complex ($BH_3{\cdot}NH_3$) and borane-dimethylamine complex ($BH_3{\cdot}NHMe_2$).[129] They are efficient not only for the deprotection of allyl carbamates and esters but also for the deprotection of allyl phenoxides, in contrast to the other allyl group scavengers mentioned (*vide supra*). With borohydride and tributyltin hydrides, the air-stable palladium divalent complex dichlorobis(triphenylphosipine)palladium(II) is assumed to be qualified as alternative to of tetrakis(triphenylphosphine)palladium(0) as it is reduced *in situ* to $Pd(PPh_3)_2$. Tributyltin hydride,[125] phenylsilane,[130] and the amine-borane complexes[129] have been used in solid-phase peptide synthesis for Aloc deprotection as these procedures lead to complete cleavage in less than 10 minutes at ambient temperature. Strong activation by palladium of the otherwise poor nucleophile tributyltin hydride is likely to take place through the formation of an intermediate palladium hydrido species. Tributylsin hydride is not indefinitely stable in the presence of palladium zerovalent complexes but decomposes to hexabutyldistannane ($Bu_3SnSnBu_3$) and H_2. This process is sufficiently slow so as not to interfere with the deprotection reaction in the majority of cases. Phenylsilane gives slower deprotection reactions compared to tributyltin hydride, but is apparently indefinitely stable in the presence of palladium

catalyst.[126] When the hydride donor and amine-borane complexes are used as allyl group scavengers, the primary products of deprotection are not free functional groups but the corresponding metallic salts.

2.5.2.6.3 Poisoning of Palladium Catalyst by Sulfur-Containing Molecules

The activity of a catalyst is reduced gradually when the unwanted, harmful components of impurities are accumulated on the catalyst's surface and slowly poison the catalyst. Poisoning is defined as a loss of catalytic activity due to the chemisorption of impurities on the active sites of the catalyst (see *Figure 2.16*). It follows that the analysis of poisoned catalysts may be complicated since the content of poison of a fully deactivated catalyst can be as low as 0.1 wt-% or even less (Forzatti & Lietti 1999). Lead (Pb), sulfur (S), phosphorus (P), zinc (Zn), calcium (Ca), and magnesium (Mg) compounds are typical catalyst poisons.

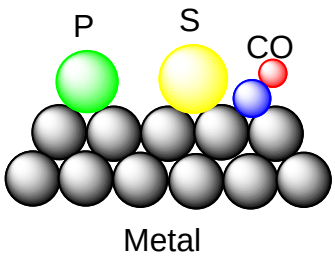


Figure 2.16 Poisoning of Catalyst

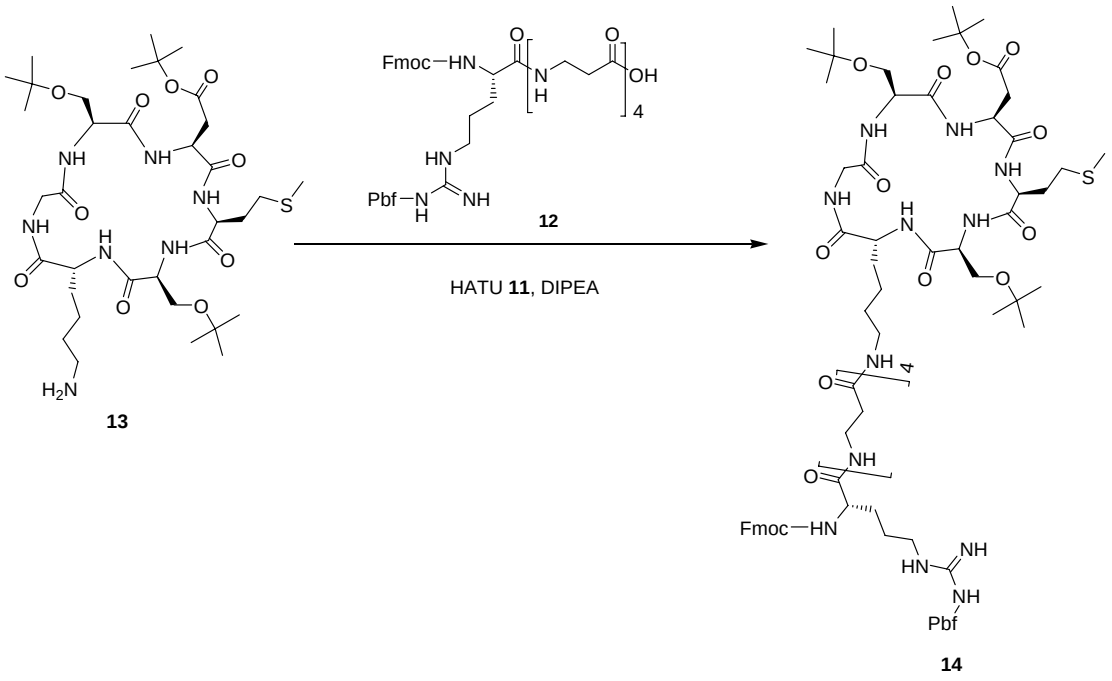
Sulfur species (H_2S, RSH, RSSR…) are poisons[131] for all catalytic processes employing reduced metals as the primary active phase. They are generally considered temporary, although their effect can be permanent depending on the process conditions, ease of regeneration, etc. Pgm (Platinum Group Metal) catalyst(s) react readily with H_2S to form sulfides. Sulfur may cause significant deactivation even at very low concentrations, due to the formation of strong metal-S bonds. Sulfur chemisorbs onto and reacts with the active catalyst sites, preventing reactant access. Furthermore, the stable metal-adsorbate bonds can lead to non-selective side reactions, which modify the surface chemistry.

Since the target peptides contain methionine residues with thioether groups on the side chains, this residue is therefore prone to cause the poisoning of the palladium in the reaction of Aloc cleavage, in which palladium (0) is applied as catalyst. This hypothesis is approved by the experimental results. When 0.1-0.2 equiv $Pd(PPh_3)_4$, according to the theoretical amount, are applied to the Aloc cleavage reactions, the Aloc protecting groups on the side chains of lysines would not be removed quantitatively even with the elongation of the reaction duration. The Aloc cleavage reaction goes to completion only when the amount of palladium catalyst is increased to 1.0 equiv. in terms of Aloc protecting groups. This phenomenon clearly substantiates the fact that sulfur containing molecules are potential palladium catalyst poisons.

One of the methodologies to avoid the poiosoning of Pd cataylsts is to use the liquid ammonia as the solvent of the Aloc deprotecting reaction instead of the routine DMF. The deprotection is carried out at -33 ºC with the reflux of liquid ammonia. 0.2 equiv $Pd(PPh_3)_4$ is enough to achieve the quantitative removal of the Aloc protecting groups with the application of 30 eqv. $PhSiH_3$ as scavenger. It is not clear whether this obvious improvement is contributed to the unique properties of liquid ammonia as solvent or due to the drastic decrease of the reaction temperature, which might influences the balance of the absorbance of sulfur-containing molecules on the surfaces of the palladium catalysts.

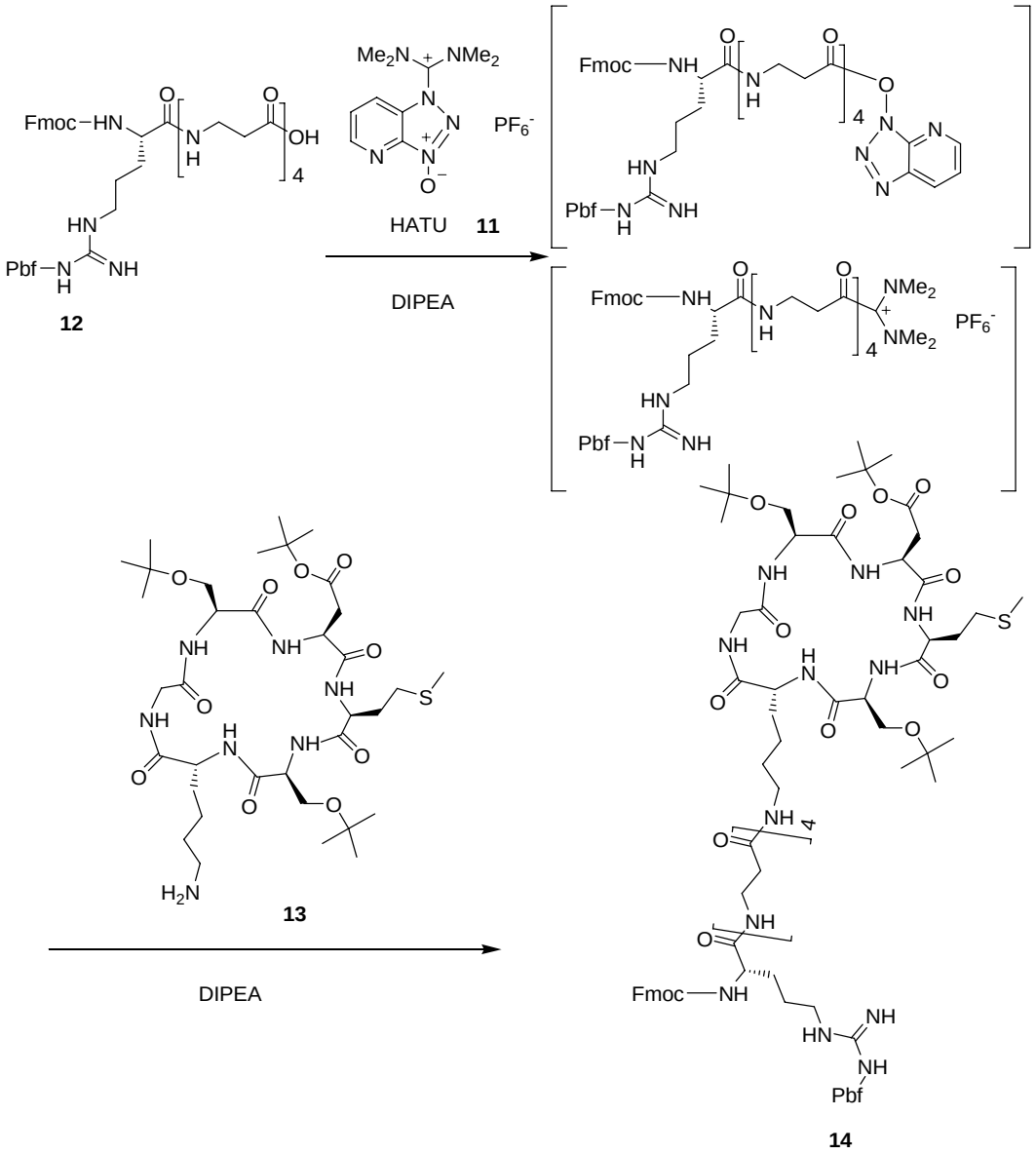
2.5.2.6.4 Strategy of Segment Condensation

The linker peptide, such as **12**, with crucial binding residues (Arg/Asp) on the flank, is supposed to be condensed with the partially protected cyclic SDMS derivative peptide **13** containing free lysine side chain in solution (*Scheme 2.13*) in order to incorporate the remote binding moieties Asp/Arg into peptide complex **14**, which was rationally designed according to the 3-dimensional spatial binding domain in natural invasin protein. This condensation reaction was fulfilled with HATU **11** as coupling reagent and DIPEA as base.



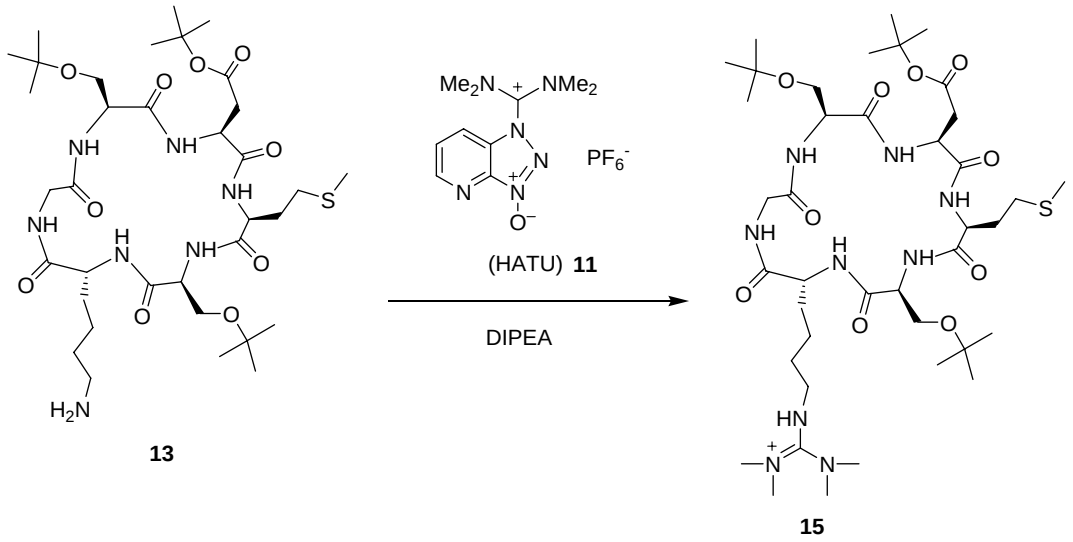
Scheme 2.13 Segment Condensation of Linker Peptide **12** with Cyclic SDMS Derivative Peptide **13** in Solution

Theoretically, this type of condensation reaction could be performed with *in-situ* or preactivation manner. However, in this project the linker peptides should be preferably preactivated at the *C*-terminus before being condensed with the partially protected SDMS derivative peptide **13** containing free lysine residue (*Scheme 2.14*).



Scheme 2.14 Segment Condensation Reaction with Preactivation of the the *C*-Terminal Component

If this condensation reaction is not to be carried out in the preactivation manner, but with the routine *in-situ* activation mode, side reaction concerning guanidinium would stand out. If *C*-terminal component peptide **12** and uronium-type coupling reagent (especially HATU **11**, which possesses stronger nucleophilicity relative to TBTU) were added simultaneously to *N*-terminal component peptide **13** without being premixed, the side product guanidinium salt adducts **15** would be formed through the reaction of the free amino group on **13** with HATU **11** (*Scheme 2.15*). This side product could be detected by MALDI-ToF-MS. It is inert and therefore resistant to the further nucleophilic attack.



Scheme 2.15 Guanidinium Salt Adducts **15** Formation through *in-situ* Activation Reaction

This side reaction could be satisfactorily circumvented with the preactivation of *C*-terminal component peptide **12** with the uronium-type coupling reagent for 5 min with the absence of *N*-terminal component peptide **13** before combining the two peptide components (*Scheme 2.14*). However, this preactivation should not process over too long time in order to minimize the racemization of the *C*-terminal residue.

The conjugates in this project, which were condensed from the partially protected cyclic SDMS derivative peptide **13** and linker peptide **12**, were achieved by preactivation as mentioned above, with the excess of linker peptide **12** by 0.5 equiv in order to achieve the maximal conversion of the cyclic peptide to the target product. The reaction was carried out at ambient temperature overnight under argon atmosphere. The condensation efficiency was monitored by MALDI-ToF-MS, the completion of the condensation being indicated by the disappearance of the signals of peptide **13**. The target peptide conjugate was then purified through RP-HPLC. Due to the difficulties in the application of RP-HPLC in these conjugate species which are highly hydrophobic, the yields of the product after RP-HPLC normally lie below 10%. The obtained pure fully protected conjugates were subjected firstly to 20% piperidine in DMF to remove the Fmoc moiety on the flank of the spacer, and subsequently applied to Reagent K to quantitatively cleave the side-chain protecting groups. The fully deprotected conjugates were purified by RP-HPLC.

2.5.3 Results and Discussions of Inhibitory Experiments

2.5.3.1 Laminin-332 as Natural Ligand to Integrin $\alpha_3\beta_1$

As the association efficiency between invasin and integrin $\alpha_3\beta_1$ is extraordinary high, we chose extracellular matrix (ECM) protein laminin-332 in our ELISA experiment as model ligand to integrin $\alpha_3\beta_1$. It is known (see Table 2.1) that laminin family belongs to the natural ligands of cell adhesion molecules among which integrins stand out.[132] The laminin-binding integrins, such as integrin $\alpha_3\beta_1$, integrin $\alpha_6\beta_1$, integrin $\alpha_6\beta_4$, and integrin $\alpha_7\beta_1$, vary in their binding preference towards the different laminin isoforms. Whereas integrin $\alpha_6\beta_1$ integrin binds to most laminin isoforms, the binding of integrin $\alpha_3\beta_1$ is more restricted to certain members of the laminin family such as laminin-332.[133- 136] It was exhibited that a synthetic peptide derived from the carboxy terminus of the laminin A chain represents a binding site for integrin $\alpha_3\beta_1$,[137] and invasin is able to thoroughly inhibit the binding of laminin-332 to integrin $\alpha_3\beta_1$,[133] suggesting sterically overlapping or identical binding sites. It is therefore reasonable to apply laminin-332 as the natural ligand to integrin $\alpha_3\beta_1$ in our ELISA experiments to measure the relative capacities of synthesized peptides/peptidomimetics as inhibitors for the adhesion between integrin $\alpha_3\beta_1$ and invasin.

2.5.3.2 Methodology of Biological Assay

2.5.3.2.1 General Introduction of ELISA

Enzyme-Linked Immuno Sorbent Assay (ELISA) is a biochemical technique used mainly in immunology to detect the existence of an antibody or antigen in the sample. In the ELISA, antigens are immobilized in the plastic surface before the correspondent antibodies are added, so that specific binding between the antigen and antibody takes place. This antibody is linked to an enzyme, to which the relative substrates are added. The degradation of substrate initiated by the enzyme leads to a visible compound which could be detected by certain spectrometer, so that the amount of antigens in the sample could be determined.

2.5.3.2.2 Sandwich ELISA in Determination of Inhibitory Capacities of Peptide/Peptidomimetic Ligand for Integrin $\alpha_3\beta_1$

Sandwich ELISA is a variant of the ELISA technique, which is metaphorized into "sandwich" because of the topology of the overlapped sample, antibody and second antibody. (see *Figure 2.17*)

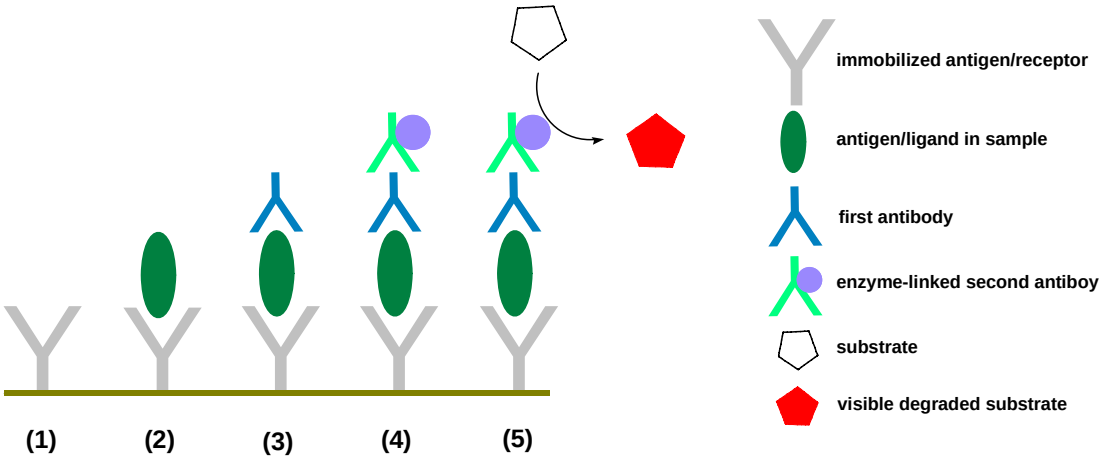


Figure 2.17 Schematic Step of Sandwich ELISA

(1) Coating: Prepare a stationary phase (microtiter plate) by immobilizing certain amount of capture antibody/receptor.

Blocking: Endcap any non-specific binding sites on the plate.

(2) Binding: Apply the antigen/ligand-containing sample to the plate, binding takes place between the immobilized capture antibody or receptor and the corresponding antigen or ligand in the sample.

Washing: Wash the plate, so that any unbounded antigen/ligand is removed.

Fixation: Apply fixation solution to the plate-well in order to fix the association between antigen/antibody or ligand/receptor.

(3) First Antibody: Apply the first antibody that binds specifically to the antigen/ligand.

Washing: Removal the superfluous unbound first antibody.

(4) Enzyme-linked Second Antibody: Apply the enzyme-linked second antibody to the plate, specific binding takes place between the first and second antibody.

Washing: Remove the superfluous unbound second antibody.

(5) Substrate: Apply the substrate to the plate, which is converted by enzyme into a visible metabolite which could be quantified by spectrometry.

Termination: Add the termination solution to the plate to stop the color-generating reaction.

Quantification: Measure the absorbance the of the plate wells to determine the presence and quantity of antigen/ligand in the sample.

2.5.3.3 Screening of Synthetic Peptide Inhibitors of the Interaction between Laminin-332 and Integrin $\alpha_3\beta_1$

As the mechanism of adhesion between invasin and integrin $\alpha_3\beta_1$ remains ambiguous, it is thus necessary to elucidate the modes of their interactions with the screening of a synthetic peptide library. The diversity of the concerned peptides could be established by the mutations of the amino acid residues in the reference peptide (see also Table 2.2 and 2.5). It is instructive to test their capacities to inhibit the adhesion of laminin-332 to integrin $\alpha_3\beta_1$ upon the screening of the individual members in the peptide library. Some preliminary binding mechanisms could be possibly figured out by analysing of the peptides with deliberate mutations at certain positions. For instance, the substitution of an ionic amino acid residue by a neutral one such as alanine could be applied to probe whether there is an ionizable domain on the corresponding receptor. Likewise, a hydrogen bond donor/acceptor residue in reference peptide could be replaced by a hydrogen acceptor/donor to check the hydrogen bond properties of the corresponding domain on the receptor. The inversion of the configuration of a crucial single amino acid residue could also be applied to probe the association manner between the functional domains on the ligand and receptor. The combinatorial mutations on components or conformational building blocks of synthetic peptides are thus instructive to explore the modes of interactions between the ligand and the receptor and thus instructive for peptide ligand optimization. The optimization of inhibitory properties is thus obtainable through

fusing the favorable elements in a single peptide/peptidomimetic molecule.

Figure 2.18 displays the inhibitory properties of the synthetic peptide library. The replaced amino acid functions as a probe to explore the interactions between the peptide inhibitor and the corresponding adhesion moiety of the receptor integrin $\alpha_3\beta_1$. The feedback of such an exploration could be systematic analyzed and the advantageous factors are supposed to be integrated for an improved inhibitory capacity. The concentrations of integrin $\alpha_3\beta_1$ and peptide inhibitor applied in ELISA are 50 nM and 1.8 mM, respectively. The inhibitions were performed at room temperature under physiological pH 7.4. The relative inhibition was calculated according to the reference of negative blank (applying EDTA in the inhibition as the supplement, blocking the adhesion of integrin $\alpha_3\beta_1$ to immobilized laminin-332, which could be regarded as that 100% inhibition takes place) and positive control (adding no peptide inhibitors in the experiment, permitting the full adhesion of integrin $\alpha_3\beta_1$ to immobilized laminin-332, which could be regarded as that no inhibition took place). For concrete procedures see also experiment part.

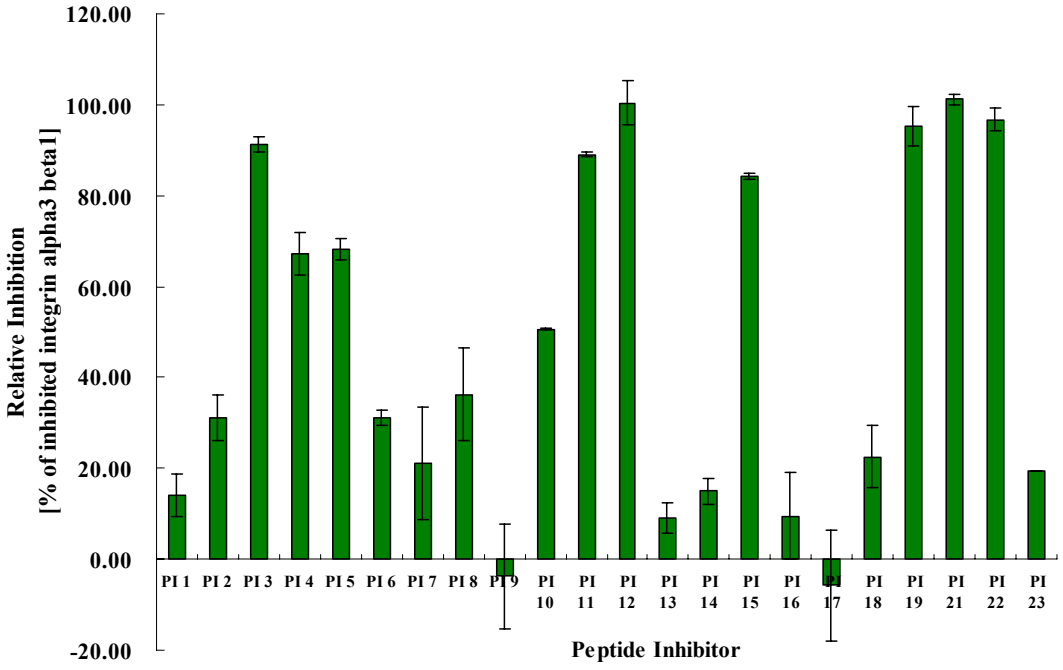


Figure 2.18 Inhibiton of the Interaction between Laminin-332 and Integrin $\alpha_3\beta_1$ by the Synthesized Peptides.

It is discernable from *Figure 2.18* that the synthesized peptide inhibitors possess divergent inhibitory capacities in terms of the association of laminin-332 with integrin $\alpha_3\beta_1$. Peptide **PI 3** and **PI 4**, in which SDMS binding sequence is resided, could be regarded in this project as the reference peptide inhibitors. The evaluation of the inhibitory properties of other peptides/peptidomimetics could be determined through the comparison to these reference peptide inhibitors.

2.5.3.3.1 Substitution of Aspartate in Reference Peptides

Through the X-ray structure of invasin and its similarity to other known integrin ligands, especially Fibronectin, the Asp911 residue in the native invasin was assumed to play a crucial role in the adhesion of invasin to integrin $\alpha_3\beta_1$. This assumption could be validated in the inhibitory experiments of the peptide derivatives with mutations in the primary sequence at the Asp postion. The Asp could be substituted by Ala, which diminished its side-chain interactions to the corresponding domain of the receptor; or by Glu, which insert an additional methylene group in the side-chain, thus extending the scope of its interactions. These operations would not change the conformations of the reference peptides basically, which sustains the validity of this methodology and the subsequent analysis. Alternatively, the configuration of Asp could be changed to its enantiomeric counterpart, namely D-Asp. This operation could elucidate the favorable configuration of the Asp upon the adhesion to the corresponding moiety of the receptor.

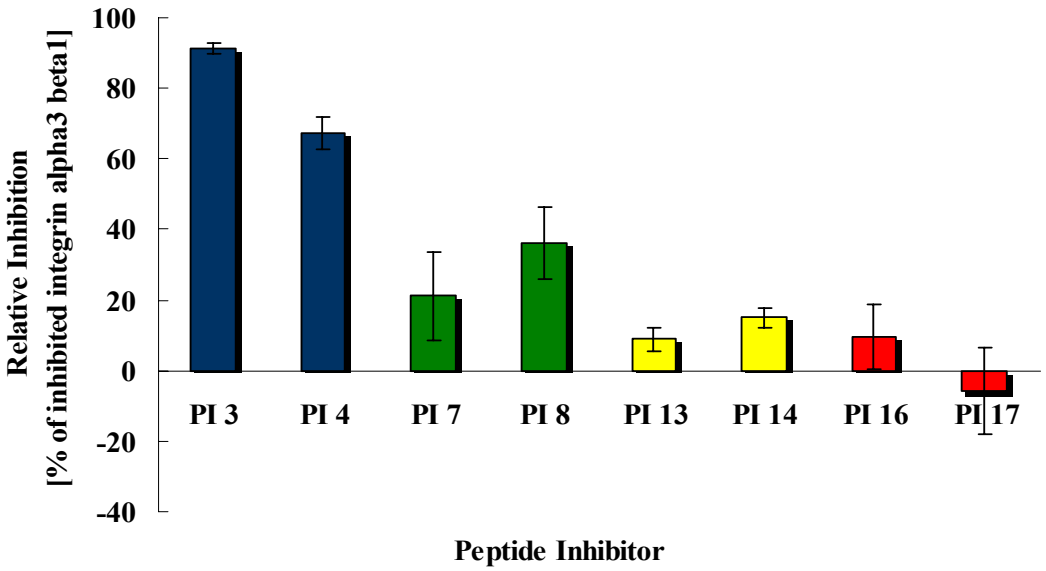


Figure 2.19 Inhibitory Results of the Peptides with Mutations at Asp in the Reference Peptides

Figure 2.19 clearly shows the importance of the Asp for the adhesion of the peptides to integrin $\alpha_3\beta_1$. The peptide derivatives with the mutation at Asp position relative to reference peptide **PI 3** and **PI 4** lost their inhibitory abilities drastically, as is shown in *Figure 2.19*.

PI 16 and **PI 17**, with the sequence cyclo-(-Ser-**Ala**-Met-Ser-D-Lys-Gly-) and cyclo-(-Ser-**Ala**-Met-Ser-Lys-Gly-), respectively, were almost thoroughly deprived of the inhibitory properties upon the alanine scans at the Asp position. These results fully elucidate the crucial roles of the side-chain of Asp in the adhesion of the ligand to integrin $\alpha_3\beta_1$, since the mutation to Ala abolished both the fully contact of the peptide inhibitor to its receptor and the ionic interactions between the two binding participants.

PI 13 and **PI 14**, with the sequence cyclo-(-Ser-**Glu**-Met-Ser-Lys-Gly-) and cyclo-(-Ser-**Glu**-Met-Ser-D-Lys-Gly-), respectively, showed obviously inferior inhibitory properties upon the Glu substitution at the Asp position in the reference peptides. This information indicated the accuracy of the steric complementary between the ligand of invasin at this position and the corresponding binding moiety

on the receptor integrin $\alpha_3\beta_1$, since the similarity of Glu and Asp is evident, no matter from the aspect of pK_a or their comformational inclinations in peptides. Substitution of Asp or Glu results in the dominant difference of the extension of the side-chain by a methylene group which elongates the interaction scope of the carboxy group on its flank. However, as the inhibition results showed, this operation hugely decreased the inhibitory properties of the reference peptides, which account for the precision of the contact between Asp position on invasin and opposite domains on integrin $\alpha_3\beta_1$. On the other hand, the introduction of an extra methylene group into the side chain of Asp increased the flexibility of the concerned molecule, which could result in the decreased activities. A flexible molecule can adopt several conformations and is thus less likely to be in its bioactive conformation. The association of Glu containing peptide inhibitors to their recptors would result in relative larger entropy cost which is adverse to tight binding. Furthermore, the increase flexibility of the ligand, which increases the chances of interacting with a variety of targets, could possibly lead to inferior selectivity. This hypothesis was nevertheless not exhibited in this experiment.

PI 7 and **PI 8**, with the sequence cyclo-(-Ser-**D-Asp**-Met-Ser-Lys-Gly-) and cyclo-(-Ser-**D-Asp**-Met-Ser-D-Lys-Gly-), respectively, indicated markedly decrease of inhibitory capacities. The inversion of the configuration of the Asp resulted in the distinct loss in the binding efficacy of reference peptides. It showed that the favorable configuration of Asp adopted upon the binding to integrin $\alpha_3\beta_1$ should be L-configuration. The change to its enantiomer led to the inappropriate directions of the aspartate side chain, which prevented the full contact between the aspartate and opposite binding domain on the integrin $\alpha_3\beta_1$. The conformational change upon the binding of D-Asp containing peptides to integrin $\alpha_3\beta_1$ would result in a bigger entropy loss which was adverse to the efficient binding.

2.5.3.3.2 Substitution of Methionine in Reference Peptides

The X-ray structure of invasin and its similarity to other native integrin ligand did not present a clear image of the functions of methionine upon the adhesion of invasin to integrin $\alpha_3\beta_1$. This defect could be remedied through the introduction of peptide derivatives with the mutation at the methionine position, such as alanine scan or oxidation of methyl sulfide side-chain of methionine into sulfoxide or even sulfone derivatives. Through the screening of these mutation species, the functions of methionine in adhesion of invasin to integrin $\alpha_3\beta_1$ could be extracted.

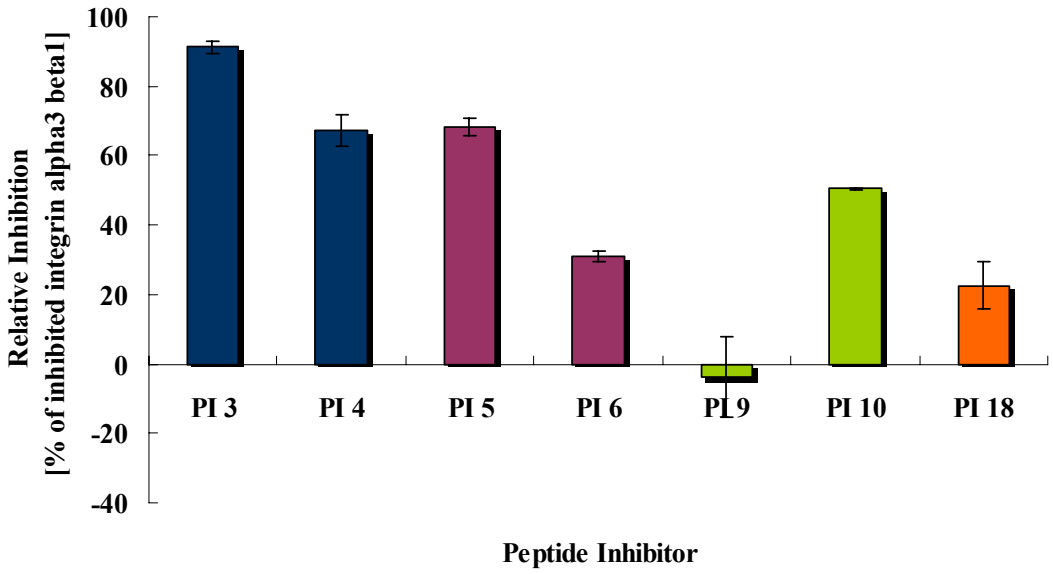


Figure 2.20 Inhibitory Results of the Peptides with Mutations at Methionine in the Reference Peptides

Figure 2.20 explained the important roles played by methionine in the binding of peptide inhibitors to integrin $\alpha_3\beta_1$. The inhibitory capacities of peptide derivatives with the mutations at methionine relative to reference peptides **PI 3** and **PI 4** were reduced to different extent, as shown in *Figure 2.20*.

PI 5 and **PI 6**, with the sequence cyclo-(-Ser-Asp-**Met(O)**-Ser-Lys-Gly-) and cyclo-(-Ser-Asp-**Met(O)**-Ser-D-Lys-Gly-), respectively, displayed slightly reduced inhibitory properties relative to their parent peptides **PI 3** and **PI 4**. The oxidation of sulfide of the side-chain of methionine in **PI 3** and **PI 4** to the sulfoxide derivatives in **PI 5** and **PI 6** might not alternate the conformation of the cyclic peptides drastically.

However, what it did change is the hydrogen-bond properties of the side-chains. The sulfide group in the side-chain of methionine could be regarded as a very weak hydrogen-bond acceptor which forms very weak or no hydrogen bonds,[138] while its oxidated state possess a much stronger hydrogen-bond capacity as the acceptor. The reduced inhibitory properties of **PI 5** and **PI 6** relative to **PI 3** and **PI 4** indicated that there is a very weak or no hydrogen-bond donor on the corresponding binding domains of integrin $\alpha_3\beta_1$. It was already explained in chapter **1.2.4.1** that a burial of unpaired polar groups upon ligand binding leads to a loss of binding affinity and should be avoided. The oxidation of sulfide side-chain of methionine to sulfoxide increased its hydrophilicity, while there was no matching component in the receptor, which led to a reduced binding affinity between the peptide inhibitors and integrin $\alpha_3\beta_1$.

PI 9 and **PI 10**, with the sequence cyclo-(-Ser-Asp-**Aib**-Ser-D-Lys-Gly-) and cyclo-(-Ser-Asp-**Aib**-Ser-Lys-Gly-), respectively, lost their binding affinities extraordinarily upon the substitution of Met by Aib in the reference peptides. The α,α-disubstituted amino acids such as Aib possess inclinations to occupy the *i*+2 position of a type-II (II') β-turn as well as *i*+1 and *i*+2 position of type-I (I') and type-III (III') β-turn.[104] The incorporation of Aib at the position originally occupied by methionine could reinforce the β-turn structure of the cyclic hexapeptide in which aspartate locked at the favorable *i*+1 position. However, as is shown by *Figure2.24*, the potential sustainment of the β-turn did not result in the optimized inhibitory capacities of the concerned pepides compared to reference peptides. On the contrary, their binding affinities to integrin $\alpha_3\beta_1$ were obviously weakened. These results showed the roles of methionine in the adhesion of peptide inhibitors to integrin $\alpha_3\beta_1$. Mutations at this place are supposed to be carried out with extreme caution.

PI 18, with the sequence cyclo-(-Ser-Asp-**Ala**-Ser-D-Lys-Gly-), exhibited a markedly reduced binding affinity compared to the reference peptide **PI 4**. These results sustained the conclusion that methionine plays an important role in the adhesion of

peptide inhibitors to integrin $\alpha_3\beta_1$. As the reason why it possessed better inhibitory capacity relative to Aib-containing counterpart **PI 9**, a`conformational research should be initiated to compare the differences between the two peptides. It was assumed that the incorporation of Aib locked the aspartate into an unfavorable type of β-turn which weakened its fully contact with the corresponding binding domains on integrin $\alpha_3\beta_1$.

2.5.3.3.3 Substitution of Serine in Reference Peptides

The function of serine in the reference peptides could be studied through the mutation operations such as alanine scan.

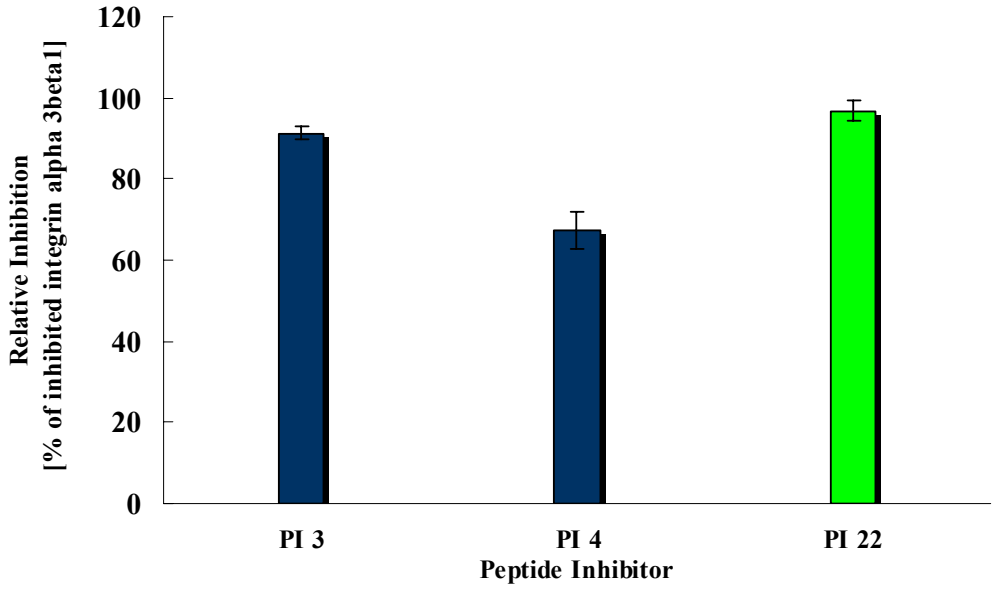


Figure 2.21 Inhibitory Results of the Peptides with Mutations at Serine in the Reference Peptides

Interestingly, the binding efficacy of the reference peptides **PI 4** was evidently increased upon the substitution by alanine, as is shown by *Figure 2.21* that peptide **PI 22**, with the sequence cyclo-(-**Ala**-Asp-Met-Ser-D-Lys-Gly-), indicating optimized inhibitory property compared to its Ser-containing counterpart **PI 4**. As the ethyl hydroxy group on the side-chain of serine can function as both hydrogen donor and acceptor, while the methyl side chain of alanine prevents the formation of hydrogen bonds, it is therefore assumed that the serine in **PI 4** might not be able to discover the

suitable hydrogen donor or acceptor on the integrin $\alpha_3\beta_1$. This unpaired polar group could possibly lead to the destabilization of the peptide/integrin complex. The substitution of serine by alanine was then able to counteract this adverse factor, so that the binding capacity was increased.

2.5.3.3.4 Substitution of Lysine in Reference Peptides

The incorporation of a lysine residue in the reference peptides was based on two reasons: on one hand, the D-lysine at its position was to induce a β-II' turn in which it occupies *i*+1 position, thus locking the aspartate in a complementary β-turn at *i*+1 position; on the other hand, the amine group on the side-chain of lysine functions as a hinge where the spacer peptide is able to be introduced. Mutations at this position was carried out with Arg, Ala, and L-Lys, in order to find out whether the side-chain of lysine plays a role on adhesion of the peptide inhibitors to integrin $\alpha_3\beta_1$, as well as the influences of the conformations of the cyclic peptide inhibitors on their binding affinities.

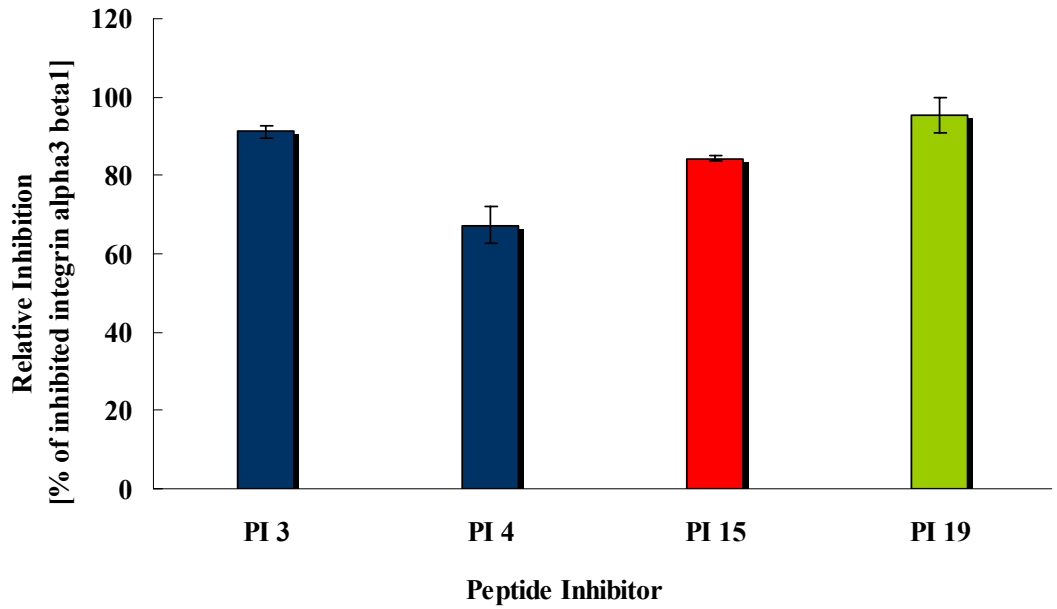


Figure 2.22 Inhibitory Results of the Peptides with Mutations at D-lysine in the Reference Peptides

Inhibitors **PI 3** and **PI 4** are diastereomers with the mere difference at the configuration of lysine. Interestingly, the introduction of D-lysine at its position did not bring forward the improved inhibitory properties. On the contrary, peptide **PI 3** which contains L-lysine at the same position in the cyclic hexapeptide displayed superior inhibitory capacity, as is shown by *Figure 2.22*. It is known that the D-amino acid preferentially occupies the *i*+1 position of a type II' β-turn in a cyclic peptide,[104] the Asp in the reference cyclic hexapeptide cyclo-(-Ser-Asp-Met-Ser-**D-Lys**-Gly-) is therefore supposed to be locked passively in a complementary β-turn. However, the generated complementary β-turn around aspartate is not necessarily the most appropriate type. As the unique amino acid residues bearing no side-chain, Gly preferentially occupies the *i*+2 position in type II β-turn and the *i*+1 positioins in type II' β-turn.[139] (see *Figure 2.23*) As is known, glycine is often regarded as a "proteinogenic D-amino acid", the incorporation of glycine into the cyclic hexapeptide exerts its influence of conformational inclination to induce β-turn. In peptide **PI 4** glycine probably occupied *i*+1 position in a type II β-turn, leading to a favourable complementary β-turn in which aspartate resides in *i*+1 position. This is the putative explanation for the better inhibitory capacity of **PI 4** relative to its D-Lys-containing counterpart **PI 3**. Another possible explanation for the discrepancy between D- and L-Lys containing peptides is that the side chain of D-Lys might induce slightly disadvantageous interactions with the receptor compared with that of the L-configuration.

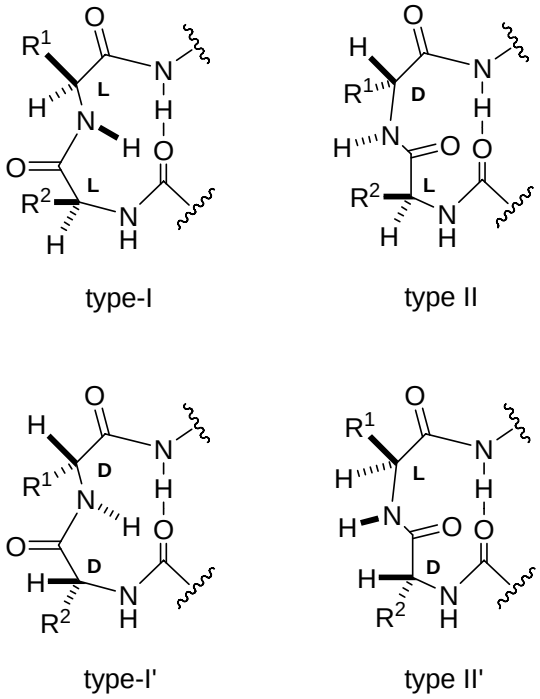


Figure 2.23 Four Most Common Types of β-turn

PI 15 and **PI 19**, with the sequence cyclo-(-Ser-Asp-Met-Ser-**Ala**-Gly-) and cyclo-(-Ser-Asp-Met-Ser-**Arg**-Gly-), respectively, were provided with the matchable inhibitory capacities compared to the L-Lys-containing reference peptide **PI 3**. Seemingly, the lysine residue in the cyclic hexapeptide inhibitors did not exert obvious influences on the adhesion of the peptides to integrin $\alpha_3\beta_1$, since the binding properties are to some extent indifferent to the mutations at this position. This results bestow us the rationality to choose the side-chain of lysine in the cyclic hexapeptide to introduce the spacers with remote synergic binding moieties at the flank.

2.5.3.3.5 Substition of Glycine in Reference Peptides

The functions of Gly in the referred cyclic hexapeptide could be probed through the *N*-alkylation, namely, substitution by sarcosine. *N*-alkyl (or imino acid) residues exhibit reduced preferences for the *trans* conformation normally assumed by secondary amides, and this effect can lead to biologically relevant β-turn structures, similar to those often induced by proline residues.[140]

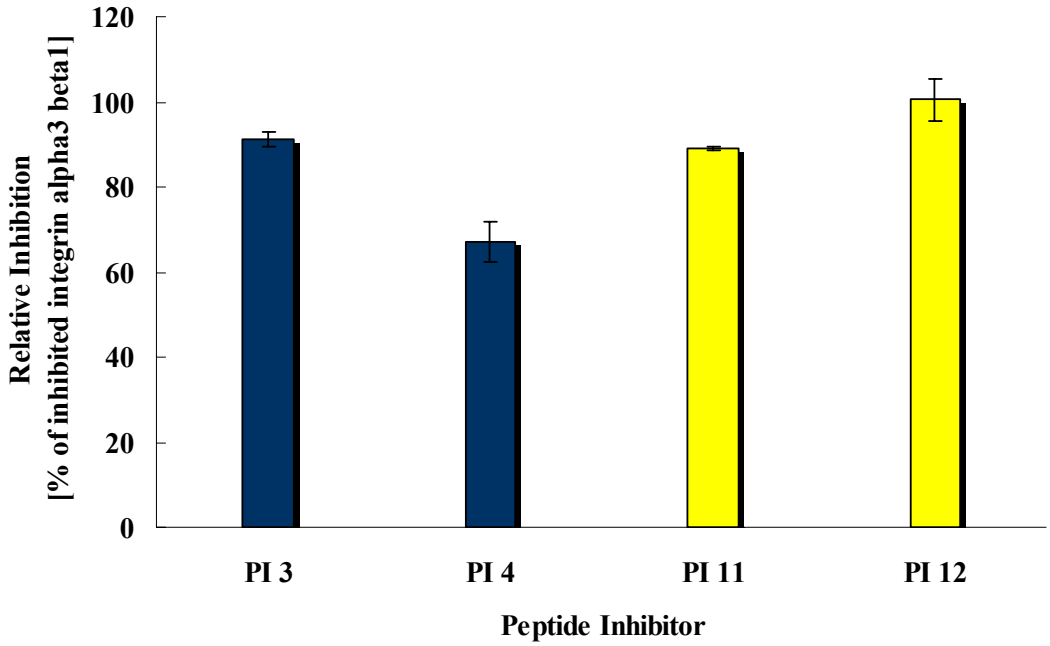


Figure 2.24 Inhibitory Results of the Peptides with Mutations at Glycine in the Reference Peptides

The substitution of glycine in the reference peptide **PI 4** by its *N*-methylated derivative sarcosine brought forward the optimized inhibitory capacity, as is shown by *Figure 2.24*. Peptide **PI 12,** with the sequence cyclo-(-Ser-Asp-Met-Ser-D-Lys-**Sar**-), exhibited markedly improvement regarding the inhibitory capacity to integrin $\alpha_3\beta_1$. *N*-methylated amino acids are known to preferentially occupy the *i*+2 position in the β-turns.[126] It is assumed that the conformational inclination of sarcosine and D-amino acid exerted cooperative influences to lock the aspartate in the favorable type of the complementary β-turn, while sarcosine and L-configuration amino acid at their positions might result in counteractive secondary structure inducing effects, as **PI 3** cyclo-(-Ser-Asp-Met-Ser-Lys-Gly) and **PI 11** cyclo-(-Ser-Asp-Met-Ser-Lys-**Sar**-) displayed similar inhibitory capacities possibly because the positive influences produced by sarcosine could be counteracted by the existence of L-type of lysine. Anyway, this experiment implied that the introduction of an *N*-alkylated amino acid at the position of glycine could improve the inhibitory capacities of the referred peptides.

2.5.3.3.6 Comparison of Inhibitory Capacities between Linear and Cyclic Peptides

One of the major advanges of cyclic peptides over their linear precursors lies in the constrained conformations with the cyclization of the linear peptides, which leads to a decreased entropy loss upon the binding of peptide to their receptor and an improved inhibitory capacity as an antagonist. In this project, the integrin recognition sequence -Ser-Asp-Met-Ser- was fused into a relatively rigid cyclic hexapeptide molecule, with the aim to lock it in a favourable conformation and to obtain a better binding affinity to integrin $\alpha_3\beta_1$. The comparison of the inhibitory capacities between the linear and cyclic peptides was performed to validate this theory. Furthermore, the recognition sequence H-Ser-Asp-Met-Ser-OH was extracted and examined in order to lower the molecular weight with the aim to achieve an improved binding affinity.

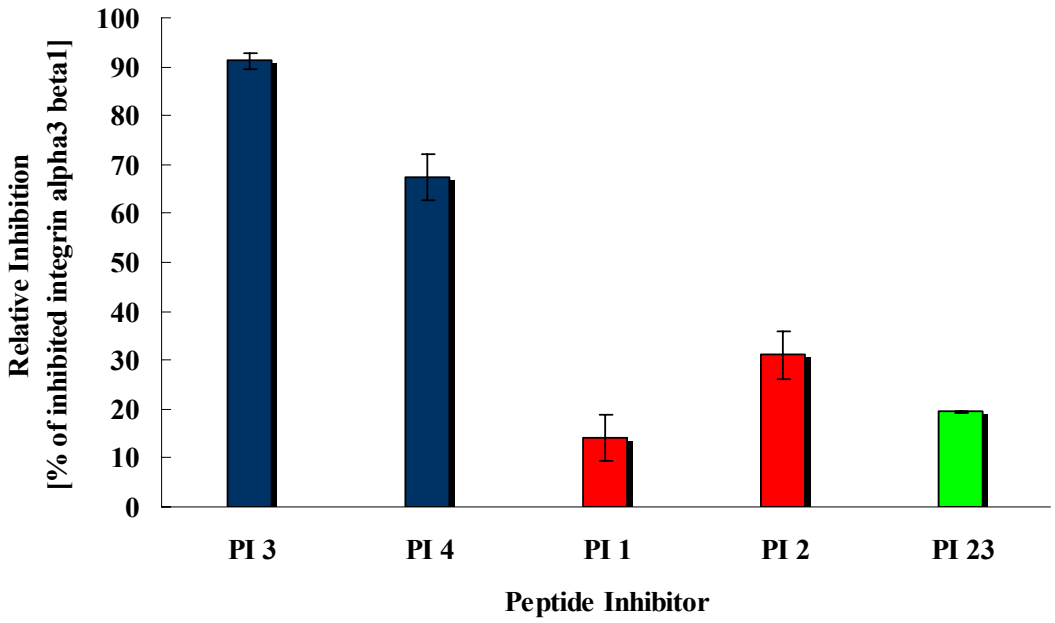


Figure 2.25 Inhibitory Results of the Cyclic and Linear Peptides

As is shown by *Figure 2.25*, the linear peptides **PI 1** H-Ser-Asp-Met-Ser-D-Lys-Gly-OH and **PI 2** H-Ser-Asp-Met-Ser-Lys-Gly-OH obviously displayed inferior inhibitory capacities compared with their cyclic counterparts **PI 3** and **PI 4**, which fully validated the rationality of the cyclization strategy. Peptide **PI 23**, with the

recognition sequence H-Ser-Asp-Met-Ser-OH, also showed evidently weakend inhibitory affinity. These results indicated that the recognition sequences should be locked in a constrained template such as cyclic peptides in order to maximize their functions as competitive ligands.

2.5.3.3.7 Repeating Recognition Sequence in the Cyclic Peptide

Since the recognition sequence -Ser-Asp-Met-Ser- begins and ends with the same amino acid residue Ser, it could be rational to design a palindromic sequence which contains two such sequences in one single molecule. It is not realistic to expect that one such molecule could bind to two receptors. However, this type of inhibitor might possess the advantage to reduce the entropy loss upon the adhesion to its receptor, since for this kind of inhibitor it is perhaps not as necessary as for the normal peptide to adjust its conformation to achieve the optimal interaction with the receptor.

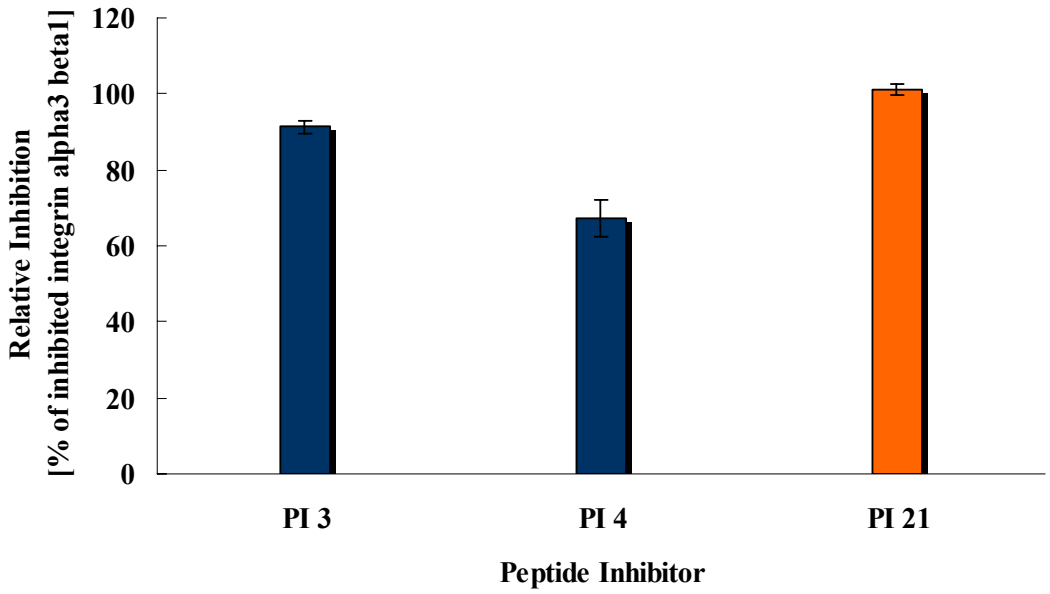


Figure 2.26 Inhibitory Results of the Cyclic Peptide with Palindromic Sequence

As is shown in *Figure 2.26*, the cyclic hexapeptide with palindromic sequence cyclo-(-Ser-Asp-Met-Ser-Asp-Met-Ser-) **PI 21** displayed a slightly improved inhibitory capacity relative to **PI 3**. It is currently not quite clear whether this improvement resides in the aforementioned entropy advantages. At least, it brought

forward a new idea and strategy to optimize the binding affinity for those antagonists that possess a recognition sequence with the same start and end residues.

2.5.3.4 Results and Discussions of Inhibitory Capacities of Peptides with Spacers

PI 24-PI 35 (Table 2.6), which bear spacers of oligo β-alanine or $(Pro\text{-}Pro\text{-}Ala)_n$ with remote synergic Asp or Arg binding domains at their flanks, as well as reference peptides **PI 3** and **PI 4** (Table 2.3) were applied to ELISA. Their capacities of inhibition of the interaction between laminin-332 and integrin $\alpha_3\beta_1$ were measured under the same conditions as those for **PI 1-23**. The results are summarized in *Figure 2.27*.

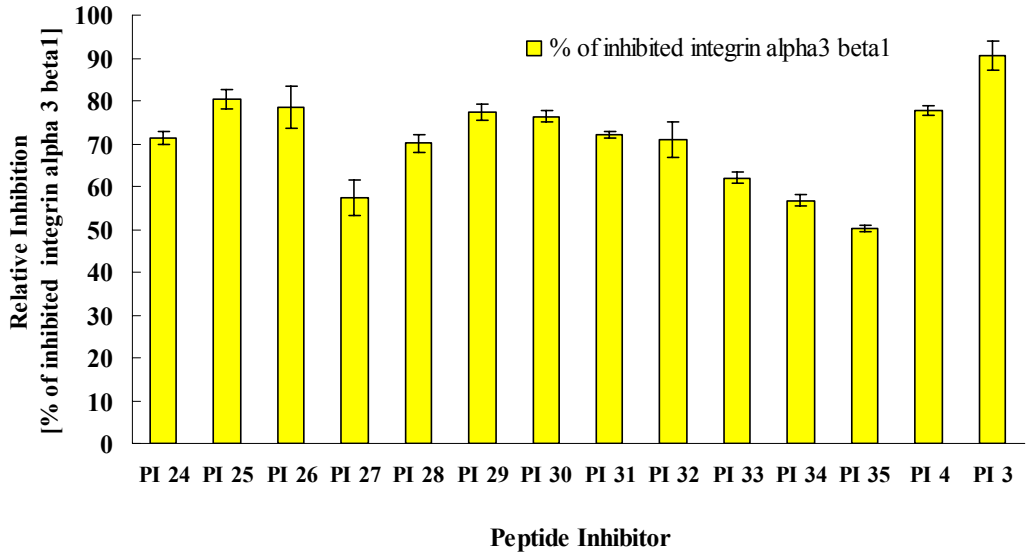


Figure 2.27 Inhibition of Laminin-332 Binding to Integrin $\alpha_3\beta_1$ by Synthetic Peptides **PI 24-35** and **PI 3-4**.

Interestingly, the peptide inhibitors **PI 24-35** with linkers bearing the synergic aspartate and arginine at the flank display weaker inhibitory properties relative to those without linkers. That is the case for both PPA-peptides and β-Ala-peptides. On average, the former peptide derivatives show higher binding affinities than the latter.

The series of linkers was designed by computer-assisted molecular modelling. The lengths of poly (Pro) linkers were calculated on the basis of their theoretical Poly Pro conformation. There should be two main conformations distributed as results of *cis*⇆ *trans* (*ω* torsion angle) (see *Figure 2.28*) and *cis*' ⇆ *trans*' (*ψ* torsion angle) equilibria, which bestows the poly (Pro) linker certain degrees of flexibility.

Amide bonds are usually found in the *trans* conformation (ω=180°) in linear peptides, whereas *cis* amide bonds (ω=0°) are observed in constrained situations such as those occurring in cyclic peptides, particularly if they are characterized by small size. A *trans* amide bond in a secondary amide is energetically more stable than a *cis* bonding by ≈ 2 kcal/mol,[141] which explains the overwhelming occurrence in linear peptides of the *trans* bonding in the crystal state and its large preponderance in solution. However, the energy difference between the two conformations markedly decreases in tertiary amides. Thus, not surprisingly, the *cis* bondings reported in the literature for linear peptides in almost all cases involve a tertiary amide in a Xxx-Yyy- sequence where Yyy is a Pro or an *N*-alkylated (Sar, MeAla, peptoid unit, ect.) residue.

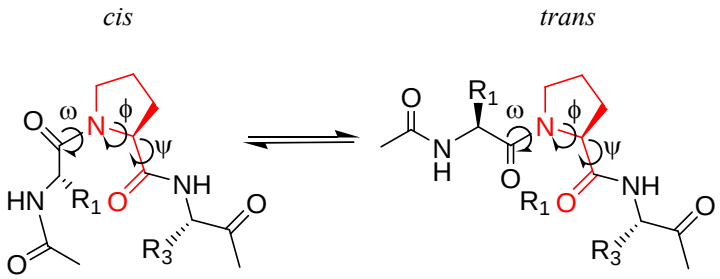


Figure 2.28 Isomerization of Proline

Pro-Pro bonds generally adopt the *trans* conformation but, in some instances, particularly when the peptide is short or the sequence is syndiotactic (L-D or D-L), *cis* bonding does occur.[142] Interestingly, the homopolymer poly (Pro) is dimorphic in that, under appropriate experimental conditions, the "all *trans*" peptide bond conformation (type II) may exhibit a transition to the "all *cis*" peptide bond conformation (type I).[141] Both helices are *left*-handed (for L-residues) and show quite similar *φ,ψ* values. The classical poly (Pro) type II is a ternary helix, which, with

appropriated side-chain replacements (e.g. an –OH function at position 4 of the ring), may be endowed with an amphiphilic character. The transition from poly (Pro) type I to type II helix implies a remarkable increase in the long dimension of the 3D-structure.

As for the φ,ψ torsion angles in Pro-based peptides, the former is almost invariant (-70±20°) owing to the restrictions imposed by the five-membered pyrrolidine ring.[141] The ψ values accessible to Pro residues either correspond to the *semi*-extended region mentioned above (*trans* conformation) or the the 3_{10}-/α- helical region (cis conformation). This latter 3D-structure, usually stabilized by 1←4 (or 1←5) intramolecular C=O···H–N hydrogen bonds, cannot be formed by Pro residue only. Pro residues, in fact, can exclusively occur at the first two (three) positions of a 3_{10}- (α-) helix, respectively, because any following Pro residue would act as a helix breaker in that it lacks the H-bonding donor (NH) functionality. The presence of a single Pro in those helices normally induces a bend in the structure. In the large majority of published examples, the *semi*-extended conformation is that preferentially adopted by Pro, indicating that this residue has an intrinsic propensity to be in the poly (Pro) structures (and at position *i*+1 of a type-II β-turn as well). This finding is especially verified in the longest homo-poly (Pro) and is related to unfavourable to unfavourable steric interactions originating between the δ-carbon of a Pro_i residue and the β-carbon of a Pro_{i-1} residue if both are folded in an α-/3_{10}-helical conformation. As for the role of the Pro pyrrolidine ring, no rationale seems to correlate the type of puckering that each residue assumes with its backbone conformation. In other words, ring conformations seem to have little control over φ,ψ. However, the flexibility of the pyrrolidine ring is expected to expand the rang of φ,ψ values available to the Pro residue, leading to a minimization of unfavourable short-range interactions.

In summary, it is confirmed that short poly (Pro) spacers and templates cannot be acritically viewed as "rigid rods", as usually considered, but rather that the *cis* ⇆ *trans* (ω torsion angle) and *cis*' ⇆ *trans*' (ψ torsion angle) equilibria, typical of the Pro

reidue. Interestingly, recent calculations have shown that even long poly (Pro) peptides may be quite flexible with a defined tendency to fold in a "worm-like" chain characterized by multiple bends.[143] It is then understandable that the molecular modelling concerning poly (Pro) had better take the *cis* ⇆ *trans* as well as *cis*' ⇆ *trans*' equilibrium into consideration. However, this exactitude would be extremely challenging for the molecular design.

Rigid molecular platforms provide well-defined distances and orientations between appropriate probes or functional groups, thus greatly facilitating a reliable and correct interpretation of experimental results based on the 3D-strucuture dependence of organic and physical chemistry processes. Peptide-based systems of different lengths present a remarkable advantage over other types of derivatized skeletons because they are easily synthetically assembled.

Oligopeptide 3D-structures based on C^{α}-trisubstituted (e.g. protein) α-amino acids of variable length have already been used as templates. However, particularly in the case of relatively short peptides, only a partially restricted mobility has been achieved. The most commonly used oligopeptides series in this context are poly (Pro), followed by $(Gly)_n$, $(Ala)_n$, and *γ*-substituted $(Glu)_n$. $(Gly)_n$ polymers are known to fold either in the ternary helix poly (Gly) I or in the antiparallel *β*-sheet conformation poly (Gly) II, while $(Ala)_n$ and *γ*-substituted $(Glu)_n$ oligomers may adopt either the *α*-helical or the *β*-sheet conformation. In addition, statistically unordered forms occur largely in the complex conformational equilibria of this type of short oligopeptides, with their population being inversely proportional to the peptide main-chain length.

Ideally, the linker peptides should adopt the favorable conformation to assist the probes on their flank to arrive at the relative functional domains. A compromise to this ideal situation is a linker with a certain flexibility, which could assure the envisaged directional interactions of the probes with their functional target, instead of a rigid linker led to a frozen but unfavorable direction. However, this compromise is achieved at the cost of entropy lost, which is adverse to tight affinity. As it is clear

that none of the peptide series discussed above can produce truly rigid backbone templates, in the past few years many groups concentrated their efforts on oligopeptides rich in the structurally severely restricted C^{α}-tetrasubstituted α-amino acids.[144] An appropriate choice of specific members of this latter class of amino acids will allow one to tailor well-determined peptides 3D-structures endowed with exactly specified, intramolecular C^{α}---C^{α} distances.

However, in the search for rigid templates/spacers, additional problems may arise from rotations about amino acid side-chain single bonds.[145] Indeed, most investigations have exploited as probes or reactive pendants: (i) flexible, unmodified, protein amino acids (His, Trp, Tyr, Cys, Met, Glu, Asp, Ser) or (ii) flexible, appropriately side-chain modified, protein amino acids (Cys, Ser, Lys, Glu, Asp, Pro, Phe). Although a limited side-chain flexibility might in general be tolerated, or may even be beneficial, that arising from protein amino acids is definitely too large, making any conclusion quite approximate. In other words, by utilizing this type of side chain, any investigation inevitably suffers a range of uncertainty even larger than that saved from the restrictions imposed by rigidification of the backbone.

As is already explained, there should be two main conformations distributed as results of *cis* ⇆ *trans* (ω torsion angle) and *cis*' ⇆ *trans*' (ψ torsion angle) equilibria. However, it is hard to estimate the population of the conformation, which makes the molecular modelling deviated from the actual conformation of the poly (Pro) linkers. The equilibria of the two conformation makes the corresponding inhibitors not as effective as the counterparts with poly (β-Ala) peptide as spacer, since the conformation of poly (β-Ala) is regarded as random coil in solution and the calculated linker length of this type serves as a statistically average value. The molecular modelling decided lengths of poly (β-Ala) linker bears therefore relatively less deviations from those of actual situations, despite the fact that poly β-Ala linkers are quite flexible and their orientations are accordingly less accurate and less controllable.

The backbone of poly (Pro-Pro-Ala) spacer is not rigid enough to maintain its length in solution; the flank synergic Arg/Asp residue might be unable to bind to the corresponding target on invasin with the accuracy assumed by the computer-assisted molecular modelling on the basis of the poly (Pro) conformation hypothesis.

A fundamental limitation for the application of geometric fitting procedure is that the complexation free engergies are the sum of enthalpic and entropic contribution.

$$\Delta G = - RT \ln K_i = \Delta H - T\Delta S$$

Positive cooperativity between different interactions in a complex will usually lead to tighter association at the expense of motional freedom and thus of entropy.[146]

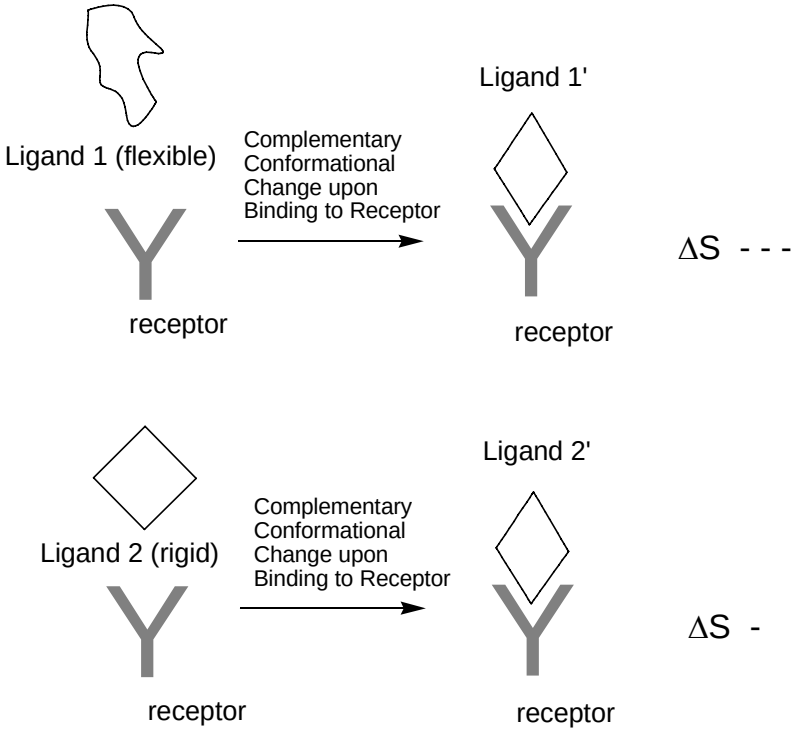


Figure 2.29 Entropy effect in Lingand Binding to Receptor

Entropy and enthalpy can often interplay and make compensation for each other upon the binding of ligand to receptor. As is already explained previously, some complexation processes are entropy-driven, while some are enthalpy-driven. The different distance dependence of non-covalent interaction mechanism necessarily modifies the size-matching requirements. For example, cyclodextrin complexes with tightly fitting and highly polarizable guest molecules form mainly by enthalpy gain,

whereas those with loose fit and aliphatic substrates in their cavity show some hydrophobic entropic contributions.

As for the complexation of the aforementioned peptide derivatives with integrin $\alpha_3\beta_1$, the incorporation of the spacer might generate a dual effect. Indeed, it promotes the synergic motifs of the inhibitor to reach their functional zone on the target protein. On the other hand, the introduction of the spacer into the peptide inhibitors enlarges their degrees of freedom，which might generate a relative bigger entropic loss when binding with the corresponding protein target explained by entropy effect (see *Figure 2.29*). Positive cooperation between different interactions in a complex will usually lead to tighter association at the expense of motional freedom and thus of entropy. For the binding with the above dual mechanism, it seems that the loss of the entropy cannot be compensated by the gain of enthalpy upon the binding, and such binding would therefore be adverse. It explains the phenomenon that peptide **PI 4** and **PI 3**, without any spacers, hold better inhibitory effects than the corresponding peptide complexes bearing poly β-Ala and poly (Pro-Pro-Ala) as spacers, which might be to some extent too flexible and are therefore more adverse upon the binding from the sense of entropy loss.

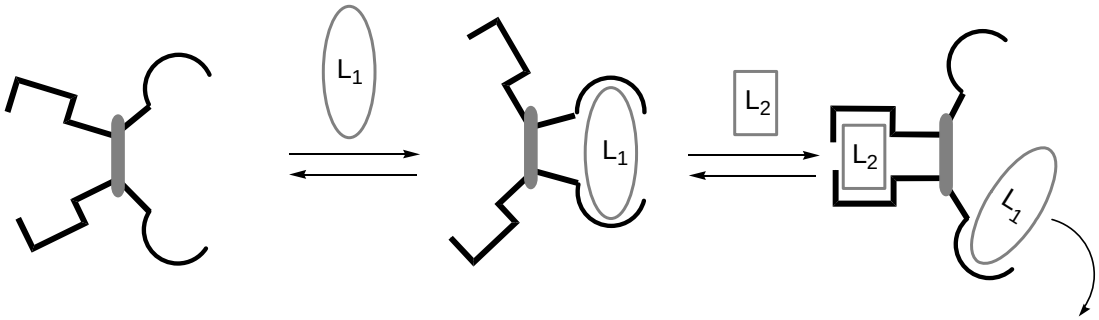


Figure 2.30 Negative Cooperativity in Synthetic Allosteric Model

Another possibility is that despite the fact that both -Ser-Asp-Met-Ser- and Asp811/Arg883 in native invasin are resided in different domains, the interface between these domains is relatively flexible and is regarded as allosteric site. Conformational changes upon the binding of invasin to its receptor could take place and the relative orientations between these two binding domains might be altered.

While in the design of artificial peptide inhibitors it would be extremely difficult to take this into account, since the crystal structure of integrin $\alpha_3\beta_1$ and that of its complex with invasin are up to now unavailable. Negative cooperativity can be realized if occupation of one binding site leads to the release of a substrate at the second binding site, which does not fit anymore after the induced conformational change (see *Figure 2.30*). However, this hypothesis was unable to be validated in this project.

In summary, the synthesized peptide conjugates applied in the biological test did not exhibite optimal inhibitory capacities to inhibite the binding of laminin332 to integrin $\alpha_3\beta_1$. However, these peptide conjugates did possess capability of inhibiting the association between integrin $\alpha_3\beta_1$ and laminin-332 to some extent, although not optimally.

2.5.3.5 Analysis of Biological Assay Error

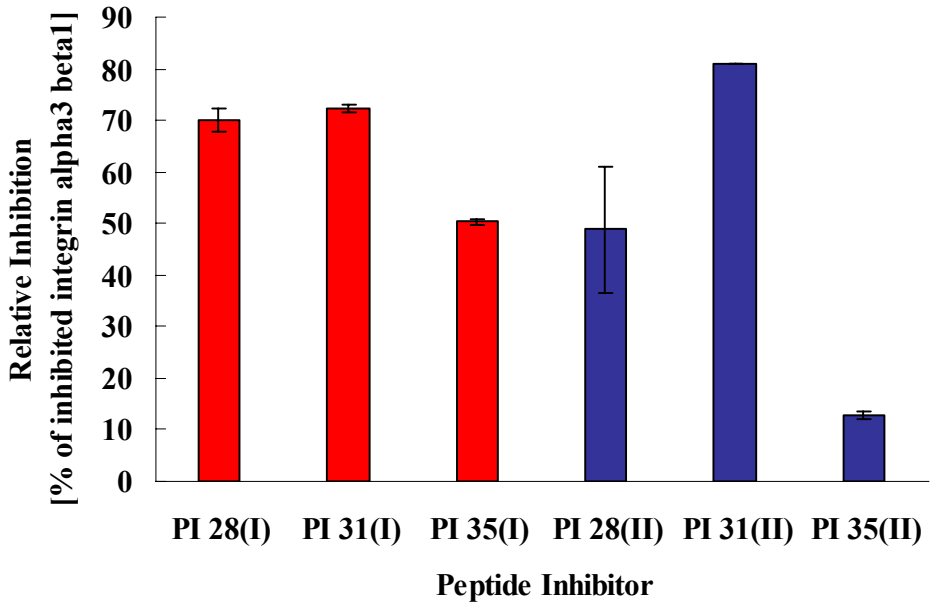


Figure 2.31 Inhibitory Capacities of Same Peptides In Two Batch Assay

It is necessary to analyze the factors of error in the biological assay. For the first batch of analysis (shown as the red bars in *Figure 2.31*), the peptides were weighed with a

balance with an accuration of 0.1 mg. As the amounts of peptides subjected to the biological assay were mostly in a couple of milligram range, the herein generated errors could be relatively large. The second batch of biological assay were associated with a balance accurated at 0.01 mg. From the comparison of two batch analysis, it is discernable that there are some discrepancies between the same peptide inhibitors in different assays as their inhibitory capacities are concerned. This might be due to the systematic weighing errors as was analyzed above. Other factors such as the different extent of integrin activations could also be attributed to the inhibitory discrepancies between the different assays. However, the trends of inhibitory capacities between the two batch analyses are comparative as is shown by the Figure. The relative inhibitions between the diverse peptides are reliable in the biological assay.

2.6 Experiments

2.6.1 General Methods

Flash Chromatography

Silica gel 60 (40-63 μm), Merck.

Thin Layer Chromatography

TLC Plate (Silica gel 60 F254) Merck; Detection: UV (254 nm, MinUVIS) Desaga.

Solvent

Solvents were technical grade and freshly distilled except otherwise mentioned; Acetonitrile for RP-HPLC was from Merck with LiChrosolv® grade; Water for RP-HPLC was processed by MiliQ water purification apparatus (resistance = 18.2 MΩ·cm^{-1}). Dichloromethane was distilled first over K_2CO_3, and then redistilled over CaH_2. DMF was distilled over ninhydrin under vacuum. THF and Et_2O were first distilled over KOH, and then redistilled over Na/benzophenone. Toluene was distilled over Na.

NMR Spectroscopy

DRX-500 (Bruker), Measure Frequency: ^{1}H-NMR: 500.13 MHz, room temperature

MALDI-ToF Mass Spectrometry

The Voyager-DE BioSpectrometry Workstaton (Applied Biosystems) MALDI-ToF Mass Spectrometer with a 1.20 m flight tube was operated with a positive mode. The ionized compounds were desorbed by LSI-Nitrogen-Laser (λ = 337 nm, Pulse Range: 3 ns, Repeat Frequency: 3 Hz). Accelerating voltage: 20 Kv, grid voltage: 94%, guide

wire: 0.05%, extraction delay time: 100 ns. 2,5-Dihydroxy benzoic acid (DHB) was used as matrix for MALDI-Tof MS measurements.

Digital Polarimeter

DIP-360 (Jasco), cell length: 10 cm, wavelength: 589 nm, room temperature

Microwave Peptide Synthesizer

CEM Microwave Peptide Synthesizer.

Fmoc Deprotection (twice): 1st time: 75 °C, 30 s, 35 w;

2nd time: 75 °C, 180 s, 35 w

Coupling: 75 °C, 300 s, 20 w.

CD Spectroscopy

CD spectroscopy was done on a J-810 Spectrometer (Jasco) in the range of 185-300 nm. Parameters: band width: 1 nm; response: 1 s; sensitivity: standard datapitch: 0.2 nm; scanning speed: 50 nm/min, optical path: 0.2 mm, accumulation: 3.

The measured CD spectra were smoothed through Means-Movement method, and converted to molar ellipticity of mean amino acid residue ($[\theta]_{MRW}$).

Analytical RP-HPLC

Pump: Thermo Separation Products P4000

Detector: Thermo Separation Products UV6000 LP, λ = 220 nm/ 254 nm

Column: Phenomenex® Jupiter Proteo C12 (4 μm, 250×4.6 mm)

Eluent A: CH_3CN/H_2O/TFA, 95:5:0.1

Eluent B: H_2O/CH_3CN/TFA, 95:5:0.1

Method I

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	1.0
5	5	95	1.0
25	50	50	1.0
30	0	100	1.0

Method II

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	1.0
3	0	100	1.0
40	100	0	1.0
50	0	100	1.0

Preparative RP-HPLC

Apparatus A

Pump: Thermo Separation Products P4000

Detector: Thermo Separation Products UV1000 LP, λ = 220 nm

Column: Phenomenex® Jupiter Proteo C12 (90 Å, 10 µm, 250×21.2 mm)

Eluent A: $CH_3CN/H_2O/TFA$ 95:5:0.1

Eluent B: $H_2O/CH_3CN/TFA$ 95:5:0.1

Apparatus B

Pump: Merck Hitachi L-7150

Detector: Merck Hitachi UV-VIS L-7420, λ = 220 nm

Cloumn: Phenomenex® Jupiter Proteo C18 (300 Å, 10 µm, 250×21.2 mm)

Eluent A: $CH_3CN/H_2O/TFA$ 95:5:0.1

Eluent B: $H_2O/CH_3CN/TFA$ 95:5:0.1

Method 1

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
5	0	100	13.0
50	10	90	13.0
55	10	90	13.0
60	0	100	13.0
65	0	100	13.0

Method 2

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
5	0	100	13.0
50	20	80	13.0
55	20	80	13.0
60	0	100	13.0
65	0	100	13.0

Method 3

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
5	0	100	13.0
50	1	99	13.0
55	1	99	13.0
60	0	100	13.0
65	0	100	13.0

Method 4

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
5	0	100	13.0
100	1	99	13.0
110	1	99	13.0
120	0	100	13.0

Method 5

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
50	5	95	13.0
55	5	95	13.0
65	0	100	13.0
70	0	100	13.0

Method 6

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate
0	0	100	13.0
50	10	90	13.0
55	10	90	13.0
65	0	100	13.0
70	0	100	13.0

Method 7

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate
0	0	100	10.0
10	0	100	10.0
20	0	100	10.0
30	0	100	10.0
40	0	100	10.0
50	0	100	10.0

Method 8

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate
0	0	100	10.0
60	20	80	10.0
65	0	100	10.0
75	0	100	10.0

Method 9

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
30	30	70	13.0
35	30	70	13.0
40	0	100	13.0

Method 10

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	70	30	10.0
30	100	0	10.0
50	100	0	10.0
60	8	92	10.0
65	8	92	10.0

Method 11

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
15	100	0	13.0
20	100	0	13.0
25	0	100	13.0
30	0	100	13.0

Method 12

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
50	50	50	13.0
55	0	100	13.0
65	0	100	13.0
70	0	100	13.0

Method 13

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
30	40	60	13.0
35	40	60	13.0
40	0	100	13.0

Method 14

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	10	90	10.0
5	10	90	10.0
40	50	50	10.0
45	50	50	10.0
50	0	100	10.0
55	0	100	10.0

Method 15

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	10.0
5	0	100	10.0
40	30	70	10.0
45	30	70	10.0
50	0	100	10.0
55	0	100	10.0

Method 16

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	30	70	13.0
40	100	0	13.0
45	100	0	13.0
50	0	100	13.0

Method 17

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	8	92	13.0
5	8	92	13.0
35	50	50	13.0
40	50	50	13.0
50	0	100	13.0

Method 18

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	20	80	13.0
40	35	65	13.0
45	35	100	13.0
50	0	100	13.0
55	0	100	13.0

Method 19

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	25	75	13.0
40	40	60	13.0
45	40	60	13.0
50	0	100	13.0
55	0	100	13.0

Method 20

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	8	92	13.0
5	8	92	13.0
35	100	0	13.0
40	100	0	13.0
50	0	100	13.0

Method 21

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	70	30	13.0
40	85	15	13.0
45	100	0	13.0
50	0	100	13.0

Method 22

Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
50	20	80	13.0
55	20	80	13.0
65	0	100	13.0
70	0	100	13.0

Method 23

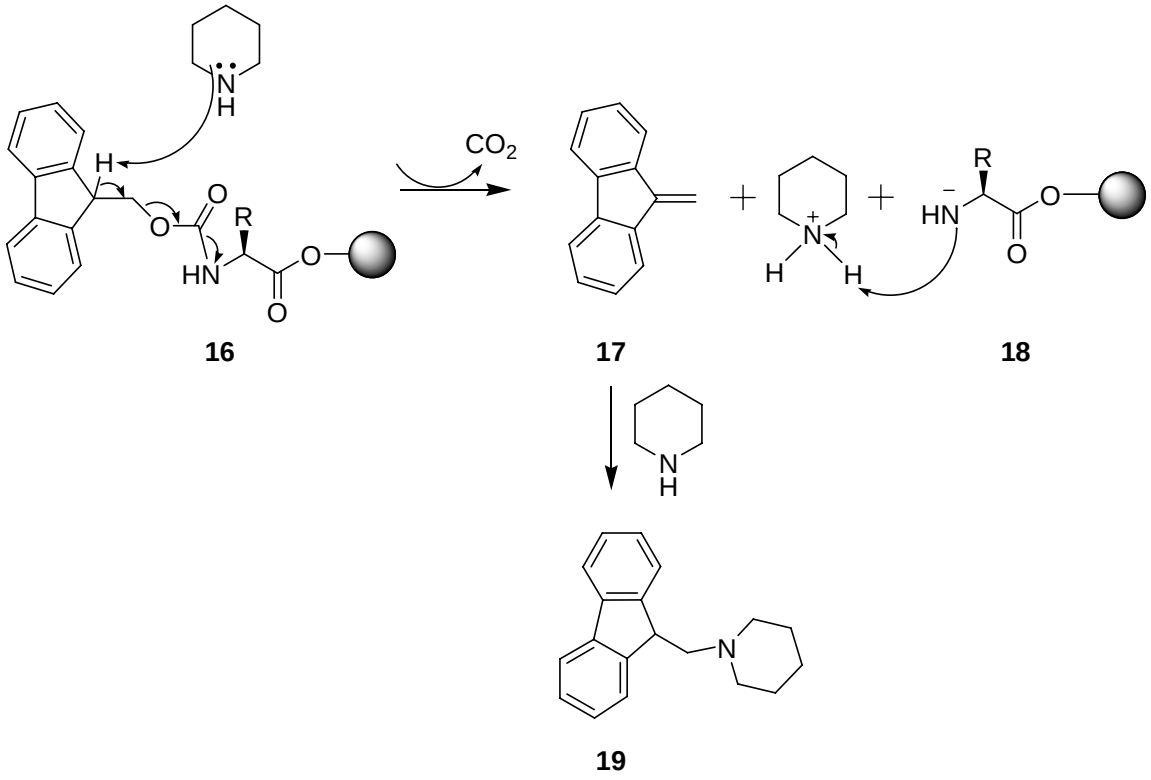
Time (min)	Eluent A (%)	Eluent B (%)	Flow Rate (ml/min)
0	0	100	13.0
30	80	20	13.0
35	80	20	13.0
40	0	100	13.0
45	0	100	13.0

2.6.2 Synthetic Experiments

2.6.2.1 General Methodology of SPPS

2.6.2.1.1 Loading of *o*-chlorotritylchloride Resin and Quantification

The *o*-chlorotritylchloride resin applied in the SPPS was first swollen in dry CH_2Cl_2. 3 equiv of Fmoc-amino acid and 3 equiv DIPEA were dissolved in dry CH_2Cl_2 (ca. 10 ml DCM/1 g amino acid). After shaking the mixture at room temperature for 5 min, additional 4.5 equiv DIPEA were added. The resin was shaken gently at room temperature for further 60 min before methanol (ca. 0.8 ml methanol/1 g resin) was introduced to the reaction system in order to trap residual unreacted Trt(2-Cl) chloride functionalities on the resin. The endcapping process took normally 15 min at room temperature. The Fmoc-amino acid loaded resin was then filtered and subsequently washed with 4×DCM, 4×DMF, 4×DCM, 4×CH_3OH in turn. The resin was dried under vacuum overnight until its weight remained constant.



Scheme 2.15 Quantification of the loading rate of Fmoc-amino acid on resin

Ca. 1 mg dried Fmoc-amino acid loaded resin **16** was treated with 3 ml 20% piperidine /DMF. The Fmoc-protecting group is cleaved under this basic condition and the dibenzofulvene intemediate **17** is captured by piperidine to generate the piperidine-dibenzofulvene-adduct **19**, which was quantified by UV absorption at 290 nm (*Scheme 2.15*). The resin loading with Fmoc-amino acid was thus calculated through the following empirical equation:

Resin Loading (mmol/g) = Abs/(weight of resin (mg) × 1.65) (Abs: absorption)

As the Fmoc group on the solid support could be possibly cleaved gradually at ambient temperature, the dibenzofulvene **17** is able to initiate a series of side reactions in the following peptide synthetic process. The Fmoc-amino acid loaded resin **16** is therefore supposed to be treated with Fmoc deprotection before storage. Fmoc-deprotecting solution pipieridine/DBU/DMF (2:2:96) was added to **16** at room temperature in order to remove the Fmoc protecting groups. The obtained N^{α}-free amino acid loaded resin **18** was then washed by 4×DCM, 4×DMF, 4×DCM, 4×CH_3OH in turn. The resin was dried under vacuum overnight until its weight remained constant.

2.6.2.1.2 On-resin Peptide Chain Elongation by Fmoc/tBu Chemistry

Except otherwise mentioned, 3 equiv Fmoc-amino acid (with respect to the N^{α}-free amino acid loaded resin **18**) were dissolved in a minimum volume of DMF (ca. 1 mmol amino acid/ml). 3 equiv uronium-type coupling reagent, e.g. TBTU, TCTU, or HATU, also dissolved in minimal volume of DMF separately, were added and then 6 equiv DIPEA were added to the reaction mixture. The duration of this preactivation process should be kept less than 5 min to prevent the formation of oxazolone intermediate that causes substantial racemization of the amino acid bearing the chiral α-carbon. The solution of activated Fmoc-amino acid was then transferred to the N^{α}-free peptidyl-resin under argon. The coupling reaction took approximately 45 minutes at room temperature, depending on the sequence. The coupling result could be controlled by MS such as MALDI-ToF-MS. Alternatively, the Kaiser Test, which

detects the primary amine functionalities, can be applied to probe if there is still unreacted amine group on the peptidyl resin. If the coupling did not go to completion, a second coupling could be carried out to enhance the efficiency. Otherwise, the uncoupled truncated peptidyl chain could be capped using Ac_2O in the presence of tertiary amine such as pyridine or triethylamine, which do not cleave the N^{α}-Fmoc temporary protecting group. Upon completion of the coupling, the peptidyl resin was washed with DMF at least four times to remove the remaining reagents attached on the resin.

The N^{α}-Fmoc temporary protecting group on the peptidyl resin was treated with 20% piperidine in DMF twice at room temperature for 5 minutes and 15 minutes, respectively. The N^{α}-free peptidyl-resin was subsequently washed with DMF at least four times to remove the remaining reagents attached to the resin.

2.6.2.1.3 Acetylation of the *N*-terminus of Peptide

3 equiv Ac_2O (in terms of N^{α}-free peptidyl resin) and 3 equiv Et_3N were dissolved in DMF, which was subsequently transferred to the N^{α}-free peptidyl resin. The mixture was agitated by argon bubbling at room temperature for approximately 30 min. The *N*-acetylation was analyzed by MALDI-ToF-MS and Kaiser Test. After the completion of the acetylation was the *N*-acetyl peptidyl resin washed by CH_2Cl_2 (4×), DMF (4×), CH_2Cl_2 (4×) in turn.

2.6.2.1.4 Cleavage of the Fully Protected Peptide from the Resin

The N^{α}-unprotected, side-chain protected peptidyl resin was washed consecutively with DMF and CH_2Cl_2. 5 ml 1% TFA in CH_2Cl_2 were then added. The system was agitated by bubbling argon through it at room temperature for approximately 5 min. The resin was then filtered and the filtrate was neutralized by pyridine and collected. Another portion of 5 ml of fresh 1% TFA in CH_2Cl_2 was added to the resin, this cycle was repeated 10 times and the the resin was washed with CH_2Cl_2 until its color converted from red to yellow. The solvent of the filtrate was removed under reduced

pressure and the yielded raw fully protected peptide was lyophilized for the next step of reaction or purification.

2.6.2.2 Microwave Synthesis of Peptide with SPPS Manner

For a 0.25 mmol peptide synthesis scale, 10 ml of 20% piperidine in DMF were added to the peptidyl resin for Fmoc removal. The reaction was performed at 75 °C for 30 seconds with 35 watt microwave. After draining and washing, further 10 ml of 20% piperidine in DMF were added, reacted with peptidyl resin at 75 °C for 180 seconds in 35 watt microwave. At the coupling reaction stage 5 ml 0.2 M amino acid DMF solution, 2 ml 0.5 M TBTU solution in DMF, and 1 ml 2 M DIPEA solution in NMP were added to the peptidyl resin, reacted at 75 °C for 300 seconds in 20 watt microwave. The cleavage of complete fully protected peptide from the resin, side-chain deprotection and consequent purification of target peptide was employed as described for manual peptide synthesis (see also chapter **2.6.2.1**).

2.6.2.3 Cyclization of N^{α}-free Side-chain Protected Linear Peptides

DMF was added to a certain amount of N^{α}-unprotected, side-chain protected peptide. The peptide solution was transferred into a syringe attached to a needle. 3 equiv uronium coupling reagent e.g. HATU were dissolved in the same amount of DMF before being transferred into another syringe of the same type. The two syringes were fixed to the double-channel-syringe-pump. The peptide solution and coupling reagent solution were added dropwise over 24 hours into a three-neck-flask containing 6 equiv DIPEA dissolved in DMF. Approximately 1 mg HATU was also added to the flask as to prevent that the peptide solution was dropped prior to HATU, which abstains the real time excess of peptide. As the oxygen can sometimes exert negative influences on the coupling reaction, and is furthermore problematic for some kind of amino acids, such as methionine or cysteine, the cyclization reaction system is preferably protected by inert gas. After the peptide and HATU solution were completely added into the flask, the solvent was then removed under reduced pressure

and the crude peptide was lyophilized for the next step of reaction or purification.

2.6.2.4 Side Chain Deprotection

The cocktail composed of mild acid such as TFA and scavengers like phenol, H_2O, thiol, and silane, Reagent K (TFA/phenol/H_2O/thioanisol/ethane-1,2-dithiol, 82.5:5:5:5:2.5) for instance, was added to the lyophilized side-chain-protected peptide to bring about deprotection. For peptides without sulfur-containing amino acids, the deprotection cocktail consisting TFA/H_2O/triisopropylsilane (92.5:2.5:2.5) can cleave the side-chain protecting groups smoothly and quantitatively at ambient temperature. For peptides containing cysteine or methionine, which are readily oxidized under the acidolytic deprotection conditions, extra protection afforded by scavengers is necessary. Under these kinds of circumstances, Reagent K is the preferable choice as a deprotection cocktail. Except the special scavengers for the methionine- or cysteine-containing peptides, the deproctection of these peptides should be carried out under inert gas. Under cautious treatment, no oxidation at the methionine residue was observed during the deprotection process of the sulfur-containing peptides.

For short- and medium-size peptides, the duration of the deprotection normally lasts no longer than 3 hours at room temperature before the quantitative reaction is obtained. For longer peptides, or those who are readily aggregating in solutions so that deprotection is to some extent impeded, longer deprotection times or slight increase of the deprotection temperature are used to achieve quantitative cleavage. For some acid-sensitive sequences and aspartate-rich peptides, the aforementioned methods could lead to decomposition of the peptide at the sensitive sequence region and the side reaction of the aspartimide formation, respectively. Hence, the real-time analysis of the deprotection reaction with the method such as MALDI-ToF-MS is necessary.

After complete deprotection of the concerned peptides, the solvent was removed under reduced pressure. A small amount of TFA was added to the peptide residues

afterwards in order to have them fully dissolved. The peptide solution was slowly added to ice-cold diethyl ether resided in an ice-water bath. The target peptides precipitate upon addition of the solution into cold ether. The suspension was stirred for approximately 30 minutes at 0 ºC before the scavenger derivative-containing diethyl ether fraction was removed after centrifugation. The solid peptide was again added to another portion of fresh ice-cold diethyl ether and stirred for 30 minutes in the ice-water bath. Likewise, the ether solution was removed by centrifugation. The obtained peptide solids were dissolved in water or water/acetonitrile, lyophilized and purified by RP-HPLC.

2.6.2.5 On-resin Coupling of α-Aminoisobutyric Acid through Azido Acid Chloride

The N^{α}-free peptidyl resin (1 equiv) was swollen in the mixture of 10 equiv NEt_3 in dry CH_2Cl_2. The concerned azidoacidchloride (N_2=Aib-Cl, 7 equiv) was cautiously and slowly added to N^{α}-free peptidyl resin. The reaction system was agitated through argon bubbling at room temperature for approximately 30-45 min. For MS analysis of the coupling results, a small amount of resin could be sampled and cleaved by HFIP (20% in CH_2Cl_2) after washing. Provided that the coupling did not go to completion, methods such as elongation of the coupling time or, alternatively, multiple coupling could be carried out. After the completion of the reaction the resin was filtered and washed by CH_2Cl_2 (4×), DMF (4×), CH_2Cl_2 (4×) orderly.

A suspension of $Sn(SPh)_2$ (10 equiv in terms of peptidyl resin) in dry CH_2Cl_2 with thiophenol (30 equiv) and NEt_3 (50 equiv) was mixed at room temperature for 5 min, which was subsequently added to the N^{α}-azido peptidyl resin in order to yield the free *N*-terminus for the next coupling. The deprotection reaction was agitated gently by argon bubbles at room temperature for approximately 20 min. N_2 was yielded upon the deprotection. Upon the completion of reaction the peptidyl resin was filtered and washed by CH_2Cl_2 (4×), DMF (4×), CH_2Cl_2 (4×).

2.6.2.6 Selective Removal of Aloc with Palladium(0) Catalyst

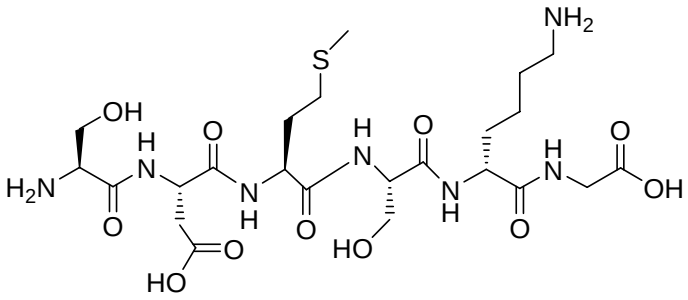
Approximately 20 μmol fully protected cyclic SDMS derivative peptide containing Lys(Aloc) residue in the sequence were dissolved in 2-3 ml dry DCM, 1 equiv $Pd(PPh_3)_4$ and 30 equiv $PhSiH_3$ were added to the peptide solution. Bubbles were formed upon the addition of palladium(0) catalyst and silicon hydride scavengers. The solution was stirred at the ambient temperature for approximately 3 hours with the protection of inert gas such as argon until the Aloc protecting group was removed quantitatively, which was monitored by MALDI-ToF-MS. Upon completion of the Aloc deprotecting reaction the solvent was removed under reduced pressure. The residue was dissolved in DMF and subjected to RP-HPLC in order to remove the palladium catalyst and purify the peptide. The obtained peptide was analyzed by MALDI-ToF-MS and analytical RP-HPLC. The portion of cyclic SDMS derivative with the free Lys residue was lyophilized and applied to the segment condensation.

2.6.2.7 Segment Condensation of Cyclic Peptide with Linker Peptides

1.1-1.5 equiv (relative to cyclic SDMS derivative peptide with unprotected Lys residue) *N*-Fmoc side-chain protected linker peptides with Asp/Arg at the flank (see also chapter **2.5.1.2.2**) were dissolved in a minimal amount of dry DMF, 2.2-3.0 equiv HATU and 4.4-6.0 equiv DIPEA (relative to cyclic SDMS derivative peptide) were added to the linker peptide solution. The solution was stirred for approximately 10-15 min at ambient temperature before being added to the DMF solution of cyclic SDMS derivative peptide with unprotected Lys residue. The combined solution was stirred at ambient temperature for approximately 30 min under argon. Further 2.2-3.0 equiv HATU and 4.4-6.0 equiv DIPEA were then added to the solution, the mixture was stirred at room temperature overnight under argon until the completion of segment condensation, as analyzed by the MALDI-ToF-MS. The solvent was then removed under the reduced pressure. The solids of peptide mixtures were dissolved in DMF and applied to RP-HPLC for purification. The residue was analyzed by MALDI-ToF-MS and analytical RP-HPLC.

2.6.2.8 Synthetic Peptide/Peptidomimetic Inhibitors

H-Ser-Asp-Met-Ser-D-Lys-Gly-OH (PI 1)



PI 1

The synthesis of fully protected peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 1** was carried out manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Deprotection of the peptide was achieved with the deprotection cocktail Reagent K.

(Except otherwise mentioned, all the yields listed in this project represent those after RP-HPLC. *vide infra*)

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 1**)

0.100g/0.11 mmol

H-Ser-Asp-Met-Ser-D-Lys-Gly-OH (**PI 1**) 0.0289g/0.0463 mmol (15.4%)

Formula: $C_{23}H_{41}N_7O_{11}S$

Molecular Mass: 623.26 g/mol

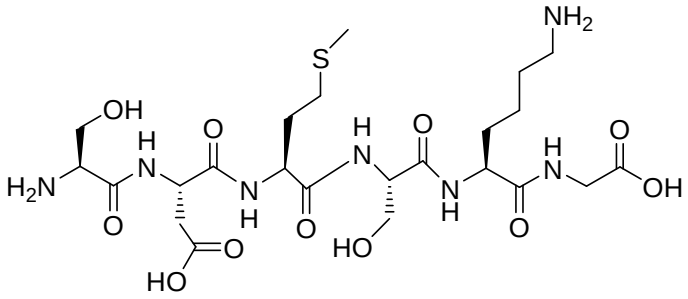
MS (MALDI-ToF): $m \cdot z^{-1}$ = 625.00 $[M+H]^+$, 646.99 $[M+Na]^+$, 662.96 $[M+K]^+$, 668.98 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{42}N_7O_{11}S]^+$ = 624.27, $[C_{23}H_{41}N_7NaO_{11}S]^+$ = 646.25, $[C_{23}H_{41}KN_7O_{11}S]^+$ = 662.22, $[C_{23}H_{40}N_7Na_2O_{11}S]^+$ = 668.23

Preparative RP-HPLC: Apparatus A, Method 1, t_R = 10.90 min

Analytical RP-HPLC: Method I, t_R = 6.63 min, 100% (220 nm)

H-Ser-Asp-Met-Ser-Lys-Gly-OH (PI 2)



PI 2

The synthesis of fully protected peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH **Prot.PI 2** was carried out manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.15 mmol), TBTU as coupling reagent. Deprotection of the peptide was achieved with the deprotection cocktail Reagent K

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH (**Prot.PI 2**)

0.098 g/0.11 mmol

H-Ser-Asp-Met-Ser-Lys-Gly-OH (**PI 2**) 0.0193 g/0.031 mmol (20.7%)

Formula: $C_{23}H_{41}N_7O_{11}S$

Molecular Mass: 623.26 g/mol

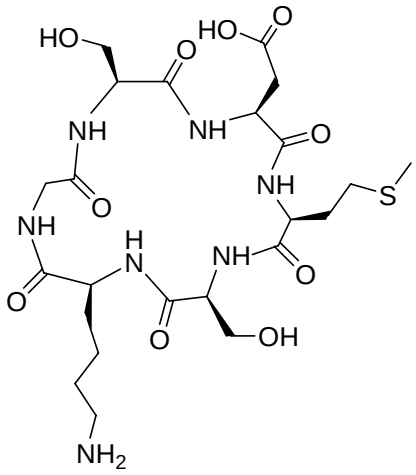
MS (MALDI-ToF): $m \cdot z^{-1}$ = 625.17 $[M+H]^+$, 647.14 $[M+Na]^+$, 663.12 $[M+K]^+$, 669.12 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{42}N_7O_{11}S]^+$ = 624.27, $[C_{23}H_{41}N_7NaO_{11}S]^+$ = 646.25, $[C_{23}H_{41}KN_7O_{11}S]^+$ = 662.22, $[C_{23}H_{40}N_7Na_2O_{11}S]^+$ = 668.23

Preparative RP-HPLC: Apparatus A, Method 1, t_R = 11.20 min

Analytical RP-HPLC: Method I, t_R = 6.35 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-Lys-Gly-) (PI 3)



PI 3

The synthesis of fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH **Prot.PI 3** was carried out manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 3** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Pro.PI 3** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH (**Prot.PI 3)**

0.1466 g/0.16 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-) (**Cyc.Pro.PI 3**)

0.068 g/0.077 mmol

cyclo-(-Ser-Asp-Met-Ser-Lys-Gly-) (**PI 3**) 0.0088 g/0.015 mmol (5.0%)

Formula: $C_{23}H_{39}N_7O_{10}S$

Molecular Mass: 605.25 g/mol

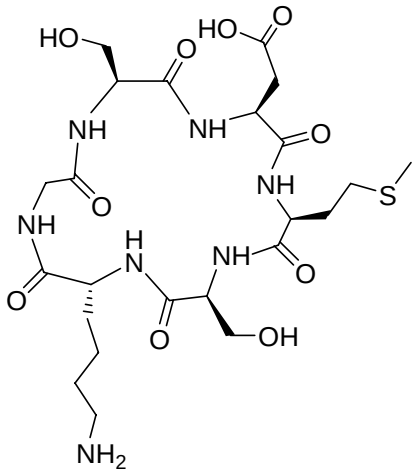
MS (MALDI-ToF): $m \cdot z^{-1}$ = 607.53 $[M+H]^+$, 629.52 $[M+Na]^+$, 645.53 $[M+K]^+$, 651.52 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_{10}S]^+$ = 606.26, $[C_{23}H_{39}N_7NaO_{10}S]^+$ = 628.24, $[C_{23}H_{39}KN_7O_{10}S]^+$ = 644.21, $[C_{23}H_{38}N_7Na_2O_{10}S]^+$ = 650.22

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 19.00 min

Analytical RP-HPLC: Method I, t_R = 11.58 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) (PI 4)



PI 4

The synthesis of fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 4** was carried out manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 4** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of fully protected cyclic peptide **Cyc.Prot.PI 4** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 4)**

0.2352 g/ 0.264 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-) (**Cyc.Prot.PI 4**) 0.1688 g/0.193 mmol

cyclo-(-Ser-Asp-Met-Ser-Lys-Gly-) (**PI 4**) 0.0190 g/0.031 mmol (10.3%)

Formula: $C_{23}H_{39}N_7O_{10}S$

Molecular Mass: 605.25 g/mol

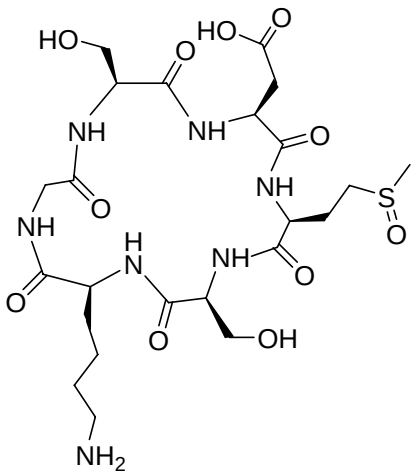
MS (MALDI-ToF): $m \cdot z^{-1}$ = 606.89 $[M+H]^+$, 628.86 $[M+Na]^+$, 644.93 $[M+K]^+$, 650.89 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_{10}S]^+$ = 606.26, $[C_{23}H_{39}N_7NaO_{10}S]^+$ = 628.24, $[C_{23}H_{39}KN_7O_{10}S]^+$ = 644.21, $[C_{23}H_{38}N_7Na_2O_{10}S]^+$ = 650.22

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 18.67 min

Analytical RP-HPLC: Method I, t_R = 11.78 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met(O)-Ser-Lys-Gly-) (PI 5)



PI 5

The synthesis of deprotected cyclic peptide **PI 3** was carried out as mentioned above. 0.03 mmol peptide **PI 3** was dissolved in DMF, into which 1 ml 30 % (v/v) H_2O_2 was injected slowly through single-channel syringe pump over 1 h at ambient temperature.

The oxidation reaction was detained at the sulfoxide level through this method, as monitored by MALDI-ToF-MS. The target oxidized derivative peptide **PI 5** was subsequently purified by RP-HPLC.

Yield: cyclo-(-Ser-Asp-Met(O)-Ser-Lys-Gly-) (**PI 5**) 0.0091 g/0.015 mmol (50.0%)

Formula: $C_{23}H_{39}N_7O_{11}S$

Molecular Mass: 621.24 g/mol

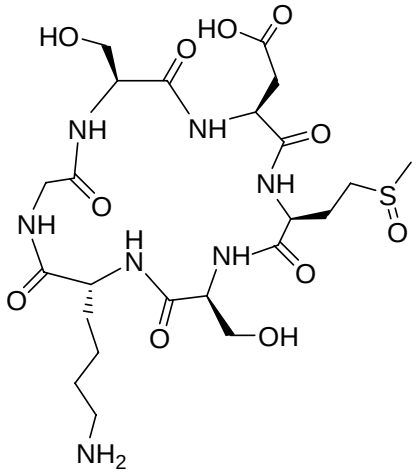
MS (MALDI-ToF): $m \cdot z^{-1}$ = 622.87 $[M+H]^+$, 644.83 $[M+Na]^+$, 660.08 $[M+K]^+$, 667.10 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_{11}S]^+$ = 622.25, $[C_{23}H_{39}N_7NaO_{11}S]^+$ = 644.23, $[C_{23}H_{39}KN_7O_{11}S]^+$ = 660.21, $[C_{23}H_{38}N_7Na_2O_{11}S]^+$ = 666.22

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 8.04 min

Analytical RP-HPLC: Method I, t_R = 4.79 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met(O)-Ser-D-Lys-Gly-) (PI 6)



PI 6

The synthesis of deprotected cyclic peptide **PI 4** was carried out as mentioned above. 0.03 mmol peptide **PI 4** was dissolved in DMF, into which 1 ml 30 % (v/v) H_2O_2 was injected slowly through single-channel syringe pump over 1 h at ambient temperature. The oxidation reaction was detained at the sulfoxide level by this method, as

monitored by MALDI-ToF-MS. The target oxidized derivative cyclic peptide **PI 6** was subsequently purified by RP-HPLC.

Yield: cyclo-(-Ser-Asp-Met(O)-Ser-Lys-Gly-) (**PI 6**) 0.0099 g/0.016 mmol (53.3%)

Formula: $C_{23}H_{39}N_7O_{11}S$

Molecular Mass: 621.24 g/mol

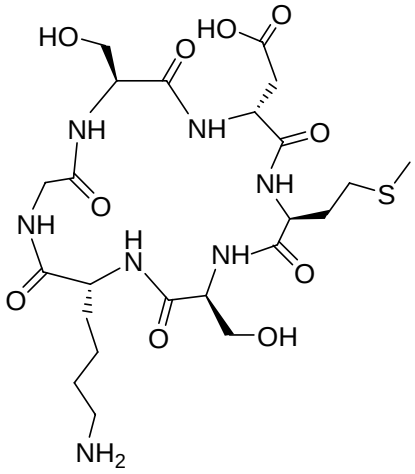
MS (MALDI-ToF): $m \cdot z^{-1}$ = 623.25 $[M+H]^+$, 645.04 $[M+Na]^+$, 661.08 $[M+K]^+$, 667.05 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_{11}S]^+$ = 622.25, $[C_{23}H_{39}N_7NaO_{11}S]^+$ = 644.23, $[C_{23}H_{39}KN_7O_{11}S]^+$ = 660.21, $[C_{23}H_{38}N_7Na_2O_{11}S]^+$ = 666.22

Preparative RP-HPLC: Apparatus A, Method 4, t_R = 8.04 min

Analytical RP-HPLC: Method I, t_R = 4.16 min, 100% (220 nm)

cyclo-(-Ser-D-Asp-Met-Ser-D-Lys-Gly-) (PI 7)



PI 7

The synthesis of protected linear peptide H-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 7** was performed manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the

precursor linear peptide **Prot.PI 7** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 7** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 7**)
0.1403 g/ 0.16 mmol
cyclo-(-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-)
(**Cyc.Prot.PI 7**) 0.0892 g/0.10 mmol
cyclo-(-Ser-D-Asp-Met-Ser-D-Lys-Gly-) (**PI 7**) 0.0242 g/0.04 mmol (13.3%)

Formula: $C_{23}H_{39}N_7O_{10}S$

Molecular Mass: 605.25 g/mol

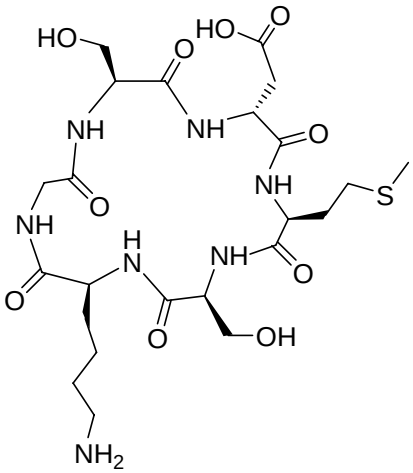
MS (MALDI-ToF): $m \cdot z^{-1}$ = 607.80 $[M+H]^+$, 629.28 $[M+Na]^+$, 645.83 $[M+K]^+$, 651.84 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_{10}S]^+$ = 606.26, $[C_{23}H_{39}N_7NaO_{10}S]^+$ = 628.24, $[C_{23}H_{39}KN_7O_{10}S]^+$ = 644.21, $[C_{23}H_{38}N_7Na_2O_{10}S]^+$ = 650.22

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 15.04 min

Analytical RP-HPLC: Method I, t_R = 11.63 min, 100% (220 nm)

cyclo-(-Ser-D-Asp-Met-Ser-Lys-Gly-) (PI 8)



PI 8

Fully protected linear peptide H-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH **Prot.PI 8** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the fully protected precursor linear peptide **Prot.PI 8** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 8** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH (**Prot.PI 8**)

0.1811g/ 0.203mmol

cyclo-(-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-) (**Cyc.Prot.PI 8**)

0.113 g/0.13 mmol

cyclo-(-Ser-D-Asp-Met-Ser-Lys-Gly-) (**PI 8**) 0.0363 g/0.06 mmol (20%)

Formula: $C_{23}H_{39}N_7O_{10}S$

Molecular Mass: 605.25 g/mol

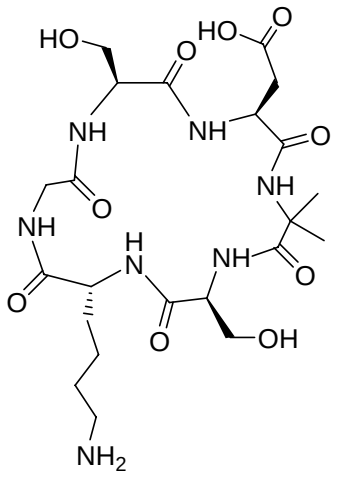
MS (MALDI-ToF): $m \cdot z^{-1}$ = 607.93 $[M+H]^+$, 629.96 $[M+Na]^+$, 645.99 $[M+K]^+$, 652.00 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_{10}S]^+$ = 606.26, $[C_{23}H_{39}N_7NaO_{10}S]^+$ = 628.24, $[C_{23}H_{39}KN_7O_{10}S]^+$ = 644.21, $[C_{23}H_{38}N_7Na_2O_{10}S]^+$ = 650.22

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 16.33 min

Analytical RP-HPLC: Method I, t_R = 11.72 min, 100% (220 nm)

cyclo-(-Ser-Asp-Aib-Ser-D-Lys-Gly-) (PI 9)



PI 9

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Aib-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 9** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.3 mmol), TBTU as coupling reagent. Aib was incorporated into the growing peptide chains on the support as the form of N_2=Aib-Cl **13** (7 equiv) in the presence of 10 equiv NEt_3. Deprotection of N^{α}-Azido group was achieved with 10 equiv $Sn(SPh)_2$, 30 equiv PhSH, and 50 equiv NEt_3. Cyclization of the precursor linear peptide **Prot.PI 9** was carried out under pseudo high dilution conditions with HATU as coupling reagent according to the aforementioned procedure. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 9** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Aib-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 9)**

0.2777 g/ 0.16 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Aib-Ser(tBu)-D-Lys(Boc)-Gly-) (**Cyc.Prot.PI 9**)

0.1208 g/0.15mol

cyclo-(-Ser-Asp-Aib-Ser-D-Lys-Gly-) (**PI 9**) 0.0772 g/0.14 mmol (46.6%)

Formula: $C_{22}H_{37}N_7O_{10}$

Molecular Mass: 559.26 g/mol

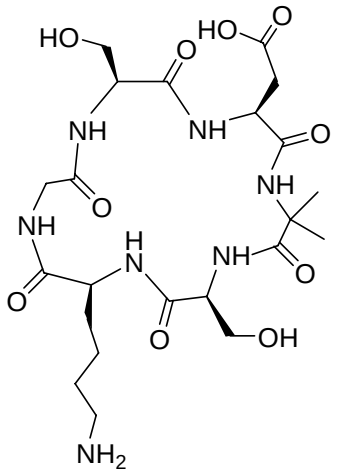
MS (MALDI-ToF): $m \cdot z^{-1}$ = 560.53 $[M+H]^+$, 582.54 $[M+Na]^+$, 598.53 $[M+K]^+$, 604.52 $[M+2Na-H]^+$

Calculated Mass: $[C_{22}H_{38}N_7O_{10}]^+$ = 560.27, $[C_{22}H_{37}N_7NaO_{10}]^+$ = 582.25, $[C_{22}H_{37}KN_7O_{10}]^+$ = 598.22, $[C_{22}H_{36}N_7Na_2O_{10}]^+$ = 604.23

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 8.34 min

Analytical RP-HPLC: Method I, t_R = 5.41 min, 100% (220 nm)

cyclo-(-Ser-Asp-Aib-Ser-Lys-Gly-) (PI 10)



PI 10

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Aib-Ser(tBu)-Lys(Boc)-Gly-OH **Prot.PI 10** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.3 mmol), TBTU as coupling reagent. Aib was incorporated into the growing peptide chains on the support as the form of N_2=Aib-Cl **13** (7 equiv) at the presence of 10 equiv NEt_3. Deprotection of N^{α}-Azido group was achieved with 10 equiv $Sn(SPh)_2$, 30 equiv PhSH, and 50 equiv NEt_3. Cyclization of the precursor linear peptide **Prot.PI 10** was carried out under pseudo high dilution conditions with

HATU as coupling reagent according to the aforementioned procedure. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 10** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Aib-Ser(tBu)-Lys(Boc)-Gly-OH (**Prot.PI 10**)

0.2887 g/ 0.16 mmol

cyclo-(-Ser(tBu)-D-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-) (**Cyc.Prot.PI 10**)

0.1308 g/0.16 mmol

cyclo-(-Ser-Asp-Aib-Ser-Lys-Gly-) (**PI 10**)

0.0606 g/0.11 mmol (36.6%) (after multiple RP-HPLC purification)

Formula: $C_{22}H_{37}N_7O_{10}$

Molecular Mass: 559.26 g/mol

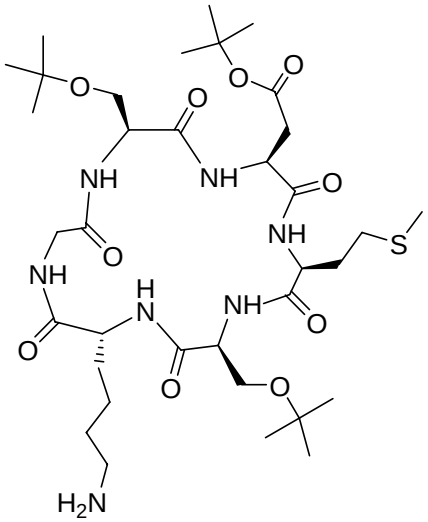
MS (MALDI-ToF): $m \cdot z^{-1}$ = 560.65 $[M+H]^+$, 582.60 $[M+Na]^+$, 598.68 $[M+K]^+$, 604.68 $[M+2Na-H]^+$

Calculated Mass: $[C_{22}H_{38}N_7O_{10}]^+$ = 560.27, $[C_{22}H_{37}N_7NaO_{10}]^+$ = 582.25, $[C_{22}H_{37}KN_7O_{10}]^+$ = 598.22, $[C_{22}H_{36}N_7Na_2O_{10}]^+$ = 604.23

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 8.00 min

Analytical RP-HPLC: Method I, t_R = 5.58 min, 100% (220 nm)

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys-Gly-) (13)



13

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Aloc)-Gly-OH **Prot.13** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.13** was carried out under pseudo high dilution conditions with HATU as coupling reagent according to the aforementioned procedure. The fully protected cyclic peptide **Cyc.Pro.13** was applied to selective deprotection with 1.0 equiv $Pd(PPh_3)_4$ and 30 equiv $PhSiH_3$ in dry DCM at room temperature for 3 h, or alternatively, with 0.1 equiv $Pd(PPh_3)_4$ and 30 equiv $PhSiH_3$ in liquid ammonia under -33^{o}C for 3 h.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Aloc)-Gly-OH (**Prot.13**)

0.2251g/ 0.25mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Aloc)-Gly-) (**Cyc.Prot.13**)

0.1886 /0.22 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys-Gly-) (**13**)

0.1393 g/0.18 mmol (60.0%)

Formula: $C_{35}H_{63}N_7O_{10}S$

Molecular Mass: 773.44 g/mol

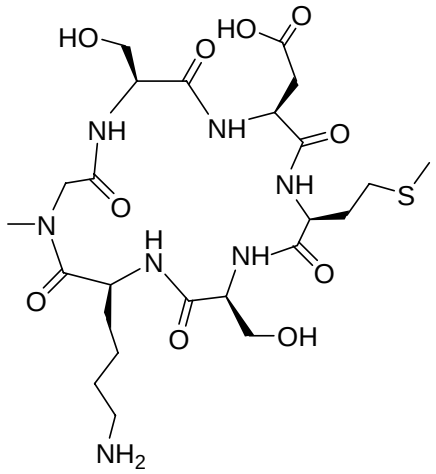
MS (MALDI-ToF): $m \cdot z^{-1}$ = 774.82 $[M+H]^+$, 796.94 $[M+Na]^+$, 812.89 $[M+K]^+$

Calculated Mass: $[C_{35}H_{64}N_7O_{10}S]^+$=774.44, $[C_{35}H_{63}N_7NaO_{10}S]^+$=796.42, $[C_{35}H_{63}KN_7O_{10}S]^+$=812.40

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 8.34 min

Analytical RP-HPLC: Method I, t_R = 5.41 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-Lys-Sar-) (PI 11)



PI 11

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Sar-OH **Prot.PI 11** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Sar-OH being loaded on the *o*-chlorotrityl resin (0.71 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 11** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 11** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Sar-OH (**Prot.PI 11**)

0.0922 g/ 0.10 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Lys(Boc)-Sar-) (**Cyc.Prot.PI 11**)

0.0617 g/0.07 mmol

cyclo-(-Ser-Asp-Met-Ser-Lys-Sar-) (**PI 11**)

0.0335g/0.05 mmol (16.6%)

Formula: $C_{24}H_{41}N_7O_{10}S$

Molecular Mass: 619.26 g/mol

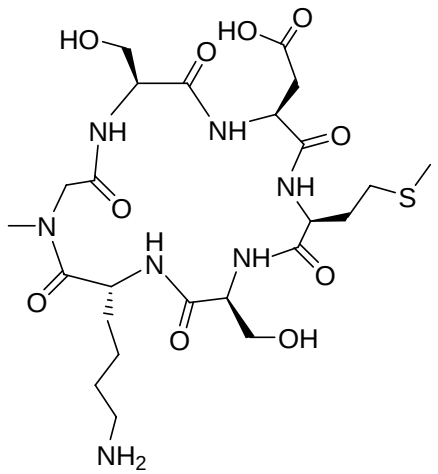
MS (MALDI-ToF): $m \cdot z^{-1}$ = 621.50 $[M+H]^+$, 643.51 $[M+Na]^+$, 659.53 $[M+K]^+$, 665.54 $[M+2Na-H]^+$

Calculated Mass: $[C_{24}H_{42}N_7O_{10}S]^+$ = 620.27, $[C_{24}H_{41}N_7NaO_{10}S]^+$ = 642.25, $[C_{24}H_{41}N_7KO_{10}S]^+$ = 658.23, $[C_{24}H_{40}N_7Na_2O_{10}S\ ]^+$ = 664.24

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 18.77 min

Analytical RP-HPLC: Method I, t_R = 12.11 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys-Sar-) (PI 12)



PI 12

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Sar-OH **Prot.PI 12** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Sar-OH being loaded on the *o*-chlorotrityl resin (0.71 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 12** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 12** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Sar-OH (**Prot.PI 12**)

0.0921 g/0.102 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)- D-Lys(Boc)-Sar-) (**Cyc.Prot.PI 12**)

0.054g/0.06 mmol

cyclo-(-Ser-Asp-Met-Ser- D-Lys-Sar-) (**PI 12**)

0.0149g/0.024 mmol (8.0%)

Formula: $C_{24}H_{41}N_7O_{10}S$

Molecular Mass: 619.26 g/mol

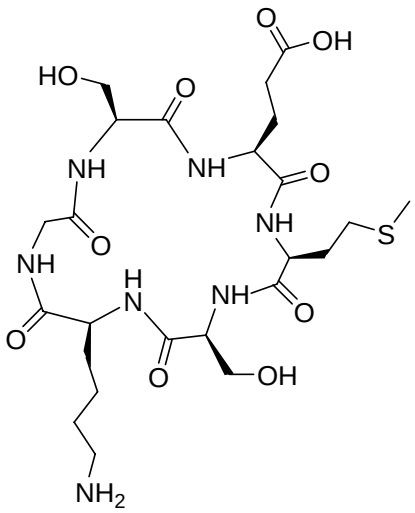
MS (MALDI-ToF): $m \cdot z^{-1}$ = 621.22 $[M+H]^+$, 643.24 $[M+Na]^+$, 659.22 $[M+K]^+$, 665.26 $[M+2Na-H]^+$

Calculated Mass: $[C_{24}H_{42}N_7O_{10}S]^+$ = 620.27, $[C_{24}H_{41}N_7NaO_{10}S]^+$ = 642.25, $[C_{24}H_{41}N_7KO_{10}S]^+$ = 658.23, $[C_{24}H_{40}N_7Na_2O_{10}S\ ]^+$ = 664.24

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 20.60 min

Analytical RP-HPLC: Method I, t_R = 12.04 min, 100% (220 nm)

cyclo-(-Ser-Glu-Met-Ser-Lys-Gly-) (PI 13)



PI 13

The fully protected linear peptide H-Ser(tBu)-Glu(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH **Prot.PI 13** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 13** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 13** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Glu(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-OH (**Prot.PI 13**)

0.1050 g/ 0.118 mmol

cyclo-(-Ser(tBu)-Glu(OtBu)-Met-Ser(tBu)-Lys(Boc)-Gly-) (**Cyc.Prot.PI 13**)

0.0565 g/0.064 mmol

cyclo-(-Ser-Glu-Met-Ser-Lys-Gly-) (**PI 13**)

0.0080 g/0.013 mmol (4.3%)

Formula: $C_{24}H_{41}N_7O_{10}S$

Molecular Mass: 619.26 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 620.78 $[M+H]^+$, 642.48 $[M+Na]^+$, 658.78 $[M+K]^+$, 664.78

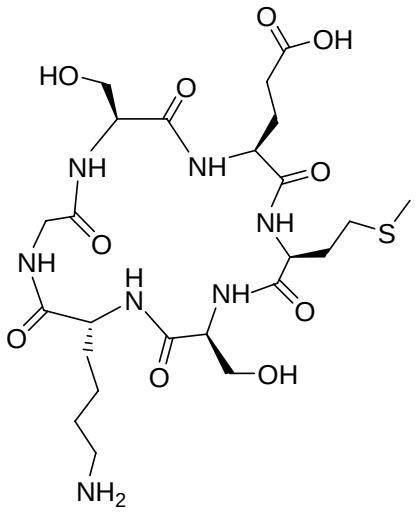
$[M+2Na-H]^+$

Calculated Mass: $[C_{24}H_{42}N_7O_{10}S]^+$ = 620.27, $[C_{24}H_{41}N_7NaO_{10}S]^+$ = 642.25, $[C_{24}H_{41}N_7KO_{10}S]^+$ = 658.23, $[C_{24}H_{40}N_7Na_2O_{10}S]^+$ = 664.24

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 22.29 min

Analytical RP-HPLC: Method I, t_R = 10.59 min, 100% (220 nm)

cyclo-(-Ser-Glu-Met-Ser-D-Lys-Gly-) (PI 14)



PI 14

The fully protected linear peptide H-Ser(tBu)-Glu(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 14** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 14** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 14** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Glu(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 14**)

0.2021 g/ 0.223 mmol

cyclo-(-Ser(tBu)-Glu(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-) (**Cyc.Prot.PI 14**)

0.1342 g/0.151 mmol

cyclo-(-Ser-Glu-Met-Ser- D-Lys-Gly-) (**PI 14**)

0.0256 g/0.041 mmol (13.7%)

Formula: $C_{24}H_{41}N_7O_{10}S$

Molecular Mass: 619.26 g/mol

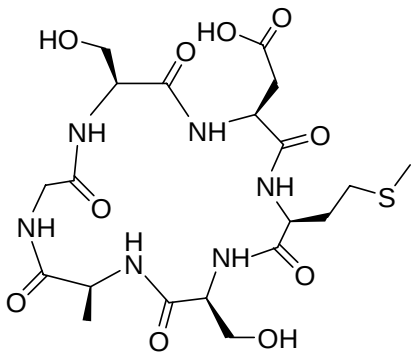
MS (MALDI-ToF): $m \cdot z^{-1}$ = 621.34 $[M+H]^+$, 643.34 $[M+Na]^+$, 659.34 $[M+K]^+$, 665.34 $[M+2Na-H]^+$

Calculated Mass: $[C_{24}H_{42}N_7O_{10}S]^+$ = 620.27, $[C_{24}H_{41}N_7NaO_{10}S]^+$ = 642.25, $[C_{24}H_{41}N_7KO_{10}S]^+$ = 658.23, $[C_{24}H_{40}N_7Na_2O_{10}S]^+$ = 664.24

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 19.82 min

Analytical RP-HPLC: Method I, t_R = 12.01 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-Ala-Gly-) (PI 15)



PI 15

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Ala-Gly-OH **Prot.PI 15** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 15** was carried out under pseudo high dilution conditions with

HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 15** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Ala-Gly-OH (**Prot.PI 15**)

0.1406 g/ 0.191 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Ala-Gly-) (**Cyc.Prot.PI 15**)

0.0449 g/0.063 mmol

cyclo-(-Ser-Asp-Met-Ser-Ala-Gly-) (**PI 15**)

0.0098 g/0.0178 mmol (5.9%)

Formula: $C_{20}H_{32}N_6O_{10}S$

Molecular Mass: 548.19 g/mol

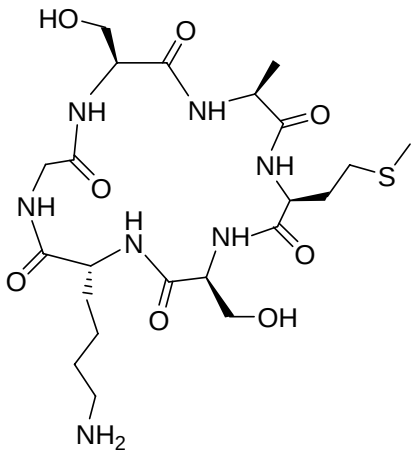
MS (MALDI-ToF): $m \cdot z^{-1}$ = 549.62 $[M+H]^+$, 571.65 $[M+Na]^+$, 587.59 $[M+K]^+$

Calculated Mass: $[C_{20}H_{33}N_6O_{10}S]^+$ = 549.20, $[C_{20}H_{32}N_6NaO_{10}S]^+$ = 571.18, $[C_{20}H_{32}N_6KO_{10}S]^+$ = 587.15

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 25.68 min

Analytical RP-HPLC: Method I, t_R = 12.46 min, 100% (220 nm)

cyclo-(-Ser-Ala-Met-Ser-D-Lys-Gly-) (PI 16)



PI 16

The fully protected linear peptide H-Ser(tBu)-Ala-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 16** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 16** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 16** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Ala-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 16**)

0.1687 g/ 0.213 mmol

cyclo-(-Ser(tBu)-Ala-Met-Ser(tBu)-D-Lys(Boc)-Gly-) (**Cyc.Prot.PI 16**)

0.1004 g/0.130 mmol

cyclo-(-Ser-Ala-Met-Ser- D-Lys-Gly-) (**PI 16**)

0.0296 g/0.053 mmol (17.7%)

Formula: $C_{22}H_{39}N_7O_8S$

Molecular Mass: 561.26 g/mol

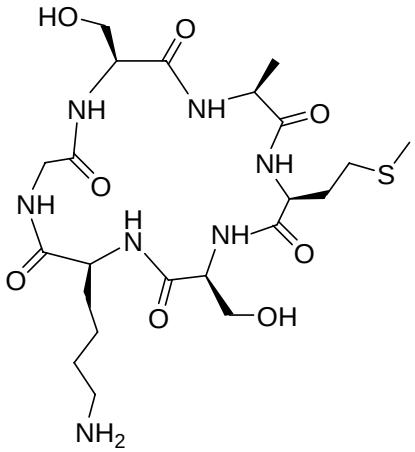
MS (MALDI-ToF): $m \cdot z^{-1}$ = 563.32 $[M+H]^+$, 585.35 $[M+Na]^+$, 601.34 $[M+K]^+$

Calculated Mass: $[C_{22}H_{40}N_7O_8S]^+$ = 562.27, $[C_{22}H_{39}N_7NaO_8S]^+$ = 584.25, $[C_{22}H_{39}N_7KO_8S]^+$ = 600.22

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 19.71 min

Analytical RP-HPLC: Method I, t_R = 12.13 min, 100% (220 nm)

cyclo-(-Ser-Ala-Met-Ser-Lys-Gly-) (PI 17)



PI 17

The fully protected linear peptide H-Ser(tBu)-Ala-Met-Ser(tBu)-Lys(Boc)-Gly-OH **Prot.PI 17** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 17** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 17** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Ala-Met-Ser(tBu)-Lys(Boc)-Gly-OH (**Prot.PI 17**)

0.1209 g/ 0.153 mmol

cyclo-(-Ser(tBu)-Ala-Met-Ser(tBu)-Lys(Boc)-Gly-) (**Cyc.Prot.PI 17**)

0.0385 g/0.0498 mmol

cyclo-(-Ser-Ala-Met-Ser-Lys-Gly-) (**PI 17**)

0.0175 g/0.031 mmol (10.3%)

Formula: $C_{22}H_{39}N_7O_8S$

Molecular Mass: 561.26 g/mol

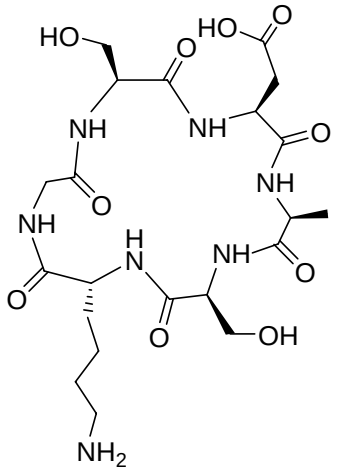
MS (MALDI-ToF): $m \cdot z^{-1}$ = 563.26 $[M+H]^+$, 585.31 $[M+Na]^+$, 601.38 $[M+K]^+$

Calculated Mass: $[C_{22}H_{40}N_7O_8S]^+$ = 562.27, $[C_{22}H_{39}N_7NaO_8S]^+$ = 584.25, $[C_{22}H_{39}N_7KO_8S]^+$ = 600.22

Preparative RP-HPLC: Apparatus A, Method 5, t_R = 24.90 min

Analytical RP-HPLC: Method I, t_R = 15.77 min, 100% (220 nm)

cyclo-(-Ser-Asp-Ala-Ser-D-Lys-Gly-) (PI 18)



PI 18

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Ala-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 18** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 18** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 18** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Ala-Ser(tBu)-D-Lys(Boc)-Gly-OH (**Prot.PI 18**)

0.1850 g/ 0.222 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Ala-Ser(tBu)-D-Lys(Boc)-Gly-) (**Cyc.Prot.PI 18**) 0.1286 g/0.158 mmol

cyclo-(-Ser-Asp-Ala-Ser- D-Lys-Gly-) (**PI 18**) 0.0181 g/0.033 mmol (11.0%)

Formula: $C_{21}H_{35}N_7O_{10}$

Molecular Mass: 545.24 g/mol

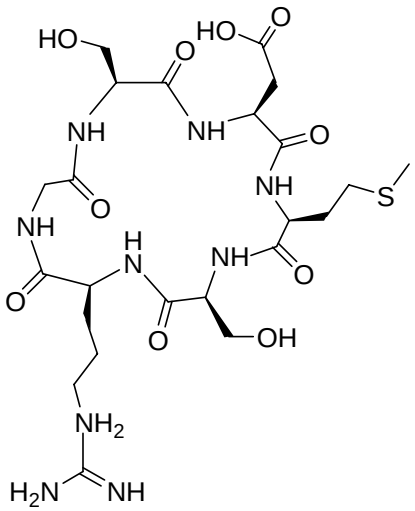
MS (MALDI-ToF): $m \cdot z^{-1}$ = 546.88 $[M+H]^+$, 568.89 $[M+Na]^+$, 584.86 $[M+K]^+$

Calculated Mass: $[C_{21}H_{36}N_7O_{10}]^+$ = 546.25, $[C_{21}H_{35}N_7NaO_{10}]^+$ = 568.23, $[C_{21}H_{35}KN_7O_{10}]^+$ = 584.21

Preparative RP-HPLC: Apparatus A, Method 4, t_R = 7.00 min

Analytical RP-HPLC: Method I, t_R = 4.36 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-Arg-Gly-) (PI 19)



PI 19

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Arg(Pbf)-Gly-OH **Prot.PI 19** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 19** was carried out under pseudo high dilution conditions with

HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 19** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Arg(Pbf)-Gly-OH (**Prot.PI 19**)

0.1820 g/ 0.173 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Arg(Pbf)-Gly-) (**Cyc.Prot.PI 19**)

0.1087 g/0.105 mmol

cyclo-(-Ser-Asp-Met-Ser-Arg-Gly-) (**PI 19**)

0.0170 g/0.0268 mmol (8.9%)

Formula: $C_{23}H_{39}N_9O_{10}S$

Molecular Mass: 633.25 g/mol

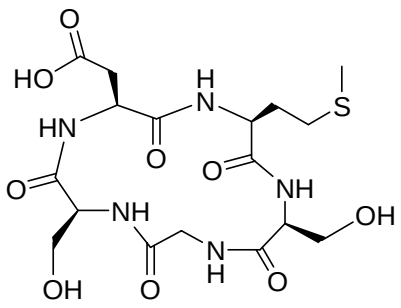
MS (MALDI-ToF): $m \cdot z^{-1}$ = 634.68 $[M+H]^+$, 656.59 $[M+Na]^+$, 672.67 $[M+K]^+$, 678.70 $[M+2Na-H]^+$

Calculated Mass: $[C_{23}H_{40}N_9O_{10}S]^+$ = 634.26, $[C_{23}H_{39}N_9NaO_{10}S]^+$ = 656.24, $[C_{23}H_{39}N_9KO_{10}S]^+$ = 672.22, $[C_{23}H_{38}N_9Na_2O_{10}S]$ = 678.23

Preparative RP-HPLC: Apparatus A, Method 2, t_R = 20.09 min

Analytical RP-HPLC: Method I, t_R = 17.81 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-Gly-) (PI 20)



PI 20

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Gly-OH **Prot.PI 20** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.998 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 20** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 20** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Gly-OH (**Prot.PI 20**)

0.1113 g/0.168 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Gly-) (**Cyc.Prot.PI 20**)

0.0144 g/0.022 mmol

cyclo-(-Ser-Asp-Met-Ser-Gly-) (**PI 20**)

0.0017 g/0.0036 mmol (1.2%)

Formula: $C_{17}H_{27}N_5O_9S$

Molecular Mass: 477.15 g/mol

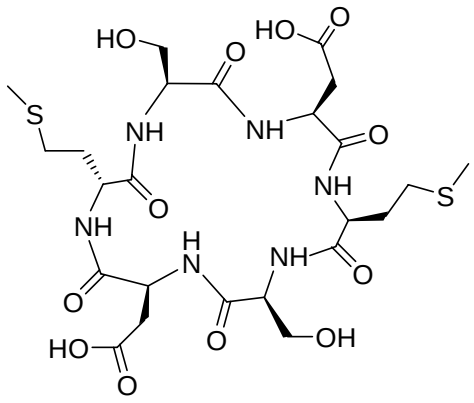
MS (MALDI-ToF): $m \cdot z^{-1}$ = 478.89 $[M+H]^+$, 500.92 $[M+Na]^+$, 516.93 $[M+K]^+$

Calculated Mass: $[C_{17}H_{28}N_5O_9S]^+$ = 478.16, $[C_{17}H_{27}N_5NaO_9S]^+$ = 500.14, $[C_{17}H_{27}KN_5O_9S]^+$ = 516.12

Preparative RP-HPLC: Apparatus A, Method 5, t_R = 20.62 min

Analytical RP-HPLC: Method I, t_R = 14.05 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-Asp-Met-) (PI 21)



PI 21

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Asp(OtBu)-Met-OH **Prot.PI 21** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Met-OH being loaded on the *o*-chlorotrityl resin (0.85 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 21** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 21** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Asp(OtBu)-Met-OH (**Prot.PI 21**)

0.1862 g/ 0.205 mmol

cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-Asp(OtBu)-Met-) (**Cyc.Prot.PI 21**)

0.0479 g/0.054 mmol

cyclo-(-Ser-Asp-Met-Ser-Asp-Met-) (**PI 21**)

0.0212 g/0.032 mmol (10.7%)

Formula: $C_{24}H_{38}N_6O_{12}S_2$

Molecular Mass: 666.20 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 667.62 $[M+H]^+$, 689.66 $[M+Na]^+$, 705.64 $[M+K]^+$

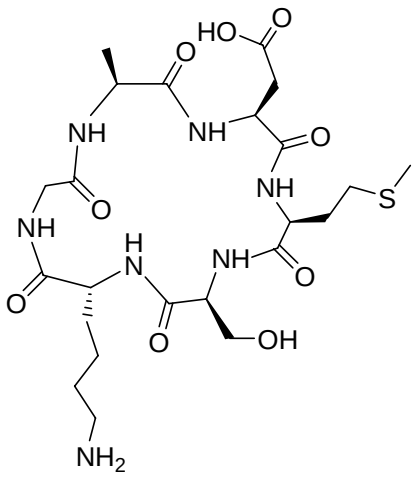
Calculated Mass: $[C_{24}H_{39}N_6O_{12}S_2]^+$ = 667.21, $[C_{24}H_{38}N_6NaO_{12}S_2]^+$ = 689.18,

$[C_{24}H_{38}N_6NaO_{12}S_2]^+ = 705.16$

Preparative RP-HPLC: Apparatus A, Method 6, $t_R = 47.40$ min

Analytical RP-HPLC: Method I, $t_R = 18.80$ min, 100% (220 nm)

cyclo-(-Ala-Asp-Met-Ser-D-Lys-Gly-) (PI 22)



PI 22

The fully protected linear peptide H-Ala-Asp(OtBu)-Met-Ser(tBu)-D-Lys(Boc)-Gly-OH **Prot.PI 22** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Gly-OH being loaded on the *o*-chlorotrityl resin (0.923 mmol/g, 0.3 mmol), TBTU as coupling reagent. Cyclization of the precursor linear peptide **Prot.PI 22** was carried out under pseudo high dilution conditions with HATU as coupling reagent. Deprotection of the fully protected cyclic peptide **Cyc.Prot.PI 22** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ala-Asp(OtBu)-Met-Ser(tBu)-Asp(OtBu)-Met-OH (**Prot.PI 22**)

0.2090 g/ 0.255 mmol

cyclo-(-Ala-Asp(OtBu)-Met-Ser(tBu)-Asp(OtBu)-Met-) (**Cyc.Prot.PI 22**)

0.1603 g/ 0.200 mmol

cyclo-(-Ala-Asp-Met-Ser-Asp-Met-) (**PI 22**)

0.0560 g/ 0.095 mmol (31.7%)

Formula: $C_{23}H_{39}N_7O_9S$

Molecular Mass: 589.25 g/mol

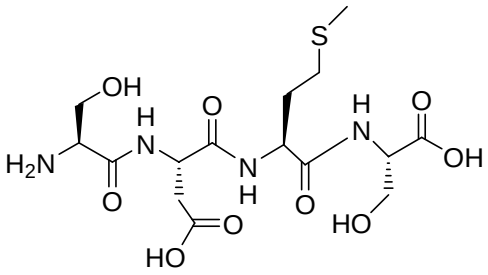
MS (MALDI-ToF): $m \cdot z^{-1}$ = 590.30 $[M+H]^+$, 612.29 $[M+Na]^+$, 628.30 $[M+K]^+$

Calculated Mass: $[C_{23}H_{40}N_7O_9S]^+$ = 590.26, $[C_{23}H_{39}N_7NaO_9S]^+$ = 612.24, $[C_{23}H_{39}KN_7O_9S]^+$ = 628.21

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 21.07 min

Analytical RP-HPLC: Method I, t_R = 17.35 min, 100% (220 nm)

H-Ser-Asp-Met-Ser-OH (PI 23)



PI 23

The fully protected linear peptide H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-OH **Prot.PI 23** was synthesized manually according to the general procedure of SPPS (Fmoc/tBu), with Fmoc-Ser(tBu)-OH being loaded on the *o*-chlorotrityl resin (0.85 mmol/g, 0.3 mmol), TBTU as coupling reagent. Deprotection of the fully protected linear peptide **Prot.PI 23** was achieved with the deprotection cocktail Reagent K.

Yield: H-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-OH (**Prot.PI 23**)

0.1758 g/ 0.29 mmol

H-Ser-Asp-Met-Ser-Asp-Met-OH (**PI 23**)

0.0876 g/ 0.20 mmol (66.7%)

Formula: $C_{15}H_{26}N_4O_9S$

Molecular Mass: 438.14 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 439.20 $[M+H]^+$, 461.15 $[M+Na]^+$, 477.15 $[M+K]^+$
Calculated Mass: $[C_{15}H_{27}N_4O_9S]^+$ = 439.15, $[C_{15}H_{26}N_4NaO_9S]^+$ = 461.13, $[C_{15}H_{26}KN_4O_9S]^+$ = 477.11

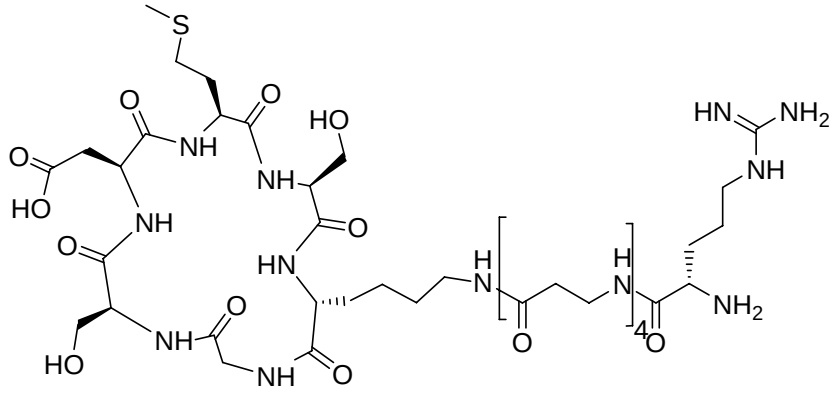
Preparative RP-HPLC: Apparatus A, Method 6, t_R = 10.50 min

Analytical RP-HPLC: Method I, t_R = 7.45 min, 100% (220 nm)

Conjugation of SDMS peptides with a remote Arg/Asp with a linker

1.5 equiv *N*-Fmoc side-chain protected linker peptides with a remote Asp/Arg at the flank (see also chapter **2.5.1.2.2**) in DMF (in relative to cyclo-(-Ser(tBu)-Asp(OtBu)-Met-Ser(tBu)-D-Lys-Gly-) **13**) was mixed with 3 equiv HATU and 6 equiv DIPEA at room temperature for approximately 15 min before being added to 1 equiv **13** in DMF. The two peptide fragments were condensed overnight at room temperature. The formation of the target peptide complex was monitored by MALDI-ToF-MS. This totally protected peptide conjugate was purified by RP-HPLC. Subsequently the Fmoc protecting group was cleaved with 2% DBU, 2% piperidine in DMF. This process took 15-20 min at room temperature before the solvent was removed at reduced pressure. The residue was then treated with reagent K to cleave the acid-labile protecting groups without further purification. The deprotecting process took approximately 30 min at room temperature. The target peptide conjugate was purified by RP-HPLC and analyzed by MALDI-ToF-MS and analytical RP-HPLC.

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Arg-(β-Ala)$_4$-)-Gly-) (PI 24)



PI 24

Yield: 2.9 %

Formula: $C_{41}H_{71}N_{15}O_{15}S$

Molecular Mass: 1045.50 g/mol

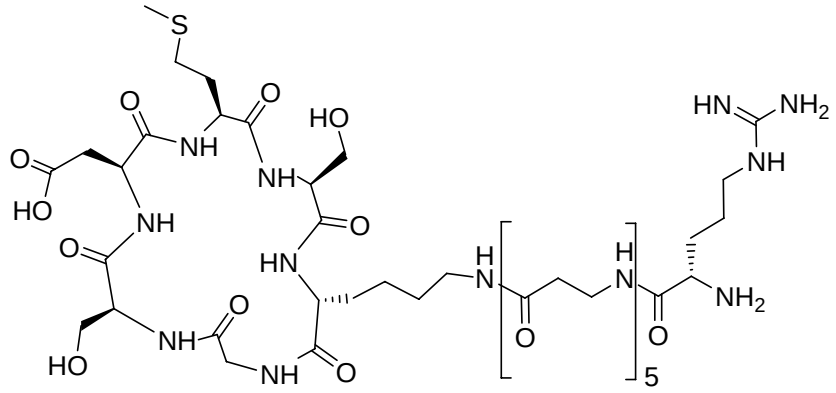
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1045.19 $[M+H]^+$, 1067.15 $[M+Na]^+$, 1083.16 $[M+K]^+$, 1089.14 $[M+2Na-H]^+$

Calculated Mass: $[C_{41}H_{72}N_{15}O_{15}S]^+$ = 1046.50, $[C_{41}H_{71}N_{15}NaO_{15}S]^+$ = 1068.49, $[C_{41}H_{71}KN_{15}O_{15}S]^+$ = 1084.46, $[C_{41}H_{70}N_{15}Na_2O_{15}S]$ = 1090.48

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 32.52min

Analytical RP-HPLC: Method I, t_R = 13.50 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Arg-(β-Ala)$_5$-)-Gly-) (PI 25)



PI 25

Yield: 3.5 %

Formula: $C_{44}H_{76}N_{16}O_{16}S$

Molecular Mass: 1116.53 g/mol

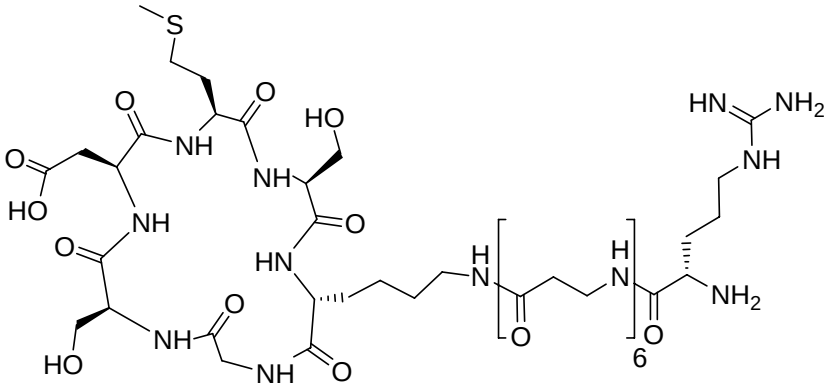
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1117.76 $[M+H]^+$, 1139.80 $[M+Na]^+$, 1155.90 $[M+K]^+$

Calculated Mass: $[C_{44}H_{77}N_{16}O_{16}S]^+$ = 1117.54, $[C_{44}H_{76}N_{16}NaO_{16}S]^+$ = 1139.52, $[C_{44}H_{76}KN_{16}O_{16}S]^+$ = 1155.50

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 33.55 min

Analytical RP-HPLC: Method I, t_R = 14.02 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Arg-(β-Ala)$_6$-)-Gly-) (PI 26)



PI 26

Yield: .4.5 %

Formula: $C_{47}H_{81}N_{17}O_{17}S$

Molecular Mass: 1187.57 g/mol

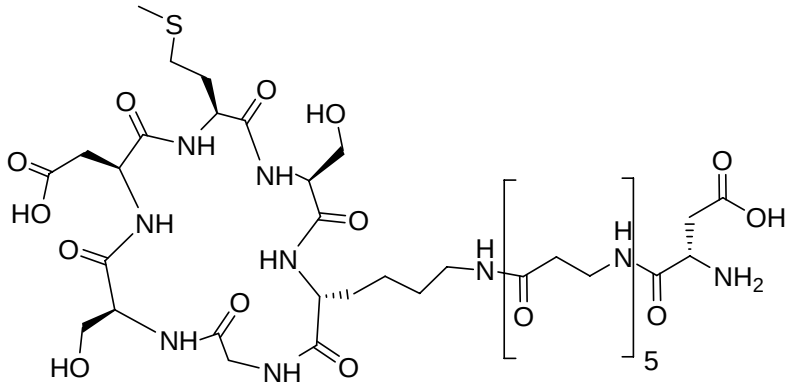
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1188.02 $[M+H]^+$, 1210.13 $[M+Na]^+$, 1226.19 $[M+K]^+$

Calculated Mass: $[C_{47}H_{82}N_{17}O_{17}S]^+$ = 1188.58, $[C_{47}H_{81}N_{17}NaO_{17}S]^+$ = 1210.56, $[C_{47}H_{81}KN_{17}O_{17}S]^+$ = 1226.54

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 34.52min

Analytical RP-HPLC: Method I, t_R = 14.46 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Asp-(β-Ala)$_5$-)-Gly-) (PI 27)



PI 27

Yield: 5.1 %

Formula: $C_{42}H_{69}N_{13}O_{18}S$

Molecular Mass: 1075.46 g/mol

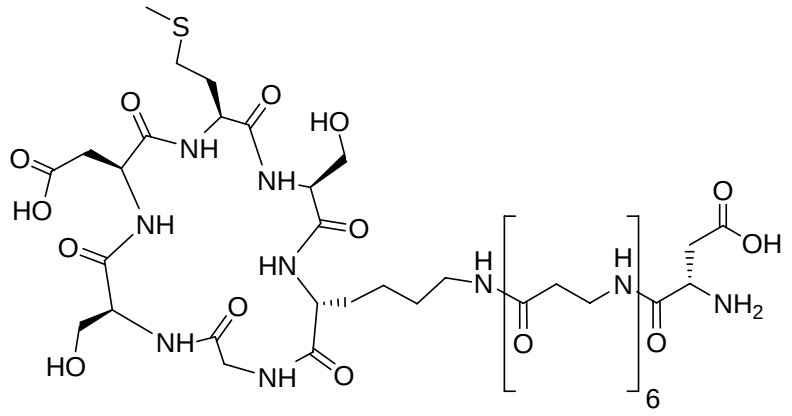
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1076.47 $[M+H]^+$, 1098.41 $[M+Na]^+$, 1114.40 $[M+K]^+$

Calculated Mass: $[C_{42}H_{70}N_{13}O_{18}S]^+$ = 1076.47, $[C_{42}H_{69}N_{13}NaO_{18}S]^+$ = 1098.45, $[C_{42}H_{69}KN_{13}O_{18}S]^+$ = 1114.42

Preparative RP-HPLC: Apparatus A, Method 7, t_R = 28.37min

Analytical RP-HPLC: Method I, t_R = 13.11 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Asp-(β-Ala)$_6$-)-Gly-) (PI 28)



PI 28

Yield: 9.8 %

Formula: $C_{45}H_{74}N14O_{19}S$

Molecular Mass: 1146.50 g/mol

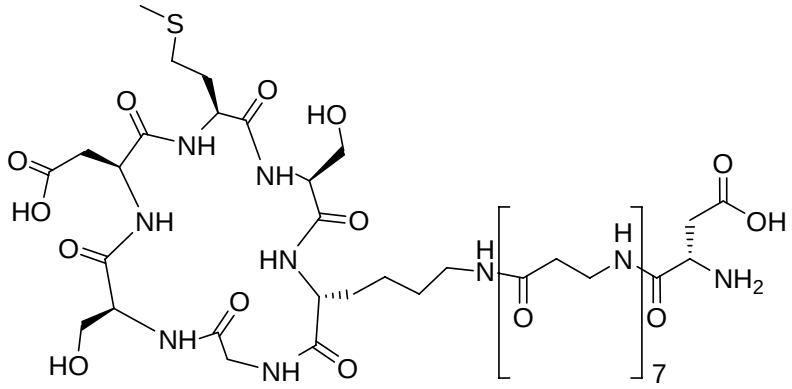
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1147.35 $[M+H]^+$, 1169.37 $[M+Na]^+$, 1186.05 $[M+K]^+$

Calculated Mass: $[C_{45}H_{75}N14O_{19}S]^+$ = 1147.51, $[C_{45}H_{74}N_{14}NaO_{19}S]^+$ = 1169.49, $[C_{45}H_{74}KN_{14}O_{19}S]^+$ = 1185.46

Preparative RP-HPLC: Apparatus B, Method 8, t_R = 40.2 min

Analytical HPLC: Method I, t_R = 13.25 min 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Asp-(β-Ala)$_7$-)-Gly-) (PI 29)



PI 29

Yield: 5.5 %

Formula: $C_{48}H_{79}N_{15}O_{20}S$

Molecular Mass: 1217.53 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 1218.44 $[M+H]^+$, 1239.92 $[M+Na]^+$, 1254.97 $[M+K]^+$

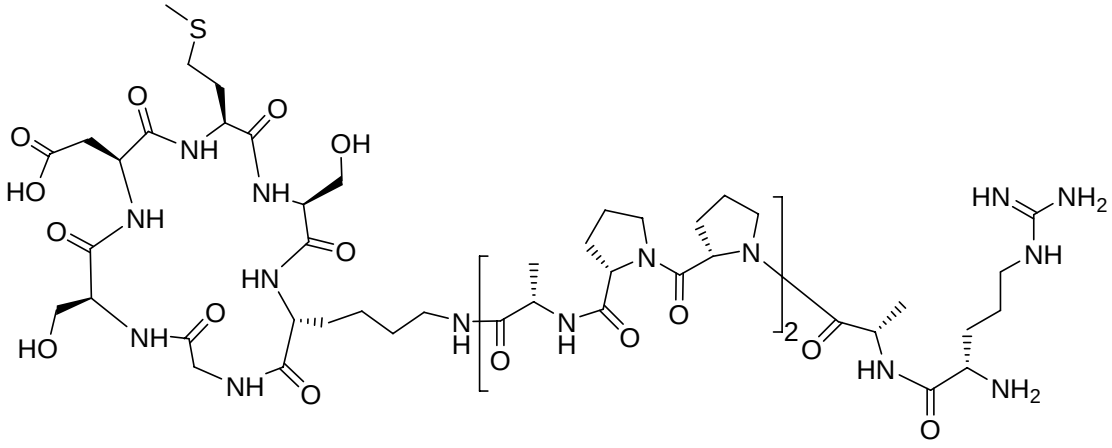
Calculated Mass: $[C_{48}H_{80}N_{15}O_{20}S]^+$ = 1218.54, $[C_{48}H_{79}N_{15}NaO_{20}S]^+$ = 1240.52, $[C_{48}H_{79}KN_{15}O_{20}S]^+$ = 1256.50

Preparative RP-HPLC: Apparatus A, Method 9, t_R = 37.79min

Analytical RP-HPLC: Method I, t_R = 17.75 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Arg-Ala-Pro-Pro-Ala-Pro-Pro-Ala)-Gly-)

(PI 30)



PI 30

Yield:6.0 %

Formula: $C_{58}H_{94}N_{18}O_{18}S$

Molecular Mass: 1362.67 g/mol

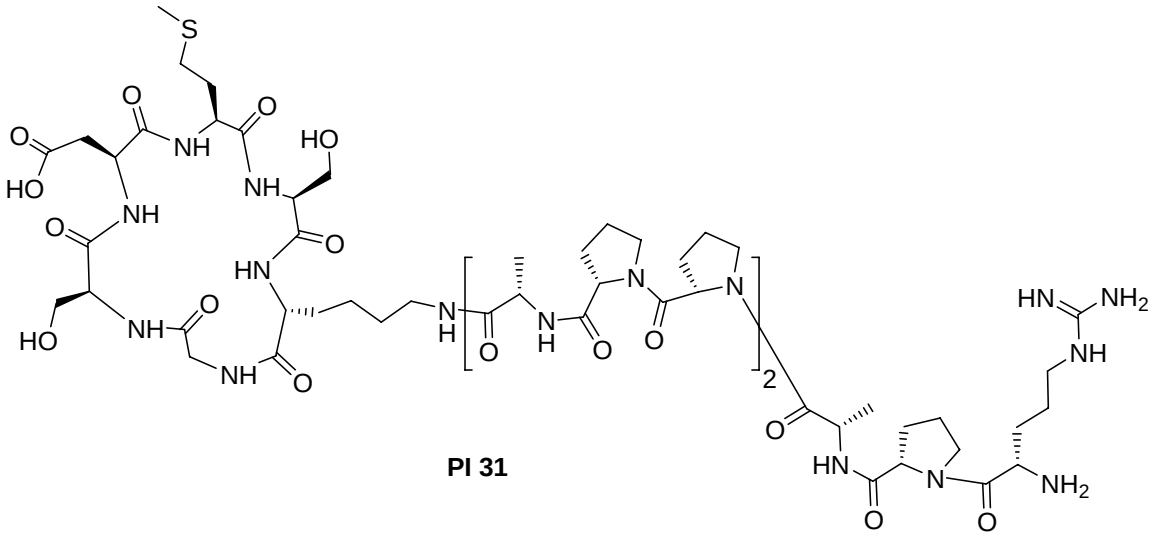
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1364.05 $[M+H]^+$, 1386.10 $[M+Na]^+$, 1402.11 $[M+K]^+$

Calculated Mass: $[C_{58}H_{95}N_{18}O_{18}S]^+$ = 1363.68, $[C_{58}H_{94}N_{18}NaO_{18}S]^+$ = 1385.66, $[C_{58}H_{94}KN_{18}O_{18}S]^+$ = 1401.64

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 38.12min

Analytical RP-HPLC: Method I, t_R = 16.11 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Arg-Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala)-Gly-) (PI 31)



PI 31

Yield: 10.3 %

Formula: $C_{63}H_{101}N_{19}O_{19}S$

Molecular Mass: 1459.72 g/mol

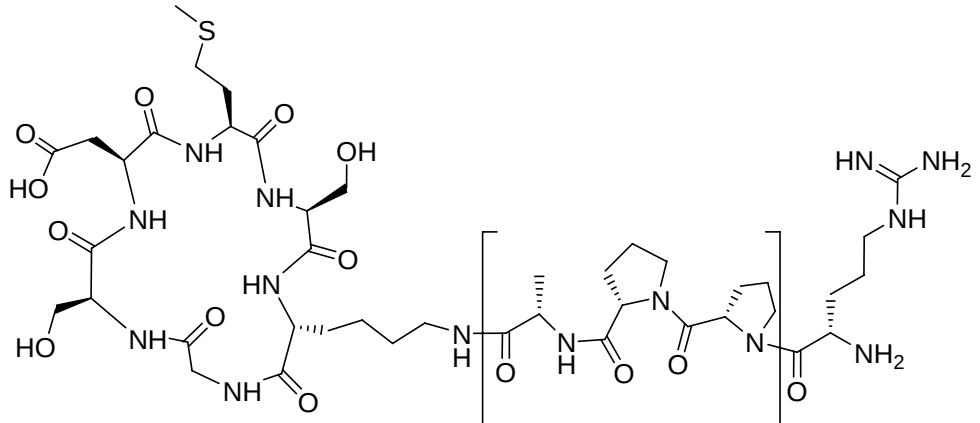
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1460.99 $[M+H]^+$, 1482.02 $[M+Na]^+$, 1499.01 $[M+K]^+$

Calculated Mass: $[C_{63}H_{102}N_{19}O_{19}S]^+$ = 1460.73, $[C_{63}H_{101}N_{19}NaO_{19}S]^+$ = 1482.71, $[C_{63}H_{101}KN_{19}O_{19}S]^+$ = 1498.69

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 39.98 min

Analytical RP-HPLC: Method I, t_R = 17.57 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Arg-Pro-Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala)-Gly-) (PI 32)



PI 32

Yield: 6.9 %

Formula: $C_{68}H_{108}N_{20}O_{20}S$

Molecular Mass: 1556.78 g/mol

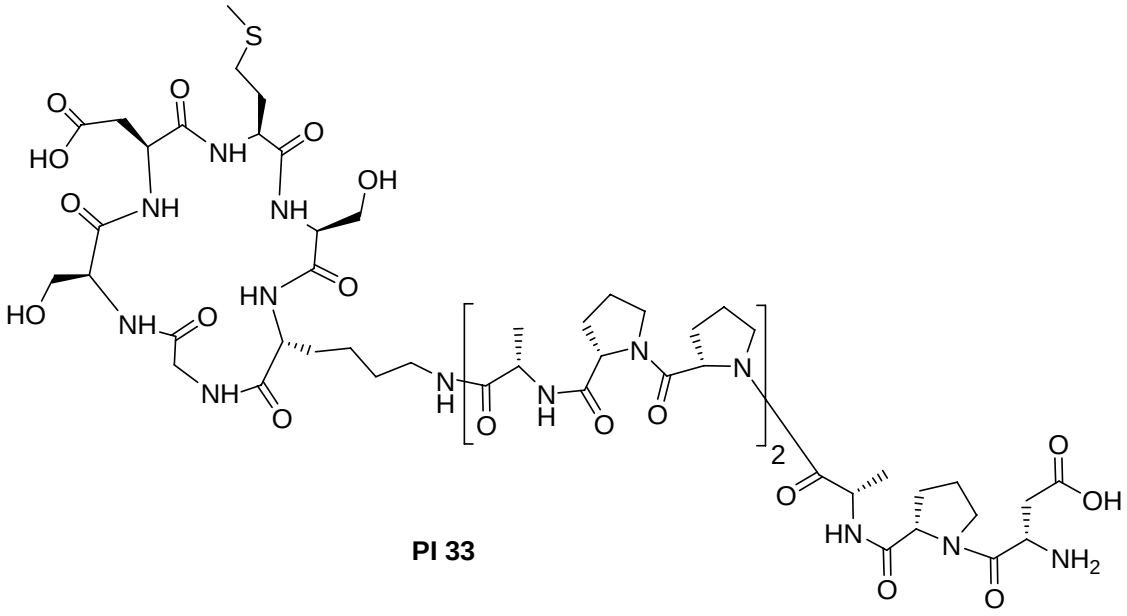
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1558.16 $[M+H]^+$, 1580.11 $[M+Na]^+$, 1596.17 $[M+K]^+$

Calculated Mass: $[C_{68}H_{109}N_{20}O_{20}S]^+$ = 1557.78, $[C_{68}H_{108}N_{20}NaO_{20}S]^+$ = 1579.77, $[C_{68}H_{108}KN_{20}O_{20}S]^+$ = 1595.74

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 40.97 min

Analytical RP-HPLC: Method I, t_R = 18.19 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Asp-Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala)-Gly-) (PI 33)



PI 33

Yield: 8.1 %

Formula: $C_{61}H_{94}N_{16}O_{21}S$

Molecular Mass: 1418.65 g/mol

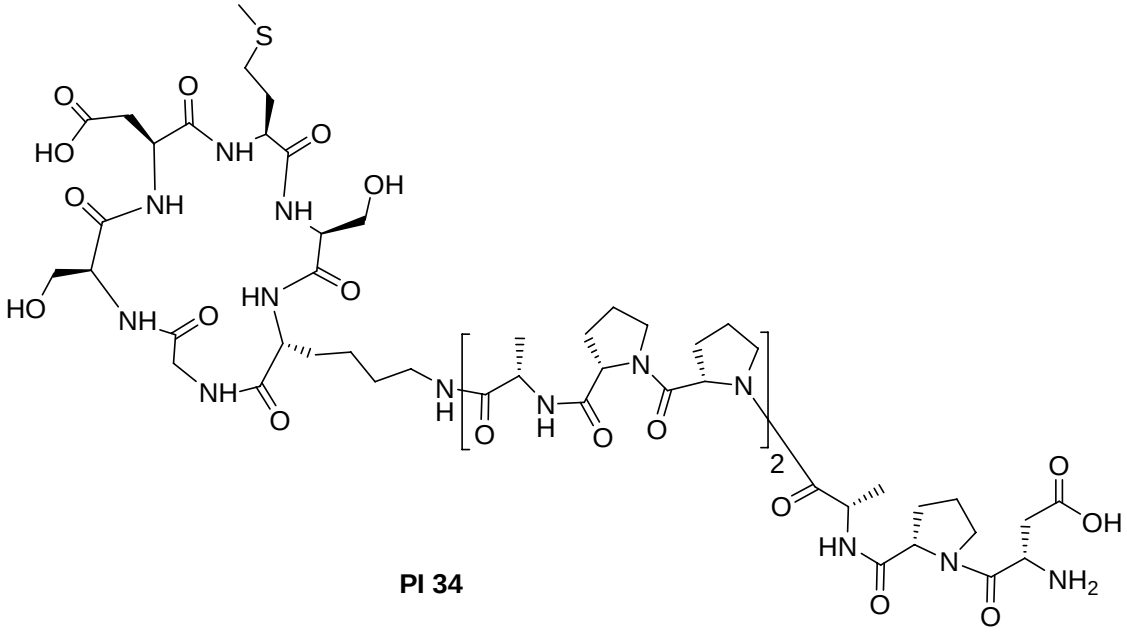
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1441.73 $[M+Na]^+$, 1457.63 $[M+K]^+$

Calculated Mass: $[C_{61}H_{94}N_{16}NaO_{21}S]^+$ = 1441.64, $[C_{61}H_{94}KN_{16}O_{21}S]^+$ = 1457.61

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 38.97 min

Analytical RP-HPLC: Method I, t_R = 17.55 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Asp-Pro-Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala)-Gly-) (PI 34)



PI 34

Yield: 7.9 %

Formula: $C_{66}H_{101}N_{17}O_{22}S$

Molecular Mass: 1515.70 g/mol

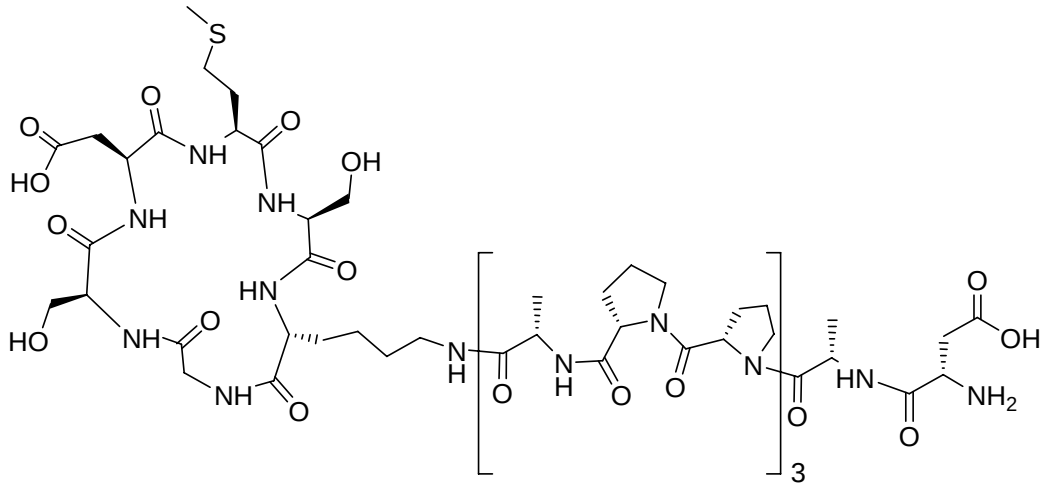
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1516.69 $[M+H]^+$, 1538.59 $[M+Na]^+$, 1554.56 $[M+K]^+$

Calculated Mass: $[C_{66}H_{102}N_{17}O_{22}S]^+$ = 1516.71, $[C_{66}H_{101}N_{17}NaO_{22}S]^+$ = 1538.69, $[C_{66}H_{101}KN_{17}O_{22}S]^+$ = 1554.67

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 39.25 min

Analytical RP-HPLC: Method I, t_R = 18.63 min, 100% (220 nm)

cyclo-(-Ser-Asp-Met-Ser-D-Lys(H-Asp-Ala-Pro-Pro-Ala-Pro-Pro-Ala-Pro-Pro-Ala)-Gly-) (PI 35)



PI 35

Yield: 9.8 %

Formula: $C_{69}H_{106}N_{18}O_{23}S$

Molecular Mass: 1586.74 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 1587.89 $[M+H]^+$, 1609.88 $[M+Na]^+$, 1625.90 $[M+K]^+$

Calculated Mass: $[C_{69}H_{107}N_{18}O_{23}S]^+$ = 1587.75, $[C_{69}H_{106}N_{18}NaO_{23}S]^+$ = 1609.73, $[C_{69}H_{106}KN_{18}O_{23}S]^+$ = 1625.70

Preparative RP-HPLC: Apparatus A, Method 6, t_R = 39.90 min

Analytical RP-HPLC: Method I, t_R = 19.11 min, 100% (220 nm)

2.6.3 Inhibition of the Interaction between Laminin-332 and Integrin $\alpha_3\beta_1$ by Synthetic Peptide/Peptidomimetic

1. Coating: 100 µl 6 µg/ml laminin-332 solutions are applied to every well of the microtiter plate for ELISA. The plate is incubated at 4 ^{o}C overnight and then washed 3 times with TBS/$MgCl_2$.
2. Blocking: The non-binding sites on the plate are endcapped with 1% BSA in TBS/$MgCl_2$ buffer at room temperature for 2 hours.
3. Integrin/Peptide Solution Preparation: The peptide samples are dissolved in TBS buffer to prepare a 20 mM solution. 248 µl 0.449 µg/ml integrin $\alpha_3\beta_1$ are dissolved in 142 µl 1 mg/ml monoclonal antibody 9EG7, 7.11 ml 1%BSA/TBS/$MgCl_2$ and 7.5 ml 1 M $MnCl_2$.

   Negative Control: 8.8 µl 0.5 M EDTA is subjected to 396 µl 55 nM integrin $\alpha_3\beta_1$ and 35.2 µl TBS/$MgCl_2$ buffer.

   Positive Control: 396 µl integrin $\alpha_3\beta_1$ is subjected to 44 µl TBS/$MgCl_2$ buffer.
4. Binding: 198 µl solution of integrin $\alpha_3\beta_1$ and 22 µl peptide solution are mixed. The final concentrations of integrin $\alpha_3\beta_1$ and peptide are 50 nM and 1.8 mM, respectively. 100 µl peptide/integrin $\alpha_3\beta_1$ mixtures are subjected to every well except positive and negative control wells. The processes of inhibition of the interaction between integrin $\alpha_3\beta_1$ and immobilized laminin-332 by synthetic peptides take 2 hours at room temperature before the solutions in every well are removed. The plate is then washed 3 times: inhibition wells with HEPES buffer; EDTA wells with 10 mM EDTA solution.
5. Fixation: 2.5% glutardialdehyde is subjected to 50 mM HEPES in 150 mM NaCl, 2 mM $MgCl_2$, 1 mM $MnCl_2$. The fixation takes approximately 10 min at room temperature. The fixation solution is then removed and the plate is washed 3 times with TBS/$MgCl_2$ buffer.
6. First Antibody: Anti β_1 is diluted by 1000 times with 1%BSA in TBS/$MgCl_2$ buffer. 100 µl anti β_1 solution are added to each well. The process takes 1.5 hour at room

temperature. The anti β_1 solution is then discarded and the plate is washed with TBS/$MgCl_2$ 3 times.

7. Second Antibody: Second antibody IgG coupled with enzyme alcaline phosphatase is diluted by 1000 times. 100 μl enzyme-linked IgG solution is added to each well. The process takes 1.5 hour at room temperature. The IgG solution is then discarded and the plate is washed with TBS/$MgCl_2$ for 3 times.
8. Substrate: 100 μl 1 mg/ml phosphylphenyl phosphate in 0.1 M glycine/NaOH containing 10 μl 1 M $Zn(OAc)_2$ and $MgCl_2$ (pH 10.2) is added to each well. It takes approximately 4-5 min before the color appears in the well.
9. Termination: 1.5 M NaOH solutions are added to each well in order to stop the color-generating reactions.
10. Quantification: The absorbance of the solution in each plate well is measured to determine the content of unbound integrin $\alpha_3\beta_1$.

2.7 Summary

In this project, integral membrane protein integrins were chosen as the research target. Integrin $\alpha_3\beta_1$, a member of the integrin family, is the receptor to several extracellular ligands such as fibronectin, invasin, and laminin. It presents a good target for the structure-based ligand design research, possessing both theoretical and applied significances. Design, screen, target and optimization of a integrin $\alpha_3\beta_1$ ligand is challenging because no crystal structure of integrin $\alpha_3\beta_1$ is yet available, and the research with regard to this integrin and its antagonists are not sufficiently in depth and lags behind that of other members of the integrin family. Invasin, whose X-ray single crystal structure is available, is one of the naturally occurring ligand of integrin $\alpha_3\beta_1$. It mediates the internalization of bacteria *Yersinia Sp.* into the mammalian cells through tight association with integrin $\alpha_3\beta_1$. Synthetic ligands which possess stronger affinities to integrin $\alpha_3\beta_1$ could serve as antagonists and thus block the adhesion of invasin to integrin $\alpha_3\beta_1$, which in turn prevents the uptake of *Yersinia*. On the basis of the similarities between invasin and other integrin $\alpha_3\beta_1$ ligands, especially fibronectin, the integrin $\alpha_3\beta_1$ binding domains on invasin had been assigned. Asp911 was considered to play an important role for the adhesion of invasin to integrin $\alpha_3\beta_1$. Furthermore, Asp811 and Arg883, which are on the same side of invasin as Asp911 and are approximately 32 Å (literature value) apart from Asp911, could assist the binding of invasin to integrin $\alpha_3\beta_1$; they are thus regarded as synergic region. There is a striking similarity between fibronectin and invasin, which could be regarded as the proof to locate the binding domain of invasin to integrin $\alpha_3\beta_1$. The crucial Asp911 was resided at a bulge contour on invasin, this factor is supposed to be taken into consideration as peptide/peptidomimetic ligand design is concerned.

Despite that -Ser-Asp-Met-Ser- sequence in the D5 domain in invasin is regarded as the binding moiety for the adhesion of invasin to integrin $\alpha_3\beta_1$, this conclusion was nevertheless drawn out of the similarities of X-ray crystal structures of invasin and

other natural ligands. This aspect was not validated by experimental results. One of the major purposes of this project is to confirm the SDMS recognition sequence by the synthetic peptide ligands containing this binding moiety. The primary sequence -Ser-Asp-Met-Ser- centered on the crucial Asp911 in invasin served as starting point in the ligand design process. Asp911 is located at *i*+1 position of a β-turn in the native sequence. This positional element should be kept constant in the synthetic peptides in order to obtain good binding affinity to integrin $\alpha_3\beta_1$. Cyclization of a peptide containing SDMS recognition sequence could lock the flexible macromolecule into a preferable conformation which leads to an improved bioactivity, this is achieved with the incorporation of special secondary structure inducing residues, such as D-amino acid, α,α-disubstituted amino acid, or *N*-alkylated amino acid. The idea of spatial screenings of peptides and peptidomimetics is an important concept, which applied the concerned peptides with diverse conformations in order to search the bioactive conformation of peptide ligand. Spatial screening was performed by the incorporation of different secondary structure inducers such as D-Lys, Aib, or Gly in this project to confirm the bioactive conformation. The cyclization could also evidently lower the flexibility of the concerned peptide, locking them in a favorable conformation as affinity to a receptor is concerned. This adavantage meets the criteria of the active ligand design.

In this project, e.g. a D-amino acid was incorporated into the cyclic hexapeptide as the secondary structure inducer. D-amino acids are known to preferentially occupy *i*+1 position of a β-turn in a cyclic hexapeptide, the crucial Asp was thus locked in *i*+1 position in a complementary β-turn. The parent peptide of this series, cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-), was successfully synthesized and taken as the reference for the screening and optimization. Residue mutations of this reference peptide were performed in addition to explore the mechanism of the adhesion of the peptide inhibitors to integrin $\alpha_3\beta_1$, optimize the synthetic ligand, and screen the lead compound. ELISA studies were carried out to analyze the affinities of the synthesized peptides to integrin $\alpha_3\beta_1$. Laminin-332 was applied as the natural ligand of

integrin $\alpha_3\beta_1$ in this project. It was confirmed that the Asp residue in the cyclic peptide plays a crucially important role for adhesion. The inhibitory capacity was almost totally lost upon its substitution by Ala. Methionine in the reference peptide was proved to be indispensable, as its substitution by Ala or Aib, as well as oxidation of the side chain to sulfoxide derivative led to a significant decrease in affinity. The serine preceeding aspartate seemed to tolerate modification, as its substitution by alanine resulted in even superior binding affinity. The lysine residue was assumed to be relevantly indifferent to mutation, as its replacement by alanine and arginine did not cause a substantial change of inhibitory capacities, thus validating the choice of this residue as the hinge of the peptide scaffold. However, it was found that peptide with L-lysine at this position displayed higher affinity than the D-lysine containing derivative. It could be probably caused by the type of the induced β-turn, the thus formed complementary β-turn, in which aspartate occupies *i*+1 position, may not match the type of the prototype, or due to the orientation of the side chain of D-lysine leading to a unfavorable interaction with the corresponding domain on the integrin $\alpha_3\beta_1$. For the L-lysine-containing peptide, glycine could also serve as the secondary structure inducer, as it is commonly regarded as "proteinogenic D-amino acid" because of its special conformational influence. The substitution of glycine by sarcosine led to an increased affinity. *N*-alkylated amino acids such as sarcosine (*N*-methyl glycine) preferentially occupy *i*+2 position of β-turn in a cyclic hexapeptide. This β-turn centered on lysine and sarcosine could stabilize the structure with the aspartate in the complementary β-turn. The combination of sarcosine and D-lysine led to one of the most efficient inhibitors in the peptide library. Cyclo-(-Ser-Asp-Met-Ser-Asp-Met-) with a palindrome sequence also exhibited a good affinity to integrin $\alpha_3\beta_1$. The combination of these favorable elements in a single cyclic peptide/peptidomimetic could be promising for an optimized second generation of integrin $\alpha_3\beta_1$ ligands.

As Asp811 and Arg883 in invasin are assumed to exert synergic influence on its association to integrin $\alpha_3\beta_1$, these binding moieties are supposed to be incorporated

into the peptide to achieve improved affinity. The peptide scaffold oligo β-alanine and $(Pro\text{-}Pro\text{-}Ala)_n$ was fused into the reference peptide in order to introduce these synergic binding element Asp811 or Arg883. Molecular modeling was applied to design the scaffold rationally. The structures adopted by the scaffold under physiological conditions were taken into consideration.

The introduction of synergic residues Asp811 and Arg883 through the peptide scaffold, however, did not generate the expected improved inhibitory capacities. This could be due to the increased entropy loss brought forward by the introduction of flexible peptide scaffold upon binding to integrin. On the other hand, the adhesion of the binding motif -Ser-Asp-Met-Ser- to its corresponding association domain on integrin could possibly lead to an induced conformational change of the protein receptor or peptide ligand, this dynamic process might result in a discrepancy between the ligand design and the actual situation. The third possible explanation is that the single amino acid Asp/Arg might not be specifically recognized by their binding domains on the protein.

In summary, synthesis of the peptide library was successfully completed through manual and microwave assisted SPPS; orthogonal deprotection was achieved by the Aloc protecting groups on the side chain of lysine; cyclization of linear protected peptides was carried out under high-pseudo-dilution condition implemented by two channel syringe pump, no dimerization was found in this project; segment condensation was smoothly and nearly quantatively converted; deprotection and subsequent purification was also fulfilled. The synthesized peptide library was applied to ELISA test as the inhibitors to the binding of laminin-332 to integrin $\alpha_3\beta_1$. Favorable elements for inhibition were screened. Mechanism of binding was preliminarily analyzed along with the screen process. The thus obtained lead compound and relative conclusions could be integrated and applied to the optimization and design of the next peptide inhibitor generation.

Optimization of the peptide complexes could be achieved if the properties of the linker were improved, namely, to increase its rigidity property upon the exact simulation with proper techniques of molecular modelling and meticulous design. The incorporation of the cooperative binding residue Asp/Arg should be more sophisticated with the binding moiety instead of single amino acid residue, in order to achieve the more positive enthalpy gain upon the binding to compensate the loss of the entropy loss. The optimization of template peptide cyclo-(-Ser-Asp-Met-Ser-D-Lys-Gly-) could also be achieved.

References

[1] Beddell, C. R., (Ed.), *The Design of Drugs to Macromolecular Targets.* J.Wiley, Chichester (1992).

[2] Peruz, M., *Protein Structure: New Approaches to Disease and Therapy.* W. H. Freeman (1992).

[3] Searle, M. S., Williams, D. H., *J. Am. Chem. Soc.,* (1992) **114**, 10690.

[4] Babine, R. E., Bender, S. L., *Chem. Rev*., (1997) **97**, 1359.

[5] Gohlke, H., Kliebe, G., *Angew. Chem., Int. Ed*., (2002) **41**, 2644.

[6] Dunitz, J. D., *Chem. Biol.*, (1995) **2**, 709.

[7] Gilli, P., Ferretti, V., Gilli, G., Brea, P. A., *J. Phys. Chem*., (1994) **98**, 1515.

[8] Williams, D. H., O'Brien, D. P., Bardsley, B., *J. Am. Chem. Soc.*, (2001) **123**, 737.

[9] Weber, P. C., Wendoloski, J. J., Pantoliano, M. W., Salemme, F. R., *J. Am. Chem. Soc.*, (1992) **114**, 3197.

[10] Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., Winter, G., *Nature*, (1985) **314**, 235.

[11] Chen, Y. W., Fersht, A. R., *J. Mol. Biol*., (1993) **234**, 1158.

[12] Connelly, P. R., Aldape, R. A., Bruzzesse, F. J., Chambers, S. P., Fitzgibbon, M. J., Fleming, M. A., Itoh, S., Livingston, D. J., Navia, M. A., Thomson, J. A., Wilson, K. P., *Proc. Natl. Acad. Sci.,* USA, (1994) **91**, 1964.

[13] Morgan, B. P., Scholtz, J. M., Ballinger, M. D., Zipkin, I. D., Bartlett, P. A., *J. Am. Chem. Soc*., (1991) **113**, 297.

[14] Shirley, B. A., Stanssens, P., Hahn, U., Pace, C. N., *Biochemistry*, (1992) **31**, 725.

[15] Obst, U., Banner, D. W., Weber, L., Diederich, F., *Chem. Biol.*, (1997) **4**, 287.

[16] McDonald, I. K., Thornton, J. M., *J. Mol. Biol.*, (1994) **238**, 777.

[17] Watson, K. A., Mitchell, E. P., Johnson, L. N., *Biochemistry*, (1994) **33**, 5745-5758.

[18] Sauter, N. K., Bednarski, M. D., Wurzburg, B. A., *Biochemistry*, (1989) **28**, 8388-8396.

[19] Janes, W., Schulz, G. E., *J. Biol. Chem.,* (1990) **256**, 10443-10445.

[20] Kubinyi, H., in *Pharmacokinetic optimization in drug research*, B. Testa, H. van de Waterbeemd, G. Folkers, R. Guy (Eds.), Wiley-VCH, Weinheim, (2001), pp 513.

[21] Schneider, H-J., Schiesteil, T., Zimmermann, P., *J. Am. Chem. Soc*., (1992) **114**, 7698.

[22] Tissot, A. C., Vuilleumier, S., Fersht, A. R., *Biochemistry*, (1996) **35**, 6786.

[23] McDonald, I. K., Thornton, J. M., *J. Mol. Bio.,* (1994) **238**, 777-793.

[24] Ben-Naim, A., *Hydrophobic Interactions*, Plenum, New York, 1980.

[25] Tanford, C., *The Hydrophobic Effect*, Wiley, New York, 1980.

[26] Richards, F. M., *Annu. Rev. Biophys. Bioeng*., (1977) **6**, 151.

[27] Sharp, K. A., Nicholls, A., Friedman, R., Honig, B., *Biochemistry*, (1991) **30**, 9696.

[28] Von Itzstein, M., Wu, W.-Y., Kok, G. B., Peggs, M. S., Dyason, J. C., Jin, B., Van, Phan T., Smythe, M. L., White, H. F., Oliver, S. W., Colman, P., Varghese, J. N., Ryan, D. M., Woods, J. M., Bethell. R. C., Hotham, V. J., Cameron, J. M., Penn, C. R., *Nature*, (1993) **363**, 418-423.

[29] (a) Brünger, A. T., *Acta. Crystallogr.*, (1998) **D54**, 905.
(b) Brünger, A. T., *Nature*, (1992) **355**, 472.
(c) Read, R. J., *Acta. Crystallogr.*, (1986) **A42**, 140.
(d) Hodel, A., Kim, S. H., Brünger, A. T., *Acta. Crystallogr.*, (1992) **A48**, 851.
(e) Laskowski, R.A., McArthur, M. W., Moss, D. S., Thornton, J. M., *J. Appl. Crystallogr.*, (1993) **26**, 283.
(f) Kraulis, P. J., *J. Appl. Crystallogr.*, (1991) **24**, 946.
(g) Merritt, E. A., Murphy, M. E. P., *Acta. Crystallogr.*, (1994) **D50**, 869.
(h) Nicholls, A., Bharadwaj, R., Honig, B., *Biophys. J.*, (1993) **64**, A166
(g) Humphries, M. J., *Biochem. Soc. Trans.*, (2000) **28**, 311-339.

[30] (a) Protein structures: Fn-III 7-10 [Protein Data Bank (PDB) code 1FNF] [Leahy, D. J., Aukhil, I., Erickson, H. P., *Cell*, (1996) **84**, 155.]
(b) Fn-III 12-14 (PDB code 1FNH) [Sharma, A., Askari, J. A., Humphries, M. J., Jones, E. Y., Stuart, D. I., *EMBO J.*, (1999) **18**, 1468.];
(c) intimin (coordinates obtained from S. Matthews) [Kelly, G., *Nature Struct. Biol.*, (1999) **6**, 313.]

(d) mannose-binding protein (PDB code 1RTM) [Weis, W. I., Drickamer, K., *Structure*, (1994) **2**, 1227.]
(e) E-selectin (PDB code 1ESL) [Graves, B. J., *Nature*, (1994) **367**, 532.]
(f) CD94 (coordinates obtained from P. D. Sun) [Boyington, J. C., *Immunity*, 170 (1999) **10**, 75.]
(g) VCAM-1 (PDB code 1VSC) [Jones, E. Y., *Nature*, (1995) **373**, 539; Wang, J., *Proc. Natl. Acad. Sci. U.S.A.*, (1995) **92**, 5714.]
(h) ICAM-1 (PDB code 1IC1) [Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M., Rossmann, M. G., *Proc. Natl. Acad. Sci. U.S.A.*, (1998) **95**, 4140.]

[31] (a) Soltis, S. M.; Stowell, M. H. B., Wiener, M. C., Phillips, G. N., Rees, D. C., *J.Appl. Crystallogr.*, **30**, 190 (1997).
(b) Otwinowski, Z., Minor, W., *Methods Enzymol.*, (1997) **276**, 307.
(c) Fortelle, D. L., Bricogne, G., *Methods Enzymol.*, (1997) **276**, 472.
(d) McRee, D. E., *Practical Protein Crystallography*, Academic Press, San Diego, CA (1993).
(e) Abrahams, J. P., Leslie, A. G. W., *Acta. Crystallogr.*, (1996) **D52**, 30.
(f) Kleywegt, G. J., Jones, T. A., *Acta. Crystallogr.*, (1997) **D52**, 826.

[32] Kate, S. A., Solé, N. A., Albericio, F., Barany, G. in *Peptides: Design, Synthesis, and Biological Activity*, Basava, C., Anantharamaiah, G. M., (Eds.), Birkhäuser, Boston, (1994) 39.

[33] Goodman, S. L., Hölzemann, G., Sulyok, G-A., G., Kessler, H., *J. Med. Chem.,* (2002) **45**, 1045-1051.

[34] Loges, S., Butzal, M., Otten, J., Schweizer, M., Fishcer, U., Bokemeyer, C., Hossfeld, D. K., Schuch, G., Fiedler, W., *Biochim. Biophy. Res. Comm.,* (2007) **357**, 1016-1020.

[35] Brooks, P. C., Montgomery, A. M., Rosenfeld, M., Reisfeld, R. A., Hu, T., Klier, G., Cheresh, D. A., *Cell*, (1994) **79**, 1157–1164.

[36] Brooks, P. C., Stromblad, S., Klemke, R., Visscher, D., Sarkar, F. H., Cheresh, D. A., *J. Clin. Investig*., (1995) **96**, 1815–1822.

[37] Brooks, P. C., Clark, R. A., Cheresh, D. A., *Science*, (1994) **264**, 569–571.

[38] Qiao, R., Yan, W., Lum, H., Malik, A. B., *Am. J. Physiol.*, (1995) **269**, 110–117.

[39] Mitjans, F., Meyer, T., Fittschen, C., Goodman, S. L., Jonczyk, A., Marshall, J. F., Reyes, G., Piulats, J., *Int. J. Cancer*, (2000) **87**, 716–723.

[40] Allman, R., Cowburn, P., Mason, M., *Eur. J. Cancer*, (2000) **36**, 410–422.

[41] Haubner, R., Finsinger, D., Kessler, H., *Angew. Chem.*, (1997) **109**, 1440 – 1456; *Angew. Chem. Int. Ed. Engl.*, (1997) **36**, 1374 – 1389.

[42] Schumann, F., Müller, A., Koksch, M., Müller, G., Sewald, N., *J. Am. Chem. Soc.*, (2000) **122**, 12009 – 12010.

[43] Bubert, C., Cabrele, C., Reiser, O., *Synlett*, (1997) 827 – 829.

[44] Beumer, R., Bubert, C., Cabrele, C., Vielhauer, O., Pietzsch, M., Reiser, O., *J. Org. Chem.*, (2000) **65**, 8960 – 8969.

[45] Beumer, R., Reiser, O., *Tetrahedron* (2001) **57**, 6497 – 6503.

[46] Zorn, C., Gnad, F., Salmen, S., Herpin, T., Reiser, O., *Tetrahedron Lett.*, (2001) **42**, 7049 – 7053.

[47] Zimmermann, D., Guthöhrlein, E., Malešević, M., Sewald, K., Wobbe, L., Heggemann, C., Sewald, N., *Chem. Bio. Chem.,* (2005) **6**, 272 –276.

[48] Urman, S., Gaus, K., Yang, Y., Strijowski, U., Sewald, N., De Pol, S., Reiser, O., *Angew. Chem.*, (2007) **119**, 4050-4053; *Angew. Chem. Int. Ed. Engl.*, (2007) **46**, 3976-3978.

[49] Ryan, K. J., Ray, C. G., *Sherris Medical Microbiology*, 4th ed., (2004) McGraw Hill, 368–70.

[50] Abe, J., Onimaru, M., Matsumoto, S., Noma, S., Baba, K., Ito, Y., Kohsaka, T., Takeda, T., *J. Clin. Invest*., (1997) **99**, 1823-1830.

[51] Abe, J., Takeda, T., *Prep Biochem Biotechnol.*, (1997) **27**, 173-208.

[52] Amromin, I., Chapnick, E. K., Morozov, V. G., *Infect. Dis. Clin. Pract.,* (2000) **9**, 236-240.

[53] Butler, T., *Yersinia* species, including plague. In: Mandell, G. L.; Bennett, J. E.; Dolin, R., eds. Mandell, Douglas and Bennett's Principles and Practice of Infectious Diseases. 4th ed. NY: Churchill Livingstone; 1999, 2406-2414.

[54] Jani, A., *Pseudotuberculosis* (Yersina) (2003).

[55] Collins, F. M., (1996). *Pasteurella, Yersinia, and Francisella.* In *Barron's Medical Microbiology (Barron S et al, eds.), 4th ed., Univ of Texas Medical Branch.*

[56] Marie, A., Tuohy, M., O'Gorman**,** M., Byington**,** C., Reid**,** B., Jackson, W. D., *Pedlatrics,* (1999) **104**, 36.

[57] Benvenga, S., Santarpia, L., Trimarchi, F., Guarneri, F., *Thyroid*, (2006) **16**, 225-236.

[58] Tomer, Y., Davies, T., *Endocr. Rev.*, (1993) **14**, 107-120.

[59] Toivanen, P., Toivanen, A., *Int. Arch. Allergy. Immunol.*, (1994) **104**, 107-111.

[60] Strieder, T., Wenzel, B., Prummel, M., Tijssen, J., Wiersinga, W., *Clin. Exp. Immunol.*, (2003) **132**, 278-282.

[61] Salyers, A. A., Whitt, D. D., *Bacterial Pathogenesis: A Molecular Approach*, 2nd ed., (2002) ASM Press, pp 207-212.

[62] Wagle, P. M., *Indian J. Med. Sci.*, (1948) **2**, 489–494.

[63] Meyer, K. F., *JAMA*, (1950) **144**, 982–985.

[64] Kilonzo, B. S., Makundi, R. H., Mbise, T. J., *Acta. Tropica.*, (1992) **50**, 323–329.

[65] Mwengee, W., Butler, T.; Mgema, S., *Clin. Infect. Dis.*, (2006) **42**, 614–621.

[66] Lee, K. K., Doig, P., Irvine, R. T., Paranchych, W., Hodges, R. S., *Mol, Microbiol.*, (1989) **3**, 1493–1499.

[67] Swanson, J., *J. Exp. Med.*, (1973) **137**, 571–589.

[68] Lund, B., Marklund, B. I., Stromberg, N., Lindberg, F., Karlsson, K. A., Normark, S., *Mol. Microbiol.*, (1988) **2**, 255–263.

[69] Bliska, J., Galan, J., Falkow, S., *Cell*, (1993) **73**, 903–920.

[70] Isberg, R. R. *Science*, (1991) **252**, 934–938.

[71] Moulder, J. W., *Microbiol. Rev.*, (1985) **49**, 298–337.

[72] Falkow, S., Isberg, R. R., Portnoy, D., *Annu. Rev. Cell Biol.*, (1992) **8**, 333–363.

[73] Takeuchi, A. *Am. J. Pathol.*, (1967) **50**, 109–136.

[74] Sansonetti, P. J., Clerc, P., Maurelli, A. T., Mournier, J., *Infect. Immun.*, (1986) **51**, 461–469.

[75] Isberg, R. R., Leong, J. M., *Cell*, (1990) **60**, 861-871.

[76] Grutzkau, A., Hanski, C., Hahan, H., Riecken, E., *Gut.*, (1990) **31**, 1011–1015.

[77] Hanski, C., Kutschka, U., Schmoranzer, H., Naumann, M., Stallmach, A., Hahn, H., Menge, H., Riecken, E., *Infect. Immun.*, (1989) **57**, 673–678.

[78] Pepe, J. C., Miller, V. L., *Proc. Natl. Acad. Sci. U. S. A.*, (1993) **90**, 6473–6477.

[79] Hamburger, Z. A., Brown, M. S., Isberg, R. R., Bjorkman, P. J., *Science*, (1999) **286**, 291-295.

[80] (a) Pepe, J. C., V. L. Miller, V. L., *Infect. Agents Dis.*, (1993) **2**, 236.
(b) Marra, A., Isberg, R. R., *Infect. Immun.*, (1997) **65**, 3412.
(c) Autenrieth, I. B., Firsching, R., *J. Med. Microbiol.*, (1996) **44**, 285.

[81] Isberg, R. R., *Mol. Biol. Med.*, (1990) **7**, 73–82.

[82] Isberg, R. R., Falkow, S., *Nature*, (1985) **317,** 262–264.

[83] Jerse, A. E., Yu, J., Tall, B. D., Kaper, J. B., *Proc. Natl. Acad. Sci. U. S. A.*, (1990) **87**, 7839–7843.

[84] Donnenberg, M. S., Calderwood, S. B., Donohue, R. A., Keusch, G. T., Kaper, J. B., *Infect. Immun.*, (1990) **58**, 1565–1571.

[85] (a) Isberg, R. R., Leong, J. M., *Cell,* (1990) **60**, 861.
(b) Isberg, R. R., G. Tran Van Nhieu, *Trends Microbiol.*, (1994) **2**, 10.

[86] Hynes, R. O., *Cell*, (1992) **69**, 11.

[87] (a) Tran Van Nhieu, G., Krukonis, E. S., Reszka, A. A., Horwitz, A. F., Isberg, R. R., *J Biol. Chem.*, (1996) **271**, 7665.
(b) Tran Van Nhieu, G., Isberg, R. R., *EMBO J.*, (1993) **12**, 1887.
[88] (a) Liu, H., Magoun, L., Leong, J. M., *Infect. Immun.*, (1999) **67**, 2045.
(b) Kenny, B., *Cell*, (1997) **91**, 511.

[89] (a) Leong, J. M., Fournier, R. S., Isberg, R. R., *EMBO J.*, (1990) **9**, 1979.
(b)Worley, M. J., Stojiljkovic, I., Heffron, F., *Mol. Microbiol.*, (1998) **29**, 1471.
(c) Diederichs, K., Freigang, J., Umhau, S., Zeth, K., Breed, J., *Protein Sci.*, (1998) **7**, 2413.
(d) Jap, B. K., Walian, P. J., *Physiol. Rev.*, (1996) **76**, 1073.
(e) Locher P. K., *Cell*, (1998) **95**, 771.
(f) Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K., Welte, W., *Science* (1998) **282**, 2215.

[90] (a) Yu, J., Kaber, J. B., *Mol. Microbiol.*, (1992) **6**, 411.
(b) Chothia, C., Lesk, A. M., *EMBO J.*, (1986) **5**, 823.
(c) Abagyan, R. A., Batalov, S., *J. Mol. Biol.*, (1997) **273**, 355.

[91] (a) Leong, J. M., Morrissey, P.E., Marra, A., Isberg, R. R., *EMBO J.*, (1995) **14**, 422.
(b) Hendrickson, W. A., Horton, J. R., LeMaster, D. M., *EMBO J.*, (1990) **9**, 1665.

[92] (a) Tran Van Nhieu, G., Isberg, R. R., *J. Biol. Chem.*, (1991) **266**, 24367.
(b) Krukonis, E. S., Dersch, P., Eble, J. A., Isberg, R. R., *J. Biol. Chem.*, (1998) **273**, 31837-43.
(c) Eble, J. A., *Biochemistry*, (1998) **37**, 10945.
(d) Takada, Y., Ylanne, J., Mandelman, D., Puzon, W., Ginsberg, M. H., *J. Cell Biol.*, (1992) **119**, 913.

[93] (a) Dersch, P., Isberg, R. R., *EMBO J.*, (1999) **18**, 1199.
(b) Su, X. D., *Science*, (1998) **281**, 991.

[94] (a) Soltis, S. M., Stowell, M. H. B., Wiener, M. C., Phillips, G. N., Rees, D. C., *J. Appl. Crystallogr.*, **30**, 190 (1997).
(b) Otwinowski, Z., Minor, W., *Methods Enzymol.*, (1997) **276**, 307.
(c) Fortelle, D. L., Bricogne, G., *Methods Enzymol.*, (1997) **276**, 472.
(d) McRee, D. E., *Practical Protein Crystallography*, Academic Press, San Diego, CA, (1993).
(e) Abrahams, J. P., Leslie, A. G. W., *Acta. Crystallogr.*, (1996) **D52**, 30.
(f) Kleywegt, G. J., Jones, T. A., *Acta. Crystallogr.*, (1997) **D52**, 826.

[95] (a) Brünger, A. T., *Acta. Crystallogr.*, (1998) **D54**, 905.
(b) Brünger, A. T., *Nature*, (1992) **355**, 472.
(c) Read, R. J., *Acta. Crystallogr.*, (1986) **A42**, 140.
(d) Hodel, A., Kim, S. H., Brünger, A. T., *Acta. Crystallogr.*, (1992) **A48**, 851.
(e) Laskowski, R. A., McArthur, M. W., Moss, D. S., Thornton, J. M., *J. Appl. Crystallogr.*, (1993) **26**, 283.
(f) Kraulis, P. J., *J. Appl. Crystallogr.*, (1991) **24**, 946.
(g) Merritt, E, A,, Murphy, M. E. P., *Acta. Crystallogr.*, (1994) **D50**, 869.
(h) Nicholls, A., Bharadwaj, R., Honig, B., *Biophys. J.*, (1993) **64**, A166
(g) Humphries, M. J., *Biochem. Soc. Trans.*, (2000) **28**, 311-339.

[96] (a) Protein structures: Fn-III 7-10 [Protein Data Bank (PDB) code 1FNF] [Leahy, D. J., Aukhil, I., Erickson, H. P., *Cell*, (1996) **84**, 155.]
(b) Fn-III 12-14 (PDB code 1FNH) [Sharma, A., Askari, J. A., Humphries, M. J., Jones, E. Y., Stuart, D. I., *EMBO J.*, (1999) **18**, 1468.];
(c) intimin (coordinates obtained from S. Matthews) [Kelly, G., *Nature Struct. Biol.*, (1999) **6**, 313.]
(d) mannose-binding protein (PDB code 1RTM) [Weis, W. I., Drickamer, K., *Structure*, (1994) **2**, 1227.]
(e) E-selectin (PDB code 1ESL) [Graves, B. J., *Nature*, (1994) **367**, 532.]
(f) CD94 (coordinates obtained from P. D. Sun) [Boyington, J. C., *Immunity*, (1999) **10**, 75.]
(g) VCAM-1 (PDB code 1VSC) [Jones, E. Y., *Nature*, (1995) **373**, 539; Wang, J., *Proc. Natl. Acad. Sci. U.S.A.*, (1995) **92**, 5714.]
(h) ICAM-1 (PDB code 1IC1) [Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M., Rossmann, M. G., *Proc. Natl. Acad. Sci. U.S.A.*, (1998) **95**, 4140.]

[97] Harpaz, Y., Chothia, C., *J. Mol. Biol.*, (1994) **238**, 528.

[98] (a) Leong, J. M., Morrissey, P. E., Isberg, R. R., *J. Biol. Chem.* (1993) **268**, 20524.
(b) Saltman, L. H., Lu, Y., Zaharias, E. M., Isberg, R. R., *J. Biol. Chem.*, (1996) **271**, 23438.

[99] (a) Bowditch, R. D., *J. Biol. Chem.*, (1994) **269**, 10856.
(b) Ugarova, T. P., *Biochemistry*, (1995) **34**, 4457.

[100] Richard B. Silverman, *The Organic Chemistry of Drug Design and Drug Action.* Second Edition, Elsevier Aademic Press, pp 241.

[101] Appelt, K., Bacquet, R. J., Bartlett, C. A., Booth, C. L. J., Freer, S. T. Fuhry, M. A. M., Gehring, M. R., Herrmann, S. M., Howland, E. F., Janson, C. A., Jones, T. R., Kan, C.-C., Kathardekar, V., Lewis, K. K., Marzoni, G. P., Matthews, D. A., Mohr, C., Moomaw, E. W., Morse, C. A., Oatley, S. J., Ogden, R.C., Reddy, M. R., Reich, S. H., Schoettlin, W. S., Smith, W. W., Varney, M. D., Villafranca, J. E., Ward, R. W., Webber, S., Webber, S. E., Welsh, K. M., White, J. *J. Med. Chem.*, (1991) **34**, 1925-1934.

[102] Greer, J., Erickson, J. W., Baldwin, J. J., Varney, M. D. *J. Med. Chem.*, (1994) **37**, 1035-1054.

[103] Bartfai, T., Langel, U., Bedecs, K., Andell, S., Land, T., Gregeresen, S., Ahren, B., Girotti, P., Consolo, S., Corwin, R. et al., Proc. *Natl. Acad. Sci.* USA, (1993) **90**, 11287-11291.

[104] Toniolo, C., Crisma, M., Formaggio, F., Valle, G., Cavicchioni, G., Precigoux, G., Aubby, A., Kamphuis, G., *Biopolymers*, (1993) **33**, 1061-1072.

[105] Wünsch, E., In *Houben Weyl*, (1974); Vol. 15/I, pp 728.

[106] Bodanszky, M., Martinez, J., In *The Peptides, Anaylsis, Structure, Biology,* Gross, E.; Meienhofer, J., Eds.; Academic: New York, (1983), Vol. 5, pp 111.

[107] Barany, G., Merrifeld, R. B., In *The Peptides, Anaylsis, Structure, Biology,* Gross, E.; Meienhofer, J., Eds.; Academic: New York, (1980), Vol. 2, p 217.

[108] Iselin, B., *Helv. Chim. Acta*, (1961) **44**, 61.

[109] Dale, J., *Angew. Chemie*, (1966) **78**, 1078-1081.

[110] Kessler, H., Haase, B. I., *J. peptide Protein Res.*, (1992) **39**, 3840.

[111] Koppfe, K.D., *J. Pharm. Sciences***,** (1972) **61**, 1345-1358**.**

[112] Izumiya, N., Kato, T., Waki, M., *Biopolymers*, (1961) **20**, 1785-1791.

[113] Brady, S. F., Paleveda, W. J., Arison, B. H., Freidinger, R. M., Nutt, R. F., Veber, D. F., In: Hruby, V.J.; Rich, D.H. (Eds.) *Peptides, Structure and Function*. Pearce Chemical Company, Rockford (1963), pp. 127-130.

[114] Schmidt. R., Neubert, K., *Int. J. peptide Protein Res.*, (1991) **37**, 502-507.

[115] Knorr, R., Trzeciak, A., Bannwarth, W., Gillessen, D., *Tetrahedron Lett*. (1969) **30**, 1927.

[116] Castro, B., Dormoy, J. R., Evin, G., Selve, C.. *Tetrehedron Lett.* (1976) **14**, 1219.

[117] Zimmer, S., Hoffmann, E., Jung, G., Kessler, H. 22nd European *Peptide Symposium,* lnterlaken 13.-l 9. September 1992, Abstract P 114.

[118] Felix, A. M., Wang, Ch.-T., Heimer E. P., Foumier, A., *Int. J. Peptide Protein Res.*, (1966) **31**, 231-238.

[119] Benoiton, N. L., Lee, Y. C., Steinauer, R., Chen, F. M. F., *Int J. Peptide Protein Res.* (1992) **40**, 559588.

[120] Malesevic, M., Strijowski, U., Bächle, D., Sewald, N., *J. Biotechnol.*, (2004) **112**, 73–77.

[121] Minami, I., Ohashi, Y., Shimizu, I., Tsuji, J., *Tetrahedron Lett.*, (1985) **26**, 2449.

[122] Merzouk, A., Guibé, F., Loffet, A., *Tetrahedron Lett.*, (1992) **33**, 477.

[123] Guibé, F., *Tetrahedron*, (1998) **54**, 2967.

[124] Tsuji, J., Mandai, T., *Synthesis*, (1996), 1.

[125] Dangles, O., Guibé, F., Balavoine, G., Lavielle, S., Marquet, A., *J. Org. Chem*., (1987) **52**, 4984.

[126] Dessolin, M., Guillerez, M.-G., Thieriet, N., Guibé, F., Loffet, A., *Tetrahedron Lett.,* (1995) **26**, 5741.

[127] Beugelmans, R., Bourdet, S., Bigot, A., Zhu, J., *Tetrahedron Lett*., (1994) **35**, 4349.

[128] Beugelmans, R., Neuville, L., Bois-Choussy, M., Chastanet, J., Zhu, J., *Tetrahedron Lett.*, (1995) **36**, 3219

[129] Gomez-Martinez, P., Dessolin, M., Guibé, F., Albericio, F., *J. Chem. Soc., Perkin Trans. 1*. (1999), 2871.

[130] Thieriet, N., Alsina, J., Giralt, E., Guibé, F., Albericio, F., *Tetrahedron Lett.*, (1997) **38**, 7275.

[131] Grove, D. E., *Platinum Metals Rev.,* (2003) **47**, 44.

[132] Belkin, A. M., Stepp, M. A., *Microsc. Res. Tech.,* (2000) **51**, 208-301.

[133] Eble, J. A., Wucherpfennig, K. W., Gauthier, L., Dersch, P., Krukonis, E., Isberg, R. R., Hemler, M. E., *Biochemistry,* (1998) **37**, 10945-10955.

[134] Delwel, G. O., Sonnenberg, A., Laminin isoforms and their integrin receptor, in: Horton, M. A. (Ed.), Adhesion Receptors as Therapeutic Targets, CRC Press, Boca Raton, (1996) pp. 9-36.

[135] Nishiuchi, R., Takagi, J., Hayashi, M., Ido, H., Yagi, Y., Sanzen, N., Tsuji, T., Yamada, M., Sekiguchi, K., *Matrix Biol.,* (2006) **25**, 189-197.

[136] Fukushima, Y., Ohnishi, T., Arita, N., Hayakawa, T., Sekiguchi, K., *J. Cancer*, (1998) **76**, 63-72.

[137] Gehlsen, K. R., Sriramarao, P., Furcht, L. T., Skubitz, A. P. N., *J. Cell Biol.,* (1992) **117**, 449-459.

[138] Silverman, R. B., *The Organic Chemistry of Drug Design and Drug Action, 2nd Ed.,* Elsevier Academic Press, pp 127.

[139] Chou, P., Y., Fasman, G., *J. Mol. Biol.* (1977) **115**, 135-175.

[140] Smith, J. A., Pease, L. G., *CRC Crit. Rev. Biochem.* (1980) **8**, 315.

[141] Crisma, M., Formaggio, F., Moretto, A., Toniolo, C., *Biopolymers*, (2006) **84**, 3-12.

[142] Benedetti, E., Bavoso, A., Di Blasio, B., Pavone, V., Pedone, C., Toniolo, C., Bonora, G. M., *Biopolymers*, (1983) **22**, 305-317.

[143] Schuler, B., Lipman, E, A., Steinbach, P. J., Kumke, M., Eaton, W. A., *Proc. Natl. Aca. Sci. USA*, (2005) **102**, 2754-2759.

[144] Toniolo, C., Crisma, M., Formaggio, F., Peggion, C., *Biopolymers*, (2001) **60**, 396-419.

[145] Saviano, M., Improta, R., Benedetti, E., Carrozzini, B., Cascarano, G. L., Didierjean, C., Toniolo, C., Crisma, M., *ChemBioChem.*, (2004) **5**, 541-544.

[146] Dunitz, J. D., *Chem. Biol*., (1995) **2**, 709.

3 Peptides in Biomineralizations

3.1 Biomineralization

Scientists and engineers have long been fascinated by the beautiful structures and functional properties of the materials formed within living organisms, especially those finely and orderly organized hard tissues of organisms such as bones, teeth, mollusc shells, which are composed of diverse minerals that are typically in close association with an organic polymeric phase, and thus are biocomposites. The nature is presenting us beautiful templates as to how these minerals selectively precipitate and subsequently orderly grow in the organism. Calcium for example, exists abundantly in the blood of organisms, while they choose bones or teeth as locations to precipitate and be precisely assembled instead of in the organs like muscle or skin. Even in different hard tissues, the pattern of the nucleation of calcium differentiates markedly from each other, leading to diverse forms of tissues and structures.

It seems that there are diverse batons behind the scene, directing and controlling the growth pattern of minerals in living organism. Sophisticated strategies have been developed by organism to direct and control the growth pattern of inorganic constituents such as calcium in their mineralized tissues as the target. Active and effective control mechanisms take place at almost all levels of structural hierarchy, ranging from the nanoscopic regime – the nucleation of a crystallite at a specific site – up to the macroscopic regime, where the biophysical properties of the mineralized tissue have to be matched to certain functions.

The mineral crystals formed by organisms, so-called biominerals, normally have shapes that are very different from the crystals produced inorganically outside the living organism. Molluscs are among the most thoroughly investigated organisms in this regard, which build concrete shells from $CaCO_3$.[1] The mollusc shell may be regarded as a microlaminate composite consisting of layers of highly oriented $CaCO_3$ crystals which are interspersed with thin sheets of an organic matrix. Crystals within separate shell

layers usually consist of either pure aragonite or pure calcite. Vaterite, when present, is usually associated with shell repair. Shell formation occurs in two principal phases. The first involves the cellular processes of ion transport and organic matrix synthesis which occur in different compartments of the molluscan mineralizing system. The second phase comprises a series of crystal nucleation and growth processes taking place in a specialized mineralization compartment, the so-called *extrapallial space* (see *Figure 3.1*).[2]

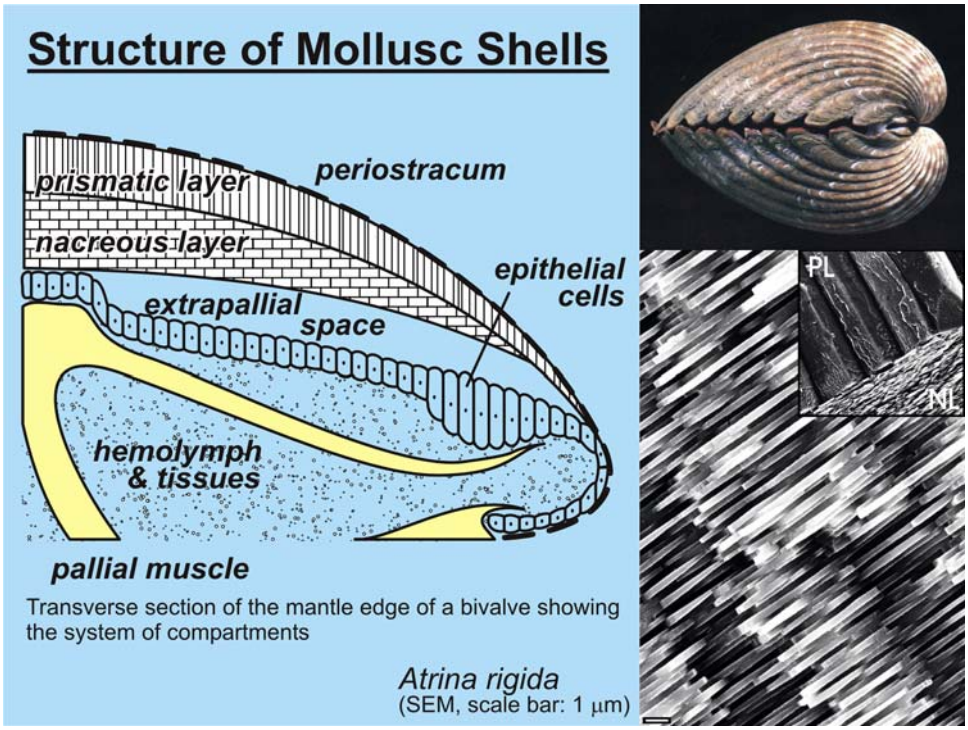


Figure 3.1 Transverse section of the mantle edge of a bivalve showing the system of compartments.[2]

The actual biological processes that direct and control the formation of a complete mineralized mollusc shell are largely unknown. However, the widely accepted consensus is that the crystal nucleation and growth events are strictly regulated by a number of highly specialized organic macromolecules. Control is believed to be exerted through specialized proteins that recognize specific crystal surfaces during the growth of the crystals. Recognition is based on molecular complementarity between the protein and the crystal structure on defined planes. The understanding of these processes is also relevant

to research on advanced materials. Biology provides an insight into unconventional strategies of a degree of sophistication yet unconceivable in artificial materials. Unfortunately, a deeper understanding of the biomineralization processes at the molecular level of structural hierarchy is hampered by the fact that our knowledge of the three-dimensional structures of macromolecules which are directly associated with the mineral layer is very poor. The conformation of these natural macromolecular templates and their properties play a crucially important role in deciding the manners of their interactions with the minerals and subsequently the assembly pattern of these crystalline biominerals.

The mimic of the naturally existing macromolecular templates that direct the assembly of the biominerals is supposed to be a suitable and promising tactic to gain a deeper understanding of this biological process. Biomimetics are defined as microstructural processing techniques that can either mimic or be inspired by biological processes. Seemingly, it would be difficult for material engineers to mimic complex cellular processes, however, material chemistry aspects of biomineralization can be studied by model systems, and utilized for biomimetic engineering. One particular aspect of interest to material chemists is the means by which these organisms use macromolecular organic constituents to mediate the growth of the mineral phase. For example, macromolecular templates are used to direct the nucleation event, manage vesicular compartments to delineate particle size and shape, and order solubilized proteins to regulate the kinetics of crystal nucleation and growth. In recent years, scientists have utilized some of these concepts to produce novel materials.

The secondary and tertiary structures of macromolecules directly involved in mineralization have been so far relatively insufficiently determined. The available information about the primary structures of peptides associated with this process is listed in Table 3.1 and 3.2. Macromolecules that were isolated from mollusc tissues have been traditionally categorized into two different classes, based on their water solubility properties. Chemical analysis showed that the water insoluble fraction mainly consists of fibrous proteins (collagen, keratin) and/or polysaccharides. These macromolecules together build a rigid framework upon which specific macromolecules from the soluble

fraction may be precisely adsorbed and assembled. The surface of this macromolecular assembly may serve as a supramolecular template for oriented nucleation of single crystals, and in fact crystallization experiments employing reconstructed matrices of purified mollusc shell macromolecules have shown that it is possible to switch between different $CaCO_3$ polymorphs[3] and to rebuild *in vitro* the gross structural features of the nacreous layer, respectively.[4]

Table 3.1 Summary of water-insoluble proteins isolated from mollusc shells

Name	Source	Characteristic sequence moif	Associated mineral	Proposed function	Ref.
MSI60	Pearl oyster protein from the nacreous layer	[$Ala_{9–13}$] and [$Gly_{3–15}$]	Aragonite	Framework protein, binding of Asp-rich sol. glycoproteins	[5]
MSI31	Pearl oyster protein from the prismatic layer	[$Gly_{3–5}$] and [Glu–Ser–Glu–Glu–Asp–X], (X = Thr or Met)	Calcite	Framework protein, binding of Asp-rich sol. glycoproteins	[5]
MSI7	Pearl oyster protein from the epithelia of the mantle	[Gly_{x-y}]	Aragonite Calcite	Framework protein, acceleration of nucleation and precipitation of CaCO3	[6]
N14, N66	Pearl oyster protein from the nacreous layer	[Asn–Gly]$_{12}$ and [Asn–Gly]$_{57}$	Aragonite	Carbonic anhydrase (N66)	[7]
N16	Pearl oyster protein from the nacreous layer	[Asn–Gly]$_6$ (as well as 4 acidic domains)	Aragonite	Control of crystal growth and morphology	[4]
Lustrin A	Abalone protein from the nacreous layer	[Gly–Ser–Ser–Ser] and [Gly–Ser] (as well as 1 basic domain)	Aragonite	Adhesion protein	[8]
Prismalin-14	Pearl oyster protein from the prismatic layer	[Asp]-and [Gly/Tyr]-rich domains, [Pro-Ile-Tyr-Arg] repeats	Calcite	Framework protein	[9]

Table 3.2 Summary of water-soluble proteins isolated from mollusc shells

Name	Source	Characteristic sequence moif	Associated mineral	Proposed function	Ref.
MSP-1	Scalop shell glycoprotein from the folia ted layer	[Asp–Gly–Ser–Asp] and [Asp–Ser–Asp]	Calcite	Induction of crystal nucleation, control of $CaCO_3$ polymorphism	[10]
Nacrein	Pearl oyster protein from the nacreous layer	[Gly–X–Asn] (X = Glu, Asn, or Asp)	Aragonite	Carbonic anhydrase, Ca-binding	[11]
Mucoperlin	Fan mussel protein from the nacreous layer	[Asp–X–Ser–Asp–X–Asp–X –Asp] (X =Val, Arg, Lys)	Aragonite	Induction of crystal nucleation, control of $CaCO_3$ polymorphism	[12]
Perlucin	Abalone protein from the nacreous layer	C-type lectin domains	Aragonite	Glycoprotein receptor	[13]
Perlustrin	Abalone protein	[Cys–X–Cys–Cys–X–X–Cys]	Aragonite	Insulin-like growth factor binding protein	[14]
Aspein	Pearl oyster highly acidic protein from the prismatic layer	[Asp_{2-10}] punctuated with [Ser–Gly] dipeptides	Calcite	Control of $CaCO_3$ polymorphism	[15]
AP7 *AP24*	Abalone protein from the nacreous layer	[Asp-Asp] [Asp-Asp-Asp-Glu-Asp]	Aragonite	Control of $CaCO_3$ polymorphism	[16]

For the induction of calcite and aragonite nucleation, systematic investigations on biological and suitably assembled artificial systems have shed some light on the structural requirements of a putative nucleation site, especially in mollusc shells. [17] The model proposes structurally pre-organized domains of acidic residues, such as aspartic acid and glutamic acid, which could serve as a supramolecular template for oriented crystal nucleation. Such highly ordered domains could result from acidic macromolecules being adsorbed on a rigid scaffold of insoluble matrix proteins (see *Figure 3.2*).[17,18] Inspired by this discovery, artificial macromolecules, which hold ordered structures by themselves without the assist of scaffold, could be served as the ideal candidates for directing the

process of nucleation of mineral crystals.

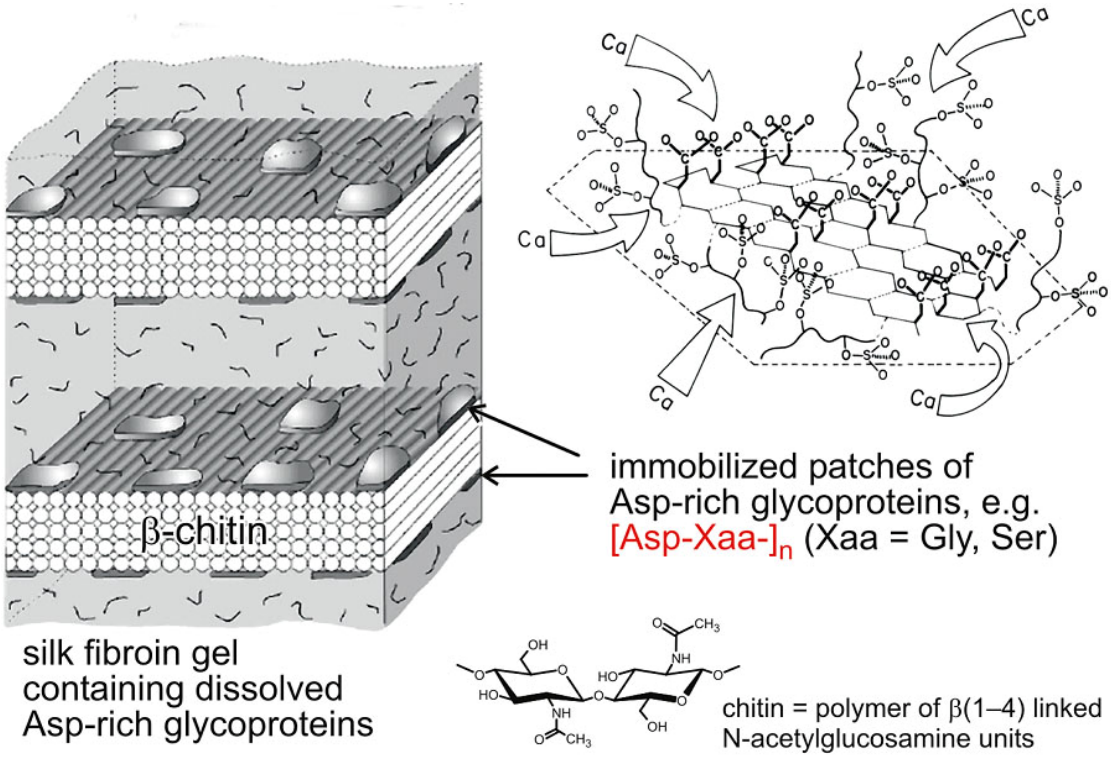


Figure 3.2 *Left*: Schematic representation of the organic matrix in the nacreous layer of *Atrina*. The β-chitin lamella are interspersed in a highly hydrated silk fibroin gel. The gel contains soluble Asp-rich glycoproteins, which can bind to the β-chitin surface by means of hydrophobic or electrostatic interactions. *Right*: Structure model of a putative nucleation site in molluscan tissues. The sulfate groups, linked to flexible oligosaccharide side chains, concentrate Ca^{2+} ions on an Asp-rich oligopeptide domain that is assumed to adapt a highly regular β sheet conformation. A first layer of Ca^{2+} ions may thus be fixed and oriented in space upon which further mineral growth ensues. [17, 18]

As an example, the interlamellar organic sheets of nacre from mollusc shells consist of thin sheets of β-chitin[19] sandwiched between thicker sheets of silk fibroin-like proteins.[20] Silk fibroin itself possesses microcrystalline domains of repeating $[Gly\text{-}Ala\text{-}Gly\text{-}Ala\text{-}Gly\text{-}Ser]_n$ units that adopt an antiparallel β-pleated sheet conformation. These domains have a highly regular and hydrophobic surface upon which acidic macromolecules are adsorbed from solution. In the course of adsorption, the acidic macromolecule has to fold into the appropriate conformation, in order to maximize its hydrophobic interactions with the silk fibroin surface. Possible candidates for acidic macromolecules interacting with silk fibroin in the described way are oligopeptides that

include sequence motifs of [Asp-Xaa]$_n$, (Xaa = Gly, Ser), which have a strong tendency to fold into a β sheet conformation in the presence of Ca^{2+} ions.[21] As a consequence, the aspartic acid residues of [Asp-Xaa]$_n$ sequences would be positioned at only one side of the β-pleated sheet, resulting in an organized two-dimensional array of carboxylate ligands.

It is tempting to assume that carboxylate residues coordinate a first layer of Ca^{2+} which would in turn become the first layer of an epitaxially growing $CaCO_3$ crystal. However, more profound analysis so far has failed to provide evidence for an epitaxial growth mechanism or a close stereochemical complementarity between the nucleating macromolecules and the incipient $CaCO_3$ crystal surface.[22] Acidic proteins extracted from different calcified tissues were shown to exert vast control on polymorph selection, texture and morphology of $CaCO_3$ crystals.[23] However, up to now only a few artificial oligopeptides have been designed and tested for specific interactions with $CaCO_3$ single crystals *in vitro*.[24]

An approach using amphiphilic peptides comprising alternating hydrophilic (Asp or Glu) and hydrophobic (Phe) amino acid rediues is adopted in order to imitate the epitopes of acid proteins from calcified tissues (see *Figure 3.3*).[22]

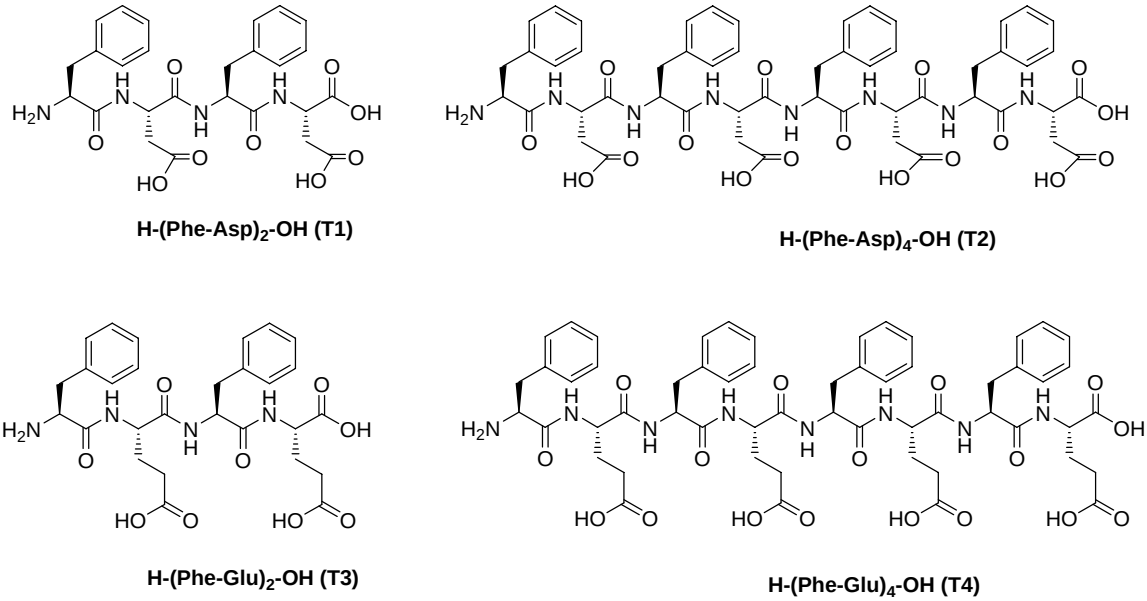


Figure 3.3 Artificial acidic peptides mimicking the potential binding epitopes of highly acidic peptides of the water-soluble fraction from mollusc shells.

Calcite crystals grown in the presence of **T 2** significantly differ in shape from the control. A specific inhibition of calcite {11.0} and {01.2} crystal faces occurs, as is shown by *Figure 3.4*.[22]

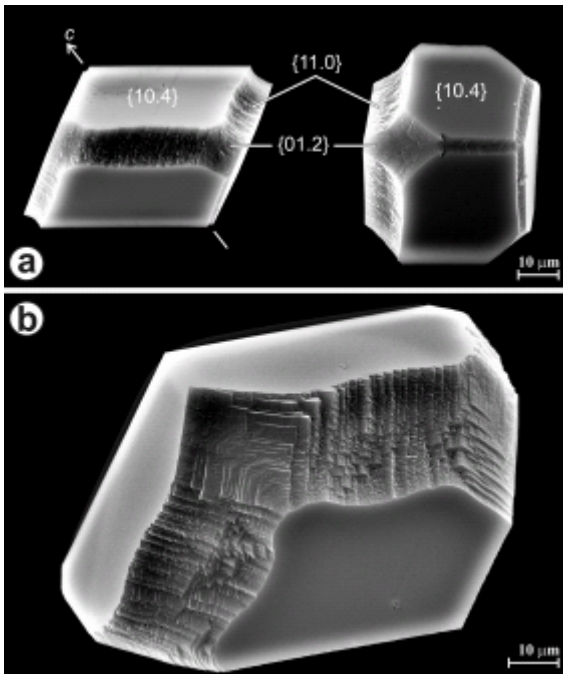


Figure 3.4 Scanning electron micrographs of calcite crystals grown in the presence of **T 2**.

The {01.2} plane is a polar plane whereas the {11.0} plane is non-polar. However, biogenic calcite crystals display these faces quite often.[25] The growth inhibition of {01.2} and {11.0} faces of calcite induced by acidic peptides containing a repeating Phe-Asp or Phe-Glu sequence motif is the first example where artificial peptides exert the same influence on the growth habit of calcite crystals as natural acidic proteins isolated from skeletal elements of various organisms. Moreover, this is the first example where single peptides selectively interact with two distinct crystal faces of calcite.

X-ray crystallographic analysis of **T 3** reveals an antiparallel β-sheet structure with formation of carboxylic acid dimers between adjacent strands (see *Figure 3.5*). The amphiphilic arrangement leads to an ordered projection of functional groups in different layers of the sheet.[26]

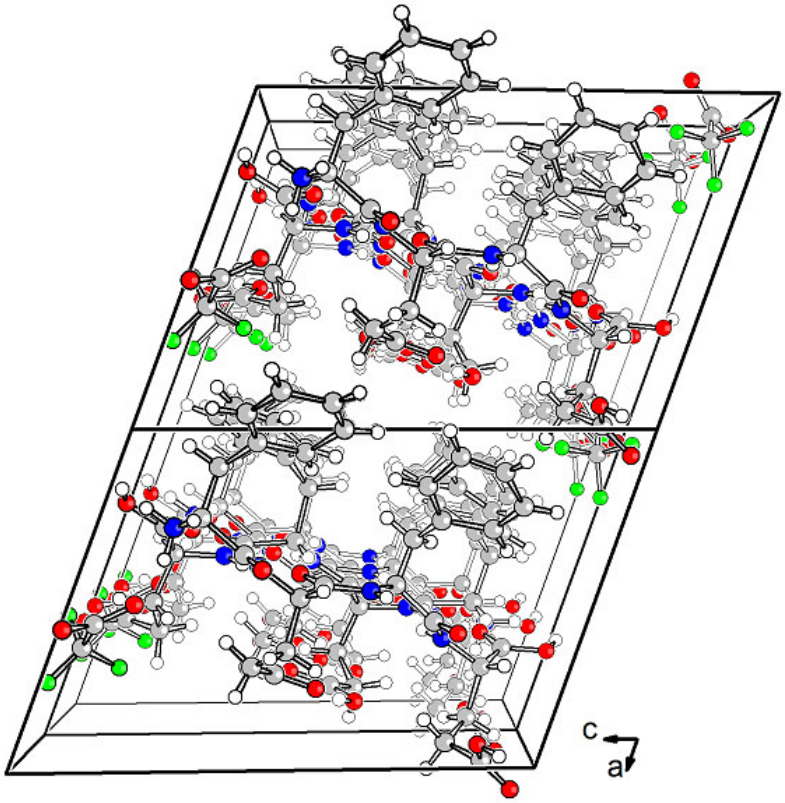


Figure 3.5 Packing plot of the crystal structure of *H-(Phe-Glu)$_2$-OH*•CF_3COOH. (Crystallographic data:triclinic, space group *P*1; *a* = 10.1259 , *b* = 11.7410, *c* = 14.7942 Å; a = 82.81, b = 70.26, g = 89.83°; Z = 2.)[26]

As already discussed before, the three dimensional structure of the macromolecule template peptides remains to be a bottle-neck, the deep research of the conformation of these β-hairpin template peptides will be extremely important to decipher the process of biomineralization.

3.2 Secondary Structure of Peptides

3.2.1 Secondary Structure Motifs

In biochemistry and structural biology, secondary structure is the general three-dimensional form of local segments of biopolymers such as proteins and nucleic acids (DNA/RNA). It does not, however, describe specific atomic positions in three-dimensional space, which are considered to be tertiary structure.

Secondary structure is formally defined by the hydrogen bonds of the biopolymer, as observed in an atomic-resolution structure. The hydrogen bonding is correlated with other structural features however, which has given rise to less formal definitions of secondary structure. For example, residues in protein helices generally adopt backbone dihedral angles in a particular region of the Ramachandran plot; thus, a segment of residues with such dihedral angles is often called a "helix", regardless of whether it has the correct hydrogen bonds.

Secondary structure in proteins consists of local inter-residue interactions mediated by hydrogen bonds. The most common secondary structures are α-helices and β-sheets. Other helices, such as the 3_{10}-helix and π-helix, are calculated to have energetically favorable hydrogen-bonding patterns but are rarely if ever observed in natural proteins except at the ends of α helices due to unfavorable backbone packing in the center of the helix. Other extended structures such as the polyproline-helix and α-sheet are rare in native state proteins but are often hypothesized as important protein folding intermediates. Tight turns and loose, flexible loops link the more regular secondary structure elements. The random coil is not a true secondary structure, but is the class of conformations that indicate an absence of regular secondary structure.

3.2.2 β-Sheet Geometry

3.2.2.1 Geometric Parameters of β-Sheet Conformation

One of the major structural elements found in globular proteins is the β-sheet. This structure is built up from a combination of several regions of the polypeptide chains, in contrast to α helix, which is built up from one continuous region. These regions, β strands, are usually from 5 to 10 residues long and are in an almost fully extended conformation with φ, ψ angles within the broad structurally allowed region in the upper left quadrant of the Ramachandran plot.

The majority of β-strands/sheets are arranged adjacent to other strands and form an extensive hydrogen bond network with their neighbors in which the NH groups in the backbone of one strand establish hydrogen bonds with the C=O groups in the backbone of the adjacent strands (see *Figure 3.6*). In the fully extended β-strands, successive side chains point straight up, then straight down, then straight up, and so on. Adjacent β-strands in a β-sheet are aligned so that their C^{α} atoms are adjacent and their side chains point in the same direction. The "pleated" appearance of β-strands arises from tetrahedral chemical bonding at the C^{α} atom; for example, if a side chain points straight up, then the bond to the C^{α} must point slightly downwards, since its bond angle is approximately 109.5°. The pleating causes the distance between C^{α}_{i} and C^{α}_{i+2} to be approximately 6 Å, rather than 7.6 Å (2 × 3.8 Å) expected from two fully extended *trans* peptide virtual bonds. The "sideways" distance between adjacent C^{α} atoms in hydrogen-bonded β strands is roughly 5 Å.

However, β-strands are rarely perfectly extended; rather, they exhibit a slight twist due to the chirality of their component amino acids. The energetically preferred dihedral angles (φ, ψ) = (–135°, 135°) (broadly, the upper left region of the Ramachandran plot) diverge somewhat from the fully extended conformation (φ, ψ) = (–180°, 180°).[27] The twist is often associated with alternating fluctuations in the dihedral angles to prevent the individual β-strands in a larger sheet from splaying apart.

The side chains point outwards from the folds of the pleats, roughly perpendicularly to the plane of the sheet; successive residues point outwards on alternating faces of the sheet.

3.2.2.2 Hydrogen Bonding Patterns

Because peptide chains have directionality conferred by their *N*-terminus and *C*-terminus, β-strands can be said to be directional, too. They are usually represented in protein topology diagrams by an arrow pointing toward the *C*-terminus. Adjacent β strands can form hydrogen bonds in antiparallel, parallel, or mixed arrangements.

In an antiparallel arrangement (see *Figure 3.6* **A**), the successive β-strands alternate directions so that the *N*-terminus of one strand is adjacent to the *C*-terminus of the next. This is the arrangement that produces the strongest inter-strand stability because it allows the inter-strand hydrogen bonds between carbonyls and amines to be planar, which is their preferred orientation. The peptide backbone dihedral angles (φ, ψ) are about (–140°, 135°) in antiparallel sheets. In this case, if two atoms C^{α}_{i} and C^{α}_{j} are adjacent in two hydrogen-bonded β-strands, then they form two mutual backbone hydrogen bonds to each other's flanking peptide groups; this is known as a close pair of hydrogen bonds.

In a parallel arrangement (see *Figure 3.6* **B**), all of the *N*-terminus of successive strands are oriented in the same direction; this orientation is slightly less stable because it introduces nonplanarity in the inter-strand hydrogen bonding pattern. The dihedral angles (φ, ψ) are about (–120°, 115°) in parallel sheets. It is rare to find less than five interacting parallel strands in a motif, suggesting that a smaller number of strands might be unstable. In this case, if two atoms C^{α}_{i} and C^{α}_{j} are adjacent in two hydrogen-bonded β-strands, then they do not hydrogen bond to each other; rather, one residue forms hydrogen bonds to the residues that flank the other (but not vice versa). For example, residue *i* may form hydrogen bonds to residues $j - 1$ and $j + 1$; this is known as a wide pair of hydrogen bonds. By contrast, residue j may hydrogen-bond to different residues all together, or to none at all.

Finally, an individual strand may exhibit a mixed bonding pattern, with a parallel strand on one side and an antiparallel strand on the other. Such arrangements are less common than a random distribution of orientations would suggest, indicating that this pattern is less stable than the antiparallel arrangement.

The hydrogen bonding of β-strands need not be perfect, but can exhibit localized disruptions known as β-bulges. The hydrogen bonds lie roughly in the plane of the sheet, with the peptide carbonyl groups pointing in alternating directions with successive residues; for comparison, successive carbonyls point in the same direction in the α-helix.

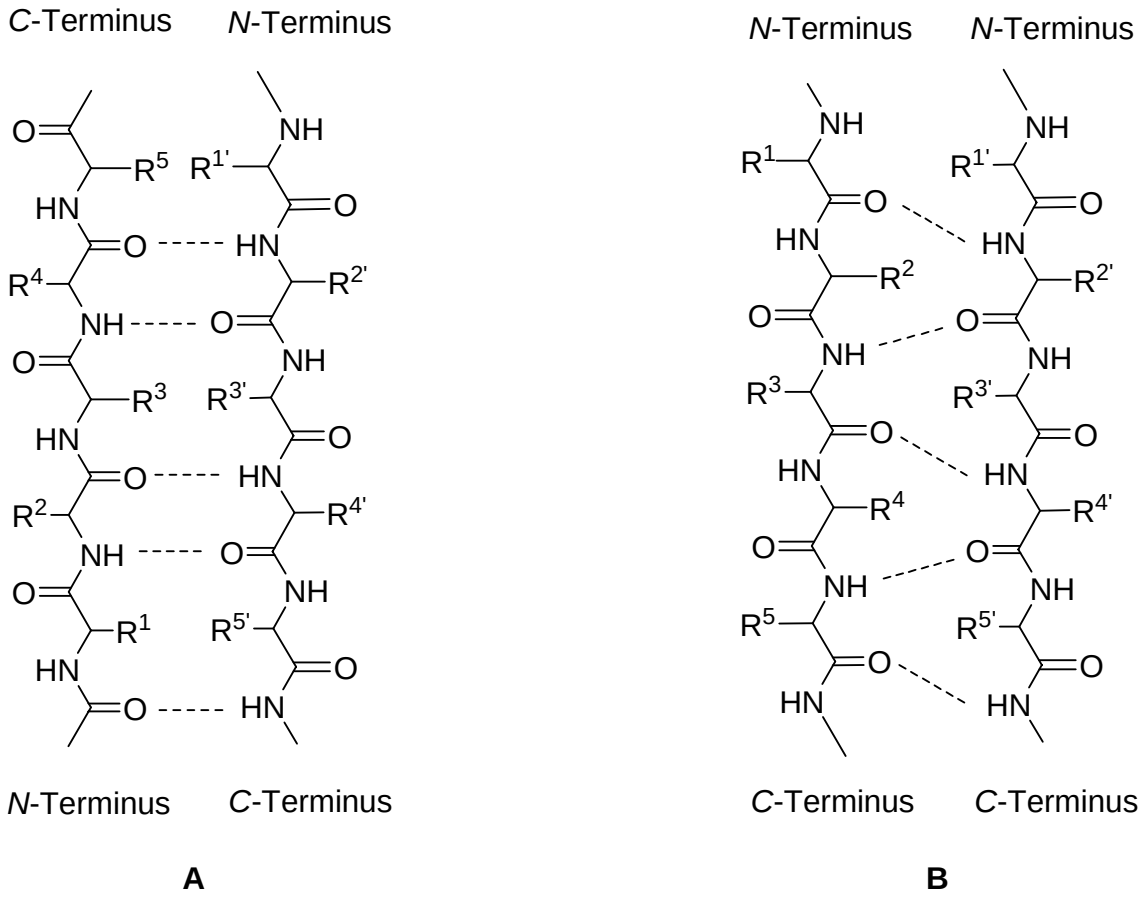


Figure 3.6 Illustration of the hydrogen bonding patterns, represented by dotted lines, in an antiparallel β-sheet **A** and a parallel β-sheet **B**.

3.2.3 β-Turn and β-Hairpin Structure

3.2.3.1 The Reverse Turn

The simplest motif involving β-strands is two adjacent antiparallel strands joined by a loop. This motif, which is named as a hairpin unit, occurs rather frequently. It is present in most antiparallel β-sheets both as an isolated ribbon and as part of more complex β-sheets. There is a strong preference for β-strands when they are adjacent in the amino acid sequence to form a hairpin motif. The lengths of the loop region between β-strands vary but are generally from two to five residues long.

Cross-strand side chain interactions have been proposed to account for observed amino acid preferences for parallel or antiparallel β-sheet structure. The side chain of residues in β-structure point at right angles alternately above and below the plane of propagating polypeptide chain. As a consequence, patterns in amino acid character in the primary sequence of a segment that adopts β-sheet structure can lead to sidedness in the character of a β-sheet. Alternating polar and nonpolar residues will creat an amphipathic sheet.

If a β-strand is folded back on itself to form an antiparallel association with the succeeding region of polypeptide, there must necessarily be residues forming a chain reversal. Accomplishing this chain reversal with two corner residues and an *i* to *i*+3 hydrogen bond leads to a conformational feature know as a β-turn. Initially recognized in silk proteins by Geddes *et al*.[28] and later defined stereochemically by Venkatachalam,[29] β-turns are widespread in proteins and both linear and cyclic peptides. The stereochemical constraints of establishing *i* to *i*+3 hydrogen bond and connecting to the incoming and outgoing β strands restricts the possible conformations of the *i*+1 and *i*+2 residues in the turn. Two major classes of β-turn meet the stereochemical criteria with all peptides bonds *trans*: type I and type II. These turns differ in the orientation of the peptide bond between the corner residues and consequently in the preferred side chain dispositions in the corner position.

In type I β-turn, both *i*+1 and *i+2* positions accommodate L-residues; however, proline preferentially fits in the *i*+2 position. Type I β-turns are the most prevalent in naturally

occurring proteins.[30,31] In a type II β-turn, the *i*+1 position can accommodate an L-residue (proline fits in this position), whereas the *i*+2 position favors a glycine, small polar L-residue, or a D-residue due to the steric clash with a side chain in the L-configuration. Proline in position *i*+1 is a strong sequence determinant for either type I or type II β-turn because of the restriction on the φ angel from the cyclic side chain.

Other β-turn geometries are also found: the mirror images of type I and type II are known as type I' and II" and are energetically equivalent if residues of the opposite chirality occupy corresponding sequence positions. For example, a D-proline residue would be favored in position *i*+1 of a type I' turn or in position *i*+2 of a type II' turn. The intervening peptide bond (between position *i*+2 and *i*+3) can adopt a *cis* conformation and still allow the turn to form an *i* to *i*+3 hydrogen bond and link two β strands. The resulting turn is called a type VI β-turn. A single turn of 3_{10}-helix[32,33] forms the requisite *i* to *i+3* hydrogen bond but has the incoming and outgoing strands significantly twisted with respect to on another. This is called a type III β-turn.

3.2.3.2 The β-Turn Motif in β-Hairpin Conformation

The stereochemical disposition of side chains in β-turn can be treated in a manner that is similar to *R*-group arrangements on cyclohexyl rings.[34] When the sidechain is oriented outward in the same plane as the polypeptide backbone, it can be called equatorial and, when it is oriented above or below the backbone plane, it can be called axial. The turn type and configuration of amino acids in the turn position will lead to predetermined side chain orientations, as summarized in Table 3.3. Assuming that the conformation of the turn sequence is tied down adequately, on can exploit turns to present side chains as reactive groups or as recognition moieties. Frequently, cyclization, either via the backbone or via side chain bridging, has been used to constrain turns and create desired handles for a range of purposes.[34]

Table 3.3 Turn Type and Configurations of Amino Acids at Corner Positions

Turn Type	Position *i*+1		Position *i*+2	
	L-Residue	D-Residue	L-Residue	D-Residue
I	eq	ax (down)	ax (up)	eq
I'	ax (up)	eq	eq	ax (down)
II	eq	ax (down)	eq	ax (down)
II'	ax (up)	eq	ax (up)	eq
III	undetermined	undetermined	undetermined	undetermined
VI	eq	eq	ax (down)	ax (down)

In summary, the category of turn structure is determined by their Ramachandaran angles of backbone at residue *i*+1 and *i*+2 (Table 3.4).[35-37] These two categories of β-turn are shown by statistic data to be the ones which are most frequently appear in the β-turn motif of the β-hairpin protein.[38,39]

Table 3.4 φ and ψ values of residue *i*+1 and *i*+2 in type I' and II' β-turn.

Turn Type	Position *i*+1		Position *i*+2	
	φ	ψ	φ	ψ
I'	+60º	+30º	+90º	0º
II'	+60º	-120º	-80º	0º

In both of type I' and type II' turns, their φ_{i+1} is positive and lies within a region of conformational space which is very poorly populated by L-amino acids, as a consequence of steric limitations.[40] Since the constraints of pyrrolidine ring formation restrict the φ value in D-Pro to be +60 ± 20º, D-Pro is thus often chosen as the residue in *de novo* design of the synthetic peptides to occupy the *i*+1 position of β-turn in order to nucleate the peptide into β-hairpin conformation. Less frequently, Asparagine residues, with a significant propensity for positive φ values are also observed at this position.[41]

The successful design of a stable β-hairpin in an apolar octapeptide Boc-Leu-Val-Val-D-Pro-Gly-Leu-Val-Val-OMe has been achieved using a centrally positioned D-Pro-Gly segment to nucleate a type II'-β turn.[42] The L-Pro-Gly peptide derivative Boc-Leu-Val-Val-Pro-Gly-Leu-Val-Val-OMe does not adopt a β-hairpin conformation despite the presence of a Pro-Gly β-turn. Early analysis of β-hairpin structures in proteins reveals that type I'/II'-turns, both of which have positive φ_{i+1} values, are favoured at the hairpin

turn positions.[38,39] A more recent analysis with a larger protein data set suggests that β-hairpins nucleated by type II β-turns are indeed found in proteins. However, strand links are in these cases significantly longer than the average hydrogen-bonded hairpin.[43] After the efficacy of D-Proline-containing loops was established in minimal β-hairpin models, the folding of a larger peptide that contained a β-hairpin was shown to be enhanced by a D-Pro-Gly loop, relative to an L-Pro-Gly loop.[44]

D-Pro-X segments promote autonomous β-hairpin folding in aqueous solution. The first example involved peptides related to the amino-terminal segment of ubiquitin.[45] A series of 16-mers was prepared, MQIFVKSXXKTITLVKV-NH, with five different XX segments: D-Pro-D-Ala, L-Pro-L-Ala, D-Pro-L-Ala, D-Pro-Gly and L-Pro-Gly.[45] NMR data indicated that each of the D-proline-containing peptides displayed β-hairpin folding with a two-residue loop at D-Pro-X, while the L-proline containing peptides appeared to be completely disordered. More recently, the l2 residue designed peptide RYVEVXGOKILQ-NH, has been shown to adopt a β-hairpin conformation in aqueous solution when X is D-Pro, but not when X is L-Pro.[46]

3.2.3.3 β-Hairpin Conformation Stabilization Factors

Self-assembly of β-hairpin conformation is driven not only by numerous intramolecular hydrogen bonds, but also by side-chain to side-chain interactions.

A β-hairpin peptide is composed of two antiparallel β strands (see *Figure 3.7*, carbonyl in the backbone not shown) linked by a short loop. When the loop contains only four residues, the two central residues correspond to positions $i + 1$ and $i + 2$ of a β-turn. The origin of the stability of isolated β hairpins in aqueous solution is unclear with contrasting opinions as to the relative importance of interstrand hydrogen bonding, hydrophobic interactions, and conformational preferences, the latter being associated largely with the turn sequence.

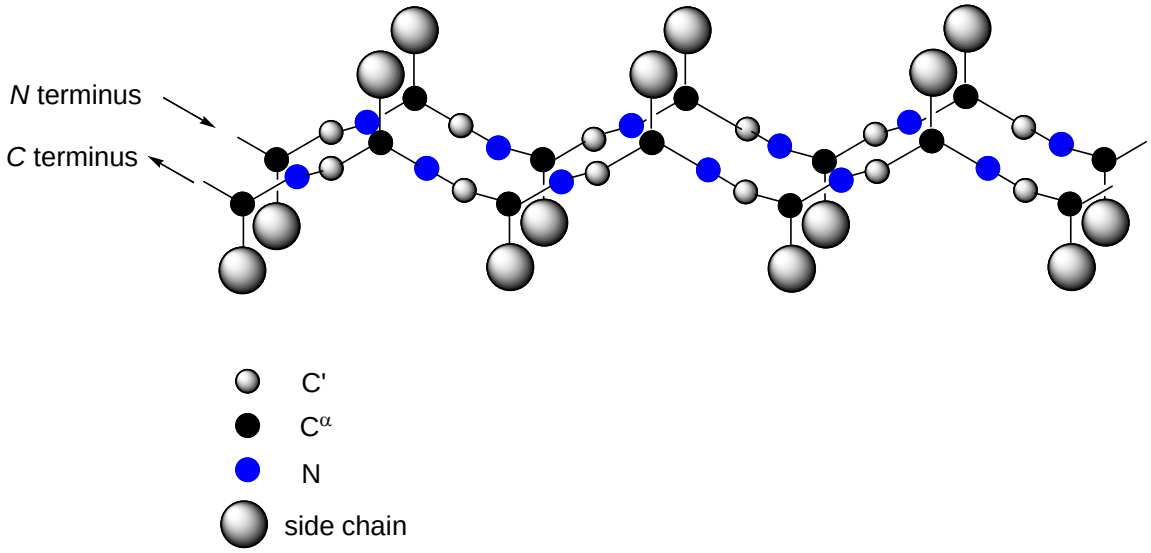


Figure 3.7 Antiparallel β-strands in β-hairpin peptide

3.2.3.3.1 The Loop Segment

Several recent studies highlight the importance of the loop residues in determining overall β-hairpin stability. de Alba *et al*.[47] examined a series of six decamers in which the central four residues were varied, but the first and last three residues were held constant. A remarkable range of β-hairpin conformations was detected by NMR within this small series. Three different loop structures are formed among these peptides, containing two, three or four residues. de Alba *et al*. concluded that "the turn residue sequence determines the turn conformation, and thereby, the other features of the β hairpin conformation, such as the pattern of interstrand residue pairing." Haque and Gellman[45] showed that a β-hairpin with natural registration could be achieved by replacing the native loop with a segment containing a D-proline residue. However, simply changing the proline configuration to L-counterpart abolished β-hairpin formation. Stranger and Gellman observed the same D-proline versus L-proline effect with a designed 12-residue sequence.[46] Griffiths-Jones *et al*.[48] have reported that an -Asn-Gly- segment can induce a modest extent of β-hairpin folding in water even when the attached strands are only two residues long.

3.2.3.3.2 Inter-strand Side chain-Side chain Interactions

Ramirez-Alvarado *et al*.[49] observed that changing two or more of the highlighted strand residues of RGITVNGKTYGR to alanine abolished β-hairpin folding. The authors concluded that "inter-strand side chain-side chain interactions are essential for β-hairpin formation." de Alba *et al*.[50] noted, however, that the strength of this conclusion is "unclear, because alanine has an intrinsically low propensity to be in a β-strand". The authors concluded that they could identify several specific interstrand side chain-side chain interactions that were either favorable or unfavorable with respect to β-hairpin formation. The interactions identified in this way were consistent with favorable and unfavorable interstrand pairings that had previously been suggested by Wouters and Curmi[51] on the basis of statistical analysis of such pairings in the crystallographic database. Maynard *et al*.[52] have reported thermodynamic analysis of the thermal unfolding of designed 16-mer Ac-KKYTVSINGKKITVSI in aqueous solution. The authors deduced that the β-hairpin conformation is stabilized by hydrophobic interactions among side chains on the two opposing β strands, based on the thermodynamic signature of the folding equilibrium, which includes a large, negative change in heat capacity at constant pressure. Interestingly, the thermodynamic signature of the folding equilibrium changes dramatically in water/methanol (1:1), relative to water, and the authors concluded that interstrand hydrogen-bond formation becomes an important stabilizing factor in the mixed solvent. Maynard *et al*.[52] note that although the loop segment plays an important role in β-hairpin formation, a favorable loop is probably not sufficient (i.e. that favorable interstrand interactions are required for β-hairpin formation).

3.3 Circular Dichroism Spectroscopy in Peptide Conformation Analysis

3.3.1 CD Spectroscopy

Circular Dichroism (CD) is very sensitive to the secondary structure of polypeptides and proteins. Circular Dichroism Spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light (rather than the commonly used absorbance of isotropic light) by a substance. It has been shown that CD spectra between 260 and approximately 180 nm can be analyzed for the different secondary structural types: α-helix, parallel and antiparallel β-sheet, turn, and other. A number of excellent review articles are available describing the technique and its application (Woody, 1985; Johnson, 1990). In fact, optical rotary dispersion (ORD, *vide infra*) data suggested a right-handed helical conformation as a major protein structural element before the Pauling and Corey model (Pauling & Corey, 1951) and Kendrew's structure of myoglobin. Modern secondary structure determination in proteins by CD are reported to achieve accuracies of 0.97 for helices, 0.75 for β-sheet, 0.50 for turns, and 0.89 for other structure types (Manavalan & Johnson, 1987).

3.3.2 Physical Principles of CD

Inherently asymmetric chromophores (uncommon) or symmetric chromophores in asymmetric environments will interact differently with right- and left-circularly polarized light resulting in two related phenomena. Circularly-polarized light rays pass an optically active medium at different velocities due to the different indices of refraction for right- and left-circularly polarized light called optical rotation or circular birefringence. The variation of optical rotation as a function of wavelength is called optical rotary dispersion (ORD). Right- and left-circularly polarized light will also be absorbed to different extents at some wavelengths due to differences in extinction coefficients for the two polarized rays called circular dichroism (CD). Optical rotary dispersion enables a chiral molecule to rotate the plane of polarized light. ORD spectra are dispersive (called a Cotton effect

for a single band) whereas circular dichroism spectra are absorptive. The two phenomena are related by the so-called König-Kramers transforms.

It can be shown that if right- and left-circularly polarized light is absorbed to different extents at any wavelength, there will be a difference in refractive indices at virtually all wavelengths. This accounts for the ability of small saturated chiral molecules to rotate the plane of polarized light of the D-line of sodium (589 nm), far away from absorptive bands. The dispersive nature of ORD is also the reason CD spectroscopy is a more sensitive analytical technique. Unlike the dispersive ORD phenomenon, circular dichroism can only occur within a normal absorption band and thus requires either an inherently asymmetric chromophore (uncommon) or a symmetric one in an asymmetric environment.

3.3.3 Secondary Structure from CD Spectra

The simplest method of extracting secondary structure content from CD data is to assume that a spectrum is a linear combination of CD spectra of each contributing secondary structure type (e.g., "pure" α-helix, "pure" β-sheet etc.) weighted by its abundance in the polypeptide conformation. The major drawback of this approach is that there is no standard reference CD spectrum for "pure" secondary structures. Synthetic homopolypeptides used to obtain reference spectra are in general, poor models for the secondary structures found in proteins. For example, the CD of an α-helix has been shown to be length dependent and no homopolypeptide system has been found that is a good example of the β-sheet structure found in proteins.

In response to these shortcomings, several methods have been developed which analyze the experimental CD spectra using a database of reference protein CD spectra containing known amounts of secondary structure (Provencher & Glöckner, 1981; Hennesey & Johnson, 1981; Manalavan & Johnson, 1987; Sreerama & Woody, 1994). These methods are in general more accurate and reliable than the novel approach outlined above.

In one method (Manalavan & Johnson, 1987), single value decomposition is used to create orthogonal CD basis vectors from CD spectra of proteins with known secondary

structure. Using the statistical technique of variable selection, unimportant variables are removed from an underdetermined system of equations allowing the solution for the important ones. The great advantage of this technique (and others like it) is that one makes no assumptions on the form of the CD from the individual secondary structural elements. One needs only to be sure that the overall structural characteristics of the analyzed protein are represented in the set of reference spectra. In this way, irregularities of secondary structure and length dependencies should be taken into account.

Two very important requirements for successful secondary structure analysis using these techniques deserve further comment. One is that the CD spectra need to be recorded from about 260 nm to at least 184 nm (and preferable 178 or below; Johnson, 1990) and the other is that an accurate protein concentration (< 10% error) is essential. The assumptions and limitations of these techniques are discussed in detail by Manning (1989).

Circular dichroism spectroscpy has also been used to determine the tertiary structure class of globular proteins. The method proposed by Venyaminov & Vassilenko (1994) claim 100% accuracy for predicting all α, α/β, and denatured proteins; 85% for α + β; and 75% for all β proteins.

Circular dichroism spectroscopy has been extensively applied to the structural characterization of peptides. The application of CD for conformational studies in peptides (like proteins) can be largely grouped into a) monitoring conformational changes (e.g., monomer-oligomer, substrate binding, denaturation, etc.) and b) estimation of secondary structural content (e.g., this peptide is 25% helical under these conditions). As already mentioned, CD is particularly well-suited to determine structural changes in both proteins and peptides. However, the absolute structural content is more difficult to derive and is prone to over-interpretation. Length-dependencies of the CD spectra of "pure" secondary structures are only one potential caveat.

The CD spectra of peptides have been reported in a number of solvent systems. The helix-promoting characteristics of trifluoroethanol (TFE) and hexafluoroisopropanol (HFIP) are well-known. A number of authors have argued that these solvents mimic a particular characteristic of the in vivo system and therefore attach relevance to structural

studies performed in these solvents. Of particular importance to this matter is a study by Waterhous & Johnson (1994) who demonstrated that peptide sequences could be induced to form α-helices or β-strands in TFE and non-micellar SDS solutions, respectively, regardless of their secondary structure in the native protein. This result underscores the difficulties in assigning relevance to structural studies in peptides using CD spectroscopy.

All-β proteins usually have a single, negative and a single, positive CD band, whose intensities are much lower than those of α helix. Their CD spectra can vary considerably among different all-β proteins. Those of regular all-β proteins resemble the CD spectrum of the β sheet of $(Lys)_n$, and usually have a single minimum between 210 and 225 nm and a stronger positive maximum between 190 and 200 nm.

3.4 Aims of the Study

The ultimate purpose to explore the exact behaviors of natural phenomenon of biomineralization could be achieved by the thorough study of biomacromolecules which direct this process. The proteins extracted from the mollucs exhibited us some hints which could be utilized as templates to investigate. The relationships between the primary and secondary structures of concerned biomacromolecules and the behaviors of the biomineralization and the patterns of thus generated crystals is one of the focus of this project. Peptides which adopt amphiphilic β-hairpin conformation could stand out as appropriate candidates that mimic the naturally occurring macromolecules that direct the process of biomineralization. The air/water interface where these amphiphilic β-hairpin peptides are located could imitate the actual enviroments of macromolecules controlling biomineralization in the living organism. The hydrophilic acidic side chains on these amphiphilic β-hairpin peptides will point to the water and hydrophobic side chains point to the air. While in organism, the acidic side chains on the macromolecules which direct the biomineralization point to the water solution containing mineral ions such as calcium; the hydrophobic residues will interact with the hydrophobic core of the rigid scaffold of insoluble matrix protein. The extended conformation of β-hairpin peptides hold ordered structures by themselves without the assist of scaffold, making them quantified as ideal candidates for organizing the process of nucleation of minerals.

As the process of biomineralization concerns the interactions between the protein templates in the tissues of molluscs and inorganic compounds which finally crystallize in the organisms, this natural phenomenon presents an ideal project to utilize the tool of supramolecular chemistry of peptides/protein to address questions in biomineralization. From the available information of the primary sequences of the peptides which participate in the process of biomineralization, mimics of these peptides should be synthetically obtained and applied. One of the most significant advantages of the application of synthetic peptides/peptidomimetics as templates to explore the

biomineralization lies in the fact that the relationship between the behaviors of biomineralization, crystal morphology and the secondary structures of the concerned peptide templates could be resolved. This achievement is more valuable when one considers that the useful information centered on this issue is specially limited and scarce. It is therefore extremely instructive for scientists and engineers who are enthusiastic in the beautiful structures and functional properties of the materials formed within living organisms. The decipherment of the structure-behavior relationship in biomineralization could not only bring forward the progress in the theoretical research field, but also the breakthrough in corresponding application domains in pharmacology and medicine.

Another main purpose of this project is to explore the process of nucleation of the peptide chain into β-hairpin conformation, the factors which stabilize/destabilize their conformation, the primary sequence dependence, and the solvent dependence. The extent of the research about the formation of β-sheet and β-hairpin lags behind that of α-helix, partially because of the difficulties in the synthetic process of these peptides. Moreover, they are not easy to handle, because of the inclination to aggregate and poor solubility in common solvents. There are still some controversial viewpoints up to now about the formation of β-sheet and β-hairpin peptides.

The β-hairpin peptides applied to the biomineralization in this project could be possibly utilized as templates to explore the factors to stabilize/destabilize the β-hairpin conformation by mutations of the amino acid residues in the primary sequences which could be crucial in determining the peptide conformation. By studying the conformational changes of the concerned peptides under different environments, such as distinctive solvents like water or halogenated alcohol, or at different pH values, the weight of factors which determine the conformational stabilities of the β-hairpin peptides could be possibly figured out. The results of the conformational analysis of concerned peptides applied in this project could potentially afford useful information related to the formation of β-hairpin structures, and could be therefore instructive for the *de novo* design of the β-sheet or β-hairpin peptides.[53]

It is forseeable that the synthesis of β-sheet derivative peptides is one of the challenging obstacles in the progress of the relative research domains, special synthetic methodologies are supposed to applied in this project. Since the intra-chain hydrogen bonds are regarded as the underlying factors that impede the smooth coupling of amio acids to the growing peptidyl chains on the resin, as well as the quantitive removal of the *N*-terminal Fmoc protecting groups, methods address the issues of intra-chain hydrogen bonds breaking should be utilized. Use of special solvent such as TFE or HFIP could be helpful but not addressing the essential problems. The incorporation of pseudo-proline analogues is efffective but not universal. Under these circumstances, the selective protections of the backbone amide groups by acid labile moity Hmb which is compatible with Fmoc/tBu chemistry of peptide synthesis could be one of the optimal choices. This methodology could be applied to all types of peptides, its incorporation and removal could be smooth and problem free. The incorporation of Hmb backbone protecting groups into the peptides which potentially adopt β-sheet or β-hairpin secondary structure could also increase their solubities, making them adaptable for chromatographic purification. Furthermore, the conformational influences the backbone- and sidechain-protecting groups bring forward are another research interest of the project.

3.5 Results and Discussion

3.5.1 Design of the Peptide Templates in Biomineralization

Calcium salts in organisms are packed and aligned cooperatively between organic matrices, such as collagens and some acidic proteins. This process is regarded as mineralization in a living body, a process known as biomineralization. During crystal formation, the organic matrix plays a role for the control of the location and pattern of nucleation. In order to imitate the natural process preferably, the synthetic macromolecules that mimic the naturally occurring protein matrices should preferentially resided in the interface between the hydrophobic and hydrophilic environments, which reproduces the water/protein biomineralization interface in organism.

Peptide adopting a β-hairpin is basically a "flat" macromolecule, despite being distorted to some extent under normal conditions. In order to exist stably between the interfaces of hydrophobic and hydrophilic bulk enviroments, the hairpin peptides should possess side-chains with distinctively different polarities on the two "faces" of the hairpin structure. This design command introduced the conception of amphiphilic hairpin peptides. If the side chains pointing upwards are exclusively hydrophobic/hydrophilic, meanwhile the side chains pointing downwards are exclusively hydrophilic/hydrophobic, that is to say, the polarity properties of upward and downward side chains are opposite, and this category of β-hairpin peptide is named amphiphilic β-hairpin peptide.

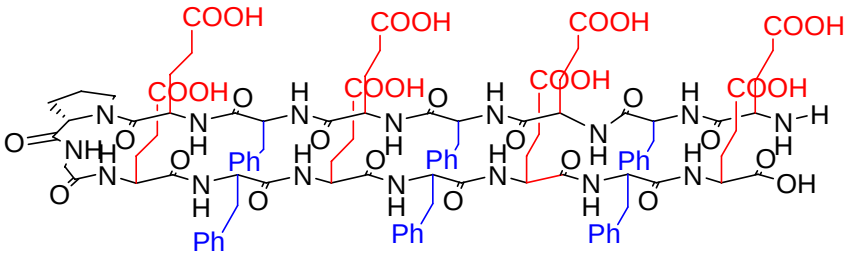


Figure 3.8 Amphiphilic β-hairpin pepide EFEFEFEpGEFEFEFE **HP 19**

Figure 3.8 shows one example of amphiphilic β-hairpin peptides with the primary sequence EFEFEFEpGEFEFEFE **HP 19**, which was applied in this project as the template to explore the behaviour of biomineralization and structure properties. It is obvious that the side chains pointing upwards exclusively originate from acidic glutamic acid, meanwhile the side chains pointing downwards are exclusively originated from hydrophobic phenylalanine. Such an amphiphilic β-hairpin peptide represents an ideal template for the research of the behaviours of molecules at the interfaces, where biomineralization most probably takes place.

It is known that the motif -D-Pro-Gly- is prone to nucleate to form a type II' β-turn at the central *i*+1 and *i*+2 position.[42] It is thus rational to introduce this motif to the β-hairpin peptide to alternate the direction of the strand and assist the formations of intrachain hydrogen bonds between the NH and CO groups on the backbone of the peptide, resulting in the β-hairpin conformation. The thus nucleated peptide would probably adopt the stable hairpin conformation at the interfaces of aqueous solution and air. The hydrophobic side-chains from phenylalanine would point to the air, while the hydrophilic side-chains of glutamate are supposed to be solvated by water molecules. This conformation adopted at the air-water interfaces affords the ideal imitation of the behaviors of the natural biomacromolecules which direct the process of biomineralization in the nature. Peptide design for the project is shown by Table 3.5.

Table 3.5 Designed Peptides for Biomineralization and Conformational Study

No.	Sequence
HP 1	EpGE
HP 2	DpGD
HP 3	DpGK
HP 4	FpGF
HP 5	FDpGKF
HP 6	FDpGDF
HP 7	DFpGFD
HP 8	FEpGEF
HP 9	FEFEpGEFEF
HP 10	FDFDpGDFDF
HP 11	FDFDFDFD
HP 12	FEFEFEFE
HP 13	EFEpGEFE
HP 14	FDFDpGDFDFpGFDFD
HP 15	DFDFDFEpGEFDFDFD
HP 16	EFEFEFDpGDFEFEFE
HP 17	DFSFDFSpGDFSFDFS
HP 18	DFDFDFDpGDFDFDFD
HP 19	EFEFEFEpGEFEFEFE
HP 20	EFSFEFSpGEFSFEFS

3.5.2 Synthesis of Peptides

3.5.2.1 Purposes of Backbone Protection in Peptide Synthesis

Successful assembly of peptides is often hampered by inherent problems such as poor solvation of the growing peptide chain during synthesis on solid supports, as well as limited solubility of fully protected peptides fragments in the solution approach. The synthesis of β-sheet peptides remains to be a tough challenge till now due to their insolubility in most of the common solutions applied in Solid Phase Peptide Synthesis (SPPS). The inter- as well as intra-molecular hydrogen bonds between the β-strand severely impede the freedom degree of the growing peptide chains and therefore prevents the incoming amino acid coupling with the peptide chains smoothly. Particular sequences that favor such phenomena have come to be known as "difficult sequences".[54] This obvious inherent drawback significantly lowers the coupling efficiency of the SPPS of β-sheet originated peptides, which generates series of truncated sequence of peptides.

Another inherent disadvantage caused by inter- or intra-chain molecular bonds during β-sheet related peptide synthesis is the insolubility of such peptides in most of the solvents, which are applied as eluents in preparative Reverse Phase High Performance Liquid Chromatography (RP-HPLC), such as water, methanol, or acetonitrile. The synthesized crude peptides are therefore inapplicable for RP-HPLC purification methodology since their high hydrophobic properties will have them precipitated in these eluents.

Many methods have been proposed to circumvent the solvation problems which are generated by inter- or intra-molecular hydrogen bonds during the β-sheet related peptide synthesis, including addition of chaotropic salts,[55-58] special solvent systems such as trifluoroethanol or hexafluoroisopropanol,[59] or so-called "magic mixtures" such as the mixture of halohydrocarbons and halo alcohols which are extremely efficient at disrupting self-association of peptide chains and enhancing peptide solubility, [60,61] elevated temperature,[62] solubilizing protecting groups,[63] and disruption of the symmetry of side-chain protecting groups.[64] These incomplete couplings can also be

improved by recouple. In such cases, the use of a different coupling method is advisable.

Despite all the endeavors the peptide chemists are striving for, the methods shown above remained only partially effective. In order to address the key factor that generates the previously discussed problems during the synthesis of β sheet related peptide synthesis, Sheppard and co-workers have described the protection of the amide bond in peptides with the 2-hydroxy-4-methoxybenzyl (Hmb) group, which is compatible with the Fmoc/tBu strategy. This protecting group has also been used in the orthogonal solid-phase synthesis of *N*-glycopeptides to avoid aspartimide formation and to improve the solubility of protected peptides, one of the main drawbacks of the convergent approach. Such a backbone protection methodology is presently believed to be the best approach, since it addresses the problem at its root, i.e. formation of hydrogen bonds. To prevent this type of interaction and, correspondingly, aggregation phenomena and precipitation of the related peptides in most applicable solvents in RP-HPLC, substitution of every backbone amide is superfluous, but in the peptide synthetic process this backbone protection will have to be initiated before the chains begin to aggregate. As an empirical rule, it has been found that such backbone protection is efficient if they are implemented at every sixth residue during the chain elongation. For shorter sequences, or intermediate fragments of approximately ten residues, placing the backbone protection around the center of the molecule is apparently sufficient although some exceptions are known.[59]

Backbone protection can also suppress aspartimide formation (see *Figure 3.9*) during the SPPS of some certain sequences that are prone to this problem, such as Asp-Asp or Asp-Gly **20**. Aspartimide derivative **21** can be generated either by acidic or basic catalysis, which can be further transformed into racemic α-peptide **22** and β-peptide **23** through hydrolysis with different ring-open manner, or into racemic piperidide of α-peptide **24** and racemic piperidide of β-peptide **25** through piperidolysis. If the amino acid residue preceding aspartate in the synthesis (residue i+1 if aspartate is taken as residue i) is backbone protected, the problem of aspartimide formation will be fully suppressed. As the amide group of residue i+1 is protected by e.g. Hmb, its nucleophilicity is extremely reduced, preventing it from attacking the carboxy group of

Asp side chain in order to avoid the aspartimide formation. Furthermore, the normally bulky backbone protecting groups will to some extent "freeze" the amide group and hence lower its reactivity. In summary, backbone protection is currently by far the most efficient method against aspartimide formation.

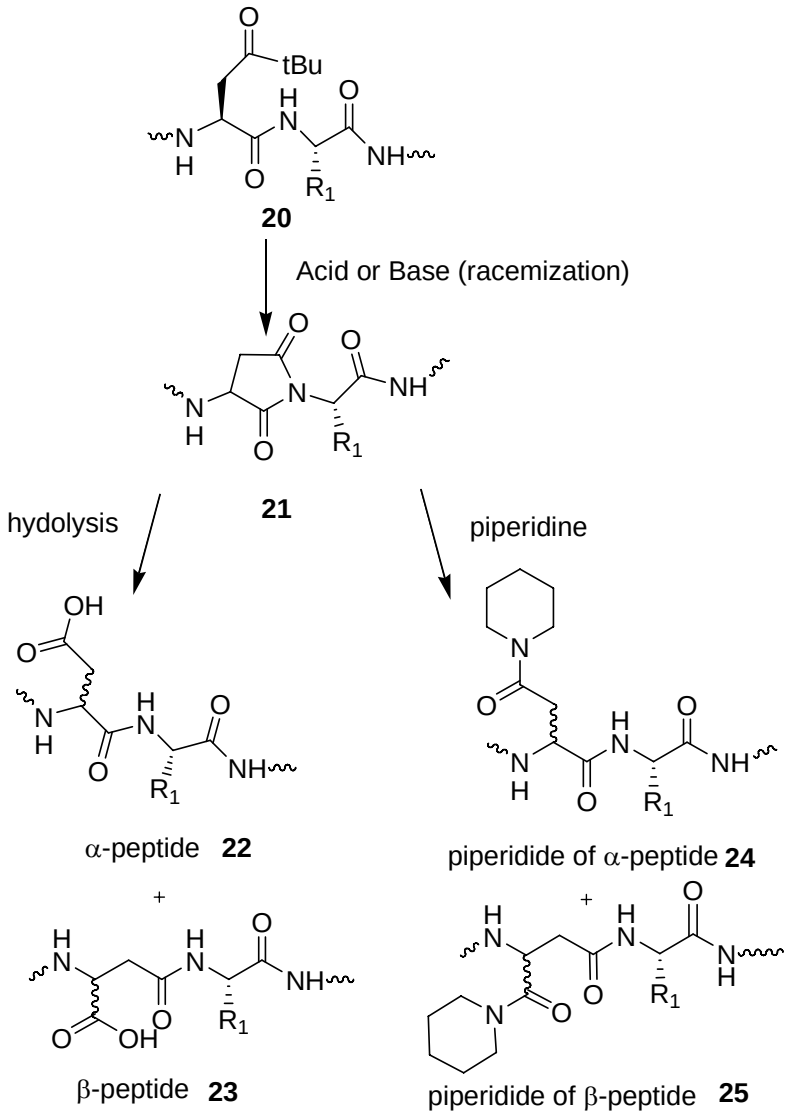


Figure 3.9 Aspartimide formation in peptide synthesis

There are two main types of reversible backbone protection based on (i) the replacement of the amide hydrogen with a protecting group and (ii) bridging the amide nitrogen to a side-chain group of the amino acid residue (so-called pseudoproline methodology). The second approach is limited to the residues such as threonine, serine or cystine, which

bears a hydroxyl or thiol functional group on the side chain, which narrows its universal application. In the first approach, amino acids are N^{α}-substituted with a group that guarantees reversible backbone protection, and these derivatives are additionally protected by a second group at the amino group to prevent aminoacylation during their incorporation into the peptide chain. As is known, introduction of methoxy groups into the benzyl moiety enhances acidolytic cleavage, the aim of solubilizing the otherwise insoluble peptides has been fully fulfilled.

For an appropriate adaptation of the benzyl-type backbone protection to the Fmoc/tBu strategy, the 2-hydroxy-4-methoxybenzyl (Hmb) group was proposed because incorporation of the 4-methoxy moiety leads to significantly enhanced acid lability.[65]

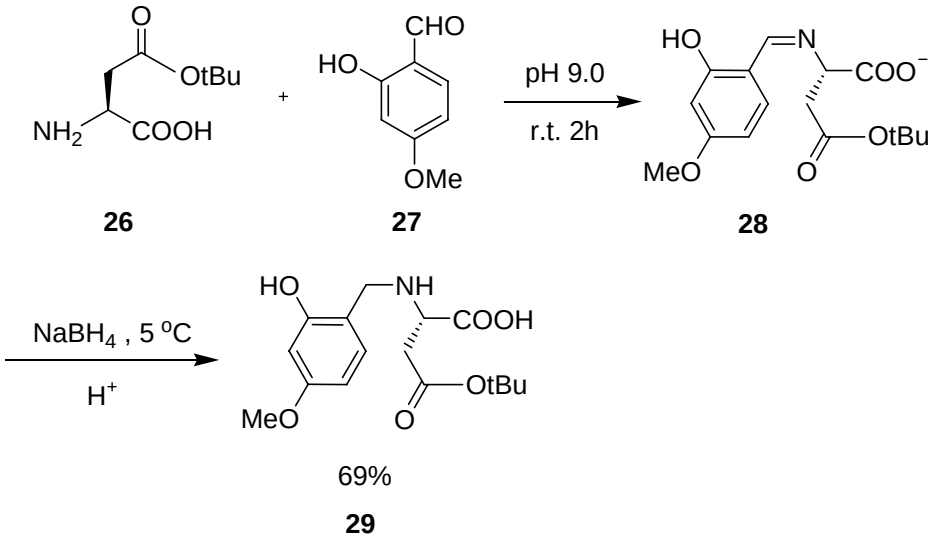
The peptide chain elongation is achieved on the solid support applying Fmoc/tBu strategy (see *Figure 2.13*). This process could be completed manually or by e.g. an automated microwave peptide synthesizer (see *Figure 2.14*), in which microwave irradiation is applied as the input energy to drive the whole process of SPPS.

3.5.2.2 Synthesis of *N,O*-bis-Fmoc-*N*-Hmb-Asp(OtBu)-OH

Since the Hmb protecting group is stable under the basic conditions for Fmoc cleavage during SPPS, it affords the perfect orthogonality to the Fmoc/tBu chemistry strategy of solid phase peptide synthesis. As the Hmb protected amino acid is sterically hindered, it could possibly bring forward the problems such as sluggish acylating rate and low efficient of coupling yields, thus causing racemization and inherent truncated sequences. With respect to this inherent limitation of the Hmb backbone protection strategy, such amino acids which are sterically unhindered will be preferably chosen for Hmb protection group. In this project, the Asp residue is regarded as the optimal acceptor of backbone protecting group since the side-chain of Asp is non-β-branched and the incorporation of Hmb group to N^{α}-amine group proven to be smoothly and efficiently.

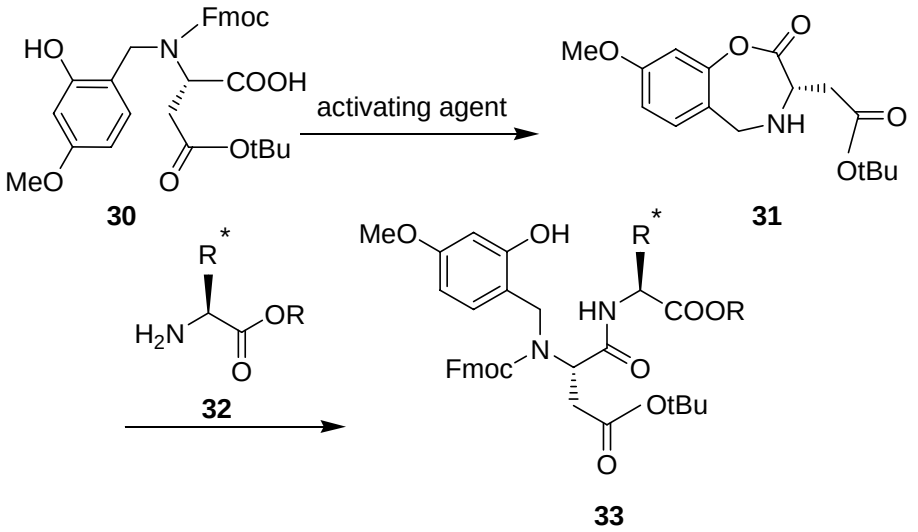
The Hmb group is smoothly incorporated into the relevant amino acid by reductive amination. The 2-hydroxy-4-methoxybenzaldehyde forms the imine with the protected N^{α}-free amino acid, which is reduced by $NaBH_4$ to give the target *N*-Hmb-Amino acid. (*Scheme 3.1*)

The reaction between H-Asp(OtBu)-OH **26** and 2-hydroxy-4-methoxybenzaldehyde **27** to give the H-(Hmb)Asp(OtBu)-OH **29** is achieved first by the nucleophilic addition of amino group of H-Asp(OtBu)-OH to the 2-hydroxy-4-methoxybenzaldehyde giving a hemiaminal –C(OH)(NHR) followed by an elimination of water to yield the imine **28**. The pH of this reaction should be kept around 9.0 to activate the nucleophilicity of the amino group. Subsequently the imine intermediate **28** is cautiously reduced by sodium borohydride under basic conditions and the product is finally neutralized with HCl to generate the target compound H-(Hmb)Asp(OtBu)-OH **29**. The reductive amination procedure is shown in *Scheme 3.1*.



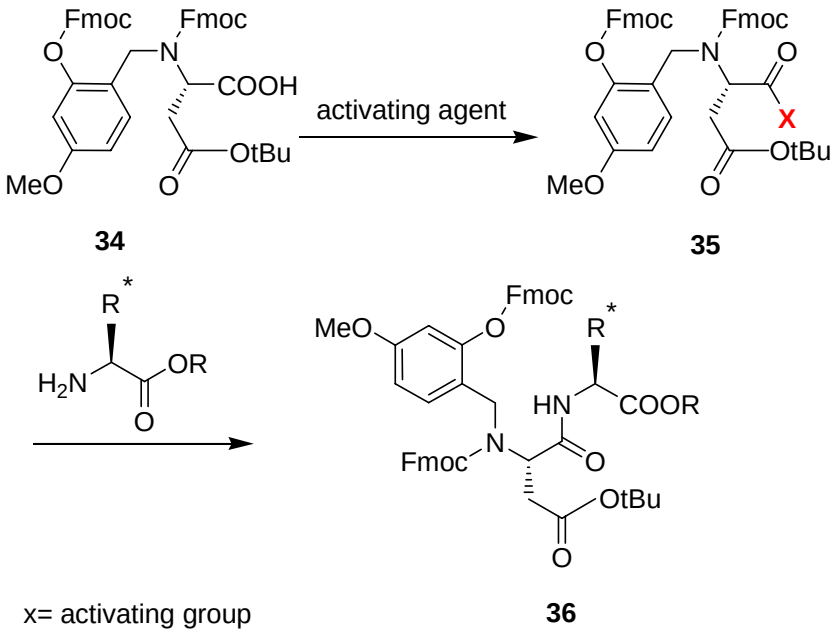
Scheme 3.1 Synthesis of *N*-Hmb-Asp(OtBu)-OH

As the N^{α}-Hmb-Asp(OtBu)-OH **29** is to be incorporated into the growing peptide chain on the solid support, the amino group of **29** requires protection with an appropriate temporary protecting group. Since the 2-hydroxy group is also nucleophilic, it can therefore also be theoretically acylated during the coupling steps. Studies of the mechanism of acylation of amino group with N^{α}-protected Hmb amino acid derivatives revealed a new type of reactive intermediate **31**, which is formed independently of the coupling procedure applied.



Scheme 3.2 4,5-Dihydro-8-methoxy-1,4-benzoxazepin-2(3H)-one derivative intermediate **31** in the activation of N^{α}-Fmoc N^{α}-(2-Hydroxy-4-methoxybenzyl) amino acid **30**.

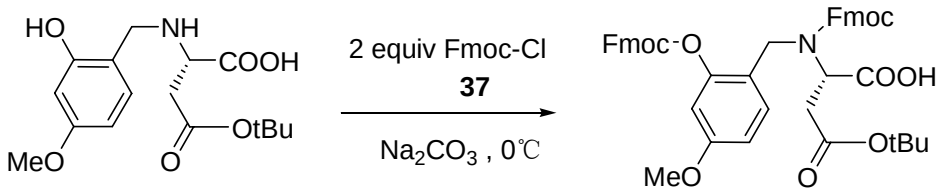
This intermediate was identified as 4,5-dihydro-8-methoxy-1,4-benzoxazepin-2(3H)- one derivative **31** (*Scheme 3.2*), which is believed to be the active component that undergoes aminolysis to produce the desired peptide **33**.[66] The acylating potency of this intermediate is, however, inferior to that of normal active esters and the coupling yields are hence low.[67] Furthermore, if the amino component of the coupling counterpart **32** is hindered and hence the coupling rate would be low, there will be some side reactions of this intermediate such as racemization. To prevent formation of the intermediate lactone **19**, the phenolic group has to be protected. For this very reason, the *N,O*-bis-Fmoc-*N*-Hmb protected amino acid is chosen over the mono-Fmoc derivatives (*Scheme 3.3*). Due to acid lability, the Hmb strategy is preferably applied in Fmoc/tBu chemistry and thus the *N,O*-bis Fmoc derivatives of Hmb amino acid **22** is the choice.



Scheme 3.3 The activation of *N,O*-bisFmoc-*N*-Hmb-Asp(OtBu)-OH **34** and its coupling to another amino acid.

3.5.2.3 Introduction of *N,O*-bis-Fmoc-*N*-Hmb-Asp(OtBu)-OH and Peptide Chain Elongation

N,O-bis-Fmoc-*N*-Hmb protected amino acids **34** are preferably prepared from H-(Hmb)Asp(OtBu)-OH **29** through the reaction with Fmoc-Cl **37** under basic conditions due to its high reactivity with amine (*Scheme 3.4*). The related azide Fmoc-N_3 and *N*-hydroxysuccinimide ester Fmoc-OSu are not sufficiently reactive for this purpose.[66]



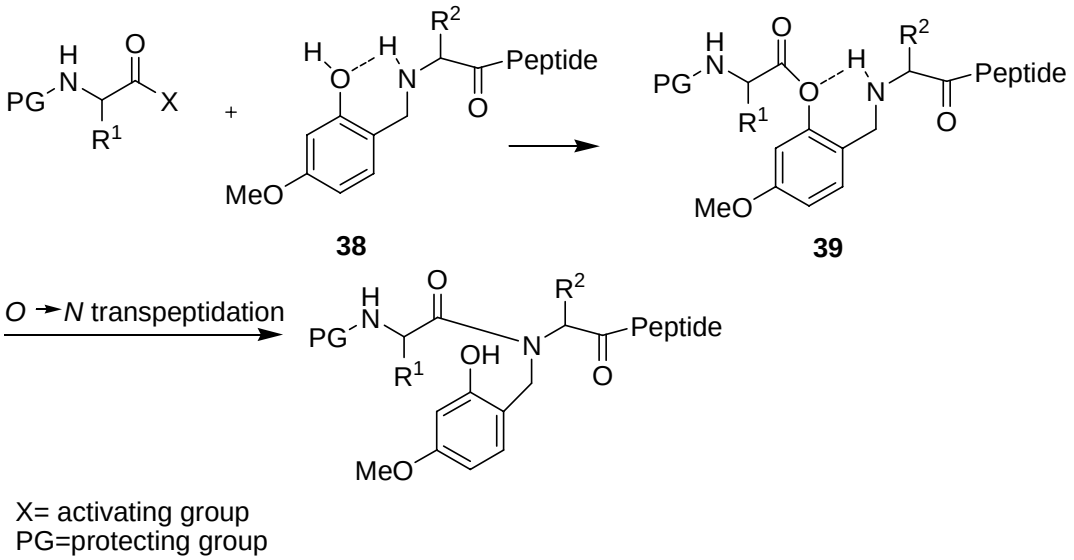
29 **34**

Scheme 3.4 Synthesis of *N,O*-bisFmoc-*N*-Hmb-Asp(OtBu)-OH **34** from *N*-Hmb-Asp(OtBu)-OH **29**

The incorporation of *N,O*-bis-Fmoc-*N*-Hmb amino acid **34** to the growing peptide chain is preferably carried out with the corresponding pentafluorophenyl (Pfp) ester[65, 68] or with carbodiimides.[65] The uronium guanidinium-type coupling reagent such as HATU is also able to fulfill such a task. The coupling efficiency is sequence dependent, as the yields strongly depend on the nature of the side chains. This can be explained by steric hindrance, especially with β-branched amino acids such as Val, Thr, and Ile. This observation suggests that only amino acids with unbranched side chains are suitable for backbone protection,[66] and among these glycine and phenylalanine are particularly recommended.[69] The acylation of *N*-terminal Hmb amino acid residues is more difficult and thus the sterically most unhindered sites are generally preferred for this acylation of N^{α}-Hmb protected amino acid derivative. HATU is a reliable candidate for this acylation reaction. UNCAs (urethane-protected α-amino acid *N*-carboxyanhydride) in a large excess appears also to be a method of choice.[70] Similarly efficient is the use of acyl fluorides.

In the course of *N*-terminal peptide extension, the phenolic group of Hmb is acylated first as it is less hindered than the secondary amino group to form the intermediate **39**.

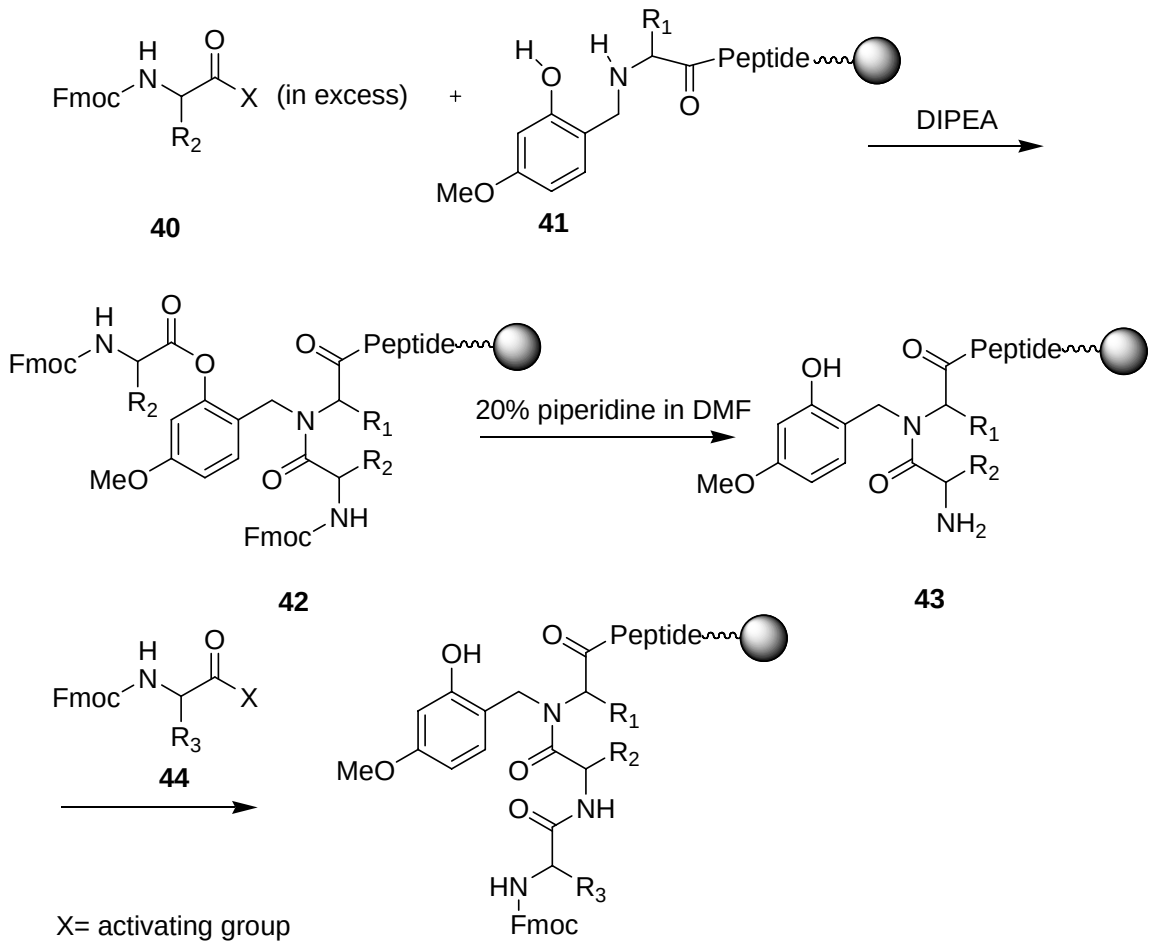
Afterwards, *O* → *N* migration, which is an intramolecular base-catalyzed reaction, takes place in a second step leading to peptide bond formation, as shown in *Scheme 3.5*.



Scheme 3.5 Aminoacylation of N^{α}-(2-Hydroxy-4-Methoxybenzyl) Amino Acid Derivatives **38**

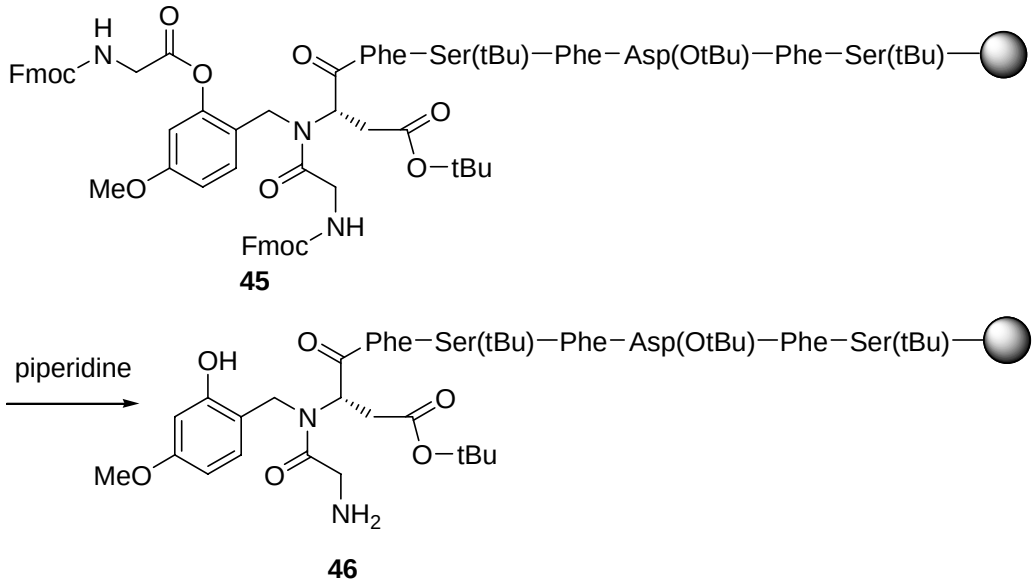
The mechanism of aminoacylation of N^{α}-Hmb-Amino acid derivative **38** serves as an argument that although N^{α}-Hmb-Amino acids or peptides are known to be sterically hindered with respect to aminoacylation, acylation of such an amino acid or peptide derivative proceeds surprisingly well.

In SPPS, since the incoming amino acid **40** and the coupling reagents are used in excess, the *N*-terminal Hmb protected amino acid residue **41** could possibly be acylated at not only the *N*-amino group, but also the 2-hydroxy position to generate the intermediate **42**. However, quantitative cleavage of this ester is assured in the subsequent Fmoc-removal with secondary amine such as piperidine, generating intermediate **43** with the free 2-hydroxy group and free *N*-terminus, which is supposed to be coupled with the next activated amino acid derivative **44** to complete the chain elongation (*Scheme 3.6*). As the Hmb-protected amino acid is not present at the *N*-terminus any more, the acylation of 2-hydroxy position will still take place at a less extent than before, assumably because of the steric hindrance. Nevertheless, it does not represent a problem in the process of chain elongation since the *O*-acylated moiety **42** is quantitatively smoothly cleaved by the standard Fmoc deproctection strategy.



Scheme 3.6 *N,O*-Acylation of *N*-Hmb-peptide in SPPS, subsequent basic cleavage of *N*-Fmoc and *O*-acidyl group, and the following peptide chain elongation.

The argument for the aforementioned mechanism is supported in this project (synthesis of **HP 17**) by the real-time control of the condensation reaction of Fmoc-Gly-OH to the growing peptide H-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-O-Trt(2-Cl)-Resin on the solid support by the correlative MALDI-ToF MS. The MALDI-ToF MS of the crude peptide Fmoc-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH contains signals at m/z = 1806.32, 1821.21, 1827.04 etc.(see *Figure 3.10*), which correspond to the *O*-acylated derivative of the aforementioned peptide (*Scheme 3.7*). While signals at m/z = 1504.20, 1526.15, 1542.07 ect. belong to target peptide Fmoc-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH.



Scheme 3.7 Piperidine treatment of *N,O*-acylation product of the *N*-terminal *N*-Hmb-protected peptide **45**

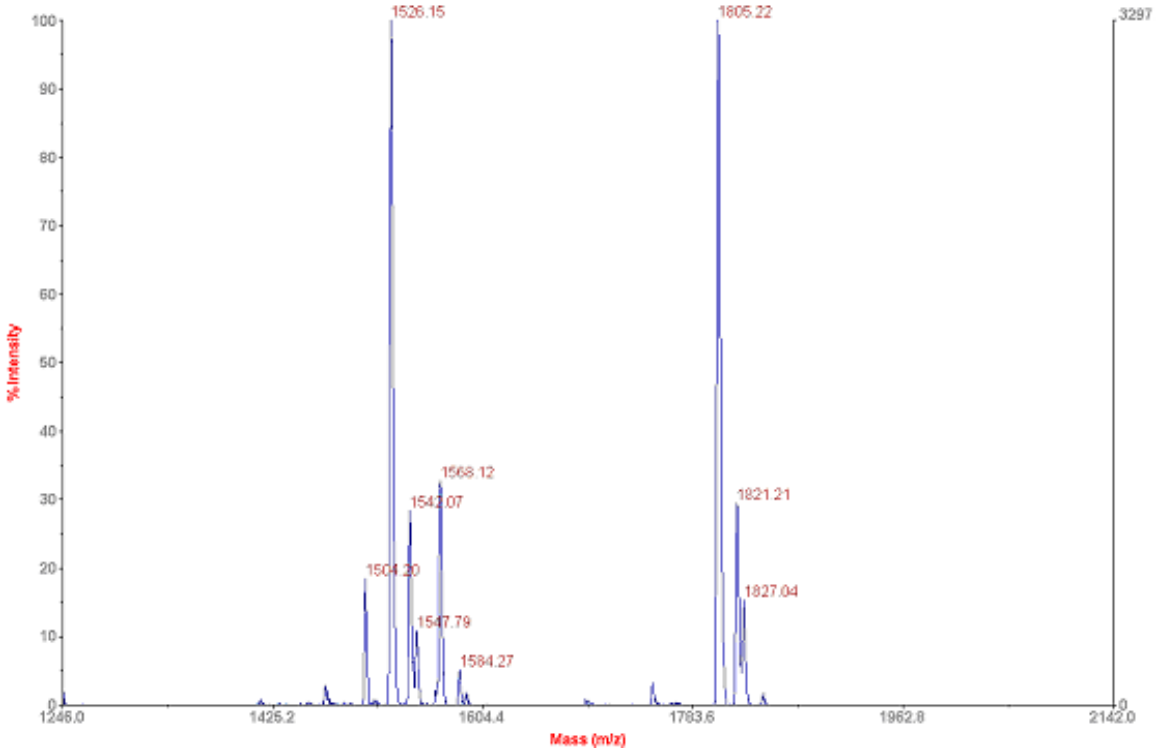


Figure 3.10 MALDI-ToF-Mass Spectrum of Fmoc-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH andd its *O*-acylated derivative.

After treatment of **45** with 20% piperidine in DMF (twice, 5min, 15min, respectively), the Fmoc-glycyl moiety on the 2-hydroxy group of Hmb is removed quantitatively,

yielding the target peptide product H-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-O-Trt(2-Cl)-Resin **46**. MALDI-ToF MS signals at m/z = 1282.61, 1304.80, 1320.78, ect. correspond to H-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH (see *Figure 3.11*).

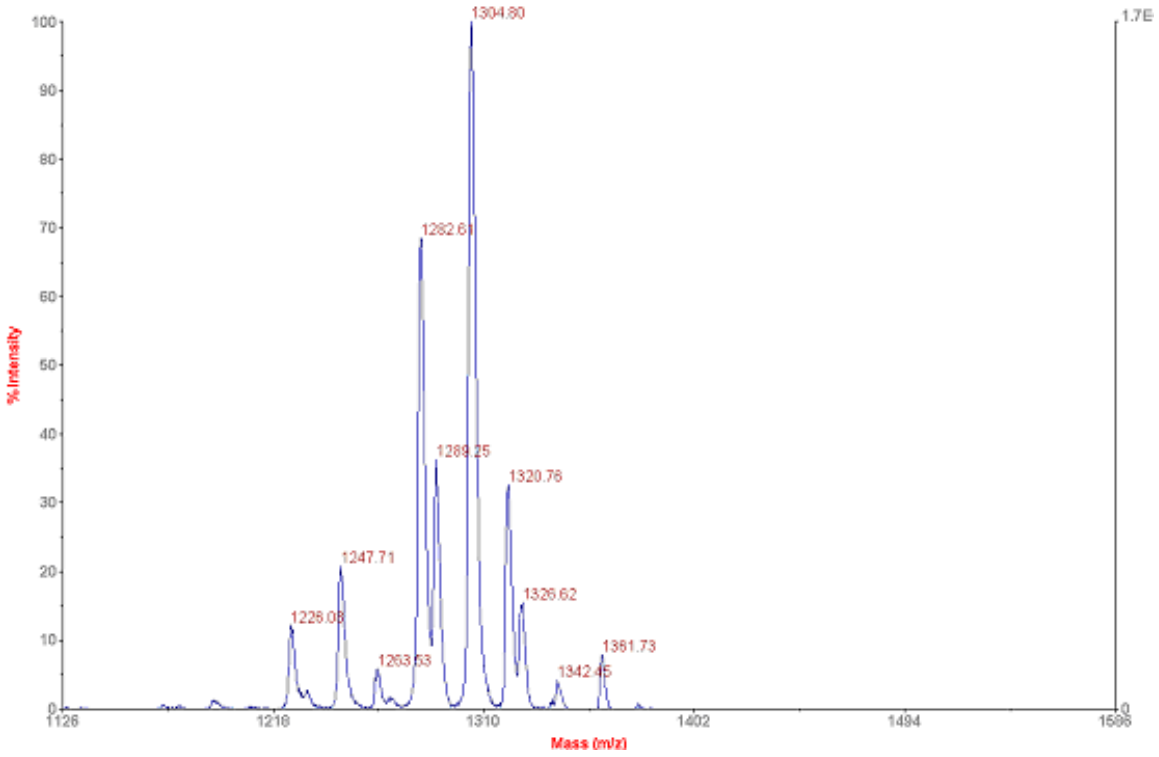


Figure 3.11 MALDI-ToF-Mass Spectrum of H-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH

After acylation of the Hmb-protected amino acid by the subsequent amino acid residue, the peptide chain elongation returned to the "normal track", i.e., the routine coupling uronium reagent such as TBTU could be applied as coupling reagent. It promotes the coupling of amino acid properly. The peptide chain elongation is able to run smoothly until the end in the synthesis of template peptide **HP 17** in this project.

The introduction of Hmb moiety significantly improves the solubility of β-sheet forming peptides in organic solvent such as DMF which is routinely applied in peptide synthesis. Purification of Hmb containing totally protected peptides is feasible through RP-HPLC, which utilizes actonitrile/H_2O or methanol/H_2O as eluent with TFA as ion pairing agent.

Clean peptides are obtainable with this purification method. This can be testified through the combination of MALDI-ToF Mass Spectrometry and analytical RP-HPLC, amino acid analysis and NMR could also be applied if necessary.

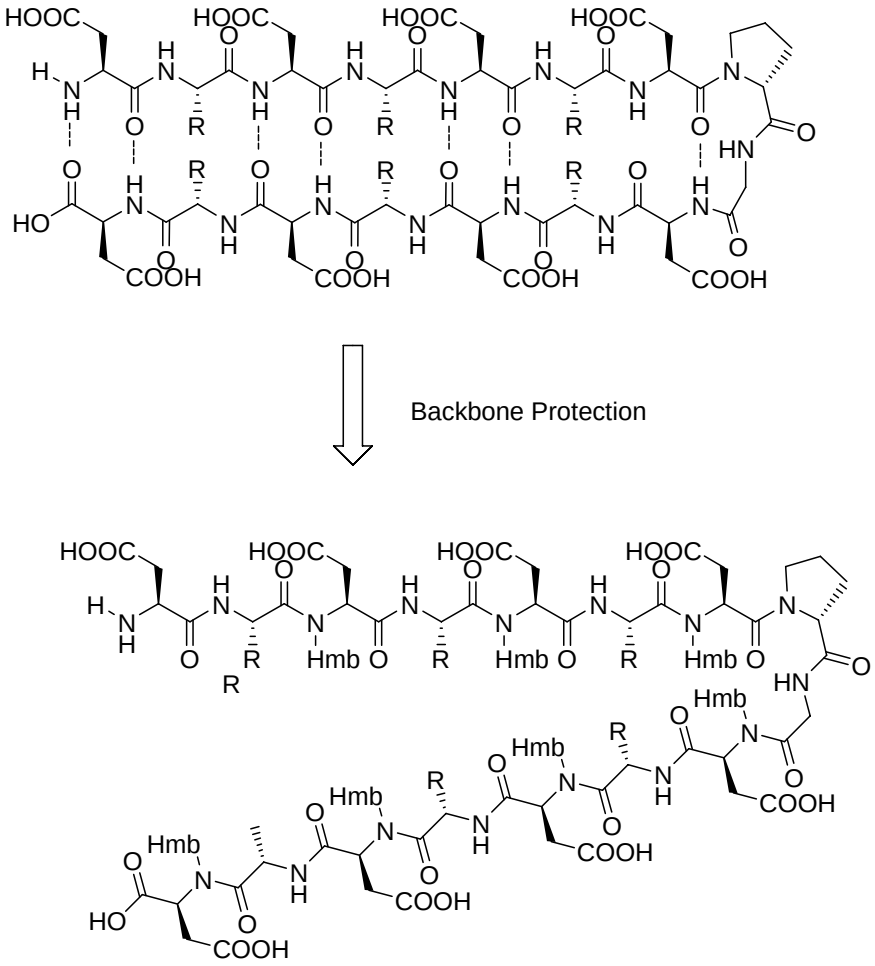


Figure 3.12 Prevention of the formation of intra-chain hydrogen bond in β-hairpin peptide through the introduction of back-bone protection groups Hmb.

The incorporation of Hmb backbone protecting groups significantly increase the solubility of the fully protected peptide in organic solvents such as DMF, probably through breaking of the intra-chain hydrogen bonds between the neighboring strand in the peptide chain (see *Figure 3.12*), thus making it feasible to purify the peptide derivative by RP-HPLC. This allows to obtain the homogenous totally protected peptide as an intermediate which can be subsequently deprotected by Reagent K (TFA/phenol/H_2O/PhSMe/ethane-1,2-dithiol/triethylsilane, 82.5:5:5:5:2.5). The formed

final products could be easily purified without having to apply RP-HPLC, thus circumventing the inherent solubility problem of β-sheet related peptide for final RP-HPLC purification. The feasibility of the Hmb-backbone protection methodology for RP-HPLC purification in this project can be proven by the comparison of the following of chromatograms (see *Figure 3.13*). On the top is the C12 Preparative RP-HPLC chromatogram of a totally protected peptide with the sequence H-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-OH **Prot. HP 19.** Its chromatogram shows a very inferior result with unacceptable resolutions. As the peptide **Prot. HP 19** bears no Hmb backbone protecting group, it shows a very limited solubility in acetonitrile/H_2O mixed solvent which is applied as the eluent of RP-HPLC, so that most of the peptide **Prot. HP 19** precipitates on the chromatographic column during the purification process. The remaining material is not eluted simultaneously and thus generating a broad front or tail on the separation, impairing the resolution as well as selectivity of the RP-HPLC. A homologous peptide with similar polarity, while being protected at some certain amide groups on the backbone, with the sequence H-Asp(OtBu)-Phe-Ser(tBu)-Phe-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH **Prot. HP 17-Hmb**, was also analized by RP-HPLC for comparison.

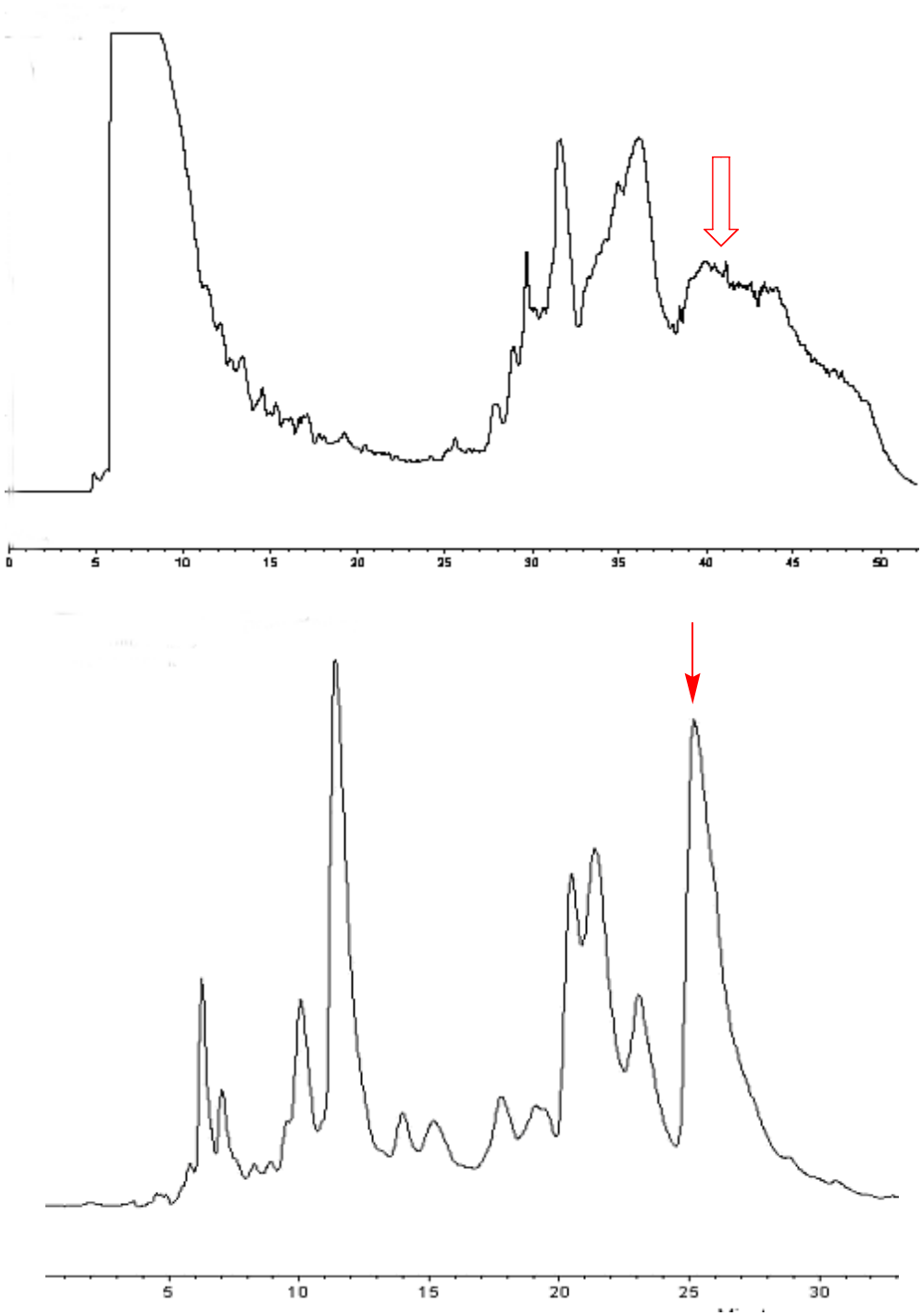


Figure 3.13 RP-HPLC of H-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-OH **Prot. HP 19** (top) and H-Asp(OtBu)-Phe-Ser(tBu)-Phe-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe- Asp(OtBu)-Phe-Ser(tBu)-OH **Prot. HP 17-Hmb** (bottom) (arrows point to target products)

The bottom chromatogram in *Figure 3.13* reflects the HPLC purification of the selectively backbone-protected peptide with the sequence H-Asp(OtBu)-Phe-Ser(tBu)-Phe-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH **Prot. HP 17-Hmb**. The otherwise so problematic HPLC purification was easily carried out upon the introduction of Hmb protecting group into the peptide sequence at some certain positions. It significantly increases the solubility of the peptide derivative, making it feasible to apply RP-HPLC to obtain the utmost purification. The last peak in the chromatogram represents the target peptide, its homogeneity could be testified through the combination of MALDI-Tof-Mass Spectroscopy (see *Figure 3.14*) and analytical RP-HPLC (see *Figure 3.15*).

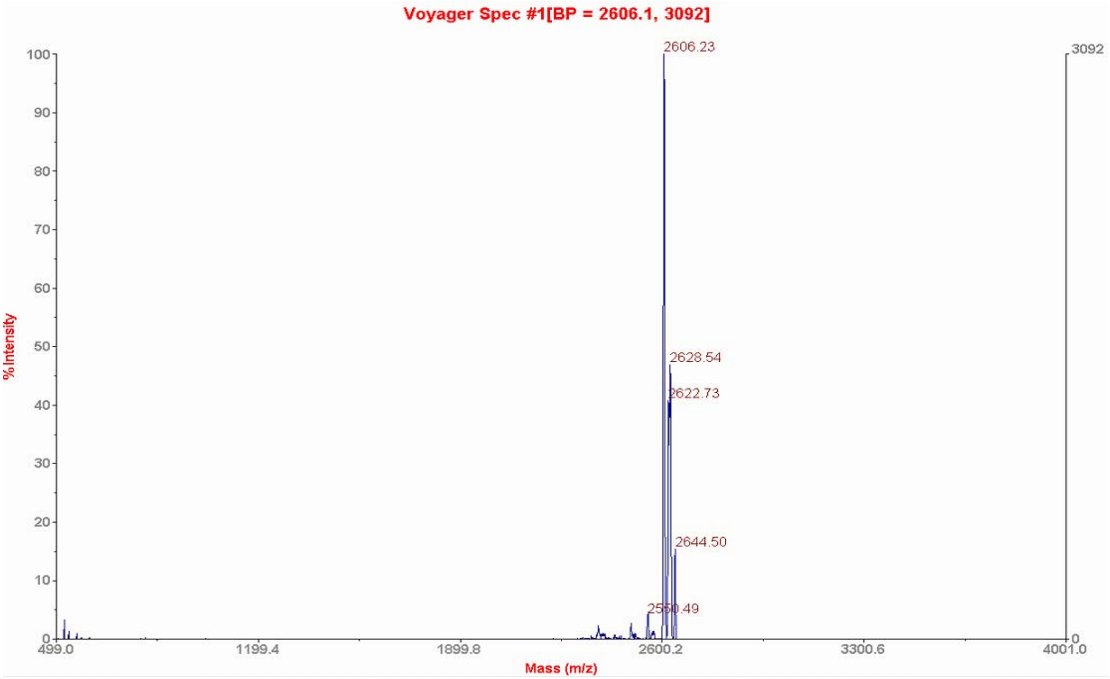


Figure 3.14 MALDI-ToF-Mass Spectrum of **Prot. HP 17-Hmb**

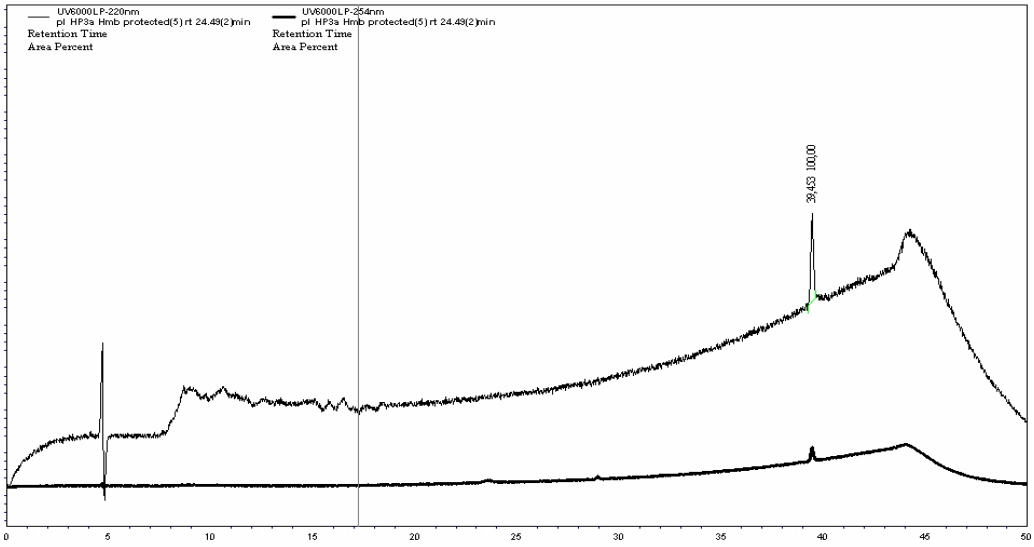
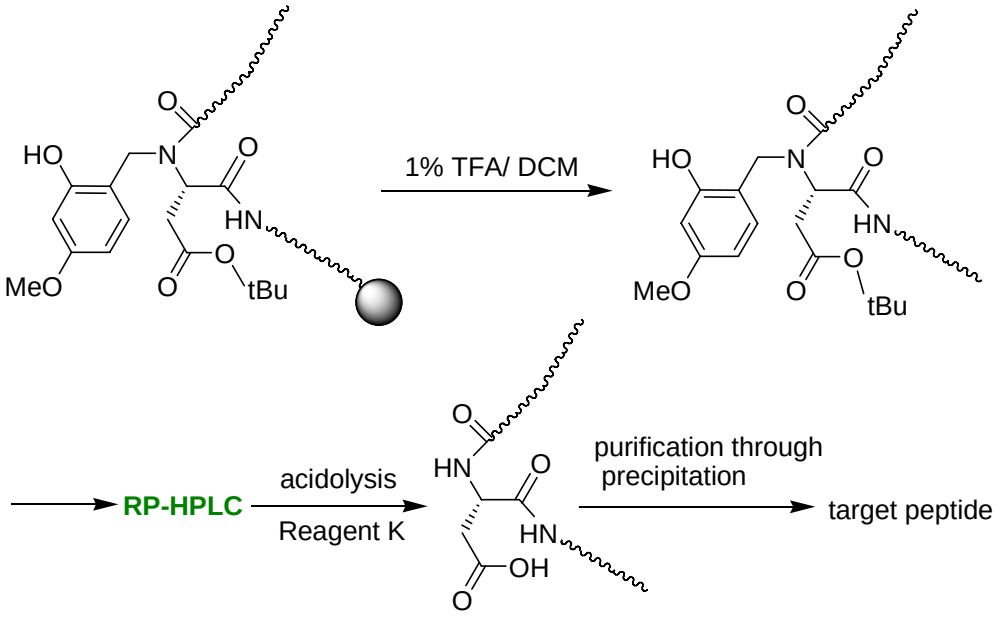


Figure 3.15 Analytical RP-HPLC chromatogram of **Prot. HP 17-Hmb**

The cleavage of Hmb moiety proceeds smoothly. The presence of the methoxy group at the 4-position of 2-hydroxy benzyl moiety significantly increases the acid lability of the benzyl-based protecting group. Cleavage of the Hmb groups can be performed with TFA, although care has to be taken to avoid alkylation of tryptophan residues, since the carbocations generated in the acidolytic cleavage are highly electrophilic and hence are able to attack the nucleophilic functionalities of the peptide. The indole side chain of Trp, thiol side chain of Cys, and imidazole side chain of His are especially prone to be attacked by these carbocations. An efficient cocktail for acidolytic removal of all acid-labile protecting groups, Reagent K (TFA/phenol/H_2O/PhSMe/ethane-1,2-dithiol, 82.5:5:5:5:2.5) is highly recommended (*Scheme 3.8*). The addition of alkyl silane such as triisopropylsilane or triethylsilane to the deprotection cocktail is also advisable. The obtained peptides could be easily purified due to the huge polarity difference of the fully deprotected peptide and the organic by-products, thus bypassing the RP-HPLC purification which might practically bring forward the previously discussed problems.



Scheme 3.8 Cleavage of the totally protected peptide from the resin, RP-HPLC purification and acidolytic deprotection.

Figure 3.16 Synthetic process of DFSFDFSpGDFSFDFS **HP 17** with backbone protecting methodology

Figure 3.16 shows the flow chart of the process of SPPS of DFSFDFSpGDFSFDFS **HP 17**by the Hmb back-bone protecting methodology.

3.5.2.4 Synthesis of Peptides Other Than HP 17

All peptides in this project (see also Table 3.5) except **HP 17** were synthesized according to routine Fmoc/tBu peptide chemistry, either by manual or microwave methodology (see also **Chapter 2.6.2**).

3.5.2.4.1 Yields Comparison of Manual and Microwave SPPS

HP 20 was utilized as template peptide for the comparison of manual and microwave SPPS in terms of synthetic efficiency for "difficult sequence". **HP 20**, with the sequence EFSFEFSpGEFSFEFS, adopts partial β-hairpin conformation. It thus presents a good reference to compare the efficacy of manual and microwave peptide synthesis. The referred manual and microwave methodology is introduced in **Chapter 2.6.2**. Yield of crude fully protected **HP 20** (lyophilized, without HPLC purification) with manual and microwave synthesis is 73.6 % and 98.0 %, respectively. Yield of fully deprotected **HP 20** (with HPLC purification) with manual and microwave synthesis is 2.32 % and 3.40 %, respectively. It can be concluded that the introduction of microwave irradiation is beneficial for the peptide with "difficult sequence" such as **HP 20**. Seemingly, microwave irradiation applied in the coupling as well as Fmoc removal step destroys the intrachain hydrogen bonds between the neighbouring β-strands, which promotes the smooth elongation of the peptide chains on resin and in turn suppresses the truncated sequences and enhances the yield of target peptide.

3.5.2.4.2 Aspartimide Formation Comparison of Manual and Microwave SPPS

Peptide **HP 15**, with the sequence DFDFDFEpGEFDFDFD, was used as a reference for the study of influences of microwave irradiation on aspartimide formation during SPPS. It was proven that the introduction of microwave irradiation during the coupling and Fmoc removal steps in SPPS could promote the aspartimide formation compared with manual synthesis in which microwave irradiation is absent. This conclusion was validated in the synthesis of **HP 15** DFDFDFEpGEFDFDFD and **HP 11** FDFDFDFD. No

aspartimide formation derivatives were found in the manual SPPS of these two peptides. While the aspartimide formation derivatives in the fully protected **HP 15** product reached 35.38 % in the microwave SPPS, with 20 % piperidine in DMF containing no HOBt as Fmoc deprotecting reagent. It was also proved that the addition of 0.1 M HOBt into 20 % piperidine/DMF as Fmoc deprotecting reagent could reduce the extent of aspartimide formation, despite not being able to fully suppress such a side reaction. In a parallel microwave SPPS of **HP 15**, in which all the parameters remained intact while 0.1 M HOBt was utilized as additive in 20 % piperidine/DMF, the extent of aspartimide formation derivatives in the fully protected **HP 15** was reduced to 15.36%.

3.5.3 Conformational Studies by CD Spectrometry

3.5.3.1 Theory Introduction of Secondary Structure Determination in Peptides

CD is a method of choice for the quick determination of protein and peptide secondary structure, both quantitatively and qualitatively. CD spectrometry is so sensitive that it can distinguish between the secondary structure such as α-helix, β-sheet, turn and random coil. Although CD spetrometry is to some extent limited in deciphering the structures of peptides compared with NMR or X-ray diffraction, CD data is instructive as preliminary guide to peptide and protein conformation and conformational transition under a wide range of conditions.[71]

The amide bonds in peptide and protein are the most important chromophores to be observed by CD spectroscopy. The elctronic transitions of amide chromophore have been characterized. The proportions of secondary structure elments α-helix, β-sheet, turn and random coil can be determined by their contributions to the curves in CD spectra of peptides or proteins. A typical α-helix is normally characterized by a negative band at 222 nm (n-π^*), a negative band at 208 nm, and a positive band at 190 nm (π-π^*). The α-helix formation is evidently length dependent. Normally short peptides do not establish stable helices in solution.[72] It had been shown that halogenated alcohol such as 2,2,2-trifluoroethanol (TFE) is capable of stabilizing helix conformation in some peptides and is hence regarded as helical inducer.[73] β-sheet is relatively less well-defined compared with α-helix,[74] and can be formed either in parallel or antiparallel manner. β-sheets display a typical negative band at 215 nm (n-π^*) and a positive band at 198 nm (π-π^*). Random coil conformations (unordered structure) are normally characterized by a strong negative band below 200 nm. Several algorithms are compiled for the computational calculations of proportions of secondary structures in peptides and proteins by fitting the observed spectrum with the combination of the characteristic absorption of the aforementioned three basic secondary structures.

3.5.3.2 Solvent Choice in CD Spectroscopy

3.5.3.2.1 Halogenated Alcohol

Halogenated alcohols such as TFE and HFIP are widely applied solvents for macromolecules in CD spectroscopy, because of their good transparency properties in the CD measurements, excellent solubilizing properties for some otherwise hardly soluble peptides such as β-sheet derivatives, and their inclination to stabilize the α-helix conformations in peptides.[73] Structures containing turns, β-hairpins, β-sheets, and hydrophobic clusters are also observable in the presence of TFE or HFIP.[75]

Halogenated alcohol could exert substantial influences on peptide conformation by the enhancement of hydrogen bonds, the disruption of water structure and lessening of the hydrophobic effect, and preferential solvation of certain groups of the peptide chain.[75]

Effect on Hydrogen Bond

TFE/HFIP is better hydrogen bond donor but a poorer hydrogen bond acceptor compared with water.[76] It has therefore been assumed that TFE/HFIP preferentially bind to oxygen atoms of the main chain carbonyl groups, serving as a hydrogen bond donor in establish the hydrogen bonds with the main chains of the peptide (carbonyl group is often able to form two hydrogen bonds). This preference of solvation by TFE/HFIP in peptide leads to enhanced intra-chain hydrogen bonding of the amide group to carbonyl since the solvent exposure of amide is markedly decreased. In other words, hydrogen bonds between the amide groups from peptide and hydroxyl groups from TFE/HFIP are weakened because TFE/HFIP is relative poorer hydrogen bond acceptor compared with water, which leads to a less extent of solvation of amides. They are thus able to form hydrogen bonds with suitable carbonyl groups on the backbones of peptide, resulting in stabilized secondary structure such as α-helix. This consensus is regarded generally as the direct influence of halogenated alcohol to stabilize the local α-helix conformations in peptides. Indirect effect on hydrogen bond could arise from the altered dielectric constant of alcohol/water mixtures. As the solvent dielectric constant is lowered from 79 of water to 27 of TFE (25 °C) electrostatic interactions become stronger.[77]

Effect on Non-polar Sidechains

Several studies[78,79] have led to the conclusion that TFE/HFIP preferentially associate with hydrophobic sites on the protein surface and emphasized the hydrophobic nature of the CF_3 group. TFE/HFIP could be modelled as mainly weakening hydrophobic-hydrophobic interactions.[80] The ability of TFE/HFIP to lessen the hydrophobic interactions between residues distant in the primary sequence of peptide accounts for ability of cosolvent to disrupt structures such as β-sheet or β-hairpin, in which hydrophobic interactions play crucially important role to stabilize the local conformation.

Effect on Solvent Strucuture

The stability of secondary structures of peptides/protein is thought to be affected by solvent structure through hydrophobic interactions. It has been proven that water structure is affected by TFE/HFIP.[81,82] In water rich mixtures, both CH_3 and CF_3 groups of the alcohols are hydrophobic, enhancing neighboring water-water interactions. It is noticed that many of the peptides that have been studied in TFE and form the most stable helices are amphiphilic in nature.[83,84] It is therefore assumed that TFE/HFIP may selectively associate with the hydrophobic surfaces of helix or sheet, thus mimicking the environments of membrane or the interior of a protein.

Effect on Turns

Turn peptides frequently contain prolines, which serves as a structure breaker for β-sheet as well as α-helix. TFE/HFIP is not able to stabilize structure in certain proline containing peptides.[75] Turn formations in longer peptides are frequently associated with the peptide chain folding back on itself, forming either helix-turn-helix or hairpin conformations. Such association is usually driven by hydrophobic interactions between residues distant in the primary sequence, which are thought to be weakened by the presence of TFE/HFIP.[85-88]

Effect on β-sheet and β-hairpin Conformation

The structures of β-sheet and β-hairpin are frequently stabilized by intra/inter-strand hydrophobic interactions, favorable electrostatic interactions, as well as hydrogen bonds

between the backbones of peptide chains. The hydrogen bonding pattern between strands seems to be entirely dependent on its own sequences. A frame shift of one residue has dramatic sequence for the pattern of hydrogen bonds, corresponding to 180° flip of one of the strands so that entirely different set of residue sidechains interact in the plane above and below the hydrogen bonded main chain.[75] It is assumed, however, that the effect of TFE/HFIP on structures of β-sheet and β-hairpin may not be mediated solely through hydrogen bonding. It should be specifically applied to individual peptide. The stability of β-sheet/β-hairpin conformation is synthetically determined by the interplay of hydrogen bonding, hydrophobic interactions, as well as electrostatic interactions. The role of TFE/HFIP in the determination of β-sheet/β-hairpin should be thus analyzed in a compositive manner, taking the weight of aforementioned factors into consideration.

3.5.3.2.2 Water Titration

Water is known to possess the inclination to break the intra/inter-chain hydrogen bonds in peptides, thus destroying local secondary structure such as α-helix. Since the forces to stabilize the β-hairpin conformation is relatively elusive to analyze, the role of water to stabilize/destabilize this structure is to be explored to some certain peptides, instead of drawing universal conclusions. Despite that water molecule could destroy the intra-chain hydrogen bonds between the neighboring backbones of β-hairpin, leading to unfolding of the local secondary structure, it could on the other hand strengthen the hydrophobic interactions in the β-hairpin peptides to reinforce this conformation, and lessen the electrostatic interactions which could be unfavorable with respect to β-hairpin conformation. The ultimate result that water brings forward to the conformation alternation of β-hairpin peptide should be in compositive manners. It depends on the fact that which factor is in the lead role. The titration of peptide HFIP solution with water is thus a choice to explore the weight of different forces in the determination of the β-hairpin conformation.

3.5.3.2.3 pH Dependence

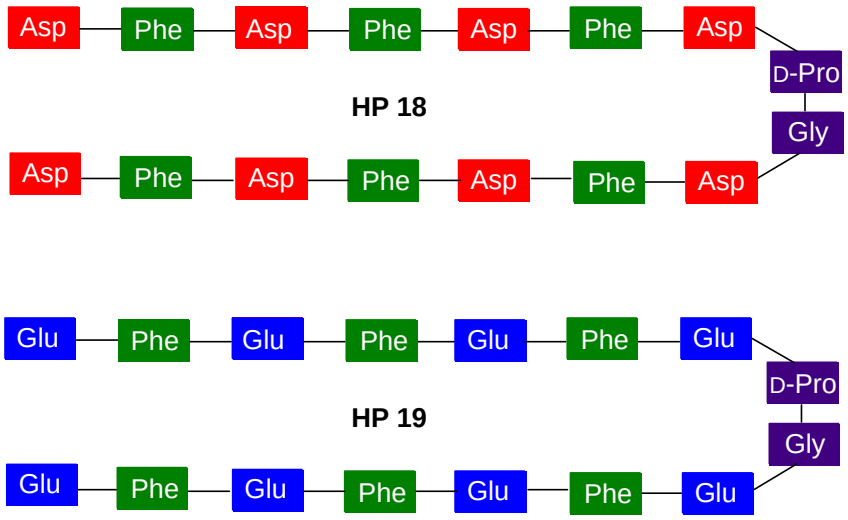
For the peptides which contain ionizable residues, electrostatic interactions would rise to an important role in determination of the secondary structures. Under this circumstance,

the pH value of the environment would be a decisive parameter with respect to β-hairpin conformation. The protonation/deprotonation of ionizable side-chain could possibly stabilize or destabilize the local conformation. The structure conversion with the change of pH value of the solution is therefore instructive to valuate the factors that determine the peptide conformation.

3.5.3.3 CD Spectroscopy Analysis

The peptide was accurately weighed and the solution were carefully preparted for the CD measurement. The parameters of CD spectrometer were adjusted according to the description of General Methods in the experiments part (see also **Chapter 2.6.1**). The measured CD values were converted to molar ellipticity of mean amino acid residue ($[\theta]_{MRW}$) (*vide infra*). The measurements were performed at the room temperature.

Conformation of H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 18) and H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19) in HFIP



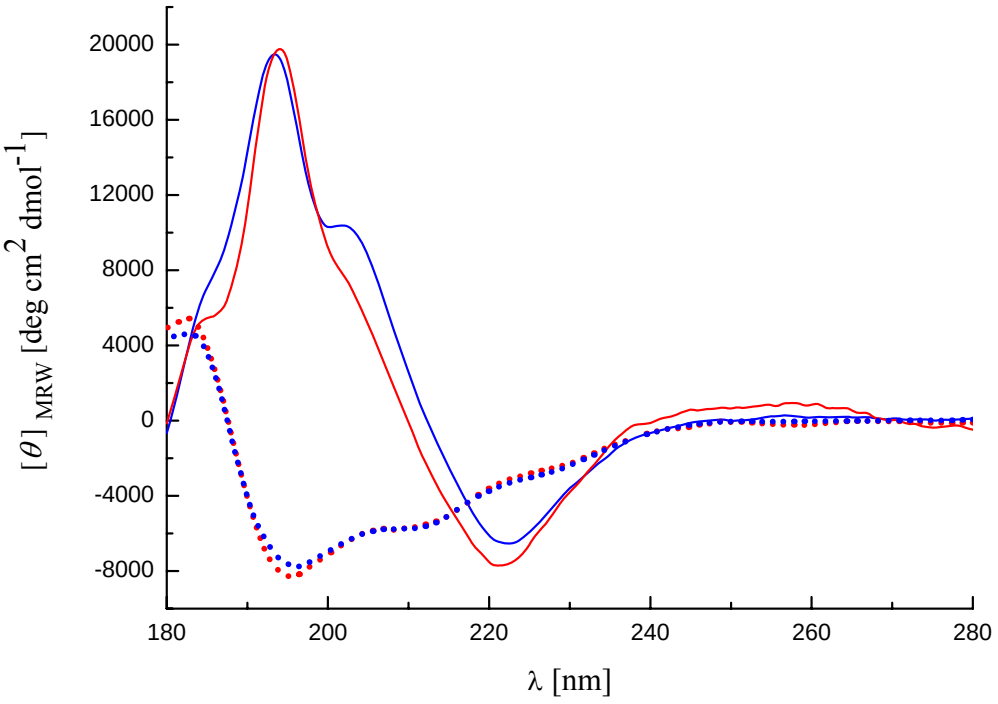


Figure 3.17 CD spectra of HFIP solution of **HP 18** (1mg/ml, blue, dotted line; 0.3 mg/ml, red, dotted line) and **HP 19** (0.3 mg/ml, blue, solid line; 0.15 mg/ml, red, solid line)

The CD spectrum of HP 18 in HFIP indicates the predominance of random coil conformation, with negative band at 195 nm and positive band at 183 nm (see *Figure 3.17*), which correspond the characteristic band of random coil. It did not show any positive band at wavelength above 190 nm, validating this conclusion. Interestingly, **HP 18** did not adopt the targeted β-hairpin conformation which was not in accordance with the design.

In contrast to the CD spectrum of peptide **HP 18**, peptide **HP 19** adopted mainly a β-hairpin conformation in HFIP, with the typical positive band at 194 nm and the negative band at 221 nm and comparable molecular ellipticities, as was shown in *Figure 3.17*. Furthermore, the increase of the concentration of **HP 19** from 0.15 mg/ml to 0.3 mg/ml stabilized the β-hairpin conformation, exhibiting an additional positive band at wavelength 202 nm, which is characteristic of a β-turn. Despite the close analogy between **HP 18** and **HP 19**, they adopt totally different conformations in HFIP.

H_2O Titration of H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19)

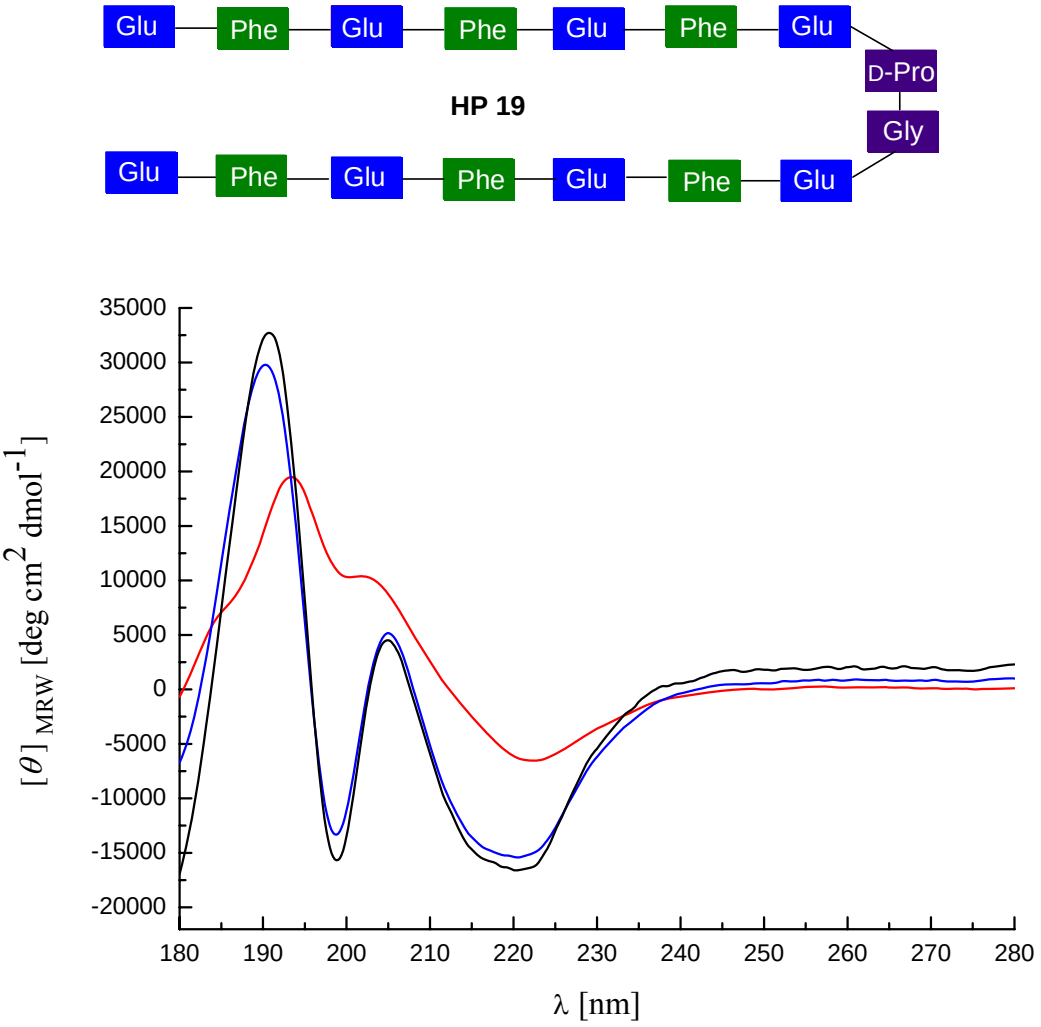


Figure 3.18 H_2O titration of **HP 19** peptide solution in HFIP: 0.3mg/ml **HP 19** in HFIP (red); 0.15mg/ml **HP 19** in HFIP/H_2O (1:1) (blue); 0.075mg/ml **HP 19** in HFIP/H_2O (3:1) (black)

As is shown in *Figure 3.18*, with the addition of water to HFIP solution of peptide **HP 19**, the positive band at 194 nm shifted to 190 nm and the intensity increased significantly. Furthermore, the positive band at around 205 nm representing β turn was also intensified. It shows that the increase of the water concentration stabilizes the β hairpin conformation

of peptide **HP 19** in the HFIP, presumably through the strengthening of the hydrophobic interactions between the side chains of phenylalanines on the neighboring strands. Another possible explanation is that water titration increase the dielectric constant of the mixture solvent, thus lowering the electrostatic interactions between the side chains of glutamate which destabilize the β-hairpin conformation.

pH Dependence of H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19)

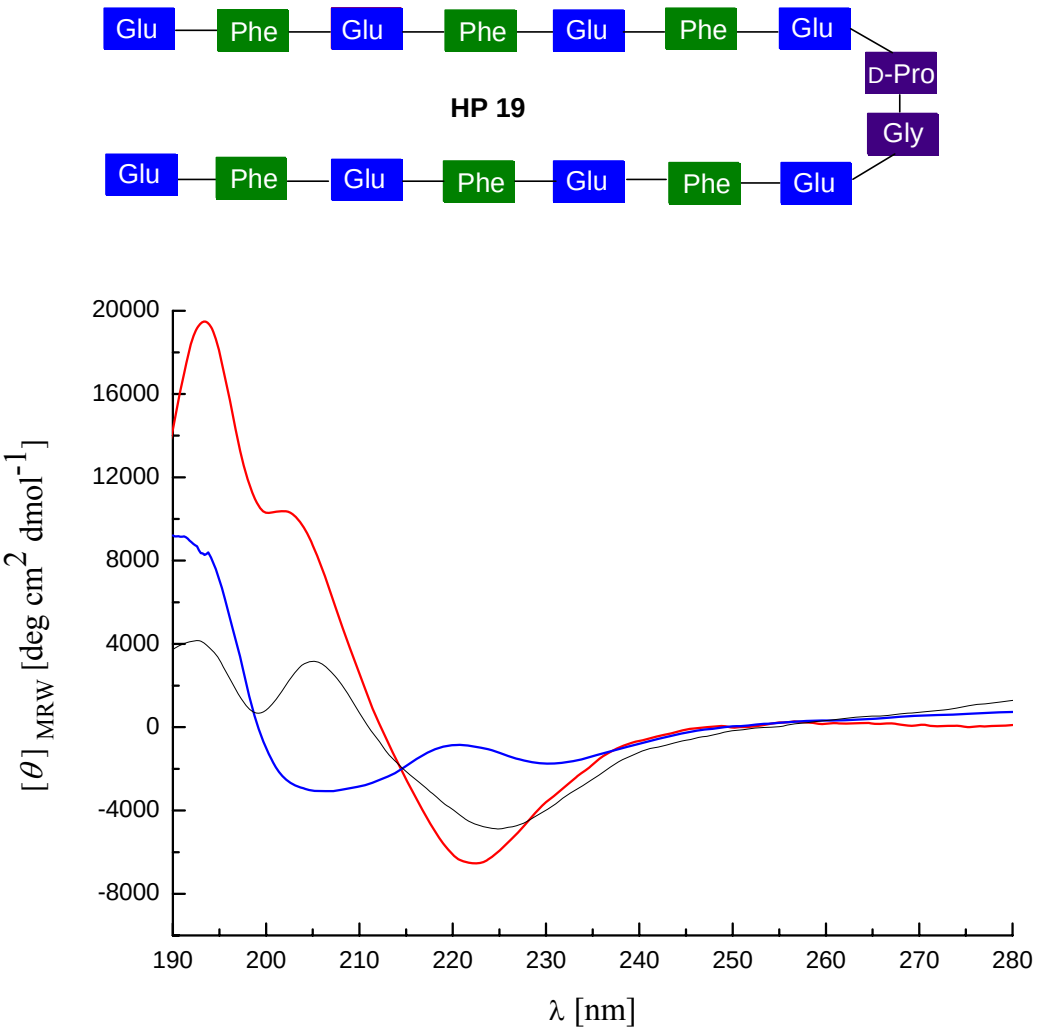


Figure 3.19 pH dependence of the conformation of peptide **HP 19**: 0.3 mg/ml **HP 19** in HFIP (red); 0.5 mg/ml **HP 19** in 2 M NaOH/HFIP (1:1) (blue); 0.25 mg/ml **HP 19** in 0.2 M HCl/HFIP (1:1) (black)

In order to examine the influence of pH on the conformation of **HP 19**, solutions with different pH values were examined. As shown in *Figure 3.19*, the addition of 0.2 M HCl to the solution of **HP 19** in HFIP decreased the intensities of the positive bands at the two typical wavelength 194 nm and 221 nm. The β-hairpin conformation of **HP 19** was destabilized to some extent by the addition of acid, while its conformation bascially remained as the β-hairpin. Since the addition of acid protonates the side chain carboxy groups and partially the backbone carbonyl group, the intra-chain hydrogen bonds between the backbone carbonyl functionalities and corresponding backbone amide groups were blocked. However, the influences of the side chain-side chain aromatic interaction as well as hydrophobic interaction seemed to exceed that of the backbone intra-chain hydrogen bond. On the contrary, the β-hairpin conformation of peptide **HP 19** was totally denaturized upon the addition of 2 M NaOH solution to the peptide HFIP solution. The pK_a of the glutamate side chain carboxyl group in is approximately 4.5.[89] Consequently, most of them will be deprotonated with the addition of 2 M NaOH. The repulsive interactions between the deprotonated carboxylate side chains of glutamate thus exceeded the factors that stabilize the β-hairpin conformation, leading to the denaturation of the peptide and loss of the original conformation. It showed that the electrostatic interactions could be significantly important for β-hairpin conformation, and under some circumstances they could exceed the forces such as hydrophobic interactions and hydrogen bond as in this experiment. In addition, the blocking of backbone hydrogen bonds by acid could lead to the destabilization of β-hairpin conformation.

Water Titration of H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 18) in HFIP

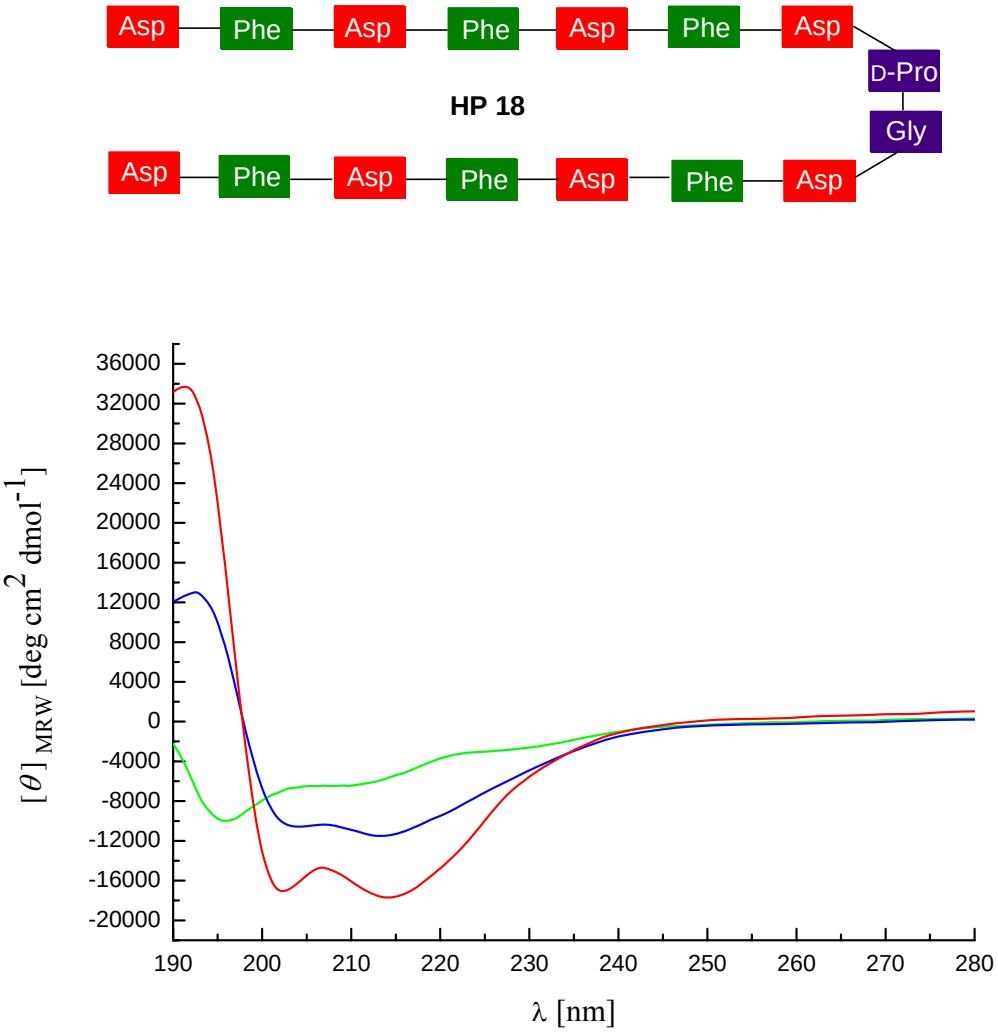


Figure 3.20 CD spectra of H_2O titration of **HP 18** in HFIP: 0.5 mg/ml **HP 18** in HFIP (green); 0.5 mg/ml **HP 18** in HFIP/H_2O (1:1) (blue); 0.33 mg/ml **HP 18** in HFIP/H_2O (1:2) (red)

As shown in *Figure 3.20*, it is obvious that upon the addition of water to the HFIP solution of peptide **HP 18**, a change of conformation took place. **HP 18** adopts mainly random coil conformation in pure HFIP, while with the increase of water concentration, signs of β-strand appeared, shown by the positive band at around 195 nm. The higher the percentage of water, the more stable the β-strand conformation was. The negative band at

around 200 nm is the characteristic of random coil, which showed that, although the β-strand conformation was stabilized by the addition of water, the inherent primary sequence of **HP 18** determined that it could not be converted to highly ordered secondary structure as **HP 19**.

H-Glu-D-Pro-Gly-Glu-OH (HP 1), H-Asp-D-Pro-Gly-Asp-OH (HP 2) in HFIP, H_2O and 0.02 M NaOH

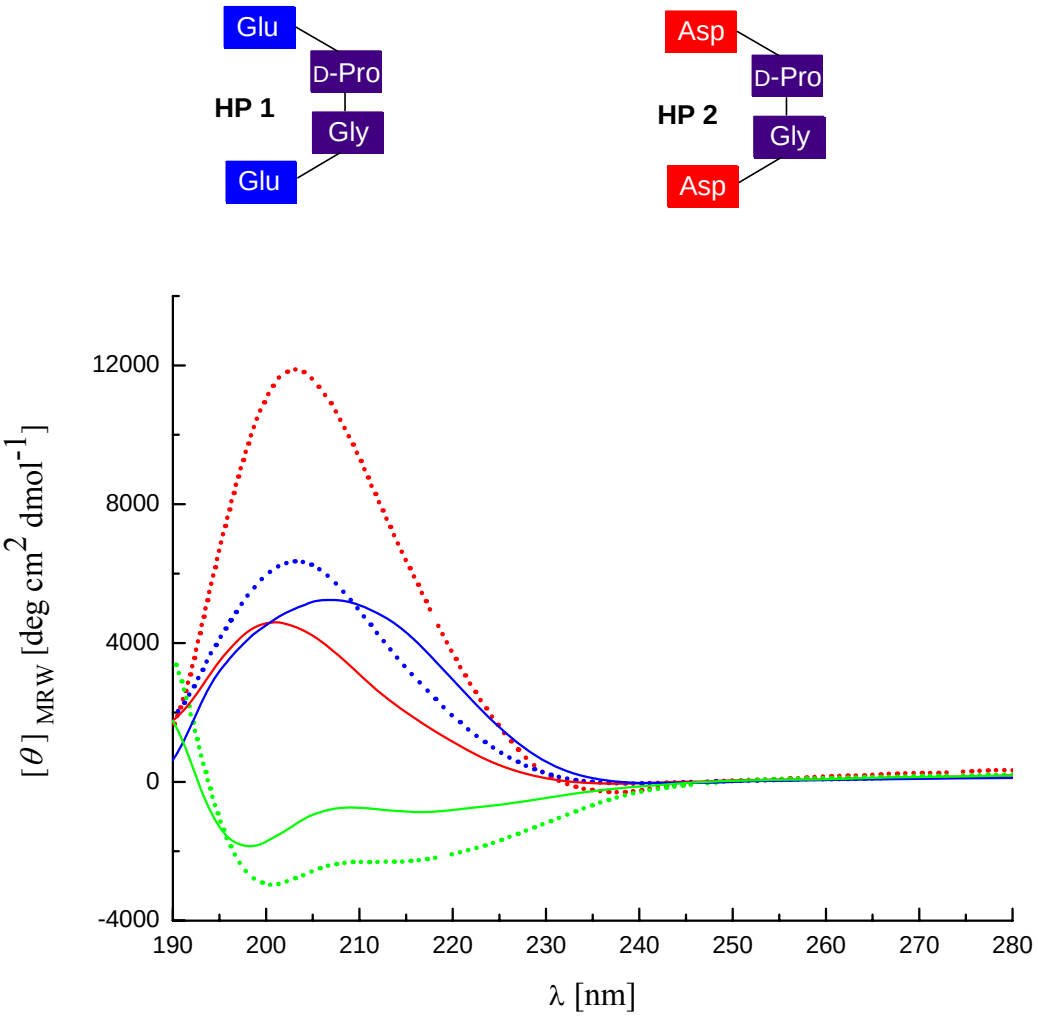


Figure 3.21 Solvent and pH dependence of the conformation of **HP 1** and **HP 2**: 2mg/ml peptide **HP 1** in H_2O (red, dotted line); 2mg/ml peptide **HP 1** in 0.02 M NaOH (blue, dotted line); 2 mg/ml peptide **HP 1** in HFIP (green, dotted line); 2 mg/ml peptide **HP 2** in H_2O (red, solid line); 2 mg/ml peptide **HP 2** in 0.02 M NaOH (blue, solid line); 2 mg/ml peptide **HP 2** in HFIP (green, solid line)

HP 1 and **HP 2**, with the sequences H-Glu-D-Pro-Gly-Glu-OH and H-Asp-D-Pro-Gly-Asp-OH, occupy the turn regions of peptide **HP 19** and **HP 18**, respectively. If cut from the parent peptides, they still adopt turn structures in H_2O, while in HFIP, the conformation shift to random coil, as is shown in *Figure 3.21*. These results were in accordance with the findings in the literature (see also Chapter **3.4.5.2**). [75] Despite the fact that HFIP is able to enhance the strength of intra-molecular hydrogen bonds, the stronger electrostatic interactions between the side chains of aspartic acid, which is caused by the low dielectric constant of HFIP, exceed the strength of the enhanced hydrogen bond interactions generated by HFIP. It seemed therefore that the repulsive electrostatic interactions between the side chains of Asp or Glu played a more important role in destabilization of the ordered structures of these peptides. Upon the addition of 0.02 M NaOH to the aqueous solution of peptide **HP 1**, the stability of the peptide was shown to be impaired, while peptide **HP 2** seems to be relatively indifferent to the pH change of the environment.

Conformation of H-Phe-D-Pro-Gly-Phe-OH (HP 4) in H_2O

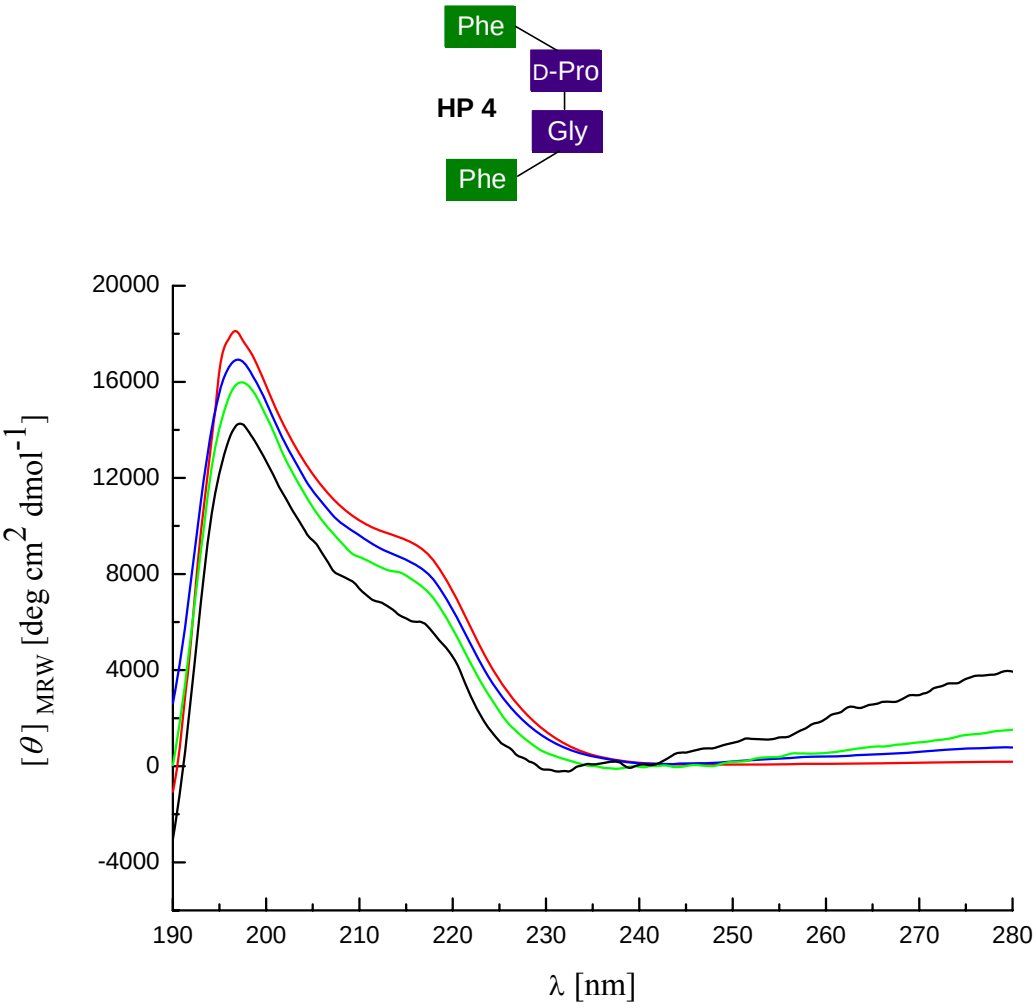


Figure 3.22 CD spectra of **HP 4** in H_2O: 2 mg/ml (red); 0.5 mg/ml (blue); 0.25 mg/ml (green); 0.1 mg/ml (black)

HP 4 adopted a turn structure in H_2O, as is shown in *Figure 3.22*, exhibiting a positive band at ~215 nm, thus bearing both the characters of turn and β-sheet. The interactions between the side chains of phenylalanine introduced the extra stabilizing factors besides the backbone hydrogen bonds. The conformation of **HP 4** remained untouched also for very small concentration of 0.1 mg/ml. The aromatic interactions seemed to be a key factor in stabilizing the β-hairpin conformation.

Conformation of H-Phe-D-Pro-Gly-Phe-OH (HP 4) in HFIP

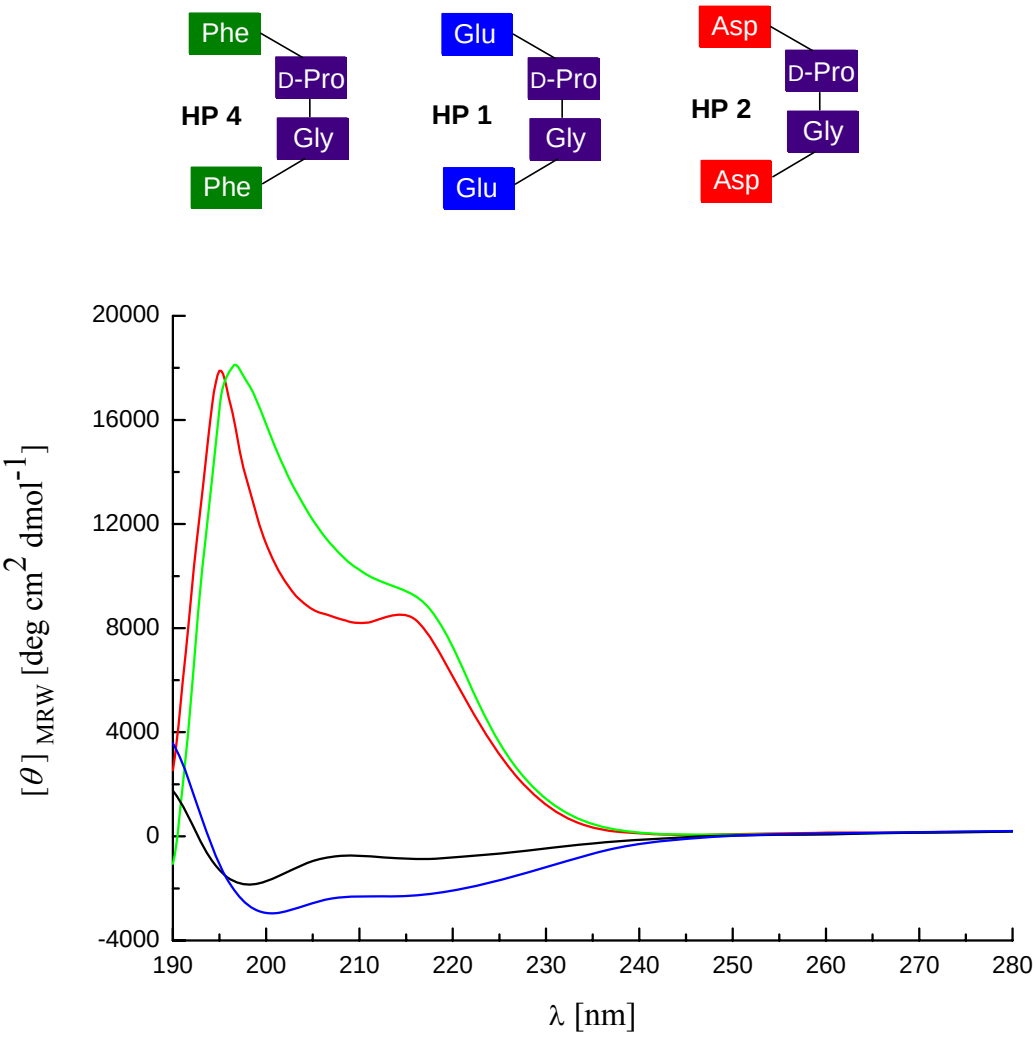


Figure 3.23 CD spectra of **HP 4** in HFIP (2 mg/ml, red) and H_2O (2 mg/ml, green) in comparison with **HP 2** (2 mg/ml, in HFIP, black) and **HP 1** (2 mg/ml, in HFIP, blue)

As shown in *Figure 3.23*, the conformation of **HP 4** was bascially indifferent to the solvent change from H_2O to HFIP, which are both polar but with different dielectric constants and thus diverse capacity to function as hydrogen bond acceptor or donor. The aromatic interactions between the side chains of phenylalanine dominated the formation of the β-turn and quasi-hairpin since the major difference between **HP 4** and **HP 1/HP 2** results from the replacement of electrostatic interactions between Asp/Asp or Glu/Glu by

aromatic interactions between Phe. They outweighed the influence of hydrogen bonds since the weakening of backbone hydrogen strength caused by addition of water did not lead to the destabilization of the β-hairpin conformation in **HP 4**. Differently, **HP 1** and **HP 2**, in which the aromatic interactions were absent, adopted distinctively different conformations in H_2O and HFIP, namely, β-turn structure in water and random coil in HFIP (see *Figure 3.20*). The absence of hydrophobic interactions in these peptides bestows the backbone hydrogen bonds crucial roles in determining the secondary structure stabilities. Under this circumstance, the solvent conversion could exert evident influences on the peptide conformation.

Conformation of H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in H_2O

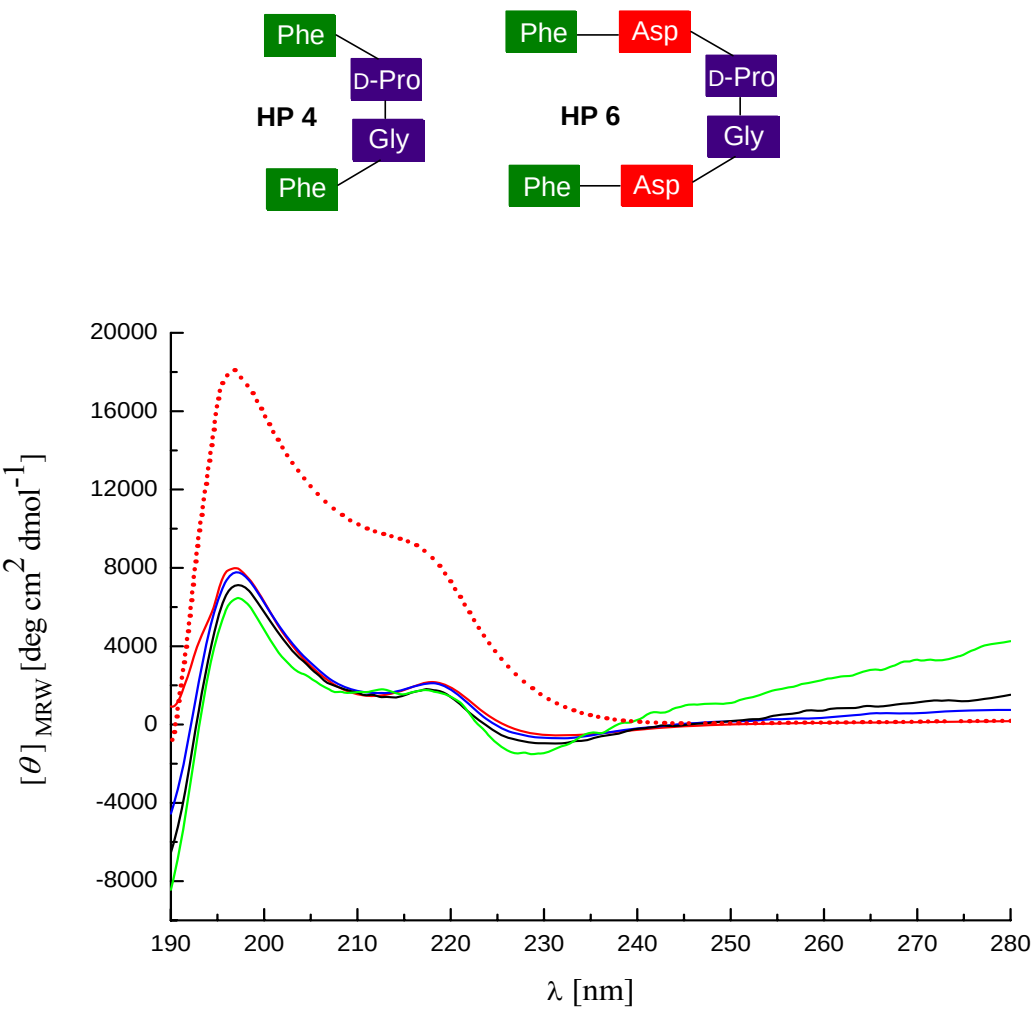


Figure 3.24 CD spectra of **HP 6** in H_2O: 2 mg/ml (red, solid line); 0.5 mg/ml (blue); 0.25 mg/ml (black); 0.1 mg/ml (green) in comparison with **HP 4** in H_2O (2 mgl/ml, red, dotted line)

As shown by *Figure 3.24*, **HP 6** adopted basically a β-hairpin conformation in HFIP. However, its conformational stability in comparison to **HP 4** was obviously weaker, as its ellipticity was markedly lower. The insertion of two aspartic acid residues compared with **HP 4** destabilized its β-hairpin conformation, since the repulsive ionic interactions (glutamate and aspartate are partially deprotonated in HFIP and water) between the side chains of aspartic acids were introduced. This experiment indicated that the "corner residues", i.e. the residues in position *i* and *i*+3 positions of a β-turn, played crucial roles

to stabilize the β-hairpin conformation. The residue swap at these positions could thoroughly alternate the pattern of hydrogen bond of the β-hairpin peptide (see also Chapter **3.4.5.2**), thus leading to different stability of the conformation.

Conformation of H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in HFIP

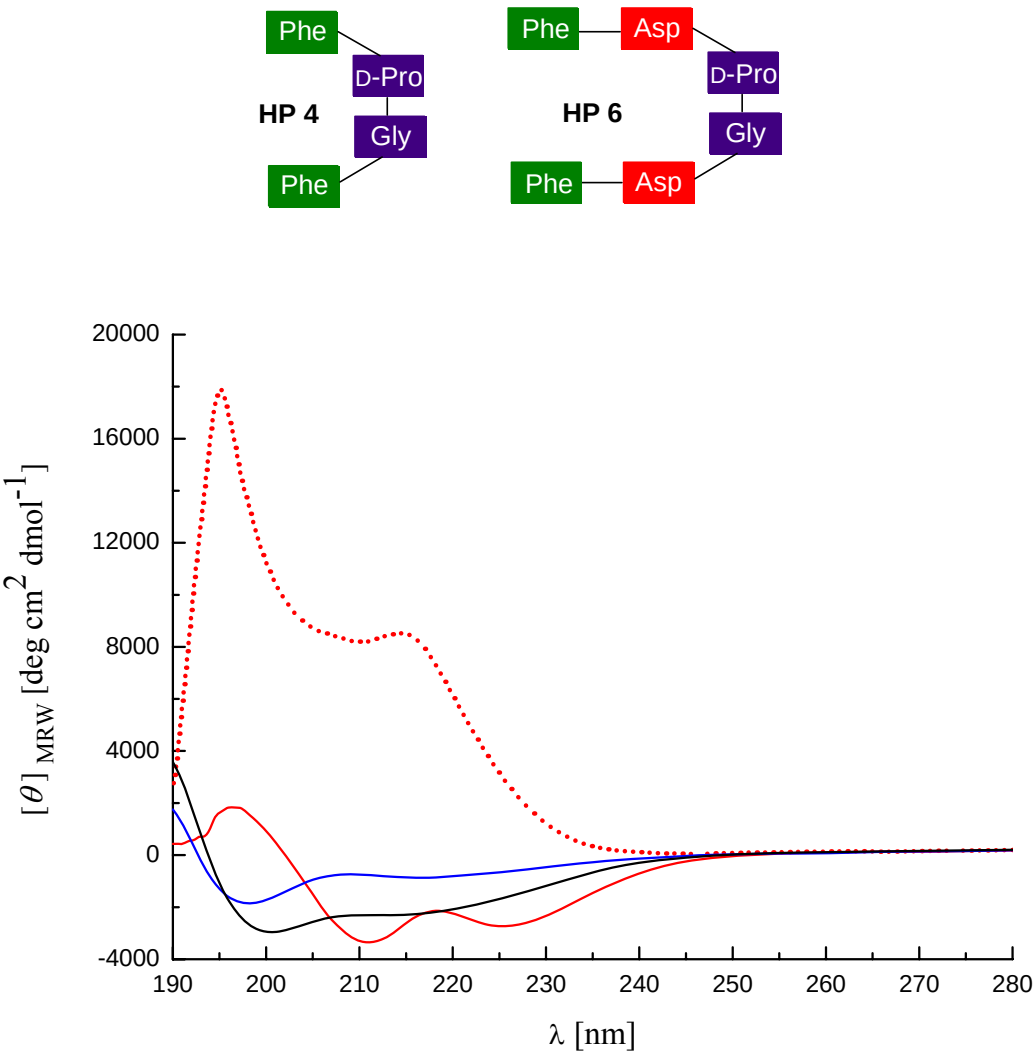


Figure 3.25 CD spectra of **HP 6** in HFIP (2 mg/ml, red, solid line); **HP 4** in HFIP (2mg/ml, red, dotted line); peptide **HP 2** in HFIP (2 mg/ml, blue); peptide **HP 1** in HFIP (2 mg/ml, black)

The contribution of the β-hairpin conformation of **HP 6** in HFIP was also inferior to that of **HP 4**, as was the case in water, shown in *Figure 3.25*. The β-hairpin conformation

proportion in the former was far more abundant than that of the latter, which could be observed from their ellipticities in CD spectra. On the other hand, the introduction of phenylalanine residues into the sequence brought forward the extra aromatic interactions which stabilized the β-hairpin conformation compared with **HP 1** and **HP 2**, in which aromatic interactions were absent, despite that the corner positions in **HP 6** were also occupied by aspartic acids which caused the repulsive electrostatic interactions inside the peptide molecule in HFIP or water.

Comparison of the Conformation of H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in H_2O and HFIP

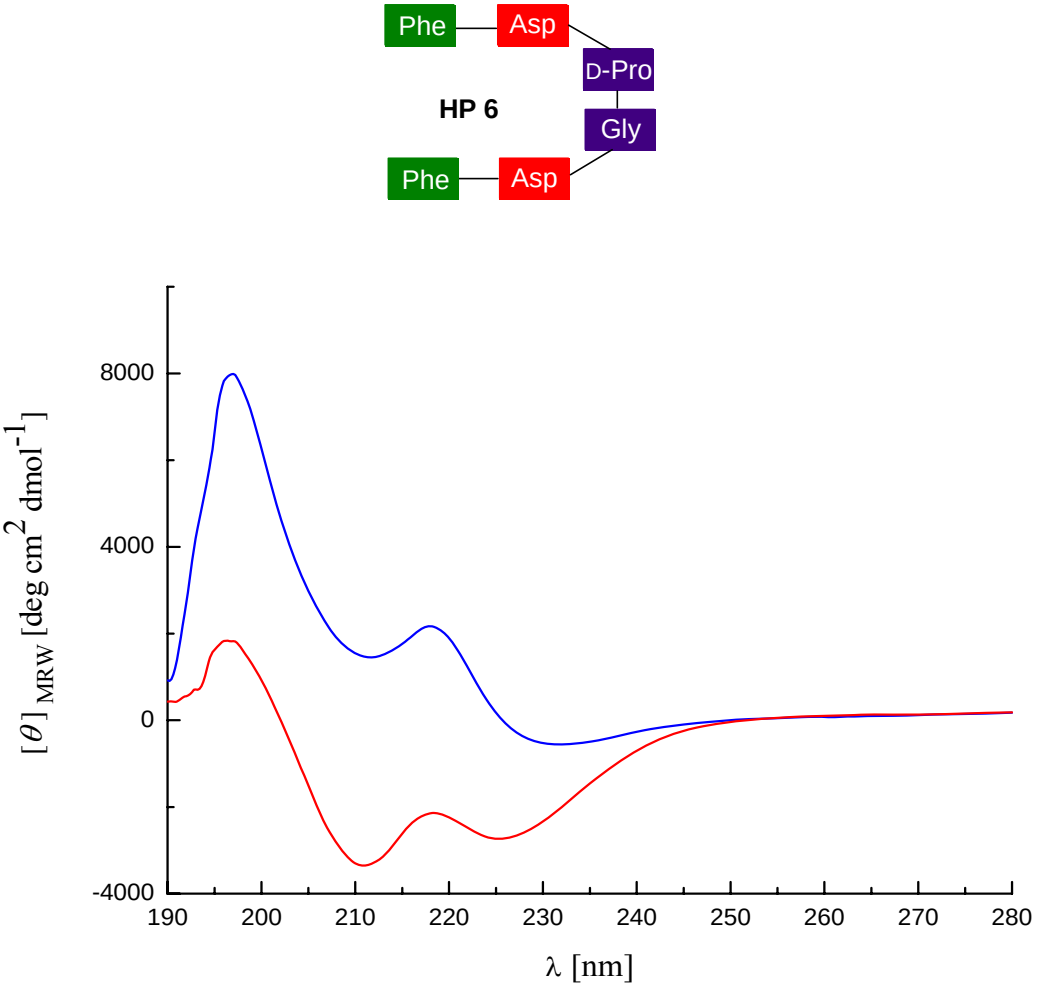
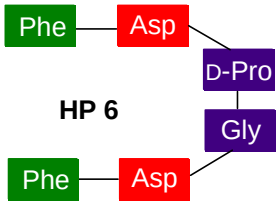


Figure 3.26 CD spectra of **HP 6** in H_2O and HFIP: **HP 6** in H_2O (2 mg/ml, blue); **HP 6** in HFIP (2 mg/ml, red)

The conformation of **HP 6** in HFIP was partially impaired compared with that in H_2O, as the positive band at ~218nm was no longer observed and a negative band was generated instead, which was shown in *Figure 3.26.* It indicated that the ordered β-hairpin conformation was partially unfolded and the conversion to random coil was taking place. This experiment again validated the conformational dependence of peptide on solvent, since different dielectric constant environment would exert diverse fluences on the forces that stabilize/destabilize the secondary conformation of peptides (see also Chapter **3.4.5.2**).

H_2O Titration of of H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in HFIP



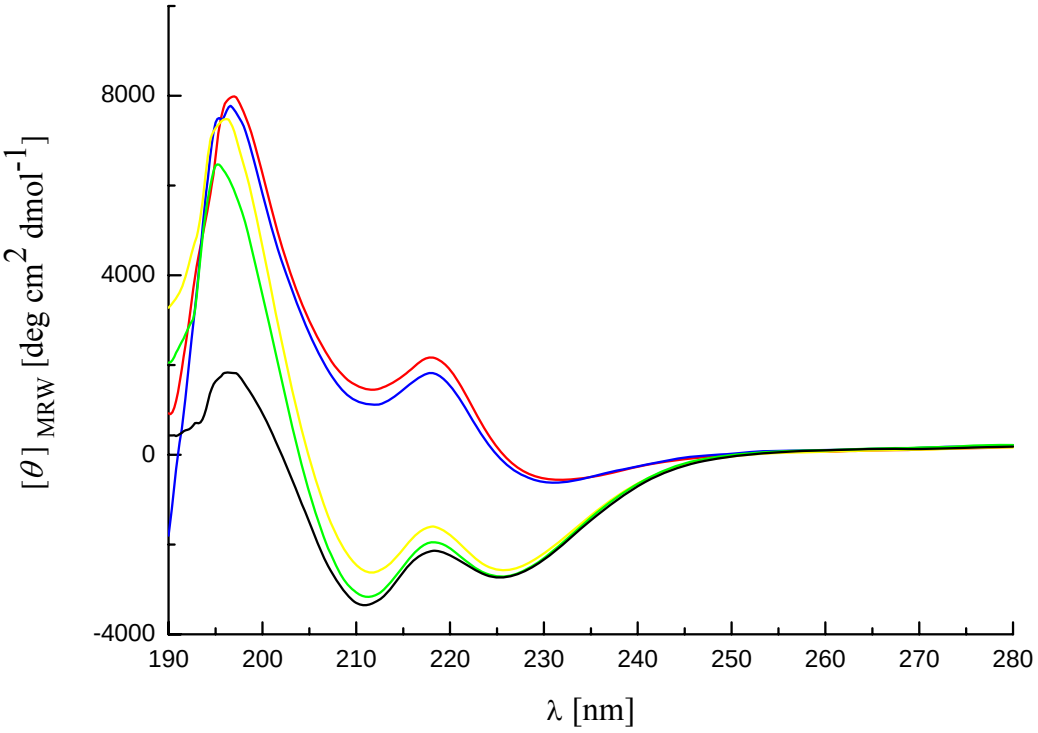


Figure 3.27 CD spectra of **HP 6** H_2O titration in HFIP: **HP 6** in H_2O (2 mg/ml, red); **HP 6** in H_2O/HFIP (19:1) (2 mg/ml. blue); **HP 6** in H_2O/HFIP (3:1) (2 mg/ml, yellow); **HP 6** in H_2O/HFIP (1:1) (2 mg/ml, green); **HP 6** in HFIP (2 mg/ml, black)

Upon the addition of water to the HFIP solution of **HP 6**, the proportion of β-hairpin conformation was gradually increased, as shown in *Figure 3.27*. **HP 6** underwent the major conformational change at a content of water between 75% and 95%. *Figure 3.26* showed again that despite that HFIP molecules were able to strengthen the intra-chain hydrogen bonds between the two strands in **HP 6**, the stronger electrostatic interactions between the side chains of aspartic acid which was caused by the low medium dielectric constant of HFIP relative to water exceeded the strength of the enhanced hydrogen bonds. Consequently, the β-hairpin conformation was stabilized upon the addition of water under this balance which favored the weakening of electrostatic interactions.

pH Dependence of H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6)

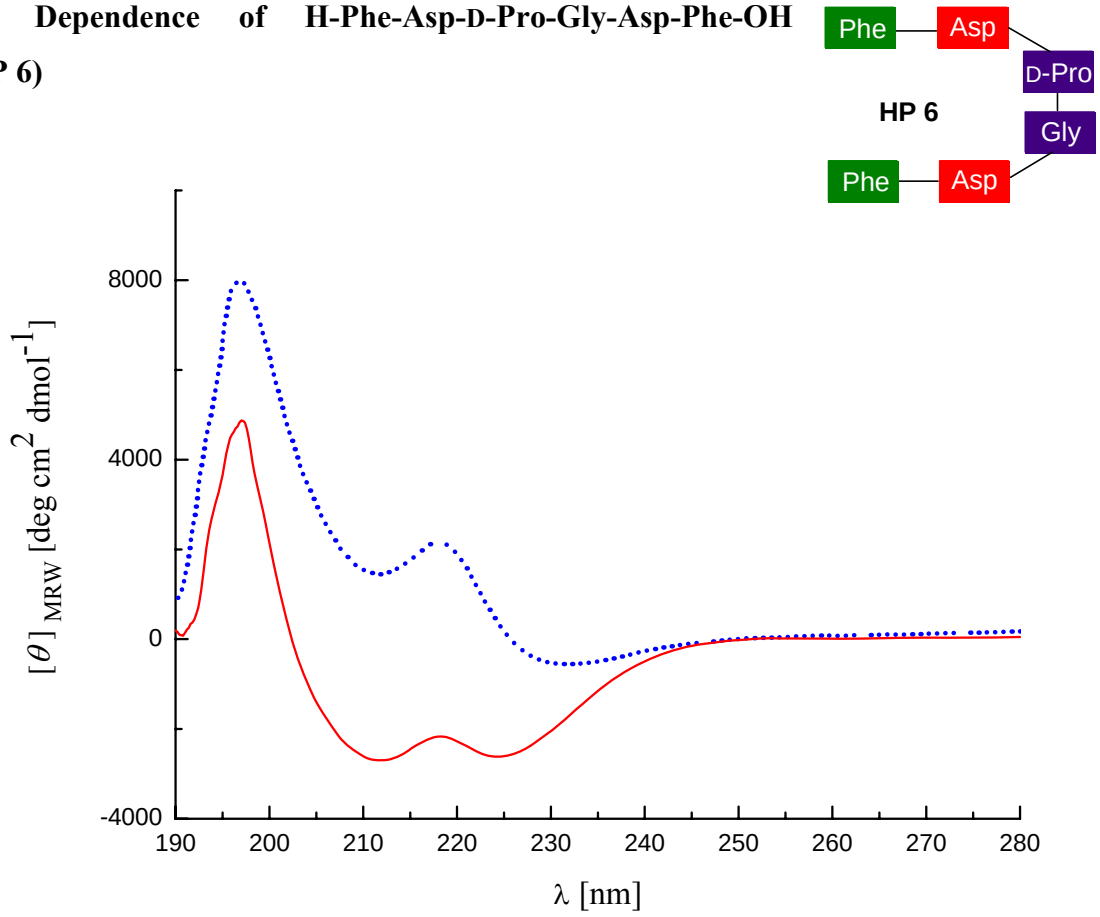


Figure 3.28 CD spectra of **HP 6**: **HP 6** in H_2O (2 mg/ml, blue, dotted line); **HP 6** in 1 M NaOH (2 mg/ml, red)

In the 1 M NaOH solution of peptide **50**, the carboxylic groups on the side chains of aspartic acid were deprotonated. Consequently, the increased repulsive ionic interaction between the side chains destabilized the orderd conformation of **HP 6** under the basic conditions, as shown in *Figure 3.28*. The positive band at ~218 nm disappeared and shifted to the negative band upon the addition of NaOH. In addition, the ellipticity of the positive band at ~197 nm was significantly decreased. This experiment indicated that the repulsive ionic interactions between the aspartate side chains led to the partial unfolding of the ordered β-hairpin conformation.

Comparison of the Conformation of H-Asp-Phe-D-Pro-Gly-Phe-Asp-OH (HP 7) and H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in HFIP

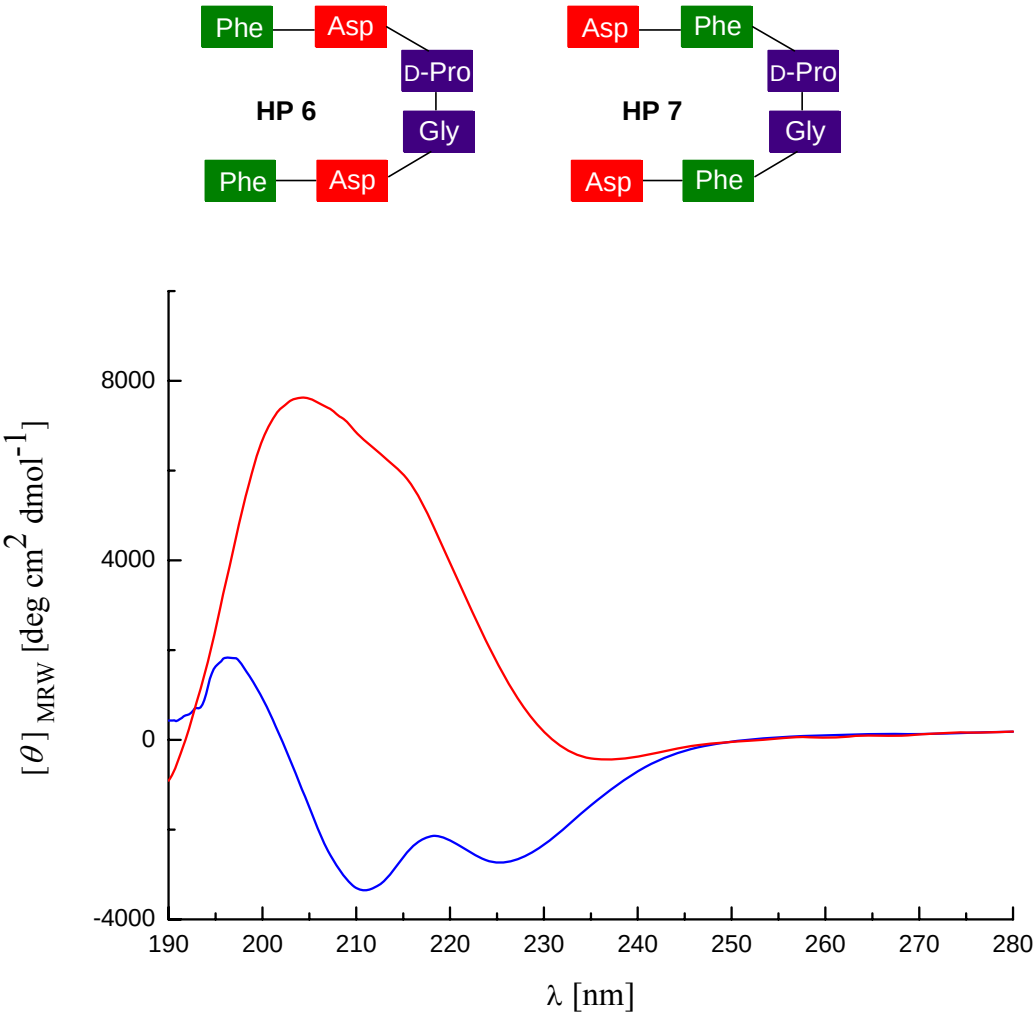


Figure 3.29 Comparison of the CD spectra oF **HP 6** and **HP 7**: 2 mg/ml **HP 6** in HFIP (blue); 2 mg/ml **HP 7** in HFIP (red)

The swap of positions of the "corner residues" with the neighbouring residues could exert a significant influence on the conformation and the stability of the conformation of the β-hairpin peptide. The difference between **HP 6** and **HP 7** lies in their "corner residues" and the neighbouring residues. The exchange of the residues at these positions might alter the pattern of the intra-chain hydrogen bond combination. As shown in *Figure 3.29*, it is obvious that the β-hairpin conformation of **HP 7** was more stable than that of **HP 6** in

HFIP. As phenylalanine residues occupy the "corner positions" of the β-hairpin peptide, its conformation was substantially stabilized by the strong aromatic interactions between the side chains of phenylalanine residues. Despite that repulsive electrostatic interactions between the aspartate side chains exist in **HP 7**, they were nevertheless excluded from the crucial "corner positions" and their unfavorable influences on the stability of the conformation of β-hairpin peptides was perhaps hence weakened. Thus **HP 7** adopted a more stable β-haipin conformation than its homologous derative **HP 6**, as the positive band appeared at ~210 nm for **HP 7** while only negative bands were observed in this region for **HP 6**. The alteration of the corner residues assumably led to a change of the pattern of hydrogen bond network in the β-hairpin peptide, which finally affected the β-hairpin conformational stability.

On the other hand, the introduction of phenylalanine residues in the sequence of **HP 6** produced the aromatic interactions which stabilized the β-hairpin conformation. As shown by the comparison of *Figure 3.29* and *Figure 3.23*, **HP 1** adopted an unordered structure in HFIP, while under the same conditions, **HP 6** adopted partial β-hairpin conformation. This comparison accounts for the positive contribution of aromatic interactions in stabilizing the β-hairpin conformation.

pH Dependence of the Conformation of H-Asp-Phe-D-Pro-Gly-Phe-Asp-OH (HP 7)

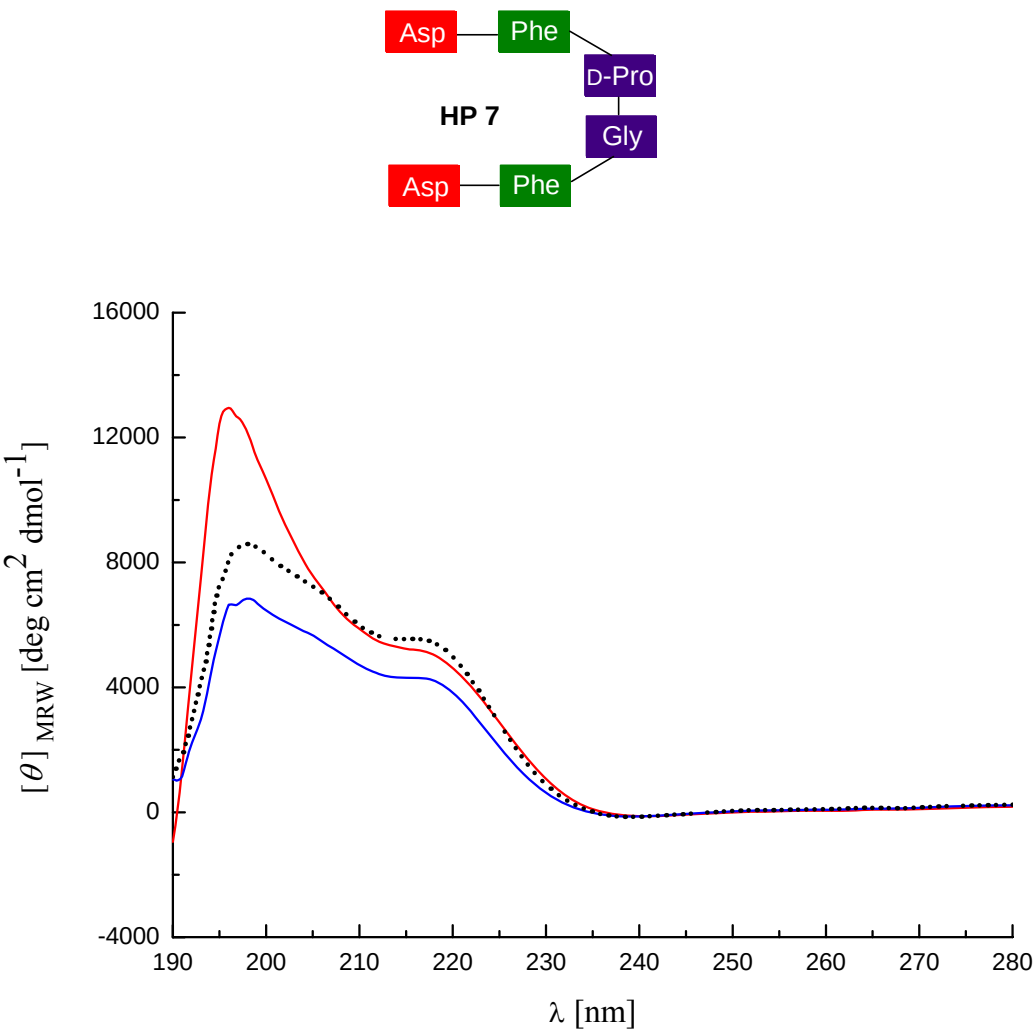


Figure 3.30 CD spectra of **HP 7** at different pH: 2 mg/ml **HP 7** in 0.02 M NaOH (red); 2 mg/ml **HP 7** in 0.2 M HCl (blue); 2 mg/ml **HP 7** in H_2O (dotted line, black)

Surprisingly, the pH dependence tendency of peptide **HP 7** was contrary to the case of **HP 6** (see *Figure 3.28*) and peptide **HP 19** (see *Figure 3.19*), as the addition of HCl decreased the β-hairpin content of **HP 7** while NaOH solution enhanced its conformational stability relative to its H_2O solution (see *Figure 3.30*). It could probably be explained from the point of hydrogen bond pattern in the β-hairpin peptide. As is already known, the pattern of hydrogen bonding between strand appears to be entirely determined by its own sequence (see also Chapter **3.2.3.2**).[75] The swap of residues on

the corner positions with those on their neighbouring positions could change the pattern of hydrogen bond networks. With this change, the protonation of the backbone carbonyl groups upon addition of HCl would drastically weaken the hydrogen bond interactions which might be more important than those in the other hydrogen bond pattern. This result showed the importance of corner residues from the viewpoint of the pattern of hydrogen bond networks.

Comparison of the Conformation of H-Phe-Glu-D-Pro-Gly-Glu-Phe-OH (HP 8) and H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in HFIP and H_2O

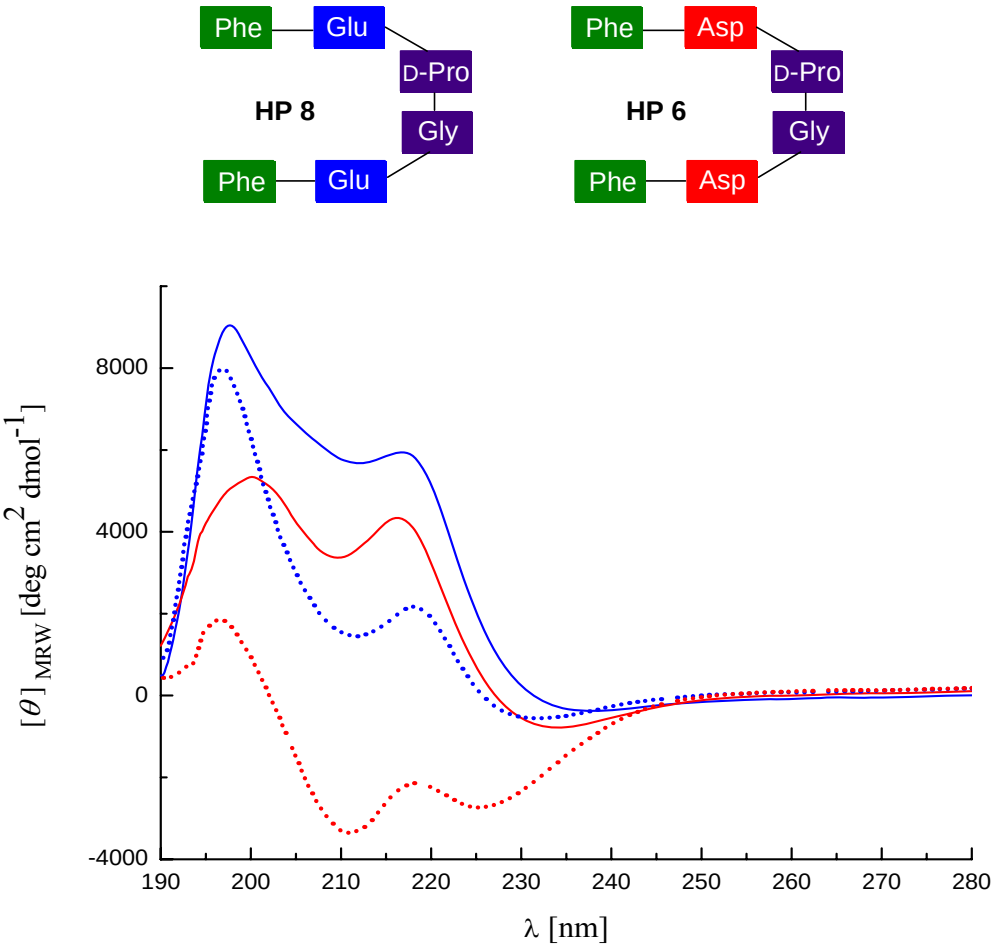


Figure 3.31 CD spectra comparison of **HP 8** and **HP 6** in HFIP and H_2O: **HP 8** in H_2O (2mg/ml, blue, solid line); **HP 8** in HFIP (2mg/ml, red, solid line); **HP 6** in H_2O (2mg/ml, blue, dotted line); **HP 6** in HFIP (2mg/ml, red, dotted line)

The substitution of aspartic acid residues in peptide **HP 6** by glutamic acids significantly stabilized the β-hairpin conformation of the concerned peptide, which was shown in *Figure 3.31*: the β-hairpin conformation of peptide **HP 8** was more stable than that of peptide **HP 6**, both in H_2O and HFIP. Furthermore, the conformation of peptide **HP 6** in HFIP was partially unordered, which was shown by the negative band near 210 nm, while peptide **HP 8** adopted β-hairpin conformation under the same conditions with two positive bands at ~200 nm and 216 nm (see *Figure 3.31*).

The conformation difference between **HP 6** and **HP 8** is similar to their parent peptide pair **HP 19** and **HP 18** (see *Figure 3.17*). However, **HP 6** does not differ from **HP 8** as much as **HP 18** from **HP 19,** with respect to the contribution of β-hairpin conformation (**HP 6** was partially unordered in HFIP with the preservation of the positive band at ~196 nm, while **HP 18** was nearly totally random coil under the condition in which **HP 19** adopted β-hairpin conformation). From this aspect one could deduce that the stabilization/destabilization effect on the peptide conformation is additive.

The differences between the peptide conformation (between **HP 18** and **HP 19**, as well as **HP 6** and **HP 8**) in terms of β hairpin stability could be ascribed to the distinctions between the aspartic acid and glutamic acid residue. As the pK_a of the carboxyl group in glutamate side-chain and aspartate side-chain are both approximately 4.5, [89] they would have the same degree of protonation under the same pH conditions, provided that the influence of the microenvironment was not taken into consideration. Since there is one more methylene group in the side chain of glutamic acid than aspartic acid, it will therefore possess more degrees of freedom that makes the side chain of glutamic acid more flexible compared to that of aspartic acid. Thus, the repulsive ionic intra-chain interactions between the glutamic acid side-chains resided on the opposite β-strand could be compensated to some extent, with the side chains of glutamic acids adopting favourable conformation, as is shown by the β-hairpin conformation of **HP 19** and **HP 8** (see *Figure 3.17* and *3.31*). While in **HP 18** and **HP 6**, the side chain of aspartic acid is relative rigid owing to the lack of one methylene group compared with glutamic acid, the side chains might thus possess relativly little room to preorganize and optimize their conformations in order to elude the repulsive ionic side-chain interactions which

destabilize their potential but not virtual β-hairpin conformations.

Another important advantage for glutamic acid-containing β-hairpin peptides is that the addition of one more methylene group compared to the aspartic acid introduces extra hydrophobic interactions with the corresponding residues on the opposite strand, the β-hairpin conformation is hence stabilized. **HP 8** also showed its property to adopt more stable β-hairpin conformation in H_2O than in HFIP. This is in agreement with the early discussed phenomenon, which the β-hairpin conformation of glutamic acid-containg peptide is more stable than that of aspartic acid-containing homologous peptide.

Comparison of the Conformation of H-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-OH (HP 9) and H-Phe-Glu-D-Pro-Gly-Glu-Phe-OH (HP 8) in HFIP and H_2O

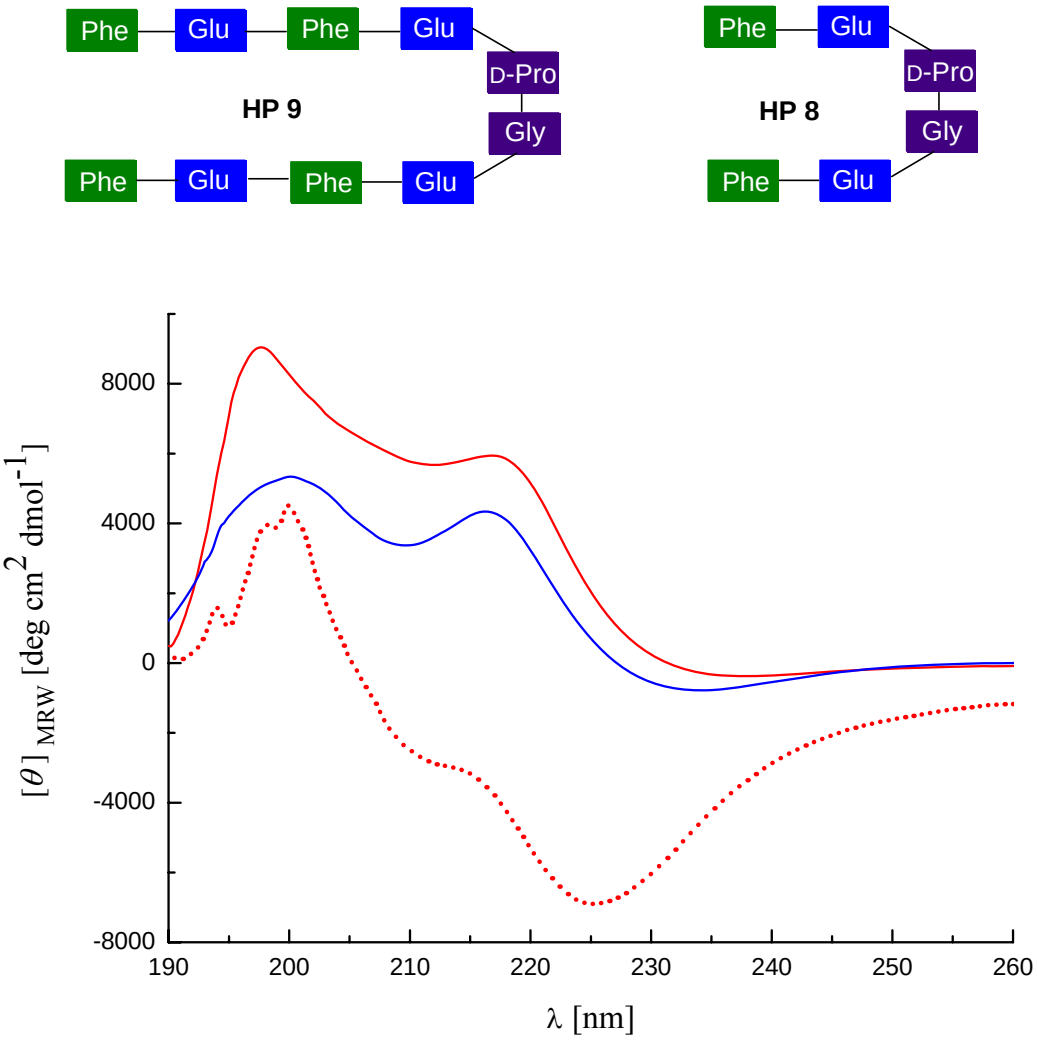


Figure 3.32 CD spectra of **HP 9** and **HP 8**: **HP 9** in HFIP (2mgl/ml, red, dotted line); **HP 8** in H_2O (2mg/ml, red, solid line); **HP 8** in HFIP (2mg/ml, blue, solid line)

HP 9 adopted partially unordered structure in HFIP (this peptide is not soluble in H_2O), as shown in *Figure 3.32*. Despite that two more phenylalanine residues and glutamic acid residues were introduced compared to the **HP 8**, the conformation of **HP 9** was not consequently stabilized by the extra aromatic interactions and hydrophobic interactions brought forward by the phenylalanine and glutamic acid. It appeared that these

advantageous factors could not match the repulsive ionic interactions between the side chains of glutamic acids. No reasonable explanation was found so far to clarify the fact that **HP 19** adopted more stable β-hairpin conformation than **HP 9** with the regular chain elongation by alternate phenylalanine and glutamic acid.

Comparison of the Conformation of H-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-OH (HP 9) and H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-OH (HP 10) in HFIP

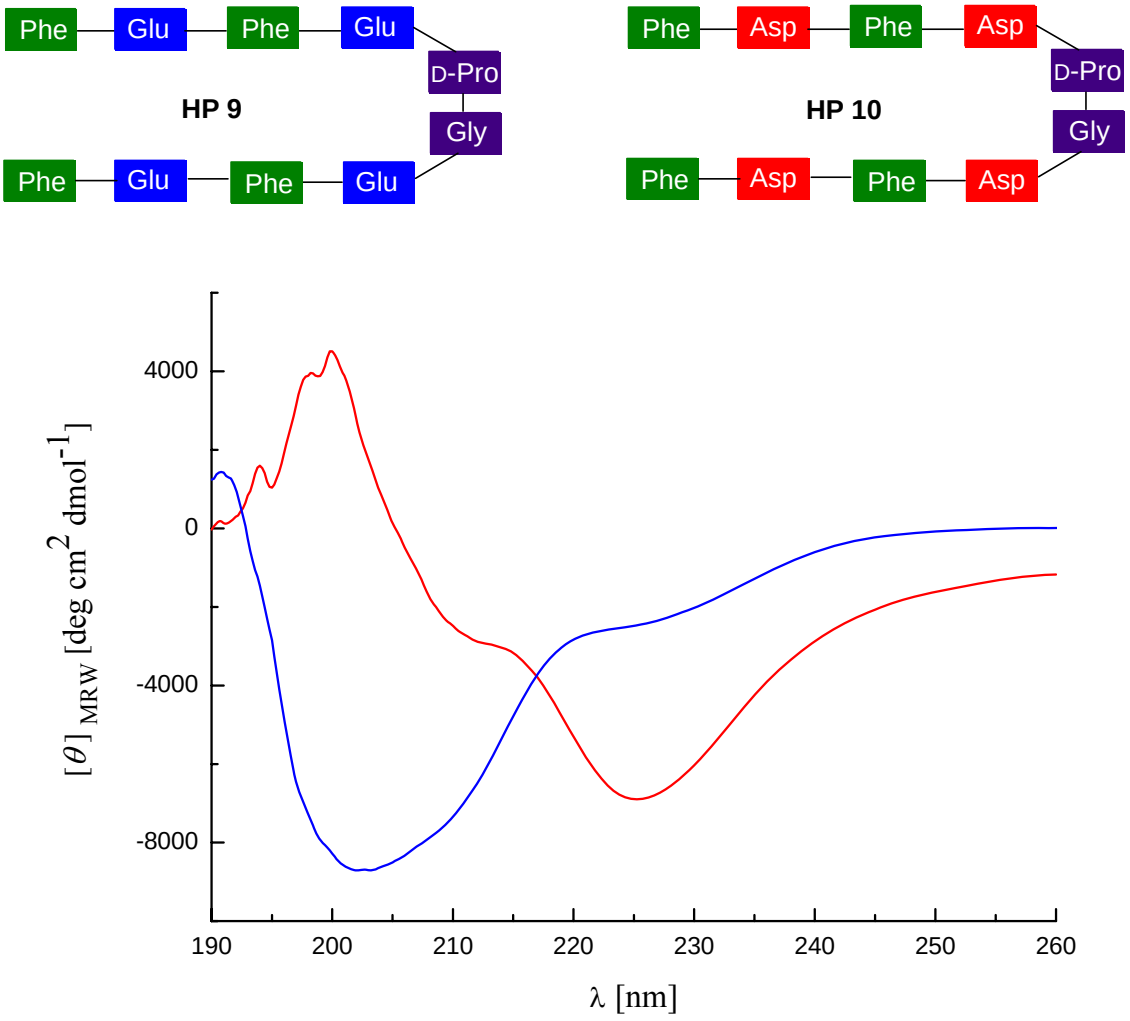


Figure 3.33 CD spectra of **HP 9** and **HP 10**: **HP 9** in HFIP (2mg/ml, red); **HP 10** in HFIP (2mg/ml, blue)

HP 9 adopted a markedly different conformation from that of **HP 10** in HFIP, as shown in *Figure 3.33*. **HP 9** was partially random coil accompanied with a β-hairpin component exhibited by the positive band at ~200 nm; while **HP 10** revealed no signs of β hairpin existence. This phenomenon could also be clarified by the same explanation for the conformational difference between **HP 6** and **HP 8** (see *Figure 3.31*) as well as **HP 18** and **HP 19** (see *Figure 3.17*). The differences between the properties of aspartic acid and glutamic acid accounted for the differences between the conformational stabilities of **HP 9** and **HP 10**.

Comparison of the Conformation of H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-OH (HP 10) and H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6) in HFIP

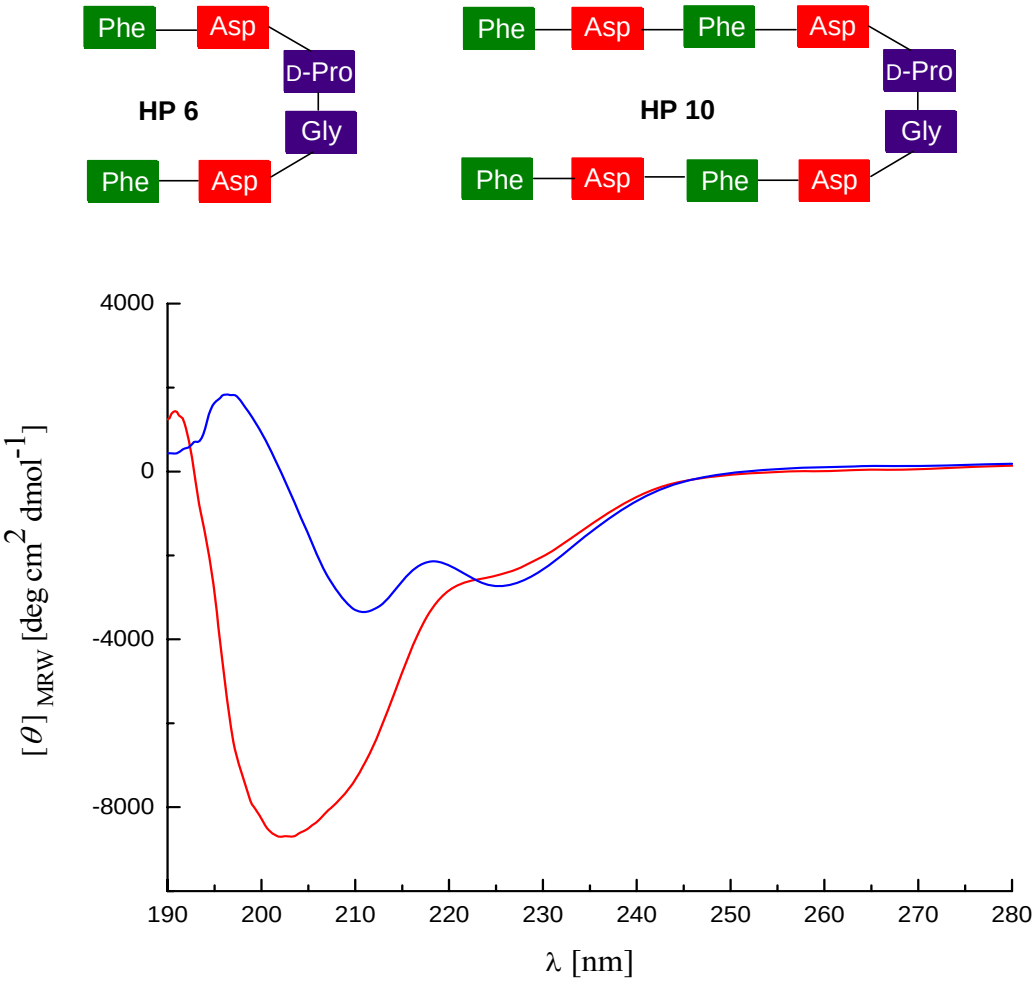
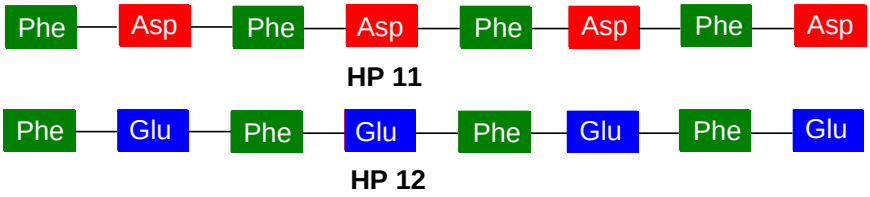


Figure 3.34 CD spectra of **HP 10** and **HP 6**: **HP 10** in HFIP (2mg/ml, red); **HP 6** in HFIP (2mg/ml, blue)

HP 6 adopted partial β-hairpin and random coil conformation as shown in *Figure 3.25* and *3.26*. With the regular chain elongation by adding alternate phenylalanine and aspartic acid, its conformational stability was weakened as was shown by the random coil features in **HP 10** (see *Figure 3.34*). The extra aromatic interactions in **HP 10** generated by the added phenylalanine residues were not sufficient to balance the repulsive electrostatic interactions between the newly introduced aspartic acid. The β-hairpin conformational stabilities of these peptides are therefore chain-length dependent.

Comparison of the Conformation of H-Phe-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 11) and H-Phe-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 12) in HFIP



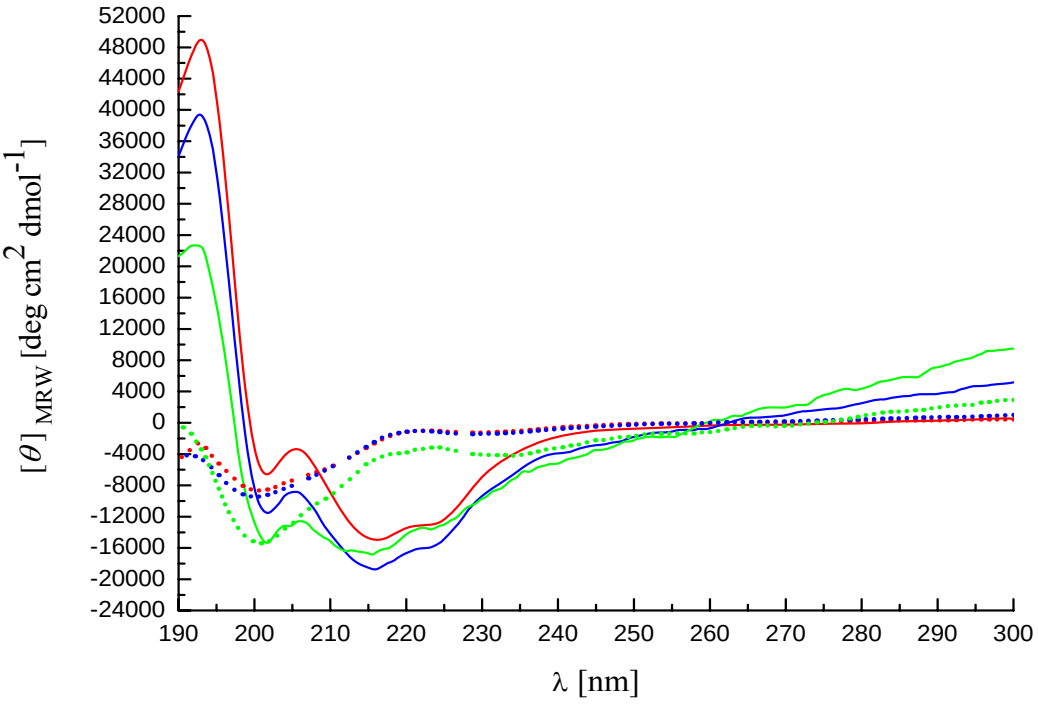


Figure 3.35 CD spectra of **HP 11** and **HP 12**: **HP 12** in HFIP (0.5mg/ml, red, solid line); **HP 12** in HFIP (0.1mg/ml, blue, solid line); **HP 12** in HFIP (0.05mg/ml, green, solid line); **HP 11** in HFIP (1mg/ml, red, dash line); **HP 11** in HFIP (0.5mg/ml, blue, dash line); **HP 11** in HFIP (0.1mg/ml, green, dash line)

Polypeptides $(FD)_n$ and $(FE)_n$ were merged in **HP 18** and **HP 19**, respectively, as imaginary β-strand part. However, **HP 18** and **HP 19** adopted very different conformations in HFIP, as was shown by *Figure 3.17*. A series of aforementioned peptides had revealed the conformational discrepancies between the Asp- and Glu-containing peptides. All the present experimental results exclusively favored the tendency that aspartic acid was not as prone as glutamic acid in terms of propelling the formation of β-strand conformation. **HP 12** and **HP 11** serve as the proper models to compare the differences between glutamic acid and aspartic acid as their potencies as individual amino acids to force the formations of β-strand, since the absence of -D-Pro-Gly- moiety makes the comparison more reasonable, reliable and straightforward.

The CD spectra of **HP 11** and **HP 12** (see *Figure 3.35*) revealed a significant difference between their conformations adopted in HFIP. **HP 12** adoped mainly β-strand conformation, with a negative band at ~216 nm and a positive band at ~193 nm, whereas **HP 11** was present the unordered structure. Furthermore, **HP 12** manifested a obvious tendency to undergo self aggregation as their CD spectra of a typical β-strand at different concentrations did not superimpose perfectly. This experiment revealed the inherent differences between the aspartic acid and glutamic acid in their tendencies to reinforce the β-strand conformation, serving as one of the most convinced proofs to explain the huge conformational discrepancies between their parent peptide **HP 18** and **HP 19**.

Comparison of the Conformation of H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19) and H-Glu-Phe-Glu-Phe-Glu-Phe-Asp-D-Pro-Gly-Asp-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 16) in HFIP

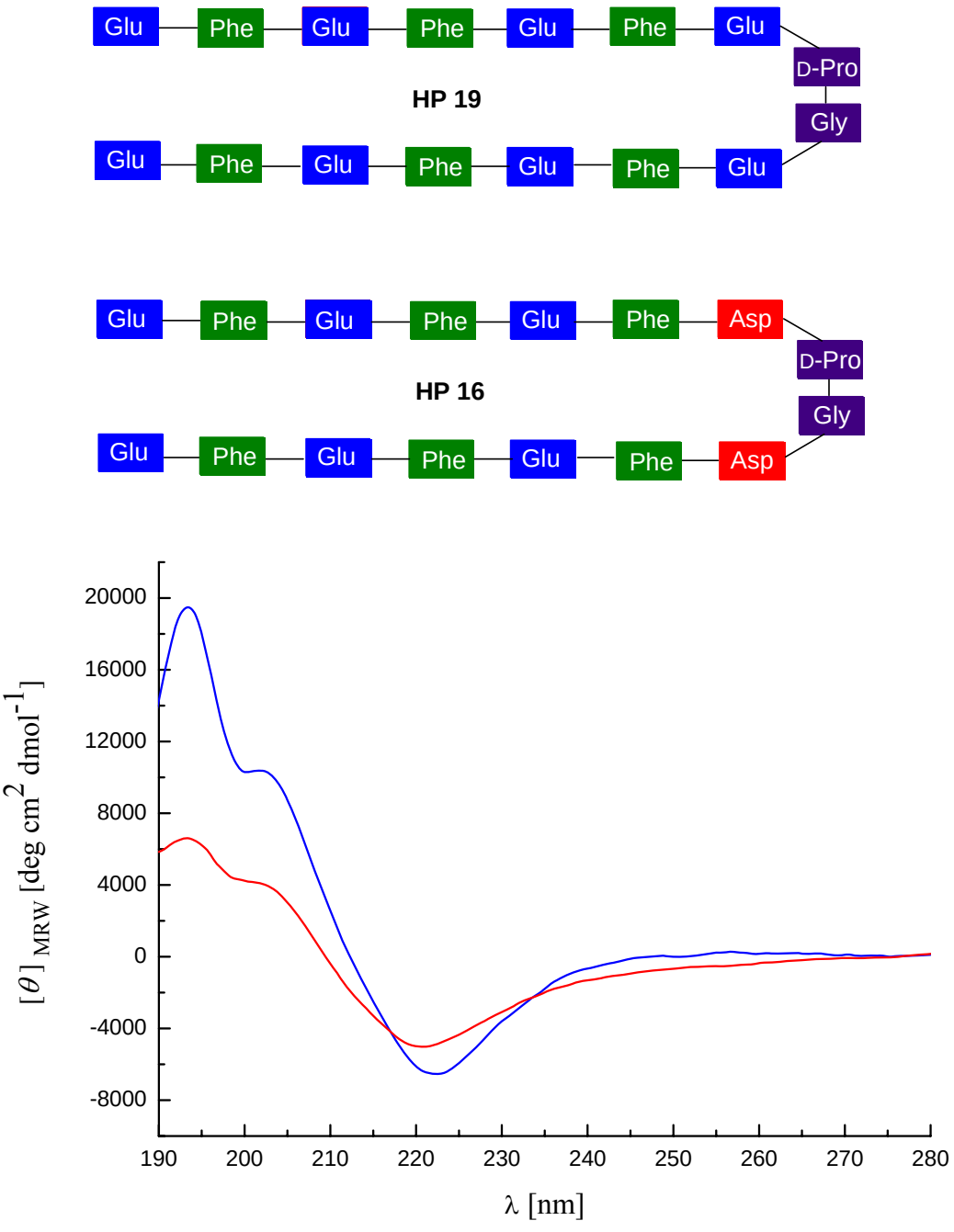


Figure 3.36 CD spectra of **HP 19** and **HP 16**: **HP 19** in HFIP (0.3mg/ml, blue); **HP 16** in HFIP (0.5mg/ml, red)

The unique difference between **HP 19** and **HP 16** lies solely in the residues resided at the "corner positions". These positions are occupied by glutamic acids in **HP 19**, while by aspartic acids in **HP 16**. The distinctions between the turn moieties have been displayed by the conformational discrepancies between **HP 2** and **HP 1** (see *Figure 3.21*). The influences of these short peptides were "transplanted" into **HP 19** and **HP 16** and therefore exerted on their conformations. As shown in *Figure 3.36*, the content of β-hairpin conformation from **HP 19** was higher than that of **HP 16**. The unique difference between these two peptides at the corner positions engendered the huge distinctions as the inclination of the conformation was concerned. This experiment revealed again the significance of the corner residues in stabilizing/destabilizing the β-hairpin conformations.

Comparison of the Conformation of H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 18) and H-Asp-Phe-Asp-Phe-Asp-Phe-Glu-D-Pro-Gly-Glu-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 15) in HFIP

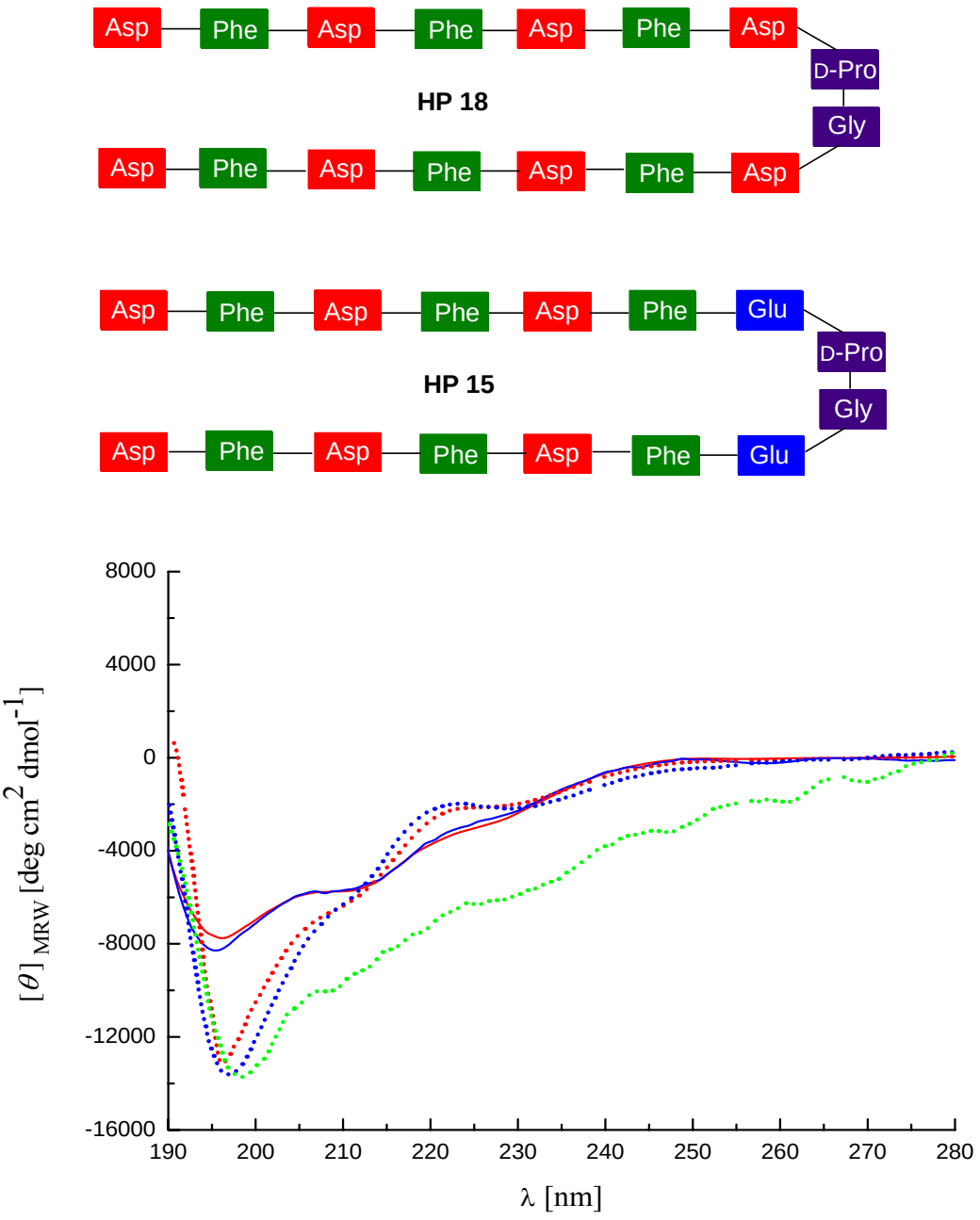


Figure 3.37 CD spectra of **HP 18** and **HP 15**: **HP 18** in HFIP (1mg/ml, red, solid line); **HP 18** in HFIP (0.3mg/ml, blue, solid line); **HP 15** in HFIP (2mg/ml, red, dotted line); **HP 15** in HFIP (0.5mg/ml, blue, dotted line); **HP 15** in HFIP (0.1mg/ml, green, dotted line)

The substitution of the "corner residues" aspartic acids by glutamic acids did not transform the conformation of peptide HP 18 from random coil to ordered structure. The total repulsive ionic interactions between the aspartate side chains could not be compensated by the tiny positive contributions ascribed to the introduction of the methylene groups from glutamic acids at the "corner positions" in terms of structure stabilization, as shown in CD spectra of **HP 15** which revealed the conformation of random coil (see *Figure 3.37*).

However, the substitution of the corner resides did generate the conformational changes, which could be deduced from the “shoulder” at the wavelength around 220 nm. It’s regarded as the appearance of β turn motif in the concerned peptide. The substitute of the corner residues from aspartic acid to glutamic acid increased the proportions of the ordered structures in the inhomogeneous peptide solution, despite not being able to fold the peptide into a fully ordered structure. It could be imagined, that if the content of orderd structure from a predominant random coil peptide solution enhanced gradually, the ellipiticities at the wavelength around 200 nm would also be increased, from negative to positive. Form the comparison of the structures of peptide **HP 15** and **HP 18**, it could be deduced that content of their ordered β-hairpin conformation is different. Furthermore, it can be assumend that β-hairpin conformation is structurally composed of β-turn and β-strand which are supposed to insert their influences separately on the stability of β-hairpin peptide.

Comparison of the Conformation of H-Glu-Phe-Ser-Phe-Glu-Phe-Ser-D-Pro-Gly-Glu-Phe-Ser-Phe-Glu-Phe-Ser-OH (HP 20) and H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19) in HFIP

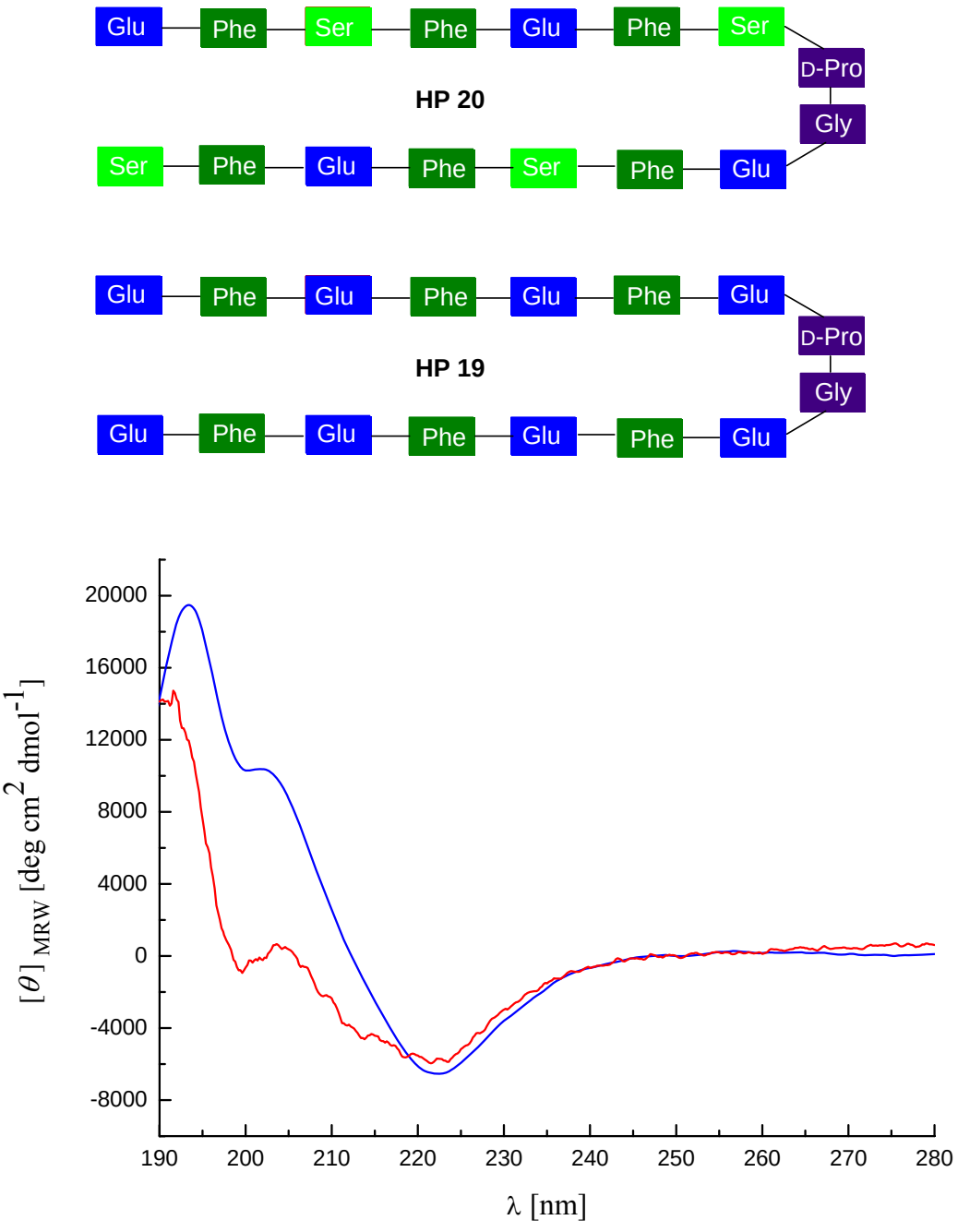


Figure 3.38 CD spectra of **HP 19** and **HP 20**: 0.3 mg/ml **HP 19** in HFIP (blue); 0.5 mg/ml **HP 20** in HFIP (red)

Glutamic acid residues were substituted in some certain positions (see also sequential difference between **HP 19** and **HP 20**) in **HP 19** by serine residues in order to reduce the interactions between the side-chains of glutamic acid residues resided in the opposite positions on the β-strands of β-hairpin peptide **HP 19**. This Glu-Glu interactions in **HP 19** were replaced by Glu-Ser interactions in **HP 20** provided that **HP 20** would also adopt β-hairpin conformation under the same condition.

As shown in *Figure 3.38*, **HP 20** still adopted mainly β-hairpin conformation in HFIP, as its counterpart peptide **HP 19**, but to a less degree, revealed by the weakening of the intensities of the positive bands at ~193 nm and 201 nm. The repulsive electrostatic interactions between the side chains of glutamic acids in **HP 19** was basically eliminated in **HP 20**, since the hydroxymethyl groups of the serine side-chains remained under most cases unionized. On the other hand, however, the hydrophobic interactions between the side chains of glutamic acids were weakened alonog with the substitution of glutamic acid by serine. This change partially accounted for the destabilization of the β hairpin conformation of **HP 20** relative to **HP 19**. This experiment revealed the importance of the hydrophobic interactions between the methylene groups in the side-chains of glutamic acids in stabilizing the β-hairpin conformation. One can therefore hypothesize that the β-hairpin conformation of peptide **HP 19** would possibly be further stabilized if the side-chains of glutamic acids elongated by one more methylene group.

Comparison of the conformation of H-Asp-Phe-Ser-Phe-Asp-Phe-Ser-D-Pro-Gly-Asp-Phe-Ser-Phe-Asp-Phe-Ser-OH (HP 17) and H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 18) in HFIP

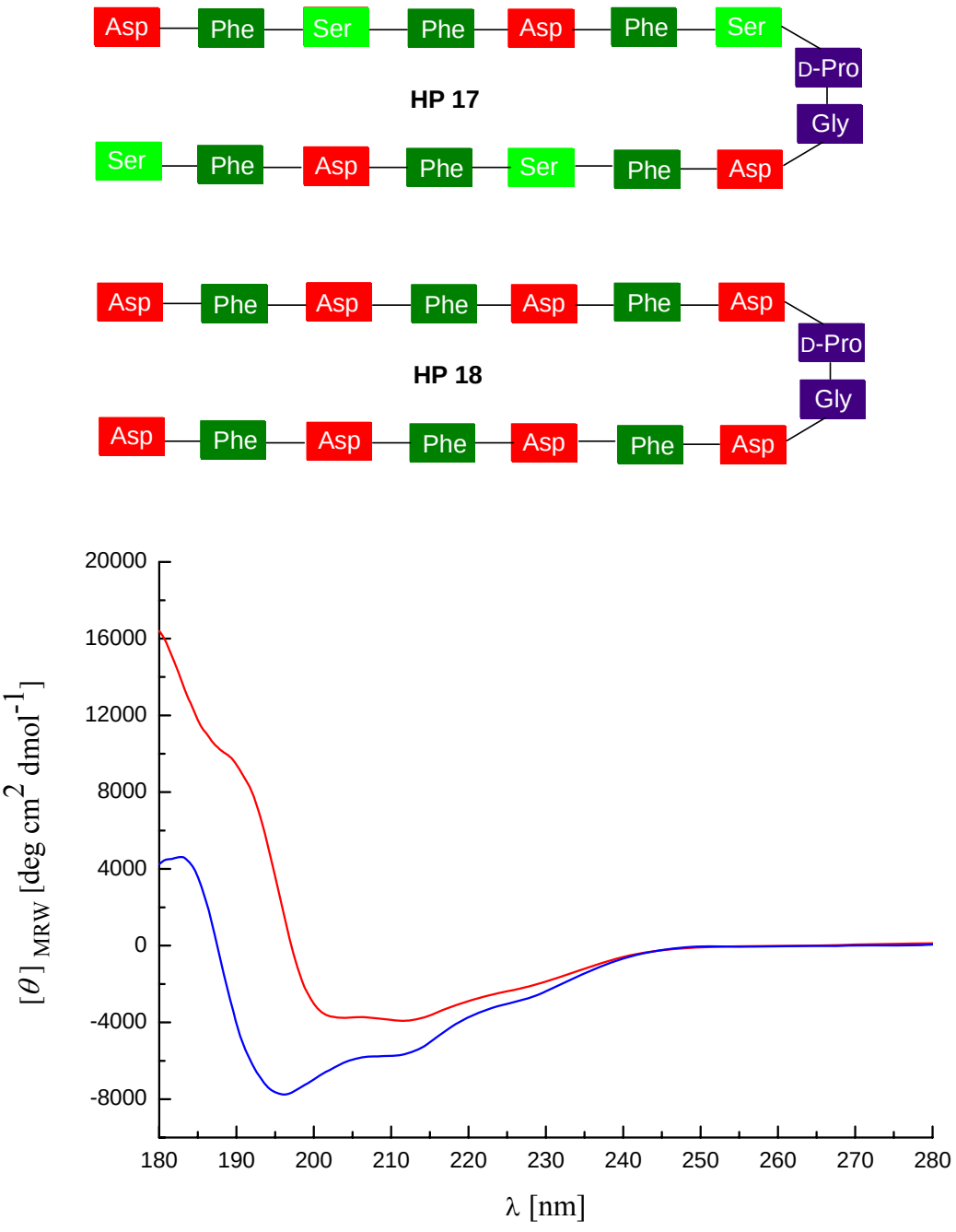
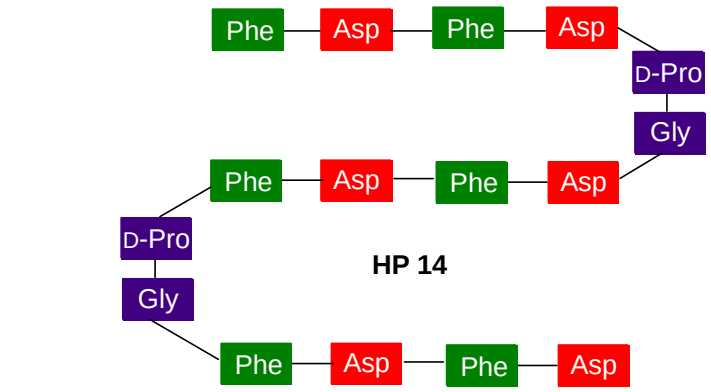


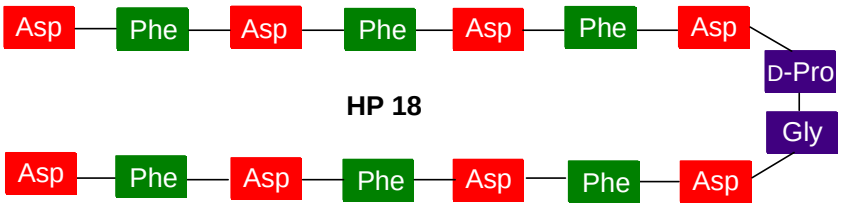
Figure 3.39 CD spectra of **HP 18** and **HP 17**: 2mg/ml **HP 17** in HFIP (red); 1 mg/ml **HP 18** in HFIP (blue)

Aspartic acid residues were substituted in some certain positions in **HP 18** by serine residues (see alsio sequential difference between **HP 18** and **HP 17**) in order to reduce the interactions between the side chains of aspartic acid residues resided in the opposite positions on the β-strands of β-hairpin peptide **HP 18**. In spite of the elimination of such a negative factor that destabilizes the β-hairpin conformation of the peptide, **HP 17** did still adopt the partial random coil conformation in HFIP. However, comparison of the CD spectra of peptide **HP 18** and **HP 17** revealed that, upon substitution of aspartic acid by serine residues, the proportion of ordered fractions in inhomogeneous peptide solution was increased, as was revealed by the weakening of the negative band and the appearance of the positive shoulder at the wavelength above 190 nm (see *Figure 3.39*). This experiment indicated again that the ionic interactions between the side chains of aspartic acid residues are one of the major factors that destabilize the potential β-hairpin conformations of peptides, and that the hydrophobic interactions brought forward by the aggregations of methylene groups from the side chains of Asp/Glu play a role in stabilizing the potential β-hairpin conformation.

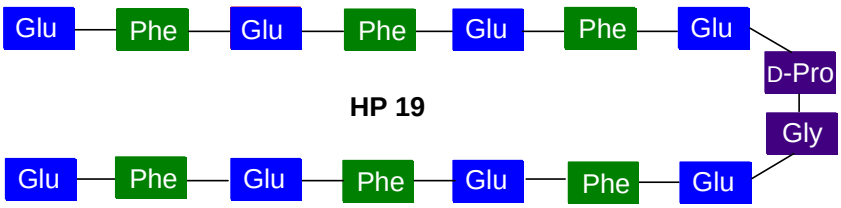
Comparison of the Conformation of H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-D-Pro-Gly-Phe-Asp-Phe-Asp-OH (HP 14) H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 18) and H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19) in HFIP



HP 14



HP 18



HP 19

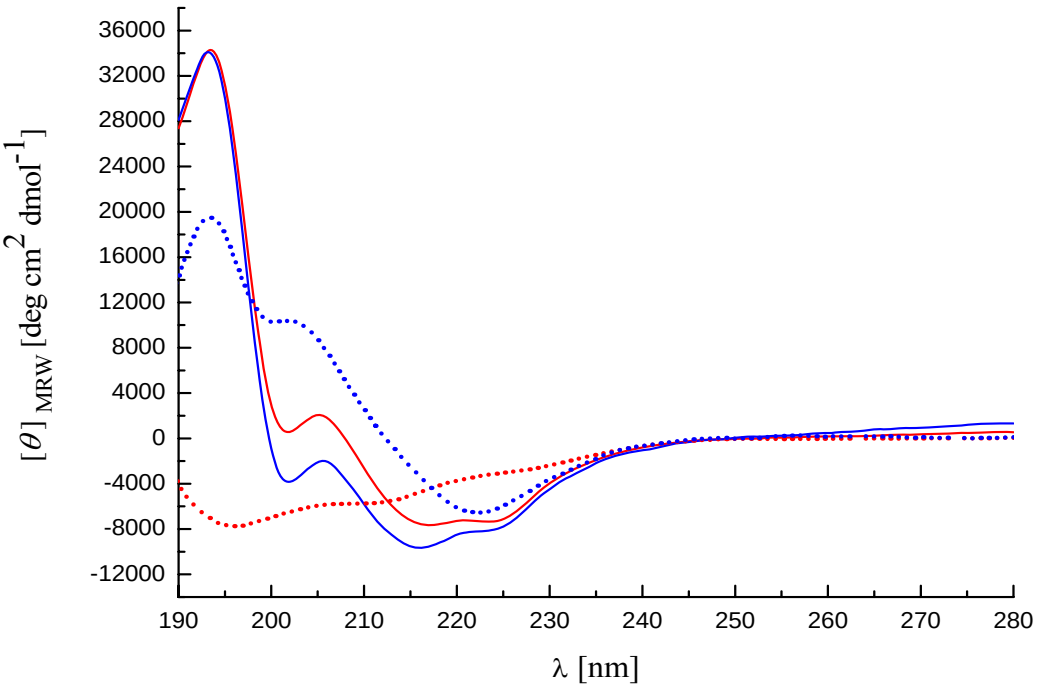


Figure 3.40 CD spectra of **HP 18**, **HP 19** and **HP 14**: **HP 14** in HFIP (0.5mg/ml, red, solid line); **HP 14** in HFIP (0.25mg/ml, blue, solid line); **HP 18** in HFIP (1mg/ml, red, dotted line); **HP 19** in HFIP (0.3mg/ml, blue, dotted line)

HP 14 contains two -D-Pro-Gly- moieties which are supposed to change the directions of the peptide chain, the hereby-designed peptide will thus theoretically be composed of two turn moieties and three antiparallel β-strands which are brought together by intra-chain hydrogen bonds and aromatic interactions as well as hydrophobic interactions between the side-chains of the residues positioned on the strands. These moieties are connected in tandem in the primary sequence to adopt the shape of "S" in terms of secondary structure. The cluster of converged aromatic groups and other hydrophobic groups are supposed to generate the additive aromatic interactions as well as hydrophobic interactions that will understandably reinforce the theoretical "double β hairpin" conformation of **HP 14**. This hypothesis was validated by the CD spectra of **HP 14** in HFIP, as shown in *Figure3.40*. Compared with the relevant "single β-hairpin" **HP 19**, which contains "advantageous" glutamic acid residues as the propensity to induce β-hairpin conformation is concerned, **HP 14** exhibited a more abundant β-hairpin content, proven by a stronger ellipticity at the

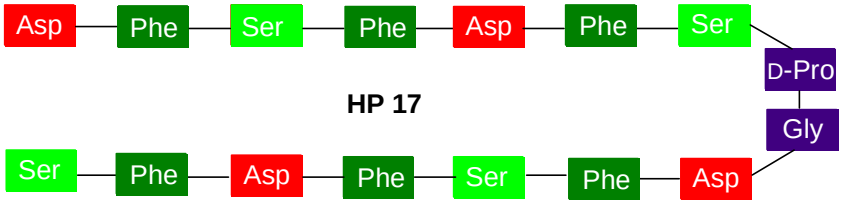
wavelength ~193 nm, despite an increase in repulsive interactions between the side chains of the aspartic acid residues in **HP 14**. The additive aromatic interactions and hydrophobic interactions "spanning over strands" aggregated by the change of the chain directions offset the negative factors which destabilize the hairpin conformation and reinforce the conformation of the potential β-hairpin conformation.

On the other hand, as far as the stability of β-hairpin conformation is concerned, **HP 14** possesses obvious advantages over **HP 18**, which is also composed of aspartic acid and phenylalanines, while having a single -D-Pro-Gly- moiety. The conformation of peptide **HP 18** was mainly unordered even at the relatively high concentration of 2 mg/ml, while **HP 14** adopted β-hairpin conformation even at the dilute concentration of 0.25 mg/ml (see *Figure 3.40*).

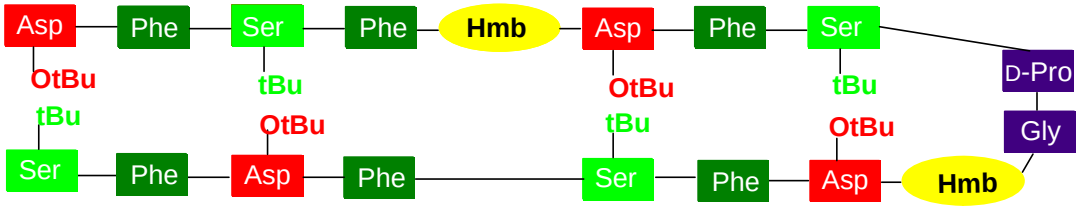
This experiment showed clearly that the aggregation of aromatic interactions and hydrophobic interactions between the strands of peptides play crucially important role in stabilizing the β-sheet (which is well known by the fact of peptide aggregations) or β-hairpin conformation. Furthermore, these forces seem to be additive. As the hydrogen bonds between the strands can not be viewed as additive **HP 14** since they could not be formed spanned over a strand, they did not offer extra contributions for the stabilization of the β-hairpin conformation.

Comparison of the Conformation of H-Asp(OtBu)-Phe-Ser(tBu)-Phe-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-*N*-Hmb-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH (Prot. HP17-Hmb) and H-Asp-Phe-Ser-Phe-Asp-Phe-Ser-D-Pro-Gly-Asp-Phe-Ser-Phe-Asp-Phe-Ser-OH (HP 17) in HFIP and ACN/H_2O (1:1)

.



HP 17



Prot. HP 17-Hmb

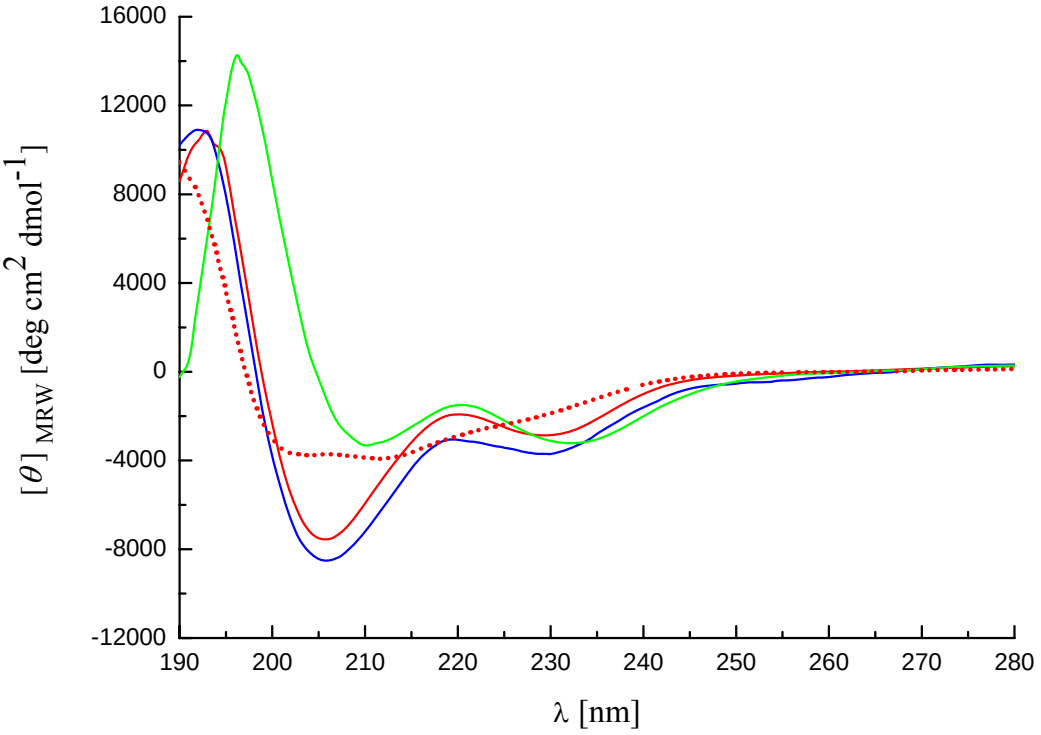


Figure 3.41 CD spectra of **Prot. HP 17-Hmb** and **HP 17**: **Prot. HP 17-Hmb** in HFIP (2mg/ml, red, solid line); 0.5 mg/ml **Prot. HP 17-Hmb** in HFIP (0.5mg/ml, blue, solid line); **Prot. HP 17-Hmb** in ACN/H_2O (1:1) (2mg/ml, green, solid line); **HP 17** in HFIP (2mg/ml, red, dotted line)

The differences of the CD spectra between **HP 17** and its side chain fully protected, backbone Hmb protected counterparts **Prot. HP 17-Hmb** were obvious as shown in *Figure 3.41*. Peptide **Prot. HP 17-Hmb** adopted partially ordered conformation containing elements of β-strand in HFIP, as was revealed by the positive band at ~193 nm. On the other hand, some fractions of this peptide adopted unordered conformation in HFIP, validated by the negative band at ~ 205 nm, while peptide **HP 17** did not reveal abundance of ordered structure. It exhibited no positive band above 190 nm, as shown in *Figure 3.41*.

The *tert*-butyl protecting groups on the side-chains of aspartic acid residues prevent the repulsive ionic interactions between the neighboring strands of **Prot. HP 17-Hmb**; meanwhile the hydrophobic interactions between the *tert*-butyl groups stabilize the

β-hairpin conformation. The cooperative influences of the elimination of ionic interactions and introduction of hydrophobic interactions significantly increase the β-hairpin content in **Prot. HP 17-Hmb**.

The backbone Hmb exerted obvious influences on the conformation of peptide **Prot. HP 17-Hmb**. The backbone protecting groups such as Hmb is a well-known functionality that prevents the formation of the secondary structures such as β-sheet through the block of the hydrogen bonds between the backbone NH and CO groups, its introduction is thus understandably able to jeopardize the β-hairpin conformation of **Prot. HP 17-Hmb**. While the side chain protecting groups in this peptide eliminated the adverse electrostatic interactions and introduced the extra hydrophobic intra-chain interactions as well. With the heterogeneous influences of *tert*-butyl side-chain protecting groups and Hmb of back-bone protecting groups on the formation of peptide **Prot. HP 17-Hmb**, and the original properties of peptide **HP 17** toward the β-hairpin conformation, the actual structure of peptide **Prot. HP 17-Hmb** decided by these factors was consequently a balanced structure between β-hairpin and random coil, exhibited by its CD spectrograms (see *Figure 3.41*).

3.6 Synthetic Experimental Section

3.6.1 General Methods (see also Chapter 2.6.1)

Solvent

Flash Chromatography

Thin Layer Chromatography

NMR Spectroscopy

MALDI-ToF Mass Spectrometry

Digital Polarimeter

Microwave Peptide Synthesizer

CD Spectroscopy

Analytical RP-HPLC

Preparative RP-HPLC

3.6.2 Peptide Synthesis (see also Chapter 2.6.2)

Loading of *o*-chlorotritylchloride Resin and Quantification

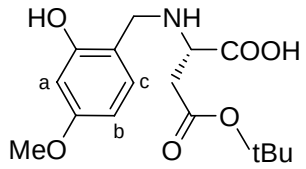
On-resin Peptide Chain Elongation by Fmoc/tBu Chemistry

Cleavage of the Fully Protected Peptide from the Resin

Side Chain Deprotection

Microwave Synthesis of Peptide with SPPS Manner

_N_-Hmb-Asp(OtBu)-OH (29)



H-Asp(OtBu)-OH · H_2O (2.36 g, 11.4 mmol) was dissolved in 20 ml H_2O and the pH was adjusted to pH 9.0 with 2 M NaOH. Finely powered 2-hydroxy-4-methoxy-benzaldehyde **27** (1.7 g, 11.4 mmlol) was added and the resulting suspension was left stirring for 2 hours at room temperature, while the pH was kept at 9-9.5 with 2 M NaOH. Then the reaction mixture was cooled to 5 ºC, and 15 ml 0.76 M aq. $NaBH_4$ (11.4 mmol) and 2 M NaOH (3 drops) were added dropwise over 20 minutes. After 1 hour at 5-10 ºC, the mixture was neutralized to pH 6-7 with 2 M HCl under vigorous stirring while the temperature was kept below 20 ºC. The product was precipitated in the cold and was collected by filtration.

Yield: 2.38 g (70%)

Formula: $C_{16}H_{23}NO_6$

Molecular Mass: 325.36 g/mol

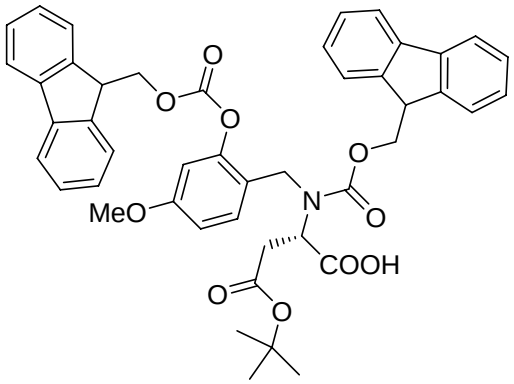
$[\alpha]_D^{21}$ = -5.8 (c = 1.0, H_2O)

MS (MALDI-ToF): $m \cdot z^{-1}$ = 326.47 $[M+H]^+$, 348.46 $[M+Na]^+$, 364.42 $[M+K]^+$

Analytical HPLC: Method I, t_R = 17.88 min, 100%

^{1}H NMR (500.15 MHz, D_6-DMSO): δ [ppm] = 1.39 (s, 9H, C(C**H**$_3$)$_3$), 2.49 (s, 1H, N**H**), 2.59 (dd, ^2J = 16.5 Hz, ^3J = 6.9 Hz, 1H, C**H**$_2$CO$_2$tBu), 2.68 (dd, ^2J = 16.5 Hz, ^3J = 5.5 Hz, 1H, C**H**$_2$CO$_2$tBu), 3.44 (t, ^3J = 6.9 Hz, 1H, NHC**H**(CH$_2$CO$_2$tBu)COO), 3.68 (s, 3H, OC**H**$_3$), 3.91, 3.92 (d, 2H, C**H**$_2$NH), 6.35-6.41 (m, 2H, C^a**H**, C^b**H**), 7.10 (d, ^3J = 8.0 Hz, 1H, C^c**H**)

_N,O_-bis-Fmoc-_N_-Hmb-Asp(OtBu)-OH (34)



2.1 equiv Fmoc-Cl **37** (3.260 g, 12.6 mmol) in 1,4-Dioxane (14 ml) were added to an ice-cold solution of *N*-Hmb-Asp(OtBu)-OH **29** (1.952g, 6 mmol) in H_2O (60 ml)/1,4-Dioxane(30 ml) containing 3.2 equiv Na_2CO_3 (2.04 g, 19.2 mmol) during 1h. The solution was stirred for 1-2 hours and a precipitate formed. The mixture was extracted with diethyl ether (3×100 ml) and the combined ethereal solutions were washed with 0.5 M HCl (3×100 ml) and dried (Na_2SO_4). After removal of the solvent, the oily residue was chromatographed (0.04-.063 mm silica gel, packed with CH_2Cl_2) with Hexane/EtOAc/CH_2Cl_2 (6:3:1) as eluent. The content of eluent was shifted to MeOH/EtOAc (2:8) as the first side-product (mono-Fmoc derivative) had been eluted. Alternatively, the target product can be purified by the RP-HPLC.

Yield: 1.90 g (41%)

Formula: $C_{46}H_{43}NO_{10}$

Molecular Mass: 769.83 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 792.14 $[M+Na]^+$, 808.05 $[M+K]^+$, 814.13 $[M+2Na-H]^+$

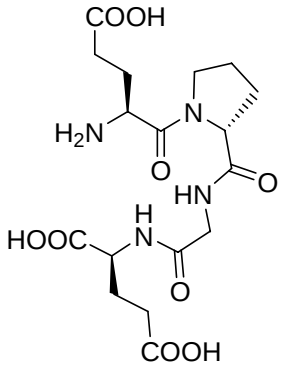
TLC: R_f = 0.21 (CH_2Cl_2/MeOH 96:4)

Preparative RP-HPLC: Apparatus B, Method 10, t_R = 20.1 min

Analytical RP-HPLC: Method II, t_R = 40.78 min, 100% (220 nm)

^{1}H NMR (500.15 MHz, $CDCl_3$): δ [ppm] = 1.36 (s, 9H, C(C**H**$_3$)$_3$), 2.44 (dd, ^{2}J = 16.5 Hz, ^{3}J = 6.0 Hz, 1H, C**H**$_2$CO$_2$tBu), 2.57 (dd, ^{2}J = 16.5 Hz, ^{3}J = 6.5 Hz, 1H, C**H**$_2$CO$_2$tBu), 3.76 (s, 3H, C**H**$_3$O), 4.17 (t, ^{3}J = 6.0 Hz, 1H, CO$_2$CH$_2$C**H**), 4.23 (t, ^{3}J = 7.0 Hz, 1H, OCOCH$_2$C**H**), 4.31,4.34 (s,s 2H, C**HH**N), 4.42-4.52 (m, 4H, OCOC**H**$_2$, NCO$_2$C**H**$_2$), 4.50 (dd, ^{3}J = 6.0 Hz, ^{3}J = 6.5 Hz, 1H, NHC**H**(CH$_2$CO$_2$tBu)COO), 6.45-7.84 (m, 19H, **H**^{Aryl})

H-Glu-D-Pro-Gly-Glu-OH (HP 1)



HP 1

The synthesis of the fully protected peptide **Prot. HP 1** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Glu(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.2 mmol), and TBTU as coupling reagent. Deprotection of peptide **Prot. HP 1** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

(Except otherwise mentioned, all the yields listed in this project represent those after RP-HPLC. *vide infra*)

Yield: H-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)-OH (**Prot. HP 1**) 0.0922 g/0.17 mmol
H-Glu-D-Pro-Gly-Glu-OH (**HP 1**) 0.0710 g/0.16 mmol (80.0%)

Formula: $C_{17}H_{26}N_4O_9$

Molecular Mass: 430.41 g/mol

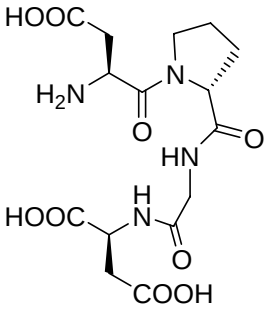
MS (MALDI-ToF): $m \cdot z^{-1}$ = 431.49 $[M+H]^+$, 453.48 $[M+Na]^+$, 469.48 $[M+K]^+$, 475.48 $[M+2Na-H]^+$

Calculated Mass: $[C_{17}H_{27}N_4O_9]^+$ = 431.43, $[C_{17}H_{26}N_4O_9Na]^+$ = 453.41, $[C_{17}H_{26}N_4O_9K]^+$ = 469.52, $[C_{17}H_{25}N_4O_9Na_2]^+$ = 475.39

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 9.67 min

Analytical RP-HPLC: Method I, t_R = 6.15 min, 100% (220 nm)

H-Asp-D-Pro-Gly-Asp-OH (HP 2)



HP 2

The synthesis of the fully protected peptide **Prot. HP 2** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Asp(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.2 mmol), and TBTU as coupling reagent. Deprotection of peptide **Prot. HP 2** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Asp(OtBu)-D-Pro-Gly-Asp(OtBu)-OH (**Prot. HP 2**) 0.0771 g/0.15 mmol
H-Asp-D-Pro-Gly-Asp-OH (**HP 2**) 0.0420 g/0.10 mmol (50.0%)

Formula: $C_{15}H_{22}N_4O_9$

Molecular Weight: 402.37

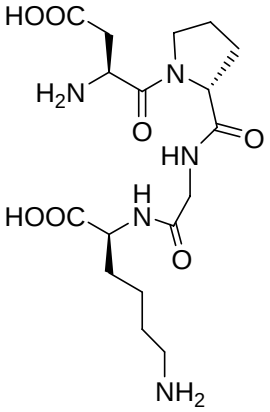
MS (MALDI-ToF): $m \cdot z^{-1}$ = 403.61 $[M+H]^+$, 425.61 $[M+Na]^+$, 441.60 $[M+K]^+$, 447.61 $[M+2Na-H]^+$

Calculated Mass: $[C_{15}H_{23}N_4O_9]^+$ = 403.37, $[C_{15}H_{22}N_4O_9Na]^+$ = 425.36, $[C_{15}H_{22}N_4O_9K]^+$ = 441.10, $[C_{15}H_{21}N_4O_9Na_2]^+$ = 447.11

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 8.42 min

Analytical RP-HPLC: Method I, t_R = 5.42 min, 100% (220 nm)

H-Asp-D-Pro-Gly-Lys-OH (HP 3)



HP 3

The synthesis of the fully protected peptide **Prot. HP 3** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Lys(Boc)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.2 mmol), and TBTU as coupling reagent. Deprotection of peptide **Prot. HP 3** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Asp(OtBu)-D-Pro-Gly-Lys(Boc)-OH (**Prot. HP 3**) 0.0865 g/0.15 mmol
H-Asp-D-Pro-Gly-Lys-OH (**HP 3**) 0.0498 g/0.12 mmol (60.0%)

Formula: $C_{17}H_{29}N_5O_7$

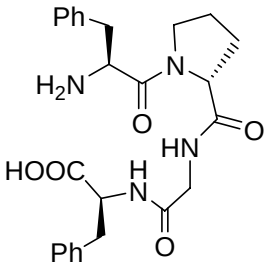
Molecular Mass: 415.20 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 416.79 $[M+H]^+$, 438.79 $[M+Na]^+$, 454.76 $[M+K]^+$, 460.78 $[M+2Na-H]^+$
Calculated Mass: $[C_{17}H_{30}N_5O_7]^+$ = 416.21, $[C_{17}H_{29}N_5O_7Na]^+$ = 438.20, $[C_{17}H_{29}N_5O_7K]^+$ = 454.17, $[C_{17}H_{28}N_5O_7Na_2]^+$ = 460.18

Preparative RP-HPLC: Apparatus A, Method 3, t_R = 7.02 min

Analytical RP-HPLC: Method I, t_R = 7.02 min, 100% (220 nm)

H-Phe-D-Pro-Gly-Phe-OH (HP 4)



HP 4

The synthesis of peptide **HP 4** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Phe-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.2 mmol), and TBTU as coupling reagent.

Yield: H-Phe-D-Pro-Gly-Phe-OH (**HP 4)** 0.0839 g/0.18 mmol (90.0%)

Formula: $C_{25}H_{30}N_4O_5$

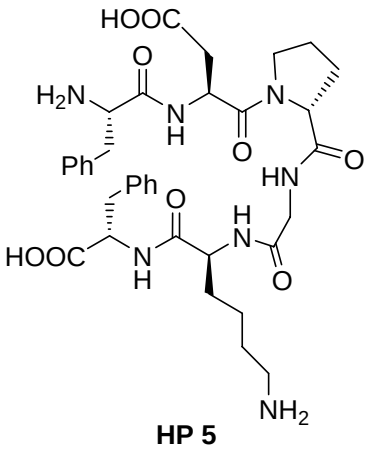
Molecular Mass: 466.22 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 467.53 $[M+H]^+$, 489.50 $[M+Na]^+$, 505.49 $[M+K]^+$, 511.52 $[M+2Na-H]^+$
Calculated Mass: $[C_{25}H_{31}N_4O_5]^+$ = 467.23, $[C_{25}H_{30}N_4O_5Na]^+$ = 489.21, $[C_{25}H_{30}N_4O_5K]^+$ = 505.19, $[C_{25}H_{29}N_4O_5Na_2]^+$ = 511.19

Preparative RP-HPLC: Apparatus A, Method 9, t_R = 29.44 min

Analytical RP-HPLC: Method I, t_R = 21.44 min, 100% (220 nm)

H-Phe-Asp-D-Pro-Gly-Lys-Phe-OH (HP 5)



HP 5

The synthesis of the fully protected peptide **Prot. HP 5** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Phe-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.2 mmol), and TBTU as coupling reagent. Deprotection of peptide **Prot. HP 5** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Phe-Asp(OtBu)-D-Pro-Gly-Lys(Boc)-Phe-OH
(**Prot. HP 5**) 0.1298 g/0.15 mmol
H-Phe-Asp-D-Pro-Gly-Lys-Phe-OH (**HP 5**) 0.0927 g/0.13 mmol (65.0%)

Formula: $C_{35}H_{47}N_7O_9$

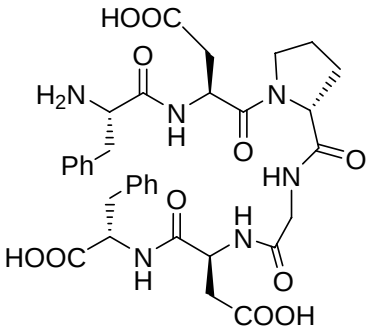
Molecular Mass: 709.34 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 710.22 $[M+H]^+$, 732.16 $[M+Na]^+$, 748.13 $[M+K]^+$, 754.14 $[M+2Na-H]^+$
Calculated Mass: $[C_{35}H_{48}N_7O_9]^+$ = 710.35, $[C_{35}H_{47}N_7O_9Na]^+$ = 732.33, $[C_{35}H_{47}N_7O_9K]^+$ = 748.31, $[C_{35}H_{46}N_7O_9Na_2]^+$ = 754.32

Preparative RP-HPLC: Apparatus A, Method 23, t_R = 21.50 min

Analytical RP-HPLC: Method I, t_R = 18.17 min, 100% (220 nm)

H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (HP 6)



HP 6

The synthesis of the fully protected peptide **Prot. HP 6** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Phe-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.3 mmol), TBTU as coupling reagent. Deprotection of

peptide **Prot. HP 6** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Phe-Asp(OtBu)-D-Pro-Gly-Asp(OtBu)-Phe-OH
(**Prot. HP 6**) 0.2183 g/0.27 mmol
H-Phe-Asp-D-Pro-Gly-Asp-Phe-OH (**HP 6**) 0.1688 g/0.24 mmol (80.0%)

Formula: $C_{33}H_{40}N_6O_{11}$

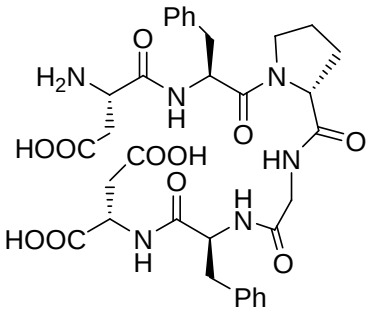
Molecular Mass: 696.28 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 697.21 $[M+H]^+$, 719.31 $[M+Na]^+$, 735.31 $[M+K]^+$, 741.33 $[M+2Na-H]^+$
Calculated MS: $[C_{33}H_{41}N_6O_{11}]^+$ = 697.28, $[C_{33}H_{40}N_6O_{11}Na]^+$ = 719.27, $[C_{33}H_{40}N_6O_{11}K]^+$ = 735.82, $[C_{33}H_{39}N_6O_{11}Na_2]^+$ = 741.69

Preparative RP-HPLC: Apparatus A, Method 11, t_R = 10.18 min

Analytical RP-HPLC: Method I, t_R = 19.24 min, 100% (220 nm)

H-Asp-Phe-D-Pro-Gly-Phe-Asp-OH (HP 7)



HP 7

The synthesis of the fully protected peptide **Prot. HP 7** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Asp(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Deprotection of peptide **Prot. HP 7** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Asp(OtBu)-Phe-D-Pro-Gly-Phe-Asp(OtBu)-OH
(**Prot. HP 7**) 0.1911 g/0.24 mmol

H-Asp-Phe-D-Pro-Gly-Phe-Asp-OH (**HP 7**) 0.1326 g/0.19 mmol (76.0%)

Formula: $C_{33}H_{40}N_6O_{11}$

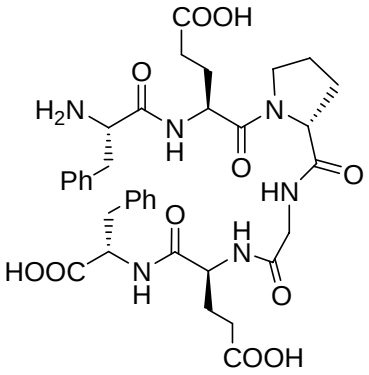
Molecular Mass: 696.28 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 697.19 $[M+H]^+$, 719.18 $[M+Na]^+$, 735.15 $[M+K]^+$
Calculated Mass: $[C_{33}H_{41}N_6O_{11}]^+$=697.28, $[C_{33}H_{40}N_6O_{11}Na]^+$=719.27, $[C_{33}H_{40}N_6O_{11}K]^+$=735.24

Preparative RP-HPLC: Apparatus A, Method 12, t_R=23.82 min

Analytical RP-HPLC: Method I, t_R = 21.12 min, 100% (220 nm)

H-Phe-Glu-D-Pro-Gly-Glu-Phe-OH (HP 8)



HP 8

The synthesis of the fully protected peptide **Prot. HP 8** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Phe-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Deprotection of the peptide **Prot. HP 8** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Phe-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)-Phe-OH
(**Prot. HP 8**) 0.2157 g/0.25 mmol
H-Phe-Glu-D-Pro-Gly-Glu-Phe-OH (**HP 8**) 0.1355 g/0.19 mmol (76.0%)

Formula: $C_{35}H_{44}N_6O_{11}$

Molecular Mass: 724.31 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 725.31 $[M+H]^+$, 747.34 $[M+Na]^+$, 763.37 $[M+K]^+$, 769.41

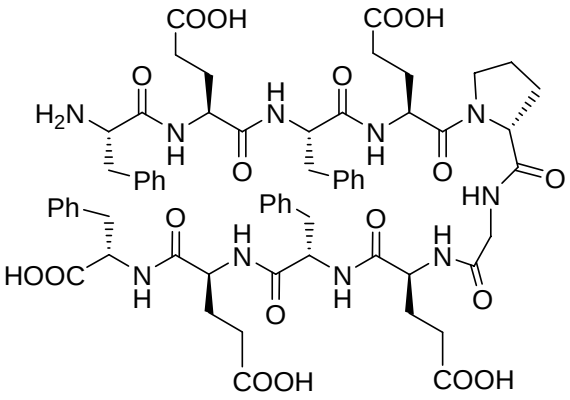
$[M+2Na-H]^+$

Calculated Mass: $[C_{35}H_{45}N_6O_{11}]^+$ = 725.31, $[C_{35}H_{44}N_6O_{11}Na]^+$ = 747.30, $[C_{35}H_{44}N_6O_{11}K]^+$ = 763.27, $[C_{35}H_{43}N_6O_{11}Na_2]^+$ = 769.28

Preparative RP-HPLC: Apparatus A, Method 12, t_R = 26.55 min

Analytical RP-HPLC: Method I, t_R = 18.90 min, 100% (220 nm)

H-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-OH (HP 9)



HP 9

The synthesis of the fully protected peptide **Prot. HP 9** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Phe-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Deprotection of the peptide **Prot. HP 9** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield: H-Phe-Glu(tBu)-Phe-Glu(tBu)-D-Pro-Gly-Glu(tBu)-Phe-Glu(tBu)-Phe-OH (**Prot. HP 9**) 0.3598 g/0.24 mmol

H-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-OH (**HP 9**) 0.2297 g/0.18 mmol (72.0%)

Formula: $C_{63}H_{76}N_{10}O_{19}$

Molecular Mass: 1276.53 g/mol

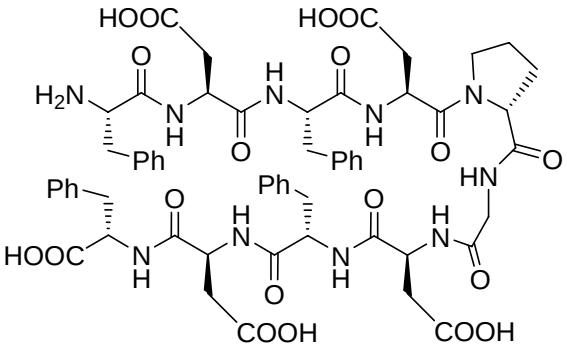
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1278.10 $[M+H]^+$, 1299.26 $[M+Na]^+$, 1315.33 $[M+K]^+$, 1321.68 $[M+2Na-H]^+$

Calculated Mass: $[C_{63}H_{77}N_{10}O_{19}]^+$ = 1277.54, $[C_{63}H_{76}N_{10}O_{19}Na]^+$ = 1299.52, $[C_{63}H_{76}N_{10}O_{19}K]^+$ = 1315.49, $[C_{63}H_{75}N_{10}O_{19}Na_2]^+$ = 1321.50

Preparative RP-HPLC: Apparatus A, Method 13, t_R = 29.01 min

Analytical RP-HPLC: Method I, t_R = 19.03 min, 100% (220nm)

H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-OH (HP 10)



HP 10

The synthesis of the fully protected peptide **Prot. HP 10** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Phe-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.3 mmol), and TBTU as coupling reagent. Deprotection of the peptide **Prot. HP 10** was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ethane-1,2-dithiol 82.5:5:5:5:2.5). Aspartimide formation took place if the applied deprotection cocktail is composed of TFA/H_2O/TIS (95:2.5:2.5).

Yield:

H-Phe-Asp(OtBu)-Phe-Asp(OtBu)-D-Pro-Gly-Asp(OtBu)-Phe-Asp(OtBu)-Phe-OH (**Prot. HP 10**) 0.2889 g/0.20 mmol

H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-OH

(**HP 10**) 0.1831 g/0.15 mmol (50.0%)

Formula: $C_{59}H_{68}N_{10}O_{19}$

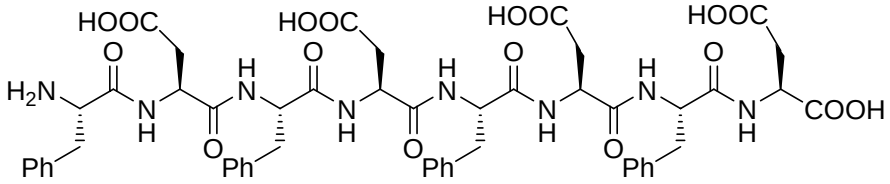
Molecular Mass: 1220.47 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 1221.26 $[M+H]^+$, 1243.77 $[M+Na]^+$, 1259.76 $[M+K]^+$, 1265.75 $[M+2Na-H]^+$
Calculated Mass: $[C_{59}H_{69}N_{10}O_{19}]^+$ = 1221.47, $[C_{59}H_{68}N_{10}O_{19}Na]^+$ = 1243.46, $[C_{59}H_{68}N_{10}O_{19}K]^+$ = 1259.43, $[C_{59}H_{67}N_{10}O_{19}Na_2]^+$ = 1265.44

Preparative RP-HPLC: Apparatus B, Method 14, t_R = 21.9 min

Analytical RP-HPLC: Method I, t_R = 19.75 min, 100% (220nm)

H-Phe-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 11)



HP 11

The synthesis of the fully protected peptide **Prot. HP 11** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Asp(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Aspartimide formation in the peptide **Prot. HP 11** was found. Deprotection of the peptide was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ethane-1,2-dithiol, 82.5:5:5:5:2.5). Aspartimide derivatives were separated from the target peptide **HP 60** through RP-HPLC operations.

H-Phe-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-OH (**Prot. HP 11**)
Yield (**HP 11**) (crude peptide): 0.3236g/0.251mmol (100%)

Formula (**Prot. HP 11**): $C_{68}H_{90}N_8O_{17}$

Formula (**Prot. HP 11** Aspartimide Formation): $C_{64}H_{80}N_8O_{16}$

Formula (**HP 11**): $C_{52}H_{58}N_8O_{17}$

Formula (**HP 11** Aspartimide Formation): $C_{52}H_{56}N_8O_{16}$

Molecular Mass (**Prot. HP 11**): 1290.64 g/mol

Molecular Mass (**Prot. HP 11** Aspartimide Formation): 1216.57 g/mol

Molecular Mass (**HP 11**): 1066.39 g/mol

Molecular Mass (**HP 11** Aspartimide Formation): 1048.38 g/mol

MS (MALDI-ToF)(**Prot. HP 11**): $m \cdot z^{-1}$ = 1291.69 $[M+H]^+$, 1313.59 $[M+Na]^+$, 1329.04 $[M+K]^+$, 1335.44 $[M+2Na-H]^+$

Calculated Mass (**Prot. HP 11**): $[C_{68}H_{91}N_8O_{17}]^+$ = 1291.65, $[C_{68}H_{90}N_8O_{17}Na]^+$ = 1313.63, $[C_{68}H_{90}N_8O_{17}K]^+$ = 1329.61, $[C_{68}H_{89}N_8O_{17}Na_2]^+$ = 1335.61

MS (MALDI-ToF)(**Prot. HP 11** Aspartimide Formation): $m \cdot z^{-1}$ = 1217.45 $[M+H]^+$, 1239.14 $[M+Na]^+$

Calculated Mass (**Prot. HP 11** Aspartimide Formation): $[C_{64}H_{81}N_8O_{16}]^+$ = 1217.58, $[C_{64}H_{80}N_8O_{16}Na]^+$ = 1239.56

MS (MALDI-ToF)(**HP 11**): $m \cdot z^{-1}$ = 1067.75 $[M+H]^+$, 1089.73 $[M+Na]^+$, 1105.35 $[M_{4\text{-}12}+K]^+$

Calculated Mass (**HP 11**): $[C_{52}H_{59}N_8O_{17}]^+$ = 1067.40, $[C_{52}H_{58}N_8O_{17}Na]^+$ = 1089.38, $[C_{52}H_{58}N_8O_{17}K]^+$ = 1105.36

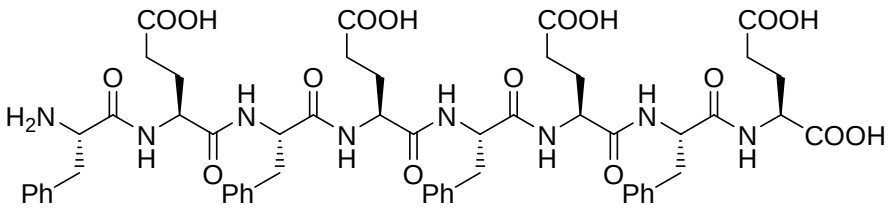
MS (MALDI-ToF)(**HP 11** Aspartimide Formation): $m \cdot z^{-1}$ = 1049.08 $[M+H]^+$, 1071.26 $[M+Na]^+$

Calculated Mass (**HP 11** Aspartimide Formation): $[C_{52}H_{57}N_8O_{16}]^+$ = 1049.39, $[C_{52}H_{56}N_8O_{16}Na]^+$ = 1071.37

Preparative RP-HPLC: Apparatus B, Method 15, t_R = 23.6 min

Analytical RP-HPLC: Method I, t_R = 23.07 min, 100% (220 nm)

H-Phe-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 12)



HP 12

The synthesis of the fully protected peptide **Prot. HP 12** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Glu(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Deprotection of peptide **Prot. HP 12** was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ethane-1,2-dithiol, 82.5:5:5:5:2.5).

Yield:

H-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-OH (**Prot. HP 12**)

(no yield data)

H-Phe-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (**HP 12**) 0.1235 mg/0.11 mmol (44.0%)

Formula: $C_{56}H_{66}N_8O_{17}$

Molecular Mass: 1122.45 g/mol

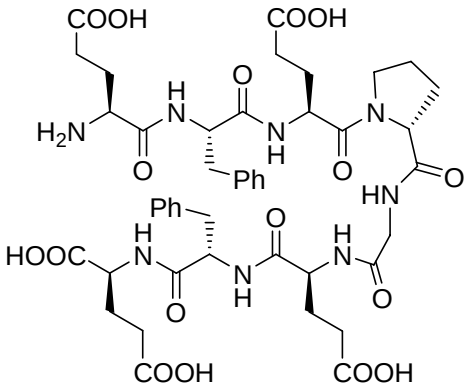
MS (MALDI-ToF): $m \cdot z^{-1}$ = 1123.51 $[M+H]^+$, 1145.47 $[M+Na]^+$, 1161.42 $[M+K]^+$, 1167.35 $[M+2Na-H]^+$

Calculated Mass: $[C_{56}H_{67}N_8O_{17}]^+$ = 1123.46, $[C_{56}H_{66}N_8O_{17}Na]^+$ = 1145.44, $[C_{56}H_{66}N_8O_{17}K]^+$ = 1161.42, $[C_{56}H_{65}N_8O_{17}Na_2]^+$ = 1167.43

Preparative RP-HPLC: Apparatus B, Method 15, t_R = 24.8 min

Analytical RP-HPLC: Method II, t_R = 19.63 min, 100% (220 nm)

H-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-OH (HP 13)



HP 13

The synthesis of the fully protected peptide **Prot. HP 13** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Glu(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Deprotection of peptide **Prot. HP 13** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5).

Yield:

H-Glu(OtBu)-Phe-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)-Phe-Glu(OtBu)-OH (**Prot. HP 13**)

0.2481 g/0.23 mmol (crude product, after lyophilization)

0.1290 g/0.11 mmol (after HPLC)

H-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-OH (**HP 13**) 0.0884 g/0.09 mmol (36.0%)

Formula: $C_{45}H_{58}N_8O_{17}$

Molecular Mass: 982.39 g/mol

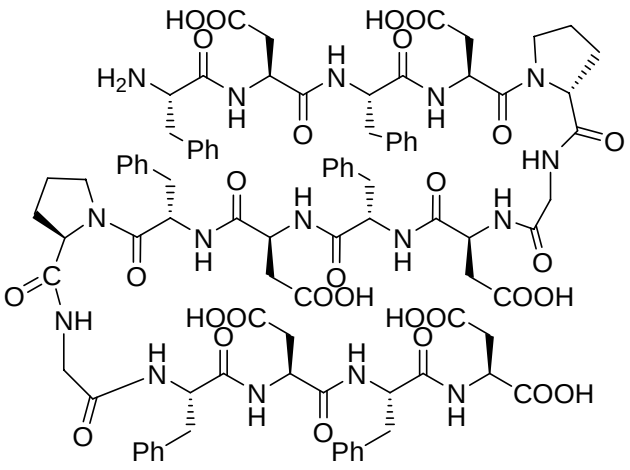
MS (MALDI-ToF): $m \cdot z^{-1}$ = 983.71 $[M+H]^+$, 1005.11 $[M+Na]^+$, 1021.14 $[M+K]^+$, 1043.08 $[M+Na+K-H]^+$

Calculated Mass: $[C_{45}H_{59}N_8O_{17}]^+$ = 983.40, $[C_{45}H_{58}N_8O_{17}Na]^+$ = 1005.38, $[C_{45}H_{58}N_8O_{17}K]^+$ = 1021.36, $[C_{45}H_{57}N_8O_{17}KNa]^+$ = 1043.34

Preparative RP-HPLC: Apparatus A, Method 13, t_R = 25.11 min

Analytical RP-HPLC: Method I, t_R = 15.05 min, 100% (220nm)

H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-D-Pro-Gly-Phe-Asp-Phe-Asp-OH (HP 14)



HP 14

The synthesis of the fully protected peptide **Prot. HP 14** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Asp(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain no HOBt as additive. Aspartimide formation in the sequence **Prot. HP 14** was found. Deprotection of peptide **Prot. HP 14** was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ ethane-1,2-dithiol, 82.5:5:5:5:2.5). The deprotection reaction at ambient temperature last 1.5 hours before the quantitative cleavage of the side-chain protecting groups was achieved. Aspartimide derivatives were separated from the target peptide by RP-HPLC operations.

Yield:

H-Phe-Asp(OtBu)-Phe-Asp(OtBu)-D-Pro-Gly-Asp(OtBu)-Phe-Asp(OtBu)-Phe-D-Pro-Gly-Phe-Asp(OtBu)-Phe-Asp(OtBu)-OH (**Prot. HP 14**) (no yield data)

H-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-D-Pro-Gly-Phe-Asp-Phe-Asp-OH (**HP 14**) 0.0190 g/0.01 mmol (4.0%)

Formula (**Prot. HP 14**): $C_{116}H_{154}N_{16}O_{29}$

Formula (**HP 14**): $C_{92}H_{106}N_{16}O_{29}$

Formula (**Prot. HP 14** Aspartimide Formation): $C_{112}H_{144}N_{16}O_{28}$

Formula (**HP 14** Aspartimide Formaiton): $C_{92}H_{104}N_{16}O_{28}$

Molecular Mass (**Prot. HP 14**): 2235.11 g/mol

Molecular Mass (**HP 14**): 1898.73 g/mol

Molecular Mass (**Prot. HP 14** Aspartimide Formation): 2161.04 g/mol

Molecular Mass (**HP 14** Aspartimide Formation): 1880.72 g/mol

MS (MALDI-ToF)(**Prot. HP 14**): $m \cdot z^{-1}$ = 2257.33 $[M+Na]^+$, 2273.77 $[M+K]^+$

Calculated Mass: $[C_{116}H_{154}N_{16}O_{29}Na]^+$ = 2258.10, $[C_{116}H_{154}N_{16}O_{29}K]^+$ = 2274.07

MS (MALDI-ToF)(**Prot. HP 14** Aspartimide Formation): $m \cdot z^{-1}$ = 2183.07 $[M+Na]^+$, 2200.34 $[M+K]^+$

Calculated Mass: $[C_{112}H_{144}N_{16}O_{28}Na]^+$ = 2184.02, $[C_{112}H_{144}N_{16}O_{28}K]^+$ = 2200.00

MS (MALDI-ToF)(**HP 14**): $m \cdot z^{-1}$ = 1899.03 $[M+H]^+$, 1921.27 $[M+Na]^+$, 1939.52 $[M+K]^+$

Calculated Mass: $[C_{92}H_{107}N_{16}O_{29}]^+$ = 1899.74, $[C_{92}H_{106}N_{16}O_{29}Na]^+$ = 1921.72, $[C_{92}H106N_{16}O_{29}K]^+$ = 1937.69

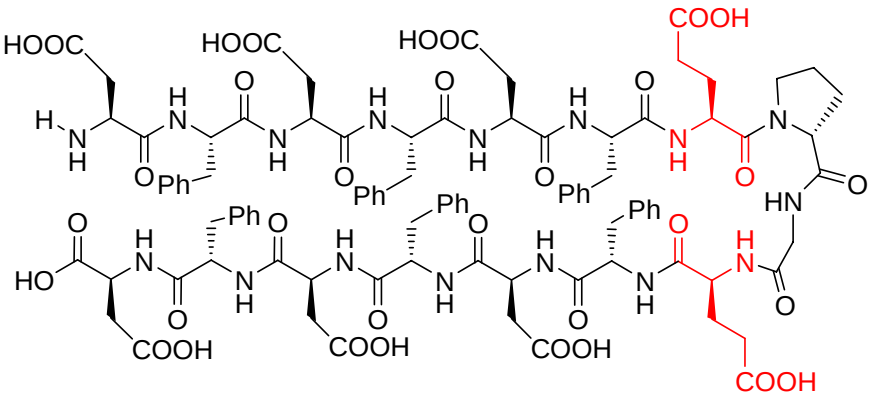
MS (MALDI-ToF)(**HP 14** Aspartimide Formation): $m \cdot z^{-1}$ = 1903.17 $[M+Na]^+$, 1919.59 $[M+K]^+$

Calculated Mass: $[C_{112}H_{144}N_{16}O_{28}Na]^+$ = 1903.71, $[C_{112}H_{144}N_{16}O_{28}K]^+$ = 1919.68

Preparative RP-HPLC: Apparatus A, Method 16, t_R = 15.58 min

Analytical RP-HPLC: Method II, t_R = 31.47 min, 100% (220 nm)

H-Asp-Phe-Asp-Phe-Asp-Phe-Glu-D-Pro-Gly-Glu-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 15 A)



HP 15 A

The synthesis of the fully protected peptide **Prot. HP 15 A** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Asp(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Aspartimide formation in **Prot. HP 15 A** was found. Deprotection of peptide **Prot. HP 15 A** was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ethane-1,2-dithiol 82.5:5:5:5:2.5). The deprotection reaction at ambient temperature took 3 hours before the quantitative cleavage of the side-chain protecting groups was achieved. Aspartimide derivatives were separated from the target peptide by RP-HPLC operations.

Yield:

H-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-OH (**Prot. HP 15 A**) (no yield data)

H-Asp-Phe-Asp-Phe-Asp-Phe-Glu-D-Pro-Gly-Glu-Phe-Asp-Phe-Asp-Phe-Asp-OH (**HP 15 A**) 0.0205 g/0.01 mmol (4.0%)

Formula (**Prot. HP 15 A**): $C_{127}H_{174}N_{16}O_{33}$

Formula (**HP 15 A**): $C_{95}H_{110}N_{16}O_{33}$

Formula (**Prot. HP 15 A** Aspartimide): $C_{123}H_{164}N_{16}O_{32}$

Formula (**HP 15 A** Aspartimide): $C_{95}H_{108}N_{16}O_{32}$

Molecular Mass (**Prot. HP 15 A**): 2451.24 g/mol

Molecular Mass (**HP 15 A**): 2002.74 g/mol

Molecular Mass (**Prot. HP 15 A** Aspartimide): 2377.17 g/mol

Molecular Mass (**HP 15 A** Aspartimide): 1984.73 g/mol

MS (MALDI-ToF)(**Prot. HP 15 A**): $m \cdot z^{-1}$ = 2452.64 $[M+H]^+$, 2474.78 $[M+Na]^+$, 2490.90 $[M+K]^+$, 2496.56.44 $[M+2Na-H]^+$

Calculated Mass (**Prot. HP 15 A**): $[C_{127}H_{175}N_{16}O_{33}]^+$ = 2452.25, $[C_{127}H_{174}N_{16}O_{33}Na]^+$ = 2474.23, $[C127H174N16O33K]^+$ = 2490.20, $[C_{127}H_{173}N_{16}O_{33}Na_2]^+$ = 2496.21

MS (MALDI-ToF)(**HP 15 A**): $m \cdot z^{-1}$ = 2003.71 $[M+H]^+$, 2024.68 $[M+Na]^+$, 2045.98 $[M+K]^+$

Calculated Mass (**HP 15 A**): $[C_{95}H_{111}N_{16}O_{33}]^+$ = 2003.75, $[C_{95}H_{110}N_{16}O_{33}Na]^+$ = 2025.73, $[C_{95}H_{110}N_{16}O_{33}K]^+$ = 2043.13

MS (MALDI-ToF)(**Prot. HP 15 A** Aspartimide Formation): $m \cdot z^{-1}$ = 2378.37 $[M+H]^+$, 2399.21 $[M+Na]^+$, 2419.10 $[M+K]^+$, 2437.08 $[M+Na+K-H]^+$

Calculated Mass (**Prot. HP 15 A** Aspartimide Formation): $[C_{123}H_{165}N_{16}O_{32}]^+$ = 2378.18, $[C_{123}H_{164}N_{16}O_{32}Na]^+$ = 2400.16, $[C_{123}H_{164}N_{16}O_{32}K]^+$ = 2416.13, $[C_{122}H_{164}N_{16}O_{32}KNa]^+$ = 2438.11

MS (MALDI-ToF)(**HP 15 A** Aspartimide Formation): $m \cdot z^{-1}$ = 1985.26 $[M+H]^+$, 2007.41 $[M+Na]^+$, 2029.27 $[M+2Na-H]^+$

Calculated Mass (**HP 15 A** Aspartimide Formation): $[C_{95}H_{109}N_{16}O_{32}]^+$ = 1985.26, $[C_{95}H_{108}N_{16}O_{32}Na]^+$ = 2007.72.16, $[C_{95}H_{107}N_{16}O_{32}Na_2]^+$ = 2029.71

HP 15 A:

Preparative RP-HPLC: Apparatus A, Method 17, t_R = 27.45min

Analytical RP-HPLC: Method II, t_R = 21.39 min, 100% (220 nm)

Analysis of the Proportion of Aspartimide Formation in the Synthesis of Prot. HP 15 A

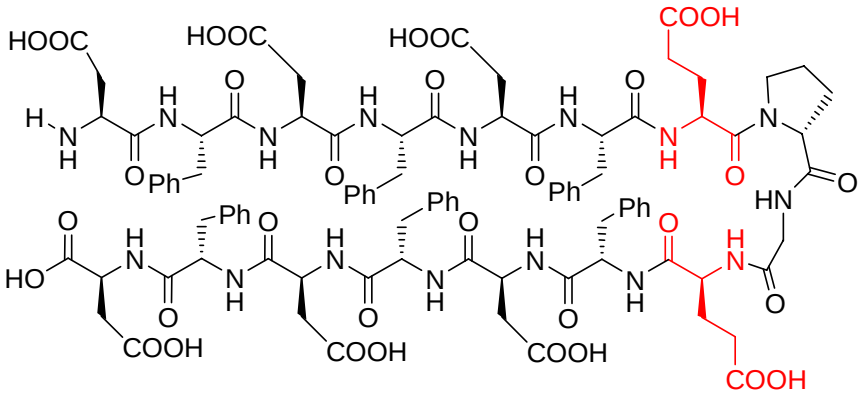
The crude peptide **Prot. HP 15 A** was cleaved from the resin; the solvent was removed under reduced pressure. The remaining residue was lyophilized before analysis by analytical RP-HPLC (Method II). 1-2 mg of crude peptide **Prot. HP 15 A** was dissolved in acetonitrile/H_2O with the addition of minimal amount of DMF in order to increase the solubility. Two fractions were obtained from the chromatography, both of which were collected and analyzed by MALDI-ToF-MS.

First fraction: (t_R = 36.63-36.92 min, 35.38 %)
MS (MALDI-ToF) : $m \cdot z^{-1}$ = 2399.69 $[M+Na]^+$, 2419.54 $[M+K]^+$, 2438.62 $[M+Na+K-H]^+$

Second fraction: (t_R = 37.94 min, 64.62 %)
MS (MALDI-ToF) : $m \cdot z^{-1}$ = 2474.98 $[M+Na]^+$, 2491.93 $[M+K]^+$, 2496.53 $[M+2Na-H]^+$

The first fraction sampled from analytical RP-HPLC was found to be **Prot. HP 15 A** with a single aspartimide, while the second fraction was the target peptide **Prot. HP 15 A**, which means that ca. 35.38 % of aspartimide formation of **Prot. HP 15 A** was caused with this method of peptide synthesis (microwave as the input energy, no HOBt additive in the Fmoc deprotection cocktail 20% piperidine/DMF).

H-Asp-Phe-Asp-Phe-Asp-Phe-Glu-D-Pro-Gly-Glu-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 15 B)



HP 15 B

The synthesis of the fully protected peptide **Prot. HP 15 B** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Asp(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), TBTU as coupling reagent. The unique difference of synthetic methodology of peptide **HP 15 B** to that of **HP 15 A** was that the N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF contained 0.1 M HOBt as additive for the synthesis of **Prot. HP 15 B**. Aspartimide formation in the sequence **Prot. HP 15 B** was also found, but to a lower extent (15.36%:35.38%) than that from the **Prot. HP 15 A**.

Deprotection of the peptide was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ethane-1,2-dithiol 82.5:5:5:5:2.5). The deprotection reaction at ambient temperature took 3 hours before the quantitative cleavage of the side-chain protecting groups was achieved. Aspartimide derivatives were separated from the target peptide by RP-HPLC.

H-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-Phe-Glu(OtBu)-D-Pro-Gly-Glu(OtBu)
-Phe -Asp(OtBu)-Phe-Asp(OtBu)-Phe-Asp(OtBu)-OH (**Prot. HP 15 B**)
H-Asp-Phe-Asp-Phe-Asp-Phe-Glu-D-Pro-Gly-Glu-Phe-Asp-Phe-Asp-Phe-Asp-OH
(**HP 15 B**) 0.0420 g/0.02 mmol (8.0%)

Formula (**Prot. HP 15 B**): $C_{127}H_{174}N_{16}O_{33}$
Formula (**HP 15 B**): $C_{95}H_{110}N_{16}O_{33}$
Formula (**Prot. HP 15 B** Aspartimide): $C_{123}H_{164}N_{16}O_{32}$
Formula (**HP 15 B** Aspartimide): $C_{95}H_{108}N_{16}O_{32}$

Molecular Mass (**Prot. HP 15 B**): 2451.24 g/mol
Molecular Mass (**HP 15 B**): 2002.74 g/mol
Molecular Mass (**Prot. HP 15 B** Aspartimide): 2377.17 g/mol
Molecular Mass (**HP 15 B** Aspartimide): 1984.73 g/mol

MS (MALDI-ToF)(**Prot. HP 15 B**): $m \cdot z^{-1}$ = 2452.47 $[M+H]^+$, 2473.67 $[M+Na]^+$, 2489.82 $[M+K]^+$, 2495.20 $[M+2Na-H]^+$, 2511.73 $[M+Na+K-H]^+$

Calculated Mass (**Prot. HP 15 B**): $[C_{127}H_{175}N_{16}O_{33}]^+$ = 2452.25, $[C_{127}H_{174}N_{16}O_{33}Na]^+$ = 2474.23, $[C_{127}H_{174}N_{16}O_{33}K]^+$ = 2490.20, $[C_{127}H_{173}N_{16}O_{33}Na_2]^+$ = 2496.21, $[C_{127}H_{173}N_{16}O_{33}KNa]^+$ = 2512.19

MS (MALDI-ToF)(**Prot. HP 15 B** Aspartimide Formation): $m{\cdot}z^{-1}$ = 2399.21 $[M+Na]^+$, 2419.70 $[M+K]^+$
Calculated Mass (**Prot. HP 15 B** Aspartimide Formation): $[C_{123}H_{164}N_{16}O_{32}Na]^+$ = 2400.16, $[C_{123}H_{164}N_{16}O_{32}K]^+$ = 2416.13

Analysis of the Proportion of Aspartimide Formation in the Synthesis of Prot. HP 15 B

The crude peptide **Prot. HP 15 B** was cleaved from the resin; the solvent was removed under the reduced pressure. The remaining residue was lyophilized analyzed by analytical RP-HPLC (Method II). 1-2 mg of crude peptide **Prot. HP 15 B** was dissolved in acetonitrile/H_2O with the addition of a minimal amount of DMF in order to increase the solubility. Two fractions were obtained from the chromatography, both of which were collected and analyzed by MALDI-ToF-MS.

First fraction: (t_R = 36.57-36.99 min, 15.36 %)
MS (MALDI-ToF) : $m{\cdot}z^{-1}$ = 2399.69 $[M+Na]^+$, 2419.54 $[M+K]^+$, 2438.62 $[M+Na+K-H]^+$

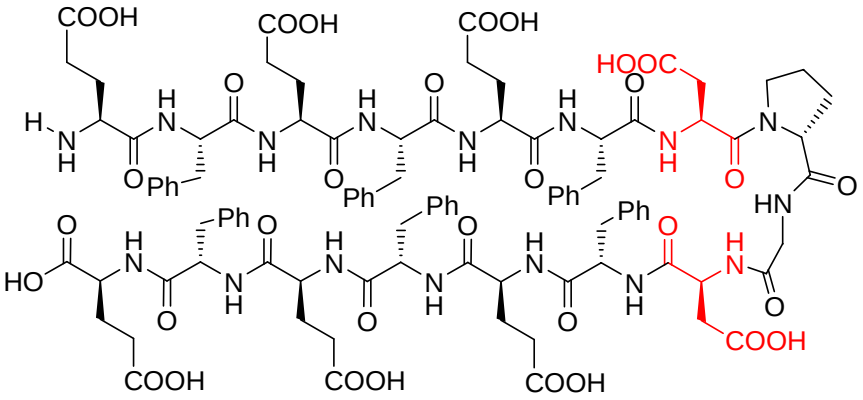
Second fraction: (t_R = 38.39 min, 84.64 %)
MS (MALDI-ToF) : $m{\cdot}z^{-1}$ = 2474.56 $[M+Na]^+$, 2491.89 $[M+K]^+$

The first fraction sampled from analytical RP-HPLC was approved to be the mono-aspartimide formation of **Prot. HP 15 B**, while the second fraction was the target peptide **Prot. HP 15 B**, which means that only ca. 15% of aspartimide formation of **Prot. HP 15 B** was observed with this method of peptide synthesis (microwave as the input energy, 0.1 M HOBt additive in the Fmoc deprotection cocktail 20% piperidine/DMF).

The addition of 0.1 M HOBt to 20 % piperidine/DMF significantly decreased the occurrence of aspartimide formation in the template peptide **Prot. HP 15** in this project,

with the microwave as the input energy in the peptide synthesis process. The explanation for this phenomenon is still to be explored.

H-Glu-Phe-Glu-Phe-Glu-Phe-Asp-D-Pro-Gly-Asp-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 16)



HP 16

The synthesis of the fully protected peptide **Prot. HP 16** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Glu(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. No aspartimide formation was found. Deprotection of peptide **Prot. HP 16** was achieved with the deprotection cocktail Reagent K (TFA/Phenol/H_2O/thiolanisole/ethane-1,2-dithiol 82.5:5:5:5:2.5). The deprotection reaction at ambient temperature took 3 hours before the quantitative cleavage of the side-chain protecting groups was achieved.

Yield:

H-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Asp(OtBu)-D-Pro-Gly-Asp(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-Phe-Glu(OtBu)-OH (**Prot. HP 16**) (no yield data)

H-Glu-Phe-Glu-Phe-Glu-Phe-Asp-D-Pro-Gly-Asp-Phe-Glu-Phe-Glu-Phe-Glu-OH (**HP 16**) 0.0895 g/0.043 mmol (17.2%)

Formula (**HP 16**): $C_{99}H_{118}N_{16}O_{33}$

Molecular Mass (**HP 16**): 2058.80 g/mol

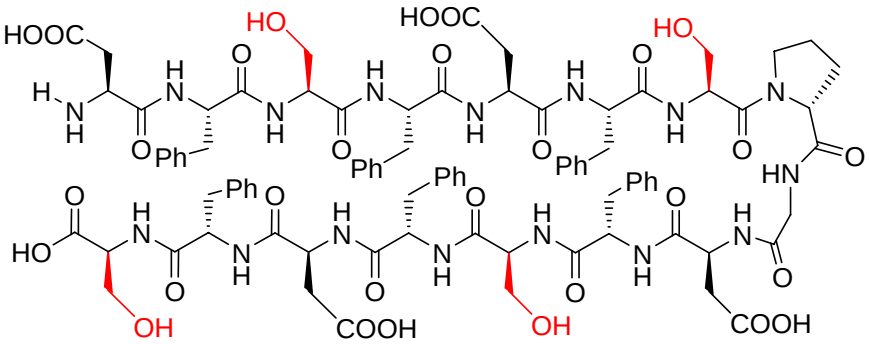
MS (MALDI-ToF)(**HP 16**): $m \cdot z^{-1}$ = 2059.98 $[M+H]^+$, 2081.96 $[M+Na]^+$, 2097.99 $[M+K]^+$

Calculated Mass (**HP 16**): $[C_{99}H_{118}N_{16}O_{33}]^+$ = 2059.81, $[C_{99}H_{118}N_{16}NaO_{33}]^+$ = 2081.79, $[C_{99}H_{118}KN_{16}O_{33}]^+$ = 2097.77

Preparative RP-HPLC: Apparatus A, Method 17, t_R = 29.56 min

Analytical RP-HPLC: Method II, t_R = 22.51 min, 100% (220 nm)

H-Asp-Phe-Ser-Phe-Asp-Phe-Ser-D-Pro-Gly-Asp-Phe-Ser-Phe-Asp-Phe-Ser-OH (HP 17)



HP 17

The synthesis of the fully protected derivative of peptide **Prot. HP 17** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Ser(tBu)-OH being loaded on the *o*-chlorotrityl resin (0.76 mmol/g, 0.25 mmol), and TBTU as coupling reagent. 4 equiv amino acid and 4 equiv TBTU were applied in each coupling step. Deprotection of peptide **Prot. HP 17** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5). However, quantitative cleavage of the side-chain protecting groups was not obtained after 8 hours deprotection reaction at ambient temperature.

Yield:

H-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-Asp(OtBu)-Phe-Ser(tBu)-Phe-Asp(OtBu)-Phe-Ser(tBu)-OH (**Prot. HP 17**)
(0.8541 g, crude product, after lyophilization)

H-Asp-Phe-Ser-Phe-Asp-Phe-Ser-D-Pro-Gly-Asp-Phe-Ser-Phe-Asp-Phe-Ser-OH
(**HP 17**) 0.0155 g/0.0083 mmol (3.3%)

Formula: $C_{89}H_{106}N_{16}O_{29}$

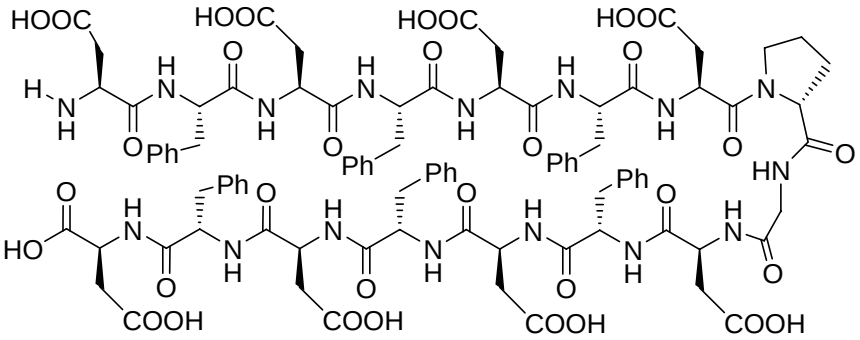
Molecular Mass: 1863.94 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 1886.83 $[M+Na]^+$, 1904.83 $[M+K]^+$, 1910.61 $[M+2Na-H]^+$
Calculated Mass: $[C_{89}H_{106}N_{16}O_{29}Na]^+$ = 1885.73, $[C_{89}H_{106}N_{16}O_{29}K]^+$ = 1903.03, $[C_{89}H_{105}N_{16}O_{29}Na_2]^+$ = 1908.91

Preparative RP-HPLC: Apparatus A, Method 18, t_R = 29.30 min

Analytical RP-HPLC: Method II, t_R = 21.97 min, 100% (220nm)

H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (HP 18)



HP 18

The synthesis of the fully protected peptide **Prot. HP 18** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Asp(OtBu)-OH being loaded

on the *o*-chlorotrityl resin (0.8 mmol/g, 0.3 mmol), and TBTU as coupling reagent. Deprotection of peptide **Prot. HP 18** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5). No aspartimide formation was found in the crude product.

Yield:
H-Asp-Phe-Asp-Phe-Asp-Phe-Asp-D-Pro-Gly-Asp-Phe-Asp-Phe-Asp-Phe-Asp-OH (**HP 18**) 0.0205 g/0.0010 mmol (0.33%)

Formula: $C_{93}H_{106}N_{16}O_{33}$

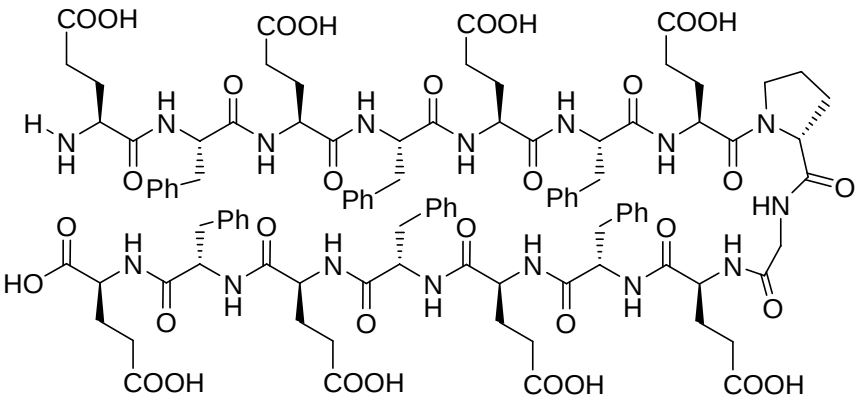
Molecular Mass: 1975.98 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 1978.37 $[M+H]^+$, 1999.42 $[M+Na]^+$, 2015.59 $[M+K]^+$, 2021.27 $[M+2Na-H]^+$, 2043.09 $[M+3Na-2H]^+$, 2064.99 $[M+4Na-3H]^+$
Calculated Mass: $[C_{93}H_{107}N_{16}O_{33}]^+$ = 1976.99, $[C_{93}H_{106}N_{16}O_{33}Na]^+$ = 1998.97, $[C_{93}H_{106}N_{16}O_{33}K]^+$ = 2015.08, $[C_{93}H_{105}N_{16}O_{33}Na_2]^+$ = 2020.95, $[C_{93}H_{104}N_{16}O_{33}Na_3]^+$ = 2042.94, $[C_{93}H_{103}N_{16}O_{33}Na_4]^+$ = 2064.93

Preparative RP-HPLC: Apparatus A, Method 18, t_R = 31.00 min

Analytical RP-HPLC: Method II, t_R = 21.18 min, 100% (220 nm)

H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH (HP 19)



HP 19

The synthesis of the fully protected peptide **Prot. HP 19** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Glu(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.3 mmol), and TBTU as coupling reagent. Deprotection of peptide **Prot. HP 19** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5). The deprotection reaction at ambient temperature took 2 hours before the quantitative cleavage of the side-chain protecting groups was achieved.

Yield:
H-Glu-Phe-Glu-Phe-Glu-Phe-Glu-D-Pro-Gly-Glu-Phe-Glu-Phe-Glu-Phe-Glu-OH
(**HP 19**) 0.0085 g/0.0041 mmol

Formula: $C_{101}H_{122}N_{16}O_{33}$

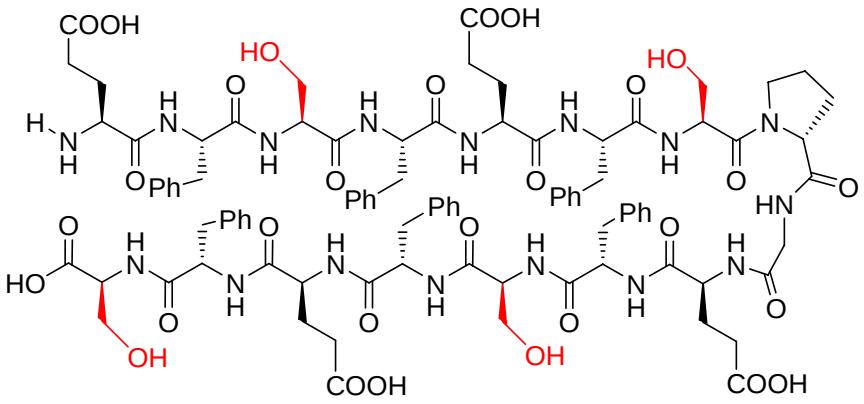
Molecular Mass: 2086.84 g/mol

MS (MALDI-ToF): $m \cdot z^{-1}$ = 2086.92 $[M+H]^+$, 2108.32 $[M+Na]^+$, 2124.35 $[M+K]^+$, 2130.26 $[M+2Na-H]^+$, 2152.09 $[M+3Na-2H]^+$, 2174.76 $[M+4Na-3H]^+$
Calculated Mass: $[C_{101}H_{123}N_{16}O_{33}]^+$ = 2087.84, $[C_{101}H_{122}N_{16}O_{33}Na]^+$ = 2109.83, $[C_{101}H_{122}N_{16}O_{33}K]^+$ = 2125.80, $[C_{101}H_{121}N_{16}O_{33}Na_2]^+$ = 2131.81, $[C_{101}H_{120}N_{16}O_{33}Na_3]^+$ = 2153.80, $[C_{101}H_{119}N_{16}O_{33}Na_4]^+$ = 2175.79

Preparative RP-HPLC: Apparatus A, Method 19, t_R = 21.38 min

Analytical RP-HPLC: Method II, t_R = 21.84 min, 100% (220 nm)

H-Glu-Phe-Ser-Phe-Glu-Phe-Ser-D-Pro-Gly-Glu-Phe-Ser-Phe-Glu-Phe-Ser-OH (HP 20-1) (Manual Synthesis)



HP 20-1

The synthesis of the fully protected peptide **Prot. HP 20-1** was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Ser(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.76 mmol/g, 0.25 mmol), and TBTU as coupling reagent. 4 equiv amino acid and 4 equiv TBTU were applied in each coupling step. Deprotection of the peptide was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5). The deprotection reaction at ambient temperature took 3 hours before the quantitative cleavage of the side-chain protecting groups was observed.

Yield:

H-Glu(OtBu)-Phe-Ser(tBu)-Phe-Glu(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-Glu(OtBu)-Phe-Ser(tBu)-Phe-Glu(OtBu)-Phe-Ser(tBu)-OH **(Prot. HP 20-1)**

(0.4357 g/0.184 mmol, crude product, after lyophilization)

H-Glu-Phe-Ser-Phe-Glu-Phe-Ser-D-Pro-Gly-Glu-Phe-Ser-Phe-Glu-Phe-Ser-OH (**HP 20-1**) 0.0112 g/0.0058 mmol (2.32%)

Formula of **Prot. HP 20-1**: $C_{125}H_{178}N_{16}O_{29}$

Molecular Mass (**Prot. HP 20-1**): 2367.29 g/mol

MS (MALDI-ToF) (**Prot. HP 20-1**): $m \cdot z^{-1}$ = 2390.68 $[M+Na]^+$, 2406.70 $[M+K]^+$,

2412.69 $[M+2Na-H]^+$

Calculated Mass (**Prot. HP 20-1**): $[C_{125}H_{178}N_{16}O_{29}Na]^+$ = 2390.28, $[C_{125}H_{178}N_{16}O_{29}K]^+$ = 2406.26, $[C_{125}H_{177}N_{16}O_{29}Na_2]^+$ = 2412.27

Formula: $C_{93}H_{114}N_{16}O_{29}$

Molecular Mass: 1918.79 g/mol

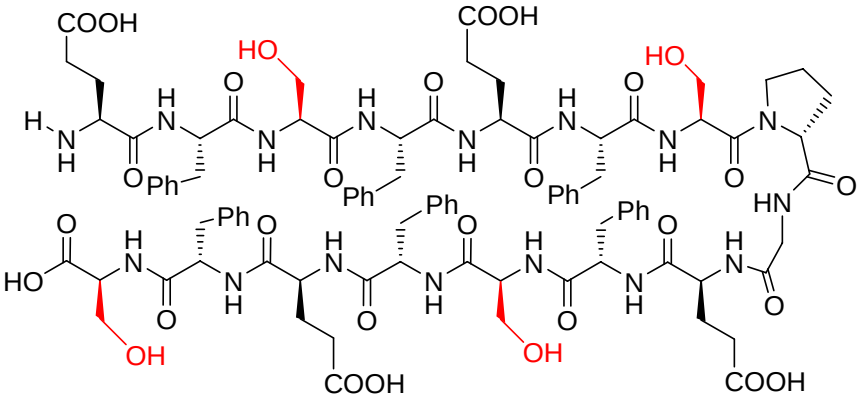
MS (MALDI-ToF) (**HP 20-1**): $m \cdot z^{-1}$ = 1941.36 $[M+Na]^+$, 1957.61 $[M+K]^+$, 1963.54 $[M+2Na-H]^+$, 1979.66 $[M+2Na-H+H_2O]^+$, 1985.58 $[M+3Na-2H]^+$

Calculated Mass (**HP 20-1**): $[C_{93}H_{114}N_{16}O_{29}Na]^+$ = 1941.78, $[C_{93}H_{114}N_{16}O_{29}NK]^+$ = 1957.77, $[C_{93}H_{113}N_{16}O_{29}Na_2]^+$ = 1963.77, $[C_{93}H_{115}N_{16}O_{30}Na_2]^+$ = 1981.78, $[C_{93}H_{112}N_{16}NO_{29}Na_3]^+$ = 1985.76

Preparative RP-HPLC: Apparatus A, Method 20, t_R = 19.66 min

Analytical RP-HPLC: Method II, t_R = 21.29 min, 100% (220 nm)

H-Glu-Phe-Ser-Phe-Glu-Phe-Ser-D-Pro-Gly-Glu-Phe-Ser-Phe-Glu-Phe-Ser-OH (HP 20-2) (Microwave Peptide Synthesizer)



HP 20-2

The comparison of peptide synthesis between manual and microwave methodology was carried out, using peptide **HP 20-1** and **HP 20-2** as the reference peptides for comparison.

The synthesis of the fully protected peptide **Prot. HP 20-2** was carried out according to the general procedure of SPPS (Fmoc/tBu), operating automatically on the CEM Microwave Peptide Synthesizer, with Fmoc-Ser(tBu)-OH being loaded on the *o*-chlorotrityl resin (0.8 mmol/g, 0.25 mmol), and TBTU as coupling reagent. The N^{α}-Fmoc-deprotecting cocktail 20% piperidine/DMF did not contain HOBt as additive. Deprotection of peptide **Prot. HP 20-2** was achieved with the deprotection cocktail TFA/H_2O/TIS (95:2.5:2.5). The deprotection reaction at ambient temperature took 3 hours before the quantitative cleavage of the side-chain protecting groups was achieved.

Yield:
H-Glu(OtBu)-Phe-Ser(tBu)-Phe-Glu(OtBu)-Phe-Ser(tBu)-D-Pro-Gly-Glu(OtBu)-Phe-Ser(tBu)-Phe-Glu(OtBu)-Phe-Ser(tBu)-OH **(Prot. HP 20-2)**
0.5796 g/0.245 mmol (raw product, after lyophilization)
H-Glu-Phe-Ser-Phe-Glu-Phe-Ser-D-Pro-Gly-Glu-Phe-Ser-Phe-Glu-Phe-Ser-OH
(**HP 20-2**) 0.0165 g/0.0086 mmol (3.4%)

Formula of **Prot. HP 20-2**: $C_{125}H_{178}N_{16}O_{29}$

Molecular Mass (**Prot. HP 20-2**): 2367.29 g/mol

MS (MALDI-ToF) (**Prot. HP 20-2**): $m{\cdot}z^{-1}$ = 2389.25 $[M+Na]^+$, 2406.22 $[M+K]^+$, 2410.93 $[M+2Na-H]^+$
Calculated Mass (**P_HP 3c-2**): $[C_{125}H_{178}N_{16}O_{29}Na]^+$ = 2390.28, $[C_{125}H_{178}N_{16}O_{29}K]^+$ = 2406.26, $[C_{125}H_{177}N_{16}O_{29}Na_2]^+$ = 2412.27

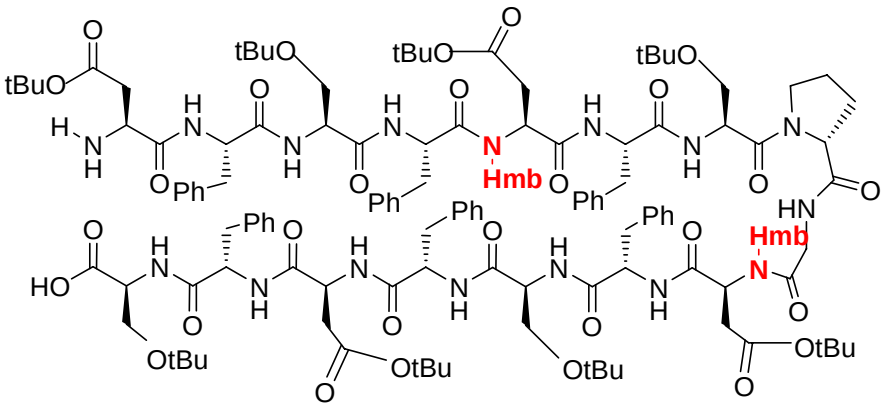
Formula (**HP 20-2**): $C_{93}H_{114}N_{16}O_{29}$

Molecular Mass (**HP 20-2**): 1918.79 g/mol

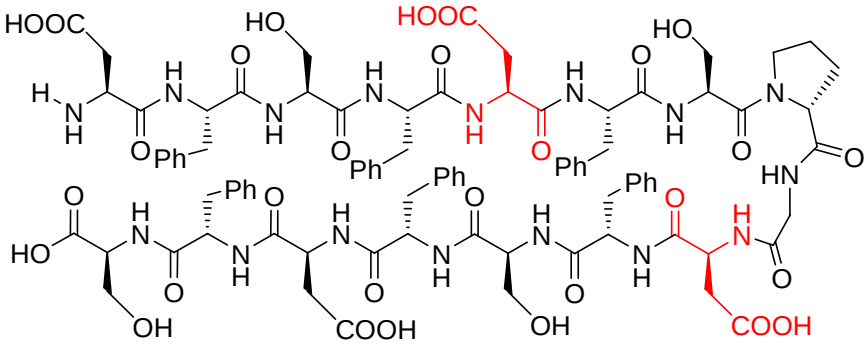
MS (MALDI-ToF) (**HP 20-2**): $m{\cdot}z^{-1}$ = 1941.56 $[M+Na]^+$, 1957.81 $[M+K]^+$, 1963.45 $[M+2Na-H]^+$

Calculated Mass (**HP 20-2**): $[C_{93}H_{114}N_{16}O_{29}Na]^+$ = 1941.78, $[C_{93}H_{114}N_{16}O_{29}NK]^+$ = 1957.77, $[C_{93}H_{113}N_{16}O_{29}Na_2]^+$ = 1963.77

H-^{1}Asp-^{2}Phe-^{3}Ser-^{4}Phe-^{5}Asp-^{6}Phe-^{7}Ser-^{8}D-Pro-^{9}Gly-^{10}Asp-^{11}Phe-^{12}Ser-^{13}Phe-^{14}Asp-^{15}Phe-^{16}Ser-OH (HP 17-Hmb)



Prot.HP 17-Hmb



HP 17-Hmb

The synthesis of the fully protected derivative of peptide **HP 17-Hmb**, namely, H-^{1}Asp(OtBu)-^{2}Phe-^{3}Ser(tBu)-^{4}Phe-*N*-Hmb-^{5}Asp(OtBu)-^{6}Phe-^{7}Ser(tBu)-^{8}D-Pro-^{9}Gly-*N*-Hmb-^{10}Asp(OtBu)-^{11}Phe-^{12}Ser(tBu)-^{13}Phe-^{14}Asp(OtBu)-^{15}Phe-^{16}Ser(tBu)-OH (**Prot. HP17-Hmb)**, was carried out according to the general procedure of manual SPPS (Fmoc/tBu), with Fmoc-Ser(OtBu)-OH being loaded on the *o*-chlorotrityl resin (0.6 mmol/g, 0.1 mmol). The residues ^{5}Asp and ^{10}Asp were incorporated into the peptide chain in the form of *N,O*-bisFmoc-*N*-Hmb-Asp(OtBu)-OH **22** (3 equiv), using 3 equiv HATU as the coupling reagent and 6 equiv DIPEA as the base. The duration of coupling

^{5}Asp and ^{10}Asp derivatives to the peptide chain on the solid support took 3 hours at room temperature. The coupling of ^{9}Gly and ^{4}Phe residue derivative to the peptide chain was also achieved through the coupling reagent HATU. For the coupling of other residues to the peptide chain, TBTU was applied as the coupling reagent. From the coupling of ^{7}Ser on, the amounts of the amino acid derivatives were increased to 4 equiv. In a similar way, the amount of TBTU and DIPEA were increased to 4 equiv and 8 equiv, respectively (with the exception of ^{5}Asp derivative with 3 equiv amino acid derivative, 3 equiv HATU, 6 equiv DIPEA). Cleavage of the peptide **Prot. HP17-Hmb** from the resin was achieved by the treatment of the peptidyl resin with 1% TFA in DCM. The solvents were removed under reduced pressure and the peptide solids were dissolved in acetonitrile and lyophilized. The hereby obtained fully protected peptide **Prot. HP17-Hmb** was dissolved in acetonitrile with the addition of minimal DMF to enhance the solubility and subsequently purified with RP-HPLC. The puried solid peptide of **Prot. HP17-Hmb** was then given to the reagent K for the cleavage of the side-chain tBu-type protecting groups as well as backbone Hmb protecting groups. The deprotection reaction took approximately 5 hours at room temperature. The solvent was then removed under reduced pressure. The obtained solid residue was dissolved in minimal TFA and slowly added to an ice-cold diethyl ether resided in ice-water bath. The target peptides precipitated upon the addition of the solution into the cold ether. This suspension was stirred for approximately 30 minutes at 0 ºC before centrifugation. The solid peptide was again added to another portion of fresh ice-cold diethyl ether and stirred for 30 minutes in the ice-water bath. Likewise, the ether solution was removed by centrifugation. The obtained peptide solid was triturated in acetonitrile/H_2O and dried by lyophilization.

Yield: **Prot. HP17-Hmb** 0.1293 g/0.05 mmol (crude products after lyophilization)

HP 17-Hmb 0.0372 g/0.02 mmol (20.0%)

Formula (**Prot. HP17-Hmb**): $C_{137}H_{186}N_{16}O_{33}$

Molecular Mass (**Prot. HP17-Hmb**): 2583.33 g/mol

MS (MALDI-ToF) (**Prot. HP17-Hmb**): $m{\cdot}z^{-1}$ = 2606.08 $[M+Na]^+$, 2622.47 $[M+K]^+$, 2627.13 $[M+2Na-H]^+$, 2644.54 $[M+Na+K-H]^+$

Calculated Mass (**Prot. HP17-Hmb**): $[C_{137}H_{186}N_{16}O_{33}Na]^+$ = 2606.32, $[C_{137}H_{186}N_{16}O_{33}K]^+$ = 2622.29, $[C_{137}H_{185}N_{16}O_{33}Na_2]^+$ = 2628.31, $[C_{137}H_{185}N_{16}O_{33}NaK]^+$ = 2644.28

Preparative RP-HPLC (**Prot. HP17-Hmb**): Apparatus A, Method 17, t_R = 24.49 min

Analytical RP-HPLC: Method II, t_R = 39.45 min, 100% (220 nm)

Formula (**HP 17-Hmb**): $C_{89}H_{106}N_{16}O_{29}$

Molecular Mass (**HP 17-Hmb**): 1863.94 g/mol

MS (MALDI-ToF) (**HP 17-Hmb**): $m \cdot z^{-1}$ = 1885.72 $[M+Na]^+$, 1903.11 $[M+K]^+$
Calculated Mass: $[C_{89}H_{106}N_{16}O_{29}Na]^+$ = 1885.73, $[C_{89}H_{106}N_{16}O_{29}K]^+$ = 1903.03,

3.7 Summary

In summary, a series of peptides designed as models for the research of β-hairpin conformation and templates for biomineralization have been synthesized by microwave SPPS or traditional manual SPPS. Moreover, DFSFDFSpGDFSFDFS **HP 17** was successfully synthesized through the Hmb backbone protection strategy. The introduction of the backbone protection groups Hmb eliminated the intra-chain hydrogen bonds between the neighbouring β-strands in the reference peptide and hence drastically increased the solubility of the fully protected peptides **Prot. HP 17-Hmb**, which could be easily dissolved in DMF/ACN mixture solvent and consequently purified by RP-HPLC. The function of Hmb to break the inter/intra-chain hydrogen bond had also been validated by the CD spectra of **Prot. HP 17-Hmb**. In conclusion, the backbone protection methodology successfully avoided the inherent problems residing in the SPPS of β-sheet related peptides. The intra/inter-chain hydrogen bonds between the neighbouring β-strands greatly impeded the flexibility and accessibility of the growing peptide chains, leading to insufficiently coupling and truncated sequences, and not being appropriate for HPLC purification. The Hmb backbone protection methodology could successfully suppress such problems.

Through the conformation studies of the synthesized peptide, it could be deduced that aspartic acids are less appropriate as components of amphiphilic β-hairpin peptides than glutamic acids. This has been proven by the fact that aspartic acids appear less frequently than glutamic acids in the database of naturally occurred β-sheet proteins and peptides. By comparison of the conformation of homologous peptides DFDFDFDpGDFDFDFD **HP 18** and EFEFEFEpGEFEFEFE **HP 19**, it was found that Asp-containing peptide **HP 18** adopted bascially random coil in HFIP, while under the same conditions, Glu-containing peptide **HP 19** adopted mainly ordered β-hairpin conformation. The distinction between the conformations of **HP 18** and **HP 19** might be based on the fact that there is one more methylene group in glutamic acid than in aspartic acid, which

makes the side chain of glutamic acid more flexible and thus being able to adopt the favourable conformation when interacting with the counterpart glutamic acids on the neighbouring β-strands. This extra flexibility enables the glutamic acids to dodge the disadvantage repulsive ionic interaction between the β-strands in the β-hairpin peptide, obviating the factors which could probably jeopardize the β-hairpin conformation. Another important factor which makes peptide **HP 19** adopt β-hairpin conformation while **HP 18** random coil might originate from the extra methylene group in glutamic acid. It introduces the additional beneficial hydrophobic interactions between the side chains of glutamic acids residing on the neighbouring β-strands in the β-hairpin peptide **HP 19**, which could perhaps hugely reinforce its stability. This result exhibited the importance of hydrophobic interactions in stabilizing the β-hairpin conformation. It could thus be imagined that if one more methylene group is inserted in the side chain of glutamic acid, that is to say, to substitute glutamic acid rediues by 2-amino adipic acids, the stability of β-hairpin conformation of such a peptide could possibly be further reinforced relative to **HP 19**. The distinction between aspartic acid and glutamic acid with respect to β-sheet stabilization has been validated by the comparison of the conformations of peptide FDFDFDFDFD **HP 11** and FEFEFEFE **HP 12,** in which turn motif -D-Pro-Gly- has been excluded and only β-strand sequences has been maintained. The conformational distinctions between peptide **HP 11** and **HP 12** clearly indicate the inherent differences between aspartic acid and glutamic acid in their inclination to stabilize the β-strand conformation.

The β-turn structure is formed by the -D-Pro-Gly- motif, a β-turn inducer. From the CD spectra of EpGE **HP 1** and DpGD **HP 2**, it could be deduced that they adopt basically β-turn conformations in H_2O. The -D-Por-Gly- motif alternated the direction of peptide and led to the folding of the chain as they are supposed to. The incorporation of β-turn motif is a crucial factor to nucleate the β-hairpin conformation in designed peptide.

As is shown by *Figure 3.21-23*, **HP 1** and **HP 2** adopted β-turn conformations in H_2O, but the ordered structures were converted to random coil in HFIP. While FpGF **HP 4**, in which phenylalanines occupy the corner positions, adopt mainly β-turn conformation both in H_2O and HFIP. This distinction could be based on the inherent differences in the

properties of H_2O and HFIP and the pattern of hydrogen bond network in the peptide. In contrast to water, halogenated alcohols such as HFIP or TFE are better proton donors and poorer proton acceptors. It was therefore proposed that HIFP/TFE is likely to bind preferentially to the mainchain carbonyl oxygen group. Such a preference leads to enhanced intra-hydrogen bond of amide groups as solvent exposure of amide is minimized. (Carbonyl oxygen is often able to accept two hydrogen bonds.) Indirect effects on H-bond result from altered dielectric constants. As the solvent dielectric constant is lowered from 79 of water to 27 (pure TFE at 25ºC), electrostatic interactions become stronger. TFE could be regarded as mainly weakening hydrophobic-hydrophobic interactions and only slightly enhancing local helical interactions. The ability of HFIP/TFE to disfavor hydrophobic interactions between residues distant in the amino acid sequence, which is found in the interior of native state protein or across strands in some β-sheets, accounts for their abilities to disrupt these structures. Structures that rely on sequence local interactions, such as helices, are preferentially stabilized. In conclusion, the increase of the strength of intra-chain hydrogen bond in peptide **HP 1** and **HP 2** caused by HFIP is not strong enough to compensate the enhancement of repulsive electrostatic interactions, which is also generated by the introduction of HFIP, between the side chains of Glu or Asp. The conformational dependence of solvent is thus a balance between the advantageous factors such as hydrophobic interactions/hydrogen bonds and disadvantageous factors such as repulsive electrostatic interactions. In addition, the pattern of hydrogen bond network could also exert influence on the stability of β-turn/hairpin conformation and their dependence of solvent, which is confirmed by the fact that **HP 4**, in which phenylalanines occupy the corner positions of the turn motif and thus possibly alternate the pattern of hydrogen bond network, adopt β-turn conformation both in H_2O and HFIP. The role of hydrogen bond network pattern, in another word the occupation of corner positions in β turn motif, in stabilizing β-hairpin conformation is further testified by the conformational differences between FDpGDF **HP 6** and DFpGFD **HP 7**, in which the positions of aspartic acids and phenylalanines are exchanged, the stability of β-hairpin conformation is superior in **HP 7** (phenylalanines occupy the corner position) to that in **HP 6** (aspartic acids occupy the corner position).

In summary, the stability of β-hairpin conformation is cooperatively determined by the primary sequence of correlative peptides, the choice of β-turn motif, the occupation of the corner residues in β-turn, the pattern of hydrogen bond network, type of the solvent, environmental pH value, and the length of the peptide. Intra-chain hydrogen bonds, electrostatic interactions and hydrophobic interactions could be regarded as the basic strength in deciding the stability of β hairpin conformation, their influence on the stability of β-hairpin conformation could be sophisticated and should be analyzed by individual case.

The comparison of manual SPPS and microwave SPPS has also been carried out in this project. It could be deduced that microwave methodology is superior to traditional manual SPPS when certain difficult sequences are involved. Microwave is assumably capable of destroying the intra/inter-hydrogen bonds during the chain elongation on the solid support, therefore the introduction of microwave irradiation during the synthesis of β-sheet or β-hairpin peptides could be beneficial, validated by the increase of the yield of target peptides and decrease of the truncated sequences. However, the effects of microwave in SPPS should be regarded as ambiguous. This kind of energy input seemed to be prone to promote the extent of some side-reactions in SPPS such as aspartimide formation. The introduction of weak acid like HOBt into Fmoc deprotecting reagent 20 % piperidine in DMF could reduce the aspartimide side products to some extent, but not able to suppress it completely. While under the same synthetic conditions, no detectable aspartimide derivatives are found in the traditional manual methodology.

References

[1] Simkiss, K., Wilbur, K. M., in: *Biomineralization*. *Cell Biology and Mineral Deposition*, San Diego: Academic Press, (1989), 230–260.

[2] Volkmer, D., in: *Encyclopedia of Separation Science* (M. Cooke, C.F. Poole, eds.), Vol. 2 (Crystallization), Academic Press, (2000), 940–950.

[3] (a) Belcher, A. M., Wu, X. H., Christensen, R. J., Hansma, P. K., Stucky, G. D., Morse, D. E., *Nature*, (1996) **381**, 56–58.
(b) Falini, G., Albeck, S., Weiner, S., Addadi, L., *Science*, (1996) **271**, 67–69.

[4] Samata, T., Hayashi, N., Kono, M., Hasegawa, K., Horita, C., Akera, S., *FEBS Letters*, (1999) **462**, 225–229.

[5] Sudo, S., Fujikawa, T., Nagakura, T., Ohkubo, T., Sakaguchi, K., Tanaka, M., Nakashima, K., Takahashi, T., *Nature* (1997), **387**, 563–564.

[6] Zhang, Y., Xie, L. P., Meng, Q. X., Jiang, T. M., Pu, R. L., Chen, L., Zhang, R. Q., *Comp. Biochem. Phys.*, B (2003) **135**, 565–573.

[7] Kono, M., Hayashi, N., Samata, T., *Biochem. Biophys. Res. Comm*., (2000) **269**, 213–218.

[8] Schen, X., Belcher, A. M., Hansma, P. K., Stucky, G. D., Morse, D. E., *J. Biol. Chem.*, (1997) **272**, 32472–32481.

[9] Suzuki, M., Murayama, E., Inoue, H., Ozaki, N., Tohse, H., Kogure, T., Nagasawa, H., *J. Biochem.*, (2004) **382**, 205–213.

[10] Sarashina, I., Endo, K., *Am. Mineral*, (1998) **83**, 1510–1515.

[11] Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T., Matsushiro, A., *Proc. Natl. Acad. Sci. USA*, (1996) **93**, 9657–9660.

[12] Marin, F., Corstjens, P., B. de Gaulejac, E. de Vrind-De Jong, Westbroek, P., *J. Biol. Chem*., (2000) **275**, 20667–20675.

[13] Mann, K., Weiss, I. M., André, S., Gabius, H.-J., Fritz, M., *Eur. J. Biochem*., (2000) **267**, 5257–5264.

[14] Weiss, I. M.. Göhring, W., Fritz, M., Mann, K., *Biochem. Biophys. Res. Commun.*, (2001) **285**, 244–249.

[15] Tsukamoto, D., Sarashina, I., Endo, K., *Biochem. Biophys. Res. Commun.*, (2004) **320**, 1175–1180.

[16] (a) Michenfelder, M., Fu, G., Lawrence, C., Weaver, J. C., Wustman, B. A., Taranto,L., Evans, J. S., Morse, D. E., *Biopolymers*, (2003) **70**, 522–533.
(b) Wustman, B. A., Morse, D. E., Evans, J. S., *Biopolymers*, (2004) **74**, 363–376.

[17] L. Addadi, S. Weiner in: *Biomineralization*, Mann S, Webb J, Williams RJP (eds.), Weinheim: VCH, **1989**, 133–156.

[18] Levi-Kalisman, Y., Falini, G., Addadi, L., Weiner, S., *J. Struct. Biol.,* (2001) **135**, 8-17.

[19] chitin is a water-insoluble (1,4)-linked 2-acetamido-2-deoxy-D-glucan.

[20] Levi-Kalisman, Y., Falini, G., Addadi, L., Weiner, S., *J. Struct. Biol*., (2001) **135**, 8–17.

[21] Bertrand, M., Brack, A., *Origin Life Evol*. B (1997) **27**, 589–598.

[22] Volkmer, D., Fricke, M., Huber, T., Sewald, N., *Chem. Commun.*, (2004) 1872–1873.

[23] (*a*) Belcher, A. M., Wu, X. H., Christensen, R. J., Hansma, P. K., Stucky, G. D., Morse, D. E., *Nature*, (1996) **381**, 56–58.
(*b*) Falini, G., Albeck, S., Weiner, S., Addadi, L., *Science*, (1996) **271**, 67–69.

[24] (*a*) DeOliveira, D. B., Laursen, R. A., *J. Am. Chem. Soc.*, (1997) **119**, 10627–10631;
(*b*) Li, C., Botsaris, G. D., Kaplan, D. L., *Cryst. Growth Des.*, (2002) **5**, 387–393.
(*c*) Kim, I. W., DiMasi, E., Evans, J. S., *Cryst. Growth Des.*, (2004) **4**, 1113–1118.

[25] (a) Young, J. R., Didymus, J. M., Brown, P. R., Prins, B., Mann, S., *Nature*, (1992) **356**, 516–518.
(b) Aizenberg, J., Hanson, J., Koetzle, T. F., Weiner, S., Addadi, L., *J. Am. Chem. Soc.*, (1997) **119**, 881–886;
(c) Mann, S., Sparks, N. H. C., *Proc. R. Soc.*, B (1998) **234**, 441–453.

[26] Volkmer, D., Fricke, M.; Huber, T., Sewald, N., *Chem. Commun.*, (2004) 1872–1873.

[27] Voet, D., Voet, J. G.., (2004). *Biochemistry*, Vol 1 3rd ed. Wiley. See esp. 227–231.

[28] Geddes, A. J., Parker, K. D., Atkins, E. D., Beighton, *E. J Mol Biol*. (1968) **32**, 343-358.

[29] Venkatachalam, C. M., *Biopolymers*, (1968) **6**, 1425-1436.

[30] Sibanda, B. L., Thornton, J. M., *Methods Enzymol.*, (1991) **202**, 59-82

[31] Sibanda, B. L., Thornton, J. M., *Nature* (1985) **316**, 170-174.

[32] Toniolo, C., *CRC Crit Rev Biochem*. (1980) **9**, 1-44.

[33] Toniolo, C., Benedetti, E., *Trends Biochem Sci*., (1991) **16**, 350-353.

[34] Rose, G. D., Gierasch, L. M., Smith, J. A., *Adv Protein Chem*., (1985) **37**, 1-109.

[35] Venkatachalam, C. M., *Biopolymers*, (1968) **6**, 1425.

[36] Wilmot, C. M., J. M. Thornton, *J. Mol. Biol.*, (1988) **203**, 221.

[37] Chou. P. Y., Fasman, G. D., *J. Mol. Biol.*, (1997) **115**, 135.

[38] Sibanda, B. L., Thornton, J. M., *Nature* (*London*), (1985) **316**, 170.

[39] Sibanda, B. L., Blundell, T. L., Thornton, J. M., *J. Mol. Biol.*, (1989) **206**, 759.

[40] J. S. Richardson, *Adv. Protein Chem.*, 1981, **34**, 167.

[41] Srinivasan, N., Anuradha, V. S., Ramakrishnan, C., Sowdhamini. R., Balaram, P., *Int. J. Pept. Protein Res.*, (1994) **44**, 112.

[42] Raghothama, S. R., Awasthi, S. K., Balaram, P., *J. Chem. Soc., Perkin Trans. 2,* (1998) 137.

[43] Gunasekaran, K., Ramakrishnan, C., Balaram, P., *Protein Eng.*, (1997) **10**,1131-1141.

[44] Struthers, M. D., Cheng, R. P., Imperiali. B., *Science* (1996) **271,** 342-345.

[45] Haque, T. S., Gellman, S. H., *J. Am. Chem. Soc.*, (1997) **119**, 2303-2304.

[46] Stanger, H. E., Gellman, S. H., *J. Am. Chem. Soc.*, **(**1998) **120**, 4236-4237.

[47] de Alba, E., Jimenez, M. A., Rico, M., *J. Am. Chem. Soc.*, (1997) **119**, 175-l 83.

[48] Griffiths-Jones, S. R., Maynard, A. J., Sharman, G. J., Searle, M. S., *J. Chem. Soc. Chem. Commun.*, (1998) **9**, 789-790.

[49] Ramirez-Alvarado, M., Blanc, F. J., Serrano, L., *Nat. Struct. Biol.*, (l996) **3**, 604-612.

[50] de Alba, E., Rico, M., Jimenez, M. A., *Protein Sci.*, (1997) **6**, 2548-2560.

[51] Wouters, M. A.; Curmi, P. M.G., *Protein Struct. Funct. Genet.*, (1995) **22**, 119-l 31.

[52] Maynard, A. J., Sharman, G. J., Searle, M. S., *J. Am. Chem. Soc.*, (1998) **120**, 1996-2007.

[53] Catherine, K. S., Regan, L., *Acc. Chem. Res.,* (1997) **30**, 153-161.

[54] Kent, S. B. H., In *Peptides, Structure and Function,* Deber, C., M.; Hruby, V, J.; Kopple, K. D., Eds,; Pierce Chemical Company: Rockford, IL, (1985); 407.

[55] Stewart, J. M., Klis, W. A., In *Innovations and Perspectives in Solid Phase Synthesis*, Epson, R., Ed.; SPCC (UK) Ltd.: Birmingham, (1990); p 1.

[56] Thaler, A., Seebach, D., Beck, A. K., *Helv, Chim. Acta*, (1991) **74**, 617.

[57] Klis, W. A., Stewart, J, M., In *Peptides: Chemistry, Structure and Biology*, River, J. E.; Marshall, G.. R., Eds.; ESCOM: Leiden, (1990); p 904.

[58] Seebach, D., Thaler, A., Beck, A. K., *Helv. Chim. Acta*, (1989) **72**, 857.

[59] Hyde, C., Johnson, T., Owen, D., Quibell. M.; Sheppard, R. C., *Int. J. Pept. Protein Res.,* (1994) **43**, 431.

[60] Rapp, W., Bayer, E., In *Peptides 1992*, Schneider, C. H,; Eberle, A. N., Eds,; ESCOM: Leiden, (1993); p 25.

[61] Zhang, L., Goldammer, C., Henkel, B., Zühl, F., Panhaus, G., Jung, G., Bayer, E., In *Innovation and Perspectives in Solid Phase Synthesis*, Epton, R., Ed.; Mayflower Worldwide Ltd: Birmingham, UK, (1994); p 711.

[62] Lloyd, D. H., Petrie, G. M., Noble, R. L., Tam, J. P., In *Peptides: Chemistry, Structure and Biology*, River, J. E.; Marshall, G.. R., Eds.; ESCOM: Leiden, (1990); p 909.

[63] Mutter, M., Oppliger, H., Zier, A., *Makromol. Chem., Rapid Commun.*, (1992) **13**, 151.

[64] Moroder, L., Gemeiner, M., Göhring, W., Jaeger, E., Thamm, P., Wünsch, E., *Biopolymer*, (1981) **20**, 17.

[65] Johnson, T., Quibell, M., Owen, D., Sheppard, R. C., *Chem. Commun*., (1993), 369

[66] Nicolás, E., Pujades, M., Bacardit, J., Giralt, E., Albericio, F., *Tetrahedron Lett*., (1997) **38**, p 2317.

[67] Offer, J., Johnson, T., Quibell, M., *Tetrahedron Lett*., (1997) **38**, 9047.

[68] Johnson, T., Quibell, M., Sheppard, R. C., *J. Pept. Sci.*, (1995) **1**, 11

[69] Quibell, M., Turnell, W. G., Jonhnson, T., *J. Chem. Soc., Perkin Trans.*, (1995) **1**, 2019.

[70] Quibell, M., Turnell, W. G., Jonhnson, T., *J. Org. Chem.,* (1994) **59**, 1745.

[71] Woody, R. W. in *Circular Dichroism, Principles and Applications,* Nakanishi, K., Berova, N., Woody, R. W., (Eds.), VCH, New York, (1994)

[72] Crisma, M., Formaggio, F., Moretto, A., Toniolo, C., *Biopolymers,* (2006) **84**, 2-12.

[73] Nelson, J. W., Kallenbach, N. R., *Proteins Struct. Funct. Genet.,* (1986) **1**, 211

[74] Sewald, N., Jakubke, H-D. in *Peptides: Chemistry and Biology*, Wiley-VCH, Weinheim, (2002).

[75] Buck, M., *Q. Rev. Biophys.,* (1998) **31**, 297-355.

[76] Mizuno. K., Kaido, H., Kimura, K., Miyamoto, K., Yoneda, N., Kawabata, T., Tsurusaki, T., Hahizume, N., Shindo, Y., *J. Chem. Soc., Faraday Trans.,* I. (1984) **80**, 879-984.

[77] Linas, M., Klein, M. P., *J. Am. Chem. Soc.* (1975) **97**, 4731-4737.

[78] Lehrmann, M. S., Mason, S. A., McIntyre, G. J., *Biochemistry,* (1995) **24**, 5862-5869.

[79] Yang, Y., Barker, S., Chen, M. J., *J. Biol. Chem.,* (1993) **268**, 9223-9229.

[80] Thomas, P. D., Dill, K. A., *Protein Science,* (1993) **2**, 2050-2056.

[81] Kovrigin, E. L., Potekhin, S. A., *Biochemistry*, (1997) **36**, 9195-9199.

[82] Westh, P., Koga, Y., *J. Phys. Chem. B,* (1997) **101**, 5755-5758.

[83] Zhou, N. E., Kay, C. M., Sykes, B. D., Hodges, R. S., *Biochemistry,* (1993) **32**, 6190-6197.

[84] Soler-Gonzalez, A, S., Fersht, A. R., *Eur. J. Biochem.,* (1997) **249**, 24-732.

[85] Sukumar, M., Gierasch, L. M., *Folding & Design,* (1997) **2**, 211-222.

[86] Shin, H. C., Merutka,G., Waltho, J. P., Wright, P. E., Dyson, H. J., *Biochemistry,* (1993 a) **32**, 6348-6355.

[87] Shin, H. C., Merutka,G., Waltho, J. P., Tennant, L. L., Dyson, H, J., Wright, P. E., *Biochemistry*, (1993 b) **32**, 6356-6364.

[88] Wray, V., Federau, T., Gronwald, W., Mayer, H., Schomburg, D., Tegge, W., Wingender, E., *Biochemistry*, (1994) **33**, 1684-1693.

[89] Cantor, C.R. and Schimmel, P.R., *Biophysical Chemistry*, Part I, W.H. Freeman, San Francisco, (1980) 49.