Universidade do Minho
Escola de Ciências

Tarsila Gabriel Castro

**Conformational Properties of Unnatural
Amino Acids in Peptidomimetics/
Foldamers: A Molecular Modelling Study**

novembro de 2015

Universidade do Minho

Escola de Ciências

Tarsila Gabriel Castro

Conformational Properties of Unnatural Amino Acids in Peptidomimetics/ Foldamers: A Molecular Modelling Study

Tese de Doutoramento em Ciências
Especialidade em Química

Trabalho realizado sob a orientação do
Doutor Manuel Melle-Franco
Doutor Nuno Miguel da Silva Micaêlo

E co-orientação do
Doutor João Carlos Ramos Nunes Marcos

novembro de 2015

STATEMENT OF INTEGRITY

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, _13 de Novembro de 2015_

Full name:

TARSILA GABRIEL CASTRO

Signature:

Tarsila Gabriel Castro

To my parents,

husband and children

"A ciência humana de maneira nenhuma nega a existência de Deus. Quando considero quantas e quão maravilhosas coisas o homem compreende, pesquisa e consegue realizar, então reconheço claramente que o espírito humano é obra de Deus, e a mais notável."

Galileu Galilei

Agradecimentos

Como uma pessoa de fé, em primeiro lugar agradeço a Deus e à intercessão de Nossa Senhora, por este ciclo tão importante na minha jornada profissional. Foram quatro anos de importante aprendizado, percorrendo distintas fases emocionas: entusiasmo, desânimo, frustação, nova carga de entusiasmo, vontade de aprender e crescer, e assim por diante. As típicas sensações e vivencias de um estudante de doutoramento. Além disso, algumas experiências pessoais tornaram a caminhada mais difícil, mas a presença constante de Deus na minha vida me levou adiante, concedendo-me a coragem e perseverança necessárias para chegar a este momento.

Um doutoramento só é possível graças a contribuição e orientação daqueles que formam e moldam o aluno. Por isso, agradeço de forma especial ao meu orientador, o Doutor Manuel Melle-Franco que me recebeu generosamente como aluna ao fim do terceiro ano de doutoramento. Com o Doutor Manuel saí da minha "zona de conforto" e vi que isso é bom e que me fez crescer. Aprendi a ser mais crítica e a ter novo olhar sobre o meu trabalho. O Doutor Manuel também foi o responsável por me inserir novamente na Química Quântica e na Química de Materiais, áreas de atuação adormecidas desde os tempos de Licenciatura. Aprendi imenso e por isso sou muito grata!

Agradeço também ao meu co-orientador, o Doutor João Carlos Marcos, que enquanto bioquímico tem um entusiasmo motivador em relação à Química Teórica. O Doutor João Carlos também foi responsável por voltar o meu interesse para o estudo do DNA, para métodos de docking e por me tornar cada vez mais interessada pela bioquímica. Acima de tudo, agradeço-o por acreditar no meu potencial e me considerar pronta a orientar outros, concedendo-me a minha primeira experiência na orientação.

Durante os três primeiros anos de Doutoramento, fui guiada pelo Doutor Nuno Micaêlo, que foi também o meu orientador na conclusão do Mestrado em Química Medicinal, na Universidade do Minho. Idealizamos o projeto de doutoramento juntos, partilhando grande entusiasmo. Com ele aprendi muitíssimo e, por isso, agradeço mais uma vez pelos desafios propostos e conhecimentos transmitidos, que me moldaram como cientista.

Passo a agradecer a todos que compõe o Centro de Química da Universidade do Minho, colegas, amigos e professores. Agradeço especialmente à colega Helena Vilaça, à Doutora Paula Margarida e ao Doutor José Alberto Martins, co-autores em trabalhos que reúnem de forma sábia a química

x

Abstract

Conformational Properties of Unnatural Amino Acids in Peptidomimetics/Foldamers: A Molecular Modelling Study

Non-canonical (unnatural) amino acids are molecules that enhance specific secondary structures and/or biological activity of peptides. Contrary to the well-known encoded (natural) amino acids, the structure and function of these residues is far for being fully understood. Nowadays, non-canonical amino acids are used to generate peptidomimetics with improved biostability and bioavailability. In this sense, we performed simulation studies of non-canonical amino acids able to induce constrained secondary structures in order to optimize peptides biological function.

Molecular Dynamics simulations were performed systematically to validate and incorporate new classes of unnatural amino acids in novel and experimentally found peptides with desirable biological function. To do this, the Gromos 54a7 Force Field was augmented with a new set of parameters based on canonical, proteinogenic amino acids, needed to model the new residues.

We study several classes of non-canonical amino acids, namely: symmetrical α,α-dialkyl glycines, asymmetrical α,α-dialkyl glycines, proline analogues, $C\alpha$ to $C\alpha$ cyclized amino acids and α,β-dehydroamino acids. These classes were chosen because very few amino acids of each class have been studied in detail. In addition, these amino acids are important examples of residues with good conformation inducer properties and/or medical applicability.

The symmetrical and asymmetrical α,α-dialkyl glycines were studied in four well-known antibiotic peptaibols. Dhg (α,α-dihexyl glycine) and Ac_6c (1-aminocyclohexane-1carboxylic acid or α,α-cyclohexyl glycine), symmetrical glycines, proved to be helical inducers in Alamethicin and Peptaibolin peptides. Also, these two examples promoted pre-organization in water, which was found to help insertion in membranes. On the other hand, the asymmetrical α,α-dialkyl glycines, like Iva (isovaline), were studied in Antiamoebin and Zervamicin peptaibols. In these studies, two amino acids analogs of Iva were found to induce improved helical secondary structure, namely α-methyl-D-leucine (MDL) and α-methyl-D-phenylalanine (MDP), which may be linked to the antibiotic properties of these peptaibols. In addition, proline analogs were also studied in Antiamoebin and Zervamicin peptaibols, which naturally contain Hyp (Hydroxyproline). Despite the known effect of prolines, which induce bends in helical secondary structures, the analog cis-3-amino-L-proline (ALP) proved to induce improved helical content in both peptaibols.

Simulation of α,β-dehydroamino acids revealed a wide range of applicability for these systems, from self-assembly peptides for drug delivery to induction of different and specific secondary structures such as β and γ turns.

Summing up, our simulation studies reveal that the incorporation of non-canonical amino acids in peptides is able to generate a large range of peptidomimetics with different structures and potential applications. Our findings show that the rational selection of unnatural residues increases membrane permeability through pre-organization in aqueous medium, stabilizes the content of desired types of secondary structure and, more generally, improves enzymatic and thermodynamic stability. This work showcases how molecular modeling can be applied to address a number of issues of interest for medicinal chemistry.

Resumo

Os aminoácidos não-canónicos (ou não-naturais) são moléculas capazes de otimizar a estrutura secundária e/ou a atividade biológica de péptidos. A estrutura e função destas moléculas ainda não é bem conhecida, por oposição aos já bem estudados aminoácidos codificados por ADN (ou canónicos). Atualmente, os aminoácidos não-canónicos são utilizados para gerar péptidos miméticos com uma melhor bioestabilidade e biodisponibilidade. Devido a estas propriedades, decidimos investigar uma série de aminoácidos não-canónicos, através de simulações moleculares, para desenvolver estruturas secundárias mais constrangidas (estruturalmente estáveis) e, por consequência, otimizar a função biológica de determinados péptidos.

Simulações de Dinâmica Molecular foram realizadas sistematicamente para validar novas classes de aminoácidos não naturais, e proceder à incorporação destas moléculas em novos péptidos ou em péptidos experimentalmente obtidos. Para alcançar este objetivo, o campo de forças Gromos 54a7 foi escolhido, e foram adicionadas parametrizações baseadas nos aminoácidos canónicos, necessárias à modelação desses novos resíduos.

Estudamos, assim, diversas classes de aminoácidos não canónicos, nomeadamente: α,α-dialquilglicinas simétricas, α,α-dialquilglicinas assimétricas, análogos de prolina, aminoácidos ciclizados de Cα a Cα e α,β-desidroamino ácidos. Estas classes foram escolhidas porque poucos representantes de cada classe foram estudados em detalhe, e os que foram, apresentaram aplicabilidade em medicina e no design de péptidos estruturalmente constrangidos.

As α,α-dialquilglicinas simétricas e assimétricas foram estudadas em quatro péptidos antibióticos amplamente estudados e conhecidos. Dhg (α,α-dihexilglicina) e Ac$_6$c (ácido 1-aminociclo-hexanocarboxílico), ambos simétricos, demonstraram ser indutores de estruturas helicoidais nos *peptaibols* Alameticina e Peptaibolin. Além disso, estes dois exemplos promovem pré-organização em água, fator que está relacionado com a inserção em membranas. Por outro lado, a assimétrica Isovalina (Iva), foi estudada nos *peptaibols* Antiamoebina e Zervamicina. Nesses estudos, dois aminoácidos análogos à Iva revelaram uma melhor capacidade de induzir estruturas secundarias helicoidais (hélices do tipo alfa ou 3$_{10}$), e uma melhor estruturação pode traduzir-se em melhorias na função antibiótica. Também nos estudos envolvendo os péptidos Antiamoebina e Zervamicina, avaliamos os análogos da prolina, um aminoácido conhecido por gerar regiões de elevada curvatura em péptidos. No entanto, o análogo ALP (cis-3-amino-L-proline) demonstrou aumentar o número de resíduos em helice nos dois péptidos.

As simulações envolvendo α,β-desidroamino ácidos revelaram que estes resíduos tem uma vasta aplicabilidade, pois os péptidos nos quais estão incorporados podem agregar-se formando géis capazes de transportar medicamentos, ou então, induzir estruturas secundárias menos comuns, como alfa e gama *turns*.

Resumindo, nossos estudos de simulação molecular revelaram que a incorporação de aminoácidos não-canónicos em péptidos é capaz de gerar um grande número de péptidos miméticos com diferentes preferências estruturais e diferentes aplicabilidades. Nossas descobertas mostraram que a correta escolha de aminoácidos não-naturais otimizam diversas características, dentro das quais se destacam: a permeabilidade em membrana, a pre-organização em meio aquoso, a estabilização de tipos específicos de estruturas secundarias e a resistência enzimática. Este trabalho destaca como a Modelação Molecular pode ser aplicada para uma melhor compreensão de um grande número de temas de interesse na Química medicinal.

Table of Contents

List of Abbreviations

A

Ace (acetyl)

Ac_nc (1-aminocycloalkane-1carboxylic acid)

Ac_3c (1-aminocyclopropane-1carboxylic acid)

Ac_4c (1-aminocyclobutane-1carboxylic acid)

Ac_5c (1-aminocyclopentane-1carboxylic acid)

Ac_6c (1-aminocyclohexane-1carboxylic acid)

Ac_7c (1-aminocycloheptane-1carboxylic acid)

AHMOD ((2S)-amino-(6R)-hydroxy-(4S)-methyl-8-oxodeca-noic acid)

Aib (a-amino isobutyric acid)

Ala or A (Alanine)

AMD ((2S)-amino-(4S)-methyldecanoic acid)

AMP (Antimicrobial Peptides)

Arg or R (Arginine)

Asn or N (Asparagine)

atm (atmosphere)

B

Bin (1,1'-binaphthyl-substituted α-aminoisobutyric acid)

Boc (*tert*-butoxycarbonyl or *tert*-butyl carbamate)

Boc_2O (*tert*-butyldicarbonate)

BS (Barrel Stave)

C

CAC (Critical Agreggation Concetration)

Cbz (carboxybenzyl)

CG (Conjugate Gradient)

CD (Circular dichroism)

CGC (Critical Gelation Concentrations)

COX-2 (cyclooxygenase-2)

CPP (Cell Penetrating Peptide)

Cys or C (Cysteine)

D

dAAs (a,a-disubstituted amino acids)

Db_zg (α,α-dibenzyl glycine)

Deg (α,α-diethyl glycine)

DHPC (dihexanoylphosphatidylcholine)

Dibg (α,α-di-isobutyl glycine)

DMAP (4-dimethylaminopyridine)

Dmg (α,α-dihydroxymethyl glycine)

DMPC (dimyristoylphosphatidylcholine)

DMSO (dimethylsulfoxide)

DNA (deoxyribonucleic acid)

DOPC (dioleoylphosphatidylcholine)

Dpg (α,α-dipropyl glycine)

DSSP (Dictionary of Secondary Structure in Proteins

$D\phi g$ (α,α-diphenyl glycine)

E

ECM (Extracellular Matrix)

F

FEP (Free Energy Pertubation)

FF (Force Field)

fs (fento seconds)

G

Gln or Q (Glutamine)

Glu or E (Glutamic Acid)

Gly or G (Glycine)

HMBC (Heteronuclear Multiple Bond Correlation)

HMQC (Heteronuclear Multiple Quantum Correlation)

Hyp (hydroxyproline)

I

Ile or I (Isoleucine)

Ind (aminoindane carboxylic acid)

Iva (isovaline or isovaleric acid)

K

K (Kelvin)

L

Leu or L (Leucine)

M

MD (Molecuar Dynamics)

MM (Molecular Modeling)

N

Nle (norleucine)

NMR (Nuclear Magnetic Ressonance)

NOE (Nuclear Overhauser)

Npx (Naproxen)

NSAID (nonsteroidal anti-inflammatory drug)

ns (nano seconds)

Nva (norvaline)

P

PBC (Periodic Boundary Conditions)

PCA (Principal Component Analysis)

PD (Parallel Displaced)

PDB (Protein Data Bank)

Phe (Phenylalanine)

PME (Particle Mesh Ewald

Pro or P (Proline)

R

RGE (arginine – glycine - glutamic acid)

RGD (arginine – glycine - aspartic acid)

RMSD (Root Mean Square Deviation)

RMSF (Root Mean Square Fluctuation)

RNA (Ribonucleic acid)

S

S (entropy, kJ/mol/K)

SAAs (Sugar Amino Acids)

SD (Steepest Descent)

Ser or S (Serine)

SS (Secondary Structure)

SPC (Simple Point Charge)

T

TEM (Transmission Electron Microscopy)

TFA (trifluoroacetic acid)

TFE (2,2,2-trifluoroethanol)

TI (Thermodynamic Integration)

Tic (1,2,3,4-tetrahydroisoquinolone)

Tle (*tert*-leucine or *tert*-butylglycine)

TMG (*N,N,N',N'*-tetramethylguanidine)

Tmt (β-methyl-2',6'-dimethyltyrosine)

Thr or T (Threonine)

Trp or W (Tryptophan)

Tyr or Y (Tyrosine)

V

Val or V (Valine)

vdW (van der Waals)

Others

2D (two-dimensional)

3D (three-dimensional)

Δ^EAbu (E-dehydroaminobutyric acid or dehydrobutyrine)

Δ^ZAbu (Z-dehydroaminobutyric acid or dehydrobutyrine)

ΔAla (dehydroalanine)

Δ^ELeu (E-dehydroleucine)

Δ^ZLeu (Z-dehydroleucine)

Δ^EPhe (E-dehydrophenylalanine)

Δ^ZPhe (Z-dehydrophenylalanine)

Δ^ETrp (E-dehydrotryptophan)

Δ^ZTrp (Z-dehydrotryptophan)

ΔVal (dehydrovaline)

Structure and Contents

This thesis consists of the compilation of published and submitted papers in international peer-reviewed journals and is organized in five sections, including ten chapters.

Section I corresponds to the **Introduction**, starting with a brief overview of the topic and the main objectives. There, a review paper, to be adapted for further submission (**Chapter I**) shows the most investigated classes of non-canonical amino acids, their occurrence and applications, primarily in the design of peptidomimetics.

Chapter I

Tarsila G. Castro, João C. Marcos, Nuno M. Micaêlo and Manuel Melle-Franco. *Non-Canonical Amino Acids as Building Blocks for Peptidomimetics: Structure Features Through Molecular Dynamics Simulations.*

The **Section II** - **Methods** complements the given information present in each Chapter of Results and Discussion Section. Chapter II – *Molecular Dynamics Methods*

Section III - **Results and Discussion** is divided in six Chapters and corresponds to the most important results achieved during the PhD work. **Chapters III** and **IV** show the α,α-dialkylglycine class incorporated in Peptaibols of different sizes, the Alamethicin and the Peptaibolin. **Chapter V** also refers to Peptaibols, but the classes investigated are proline analogs and D-amino acids, incorporated on Zervamicin and Antiamoebin. The **Chapter VI and VII** address the α,β-dehydro amino acids through different approaches.

Chapter III

Tarsila G. Castro and Nuno M. Micaêlo. *Modeling of Peptaibol Analogues Incorporating Nonpolar α,α- Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation.* dx.doi.org/10.1021/jp4074587 | J. Phys. Chem. B 2014, 118, 649–658.

Chapter IV

Tarsila G. Castro and Nuno M. Micaêlo. *Conformational and Thermodynamic Properties of Non-Canonical α,α-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies.* dx.doi.org/10.1021/jp505400q | J. Phys. Chem. B 2014, 118, 9861–9870.

Chapter V

Tarsila G. Castro, Nuno M. Micaêlo and Manuel Melle-Franco. *The Secondary Structure of Antiamoebin I and Zervamicin II Peptaibols Incorporating D-Amino Acids and Proline Analogues. A Modeling Study, 2015, submitted.*

Chapter VI

H. Vilaça, G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira. New self-assembled supramolecular hydrogels based on dehydropeptides, *J. Mater. Chem. B,* 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Chapter VII

Helena Vilaça, **Tarsila G.Castro**, Loly Torres Pérez, Ashkan Dehsorkhi, Cristóvão F. Lima, Catarina Gonçalves, Manuel Melle-Franco, Loic Hilliou, Miguel Gama, Ian W. Hamley, José A. Martins, Paula M. T. Ferreira. Self-healing RGD dehydropeptide hydrogel, 2015, submitted.

The **Section IV - Conclusions** includes **Chapter IX**, which summarizes the most important findings about the non-canonical amino acids under investigation, through Molecular Modeling Studies, presented in the previous sections. Also, the possibilities for further research and practical applications are discussed.

Finally, the **Section V** consists of an **Appendix** section, where the new topologies parameters of Chapters III to VII are shown. For the published articles each appendix containing the supplementary material/supporting information.

We also add as Appendix, one paper in progress that requires future simulations and analysis to better understand the results obtained to date: **Tarsila G. Castro**, Nuno M. Micaêlo and Manuel Melle-Franco; *Conformational Properties of the Non-canonical Cyclic Ac_nc Amino Acids: A Molecular Modeling Study*, 2015).

INTRODUCTION

Overview

Peptidomimetics are molecules that mimic the three-dimensional structure of a natural peptide and retain the capacity to interact with biological targets and generate the same biological effect. Peptidomimetics are designed to circumvent some of the problems associated with natural peptides, like stability against proteolysis and poor bioavailability. Nowadays, academic research labs and small biotech companies are emerging with rational design strategies to discover novel therapeutic peptides such as: antibiotics, anticancer, neuromodulator, opioid, hormones, vaccines, radiolabeled peptides and self-assembled peptides for bioengineering.

Usually, peptidomimetics are composed by non-canonical (unnatural) amino acids that enhance specific secondary structures and/or its biological activity. Contrary to the well-known encoded (natural) amino acids, the structure and function of these residues is far for being fully understood, limiting our capacity in the rational design of novel peptidomimetics.

The rational design of new peptidomimetics is highly dependent on our knowledge about the structure-function relation properties of the non-canonical amino acids. Relative few theoretical and structural studies that elucidate the conformational properties of peptidomimetics are found in literature and NMR/X-ray structures of these molecules are almost absent from the protein data bank (PDB).

In this sense, our intention for these four years of PhD, was to fill the gap regarding the knowledge of the structure-function properties of new non-canonical amino acids. This was accomplished using molecular modelling methodologies. The outcome of this research is going to enable a more rational understanding of the conformational preference of peptidomimetics and model peptides bearing unnatural AA and, will make possible the future development of more effective peptidomimetics and the design of novel foldamers. From a molecular modelling perspective, this study provides the scientific community a ready-to-use large library of validated new unnatural AA parameterizations for the GROMOS biomolecular force field.

Objectives

The main objective of this work was to create a library of non-canonical amino acids and their incorporation into a force field, so that the simulation of these molecules would be readily available to the scientific community.

It was intended to model and parameterize several classes of non-canonical amino acids following a common criteria and nomenclature and to validate the proposed structures of these molecules with molecular dynamics simulations in different media (individually, inserted into peptides, in aqueous medium, in a membrane, in organic solvents, etc.).

The simulations conducted aimed at classifying different unnatural amino acids according to their structural and functional characteristics. It is expected to bring together amino acids that induce α-helices, β-turns, β-hairpin and β-sheets, and that these properties will improve biological function.

Also, we aimed to explain the structural and functional differences resulting from the insertion of non-canonical amino acids in peptides with well-known activity and/or that have been already synthesized and characterized experimentally. Ultimately, the results obtained, make possible to generalize structural restrictions and suggest foldamers, which is relevant for the design of peptidomimetics.

Structurally, we expected to confirm the general type of secondary structure that each non-canonical amino acid promotes. Some articles in the literature already indicate predominant conformational features for some of the amino acids under study. However, we thought necessary to confirm these characteristics in a broader context. In this sense, analytical techniques such as Root Mean Square Deviation (RMSD) Root Mean Square Fluctuation (RMSF), Ramachandran Plots, Secondary Structure (SS) and Hydrogen Bond were implemented to analyze the different possible secondary structures and the level of conformational restriction that these amino acids induce to the peptides in which they are inserted. Some of these analysis tools have to be modified or adapted to be applicable to amino acid changes in the main chain, since most programs were developed to recognize only peptide bonds and typical torsion angles.

Finally, after individual validation and in case studies, the non-canonical amino acids will be incorporated into peptides composed of natural amino acids in order to propose new foldamers (peptides that have a tendency to adopt a specific compact conformation).

Chapter I

Non-Canonical Amino Acids as Building Blocks for Peptidomimetics:
Structure Features Through Molecular Dynamics Simulations

1. Introduction

This review focuses on the major differences between encoded and non-canonical amino acids, which give to the latter the ability to be successfully incorporated into peptides, generating peptidomimetics for medical use. Most of the findings about non-canonical amino acids to date are based on experimental studies. Driven by this fact, we gather here results for some classes of residues, concerning the structure and function of these molecules, from experiments and/or molecular simulations.

Peptides and proteins have been exhaustively studied in the past decades, especially peptides, due to its great potential as drugs. These entities play important functions, as hormones, neurotransmitters, inhibitors, etc., which are essential for human life.[1-7] However, the use of peptides as drugs has major drawbacks with bioavailability and biostability.[2, 8-10]

The degradation by proteases and problems concerning nonselective molecular receptors due the high inherent flexibility are some of the disadvantages of natural peptides. In adittion, pharmacokinetics, the relation on how the human body impacts petides is also a process that does not favor the use of these molecules as drugs. The pharmacokinetic process consists of different stages, namely: absorption, distribution, metabolization and excretion. In these phases peptides have common problems like poor oral availability, poor transport properties (through cell membranes) and rapid excretion through the liver and kidneys. The enzymatic stability of a peptide is related to several factors as amino acids composition, secondary structure, flexibility, lipophilicity, among others.[8-10]

To overcome the problems mentioned above, protein-like sequences, called peptidomimetics have been designed and tested.[2, 4-5, 11-13] The most common way to generate peptidomimetics is through modifications of the native/encoded amino acids, so that the new peptide has a similar secondary structure and maintains or improves biological function. For instance, the hydrolysis of peptide bonds by proteases can be obstructed through the introduction of atypical moieties, as D-amino acids, non-canonical amino acids or by introducing a N-alkyl group.[9, 14-16]

The rational design of new peptidomimetics is highly dependent on our knowledge about the structure-function relation properties of non-canonical amino acids. Only few theoretical and structural studies about the conformational properties of unnatural amino acids and peptidomimetics are available to date. In this sense, our work tries to create a new non-

canonical amino acid library, suggesting alternatives with better characteristics as foldamers and its possible applicability.

1.1. Amino Acids and Peptides

Encoded amino acids are organic molecules presenting a carboxylic (COOH) and an amine (NH_2) groups linked to a chiral carbon atom, named Cα. They are the fundamental building units of peptides and proteins, i.e, when two or more amino acids are linked through amide bonds (peptide bond). Two amino acids link through a condensation reaction releasing a water molecule. There are 20 natural amino acids encoded by DNA, which constitute most known proteins and enzymes.
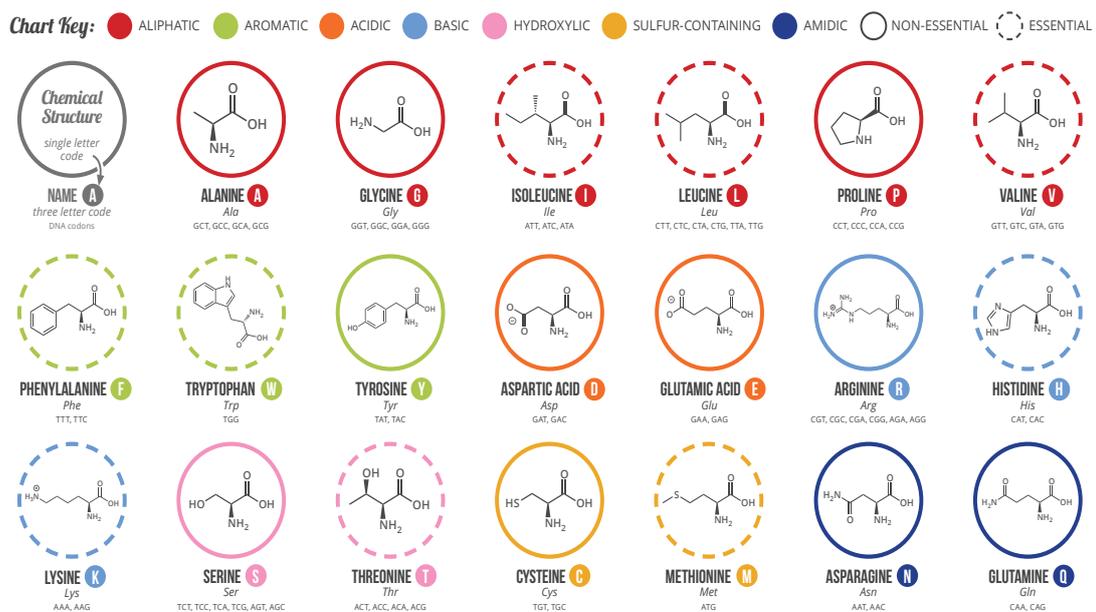
Figure 1. The 20 encoded amino acids in the human genetic code. The chart key helps to illustrate the different fundamental properties. Essential amino acids must be obtained from the diet while the non-essential ones can be synthesized in the human body. A most general classification divides the amino acids as nonpolar (Gly, Ala, Leu, Ile, Val, Cys, Met, Pro, Phe, Trp), polar uncharged (Ser, Thr, Tyr, Asn, Gln), acidic charged (Asp, Glu) and basic charged (Lys, Arg, His).[17]

Exceptions to the 20 canonical amino acids were reported.[18] For example, Hydroxyproline (Hyp) and Hydroxylysine occur on protein collagen. They are produced by hydroxylation of the amino acids proline and lysine, respectively, by the correspondent hydroxylase enzyme, as a post-translational modification.[19-21] The α-aminoadipic acid present on corn proteins is another example and it is an intermediate in the lysine metabolism.[22] The penicilamine is an α-amino acid metabolite of penicillin, similar to Cysteine, and it is used to treat arthritis.[23] Ornithine participates in the urea cycle, as one of the products of the action of the enzyme arginase on L-

arginine.[24] Citrulline, naturally found in watermelon, is an amino acid derived from arginine.[25] The structures of these amino acids are shown in Figure 2.

Importantly, although by definition there are 20 amino acids encoded by DNA, there are two other residues that are proteinogenic: selenocysteine (Sec)[26] and pyrrolisine.[27] Both amino acids appear in proteins of Archea organisms [28] However Sec is a naturally found residue in all kingdoms of life as the building block of selenoproteins. Sec is considered the 21st amino acid and has been found in 25 human selenoproteins and selenoenzymes.[29] Sec is encoded by a UGA codon, which is normally a stop codon, but acts by performing a translational recoding.[30] Pyrrolysine is incorporated during translation by the genetic code, just like standard amino acids. It is encoded in mRNA by the UAG codon, which in most organisms is a stop codon, similar to UGA.
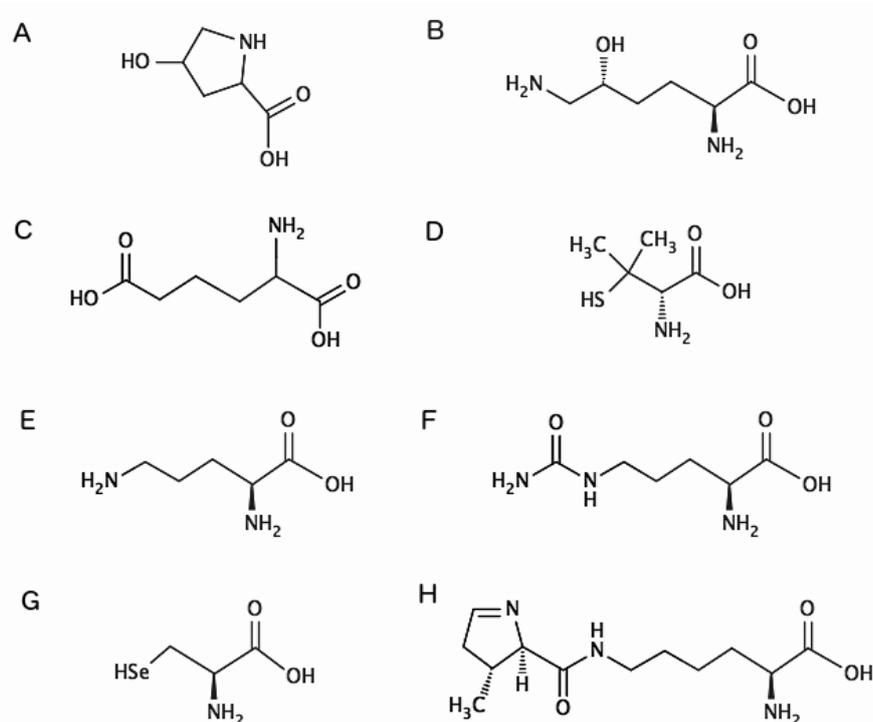
Figure 2. Two-dimensional structures of some non-canonical amino acids naturally found in nature or in post-translational processes. (A) hydroxyproline, (B) hydroxylysine, (C) α-aminoadipic acid, (D) penicilamine, (E) L-ornithine, (F) citrulline, (G) selenocysteine and (H) pyrrolysine.

Peptides are amino acids polymers, which when short are classified as oligopeptides and when larger, are polypeptides. The exact terminology, in accordance with the length, is quite variable.[31] Some sources consider oligopeptides sequences of 2-10 amino acids, other 2-20 and yet, 2-40 residues are also reported.[32] Oligopeptides may also be classified based on molecular structure, for instance: aeruginosins, cyanopeptolins, microcystins, microviridins,

microginins, anabaenopeptins and cyclamides.[33-35] Polypeptides are peptides that contain longer, continuous, and linear peptide chains. All proteins are polypeptides, but the reverse is not true since a protein has a specific sequence generated by a gene. Peptides are the building blocks of proteins, which have a fundamental biological function; they make up the living organisms. However, small naturally occurring peptides, may present, alone, important biological functions.[36-37] Some vertebrate hormones such as insulin, glucagon, and corticotropin comprise less than 50 amino acid residues. Examples of small naturally occurring peptides are the hormones oxytocin,[38] thyrotropin[39] and enkephalin.[40] Also, certain fungi are highly toxic and contain peptides, as amanitin, with important uses in medicine.[41]

1.2. Encoded Amino Acids Properties

Any substance that contains a carbon atom with four different substituents occurs in the form of two optical isomers, i.e. present optical activity to rotate the polarization plane of light to the right (clockwise) or leftwards (counter clockwise). Nineteen of the amino acids are chiral and found in the configuration L. The only exception is Gly, which does not have any carbon atoms with different substituents; it is an achiral molecule. L and D configurations refer to L and D configurations of glyceraldehyde. Nine of the nineteen chiral L-amino acids commonly found in proteins are dextrorotatory.[42-43]
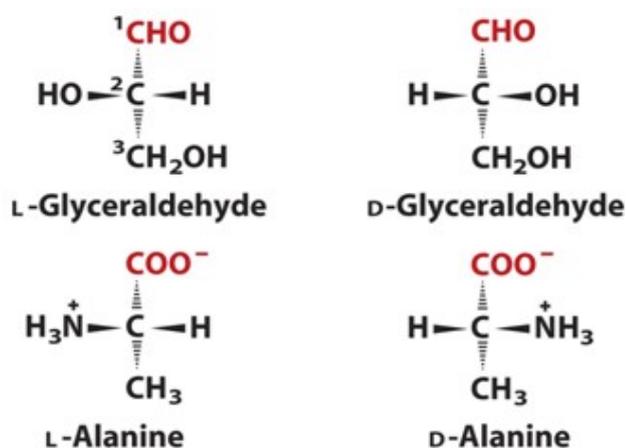
Figure 3. Fischer projection of L and D glyceraldehyde, and L and D Alanine.

The R/S nomenclature system is a more general method for denoting enantiomers. This classification method does not involve a reference molecule such as glyceraldehyde, instead, it labels each chiral center as R or S according to a system, which assigns a priority based on atomic number to each substituent. In the case of amino acids, if the center is oriented so that the H atom is pointed away from a viewer, the viewer will then see two possibilities: the

decreasing priority of the remaining three substituents in a clockwise direction, is labeled R (for Rectus, Latin for right). However, if it decreases in a counter-clockwise direction, it is S (for Sinister, Latin for left).

For most amino acids, the L form corresponds to an S absolute stereochemistry. Only L-cysteine is (R)-cysteine, but this only reflects the fact that the sulfur atom has a higher priority than a carbon atom, and does not reflect a real difference in 3D structure.

The geometry characteristics of the encoded amino acids residues are normally obtained from crystal structures of related molecules. Bond lengths and bond angles are essentially invariant among the 20 amino acids.[44-45] Only the backbone N—Cα—C angle, τ, varies and causes variation on the angle of the tetrahedral center. In other words, despite the Cα being tetrahedral, which would give 110°, τ can sometimes stretch to larger values in order to accommodate other strains in the structure.[46-47]
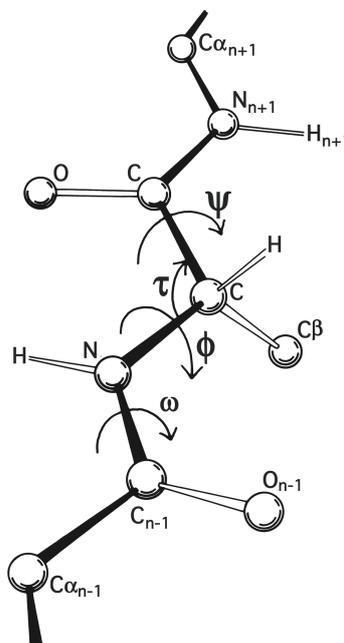
Figure 4. Standard peptide representation, with indication of angle t (N—Cα—C), and dihedrals, ω (C-N), φ (N-Cα) and ψ (Cα-C).[48]

The peptide bond (C-N) restricts the dihedral angle ω to values very close to 180°, generating the typical *trans* configuration. In peptides containing proline the *cis* form can be found, with ω = 0°. The distance between the Cα atoms in the *trans* and *cis* isomers is approximately 3.8 and 2.9 Å, respectively. The proline (Pro) ring is not completely flat and also induces stronger stereochemical constraints due to the lack of the flexible backbone NH, necessary for the formation of hydrogen bonds. These unique properties of Pro disrupts helical secondary structure (SS) and promotes turn SS in peptide chains.[49]

Another encoded amino acid with unique characteristics is glycine (Gly). Due to its small size and flexibility, Gly can assume conformations normally forbidden by close contacts of the β-carbon on other residues. Also, the lack of chirality allows that this amino acid adopts both right-handed and left-handed conformations.

The parameters φ and ψ are the most important for amino acids structure, and consequently, for peptide conformation. They are the backbone dihedrals, and in theory, the average φ and ψ values for α-helices and β-sheets are predicted to be between -57, -47 and -80, +150, respectively. However, for experimental structures these values were found to be different.[50-51]

Another important characteristic of natural amino acids is that they are amphoteric molecules (can behave as acids, as well as bases) and zwitterionic varying with the pH. They are neutral molecules at physiological pH (≈7.4), yet carry a positive and a negative electrical charge. Figure 5 shows the intramolecular proton transfer that generates a zwitterion. This form exists in the solid state[52] and in water solution. In rare cases the zwitterion form is also stable in the gas phase, like for the residue Arginine.[53]
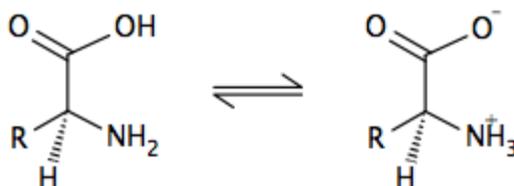
Figure 5. Graphical representation of amino acid isomers. The isomer on the right is a zwitterion.

The zwitterionic form is pH dependent. At physiological pH a carboxylate group and a protonated amine occur simultaneously. At low pH values, an acidic medium, a hydrogen ion is added to the carboxylate group, resulting in a global net charge of +1 (still present on the amine). On the other hand, at high pH values, a basic medium, a hydrogen ion is removed from the amine group, by the excess base, turning the global net charge to -1.
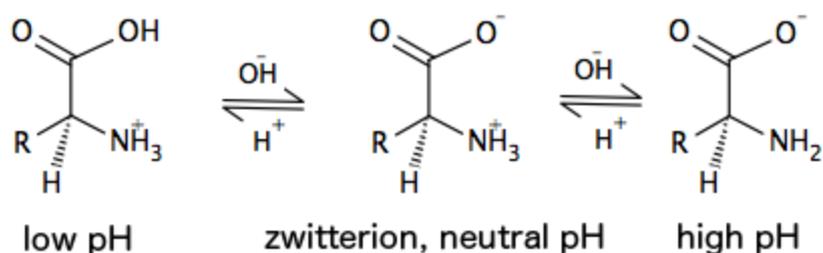
Figure 6. Representative scheme of the change of zwitterion form to positively or negatively charge amino acids according to the pH.

The amino acids arginine, lysine and histidine are positively charged at physiological pH, while aspartate and glutamate are negatively charged in the same conditions. Due to the characteristic of presenting a third pK_a, they are named triprotic, with the third value associated with the ionizable functional group on the side chain. Amino acid backbone modifications or alterations on amine and carboxylic acid termini may change the zwitterionic nature of encoded amino acids.

1.3. Peptide Profile and Biological Function

Peptides SS and other properties, such as hydrophobicity or polar profile are directly related with their function. Cell-penetrating peptides (CPP) and antimicrobial peptides (AMP), both membrane active peptides are examples where the amino acid content, and the resulting peptide properties, relates directly with their function. The CPPs present great potential as drug delivery peptides and the AMP of antibiotic candidates.[54-55]

Penetratin is a well studied CPP, which acts as antifungal and adopts a helical SS in an environment of low polarity (interior of cell membranes).[56] Penetratin analogues should conserve the SS to achieve the same or optimized function.[57]

Temporin A is a small, highly hydrophobic AMP, found in the skin of the European red frog. This peptide proved to be active against both Gram-positive and Gram-negative bacteria, with the advantage of not being toxic to human red blood cells at the concentrations required to kill bacteria.[58] Wade and co-workers reported the insertion of D-amino acids to generate an antibiotic analogue that would resist enzymatic proteolysis.[58]

Peptaibols belong to the class of AMPs and are peptides rich in the non-canonical amino acid Aib (α-amino isobutyric acid). Many peptaibols interact with cell membranes through a barrel-stave channel model. They are mostly helical entities, which allow the optimal channel formation necessary for biological function. We reported the structural properties of a series of non-coded amino acids inserted in two different peptaibols, Peptaibolin and Alamethicin,[59-60] obtaining improvements on peptide conformation stability and function.

We also worked, recently, on peptide hydrogelators carrying α,β-dehydroamino acids.[61] These peptides can be used for drug delivery, due their capability to self-assemble as a hydrogel. We proved that the aggregation process occurs due to the non-canonical ΔPhe, which interacts with the Npx (naproxen) group also present in our model peptides, through π–π interactions.[61]

The peptides mentioned above are only a few examples where biological function is directly connected to the SS or to amino acids content, or even, to the type of inter and intramolecular interactions which these residues can perform.

2. Peptidomimetics Design

2.1. Structural Properties of Non-Canonical Amino Acids

Non-canonical amino acids are organic molecules also containing an amine and a carboxylic acid group but are not directly encoded by the DNA. However, as mentioned before, some residues are found in nature. In addition, a large array of non-canonical amino acids can be synthesized.[62]

The incorporation of non-canonical amino acids into peptides is one of the approaches to generate peptidomimetics that overcome the problems previously mentioned concerning the pharmacokinetics and enzymatic stability of natural peptides as drugs. In fact, the replacement of natural amino acids often results in higher activity and increased biological stability.[4, 63-64]

Figure 7 summarizes the most common natural and artificial modifications applied to encoded amino acids, used to generate peptidomimetics.
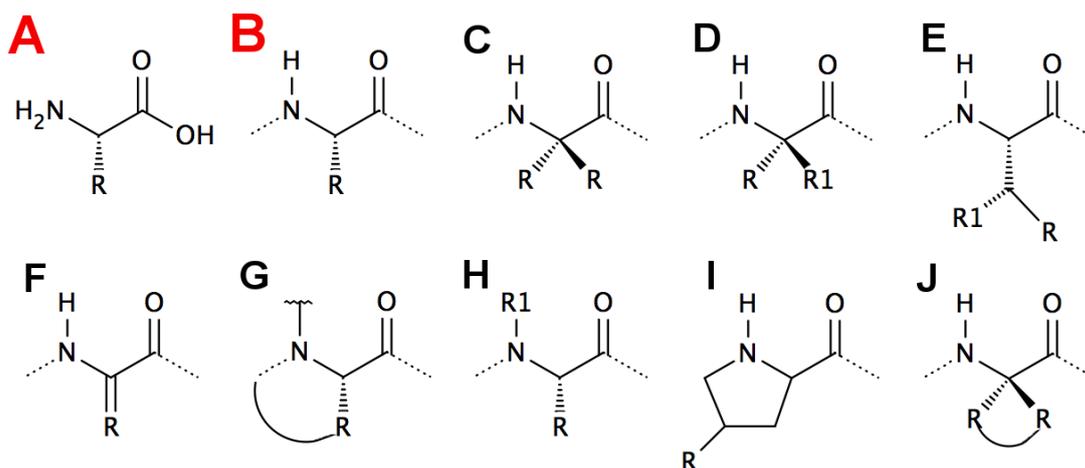
Figure 7. (A) and (B) are 2D representations of the basic structure of an encoded amino acid. (C) is a symmetrical α,α-dialkyl glycine, (D) is an asymmetrical α,α-dialkyl glycine, (E) is a β-substituted amino acid, (F) is a α,β-dehydroamino acid, (G) represents a N-cyclization, (H) represents a N-alkylation, (I) represents proline analogues and (J) represents cyclized amino acids (known as Ac_nc).

2.1.1. Symmetrical α,α -dialkyl glycines

The most widely studied class of non-canonical amino acids is probably the class of α,α-dialkyl glycines (Figure 8). This type of residue is found in many natural occurring peptides, especially, in antimicrobial peptides.[65-67] The Aib (α-aminoisobutyric acid) is the prototype of this class, and known to restrict the dihedral angles to generate α-helical conformations.[68-70]

Aib was successfully incorporated in peptides as enkephalin, bradykinin and angiotensin II,[71] generating active and constrained peptidomimetics. Also, Ac_6c (1-aminocyclohexane-1-carboxylic acid) has been tested on enkephalin and endomorphin peptides, to achieve peptidomimetics with large activity in vivo.[72-73] Ac_6c is both an α,α-dialkyl glycine (because it is alkyl disubstituted at Cα) and a residue of Ac_nc residues, where the chains attached to the Cα are involved in a Cα to Cα cyclization.

Our studies regarding the incorporation of α,α-dialkyl glycines suggest that some residues of this class are more capable of inducing α-helical conformations and promoting spontaneous membrane permeation than the native Aib in peptaibolin or helical structures in Alamethicin. The best results were obtained for Dhg (α,α-dihexyl glycine) and Ac_6c (α,α-cyclohexyl glycine).[59-60]
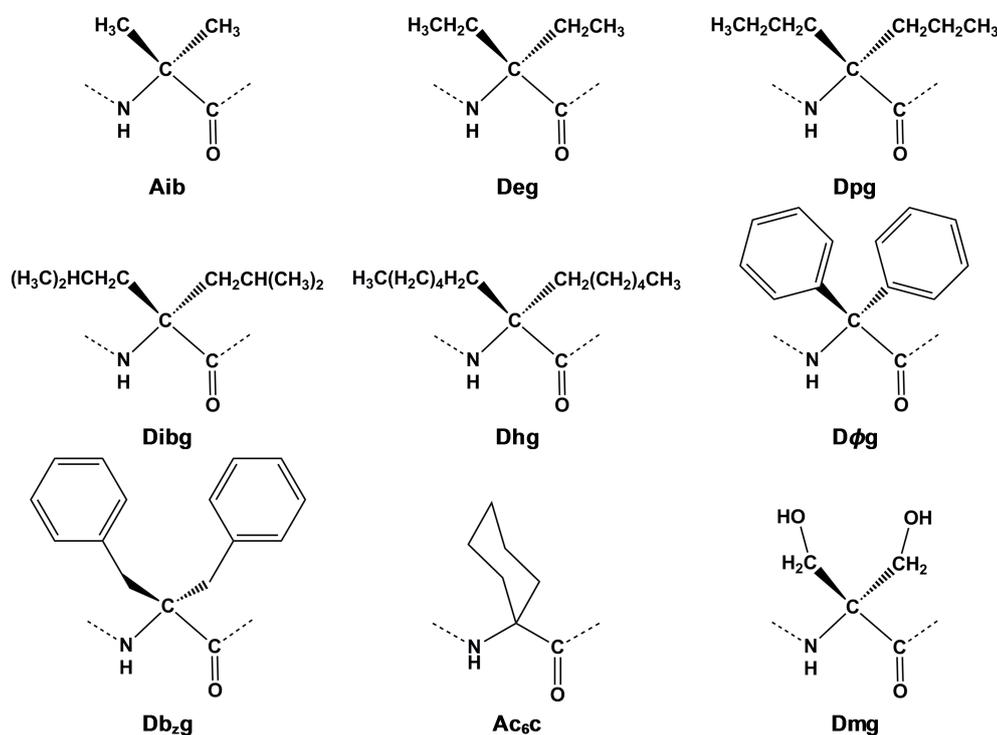
Figure 8. Two-dimensional structures of α,α-dialkyl glycines: α-amino isobutyric acid (Aib), α,α-diethyl glycine (Deg), α,α-dipropyl glycine (Dpg), α,α-di-isobutyl glycine (Dibg), α,α-dihexyl glycine (Dhg), α,α-diphenyl glycine (DΦg), α,α-dibenzyl glycine (Db_2g), α,α-cyclohexyl glycine (Ac_6c), and α,α-dihydroxymethyl glycine (Dmg).

2.1.2. Asymmetrical D-α,α-dialkyl glycines

The disubstituted amino acids can also be asymmetrical molecules, where the substituents attached to the Cα are different. The presence of two different alkyl groups makes the carbon chiral, and consequently L or D. The best-known amino acid of this class is Iva (isovaline), and it is typically found in peptaibols on D arrangement.[74-76]

Ross and co-workers[77] reported in 1993 the synthesis of a-amino acids, including three asymmetrical α,α-dialkyl glycines. Mendel and co-workers[78] reported the protein biosynthesis with conformational restricted residues, addressing different classes of amino acids, which included Iva and other asymmetrical disubstituted amino acids This approach successfully generated peptides with well-defined secondary structures.
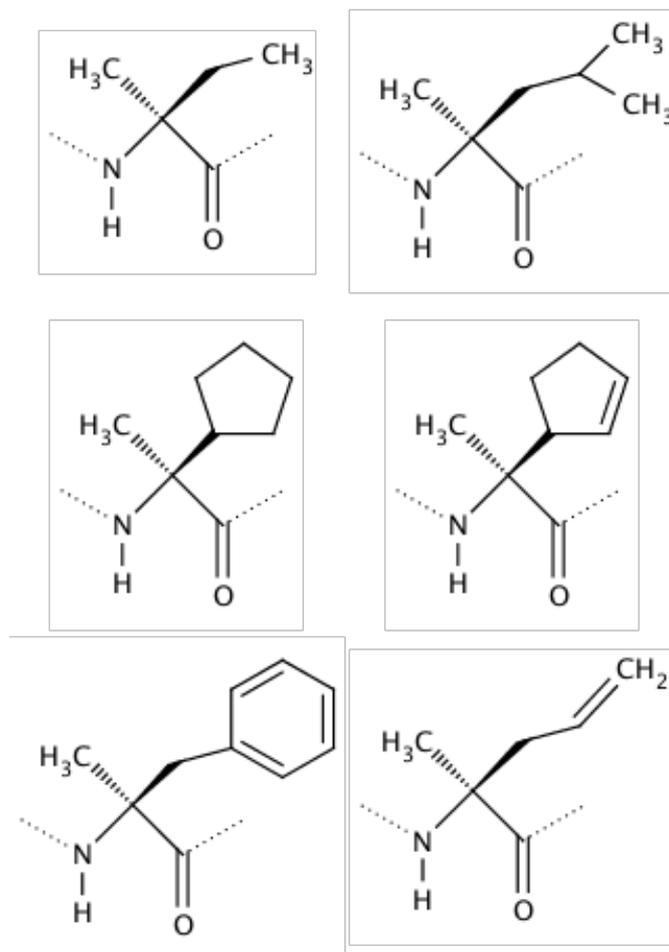
Figure 9. Two-dimensional structures of D-Iva (top of left column) and known asymmetric α,α-dialkyl glycines.

2.1.3. Cα to Cα cyclized amino acids - Ac$_n$c residues

Cyclized Ac$_n$c residues have been widely studied over the past decades through experimental and theoretical methods.[79-87] The conformational preferences of these residues vary according to the cycle. Previous experimental and theoretical results indicate that the Ac$_n$c with cycles with more than 3 atoms (n = 4–12) explore, mostly, a main chain geometry similar to Aib ($\varphi,\psi \approx \pm60^{\circ}, \pm30^{\circ}$) which is typical of 3$_{10}$-helix or α-helix SS [79, 87-93]. The residues Ac$_5$c (1-aminocyclopentane-1-carboxylic acid) and Ac$_6$c (1-aminocyclohexane-1-carboxylic acid) have been found to originate γ-turn conformations in small peptides [81, 94]. On the other hand, Ac$_3$c (1-aminocyclopropane-1-carboxylic acid) is the only member of Ac$_n$c family that prefers

molecular geometries in the *bridge* region ($\varphi, \psi \approx \pm90^{\circ}, 0^{\circ}$) and this particularity [81, 95-98] has been the subject of experimental and theoretical studies over the past decades.[84, 88, 95, 99-101]
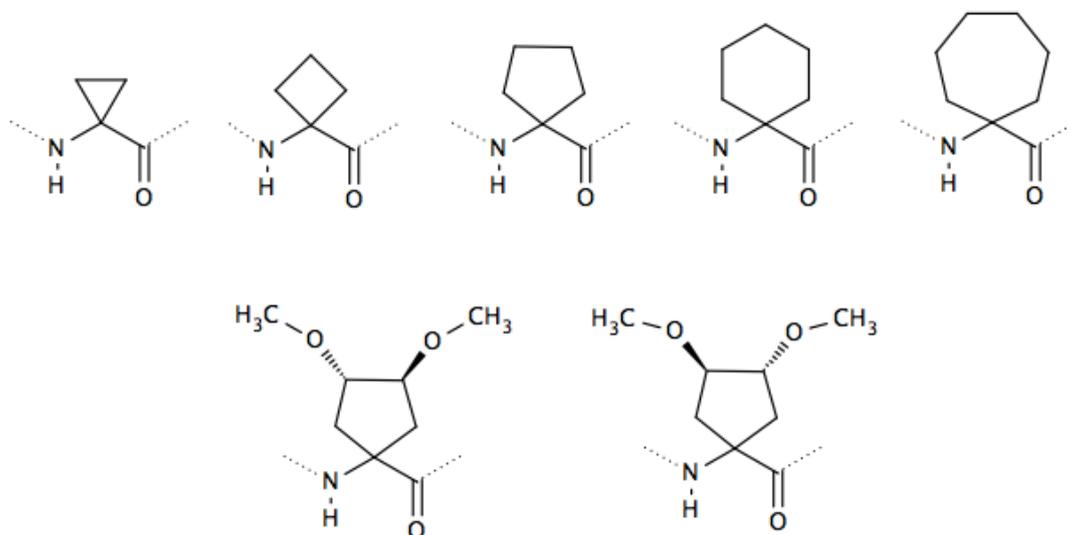
Figure 10. Two-dimensional structures of non-canonical Ac_nc (1-aminocycloalkane-1-carboxylic acids) residues, where n refers to the size of the cycle: Ac_3c, Ac_4c, Ac_5c, Ac_6c, Ac_7c, (S,S)-Ac_5c^{dOM} and (R,R)-Ac_5c^{dOM} .

2.1.4. Proline Analogues

Proline analogues represent a class with unique conformational features, since the natural Pro residue is known to disrupt or prevent α-helix SS and favors the formation of β-turn structures. Amino acids analogs of proline have been studied experimentally and theoretically, to understand structure preference and applications.[49, 102-105] Pro derivatives have been found in proteins of microbial and marine species.[9]
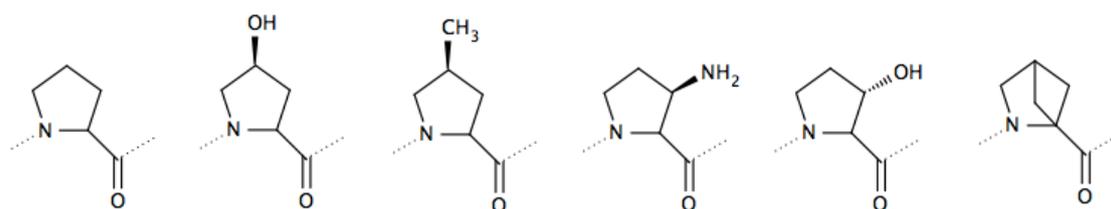
Figure 11. Two-dimensional structures of the encoded amino acid Pro and proline analogues. From left to right: L-Pro, 4-hydroxy-L-proline (Hyp), cis-4-methyl-L-proline, cis-3-amino-L-proline, trans-3-hydroxy-L-proline and 2,4-methyl-proline.

2.1.5. β-substituted and planar amino acids

β-substituted amino acids have been used to generate more potent peptidomimetics of naturally occurring peptide hormones, as opioid peptides, angiotensin or somatostatin.[9, 106]

Natural amino acids as Phe, Trp and Tyr are found in the pharmacophore of many peptide hormones. The addition of alkyl groups on the β position has been proved to be a powerful strategy to rigidify the residue and optimize the activity.[4, 107-110]
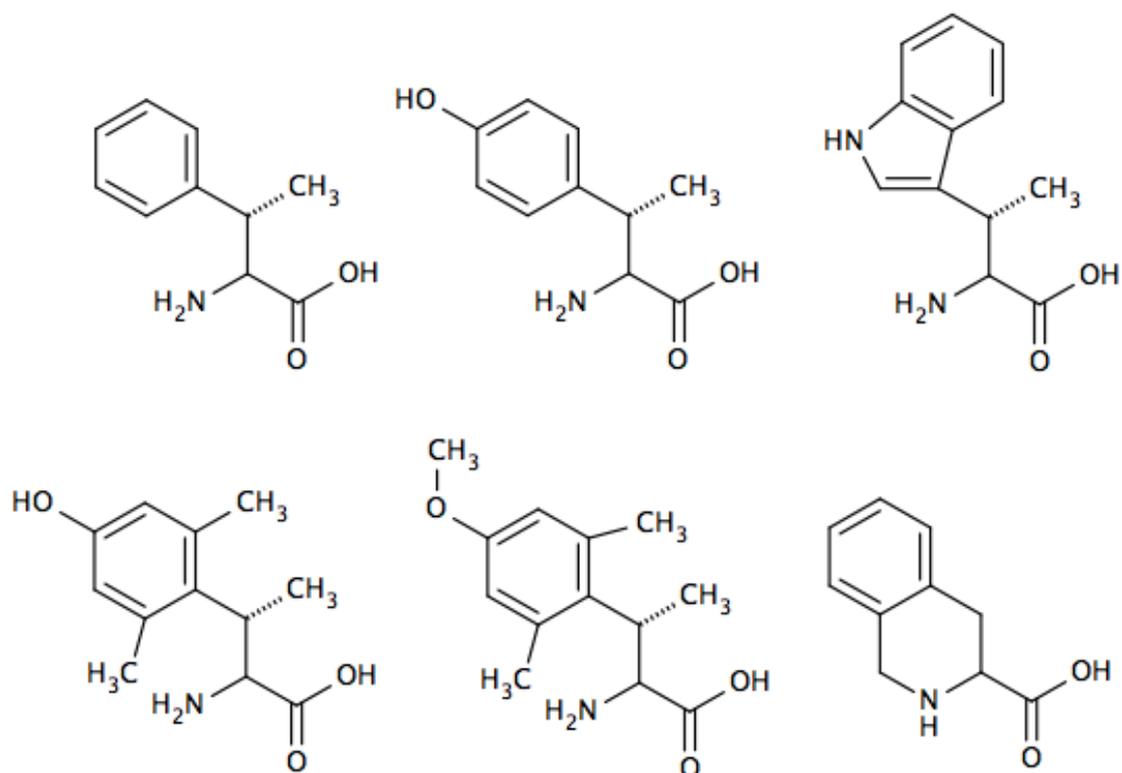
Figure 12. Structures of selected examples of non-canonical β-substituted amino acids. From left to right, and top to bottom: β-MePhe, β-MeTyr, β-MeTrp, Tmt (β-methyl-2',6'-dimethyltyrosine), β-methyl-2',6'-dimethyl-4'metoxytyrosine, Tic (1,2,3,4-tetrahydroisoquinoline).

2.1.6. α,β-dehydroamino acids

α,β-Dehydroamino acids are non-canonical amino acids naturally found in peptides.[111-113] The lack of asymmetry due to the planar hybridization sp^2 of the Cα carbon, separates this class of amino acids from the encoded ones. In addition, these residues can present β-substituents as isomers Z and E, and the possibility of π-electron conjugation. All these properties contribute to a very specific constrain which influences the bioactivity and applications of the dehydropeptides.

The conformational properties of peptides carrying α,β-dehydroamino acids have been extensively reviewed.[111-118] The three residues, dehydroalanine (ΔAla), dehydrobutyrine (ΔAbu) and dehydrophenylalanine (ΔPhe) are the most investigated.[119-126]
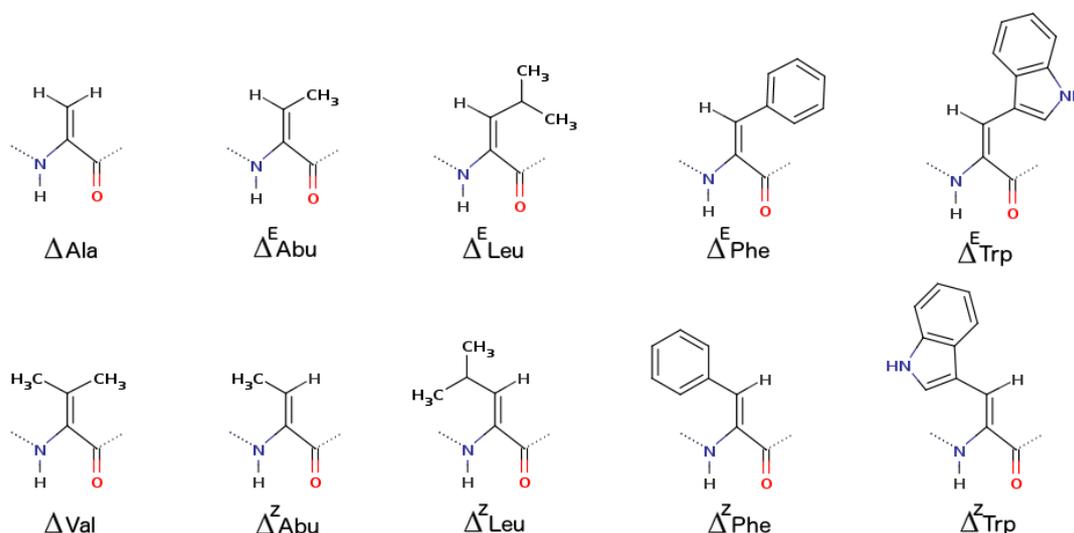
Figure 13. Two-dimensional structures of non-canonical α,β-dehydroamino acids: dehydroalanine (ΔAla), dehydrobutyrine (ΔAbu), dehydroleucine (ΔLeu), dehydrophenylalanine (ΔPhe), dehydrotryptophan (ΔTrp) and dehydrovaline (ΔVal). Those who present Z/E forms are: ΔAbu, ΔLeu, ΔPhe and ΔTrp.

This type of residue favors the formation of β-turns. In small peptides, when the dehydroamino acid is placed in the second position, especially ΔPhe, β or γ turns are the most probable arrangements. In intermediate or long peptides, sequential placement or sequential repeats of ΔPhe, induce repeated β-turns that can be accommodated in a 3_{10}-helix.[127]

2.1.7. Others Side Chain Modified Amino Acids

The amino acids and applications already mentioned show that non–canonical constrained amino acids have acquired considerable importance in the design of bioactive peptidomimetics. Figures 14 and 15 show selected examples of non-canonical residues that differ from the classes addressed above.

The amino acids Bin and Bip are reported to combine structural features of both Db_zg and Ac_7c residues.[128-130] In fact, Bip and Bin can be considered turn/helix inducers and due the characteristic of being rigid structures diminish physiological vulnerability.[128-131]

Daf is another example of a rigid amino acid that imposes geometrical constrains when inserted into a peptide. This residue possesses the unique property of also being a ligand that can coordinate metal atoms. This fact is very important allowing a broad spectrum of applications: metal-binding sites on proteins, peptide-based electronic devices and molecular switches.[132-133] The expected conformations for Daf would be β-bends and α/3_{10}-helix forms, since this residue can be classified as a α,α-disubstituted glycine, similar to Aib or Ac_7c.

However, a C5 conformation (fully extended form) was characterized experimentally, with a tendency to form a helical structure.[132]
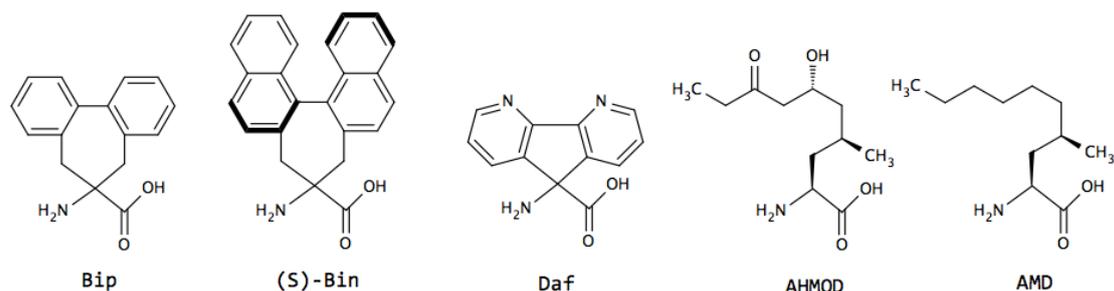
Figure 14. Selected examples of unnatural amino acids: Bip (2′,1′:1,2;1″,2″:3,4-dibenzcyclohepta-1,3-diene-6-amino-6-carboxylic acid), Bin (1,1′-binaphthyl-substituted α-aminoisobutyric acid), Daf (9-amino-4,5-diazafluorene-9-carboxylicacid), AHMOD ((2S)-amino-(6R)-hydroxy-(4S)-methyl-8-oxodeca-noic acid) and AMD ((2S)-amino-(4S)-methyldecanoic acid).

AHMOD and AMD are both naturally found on culicinin peptaibols. Culicinins are peptides isolated from the fungus *Culicinomyces clavisporus*.[134] Importantly, culicinin D was found to exhibit potent antitumor activity.[134-135] The spatial structure of Culicinins is a right-handed helix, with a tighter N-terminus, forming a 3_{10}-helix conformation.[134]
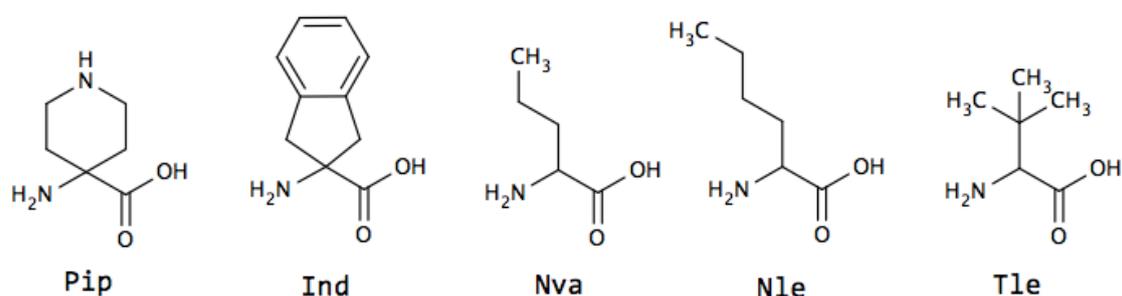
Figure 15. Two-dimensional structures of the non-canonical amino acids: Pip (4-aminopiperidine-4-carboxylic acid), Ind (aminoindane carboxylic acid), Nva (norvaline or 2-Aminopentanoic acid), Nle (norleucine or (2S)-2-aminohexanoic acid) and Tle (*tert*-leucine or *tert*-butylglycine).

The non-canonical amino acids norvaline (Nva), norleucine (Nle) and *tert*-leucine (Tle), are hydrophobic residues. Nva and Nle proved to be helical stabilizing amino acids.[136-137] Nva and Nle are found in small amounts in some bacterial strains.[138] Nva Norvaline is known to promote tissue regeneration and muscle growth,[139] while Nle can act as an isostere of methionine.[140] In contrast, Tle does not induce the same constrain observed for Nva and Nle, varying with the environment and amino acid content of the peptide in which is inserted.[141-142]

Pip is a naturally occurring amino acid found on Efrapeptin peptides, which are produced by fungi of the species *Tolypocladium*.[143] This class of peptides has antifungal, insecticidal, and mitochondrial ATPase inhibitory activities. The right-handed α-helical structure cannot be adopted by Pip-rich peptides. For efrapeptin, for example, the dominant structure is a 3_{10}-helix.[143] Pip was also reported to increase water solubility of peptides.[141, 144] The non-canonical residue Ind has a stabilizing effect towards the formation of $\alpha/3_{10}$-helices.[141, 145-146]

A recent review by Rogers and Suga shows that genetic code reprograming can generate functional non-proteinogenic amino acids.[147] Selected examples are Phe-like residues, Lys-like, peptoids, D-stereochemistry, N-Acyl and N-Acetyl.

2.2. Backbone Modifications

Peptide backbone plays an important role on peptide stabilization. Modifications on the peptide backbone are another approach to generate peptidomimetics more conformationally constrained and thus more stable. Many types of backbone modifications have been performed and tested.[3-4, 9, 11, 63, 148] Basically, a backbone can suffer alteration by isosteric or isoelectronic substitutions, resulting in several types of mimetics. The isosteric modification consists in maintaining the same number of valence electrons, but can differ in the number of atoms and atom types, while a isoelectronic substitution refers to two atoms, ions or molecules that have to present the same electronic structure and/or same number of valence electrons, but also, the same structure (number of atoms and connectivity).[2-3, 9, 12, 148]

Figure 8 exemplifies the most important peptide backbone modifications. We can cite, the replacement of $C\alpha$, the backbone extension, carbonyl replacement, etc.

Detailing some types of backbone modifications, we have the azapeptides, where a N atom replaces the $C\alpha$. The peptides generated through this transformation can be therapeutically applied as inhibitors of cysteine proteases.[4, 149]

The depsipeptides are also very important, and the result of the replacement of amide to ester bond. Two remarkable examples are the depsipeptides extract from marine invertebrates, Didemnin B and Plitidepsin (dehydrodidemnin B).[150]
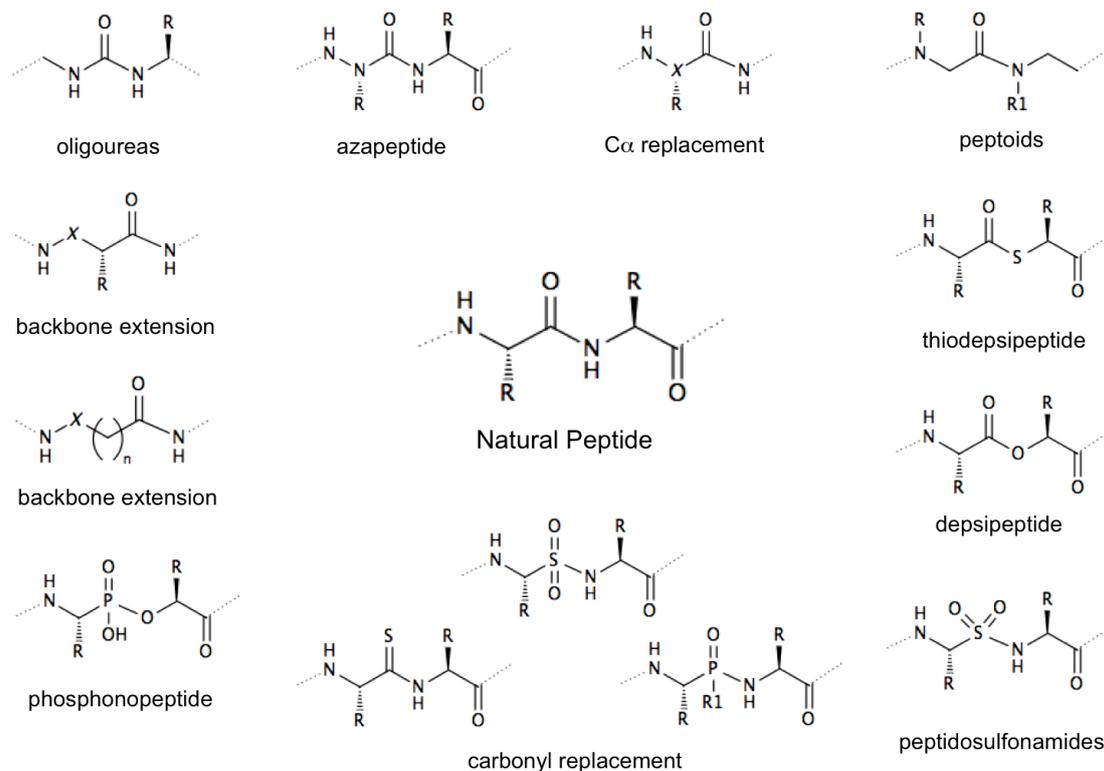
Figure 16. Two-Dimensional structure of a natural peptide (center) surrounded by known types of peptide backbone modifications.

Didemnin B has potent biological activity. One example is the strong antiviral effect against DNA and RNA viruses such as herpes simplex virus type 1.[151] Importantly, this peptide is a strong drug candidate to treat small cell lung cancer.[152]
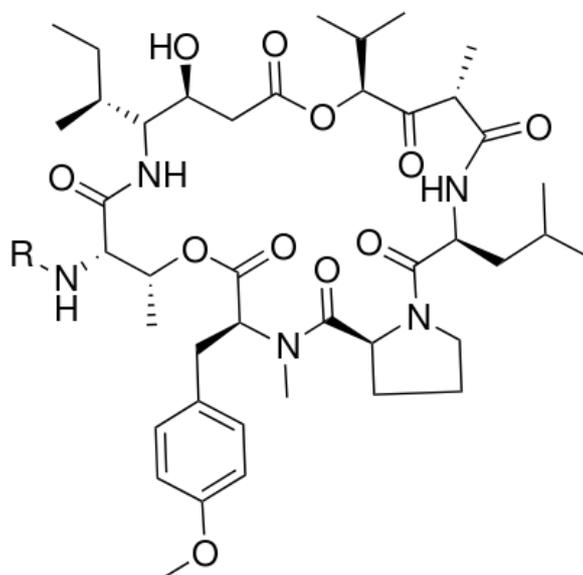
Figure 17. Didemins general structure. Didemin B corresponds to R=Lac-Pro-N-Me-L-Leu.

Plitidepsin is a depsipeptide that carries a β-hydroxy-γ-amino acid, another example of a non-canonical residue. This peptide presents potent activity against antimyeloma in vitro and in vivo.[153]
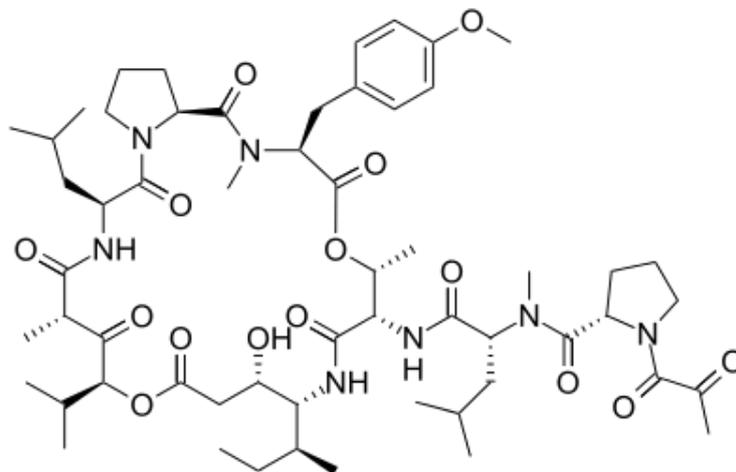
Figure 18. Structure of the depsipeptide Plitidepsin (Aplidin).

2.2.1. Retro-Inverso Peptidomimetics

Retro-inverso peptides are generated when the amino acid sequence is reversed, i.e. reverse amide peptide bonds, and the α-center chirality of the amino acid subunits is inverted as well, for D-amino acids. The use of these peptides is another approach to design peptidomimetics more resistant to proteolytic degradation, but not always increase the pharmacological potency.[4, 154]
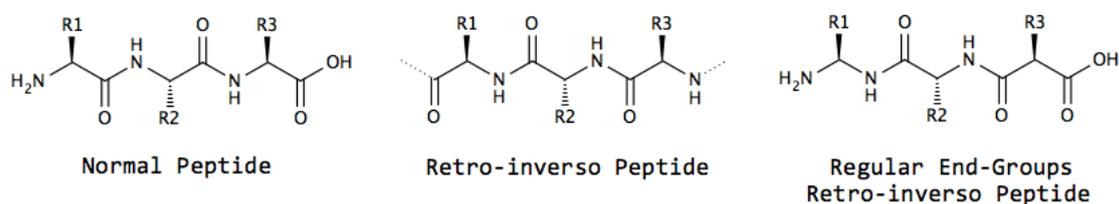
Figure 19. Structures that show the difference among a normal peptide, a retro-inverso peptide and a retro-inverso peptide with regular end-groups.

The retro-inverso peptides with regular terminal groups are more able to link to native peptides, generating potent peptidomimetics, or to be embedded in a large normal peptide to achieve the same goal. One example of this is the peptide Tuftsin, which in its normal state is completely degraded *in vivo* in about 8 minutes. However, when in retro-inverso peptide form, only 2% of hydrolysis is observed after 50 minutes, and with the retention of bioactivity.[155]

2.2.2. Sugar Amino Acids

Sugar amino acids (SAAs) are building blocks with applicability as drugs or in peptide design.[156-157] Chemically they are monosaccharide derivatives containing an amine and a carboxylic acid group. Risseeuw, Fleet and co-workers have published two important compendiums on this field.[158-159] Recently, the synthesis and design of peptidomimetics carrying SAAs was explored by Tian and co-workers.[160] Other reviews and papers show the use of SAAs scaffolds in drug design and peptidomimetics.[157, 161]
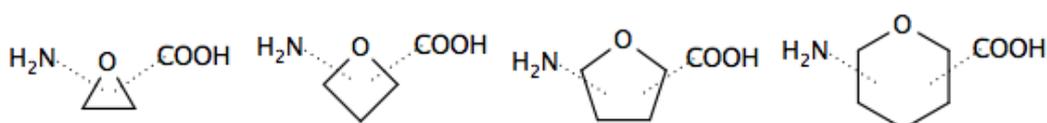
Figure 20. Two-dimensional structures of the types of SAAs according to Fleet's compendium.

3. Conclusions

This review focused on four important topics: the difference between encoded and non-proteinogenic amino acids, the relation between peptide secondary structure and biological function, the most relevant non-canonical amino acid classes and the most common peptide backbone structure modifications. Above all, the non-canonical amino acids were emphasized in more detail because it is our main field of research.

In Table 1, the conformational preferences of the non-canonical amino acids that stand out within their class are summarized.

Table 1. Conformational preferences of the non-canonical amino acids adressed in this study.

Non-canonical Amino Acid Class	Highlights	Conformational Preferences
Symmetric α,α-dialkyl glycines	Aib	3_{10}-helix or α-helix
	Dhg	α-helix
	Iva	3_{10}-helix or α-helix
Asymmetric α,α-dialkyl glycines (D-amino acids)	MDL	3_{10}-helix or α-helix
	MDP	3_{10}-helix or α-helix
	Ac_3c	bridge region ($\varphi,\psi \approx \pm 90^\circ$, 0°)
Cα to Cα cyclized amino acids - Ac_nc	Ac_6c	3_{10}-helix or α-helix
	(R,R) Ac_5c^{dOM}	3_{10}-helix or α-helix
Proline Analogues	Hyp	β-turn, bend
α,β- dehydroamino acids	Δ^zPhe	β-turn or γ-turn

The incorporation of non-canonical amino acids into known peptides proves that it is possible to optimize the characteristics of native peptides and obtain novel molecules with improved activity and stability. The selected examples mentioned here illustrate different ways to generate peptidomimetics, but in general, this can be done by incorporating non-canonical amino acids or by changes in the backbone. Also, we have reported how the application of these strategies successfully generates active peptidomimetics.

We believe that in the future peptidomimetics will manifest a large variety of applications in medicinal chemistry, biotechnology and nanotechnology fields. The rational design of these molecules will produce new bio devices, biosensors and other biomaterials capable to perform specific interactions with physiological environments.

References

(1)   Antosova, Z.; Mackova, M.; Kral, V.; Macek, T., Therapeutic application of peptides and proteins: parenteral forever? *Trends in Biotechnology* **2009,** *27*, 628-635.

(2)   Avan, I.; Hall, C. D.; Katritzky, A. R., Peptidomimetics via modifications of amino acids and peptide bonds. *Chemical Society Reviews* **2014,** *43*, 3575-3594.

(3)   Gentilucci, L.; Tolomelli, A.; Squassabia, F., Peptides and peptidomimetics in medicine, surgery and biotechnology. *Curr Med Chem* **2006,** *13*, 2449-66.

(4)   Grauer, A.; König, B., Peptidomimetics – A Versatile Route to Biologically Active Compounds. *European Journal of Organic Chemistry* **2009,** *2009*, 5099-5111.

(5)   Kieber-Emmons, T.; Murali, R.; Greene, M. I., Therapeutic peptides and peptidomimetics. *Current Opinion in Biotechnology* **1997,** *8*, 435-441.

(6)   Lien, S.; Lowman, H. B., Therapeutic peptides. *Trends in Biotechnology* **2003,** *21*, 556-562.

(7)   Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M., Synthetic therapeutic peptides: science and market. *Drug Discovery Today* **2010,** *15*, 40-56.

(8)   Bocci, V., Catabolism of therapeutic proteins and peptides with implications for drug delivery. *Advanced Drug Delivery Reviews* **1989,** *4*, 149-169.

(9)   Gentilucci, L.; De Marco, R.; Cerisoli, L., Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr Pharm Des* **2010,** *16*, 3185-203.

(10)   Lee, V. H. L.; Yamamoto, A., Penetration and enzymatic barriers to peptide and protein absorption. *Advanced Drug Delivery Reviews* **1989,** *4*, 171-207.

(11)   Ripka, A. S.; Rich, D. H., Peptidomimetic design. *Curr Opin Chem Biol* **1998,** *2*, 441-52.

(12)   Ruzza, P., Peptides and Peptidomimetics in Medicinal Chemistry. In *Medicinal Chemistry and Drug Design*, Ekinci, D., Ed. InTech: 2012.

(13)   Vagner, J.; Qu, H.; Hruby, V. J., Peptidomimetics, a synthetic tool of drug discovery. *Current Opinion in Chemical Biology* **2008,** *12*, 292-296.

(14)   Connor, R. E.; Tirrell, D. A., Non‐Canonical Amino Acids in Protein Polymer Design. *Polymer Reviews* **2007,** *47*, 9-28.

(15)   De Zotti, M.; Formaggio, F.; Crisma, M.; Peggion, C.; Moretto, A.; Toniolo, C., Handedness preference and switching of peptide helices. Part I: Helices based on protein amino acids. *J Pept Sci* **2014,** *20*, 307-22.

(16)   Odar, C.; Winkler, M.; Wiltschi, B., Fluoro amino acids: A rarity in nature, yet a prospect for protein engineering. *Biotechnology Journal* **2015,** *10*, 427-446.

(17)   A Brief Guide to the Twenty Common Amino Acids. http://www.compoundchem.com/2014/09/16/aminoacids/ (accessed 11/10/2015).

(18)   Weber, A.; Miller, S., Reasons for the occurrence of the twenty coded protein amino acids. *Journal of Molecular Evolution* **1981,** *17*, 273-284.

(19)   Blumenkrantz, N.; Ullman, S.; Asboe-Hansen, G., Parallel studies on collagen hydroxyproline and hydroxylysine in human skin biopsies. *Acta Derm Venereol* **1976,** *56*, 93-8.

(20)   Gorres, K. L.; Raines, R. T., Prolyl 4-hydroxylase. *Critical Reviews in Biochemistry and Molecular Biology* **2010,** *45*, 106-124.

(21)   Szpak, P., Fish bone chemistry and ultrastructure: implications for taphonomy and stable isotope analysis. *Journal of Archaeological Science* **2011,** *38*, 3358-3372.

(22)   Zhu, X.; Tang, G.; Galili, G., The catabolic function of the alpha-aminoadipic acid pathway in plants is associated with unidirectional activity of lysine-oxoglutarate reductase, but not saccharopine dehydrogenase. *Biochemical Journal* **2000,** *351*, 215-220.

(23)   Peisach, J.; Blumberg, W. E., A mechanism for the action of penicillamine in the treatment of Wilson's disease. *Mol Pharmacol* **1969,** *5*, 200-9.

(24)   Sugino, T.; Shirai, T.; Kajimoto, Y.; Kajimoto, O., L-ornithine supplementation attenuates physical fatigue in healthy volunteers by modulating lipid and amino acid metabolism. *Nutr Res* **2008,** *28*, 738-43.

(25)   Fearon, W. R., The carbamido diacetyl reaction: a test for citrulline. *Biochemical Journal* **1939,** *33*, 902-907.

(26)   Johansson, L.; Gafvelin, G.; Arnér, E. S. J., Selenocysteine in proteins—properties and biotechnological use. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2005,** *1726*, 1-13.

(27)   Srinivasan, G.; James, C. M.; Krzycki, J. A., Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA. *Science* **2002,** *296*, 1459-1462.

(28)   Rother, M.; Krzycki, J. A., Selenocysteine, Pyrrolysine, and the Unique Energy Metabolism of Methanogenic Archaea. *Archaea* **2010,** *2010*, 14.

(29)   Schmidt, R. L.; Simonović, M., Synthesis and decoding of selenocysteine and human health. *Croatian Medical Journal* **2012,** *53*, 535-550.

(30)   Baranov, P. V.; Gesteland, R. F.; Atkins, J. F., Recoding: translational bifurcations in gene expression. *Gene* **2002,** *286*, 187-201.

(31)   Jones, J. H., Words derived from the noun peptide. *Journal of Peptide Science* **2006,** *12*, 79-81.

(32)   Nomenclature and Symbolism for Amino Acids and Peptides. *European Journal of Biochemistry* **1984,** *138*, 9-37.

(33)   Portmann, C.; Blom, J. F.; Gademann, K.; Jüttner, F., Aerucyclamides A and B: Isolation and Synthesis of Toxic Ribosomal Heterocyclic Peptides from the Cyanobacterium Microcystis aeruginosa PCC 7806. *Journal of Natural Products* **2008,** *71*, 1193-1196.

(34)   Dawson, R. M., the toxicology of microcystins. *Toxicon* **1998,** *36*, 953-962.

(35)   Agha, R.; Cirés, S.; Wörmer, L.; Quesada, A., Limited Stability of Microcystins in Oligopeptide Compositions of Microcystis aeruginosa (Cyanobacteria): Implications in the Definition of Chemotypes. *Toxins* **2013,** *5*, 1089.

(36)   ALBERT, L. L.; DAVID, L. N.; COX, M. M., *Princípios de Bioquímica*. 4 ed.; Sarvier: 2007.

(37)   DAVID, L. N.; COX, M. M., *Princípios de Bioquímica de Lehninger*. 5 ed.; Artmed: 2011.

(38)   Lee, H. J.; Macbeth, A. H.; Pagani, J. H.; Young, W. S., 3rd, Oxytocin: the great facilitator of life. *Prog Neurobiol* **2009,** *88*, 127-51.

(39)   Urayama, A.; Yamada, S.; Kimura, R.; Zhang, J.; Watanabe, Y., Neuroprotective effect and brain receptor binding of taltirelin, a novel thyrotropin-releasing hormone (TRH) analogue, in transient forebrain ischemia of C57BL/6J mice. *Life Sciences* **2002,** *72*, 601-607.

(40)   Marcotte, I.; Separovic, F.; Auger, M.; Gagné, S. M., A Multidimensional 1H NMR Investigation of the Conformation of Methionine-Enkephalin in Fast-Tumbling Bicelles. *Biophysical Journal* **2004,** *86*, 1587-1600.

(41)   Hallen, H. E.; Luo, H.; Scott-Craig, J. S.; Walton, J. D., Gene family encoding the major toxins of lethal Amanita mushrooms. *Proceedings of the National Academy of Sciences* **2007,** *104*, 19097-19101.

(42)   In *Pure and Applied Chemistry*, 1984; Vol. 56, p 595.

(43)   Voet, D.; Pratt, C. W.; Voet, J. G., *Principles of Biochemistry*. 4th Edition International Student Version ed.; John Wiley & Son Inc: 2012.

(44)   Engh, R. A.; Huber, R., Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallographica Section A* **1991,** *47*, 392-400.

(45)   Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A., Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *The Journal of Physical Chemistry* **1975,** *79*, 2361-2381.

(46)   Touw, W. G.; Vriend, G., On the complexity of Engh and Huber refinement restraints: the angle τ as example. *Acta Crystallographica Section D: Biological Crystallography* **2010,** *66*, 1341-1350.

(47)   Karplus, P. A., Experimentally observed conformation-dependent geometry and hidden strain in proteins. *Protein Science : A Publication of the Protein Society* **1996,** *5*, 1406-1420.

(48)   Richardson, J. S., The anatomy and taxonomy of protein structure. *Adv Protein Chem* **1981,** *34*, 167-339.

(49)   Baures, P. W.; Ojala, W. H.; Gleason, W. B.; Johnson, R. L., Conformational analysis of homochiral and heterochiral diprolines as β-turn-forming peptidomimetics: unsubstituted and substituted models. *J. Pept. Res.* **1997,** *50*, 1-13.

(50)   Hovmoller, S.; Zhou, T.; Ohlson, T., Conformations of amino acids in proteins. *Acta Crystallographica Section D* **2002,** *58*, 768-776.

(51)   Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V., Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **1963,** *7*, 95-99.

(52)   Jonsson, P.-G.; Kvick, A., Precision neutron diffraction structure determination of protein and nucleic acid components. III. The crystal and molecular structure of the amino acid [alpha]-glycine. *Acta Crystallographica Section B* **1972,** *28*, 1827-1833.

(53)   Price, W. D.; Jockusch, R. A.; Williams, E. R., Is Arginine a Zwitterion in the Gas Phase? *Journal of the American Chemical Society* **1997,** *119*, 11988-11989.

(54)   Wadhwani, P.; Epand, R. F.; Heidenreich, N.; Bürck, J.; Ulrich, A. S.; Epand, R. M., Membrane-Active Peptides and the Clustering of Anionic Lipids. *Biophysical Journal* **2012,** *103*, 265-274.

(55)   Wang, Q.; Hong, G.; Johnson, G. R.; Pachter, R.; Cheung, M. S., Biophysical Properties of Membrane-Active Peptides Based on Micelle Modeling: A Case Study of Cell-Penetrating and Antimicrobial Peptides. *The Journal of Physical Chemistry B* **2010,** *114*, 13726-13735.

(56)   Garibotto, F. M.; Garro, A. D.; Rodríguez, A. M.; Raimondi, M.; Zacchino, S. A.; Perczel, A.; Somlai, C.; Penke, B.; Enriz, R. D., Penetratin analogues acting as antifungal agents. *European Journal of Medicinal Chemistry* **2011,** *46*, 370-377.

(57)   Bahnsen, J. S.; Franzyk, H.; Sandberg-Schaal, A.; Nielsen, H. M., Antimicrobial and cell-penetrating properties of penetratin analogs: Effect of sequence and secondary structure. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2013,** *1828*, 223-232.

(58)   Wade, D.; Silberring, J.; Soliymani, R.; Heikkinen, S.; Kilpeläinen, I.; Lankinen, H.; Kuusela, P., Antibacterial activities of temporin A analogs. *FEBS Letters* **2000,** *479*, 6-9.

(59)   Castro, T. G.; Micaêlo, N. M., Modeling of Peptaibol Analogues Incorporating Nonpolar α, α-Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation. *J. Phys. Chem. B* **2014,** *118*, 649-658.

(60)   Castro, T. G.; Micaelo, N. M., Conformational and Thermodynamic Properties of Non-Canonical alpha,alpha-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies. *J. Phys. Chem. B* **2014,** *118*, 9861-9870.

(61)   Vilaca, H.; Pereira, G.; Castro, T. G.; Hermenegildo, B. F.; Shi, J.; Faria, T. Q.; Micaelo, N.; Brito, R. M. M.; Xu, B.; Castanheira, E. M. S.; Martins, J. A.; Ferreira, P. M. T., New self-assembled supramolecular hydrogels based on dehydropeptides. *Journal of Materials Chemistry B* **2015,** *3*, 6355-6367.

(62)   Ager, D. J.; Fotheringham, I. G., Methods for the synthesis of unnatural amino acids. *Current opinion in drug discovery & development* **2001,** *4*, 800-7.

(63)   Giannis, A.; Rübsam, F., Peptidomimetics in Drug Design. In *Advances in Drug Research*, Bernard, T.; Urs, A. M., Eds. Academic Press: 1997; Vol. Volume 29, pp 1-78.

(64)   Hruby, V. J.; al-Obeidi, F.; Kazmierski, W., Emerging approaches in the molecular design of receptor-selective peptide ligands: conformational, topographical and dynamic considerations. *Biochemical Journal* **1990,** *268*, 249-262.

(65)   Chugh, J. K.; Wallace, B. A., Peptaibols: models for ion channels. *Biochemical Society Transactions* **2001,** *29*, 565-570.

(66)   Degenkolb, T.; Brückner, H., Peptaibiomics: Towards a Myriad of Bioactive Peptides Containing C α -Dialkylamino Acids? *Chemistry & Biodiversity* **2008,** *5*, 1817-1843.

(67)   Duclohier, H., Peptaibiotics and Peptaibols: An Alternative to Classical Antibiotics? *Chemistry & Biodiversity* **2007,** *4*, 1023-1026.

(68)   Karle, I. L.; Balaram, P., Structural characteristics of .alpha.-helical peptide molecules containing Aib residues. *Biochemistry* **1990,** *29*, 6747-6756.

(69)   Mendel, D.; Ellman, J.; Schultz, P. G., Protein-Biosynthesis with Conformationally Restricted Amino-Acids *J. Am. Chem. Soc.* **1993,** *115*, 4359-4360.

(70)   Toniolo, C.; Crisma, M.; Formaggio, F.; Valle, C.; Cavicchioni, G.; Precigoux, G.; Aubry, A.; Kamphuis, J., Structures of peptides from α -amino acids methylated at the α -carbon. *Biopolymers* **1993,** *33*, 1061-1072.

(71)   Sagan, S.; Karoyan, P.; Lequin, O.; Chassaing, G.; Lavielle, S., N- and Calpha-methylation in biologically active peptides: synthesis, structural and functional aspects. *Curr Med Chem* **2004,** *11*, 2799-822.

(72)   Doi, M.; Asano, A.; Komura, E.; Ueda, Y., The structure of an endomorphin analogue incorporating 1-aminocyclohexane-1-carboxlylic acid for proline is similar to the β -turn of Leu-enkephalin. *Biochemical and Biophysical Research Communications* **2002,** *297*, 138-142.

(73)   Toniolo, C.; Crisma, M.; Valle, G.; Bonora, G. M.; Polinelli, S.; Becker, E. L.; Freer, R. J.; Sudhanand; Rao, R. B.; Balaram, P.; et al., Conformationally restricted formyl methionyl tripeptide chemoattractants: a three-dimensional structure-activity study of analogs incorporating a C alpha,alpha-dialkylated glycine at position 2. *Pept Res* **1989,** *2*, 275-81.

(74)   Chugh, J. K.; Wallace, B. A., Peptaibols: Models for Ion Channels. *Biochem. Soc. Trans.* **2001,** *29*, 565-570.

(75)   Shenkarev, Z. O.; Balashova, T. A.; Efremov, R. G.; Yakimenko, Z. A.; Ovchinnikova, T. V.; Raap, J.; Arseniev, A. S., Spatial Structure of Zervamicin IIB Bound to DPC Micelles: Implications for Voltage-Gating. *Biophys. J.* **2002,** *82*, 762-771.

(76)   Snook, C. F.; Woolley, G. A.; Oliva, G.; Pattabhi, V.; Wood, S. P.; Blundell, T. L.; Wallace, B. A., The structure and function of antiamoebin I, a proline-rich membrane-active polypeptide. *Structure* **1998,** *6*, 783-792.

(77)   Roos, E. C.; Lopez, M. C.; Brook, M. A.; Hiemstra, H.; Speckamp, W. N.; Kaptein, B.; Kamphuis, J.; Schoemaker, H. E., Synthesis of .alpha.-substituted .alpha.-amino acids via cationic intermediates. *J. Org. Chem.* **1993,** *58*, 3259-3268.

(78)   Mendel, D.; Ellman, J.; Schultz, P. G., Protein biosynthesis with conformationally restricted amino acids. *J. Am. Chem. Soc.* **1993,** *115*, 4359-4360.

(79)   Benedetti, E.; Di Blasio, B.; Iacovino, R.; Menchise, V.; Saviano, M.; Pedone, C.; Maria Bonora, G.; Ettorre, A.; Graci, L.; Formaggio, F.; Crisma, M.; Valle, G.; Toniolo, C., Conformational restriction through Calpha i - Calpha i cyclization: 1-aminocycloheptane-1-carboxylic acid (Ac7c). *J. Chem. Soc., Perkin Trans. 2* **1997**, 2023-2032.

(80)   Toniolo, C., Conformationally Restricted Peptides Through Short-Range Cyclizations. *International Journal of Peptide and Protein Research* **1990,** *35*, 287-300.

(81)   Aschi, M.; Lucente, G.; Mazza, F.; Mollica, A.; Morera, E.; Nalli, M.; Paradisi, M. P., Peptide backbone folding induced by the C-alpha-tetrasubstituted cyclic alpha-amino acids 4-amino-1,2-dithiolane-4-carboxylic acid (Adt) and 1-aminocyclopentane-1-carboxylic acid (Ac(5)c). A joint computational and experimental study. *Org. Biomol. Chem.* **2003,** *1*, 1980-1988.

(82)   Bardi, R.; Piazzesi, A. M.; Toniolo, C.; Sukumar, M.; Balaram, P., Stereochemistry of peptides containing 1-aminocyclopentanecarboxylic acid (Acc5): Solution and solid-state conformations of Boc-Acc5-Acc5-NHMe. *Biopolymers* **1986,** *25*, 1635-1644.

(83)   Barone, V.; Fraternali, F.; Cristinziano, P. L.; Lelj, F.; Rosa, A., Conformational behavior of α, α-dialkylated peptides: Ab initio and empirical results for cyclopropylglycine. *Biopolymers* **1988,** *27*, 1673-1685.

(84)   Crisma, M.; Bonora, G. M.; Toniolo, C.; Barone, V.; Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Santini, A.; Fraternali, F.; Bavoso, A.; Lelj, F., Structural versatility of peptides containing Cα, α-dialkylated glycines: conformational energy computations, i.r. absorption and 1H n.m.r. analysis of 1-aminocyclopropane-1-carboxylic acid homopeptides. *Int. J. Biol. Macromolec.* **1989,** *11*, 345-352.

(85)   Demizu, Y.; Doi, M.; Kurihara, M.; Okuda, H.; Nagano, M.; Suemune, H.; Tanaka, M., Conformational studies on peptides containing alpha,alpha-disubstituted alpha-amino acids: chiral cyclic alpha,alpha-disubstituted alpha-amino acid as an alpha-helical inducer. *Org. Biomol. Chem.* **2011,** *9*, 3303-3312.

(86)   Di Blasio, B.; Lombardi, A.; Nastri, F.; Saviano, M.; Pedone, C.; Yamada, T.; Nakao, M.; Kuwata, S.; Pavone, V., Conformation of diastereomeric peptide sequences: Structural analysis of Z-D-Val-Ac6c-Gly-L-Phe-OMe. *Biopolymers* **1992,** *32*, 1155-1161.

(87)   Gatos, M.; Formaggio, F.; Crisma, M.; Toniolo, C.; Bonora, G. M.; Benedetti, Z.; Di Blasio, B.; Iacovino, R.; Santini, A.; Saviano, M.; Kamphuis, J., Conformational Characterization of the 1-Aminocyclobutane-1-carboxylic Acid Residue in Model Peptides. *J. Pept. Sci.* **1997,** *3*, 110-122.

(88)   Ballano, G.; Zanuy, D.; Jiménez, A. I.; Cativiela, C.; Nussinov, R.; Alemán, C., Structural Analysis of a β-Helical Protein Motif Stabilized by Targeted Replacements with Conformationally Constrained Amino Acids. *J. Phys. Chem. B* **2008,** *112*, 13101-13115.

(89)   Gatos, M.; Formaggio, F.; Crisma, M.; Valle, G.; Toniolo, C.; Bonora, G. M.; Saviano, M.; Iacovino, R.; Menchise, V.; Galdiero, S.; Pedone, C.; Benedetti, E., Conformational characterization of peptides rich in the cycloaliphatic Cα, α-disubstituted glycine 1-amino-cyclononane-1-carboxylic acid. *J. Pept. Sci.* **1997,** *3*, 367-382.

(90)   Moretto, A.; Formaggio, F.; Crisma, M.; Toniolo, C.; Saviano, M.; Benedetti, E.; Iacovino, R.; Vitale, R. M., Ac10c: a medium-ring, cycloaliphatic Cα, α-disubstituted glycine. Incorporation into model peptides and preferred conformation. *J. Pept. Res.* **2001,** *57*, 307-315.

(91)   Santini, A.; Di, B. B.; Galdiero, S.; R., I.; Pedone, C.; Benedetti, E.; Crisma, M.; Toniolo, C., Molecular and crystal structures of two terminally-blocked Ac5c homo-oligopeptides. In *Zeitschrift für Kristallographie*, 1996; Vol. 211, p 616.

(92)   Saviano, M.; Iacovino, R.; Benedetti, E.; Moretto, V.; Banzato, A.; Formaggio, F.; Crisma, M.; Toniolo, C., Preferred conformation of peptides based on cycloaliphatic C α, α-disubstituted glycines: 1-amino-cycloundecane-1-carboxylic acid (Ac11c). *J. Pept. Sci.* **2000,** *6*, 571-583.

(93)   Saviano, M.; Iacovino, R.; Menchise, V.; Benedetti, E.; Bonora, G. M.; Gatos, M.; Graci, L.; Formaggio, F.; Crisma, M.; Toniolo, C., Conformational restriction through C αi $\leftrightarrow$ C αi cyclization: Ac12c, the largest cycloaliphatic C α, α-disubstituted glycine known. *Biopolymers* **2000,** *53*, 200-212.

(94)   Paradisi, M. P.; Torrini, I.; Zecchini, G. P.; Lucente, G.; Gavuzzo, E.; Mazza, F.; Pochetti, G., Gamma-turn conformation induced by alpha,alpha-disubstituted amino acids with a cyclic 6-membered side-chain. *Tetrahedron* **1995,** *51*, 2379-2386.

(95)   Zimmerman, S. S.; Pottle, M. S.; Némethy, G.; Scheraga, H. A., Conformational Analysis of the 20 Naturally Occurring Amino Acid Residues Using ECEPP. *Macromolecules* **1977,** *10*, 1-9.

(96)   Rodriguez-Ropero, F.; Zanuy, D.; Casanovas, J.; Nussinov, R.; Aleman, C., Application of 1-aminocyclohexane carboxylic acid to protein nanostructure computer design. *J. Chem Inf. Model.* **2008,** *48*, 333-343.

(97)   Alemán, C., Conformational Properties of α-Amino Acids Disubstituted at the α-Carbon. *J. Phys. Chem. B* **1997,** *101*, 5046-5050.

(98)   Zanuy, D.; Ballano, G.; Jimenez, A. I.; Casanovas, J.; Haspel, N.; Cativiela, C.; Curco, D.; Nussinov, R.; Aleman, C., Protein Segments with Conformationally Restricted Amino Acids Can Control Supramolecular Organization at the Nanoscale. *J. Chem Inf. Model.* **2009,** *49*, 1623-1629.

(99)   Headley, A. D.; Ganesan, R.; Nam, J., The effect of the cyclopropyl group on the conformation of chemotactic formyl tripeptides. *Bioorg. Chem.* **2003,** *31*, 99-108.

(100)   Jiménez, A. I.; Vaquero, V.; Cabezas, C.; López, J. C.; Cativiela, C.; Alonso, J. L., The Singular Gas-Phase Structure of 1-Aminocyclopropanecarboxylic Acid (Ac3c). *J. Am. Chem. Soc.* **2011,** *133*, 10621-10628.

(101)   Gomez-Catalan, J.; Aleman, C.; Perez, J. J., Conformational profile of 1-aminocyclopropanecarboxylic acid. *Theor. Chem. Acc.* **2000,** *103*, 380-389.

(102)   Caumes, C.; Delsuc, N.; Azza, R. B.; Correia, I.; Chemla, F.; Ferreira, F.; Carlier, L.; Luna, A. P.; Moumne, R.; Lequin, O.; Karoyan, P., Homooligomers of substituted prolines and [small beta]-prolines: syntheses and secondary structure investigation. *New J. Chem.* **2013,** *37*, 1312-1319.

(103)   Bach, T.; Takagi, H., Properties, metabolisms, and applications of l-proline analogues. *Appl. Microbiol. Biotechnol.* **2013,** *97*, 6623-6634.

(104)   Tsogoeva, S. B.; Jagtap, S. B.; Ardemasova, Z. A., 4-trans-Amino-proline based di- and tetrapeptides as organic catalysts for asymmetric C–C bond formation reactions. *Tetrahedron: Asymmetry* **2006,** *17*, 989-992.

(105)   Torino, D.; Mollica, A.; Pinnen, F.; Feliciani, F.; Spisani, S.; Lucente, G., Novel chemotactic For-Met-Leu-Phe-OMe (fMLF-OMe) analogues based on Met residue replacement by 4-amino-proline scaffold: Synthesis and bioactivity. *Bioorg. Med. Chem.* **2009,** *17*, 251-259.

(106)   Seebach, D.; Beck, A. K.; Bierbaum, D. J., The World of β- and γ-Peptides Comprised of Homologated Proteinogenic Amino Acids and Other Components. *Chemistry & Biodiversity* **2004,** *1*, 1111-1239.

(107)   Haskell-Luevano, C.; Toth, K.; Boteju, L.; Job, C.; Castrucci, A. M. d. L.; Hadley, M. E.; Hruby, V. J., β-Methylation of the Phe7 and Trp9 Melanotropin Side Chain Pharmacophores Affects Ligand–Receptor Interactions and Prolonged Biological Activity. *Journal of Medicinal Chemistry* **1997,** *40*, 2740-2749.

(108)   Jiao, D.; Russell, K. C.; Hruby, V. J., Locally constrained tyrosine analogues with restricted side chain dynamics. *Tetrahedron* **1993,** *49*, 3511-3520.

(109)   Kover, K. E.; Jiao, D.; Fang, S.; Hruby, V. J., Conformational Properties of the Unnatural Amino Acid .beta.-Methylphenylalanine in a Linear Octapeptide System; Correlations of 13C-NMR Chemical Shifts with the Side-Chain Stereochemistry of These Amino Acid Residues. *The Journal of Organic Chemistry* **1994,** *59*, 991-998.

(110)   Mosberg, H. I.; Omnaas, J. R.; Lomize, A.; Heyl, D. L.; Nordan, I.; Mousigian, C.; Davis, P.; Porreca, F., Development of a model for the delta opioid receptor pharmacophore. 2. Conformationally restricted Phe3 replacements in the cyclic delta receptor selective tetrapeptide Tyr-c[D-Cys-Phe-D-Pen]OH (JOM-13). *J Med Chem* **1994,** *37*, 4384-91.

(111)   Singh, T. P.; Kaur, P., Conformation and design of peptides with α, β-dehydro-amino acid residues. *Progress in Biophysics and Molecular Biology* **1996,** *66*, 141-165.

(112)   Singh, T. P.; Narula, P.; Patel, H. C., [alpha],[beta]-Dehydro residues in the design of peptide and protein structures. *Acta Crystallographica Section B* **1990,** *46*, 539-545.

(113)   Siodłak, D., α, β-Dehydroamino acids in naturally occurring peptides. *Amino Acids* **2015,** *47*, 1-17.

(114)   Busetti, V.; Crisma, M.; Toniolo, C.; Salvadori, S.; Balboni, G., α, β-Dehydro-amino acid residues in the design of peptide structures. Molecular and crystal structures of two folded dehydro peptides. *International Journal of Biological Macromolecules* **1992,** *14*, 23-28.

(115)   Ciszak, E. W. A.; PietrzyÑSki, G.; Rzeszotarska, B., Conformational investigation of α, β-dehydropeptides. *International Journal of Peptide and Protein Research* **1992,** *39*, 218-222.

(116)   English, M. L.; Stammer, C. H., Enzyme Stability of Dehydropeptides. *Biochemical and Biophysical Research Communications* **1978,** *83*, 1464-1467.

(117)   Pieroni, O.; Fissi, A.; Jain, R. M.; Chauhan, V. S., Solution structure of dehydropeptides: A CD investigation. *Biopolymers* **1996,** *38*, 97-108.

(118)   Scaloni, A.; Barra, D.; Bossa, F., Sequence Analysis of Dehydroamino Acid Containing Peptides. *Analytical Biochemistry* **1994,** *218*, 226-228.

(119)   Buczek, A.; Siodłak, D.; Bujak, M.; Broda, M. A., Effects of Side-Chain Orientation on the Backbone Conformation of the Dehydrophenylalanine Residue. Theoretical and X-ray Study. *The Journal of Physical Chemistry B* **2011,** *115*, 4295-4306.

(120)   Desai, P.; Coutinho, E., Effect of Stereochemistry (Z and E) and Position of α, β-Dehydrophenylalanine (ΔPhe) on β-turn Stability. *Molecular modeling annual* **2000,** *6*, 595-599.

(121)   Dey, S.; Mitra, S. N.; Singh, T. P., Design of peptides using α, β-dehydro-residues: Synthesis, crystal structure and molecular conformation of Boc-L-Val-ΔPhe-ΔPhe-L-Val-OCH3. *Biopolymers* **1996,** *39*, 849-857.

(122)   Gross, E.; Morell, J. L., Presence of dehydroalanine in the antibiotic nisin and its relation to activity. *Journal of the American Chemical Society* **1967,** *89*, 2791-2792.

(123)   Nandel, F. S.; Malik, N.; Singh, B.; Jain, D. V. S., Conformational structure of peptides containing dehydroalanine: Formation of β-bend ribbon structure. *International Journal of Quantum Chemistry* **1999,** *72*, 15-23.

(124)   Nandel, F. S.; Sahrawat, T. R., Conformational study of poly- Δ Abu peptides and construction of amphipathic nanostructure. *Peptide Science* **2009,** *92*, 44-51.

(125)   Seebeck, F. P.; Szostak, J. W., Ribosomal Synthesis of Dehydroalanine-Containing Peptides. *Journal of the American Chemical Society* **2006,** *128*, 7150-7151.

(126)   Siodłak, D.; Grondys, J.; Lis, T.; Bujak, M.; Broda, M. A.; Rzeszotarska, B., The conformational properties of dehydrobutyrine and dehydrovaline: theoretical and solid-state conformational studies. *Journal of Peptide Science* **2010,** *16*, 496-505.

(127)   Rajashankar, K. R.; Ramakumar, S.; Chauhan, V. S., Design of a helical motif using .alpha.,.beta.-dehydrophenylalanine residues: crystal structure of Boc-Val-.DELTA.Phe-Phe-Ala-Phe-.DELTA.Phe-Val-.DELTA.PHe-Gly-OCH3, a 310-helical nonapeptide. *Journal of the American Chemical Society* **1992,** *114*, 9225-9226.

(128)   Formaggio, F.; Crisma, M.; Toniolo, C.; Tchertanov, L.; Guilhem, J.; Mazaleyrat, J.-P.; Gaucher, A.; Wakselman, M., Bip: a C α -Tetrasubstituted, Axially Chiral α -Amino Acid. Synthesis and Conformational Preference of Model Peptides. *Tetrahedron* **2000,** *56*, 8721-8734.

(129)   Mazaleyrat, J.-P.; Gaucher, A.; Šavrda, J.; Wakselman, M., Novel α , α -disubstituted α -aminoacids with axial dissymmetry and their N- or C-protected derivatives. *Tetrahedron: Asymmetry* **1997,** *8*, 619-631.

(130)   Mazaleyrat, J.-P.; Gaucher, A.; Wakselman, M.; Tchertanov, L.; Guilhem, J., A new chiral α -aminoacid with only axial dissymmetry: Synthesis and X-ray analysis of a 1,1′ -binaphthyl-substituted α -aminoisobutyric acid (Bin) and of its biphenyl analogue (Bip). *Tetrahedron Letters* **1996,** *37*, 2971-2974.

(131)   Mazaleyrat, J.-P.; Wright, K.; Gaucher, A.; Toulemonde, N.; Wakselman, M.; Oancea, S.; Peggion, C.; Formaggio, F.; Setnička, V.; Keiderling, T. A.; Toniolo, C., Induced Axial Chirality in the Biphenyl Core of the C α -Tetrasubstituted α -Amino Acid Residue Bip and Subsequent Propagation of Chirality in (Bip)n/Val Oligopeptides. *Journal of the American Chemical Society* **2004,** *126*, 12874-12879.

(132)   Mazaleyrat, J.-P.; Wright, K.; Wakselman, M.; Formaggio, F.; Crisma, M.; Toniolo, C., 9-Amino-4,5-diazafluorene-9-carboxylic Acid (Daf), a New C α , α -Disubstituted Glycine Containing a Spatially Constrained Bipyridine-Like Ligand for Transition Metals – Synthesis and Evaluation of Peptide-Coupling Conditions at its C- and N-Termini. *European Journal of Organic Chemistry* **2001,** *2001*, 1821-1829.

(133)   Regan, L., The Design of Metal-Binding Sites in Proteins. *Annual Review of Biophysics and Biomolecular Structure* **1993,** *22*, 257-281.

(134)   He, H.; Janso, J. E.; Yang, H. Y.; Bernan, V. S.; Lin, S. L.; Yu, K., Culicinin D, an Antitumor Peptaibol Produced by the Fungus Culicinomyces clavisporus, Strain LL-12I252. *Journal of Natural Products* **2006,** *69*, 736-741.

(135)   Ko, K.-Y.; Wagner, S.; Yang, S.-H.; Furkert, D. P.; Brimble, M. A., Improved Synthesis of the Unnatural Amino Acids AHMOD and AMD, Components of the Anticancer Peptaibol Culicinin D. *The Journal of Organic Chemistry* **2015,** *80*, 8631-8636.

(136)   Lyu, P. C.; Sherman, J. C.; Chen, A.; Kallenbach, N. R., Alpha-helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proceedings of the National Academy of Sciences of the United States of America* **1991,** *88*, 5317-20.

(137)   Padmanabhan, S.; Baldwin, R. L., Straight-chain non-polar amino acids are good helix-formers in water. *Journal of Molecular Biology* **1991,** *219*, 135-137.

(138) Soini, J.; Falschlehner, C.; Liedert, C.; Bernhardt, J.; Vuoristo, J.; Neubauer, P., Norvaline is accumulated after a down-shift of oxygen in Escherichia coli W3110. *Microbial Cell Factories* **2008,** *7*, 30.

(139) Ming, X.-F.; Rajapakse, A.; Carvas, J.; Ruffieux, J.; Yang, Z., Inhibition of S6K1 accounts partially for the anti-inflammatory effects of the arginase inhibitor L-norvaline. *BMC Cardiovascular Disorders* **2009,** *9*, 12.

(140) Clementi, M. E.; Misiti, F., Substitution of methionine 35 inhibits apoptotic effects of Abeta(31-35) and Abeta(25-35) fragments of amyloid-beta protein in PC12 cells. *Medical science monitor : international medical journal of experimental and clinical research* **2005,** *11*, Br381-5.

(141) Andrews, M. J. I.; Tabor, A. B., Forming stable helical peptides using natural and artificial amino acids. *Tetrahedron* **1999,** *55*, 11711-11743.

(142) Cornish, V. W.; Kaplan, M. I.; Veenstra, D. L.; Kollman, P. A.; Schultz, P. G., Stabilizing and Destabilizing Effects of Placing .beta.-Branched Amino Acids in Protein .alpha.-Helixes. *Biochemistry* **1994,** *33*, 12022-12031.

(143) Jost, M.; Weigelt, S.; Huber, T.; Majer, Z.; Greie, J.-C.; Altendorf, K.; Sewald, N., Synthesis, and Structural and Biological Studies of Efrapeptin C Analogues. *Chemistry & Biodiversity* **2007,** *4*, 1170-1182.

(144) Wysong, C. L.; Yokum, T. S.; Morales, G. A.; Gundry, R. L.; McLaughlin, M. L.; Hammer, R. P., 4-Aminopiperidine-4-carboxylic Acid: A Cyclic α, α-Disubstituted Amino Acid for Preparation of Water-Soluble Highly Helical Peptides. *The Journal of Organic Chemistry* **1996,** *61*, 7650-7651.

(145) Toniolo, C., Structure of conformationally constrained peptides: From model compounds to bioactive peptides. *Biopolymers* **1989,** *28*, 247-257.

(146) Valle, G.; Crisma, M.; Toniolo, C.; Sudhanand; Rao, R. B.; Sukumar, M.; Balaram, P., Stereochemistry of peptides containing 1-aminocycloheptane-1-carboxylic acid (Ac7c). *International Journal of Peptide and Protein Research* **1991,** *38*, 511-518.

(147) Rogers, J. M.; Suga, H., Discovering functional, non-proteinogenic amino acid containing, peptides using genetic code reprogramming. *Organic & Biomolecular Chemistry* **2015,** *13*, 9353-9363.

(148) Cudic, P.; Stawikowski, M., Pseudopeptide Synthesis via Fmoc Solid-Phase Synthetic Methodology. *Mini-Reviews in Organic Chemistry* **2007,** *4*, 268-280.

(149) Magrath, J.; Abeles, R. H., Cysteine protease inhibition by azapeptide esters. *Journal of Medicinal Chemistry* **1992,** *35*, 4279-4283.

(150) Newman, D. J.; Cragg, G. M., Marine Natural Products and Related Compounds in Clinical and Advanced Preclinical Trials†. *Journal of Natural Products* **2004,** *67*, 1216-1238.

(151) Rinehart, K. L.; Kishore, V.; Bible, K. C.; Sakai, R.; Sullins, D. W.; Li, K.-M., Didemnins and Tunichlorin: Novel Natural Products from the Marine Tunicate Trididemnum solidum. *Journal of Natural Products* **1988,** *51*, 1-21.

(152) Shin, D. M.; Holoye, P. Y.; Forman, A.; Winn, R.; Perez-Soler, R.; Dakhil, S.; Rosenthal, J.; Raber, M. N.; Hong, W. K., Phase II clinical trial of didemnin B in previously treated small cell lung cancer. *Investigational new drugs* **1994,** *12*, 243-9.

(153) Mitsiades, C. S.; Ocio, E. M.; Pandiella, A.; Maiso, P.; Gajate, C.; Garayoa, M.; Vilanova, D.; Montero, J. C.; Mitsiades, N.; McMullan, C. J.; Munshi, N. C.; Hideshima, T.; Chauhan, D.; Aviles, P.; Otero, G.; Faircloth, G.; Mateos, M. V.; Richardson, P. G.; Mollinedo, F.; San-Miguel, J. F.; Anderson, K. C., Aplidin, a marine organism-derived compound with potent antimyeloma activity in vitro and in vivo. *Cancer research* **2008,** *68*, 5216-25.

(154)   Li, C.; Pazgier, M.; Li, J.; Li, C.; Liu, M.; Zou, G.; Li, Z.; Chen, J.; Tarasov, S. G.; Lu, W.-Y.; Lu, W., Limitations of Peptide Retro-inverso Isomerization in Molecular Mimicry. *Journal of Biological Chemistry* **2010**, *285*, 19572-19581.

(155)   Verdini, A. S.; Silvestri, S.; Becherucci, C.; Longobardi, M. G.; Parente, L.; Peppoloni, S.; Perretti, M.; Pileri, P.; Pinori, M., Immunostimulation by a partially modified retro-inverso-tuftsin analog containing Thr1.sum..psi.[NHCO](R,S)Lys2 modification. *Journal of Medicinal Chemistry* **1991**, *34*, 3372-3379.

(156)   Rjabovs, V.; Turks, M., Tetrahydrofuran amino acids of the past decade. *Tetrahedron* **2013**, *69*, 10693-10710.

(157)   Jensen, K. J.; Brask, J., Carbohydrates in peptide and protein design. *Peptide Science* **2005**, *80*, 747-761.

(158)   Risseeuw, M. D. P.; Overhand, M.; Fleet, G. W. J.; Simone, M. I., A compendium of sugar amino acids (SAA): scaffolds, peptide- and glyco-mimetics. *Tetrahedron: Asymmetry* **2007**, *18*, 2001-2010.

(159)   Risseeuw, M.; Overhand, M.; Fleet, G. J.; Simone, M., A compendium of cyclic sugar amino acids and their carbocyclic and heterocyclic nitrogen analogues. *Amino Acids* **2013**, *45*, 613-689.

(160)   Tian, G.-Z.; Wang, X.-L.; Hu, J.; Wang, X.-B.; Guo, X.-Q.; Yin, J., Recent progress of sugar amino acids: Synthetic strategies and applications as glycomimetics and peptidomimetics. *Chinese Chemical Letters* **2015**, *26*, 922-930.

(161)   Gruner, S. A. W.; Locardi, E.; Lohof, E.; Kessler, H., Carbohydrate-Based Mimetics in Drug Design: Sugar Amino Acids and Carbohydrate Scaffolds. *Chemical Reviews* **2002**, *102*, 491-514.

METHODS

Chapter II

Molecular Dynamics Methods

1.Molecular Dynamics Simulations

The need to study computationally the structure of matter at the molecular level spurred the development of a research area now known as Molecular Modeling (MM). Any molecular system can be understood in great depth if we are able to describe in detail their molecular interactions and sample the distribution of its conformational states and their energies. In this sense, MM is a set of theoretical methods supported by the fundamentals of physics, which are implemented to model, visualize and simulate the behavior of molecular systems.

Biological systems, as the investigated in the present work, should be seen through a dynamic perspective. The properties of a system arise from an average of the different conformational states explored and this type of average can be obtained recurring to Molecular Dynamics (MD) techniques and using a force field (FF) that describes the physical reality of our systems.

MD simulations had its primordium in 1957 with Alder and Wainwright,[1] which studied a system of rigid spheres. In this system particles move with constant velocity between perfectly elastic collisions. The first application of MD for the study of materials was made by Vineyard et al.[2] that investigated the damage process in a material by radiation using a short-range repulsive potential and a potential responsible for the cohesion of the crystal. Rahman[3] was the first to investigate a system under continuous potentials, describing liquid argon through MD. It was surprising to observe how a system with a reduced number of particles could satisfactorily reproduce the thermodynamic properties of real systems. Rahman also performed, in collaboration with Stilinger, the first simulation of a molecular liquid: water.[4] Many papers and books talk about the history and evolution of this area.[5-7]

Recently, Martin Karplus, Michael Levitt and Arieh Warshel, received the 2013 Nobel Prize for *The development of multiscale models for complex chemical systems*. They developed in the 1970s powerful programs to understand and predict chemical processes, being the pioneers in biomolecules simulation.[8-9]

With the increase of computational power, quantum mechanics would raise the molecular mechanics to a new level, determining the inter-atomic forces directly through the explicit consideration of the relevant electron orbitals. However, molecular mechanics produce satisfactory results even disregarding any quantum effects and this is possible due to several factors. A key factor is the portability of force fields, that is, parameters derived for small molecules can be translated into similar large macromolecules. Another important property of classical molecular mechanics is that the potential energy can be defined in terms of the atomic nuclei coordinates, which is only possible thanks

to the Born-Oppenheimer approximation.[10-11] This approximation is based on the considerable difference between the electron mass m_e and the mass of a proton m_p (m_p = 1836 m_e), which means that the first can fit almost instantly to the second without influencing it. In other words, as nuclei are much heavier than electrons, we can consider nuclei as point particles that follow classical Newtonian dynamics. The Born-Oppenheimer approximation is a concept used in quantum chemistry, although, it is implicitly used in MD simulations to justify each atom following Newtonian physics.

MD is a deterministic method (since it follows physical laws and no randomness is involved), by which sets of atomic positions are derived in sequence, applying Newton's equation of motion. The atoms motion are described according to Newton's second law.[6]

$$F(i,t_j) = m(i) \times a(i,t_j) \tag{1.1}$$

In Equation (1.1) $F(i,t_j)$ corresponds to the force acting in particle i at the moment t_j, $a(i,t_j)$ is the acceleration and $m(i)$ is the mass of particle i. From this equation the Newton equation of motion for a system of N particles is derived:

$$\frac{\partial^2 r(i,t_j)}{\partial t_j^2} = \frac{F(i,t_j)}{m(i)} \qquad i = 1,...,N \tag{1.2}$$

where $r(i,t_j)$ is the position of particle i at the moment t_j. The $F(i,t_j)$ on Equations (1.1) e (1.2) can be obtained from the gradient of the potential computed with a FF, for each particle i at the moment t_j.

$$F(i,t_j) = -\nabla V_{\{r(i,t_j)\}} \tag{1.3}$$

$$\text{and } \nabla V_{\{r(i,t_j)\}} = \frac{\partial V}{\partial r(i,t_j)} \tag{1.4}$$

In practice, for each system of particles containing a total of N atoms, the forces acting on each particle are added obtaining a resultant vector force, and hence the instantaneous acceleration from which we determine the new position and velocity of the atom in the immediately subsequent time. Continuous potential means that each particle will have its force changed with every change in its position or its neighbor's position. This situation demands that the Newtonian equations are integrated by using a differential method with finite elements performed by specific integration algorithms. The

continuous potential requires that the equations of motion be integrated by breaking the calculation into, very short, time steps (commonly in a range of 1 fs to 10 fs).

For most MD applications, Verlet-like algorithms are perfectly adequate.[12] However, sometimes it is convenient to employ a higher-order algorithm.[6] For this work we used the *leap-frog* algorithm[6, 13] which is an improved implementation of the Verlet algorithm.[12] This algorithm is less prone to numerical errors and is capable of coupling the system to a thermal bath by scaling the velocities.[6, 13]

$$r(t + \delta t) = r(t) + \delta t v(t + \tfrac{1}{2}\delta t) \qquad (1.5.)$$

$$v\left(t + \tfrac{1}{2}\delta t\right) = v\left(t - \tfrac{1}{2}\delta t\right) + \delta t a(t) \qquad (1.6)$$

$r(t)$ it is the position of particle i at the moment t, $r(t + \delta t)$ corresponds to the new position of atom i at the moment $t + \delta t$, and v and a are the velocity and acceleration. On this algorithm, the velocity is included on the determination of the new atom positions yielding higher numeric precision, due to the fact that there is no need to use the δt^2 term present on the Verlet algorithm (equations 1.7 and 1.8).

$$r(t + \delta t) = r(t) + \delta t v(t) + \tfrac{1}{2}\delta t^2 a(t) + \cdots \qquad (1.7)$$

$$r(t - \delta t) = r(t) - \delta t v(t) + \tfrac{1}{2}\delta t^2 a(t) - \cdots \qquad (1.8)$$

Adding the equations 1.7 and 1.8 we obtained:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \qquad (1.9)$$

On Verlet algorithms, the position and acceleration at time t, and the position from the previous step, $r(t - \delta t)$ are used to calculate the new atom positions at $t + \delta t$, $r(t + \delta t)$. This is an algorithm of simple implementation, but there are some drawbacks, like: the position $r(t + \delta t)$ it is only obtained by adding the term $\delta t^2 a(t)$, which causes loss of precision. Other disadvantage is that the velocity term is not explicit making hard the determination of velocities until the positions been computed at next step. The velocities can be obtained according with equation 1.10:

$$v(t) = [r(t + \delta t) - r(t - \delta t)]/2\delta t \qquad (1.10)$$

On the other hand, the *leap-frog* algorithm uses the velocity explicitly and, in consequence, the velocities and new positions are calculated together, simplifying the motion integration.

Nowadays, MD simulations are essential tools for understanding the physical basis of the structure and function of biological macromolecules. Moreover, they can provide great detail relating to the dynamic properties of model systems in more detailed manner than most experimental techniques.[14]

Computer simulations used in this study are the basis for explaining the properties of our case studies and to investigate if the biological function of the new non-canonical amino acids/peptides are related to their structural features, providing detailed conformations and other properties that determine the behavior of systems in time and space.[15]

The computational package GROMACS[16-18] was the chosen MD program for this work. We used the GROMOS 54a7 FF[19-20] available on GROMACS for all simulations performed. It is known that this program is very suitable for modeling biomolecules such as proteins and lipids. GROMACS is extremely fast to calculate the nonbonding interactions, which typically dominate the simulations, and therefore, is suitable for the study of biological systems involving large numbers of particles.[21]

2.Force Field

A biomolecular force field (FF) refers to a set of common parameters used to calculate the potential energy of a system. These parameters are derived from experimental data and quantum mechanical calculations to describe the physical reality of a system of atoms.

The potential energy calculated with a FF considers the bonded and non-bonded terms present in a molecular system. The bonded terms include the energy contribution derived from covalent bonds, such as bond, angles and torsions (proper and non-proper dihedrals). The non-bond parameters refer to terms that describe the long-range electrostatic and van der Waals (vdW) forces between different molecules or among atoms linked to a distance of more than three covalent bonds.[12]

The potential energy of a system is described on equation (2.1) and in Figure 1.

$$V = V_{bonds} + V_{angles} + V_{torsions} + V_{vdW} + V_{electrostatic} \qquad (2.1)$$

Detailing the above equation, we have the Equation 2.2. Some FF may have additional terms, but invariably contains the five components showed in this equation.

$$V(r^N) = \sum_{bonds} \frac{k_i}{2}\left(l_i - l_{i,0}\right)^2 + \sum_{angles} \frac{k_i}{2}\left(\theta_i - \theta_{i,0}\right)^2 + \sum_{torsions} \frac{V_n}{2}\left(1 + \cos(n\omega - \gamma)\right)$$
$$+ \sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right) \qquad (2.2)$$

$V(r^N)$ refers to the potential energy due to the position of the N particles (atoms) of the system. Each contribution is also represented on Figure 1. The first two terms (bonds and angles) are modeled by a harmonic potential. The third term denotes a torsional potential that models how the energy change when a bond rotates. Finally, the fourth term includes the two non-bonded contributions. The

electrostatic term is modeled through a Coulomb potential and Lennard-Jones potential is used for van der Waals (vdW) forces. The vdW interactions are usually truncated at a particular cut-off distance to reduce the number of calculations.

The constants showed in Equations 2.2, for each term, correspond to the force constants representing equilibrium parameters for the different types of physical interactions described. These constants are defined in the topology file for each FF, in our case for the FF G54a7 (see in Appendix V).

Also, the topology file describes all physical parameters of a molecule, such as, atom types, net charge, bond distance, angles, dihedrals and exclusions. For this work, it was necessary to define new topology parameters for the classes of non-canonical amino acids studied. Since most of the new amino acids are similar to natural amino acids, we derive the parameters on the existing topologies in FF G54a7, changing the required angles and dihedrals, to fit experimental or theoretical data concerning the geometry of these molecules.
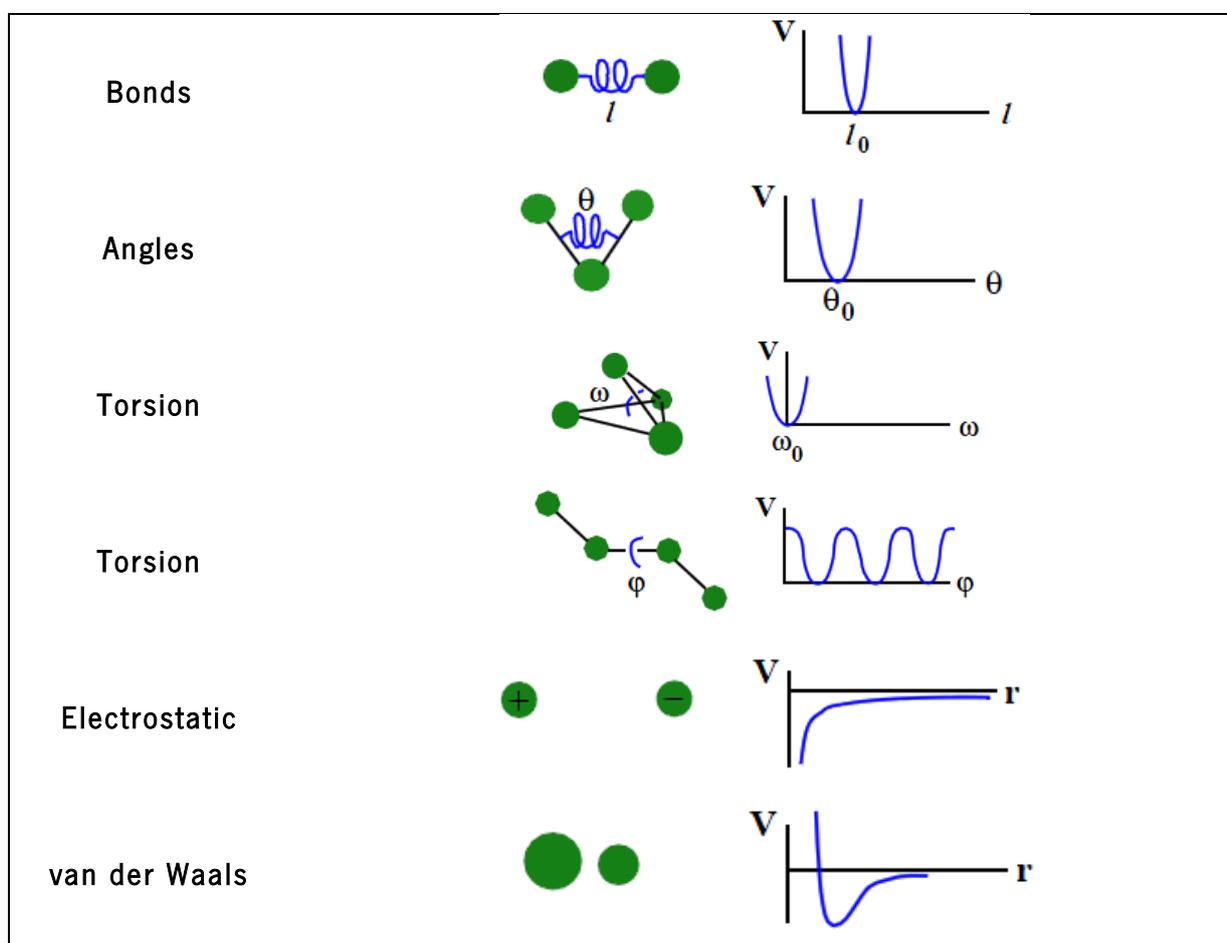
Figure 1. Schematic representation of the potential energy functions (V) common in molecular force fields. Figure adapted from Steinbach.[22]

3. Molecular Dynamics Protocol

To perform a MD simulation a series of procedures and techniques has to be chosen and applied to produce proper sampling. The standard protocol associated to a MD simulation of biomolecules can be summarized as follow.

3.1. Starting point

The spatial coordinates of the starting structure are generally obtained experimentally (from X-ray or NMR techniques). We can also use theoretical techniques like homology modeling.

It is necessary to insert the starting structure into a solvent such as water, ethanol, organic solvents or a membrane. This is necessary to reproduce physiological conditions or to observe the system behave in solvents commonly used in experimental techniques like Circular Dichroism (CD) or spectroscopy techniques. Furthermore, the molecules of the media to be used have to be equilibrated. The use of explicit solvent molecules introduces a high degree of realism in the simulation of biomolecules.

From this step, it is important to define Periodic Boundary Conditions (PBC), where the central cell is surrounded by replicas of itself.[6, 15] This procedure enables a simulation with a small number of particles and minimize surface effects that would occur if the system interacts with the void.
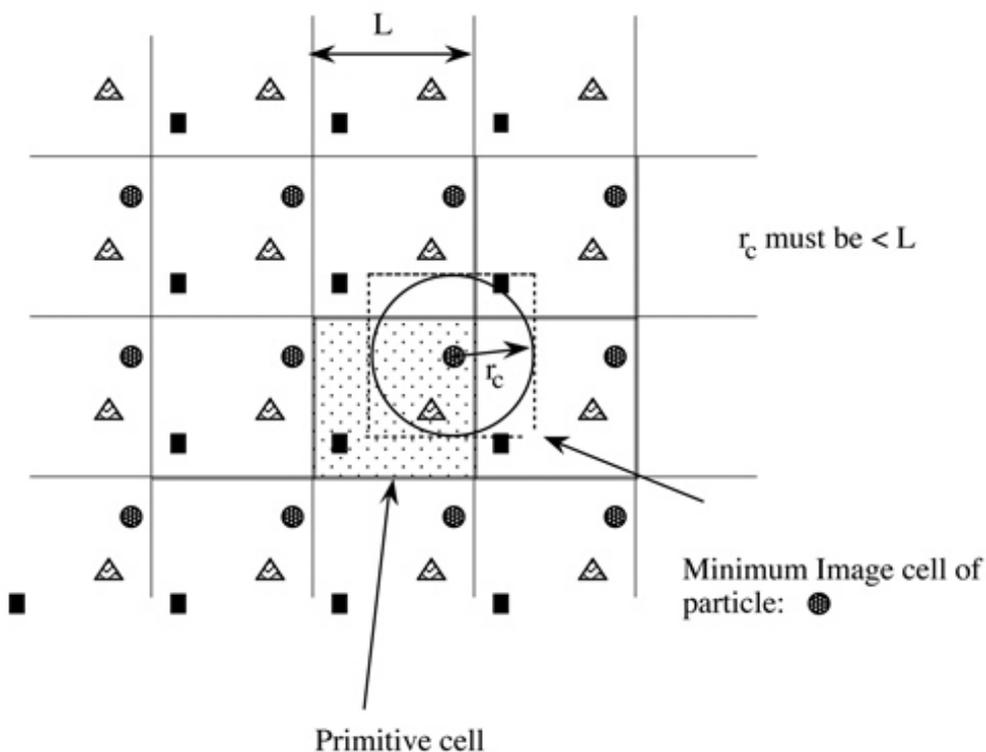
Figure 2. Periodic Boundary Conditions representation in two dimensions showing the primitive cell and where L is the size of the box and r_c is the cut-off.

To apply PBC we have to deal with the non-bonded interactions between the atoms of the central cell and the atoms of the surrounding images. It is necessary to use a box size that prevents that a particle interacts with its own image.

3.2. Energy Minimization

In this stage, the system geometry is optimized to obtain a structure that corresponds to a lower energy state. Typically, energy minimization techniques find a local minimum with respect to the starting point. During energy minimization, the geometry is changed so that the energy of the molecule is reduced, step-by-step as shown in Figure 3.
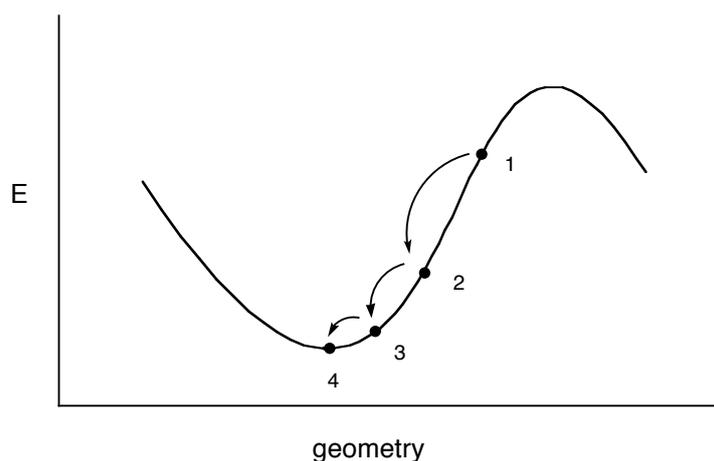
Figure 3. Description of energy minimization process, where the Energy reaches a local minimum according to changes in the geometry.

There are several methods that perform geometry optimization to find the minimum, but the most relevant on MD simulations are: the Newton-Raphson, the Steepest Descent and the Conjugate Gradient methods.

The algorithm that we used this step was the *Steepest Descent* (SD). SD is a largely used method, due to its easy implementation and because it is very efficient for structures that are far from the minimum.[6, 23] This method searches for a minimum starting in the direction that points to the largest decrease. The algorithm stops when the determined number or when a convergence criterion (related to the norm of potential gradient) is reached. The SD method is illustrated for a system with two geometrical coordinates in Figure 4.
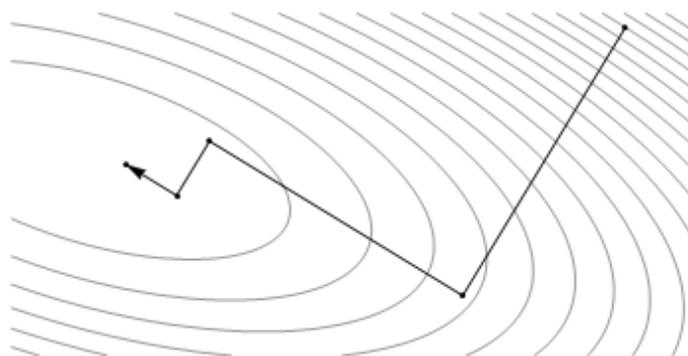
Figure 4. Graphical representation of the SD method for a system with only two geometrical coordinates.

A more common method for energy minimization is the Conjugate Gradient (CG) method, which starts similarly to the SD method: the direction in which the geometry is first minimized is in the direction in which the gradient is largest (steepest) from the initial point. Then, the algorithm proceeds iteratively along a direction perpendicular or conjugate to the current direction, reaching more rapidly the minimum, since it avoids some of the oscillations typically observed in SD. This algorithm progresses slowly in the first steps of energy minimization, but when near to the minimum can be more efficient than the SD method.

The Newton-Raphson method is based on a Taylor series expansion of the potential energy surface at the current geometry. This procedure is iterated until the parameter values stabilizes. This method is more computationally expensive of the previous methods mentioned to perform energy minimization, since it needs an estimate of the hessian.

For all simulations performed in this thesis, we use the *Reaction-Field* method[24] for the long–range electrostatic interactions. This method assumes the existence of a continuous environment, beyond a certain cutoff radius, typically 1.4 nm. Also, a dielectric constant that describes the solvent used is necessary for this algorithm. The van der Waals interactions were also truncated with twin-range cutoffs of 0.8 and 1.4 nm.

Another common method to compute long-range interactions is the Ewald summation. This method was originally designed to compute long-range interactions on crystals, because the sum is over an infinite number of periodic images. Due to PBC, particle-mesh Ewald (PME) is now widely used in biological MD simulations, but only for small systems, since the reciprocal sum increases with the number of particles in the system and this is computationally expensive.[25]

3.3. Initializing and Equilibration

At this point, we assign initial velocities for the atoms on the system and do the first integration of the equations of motion. The initial velocity of each atom it is not known, because of that, it is necessary to generate the initial velocities according to the temperature. Initial velocities are random but follow a Maxwell-Boltzmann distribution.[6] The initial velocities v_i for each atom on the system are:

$$p(v_i) = \sqrt{\frac{m_i}{2\pi kT}} \exp\left(-\frac{m_i v_i^2}{2kT}\right)$$

$$i = 1,...,3N$$

(3.3)

where k is the Boltzmann constant, T it is the absolute temperature and m_i is the mass of the atom. Once the initial velocities are defined, the potential energy of the system is calculated and it begins to integrate the Newton's equations of motion (1.2) for each particle, which will determine the trajectory of each atom.

The integration of the motion of a particle may be achieved using various algorithms. In this work we used the *leap-frog* algorithm[13] which is a modification of the Verlet algorithm,[12] as mentioned before.

In this stage, the systems should reach an equilibrium, which implies that a set of properties become stable. It is possible to monitor the equilibration through analysis like following the root mean square deviation.

The equilibration dynamics performed on this stage can be done using position restraints techniques, to impose some restrictions to atoms positions. Only on the studies presented on Chapters III, IV and V, we use this method due the peptide length and flexibility. Generally, this procedure should follow three steps: in the first one, only all heavy atoms are constrained. In the second, the restrictions are imposed to the atoms on the main-chain. On the final step the system is free and the atoms of the peptides can interact with the molecules of the environment and accommodate better in the solvation layers. These steps are necessary to relax properly the high and lower frequency modes and to avoid close contacts. These position restraints can also be applied on the minimization stage.

We pretend to analyze our systems in the NPT (isobaric-isothermic) ensemble to reproduce solution conditions. For this purpose we use, from this stage, the Berendsen thermostat and barostat algorithms[26] that serve to guarantee that the biomolecule and the solvent are under the same temperature and pressure along the simulation.

The LINCS algorithm[27-28] is a method used to maintain bonds and angle constrained and to eliminate the vibrational modes of higher frequency, allowing larger integration time steps. The SETTLE algorithm[24] fulfills the same function for water molecules.

3.4. Production Run

After initializing the system, a MD simulation is performed during the necessary time to guarantee a good sampling of the conformational states explored by the system. The total time should be superior to the relaxation time of the properties to be analyzed. Along the simulation, the positions of the atoms are recorded at fixed time intervals to form the simulation trajectory. In addition, the velocities and forces can be registered.

3.5. Analysis

The trajectories obtained are analyzed in order to calculate a variety of useful structural properties, such as deviation from an initial structure, number of hydrogen bonds, the area exposed to the solvent, the secondary structure type, and other properties.

4. Common Structural Analysis to evaluate MD Simulations

The most common tools/programs used to analyze MD Simulations performed on this work are: RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuation), Hydrogen Bond analysis, Secondary Structure (SS) Analysis, Ramachandran plots and thermodynamic analysis, among others.[23] All the programs necessary to perform these analyses are available on GROMACS package.

The RMSD is typically used to analyze the structural stability of peptides and proteins, by following the changes along the simulation against the experimental starting structure. We also can monitor the structural properties or a secondary structure preference using a DSSP (Dictionary of Secondary Structure in Proteins) method. This approach can tell us the SSs populated by a system on a MD simulation.

The Hydrogen Bond analysis is also important to characterize our peptides. We can evaluate the type of intramolecular hydrogen bonds most common for a system or analyze the intermolecular hydrogen bonds with the solvent.

The Ramachandran plots were very useful on our studies, as a tool capable to analyze each an amino acid of interest on the peptide under study, to understand the dihedral preferences of different non-canonical amino acids.

References

(1)   Alder, B. J.; Wainwright, T. E., Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* **1957,** *27*, 1208-1209.

(2)   Gibson, J. B.; Goland, A. N.; Milgram, M.; Vineyard, G. H., Dynamics of Radiation Damage. *Physical Review* **1960,** *120*, 1229-1253.

(3)   Rahman, A., Correlations in the Motion of Atoms in Liquid Argon. *Physical Review* **1964,** *136*, A405-A411.

(4)   Rahman, A.; Stillinger, F. H., Molecular Dynamics Study of Liquid Water. *The Journal of Chemical Physics* **1971,** *55*, 3336-3359.

(5)   Frenkel, D.; Smit, B., Chapter 4 - Molecular Dynamics Simulations. In *Understanding Molecular Simulation (Second Edition)*, Frenkel, D.; Smit, B., Eds. Academic Press: San Diego, 2002; pp 63-107.

(6)   Leach, A., *Molecular Modelling: Principles and Applications (2nd Edition)*. Prentice Hall: 2001.

(7)   Rino, J. P.; Studart, N., Um potencial de interação para o estudo de materiais e simulações por dinâmica molecular. *Química Nova* **2001,** *24*, 838-845.

(8)   Hodak, H., The Nobel Prize in Chemistry 2013 for the Development of Multiscale Models of Complex Chemical Systems: A Tribute to Martin Karplus, Michael Levitt and Arieh Warshel. *Journal of Molecular Biology* **2014,** *426*, 1-3.

(9)   Warshel, A.; Karplus, M., Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *Journal of the American Chemical Society* **1972,** *94*, 5612-5625.

(10)   Barnett, R. N.; Landman, U., Born-Oppenheimer molecular-dynamics simulations of finite systems: Structure and dynamics of (H2O)2. *Physical Review B* **1993,** *48*, 2081-2097.

(11)   Kubach, C.; Rougeau, N., On the use of a Born-Oppenheimer type separation in the treatment of the dynamics of molecular collisions. *Journal of Molecular Structure: THEOCHEM* **1998,** *424*, 171-180.

(12)   Verlet, L., Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* **1967,** *159*, 98-103.

(13)   Hockney, R. W., *Methods in Computational Physics*. Academic Press, New York: 1970; Vol. 9, p 135–211.

(14)   Karplus, M.; McCammon, J. A., Molecular dynamics simulations of biomolecules. *Nat Struct Mol Biol* **2002,** *9*, 646-652.

(15)   van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; Yu, H. B., Biomolecular Modeling: Goals, Problems, Perspectives. *Angewandte Chemie International Edition* **2006,** *45*, 4064-4092.

(16)   Bekker, H.; Berendsen, H. J. C.; Dijkstra, E. J.; Achterop, S.; Vondrumen, R.; Vanderspoel, D.; Sijbers, A.; Keegstra, H.; Reitsma, B.; Renardus, M. K. R., *GROMACS - A Parallel Computer for Molecular-Dynamics Simulations* World Scientific Publ Co Pte Ltd: Singapore, 1993; p 252-256.

(17)   Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory. Comput.* **2008,** *4*, 435-447.

(18)   Lindahl, E.; Hess, B.; van der Spoel, D., GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001,** *7*, 306-317.

(19)   Hermans, J.; Berendsen, H. J. C.; Van Gunsteren, W. F.; Postma, J. P. M., A consistent empirical potential for water–protein interactions. *Biopolymers* **1984,** *23*, 1513-1518.

(20)   Huang, W.; Lin, Z. X.; van Gunsteren, W. F., Validation of the GROMOS 54A7 Force Field with Respect to beta-Peptide Folding. *J. Chem. Theory Comput.* **2011,** *7*, 1237-1243.

(21)   Adcock, S. A.; McCammon, J. A., Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews* **2006,** *106*, 1589-1615.

(22)   Steinbach,   P.   J.   Introduction   to   Macromolecular   Simulation https://www.biophysics.org/Portals/1/PDFs/Education/steinbach.pdf (accessed September 2015).

(23)   Spoel, D. v. d.; Linddahl, E.; Hess, B.; Buuren, A. R. v.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; Drunen, R. v.; Berendsen, H. J. C., *Gromacs User Manual version 4.5.4*. 2010.

(24)   van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C., A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *The Journal of Chemical Physics* **1998,** *108*, 10220-10230.

(25)   Zhou, R., *Molecular Modeling at the Atomic Scale: Methods and Applications in Quantitative Biology*. CRC Press: 2014.

(26)   Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **1984,** *81*, 3684-3690.

(27)   Hess, B., P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008,** *4*, 116-122.

(28)   Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18*, 1463-1472.

RESULTS AND DISCUSSION

Chapter III

Modeling of Peptaibol Analogues Incorporating Nonpolar α,α- Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation.

Modeling of Peptaibol Analogues Incorporating Nonpolar α,α-Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation

Tarsila G. Castro and Nuno M. Micaêlo*

Departamento de Química, Escola de Ciências, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal

S *Supporting Information*

ABSTRACT: In this study, we investigate the effect of nine noncanonical α,α-dialkyl glycines on the structure, dynamics, and membrane permeation properties of a small peptaibol, peptaibolin. The noncanonical amino acids under study are Aib (α-amino isobutyric acid), Deg (α,α-diethyl glycine), Dpg (α,α-dipropyl glycine), Dibg (α,α-di-isobutyl glycine), Dhg (α,α-dihexyl glycine), DΦg (α,α-diphenyl glycine), Db_2g (α,α-dibenzyl glycine), Ac_6c (α,α-cyclohexyl glycine), and Dmg (α,α-dihydroxymethyl glycine). It is hypothesized that these amino acids are able to induce well-defined secondary structures in peptidomimetics. To investigate this hypothesis, we designed new peptaibolin peptidomimetics by replacing the native Aib positions with a new α,α-dialkyl glycine. We show that Dhg and Ac_6c noncanonical amino acids are able to induce α-helix secondary structures of peptaibolin in water, which are not present in the native structure. We also demonstrate that the α,α-dialkyl glycines increase the membrane permeability of peptaibolin in 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) membranes. However, there is no apparent correlation between increased helicity and membrane permeability. In summary, we show that some α,α-dialkyl glycines under study induce the formation of α-helix secondary structures in peptaibolin and promote spontaneous membrane permeation. Our findings increase the knowledge of the membrane permeability and folding of peptides incorporating α,α-dialkyl glycines.
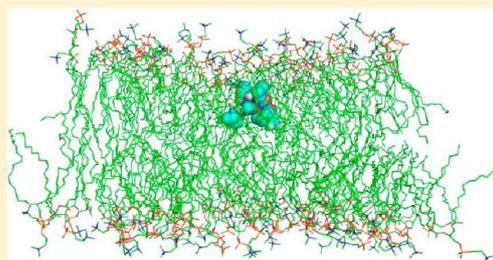
INTRODUCTION

Peptaibols are a family of membrane-active peptides biosynthesized by soil fungi.[1,2] The native sequence of these peptides incorporates the symmetric α,α-dialkyl glycine Aib (α-aminoisobutyric acid), a C-terminal alcohol having a length of 5 to 20 residues. Noncanonical amino acids Hyp (imino acid hydroxyproline) and Iva (isovaleric acid) are also very common in the peptaibol sequence.[1,3] This unique family of peptides has been investigated for the past four decades because of its antibacterial and antifungal properties and potential clinical applications. Furthermore, these peptides are very useful for investigating transmembrane ion transport through model lipid membranes, cells, and organelles.[4−7] In our study, we focus on peptaibolin, the smallest peptaibol reported.[1] Peptaibolin was isolated from two fungal strains, *Sepedonium sp.* HKI-0117 and *Sepedonium ampullosporum* HKI-0053, and characterized as a particular α-helical peptide. The structure is characterized by an N-terminus on an intramolecular three-center double hydrogen bond forming a type-III β-turn (C_{10}-ring structure) fused with an α-turn (C_{13}-ring structure). Acyl C=O0 (atom numbers in Figure 1) is the acceptor of two hydrogen bonds, with N3−H and N4−H being the donor groups. This structural motif develops into an additional α-turn (C=O1-H-N5), giving rise to an incipient α-helix. Also, C=O2-H-O was related as the fourth hydrogen bond involved in this unusual α-helix.

However, less information regarding membrane interaction and permeability has been reported.[8,9]

Peptaibolin is a leucine-based peptide carrying two Aib residues in its sequence. Aib is an important example of a noncanonical amino acid (α,α-dialkyl glycine) that occurs naturally in some peptides but is not encoded by DNA.[10] The α,α-dialkyl glycines are disubstituted amino acids at the Cα, and it is proposed that the double substitution induces a constrained conformation of the ψ and φ main-chain dihedrals.[11] Also, the steric hindrance caused by the second alkyl group attached to Cα contributes to the constrained peptide.[10] The incorporation of noncanonical amino acids has been extensively used in the design of peptidomimetics with biomedical applications.[12−14] In fact, it is shown that these amino acids are capable of inducing specific types of secondary structure that are correlated with improved peptide function. Furthermore, the insertion of these noncanonical amino acids increases the enzymatic resistance under physiological conditions.[15−19]

In this work, we investigate the conformational and membrane permeation properties of peptaibolin incorporating eight noncanonical amino acids in the native Aib positions. The
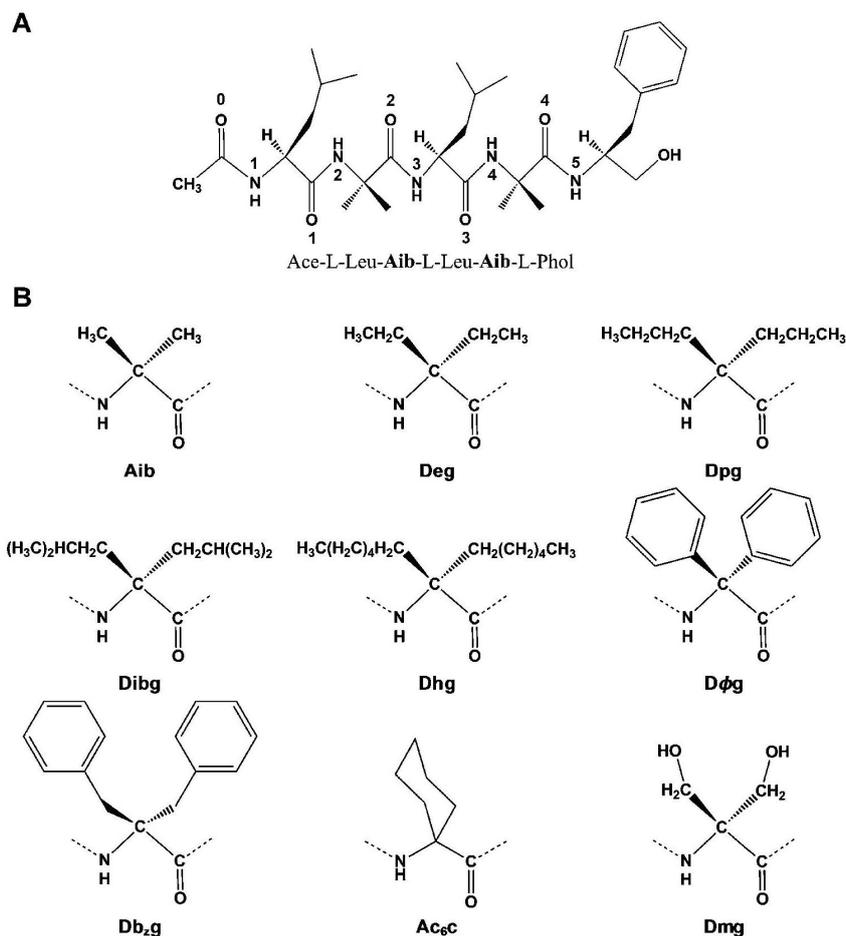
Figure 1. Two-dimensional structure of (A) peptaibolin and (B) α,α-dialkyl glycines under study: α-amino isobutyric acid (Aib), α,α-diethyl glycine (Deg), α,α-dipropyl glycine (Dpg), α,α-di-isobutyl glycine (Dibg), α,α-dihexyl glycine (Dhg), α,α-diphenyl glycine (DΦg), α,α-dibenzyl glycine (Db$_z$g), α,α-cyclohexyl glycine (Ac$_6$c), and α,α-dihydroxymethyl glycine (Dmg).

α,α-dialkyl glycines studied in this model peptide are α,α-diethyl glycine (Deg), α,α-dipropyl glycine (Dpg), α,α-di-isobutyl glycine (Dibg), α,α-dihexyl glycine (Dhg), α,α-diphenyl glycine (DΦg), α,α-dibenzyl glycine (Db$_z$g), α,α-cyclohexyl glycine (Ac$_6$c), and α,α-dihydroxymethyl glycine (Dmg). This series includes five nonpolar aliphatic amino acids of different sizes and volumes (Aib, Deg, Dpg, Dibg, and Dhg), one cyclic amino acid (Ac$_6$c), two amino acids with aromatic side chains (DΦg and Db$_z$g), and one polar aliphatic amino acid (Dmg). The smallest residues Deg and Dpg are known to induce both fully extended C$_5$ conformations and helical conformations in crystal structures.[20] Dibg has already been synthesized,[21] but there are few reports concerning its structure. Experimental and theoretical investigations indicate that DΦg and Db$_z$g induce C$_5$ and C$_7$ backbone conformations.[22−25] Previous results for cyclic amino acids such as Ac$_6$c (Ac$_n$c, $n > 3$) suggest that C$\alpha \leftrightarrow$ Cα cyclization constrains the main-chain dihedrals even more than double substitution at the Cα in Aib.[26,27] However, it is reported that the Ac$_6$c residue has folding properties similar to those of Aib. Our studies investigate the structural properties of all of these noncanonical amino acids using peptaibolin and determine which amino acids

have a greater tendency to induce α-helical secondary structures (SS) in this peptide. Furthermore, we also investigate the membrane (1-palmitoyl-2-oleoylphosphatidylcholine, POPC) permeability of peptaibolin incorporating each noncanonical amino acid.

■ MATERIALS AND METHODS

Noncanonical Amino Acid Force Field Parameters. The 3D structure of the α,α-dialkyl glycines was designed with the program PyMOL.[28] The GROMOS topologies (bonded and nonbonded parameters) for each noncanonical amino acid were based on the corresponding amino acid parametrized with the GROMOS 54a7 force field (FF).[29,30] The topologies for each α,α-dialkyl glycine under study are shown in the Supporting Information.

System Preparation. The peptaibolin experimental structure, characterized by Crisma and co-workers,[8] is not available in a database but was kindly provided by these researchers for analysis and comparison. We designed a peptaibolin α-helix structure with PyMOL and nine peptide analogues by replacing the native Aib positions with a α,α-dialkyl glycine and ALA. Aib and ALA were used as control
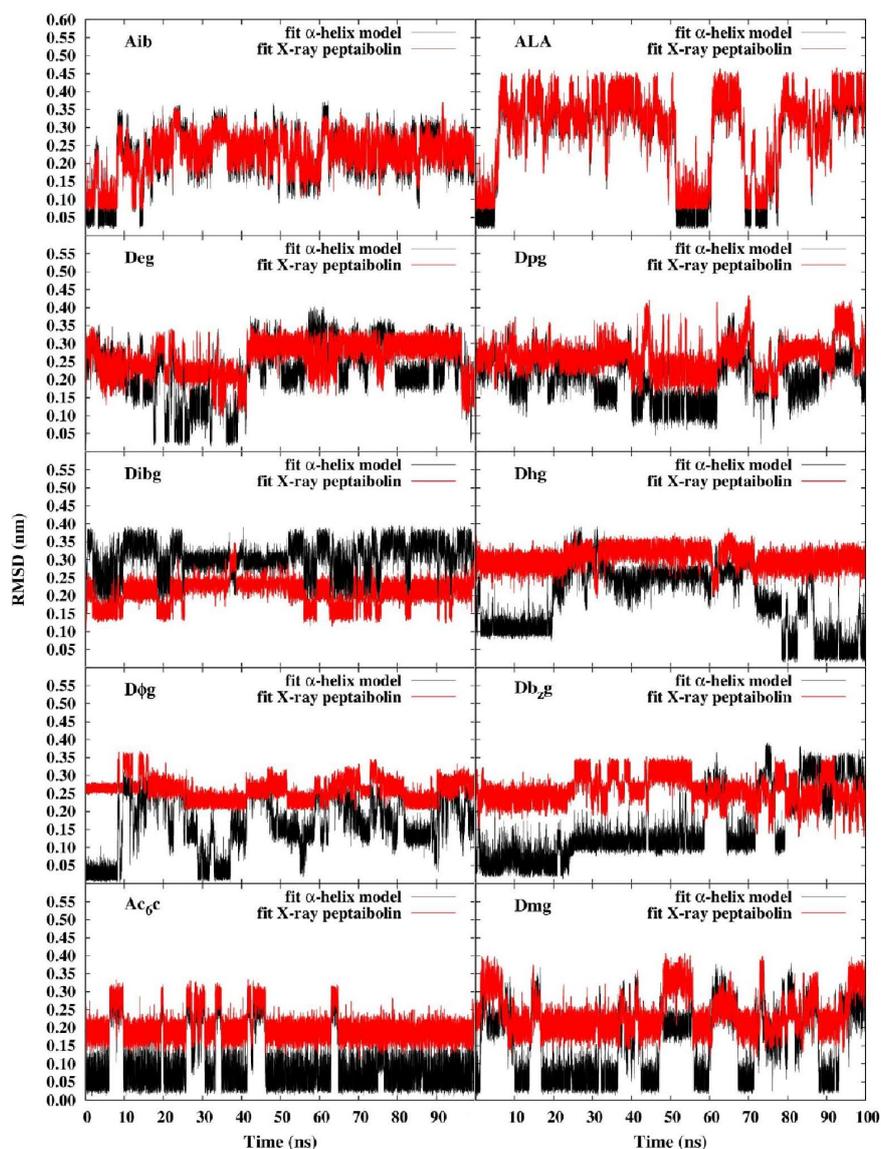
Figure 2. Cα rmsd of peptaibolin and analogues fitting the starting α-helix model vs Cα rmsd of peptaibolin and analogues fitting the X-ray structure of peptaibolin.

amino acids to compare with the properties of the new α,α-dialkyl glycines. These peptidomimetics were named with the three letter code for each α,α-dialkyl glycine that was inserted into the peptide (Figure 1).

Peptaibolin and its analogues were modeled in water, with the simple point charge (SPC) water model,[31] using dodecahedral boxes with a layer of at least 1 nm between the peptides and the walls in all three directions. The systems have about 800−1200 water molecules.

In-membrane simulations were done using a POPC membrane composed of 128 phospholipids, previously equilibrated with water.[32] Each peptide (minimized structure) was manually placed on the surface of the membrane with three different orientations: N-terminus close to the polar heads of the phospholipids, C-terminus near the polar heads, and

peptide parallel to the membrane (Figure 6). It was necessary to remove a small number of water molecules to create a cavity for peptide insertion in the aqueous phase of the membrane system. This procedure yielded 30 different peptide−POPC systems (10 peptides with 3 orientations).

Molecular Dynamics Simulations. All simulations were performed using GROMACS, version 4.0.5.[33] The reaction field method, with a cutoff of 1.4 nm and a dielectric constant of 54 for water,[31,34] was used for the treatment of long-range interactions. The van der Waals interactions were also truncated with twin-range cutoffs of 0.8 and 1.4 nm. The LINCS algorithm[35] was used to constrain the chemical bonds of the peptides and SETTLE algorithm[36] in the case of water. The systems were simulated in the isothermal−isobaric ensemble. The temperature (300 K) and pressure (1 atm)

were controlled using the Berendsen algorithms[37] with coupling constants of $\tau_T = 0.1$ ps and $\tau_P = 0.5$ ps, respectively. In POPC, these parameters were $\tau_T = 0.2$ ps and $\tau_P = 0.5$ ps.

For the peptides in water, three steps of energy minimization were performed. In the first two steps, position restraints were applied to all heavy atoms of the peptides and afterward on the main chain, with a force constant of 1000 kJ·mol^{-1}·nm^{-2}. In the third step of energy minimization, no position restraints were applied. In the peptide POCP systems, one step of energy minimization was done without position restraints.

In water, molecular dynamics (MD) simulations of 100 ns length were done. In peptide−POPC systems, 150 ns of MD simulations was used for each peptide orientation. For both systems, conformations were recorded every 1 ps. It is important to note that no force was applied to peptides to initiate the process of insertion into the membrane.

Analysis. For the peptides in water, the rmsd (root-mean-square deviation) fitted against a conformation modeled with an ideal α-helical secondary structure, Ramachandran plots, and central structure analysis was obtained over all of the conformations from the 100 ns simulations. The central structure of the peptides is obtained from an rmsd matrix that calculates the conformation that minimizes the rmsd variance against all of the conformations of a trajectory, indicating the most representative structure of the simulation. Quantitatively, we calculate a folding free energy using the rmsd analysis, assuming the conformations under the 0.15 nm cutoff to be folding states. For our isothermal−isobaric ensemble, it is possible to estimate the folding free energy as $\Delta G = -RT$ $\ln(N_{\text{folded conformations}}/N_{\text{unfolded conformations}})$. We also calculate a conformational entropy using the quasi-harmonic approximation[38] and the percentage of frames presenting zero, one, two, or three hydrogen bonds of a typical α-helix ($i \rightarrow i + 4$) secondary structure via a hydrogen bond analysis.

For the peptide−POPC systems, we calculate the distance traveled by each peptide into the membrane, that is, the distance traveled along the z axis from the starting point in the aqueous phase into the interior of the membrane. In this analysis, we considered that the origin of this reference ($z = 0$) is the plane that passes through the middle of the membrane bilayer.

■ RESULTS AND DISCUSSION

α-Helical Preorganization in Water. Figure 1 shows the 2D structure and sequence (Figure 1A) of the native peptaibolin. Labels in Figure 1A highlight the oxygen and

Table 1. Folding Free Energy (ΔG) of Peptaibolin and Analogues in Water Using an rmsd Criterion of 1.5 nm to Define Folded and Unfolded States (Materials and Methods)

peptides	ΔG (kJ/mol)
peptaibolin (Aib)	5.17
ALA	3.56
Deg	4.65
Dpg	3.56
Dibg	17.79
Dhg	1.56
DΦg	0.82
Db$_z$g	−1.56
Ac$_6$c	−4.08
Dmg	0.58

Table 2. Percentage of Conformations with Zero, One, Two, or Three Hydrogen Bonds Involved in a Typical $i \rightarrow i + 4$ α-Helix[a]

peptides	0	1	2	3
peptaibolin (Aib)	81.42	12.24	5.49	0.85
ALA	78.42	9.37	10.77	1.44
Deg	82.38	13.73	3.67	0.23
Dpg	72.20	27.62	0.18	0.00
Dibg	100.00	0.00	0.00	0.00
Dhg	81.84	8.62	9.19	0.34
DΦg	77.73	13.25	9.02	0.00
Db$_z$g	76.88	12.60	10.30	0.22
Ac$_6$c	39.19	30.71	25.34	4.76
Dmg	61.82	25.13	12.54	0.52

[a]See also Figure 1.
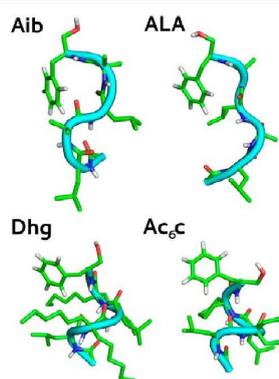
Figure 3. Central structures of peptaibolin and peptidomimetics carrying ALA, Dhg, and Ac$_6$c from the last 25 ns in water. The coloring of the atoms follows the convention: green for carbon, blue for nitrogen, red for oxygen, white for hydrogen, and cyan to highlight the cartoon that defines the SS. The water molecules were omitted for better visualization.

nitrogen atoms expected to form a typical α-helix ($i \rightarrow i + 4$). The α,α-dialkyl glycines that were inserted into the peptaibolin native Aib positions are also shown in Figure 1B.

We calculate two types of peptide Cα rmsd's (Figure 2) for all conformations for each simulation. The first one fits the peptide Cα against the initial structure modeled as an α-helix, and the other one fits the peptide Cα against the experimental structure provided by Crisma and co-workers.[8] The former identifies which noncanonical amino acids promote α-helix conformations, and the latter identifies the preference to induce a more nativelike conformation, similar to the X-ray structure.

Figure 2 shows the time evolution of the Cα rmsd of peptaibolin and its mimetics carrying ALA or a new α,α-dialkyl glycine residue. Every panel has the rmsd fit against a modeled α-helix (black traces) superimposed on the rmsd fit against the X-ray structure (red traces) for each peptide under investigation. For the black traces, the fitting of the conformations was done against the initial structure modeled in the α-helix, thus lower rmsd values (<0.15 nm) are an indication that the peptide structure is close to an ideal α-helix. In this case, the rmsd analysis reveals that the α,α-dialkyl glycines have a different propensity to maintain the peptide structure close to a typical α-helix conformation. Peptaibolin carrying Dpg and Dibg seem to populate conformations

Article
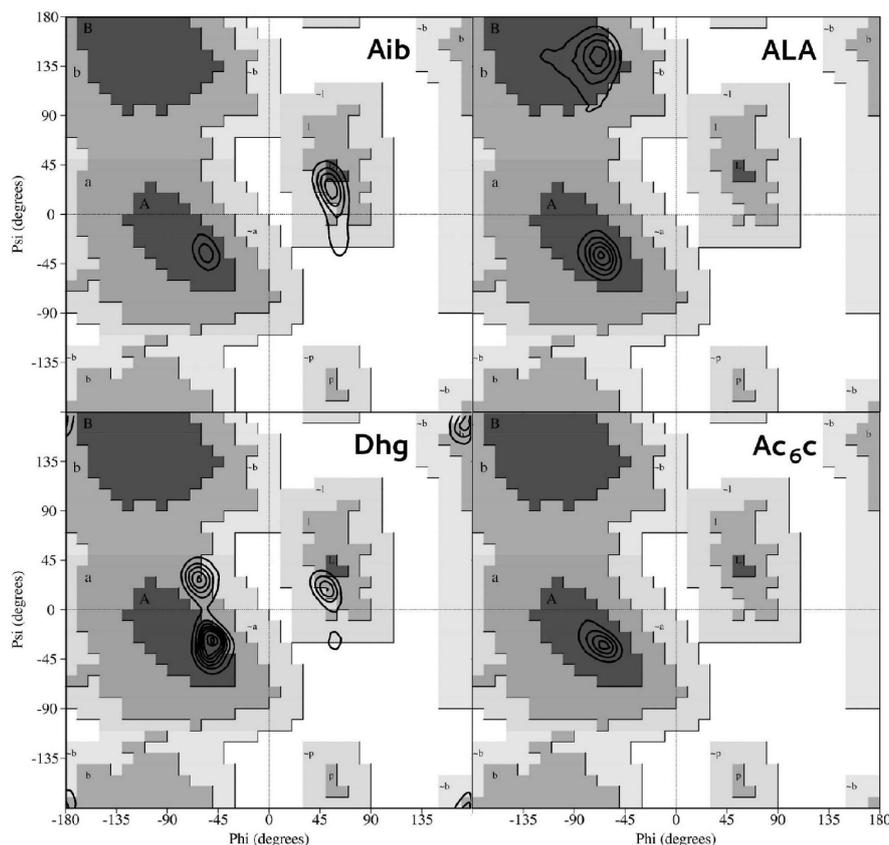
Figure 4. Probability density contours of φ and ψ pairs for amino acids AIB, ALA, Dhg, and Ac$_6$c in water. These contours are superimposed on the Ramachandran diagram in which region A corresponds to typical dihedrals of the right α-helix, B corresponds to the β-sheet space, and L corresponds to the left α-helix region. The contours levels used were 0.002−0.012 spaced by 0.002 for peptaibolin, 0.001−0.006 spaced by 0.001 for ALA, 0.001−0.009 spaced by 0.001 for Dhg, and 0.005−0.025 spaced by 0.005 for Ac$_6$c.

Table 3. Peptide Conformational Entropy (S) in Water, Estimated with the Quasi-Harmonic Approximation Method (Materials and Methods)

peptides	S (kJ/mol/K)
peptaibolin (Aib)	1.67
ALA	1.73
Deg	1.88
Dpg	2.10
Dibg	1.88
Dhg	2.78
DΦg	2.26
Db$_z$g	2.35
Ac$_6$c	1.30
Dmg	1.92

dissimilar to the initial structure. However, Deg, Dhg, Dϕg, Db$_z$g, and Dmg peptides have some conformations close to the α-helix structure at several time intervals during the simulation, indicating that these substitutions promote the formation of this type of secondary structure (SS). The peptide with Ac$_6$c is the one that remains close to an α-helix secondary structure during most of the 100 ns simulation time. Ac$_6$c is the only Cα cyclized amino acid of this series. Previous results suggest that the Cα ↔ Cα cyclization constrains the main-chain dihedrals

even more than the constraint resulting from the double substitution at Cα, as in Aib.[26,27]

The red traces show the Cα rmsd fit to the peptaibolin X-ray structure (Figure 2). We highlight the peptide-containing Dibg residue; it is the only one that presents conformations more similar to the native peptaibolin X-ray structure compared to an ideal α-helix. For the peptides more similar to sample α-helical structures, like the ones containing Dhg, DΦg, Db$_z$g, Ac$_6$c, and Dmg, we see a shift in the rmsd's (fit against the native structure) to higher values. This fact is an indication that these α,α-dialkyl glycines tend to induce structures closer to typical α-helical conformations than to the native peptaibolin X-ray structure.

We estimate a folding free energy (ΔG) for the modeled peptaibolin and analogues (Table 1) using the rmsd data fitted against an ideal α-helical peptaibolin reference structure.

Table 1 demonstrates quantitatively the same trend observed on the black rmsd traces shown in Figure 2: the peptides carrying Ac$_6$c and Db$_z$g demonstrate the preference to induce α-helical folded states of peptaibolin; DΦg and Dmg also have a reasonable number of folding states, resulting in a folding ΔG close to zero. The peptide carrying Dibg has the highest free energy, indicating that this residue induces a significant deviation from the initial α-helix structure.
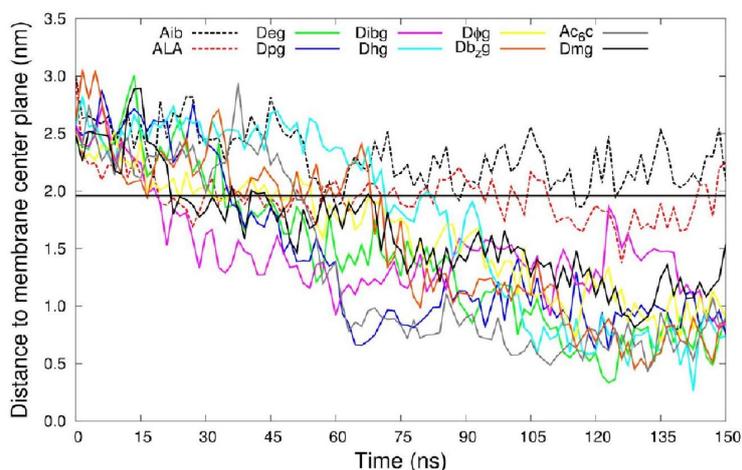
Figure 5. Center of mass position of each amino acid along the *z*-axis component of the simulation box. Peptides migrate spontaneously from the aqueous environment toward the center of the membrane during the 150 ns simulation time. The line at 1.96 nm corresponds to the water/membrane interface.

Hydrogen bond analysis was employed to check if the peptides have the intramolecular interactions expect to form an α-helix (in Figure 1, C=O0-H-N4, C=O1-H-N5, and C=O2-H-O). Table 2 shows, for each peptide, the percentage of conformations sampled over the entire 100 ns simulation with zero, one, two, or three hydrogen bonds.

The analysis of the helicity based on the backbone hydrogen bonds agrees with the previous analysis. Residue Ac_6c is the most capable of sample conformations with an α-helix structure; the percentage of frames with one, two, or three $i \rightarrow i + 4$ hydrogen bonds corresponds to 61% of the sampled conformations. The peptides carrying Dmg, Db_zg, and DΦg also present a high percentage of conformations with $i \rightarrow i + 4$ hydrogen bonds.

Figure 3 shows the central structures obtained for the last 25 ns of simulation in water for peptaibolin and the peptidomimetics with ALA, Dhg, and Ac_6c. (Supporting Information Figure S1 shows the central structure of peptaibolin, ALA, and Dhg at other time intervals.)

The central structures shown in Figure 3 clearly reveal that the peptides carrying Dhg and Ac_6c have an α-helical conformation whereas Aib and ALA promote peptide unfolding. Dhg is the bulkiest aliphatic noncanonical amino acid and Ac_6c is the only one cyclized at the Cα, suggesting that these characteristics seem to promote the folding of the peptide into helical SS. Ramachandran plots (Figure 4) help us to understand the individual geometrical properties of the main-chain dihedrals for each amino acid incorporated into peptaibolin. The distribution was calculated from the ψ and φ angles recorded from the two Aib positions of peptaibolin, replaced by ALA or by any α,α-dialkyl glycine. A total of 200 000 points were used to calculate the probability densities shown in Figure 4 (1 peptide × 2 residue positions × 100 000 conformations) and are displayed as probability density contours. We show the main-chain dihedral geometry sampled by Aib, ALA, Dhg, and Ac_6c. The Ramachandran diagrams for the other noncanonical amino acids discussed in this work are shown in the Supporting Information (Figure 2S).

Figure 4 shows, as expected, that ALA has greater conformational freedom than Aib. ALA sampled conformations corresponding to the β-sheet and right α-helix regions, whereas Aib explores conformations only of right and left helices. The double substitution at the Aib Cα eliminates the β-sheet conformations present in ALA, restraining the conformational space toward right- and left-helical SS. Dhg main-chain dihedrals show a distribution of angles in helical regions but can also explore extended conformations ($\varphi \approx \pm 180°$, $\psi \approx \pm 180°$). The Ac_6c Ramachandran plot clearly shows that the dihedral pairs are exclusively constrained in the right α-helix space. This fact is clear evidence of the greater tendency of the peptide to adopt an α-helix SS as observed from the previous rmsd data (Figure 2). Although we cannot establish a rule stating that residues that populate main-chain dihedral angles of the Ramachandran plot typical of α-helices will lead to peptides rich in α-helical SS, we consider that this fact can be an indication of this effect.

The conformational entropy (S) of peptaibolin and its mimetics were estimated using a quasi-harmonic approximation.[38]

The conformational entropy indicates that the peptidomimetic containing Ac_6c has the lowest conformation entropy among the peptides under study. This fact indicates that the peptide is highly constrained, leading to the conclusion that Ac_6c not only induces α-helix SS in peptaibolin but also imposes lower conformation freedom on this peptide. The peptides containing the residues with short side chains—Aib, ALA, Deg, and Dmg—have conformational entropy lower than for the peptides carrying large, bulky side chains such as Dpg, DΦg, and Db_zg with the exception of Dibg, which seems to be restrained in an unfolded state.

Spontaneous Insertion in POPC Membranes. Experimentally, peptaibolin shows activity against gram-positive bacteria, as reported by Hulsmann and co-workers, but to date, no mechanism of insertion into membranes has been reported.[1,9] Our membrane permeation study of peptaibolin and analogues brings new knowledge about the insertion mechanism into membranes and, most importantly, about the effect of each noncanonical amino acid on this process. Our modeling experiments were done on a POPC membrane previously equilibrated in water. We emphasize that no

654
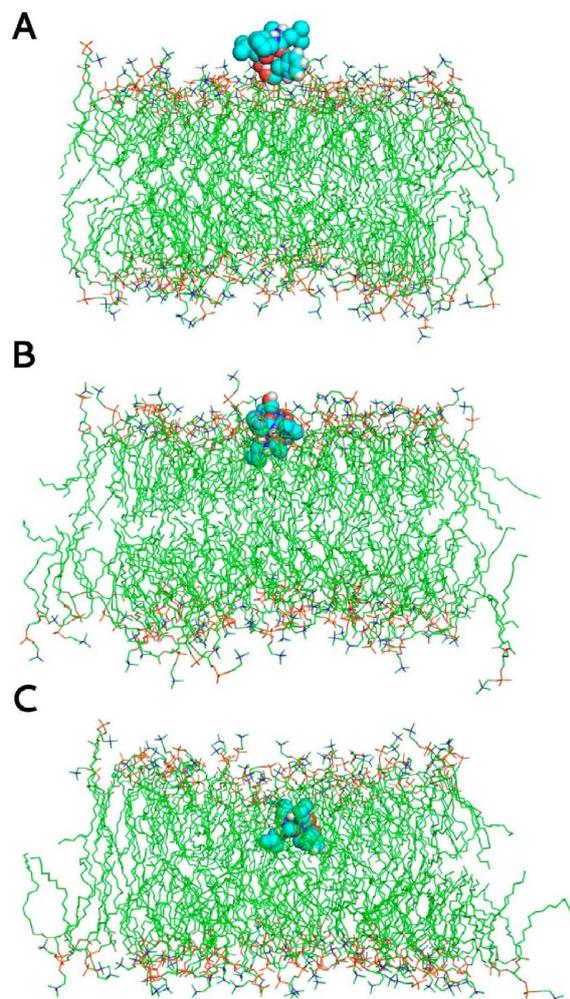
Figure 6. Snapshots of the spontaneous insertion of the Ac$_6$c analogue into a POPC membrane. Water molecules were omitted for better visualization. Snapshots: (A) $t = 0$, (B) $t = 55$ ns, and (C) $t = 150$ ns.

potential was applied to the peptides to promote membrane insertion. All peptides were manually placed on the surface of the membrane with three different orientations, close to the polar heads of the phospholipids.

Figure 5 shows the spontaneous insertion of the noncanonical peptidomimetics into a POPC membrane. This figure shows the distance traveled for each peptide (one peptide orientation only). We did not observe a specific tendency for one specific replicate. We observe that peptaibolin and the peptidomimetics incorporating ALA are preferentially located on the surface of the membrane. Our modeling experiment shows that, apparently, there is no spontaneous insertion of these two peptides in the POPC membrane. Although there is no experimental information regarding the membrane permeation properties of peptaibolin, we cannot conclude that this process will not occur in in vitro experiments or under physiological conditions. However, the peptidomimetics incorporating the new noncanonical amino acids permeate the POPC membrane in a spontaneous way. The peptide with

Dmg, a polar α,α-dialkyl glycine, permeates the membrane and remains on the POPC water interface close to the polar heads of phospholipids. However, the peptide carrying the apolar noncanonical residues shows a greater tendency to enter the POPC membrane. This fact is somehow expected from the apolar characteristics of these residues. It must be emphasized that we cannot conclude that greater permeation will lead to improved antimicrobial activity without further experimental studies. Nevertheless, it is evident so far that peptaibolin incorporating Ac$_6$c was demonstrated to be able to permeate POPC membranes in a spontaneous way, in addition to his greater tendency to adopt a rigid α-helical SS. In light of the current models of action of these peptides, the helical form is considered to be the ideal conformation for the biological activity and the formation of barrel-stave-type channels;[39] consequently, we can assume that the preorganization in the α-helix of this peptide might help to reduce the structural rearrangement necessary to penetrate the membrane and formation of membrane channels. Figure 6 displays three frames corresponding to the spontaneous insertion of the peptide with Ac$_6$c into a POPC membrane at three simulation times: $t = 0$, 55, and 150 ns.

To investigate if the peptides that translocate into the membrane environment also adopt conformations close to an α-helical structure and if there is any correlation between the α-helical conformation and membrane permeation, we calculated the rmsd (fitted against the model α-helix conformation) and plotted this data as a function of the distance traveled toward the center of the POPC membrane (Figure 7). It is clear that peptaibolin and the peptide replaced with ALA remain on the surface of the membrane (Figure 7, top panels). However, α,α-dialkyl glycines Deg, Dpg, Dibg, Dhg, Db$_2$g, and Dmg promote the peptide insertion into the POPC membrane, as observed before. However, the conformations of these peptides that are found inside the membrane environment ($d < 1.96$ nm) have high rmsd values, indicating that they are quite dissimilar from the initial α-helix structure used for fitting. The peptides carrying DΦg and Ac$_6$c also permeate the POPC membrane and have low Cα rmsd values, indicating a more nativelike conformation similar to the α-helix. These results clearly indicate that the Ac$_6$c noncanonical residue promotes the folding of peptaibolin into helical SS in aqueous and membrane environments; however, we cannot establish a correlation between α-helical preorganization and membrane permeability. We have to clarify that current models of membrane disruption by these peptides, such as the barrel stave-type channels,[39] imply that more than one peptide monomer is necessary to disrupt the membrane. This suggests that these peptides could adopt α-helical structures in a cooperative way in the presence of more monomers. This aspect will be the subject of future studies.

■ CONCLUSIONS

Our findings on the insertion of α,α-dialkyl glycines in small peptaibols suggest that some noncanonical residues are more capable of inducing α-helical conformations and promoting spontaneous membrane permeation than the native Aib in peptaibolin. However, there is no correlation between the acquired helicity and membrane permeability because some of the peptides permeated the membrane without adopting α-helical conformations in either an aqueous or membrane environment. We demonstrate that Dhg and Ac$_6$c are able to maintain a structure closer to the reference one modeled in the
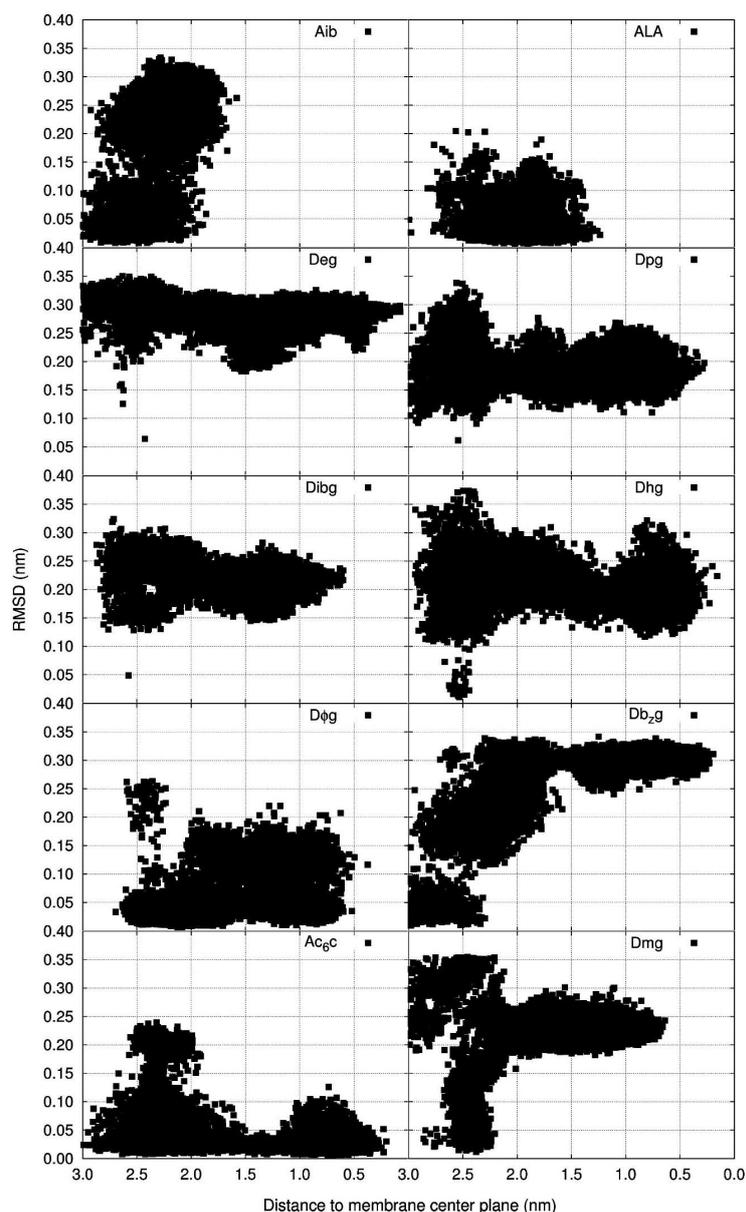
Figure 7. Cα rmsd of peptaibolin and analogues along the spontaneous insertion into the POPC membrane. Fitting of Cα relative to the minimized structure of each peptide.

α-helix. In fact, Ac_6c is the most efficient residue in constraining the conformations of the peptide close to the modeled structure in the α-helix. Furthermore, Ramachandran plots show that the dihedral pair of this residue explores only the right α-helix regions. These facts could be explained by the Cα cyclization of this residue that constrains its main-chain dihedrals. Furthermore, all nonpolar noncanonical amino acids promoted spontaneous membrane peptide permeation. This fact is somehow expected because of the apolar characteristics of these residues. Only peptaibolin and the corresponding peptidomimetic incorporating ALA did not permeate the membrane in our simulations; however, this result must be further validated experimentally. We emphasize that the peptide incorporating the Ac_6c residue also induces nativelike conformations inside the membrane environment. Maintaining the helicity of this peptide in the membrane environment could improve the antimicrobial properties of this peptide, but further experimental studies will be required to validate this hypothesis. If seems thus far that the Ac_6c noncanonical residue is a versatile residue that is capable of inducing helical conformations in short peptides, inducing lower conformational flexibility, and improving membrane permeability.

We propose that peptaibolin peptidomimetics incorporating Ac_6c residues might improve their structure and antimicrobial function, and for this reason, further investigations of peptides incorporating $C\alpha$-cyclized residues will be presented in the near future.

■ ASSOCIATED CONTENT

⑤ Supporting Information

All parametrizations for the new amino acids discussed in this article. Central structures of peptaibolin and analogues with ALA and Dhg at different time intervals and Ramachandran plots for noncanonical amino acids Deg, Dpg, Dibg, DΦg, Db$_2$g, and Dmg. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: micaelo@quimica.uminho.pt. Phone: +351 253 604 384.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Chugh, J. K.; Wallace, B. A. Peptaibols: Models for Ion Channels. *Biochem. Soc. Trans.* **2001**, *29*, 565−570.

(2) Whitmore, L.; Wallace, B. A. The Peptaibol Database: A Database for Sequences and Structures of Naturally Occurring Peptaibols. *Nucleic Acids Res.* **2004**, *32*, D593−D594.

(3) Whitmore, L.; Wallace, B. A. Analysis of Peptaibol Sequence Composition: Implications for in vivo Synthesis and Channel Formation. *Eur. Biophys. J. Biophys. Lett.* **2004**, *33*, 233−237.

(4) Lippe, C. Thiourea and Potassium Permeability of Phospholipid Bilayer Membranes as Affected by Enniatin B. *Nature* **1968**, *218*, 196−197.

(5) Menestrina, G.; Voges, K.-P.; Jung, G.; Boheim, G. Voltage-Dependent Channel Formation by Rods of Helical Polypeptides. *J. Membr. Biol.* **1986**, *93*, 111−132.

(6) Mueller, P.; Rudin, D. O. Action Potentials Induced in Biomolecular Lipid Membranes. *Nature* **1968**, *217*, 713−719.

(7) Ovchinni, Ya Membrane Active Complexones - Chemistry and Biological Function. *FEBS Lett.* **1974**, *44*, 1−21.

(8) Crisma, M.; Barazza, A.; Formaggio, F.; Kaptein, B.; Broxterman, Q. B.; Kamphuis, J.; Toniolo, C. Peptaibolin: Synthesis, 3D-Structure, and Membrane Modifying Properties of the Natural Antibiotic and Selected Analogues. *Tetrahedron* **2001**, *57*, 2813−2825.

(9) Hulsmann, H.; Heinze, S.; Ritzau, M.; Schlegel, B.; Grafe, U. Isolation and Structure of Peptaibolin, a New Peptaibol from Sepedonium Strains. *J. Antibiot.* **1998**, *51*, 1055−1058.

(10) Mendel, D.; Ellman, J.; Schultz, P. G. Protein-Biosynthesis with Conformationally Restricted Amino-Acids. *J. Am. Chem. Soc.* **1993**, *115*, 4359−4360.

(11) Fu, Y. W.; Hammarstrom, L. G. J.; Miller, T. J.; Fronczek, F. R.; McLaughlin, M. L.; Hammer, R. P. Sterically Hindered $C^{\alpha,\alpha}$-Disubstituted α-Amino Acids: Synthesis from α-Nitroacetate and Incorporation into Peptides. *J. Org. Chem.* **2001**, *66*, 7118−7124.

(12) Gentilucci, L.; Tolomelli, A.; Squassabia, F. Peptides and Peptidomimetics in Medicine, Surgery and Biotechnology. *Curr. Med. Chem.* **2006**, *13*, 2449−2466.

(13) Grauer, A.; Konig, B. Peptidomimetics - A Versatile Route to Biologically Active Compounds. *Eur. J. Org. Chem.* **2009**, 5099−5111.

(14) Vagner, J.; Qu, H. C.; Hruby, V. J. Peptidomimetics, a Synthetic Tool of Drug Discovery. *Curr. Opin. Chem. Biol.* **2008**, *12*, 292−296.

(15) Oh, J. E.; Lee, K. H. Synthesis of Novel Unnatural Amino Acid as a Building Block and Its Incorporation into an Antimicrobial Peptide. *Bioorg. Med. Chem.* **1999**, *7*, 2985−2990.

(16) Ressurreicao, A. S. M.; Bordessa, A.; Civera, M.; Belvisi, L.; Gennari, C.; Piarulli, U. Synthesis and Conformational Studies of Peptidomimetics Containing a New Bifunctional Diketopiperazine Scaffold Acting as a β-Hairpin Inducer. *J. Org. Chem.* **2008**, *73*, 652−660.

(17) Feytens, D.; Cescato, R.; Reubi, J. C.; Tourwe, D. New sst(4/5)-Selective Somatostatin Peptidomimetics Based on a Constrained Tryptophan Scaffold. *J. Med. Chem.* **2007**, *50*, 3397−3401.

(18) Mallareddy, J. R.; Borics, A.; Keresztes, A.; Kover, K. E.; Tourwe, D.; Toth, G. Design, Synthesis, Pharmacological Evaluation, and Structure−Activity Study of Novel Endomorphin Analogues with Multiple Structural Modifications. *J. Med. Chem.* **2011**, *54*, 1462−1472.

(19) Whitby, L. R.; Ando, Y.; Setola, V.; Vogt, P. K.; Roth, B. L.; Boger, D. L. Design, Synthesis, and Validation of a beta-Turn Mimetic Library Targeting Protein−Protein and Peptide−Receptor Interactions. *J. Am. Chem. Soc.* **2011**, *133*, 10184−10194.

(20) Rodrigues, L. M.; Fonseca, J. I.; Maia, H. L. S. Synthesis and Conformational Investigation of Tetrapeptide Analogues of the Fragment B23-B26 of Insulin. *Tetrahedron* **2004**, *60*, 8929−8936.

(21) Fu, Y.; Etienne, M. A.; Hammer, R. P. Facile Synthesis of α,α-Diisobutylglycine and Anchoring Its Derivatives into PAL-PEG-PS Resin. *J. Org. Chem.* **2003**, *68*, 9854−9857.

(22) Casanovas, J.; Zanuy, D.; Nussinov, R.; Alemán, C. Intrinsic Conformational Characteristics of α,α-Diphenylglycine. *J. Org. Chem.* **2007**, *72*, 2174−2181.

(23) Casanovas, J.; Nussinov, R.; Alemán, C. Intrinsic Conformational Preferences of $C^{\alpha,\alpha}$-Dibenzylglycine. *J. Org. Chem.* **2008**, *73*, 4205−4211.

(24) Pavone, V.; Lombardi, A.; Saviano, M.; De Simone, G.; Nastri, F.; Maglio, O.; Omote, Y.; Yamanaka, Y.; Yamada, T. Conformational Behavior of $C^{\alpha,\alpha}$-Diphenyl Glycine: Extended Conformation in Tripeptides Containing Consecutive Dϕg Residues. *Biopolymers* **2000**, *53*, 161−168.

(25) Valle, G.; Crisma, M.; Bonora, G. M.; Toniolo, C.; Lelj, F.; Barone, V.; Fraternali, F.; Hardy, P. M.; Langrangoldsmith, A.; Maia, H. L. S. Structural Versatility of Peptides from $C^{\alpha,\alpha}$ Disubstituted Glycines - Preferred Conformational of the $C^{\alpha,\alpha}$-Dibenzylglycine Residue. *J. Chem. Soc., Perkin Trans. 2* **1990**, 1481−1487.

(26) Rodriguez-Ropero, F.; Zanuy, D.; Casanovas, J.; Nussinov, R.; Aleman, C. Application of 1-Aminocyclohexane Carboxylic Acid to Protein Nanostructure Computer Design. *J. Chem. Inf. Model.* **2008**, *48*, 333−343.

(27) Alemán, C. Conformational Properties of α-Amino Acids Disubstituted at the α-Carbon. *J. Phys. Chem. B* **1997**, *101*, 5046−5050.

(28) Schrödinger. *The PyMOL Molecular Graphics System*, 1.3r1; 2010.

(29) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J. Biophys. Lett.* **2011**, *40*, 843−856.

(30) Huang, W.; Lin, Z. X.; van Gunsteren, W. F. Validation of the GROMOS 54A7 Force Field with Respect to β-Peptide Folding. *J. Chem. Theory Comput.* **2011**, *7*, 1237−1243.

(31) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(32) Tieleman, P. Palmitoyl-linoleylphosphatidylcholine Bilayer with 128 Lipids. http://moose.bio.ucalgary.ca/index.php?page=Structures_and_Topologies.

(33) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.

(34) Smith, P. E.; Vangunsteren, W. F. Consistent Dieletric-Properties of the Simple Point-Charge and Extended Simple Point Charge Water Models at 277 and 300K. *J. Chem. Phys.* **1994**, *100*, 3169−3174.

(35) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(36) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for Use with a Reaction Field. *J. Chem. Phys.* **1998**, *108*, 10220−10230.

(37) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(38) Baron, R.; Gunsteren, W. F. v.; Hünenberger, P. H. Estimating the Configurational Entropy from Molecular Dynamics Simulations: Anharmonicity and Correlation Corrections to the Quasi-Harmonic Approximation. *Trends Phys. Chem.* **2006**, *11*, 87−122.

(39) Shai, Y. Mode of Action of Membrane Active Antimicrobial Peptides. *Biopolymers* **2002**, *66*, 236−248.

Chapter IV

Conformational and Thermodynamic Properties of Non-Canonical
α,α-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies

Conformational and Thermodynamic Properties of Non-Canonical α,α-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies

Tarsila G. Castro and Nuno M. Micaêlo*

Departamento de Química, Escola de Ciências, Universidade do Minho, Largo do Paço, Braga 4704-553, Portugal

ⓢ *Supporting Information*

Aib ---> Dhg
2, 4, 6, 9, 11,
14, 17 and 18

Native Alamethicin with Aib   Alamethicin peptidomimetic with Dhg
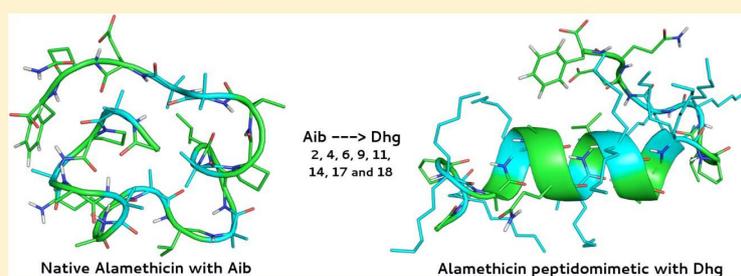
ABSTRACT: In this work, we investigate the structure, dynamic and thermodynamic properties of noncanonical disubstituted amino acids (α,α-dialkyl glycines), also known as non-natural amino acids, in the peptaibol Alamethicin. The amino acids under study are Aib (α-amino isobutyric acid or α-methyl alanine), Deg (α,α-diethyl glycine), Dpg (α,α-dipropyl glycine), Dibg (α,α-di-isobutyl glycine), Dhg (α,α-dihexyl glycine), DΦg (α,α-diphenyl glycine), Db$_2$g (α,α-dibenzyl glycine), Ac$_6$c (α,α-cyclohexyl glycine), and Dmg (α,α-dihydroxymethyl glycine). It is hypothesized that these amino acids are able to induce well-defined secondary structure in peptidomimetics. To test this hypothesis, new peptidomimetics of Alamethicin were constructed by replacing the native Aib positions of Alamethicin by one or more new α,α-dialkyl glycines. Dhg and Ac$_6$c demonstrated the capacity to induce well-defined α-helical structures. Dhg and Ac$_6$c also promote the thermodynamic stabilization of these peptides in a POPC model membrane and are better alternatives to the Aib in Alamethicin. These noncanonical amino acids also improved secondary structure properties, revealing preorganization in water and maintenance of α helical structure in POPC. We show that it is possible to optimize the helicity and thermodynamic properties of native Alamethicin, and we suggest that these amino acids could be incorporated in other peptides with similar structural effect.

◼ INTRODUCTION

α,α-Dialkyl glycines are noncanonical amino acids where the Cα is substituted with two alkyl side chains. This substitution can be symmetrical or not. It is proposed that the double substitution at the Cα of α,α-dialkyl glycines induces a more constrained conformation of the φ and ψ main-chain dihedral angle pair. Consequently, these amino acids should explore a more restrictive range of dihedral angles of the Ramachandran space corresponding to the helical secondary structure conformation, observed for the natural amino acids encoded by DNA.[1] Such structural arrangement, combined with the steric hindrance caused by the presence of the second alkyl group attached to Cα, leads to the formation of constrained peptides. These properties can be very important to evaluate the foldamer potential—capability to induce always the same secondary structure, independent of the amino acids sequence or the solvent used—of these amino acids in the modeling of peptides with a particular secondary structure.[2]

In fact, noncanonical amino acids already have a relevant role on the conformation and design of peptidomimetics with biomedical applications.[3−6] It is also shown that the incorporation of noncanonical amino acids is capable to induce specific types of secondary structure in peptides with a significant increase on the bioavailability and stability in physiological conditions.[7−12] This type of amino acids may occur naturally in some peptides but are not encoded by DNA. Known natural examples of this class of amino acids are Aib (α-amino isobutyric acid) and IVA (isovaline or isovaleric acid).[2] Aib occurs naturally in peptaibols with antibiotic activity such as Alamethicin, Zervamicin, and Antiamoebin I.[13,14] In these three peptides, this amino acid is responsible for the formation of α-helical structures. This type of arrangement is essential for their insertion into lipid bilayers of cell membranes and formation of barrel stave type channels.[15−17] In a recent modeling study done by us, we also reported the α-helical preorganization of a small peptide Peptaibolin, as a result of the incorporation of this class of amino acids.[18]

93

Tarsila G. Castro and Nuno M. Micaêlo.:dx.doi.org/10.1021/jp505400q | J. Phys. Chem. B 2014, 118, 9861–9870.
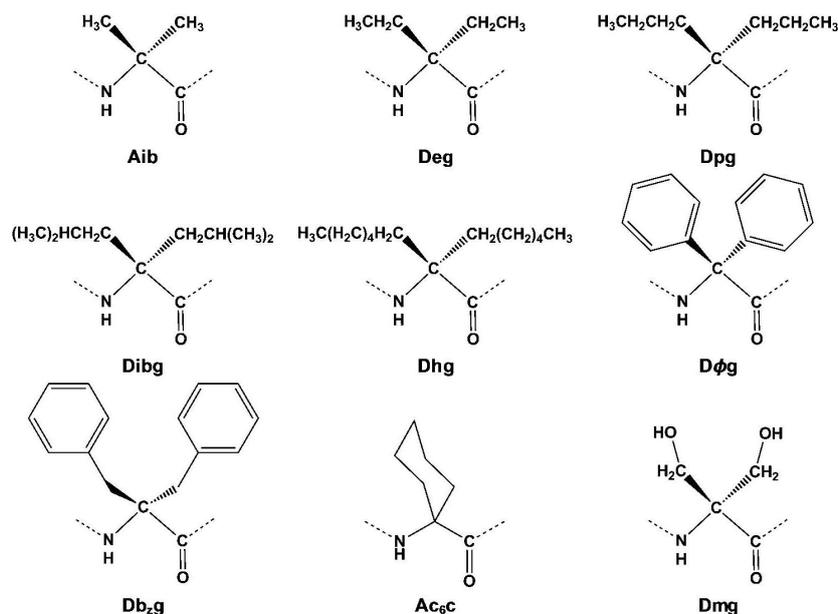
Figure 1. Two-dimensional structure of the α,α-dialkyl glycines studied in this work: α-amino isobutyric acid (Aib), α,α-diethyl glycine (Deg), α,α-dipropyl glycine (Dpg), α,α-di-isobutyl glycine (Dibg), α,α-dihexyl glycine (Dhg), α,α-diphenyl glycine (DΦg), α,α-dibenzyl glycine (Db$_z$g), α,α-cyclohexyl glycine (Ac$_6$c), and α,α-dihydroxymethyl glycine (Dmg).

In this work, we evaluate the structural and thermodynamic effects of replacing Aib by symmetric α,α-dialkyl glycines in Alamethicin, a peptaibol with well-known conformational structure and function, in order to establish if the noncanonical amino acids increase the helicity of novel Alamethicin peptides compared to the native structure. The new α,α-dialkyl glycines studied in this paper are α,α-diethyl glycine (Deg), α,α-dipropyl glycine (Dpg), α,α-di-isobutyl glycine (Dibg), α,α-dihexyl glycine (Dhg), α,α-diphenyl glycine (DΦg), α,α-dibenzyl glycine (Db$_z$g), α,α-cyclohexyl glycine (Ac$_6$c), and α,α-dihydroxymethyl glycine (Dmg). These α,α-dialkyl glycines form a heterogeneous group of amino acids that enable the design of peptidomimetics with different degrees of amphipathicity and structural behavior. This series of peptides includes five nonpolar aliphatic amino acids of different size and volume (Aib, Deg, Dpg, Dibg, Dhg), one cyclic amino acid (Ac$_6$c), two amino acids with aromatic side chains (DΦg and Db$_z$g), and one polar aliphatic amino acid (Dmg).

The incorporation of Aib in peptides has been extensively investigated in the past decades due to its ability to induce α-helix conformation,[19–22] and it has been observed that Aib stabilizes a type II β-turn in small peptides (2–4 residues) and 3_{10}-helix in peptides with 4–6 residues.[23,24] Deg and Dpg induce both fully extended C$_5$ conformations and helical conformations in crystal structures.[23,24] Dibg has also been synthesized previously,[25] but there are few results about its structure. It is expected that, in the same way as the others amino acids with two or more carbons in each side chain (Deg, Dpg, Db$_z$g and DΦg), Dibg also prefers an extended conformation of C$_5$ type.[26]

Experimental and theoretical investigations indicate that the noncanonical amino acids DΦg and Db$_z$g induce C$_5$ and C$_7$ backbone conformations.[26–29] Previous results about cyclic amino acids such as Ac$_6$c (Ac$_n$c, $n > 3$) suggest that the C$\alpha \leftrightarrow$ Cα cyclization constrains the main chain dihedrals, even more

than the constraint resulting of the double substitution at the Cα, as in Aib.[30,31]

Our case study, Alamethicin, is a peptaibol with known antimicrobial activity isolated from the fungus *Trichoderma viride*, and its structure was studied by X-ray diffraction by Fox and Richards.[32] It consists of a sequence with 19 residues (Ac-Aib-L-Pro-Aib-L-Ala-Aib-L-Ala-L-Gln-Aib-L-Val-Aib-Gly-L-Leu-Aib-L-Pro-L-Val-Aib-Aib-L-Glu-L-Gln-Phe), including eight Aib residues at positions 1, 3, 5, 8, 10, 13, 16, and 17.[33] This peptide can be an alternative to conventional antibiotics,[34,35] affecting the membrane permeability and leading to cell death due to osmotic shock and leakage of intracellular material.[16,36,37]

Modeling studies done by Tieleman and co-workers employing molecular dynamics simulations, investigated the structural and dynamic properties of Alamethicin in water, methanol, and the phosphatidylcholine bilayer membrane.[38,39] The authors have found substantial loss of structure in aqueous environment, especially at the C-terminal segment of the peptide. Furthermore, the formation of channels was investigated in three studies from Tieleman and co-workers that evaluate the most stable Alamethicin bundles consisting of 4, 5, 6, 7, or 8 helices.[39–41] It was observed in the bilayer and methanol that Alamethicin underwent partial loss of structure about its central Gly-X-X-Pro sequence motif.

Alamethicin has been the most investigated Peptaibol. Before the contributions of Tieleman and co-workers, MD simulations had been employed to analyze this Peptaibol in other solvent environments.[42,43] Modeling studies in methanol and chloroform suggest that the Alamethicin structure is mostly α-helical, but it can present some residues organized in the 3_{10}-helix form, from the tenth residue.[36,43–46] Also, the C-terminal loses its initial helical structure, presenting more flexibility than the N-terminal and the central part of the peptide.[36,38,39,43–46]

Alamethicin channels were predominantly investigated by theoretical approaches.[47,48] Fox and Richards[32] and other

researchers[49−53] suggest the formation of a barrel-stave channel with Alamethicin monomers using experimental techniques. Further insights were also obtained through molecular dynamics methods in different membrane models (POPC, DMPC, DOPC, DMPC/DHPC).[40,54−57]

MATERIALS AND METHODS

Noncanonical Amino Acid FF Parameters. The three-dimensional structure of the new noncanonical amino acids (Figure 1) were designed with the program Pymol.[58] The GROMOS topologies (bonded and nonbonded parameters) for each amino acid were transferred from the corresponding natural amino acids parametrized with the GROMOS 54a7 force field (FF).[59,60] (Supporting Information) provides the bonded and nonbonded parameters of each noncanonical amino acid under study using the FF GROMOS 54a7 syntax (Table 1S).

System Preparation. The X-ray structure of Alamethicin used in this study is available in the Protein Data Bank,[61] with the code 1AMT.[32] We created nine Alamethicin peptidomimetics by replacing all eight Aib residues by one of the new α,α-dialkyl glycines (Figure 1) and Ala. These peptidomimetics were named by the acronym of the new α,α-dialkyl glycine that was inserted.

The new Alamethicin peptidomimetics were modeled in water with the simple point charge (SPC) water model,[62] ethanol, and POPC membranes. These three solvents allow the evaluation of the peptidomimetics structure in solvents of different polarities and molecular environments. In water, Alamethicin and its mimetics were simulated in a dodecahedral box considering a hydration layer of at least 1 nm between the peptide and the walls, in all three directions. Thus, the systems have about 3300−3500 water molecules. In ethanol, the systems were modeled in a cubic box, with dimensions of 7 × 7 × 7 (nm) and containing approximately 3300 molecules of ethanol. In both media, the systems were neutralized with the addition of two Na$^+$ ions.

Peptide simulations in membrane were done using a POPC membrane composed of 128 phospholipids, previously equilibrated with water.[63] Each peptide (Alamethicin and peptidomimetics) was manually inserted in a transmembrane orientation into the equilibrated POPC membrane. It was necessary to remove three phospholipids of each monolayer to minimize collisions with the peptides. This procedure yielded 10 different peptide POPC systems; system 1: the native Alamethicin; system 2: the Alamethicin analog carrying Ala in the native Aib positions; systems 3 to 10: eight Alamethicin mimetics resulting from the insertion of the eight noncanonical α,α-dialkyl glycines.

Molecular Dynamics Simulations. All simulations were performed using the GROMACS 4.0.5 version.[64,65] For the treatment of long-range interactions, we used the Reaction Field method, with a cutoff of 1.4 nm and dielectric constant of 54 for SPC water model[62,66] and 24.3 for ethanol.[67−69] The van der Waals interactions were also truncated with a twin-range cutoff of 0.8 and 1.4 nm. The algorithm LINCS[70,71] was used to constrain the chemical bonds of the peptides and the algorithm SETTLE[72] in the case of water. The pressure and temperature Berendsen algorithms were used to control the temperature and pressure at 300 K and 1 atm, respectively.[73] In water and ethanol, we used a coupling constant of $\tau_T = 0.1$ ps and $\tau_P = 0.5$ ps, respectively, and in POPC these parameters were $\tau_T = 0.2$ ps and $\tau_P = 1.0$ ps.

In all systems (peptide in water, ethanol, and POPC membranes), three steps of energy minimization were performed. In the first two steps of energy minimization, position restraints (with force constant of 1000 kJ·mol^{-1}·nm^{-2}) were applied to all heavy atoms of the peptide and afterward on the main chain. In the third step of energy minimization, no position restraints were applied. Two molecular dynamics simulations of 100 ps were done with position restraints (force constant of 1000 kJ·mol^{-1}·nm^{-2}) on the heavy atoms and afterward on the main chain. The systems were equilibrated and sampled using 100 ns molecular dynamics simulations with an integration interval of 2 fs. To ensure a better sampling of the conformational states of these peptides in water and ethanol, five replicates of each system were done. Conformations were recorded every 1 ps.

Free Energy Calculations. To evaluate the relative free energy cost ($\Delta\Delta G^{1-2}$, see Figure 2) of replacing each Aib

$$\alpha,\alpha\text{-dialkyl glycine}^{vacuo} \xrightarrow{\Delta G_1} \text{Aib}^{vacuo}$$

$$\Delta G_3 \downarrow \qquad\qquad \downarrow \Delta G_4$$

$$\alpha,\alpha\text{-dialkyl glycine}^{Alm} \xrightarrow{\Delta G_2} \text{Aib}^{Alm}$$

$$\Delta G_1 + \Delta G_4 - \Delta G_2 - \Delta G_3 = 0$$
$$\Delta G_1 - \Delta G_2 = \Delta G_3 - \Delta G_4$$
$$\Delta\Delta G^{3-4} = \Delta G_3 - \Delta G_4$$
$$\Delta\Delta G^{1-2} = \Delta G_1 - \Delta G_2$$

$$\Delta\Delta G^{3-4} = \Delta\Delta G^{1-2} < 0 \longrightarrow \alpha,\alpha\text{-dialkyl glycine}$$
$$\Delta\Delta G^{3-4} = \Delta\Delta G^{1-2} > 0 \longrightarrow \text{Aib}$$
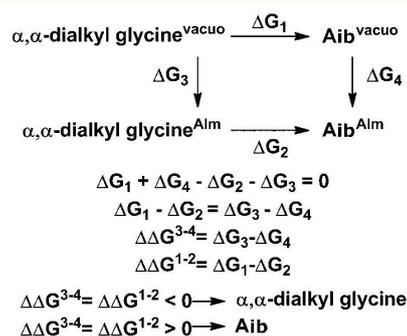
Figure 2. Thermodynamic cycle and equations used to determine the relative free energy of insertion ($\Delta\Delta G^{3-4}$) of a new α,α-dialkyl glycine relative to every Aib position in the native Alamethicin. ΔG_1 and ΔG_2 are the free energy resulting from the conversion of an α,α-dialkyl glycine into Aib in vacuo and in the Alamethicin peptide, respectively. ΔG_3 and ΔG_4 correspond to the free energy of incorporating an α,α-dialkyl glycine or Aib, respectively, into the Alamethicin peptide.

position of Alamethicin inserted on a POPC membrane, by a new noncanonical amino acid, we performed free energy perturbation (FEP) experiments using the Thermodynamic Integration (TI) technique.[74−76] This $\Delta\Delta G^{1-2}$ will measure and allow the comparison of the relative thermodynamic stability of each new noncanonical amino acid at every Aib position and indicate its contribution for the peptide stability. Negative $\Delta\Delta G^{1-2}$ indicates that it is thermodynamically favorable to replace Aib by a given noncanonical amino acid; a positive value indicates the opposite. In this approach, 21 intermediate Hamiltonian states separating the initial and final state were simulated using a coupling parameter λ. The relative free energy was given by the integration of the Hamiltonian derivative relative to the coupling parameter (λ) that connects the initial and final states. The trapezoidal rule was employed for this integration.

The TI experiments of Alamethicin in membrane, comprised the alchemical mutation of each eight new α,α-dialkyl glycines and, in separate FEP calculations, into native Aib residue (ΔG_2, see Figures 1 and 2) . The same alchemical transformation of each α,α-dialkyl glycine into Aib residue was also made in vacuo, to complete the necessary thermodynamic cycle (ΔG_1, see Figures 1 and 2). We choose to use vacuo in order to obtain a relative free energy solely correlated to the mutation of Aib
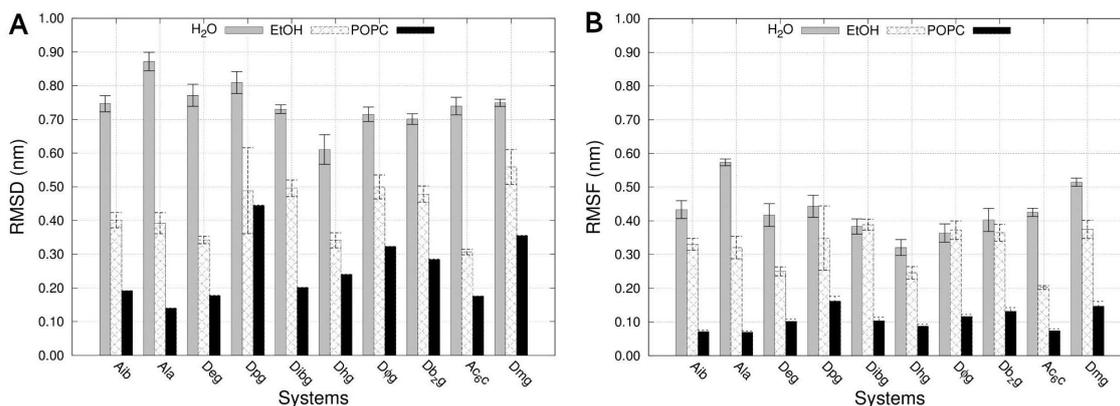
Figure 3. (A) Cα RMSD and (B) Cα RMSF averages for Alamethicin and peptidomimetics, in water, ethanol, and POPC. Fitting of Cα relative to the experimental X-ray structure of Alamethicin in α-helix. Average values obtained from five replicate simulations (in water and ethanol), including standard deviation error bars.

into each noncanonical amino acid without any solvent effect. For the purpose of this evaluation, this approach is sufficient. With these experiments, it was possible to evaluate the free energy of insertion of each α,α-dialkyl glycine in every Alamethicin position in a membrane environment. In total, with this protocol, it was possible to study the thermodynamic properties of 72 Alamethicin mimetics inserted on the membrane. The coupling parameter λ was varied from 0 to 1, with incremental steps of 0.05 λ for each simulation, resulting in 21 simulations for each of the 72 new systems (these systems comes from individual substitution of the 8 Aib positions for one of the 8 new α,α-dialkyl glycines or Ala, resulting in 9 residues for 8 possible positions = 72 different peptides). We used an integration interval of 2 fs and simulations of 10 ns sampling for each of the 21 λ points, resulting in a total sampling time of 210 ns.

Analysis. RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuation), SS (Secondary Structure Analysis), and Ramachandran plots analysis were performed over all the conformations from the 100 ns simulations.[65] All measurements are averaged over five replicates, and the corresponding standard deviation is presented. Average structures (central conformations) shown on figures are the ones that minimize the RMSD variance when used for fitting against all other conformations of the trajectory.

■ RESULTS AND DISCUSSION

The 2-dimensional structures of all α,α-dialkyl glycines investigated in this work are shown in Figure 1. Aib and Ala were used as reference residues. Deg, Dpg, Dibg, and Dhg are noncanonical amino acids with nonpolar, aliphatic side chains. DΦg and Db$_z$g are disubstituted amino acids with aromatic side chains, Ac$_6$c has a cyclic side chain and Dmg is the only noncanonical amino acid under study with aliphatic, polar side chain. Some noncanonical amino acids of this collection are similar to natural amino acids. That is the case for amino acids Aib, Dibg, Dmg, and Db$_z$g, which are similar to alanine, leucine, serine, and phenylalanine, respectively. These disubstituted amino acids lose their chirality because of the addition of the second symmetrical side chain.

Figure 3 shows the Cα RMSD and RMSF averages of each Alamethicin system in three different environments (water,

ethanol, and POPC) relative to the experimental X-ray structure of Alamethicin in α-helix. In Figure 3A, the peptide with the highest RMSD relative to the native α-helix structure is the one substituted by Ala in all Alamethicin Aib positions. This result is a first indication that the Ala amino acid in Alamethicin is unable to promote preorganization in α-helix. Furthermore, the native form of Alamethicin loses a substantial part of its helical structure in water, as observed by Tieleman and co-workers.[38] On the other hand, in water, Dhg appears to promote a conformation more native-like relative to the reference structure in α-helix than Aib, indicating that this amino acid may induce α-helical conformations. The peptidomimetics containing the amino acids Dibg, DΦg, Db$_z$g, Ac$_6$c, and Dmg were shown to be poorly structured, similarly to native Alamethicin in water. Also, in all solvents, Alamethicin substituted by Dmg does not seem to have conformations similar to the native structure in α-helix.

All Alamethicin peptidomimetics structure (including Ala) are more native-like when solvated in ethanol (Figure 3A). This fact seems to correlate with the preference of these peptides to adopt α-helical structures in low polar environments. Ethanol was used here to evaluate the structure of the peptides in a media with an intermediate dielectric constant between the water and POPC membrane. We observe that Alamethicin substituted by Aib, Ala, Dpg, Dibg, DΦg, and Db$_z$g present high RMSD values in ethanol, and the ones carrying the residues Ac$_6$c, Dhg, and Deg have low deviation in ethanol. Note that Dhg induces low RMSD both in water and in ethanol.

In POPC (Figure 3A), it is observed that the less bulky residues, such as Aib, Ala, Deg and Ac$_6$c promote lower RMSD in comparison to the amino acids with large side chains. This behavior might be related to the fact that these small residues are well-arranged between phospholipids chains, and therefore, they do not suffer large structural rearrangement when inserted on the membrane. On the other hand, the peptide-containing residues with longer and bulky side chains must need to rearrange these amino acids between the phospholipid chains to minimize steric hindrance, causing some structural perturbation on the native helical conformation of the peptides.

The conformational sampling of the peptides under study was evaluated using a RMSF analysis. Figure 3B shows that, in water, most peptides bearing a noncanonical disubstituted

amino acid are more constrained than the peptide with Ala in all Aib positions. It is also apparent that the amino acid that exhibits the lowest RMSF in aqueous environment is Dhg. The other peptides have similar RMSF relative to the native Alamethicin. The RMSF behavior of the peptides with Dibg, DΦg, and Db$_z$g, in ethanol is very similar to what is observed in water. However, the general trend seems to be the reduction of the amplitude of the peptide RMSF in this medium.

In POPC, peptides containing residues of Aib, Ala, Dhg, and Ac$_6$c suffer the smallest structural fluctuations. Despite the large and bulky side chain of Dhg, this amino acid is capable to induce lower fluctuations when the peptide is inserted in the membrane, similarly to what is observed for the smallest amino acids discussed in this work, Ala and Aib. These facts suggest that peptides bearing Dhg are, in general, more conformationally restrained in different environments than most of the other residues under study.

The secondary structure analysis was used to determine and quantify the type of secondary structure conformations explored by these peptides and the number of residues that are involved in a particular type of secondary structure. Figure 4
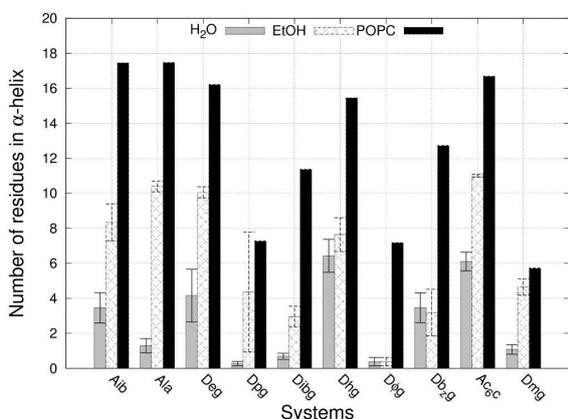
Figure 4. Average number of residues in α-helix of all peptides in water, ethanol, and POPC. Average values obtained from five replicate simulations (in water and ethanol), including standard deviation error bars.

shows the average number (over five replicates) of residues involved in α-helix throughout the 100 ns simulation, obtained for each peptide investigated in water, ethanol, and POPC.

It was noted earlier that Ac$_6$c and Dhg are more capable of inducing Alamethicin conformations closer to the native X-ray structure in α-helix than peptides with Aib or Ala (Figure 3A). This fact is confirmed by our secondary structure analysis. We show in Figure 4 that, in water, the peptides with higher number of amino acids in α-helix are those containing Ac$_6$c and Dhg.

In water, the analogue containing Ala presents an average of less than two residues in α-helix. The analogues containing Deg and Db$_z$g have an average of four residues in α-helix, whereas the analogues containing Dpg, Dibg, DΦg, and Dmg have a residual number of amino acids in α-helix.

In POPC, most of the peptides show a high number of residues in α-helix. This observation agrees with the results of Tieleman and co-workers about the capability of the native Alamethicin in maintaining or reorganizing the conformational

structure when near or inserted in a membrane.[38] Only the amino acids Dpg, DΦg, and Dmg have lower tendency to induce this type of secondary structure in this medium. Ethanol behaves as a medium with intermediate properties between water and POPC, because there is a considerable increase in the number of residues in α-helix, relative to those seen in water.

To illustrate the previous analyses, Figure 5 shows the central structures (see Material and Methods) of all peptides in water. This analysis provides the most representative structure of the ensemble of conformations sampled during the simulation. It is clear that the most of them have lost their initial helical structure, except for the case of Dhg (Figure 5F).
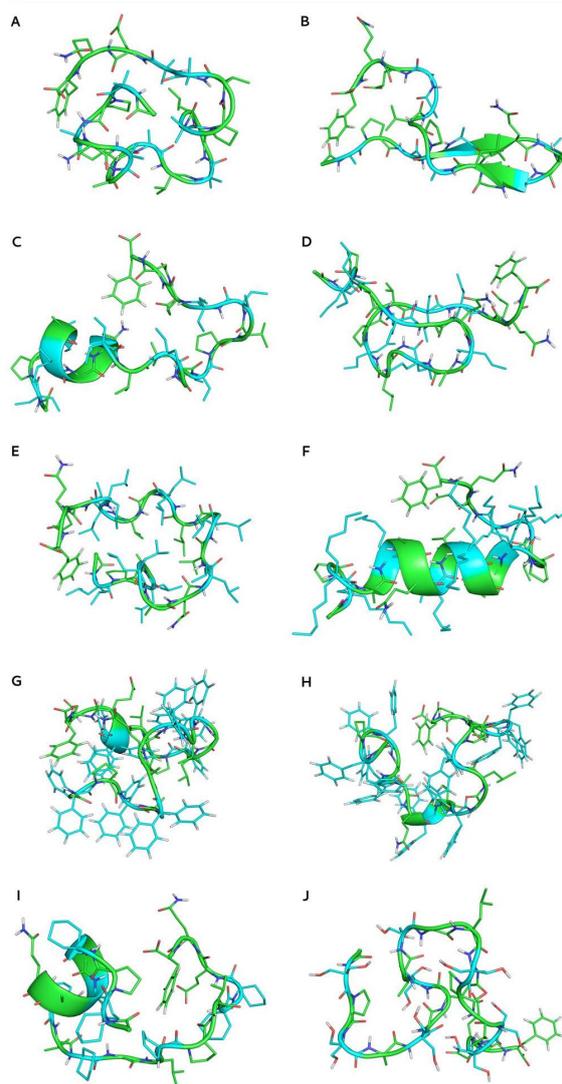
Figure 5. Central structures of one replicate of Alamethicin and analogues in water: (A) Aib, (B) Ala, (C) Deg, (D) Dpg, (E) Dibg, (F) Dhg, (G) DΦg, (H) Db$_z$g, (I) Ac$_6$c, and (J) Dmg. The coloring of the atoms follows the convention: green for carbon, blue for nitrogen, red for oxygen, white for hydrogen, and cyan to highlight amino acid of interest. The water molecules were omitted for better visualization, and peptides show the cartoon that defines its secondary structure.
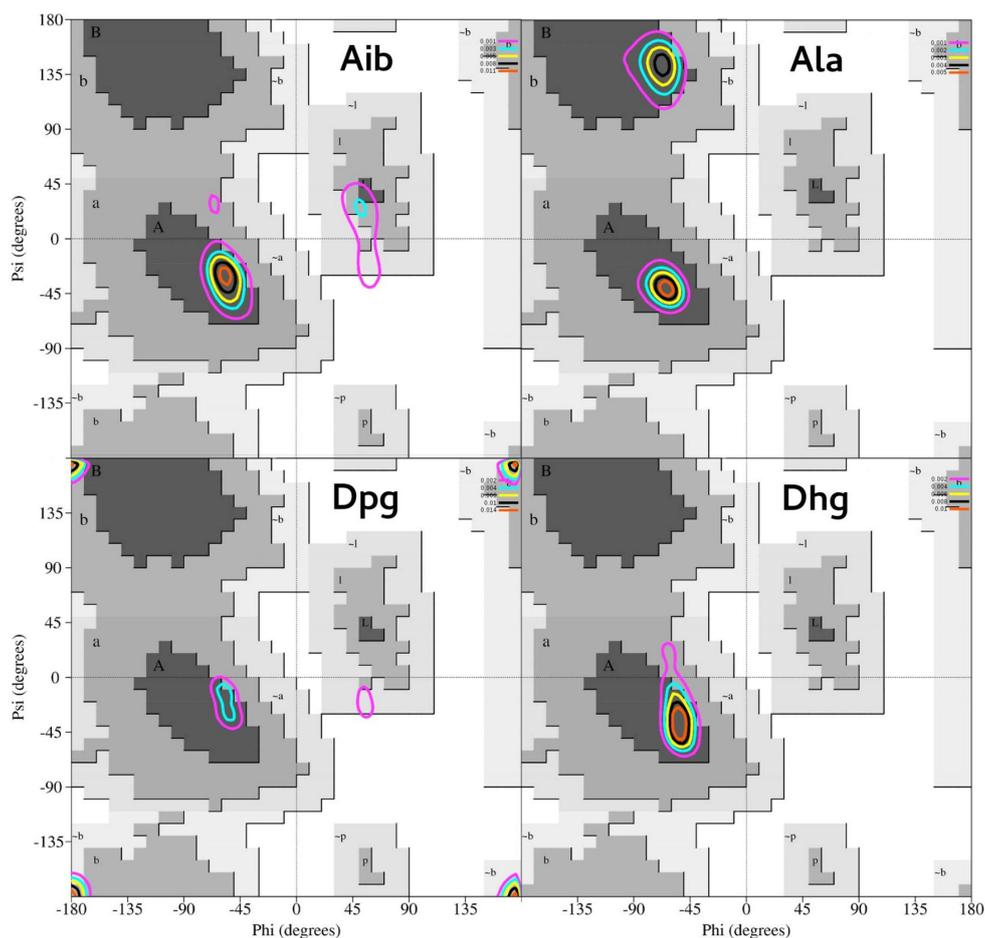
Figure 6. Probability density contours of φ and ψ pairs for the amino acids Aib, Ala, Dpg, and Dhg in water. These contours are superimposed on the Ramachandran diagram in which region (A) corresponds to typical dihedrals of right α-helix, (B) corresponds to β-sheets space, and (L) to left α-helix region.

The conformations observed in Figure 5 confirm the previous observation made on the RMSD analysis and number of residues in the α-helix (Figures 3A and 4). It is visible clear that the Dhg residue induces α-helical conformations of Alamethicin in water (Figure 5F), and the native peptide has lost its helical conformation (Figure 5A). Experiments conducted by Tieleman et al. demonstrate that the loss of structure in water for the native Alamethicin (Figure 3A) can be explained by the hydrophobic effect. The authors also demonstrate that the peptide reorganizes into a α-helical conformation at the membrane water interface.[38] A similar behavior for the new peptidomimetics is expected, and in fact, the implemented analysis has shown that a great part of these analogues are unstructured in water but maintain a helical conformation when inserted in the membrane. It is interesting to note that the analogue containing Ala (Figure 5B) completely loses the initial structure in α helix and promotes the formation of an antiparallel β-sheet during the simulation in water.

The conformations of analogues containing Dpg, Dibg, DΦg, and Dmg (Figure 5D, E, G, and J, respectively) suggest the formation of random coil structures. It is clear that the peptide

containing Dhg (Figure 5F) has most of its residues in α-helix (approximately half of the residues of this peptide), and only the amine and carboxyl terminal are unstructured. This α-helix preorganization suggests that the insertion of the peptide in the membrane will be thermodynamically less costly compared to a unstructured peptide. In principle, some peptides (for example, Ala, Deg, or Dibg) will require a higher reorganization cost to adopt an α-helix conformation in POPC (see Figure 4). The peptide containing the cyclic amino acid Ac$_6$c (Figure 5I) has also a α-helix region at its N-terminal. It is, therefore, an amino acid less bulky than the Dhg residue that also promotes the formation of α-helical conformations.

Within the context of structural analysis, we investigated the dihedral angle pairs Psi (ψ) and Phi (φ) of the Ramachandran space for each α,α-dialkyl inserted in Alamethicin.[77] This type of analysis is essential to understand the backbone degrees of freedom and secondary structure of each amino acid compared to the natural amino acids. Disubstituted amino acids have two symmetric and sometimes bulky side chains constraining the amino acid structure around the Cα. Therefore, the dihedral conformational space of these disubstituted amino acids might adopt conformations that lie outside of the classical

Article

Ramachandran plot regions of canonical amino acids. Figure 6 shows the probability density of the φ and ψ pairs of four amino acids of interest obtained from our simulation of Alamethicin in water, and they are Ala, Aib, Dpg, and Dhg. The distribution was calculated from the ψ and φ angles recorded from the eight Alamethicin Aib positions replaced by Ala and the α,α-dialkyl glycines. A total of 4 000 000 points were used to calculate the probability densities shown in Figure 6 (5 replicate simulations $\times$ 8 residue positions $\times$ 100 000 conformations). The Ramachandran diagrams for the others amino acids discussed in this work are shown in the Supporting Information (Figure 1S).

Figure 6 shows that the highest density of φ and ψ dihedral pairs obtained for Aib (Figure 6 − upper left) inserted in Alamethicin are in regions corresponding, as expected, to left and right α-helices, with higher preference for the right α-helix region. Aib is a symmetric amino acid, where the carbon alpha has no chirality, and consequently, it is neither an L- or D-amino acid. This fact has clear consequences on the propensity to sample both left and right regions of the Ramachandran plot. Another important observation is that the probability density observed for Aib, when compared to those obtained by Ala (Figure 6 − upper right), also confirms that the Aib is more suitable for constraining the peptide structure in α-helix. Note that, as expected, Ala explores dihedral pairs in the region of β sheets. We show that the double-methylation at Cα on Aib eliminates completely the conformations in the β-sheet Ramachandran space.

The Dhg residue (Figure 6 − lower right) has φ and ψ pairs only in the right α-helix region, suggesting that it is not possible to establish a correlation between the lack of chirality with the propensity to induce both left and/or right α-helices. The new amino acid Dpg (Figure 6 − lower left) has dihedral angles scattered at 180°, indicating the possibility of the arrangement in extended conformation as previously suggested by Valle and co-workers for amino acids with two or more carbons in the branched side chain.[23,24,26] The results presented by the Aib, Dpg, and Dhg residues clearly indicate that disubstituted amino acids constitute a diverse class of new residues with great conformational variability that are not exclusively in α-helix conformations.

So far, the structural and dynamics findings presented above, suggested that some new noncanonical amino acids, such as Dhg and Ac$_6$c, are able to induce peptides to adopt helical secondary structures compared to Aib in native Alamethicin. However, from a thermodynamic point of view, it is important to evaluate the relative free energy cost of replacing each Aib for a new α,α-dialkyl glycine. This aspect is relevant, taking into account the function of Alamethicin in the insertion and disruption of cell membranes. We evaluated the relative free energy cost of replacing each Aib position by a new α,α-dialkyl glycine in Alamethicin inserted in a membrane environment. This was accomplished using the thermodynamic cycle shown in Figure 2 (see Material and Methods).

In this thermodynamic cycle, ΔG_1 corresponds to the free energy associated with the transformation of an α,α-dialkyl glycine to Aib in vacuo, whereas ΔG_2 refers to the free energy of the transformation of an α,α-dialkyl glycine to Aib insert in Alamethicin. In this thermodynamic cycle, we do not want to evaluate any solvent effect, and for this reason, we chose to close the thermodynamic cycle with an alchemical trans- formation in vacuo. In this thermodynamic cycle, if $\Delta\Delta G^{3-4} < 0$, it is thermodynamically favorable to replace Aib by a new

α,α-dialkyl glycine in this position, and if $\Delta\Delta G^{1-2} > 0$, it is preferable to maintain the native Aib. The relative free energy values of insertion of a new α,α-dialkyl glycine at each position previously occupied by Aib are show in Figure 7.
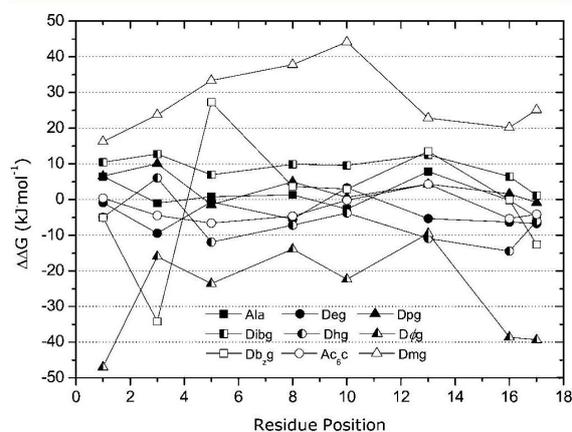
Figure 7. Relative free energy of insertion ($\Delta\Delta G^{3-4}$) of the new non- natural amino acids in Alamethicin, in each position previously occupied by Aib (1, 3, 5, 8, 10, 13, 16, and 17), in the POPC membrane. Negative free energy values indicate the preference toward the noncanonical amino acid.

In Figure 7, it is observed that for all positions of interest (1, 3, 5, 8, 10, 13, 16, and 17), we found alternative residues that are thermodynamically more favorable than Aib. DΦg seems to be one of the best amino acids to replace the native positions belonging to Aib in six out of the eight positions. It is also important to note that in the position 10, the amino acids Dhg, Ala, Dpg, Db$_z$g, and Deg lead to similar relative free energies. This thermodynamic data can also be used to design a novel Alamethicin peptidomimetics with improved thermodynamic stability in membrane environments.

Taking into account only the thermodynamic data, the best options for replacing Aib at the eight positions are DΦg, Db$_z$g, DΦg, DΦg, Dhg, DΦg, and DΦg. However, to suggest a novel peptidomimetic for Alamethicin, we may also take into account the structural and dynamics properties previously observed for each new α,α-dialkyl glycine. The previous analysis indicates that the choice of DΦg at positions 1, 5, 8, 10, 16, and 17 might not be the best option because it did not promote good α- helical preorganization of Alamethicin in water and in the POPC membrane.

We can suggest that the amino acid sequence to replace all the Aib positions that combine the best thermodynamic and α- helical propensity in Alamethicin are Dhg, Deg, Dhg, Dhg, Aib, Dhg, Dhg, and Deg, replacing positions 1, 3, 5, 8, 10, 13, 16, and 17 in Alamethicin. This suggestion of amino acids reflects the best combination of structural characteristics and thermodynamic properties for Alamethicin. In this regard, this new peptide was modeled and evaluated under the same simulation conditions (in water and POPC). A summary of the structural properties of this peptide is compared with some of the previous Alamethicin peptidomimetics in Figure 8. Figure 8A shows that the structure of the new peptide is more native- like than Alamethicin with Aib, Deg, or Ac$_6$c and it clearly adopts conformations with α-helix organization (Figure 8C). It is also interesting to note that the RMSD of this new peptide

99

Tarsila G. Castro and Nuno M. Micaêlo.:dx.doi.org/10.1021/jp505400q | J. Phys. Chem. B 2014, 118, 9861–9870.
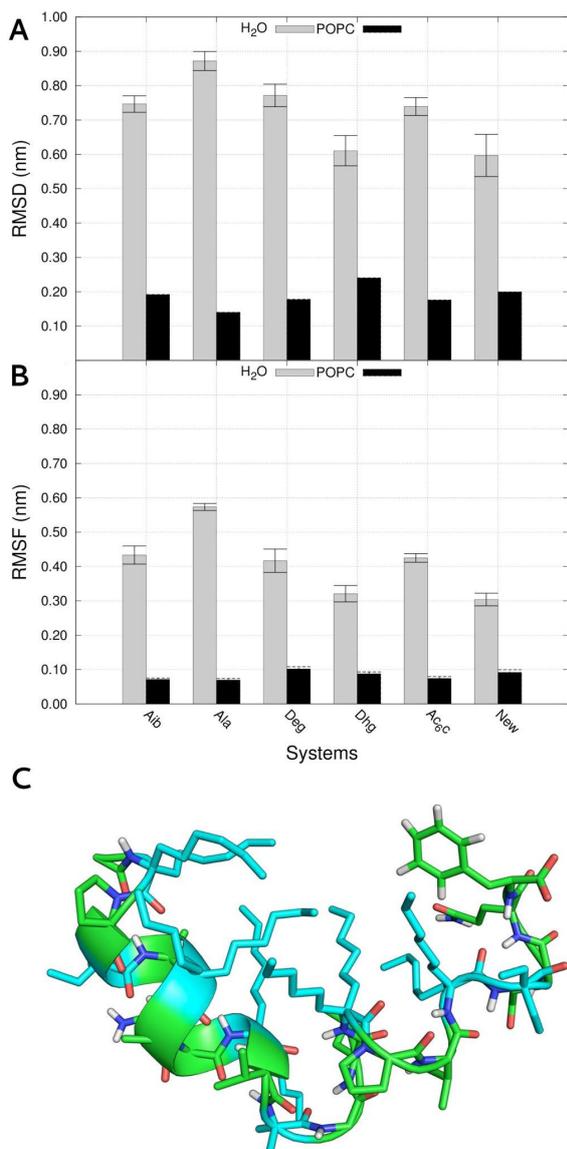
Figure 8. (A) Cα RMSD and (B) Cα RMSF averages for Alamethicin, peptidomimetics with Ala, Deg, Dhg, and Ac$_6$c and for the new Alamethicin peptide (New), in water and POPC. Fitting of Cα relative to the experimental X-ray structure of Alamethicin in α-helix. Average values obtained from five replicate simulations (in water), including standard deviation error bars. (C) Central structure of one replicate of the new Alamethicin analogue.

and the Alamethicin substituted by Dhg are statistically equivalent. This means that the expected conformational constraining imposed by the Dhg residue was observed on this new peptidomimetic. In POPC (Figure 8A), it is apparent that the structural deviation of the new peptide is equivalent to the other Alamethicin analogues. A similar observation can be made for the dynamic properties compared to Alamethicin with Aib, Deg, Dhg, and Ac$_6$c residues (Figure 8B). The chosen amino seems to favor the stabilization of the peptide in α helix, otherwise, we would have found higher values of RMSD and

RMSF, as well as a less folded central structure. We cannot tell yet if there are any positive, negative or cancelation of both correlations effects between the chosen residues. However, the low RMSD and RMSF seem to indicate the absence of negative correlations.

CONCLUSIONS

The modeling studies of the α,α-dialkyl glycines using Alamethicin as a model peptaibol provided significant results about the structure and function of the new noncanonical amino acids here proposed. In water, Dhg, DΦg, and Db$_z$g impose more constrained and helical structures than Aib. The lack of chirality around Cα and bulky side chains of these amino acids must be responsible for this effect. In ethanol, Deg, Ac$_6$c, and Dhg are the amino acids that induce higher peptide helicity. In POPC, Ala, Deg, and Ac$_6$c rendered analogues with structural behavior similar to native Alamethicin. In this environment, it was noted that smaller amino acids are well-arranged between phospholipids chains, and therefore, the peptides do not suffer large structural rearrangement.

The analyses implemented indicate that Dhg and Ac$_6$c are the ones more capable of inducing α-helix conformations in Alamethicin, in all solvents under study. Moreover, they seem to improve the thermodynamic stabilization of the peptides in the membrane. This result is consistent with our prior study in Peptaibolin[18] where these residues were the most capable to induce helical conformations. This indicates that these residues seem to have a foldamer profile; however, further experiments using different peptides are required to propose a definitive conclusion about their foldamer role.

Ramachandran analysis demonstrated that the disubstituted amino acids do not only induce α-helix conformations. Dpg and Db$_z$g may prefer an extended conformation, and this fact agrees with previous results.[23,24,26] The α,α-dialkyl glycines show to have different propensities to induce secondary structures, particularly right and left α-helix, β-sheets, and planar structures. These residues seem to constitute a class of amino acids with great conformational variability, not restricted to α-helical conformations. The foldamer potential of this class of amino acids needs to be further evaluated in future studies using other peptides. This is necessary to evaluate if other peptide context affects the conformational properties observed in this study. The relative free energy of replacing Aib by a new α,α-dialkyl glycine suggests that there are better alternatives to Aib in almost all positions previously occupied by Aib, except at position 11. We proposed a new analogue of Alamethicin that combines the best structural, dynamic, and thermodynamics properties. The modeling of Alamethicin analogues by inserting new α,α-dialkyl glycines suggests that it is possible to optimize the characteristics of native Alamethicin and obtain novel peptides that may have improved antibiotic activity.

ASSOCIATED CONTENT

Ⓢ Supporting Information

All parametrizations for the new amino acids discussed in this article and Ramachandran plots are available as Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: micaelo@itqb.unl.pt. Tel.: +351 253 604 384.

The Journal of Physical Chemistry B

■ **REFERENCES**

(1) Fu, Y. W.; Hammarstrom, L. G. J.; Miller, T. J.; Fronczek, F. R.; McLaughlin, M. L.; Hammer, R. P. Sterically hindered C-alpha,C-alpha-disubstituted alpha-amino acids: Synthesis from alpha-nitro-acetate and incorporation into peptides. *J. Org. Chem.* **2001**, *66*, 7118−7124.

(2) Mendel, D.; Ellman, J.; Schultz, P. G. Protein-biosynthesis with conformationally restricted amino-acids. *J. Am. Chem. Soc.* **1993**, *115*, 4359−4360.

(3) Gentilucci, L.; Tolomelli, A.; Squassabia, F. Peptides and peptidomimetics in medicine, surgery and biotechnology. *Curr. Med. Chem.* **2006**, *13*, 2449−2466.

(4) Giannis, A. Peptidomimetics for receptor ligand discovery, development, and medical perspectives. *Angew. Chem., Int. Ed.* **1993**, *32*, 1244−1267.

(5) Grauer, A.; Konig, B. Peptidomimetics—A Versatile Route to Biologically Active Compounds. *Eur. J. Org. Chem.* **2009**, 5099−5111.

(6) Vagner, J.; Qu, H. C.; Hruby, V. J. Peptidomimetics, a synthetic tool of drug discovery. *Curr. Opin. Chem. Biol.* **2008**, *12*, 292−296.

(7) Ballet, S.; Feytens, D.; Buysse, K.; Chung, N. N.; Lemieux, C.; Tumati, S.; Keresztes, A.; Van Duppen, J.; Lai, J.; Varga, E.; Porreca, F.; Schiller, P. W.; Broeck, J. V.; Tourwe, D. Design of Novel Neurokinin 1 Receptor Antagonists Based on Conformationally Constrained Aromatic Amino Acids and Discovery of a Potent Chimeric Opioid Agonist-Neurokinin 1 Receptor Antagonist. *J. Med. Chem.* **2011**, *54*, 2467−2476.

(8) Feytens, D.; Cescato, R.; Reubi, J. C.; Tourwe, D. New sst(4/5)-selective somatostatin peptidomimetics based on a constrained tryptophan scaffold. *J. Med. Chem.* **2007**, *50*, 3397−3401.

(9) Mallareddy, J. R.; Borics, A.; Keresztes, A.; Kover, K. E.; Tourwe, D.; Toth, G. Design, Synthesis, Pharmacological Evaluation, and Structure−Activity Study of Novel Endomorphin Analogues with Multiple Structural Modifications. *J. Med. Chem.* **2011**, *54*, 1462−1472.

(10) Oh, J. E.; Lee, K. H. Synthesis of novel unnatural amino acid as a building block and its incorporation into an antimicrobial peptide. *Bioorgan. Med. Chem.* **1999**, *7*, 2985−2990.

(11) Ressurreicao, A. S. M.; Bordessa, A.; Civera, M.; Belvisi, L.; Gennari, C.; Piarulli, U. Synthesis and conformational studies of peptidomimetics containing a new bifunctional diketopiperazine scaffold acting as a beta-hairpin inducer. *J. Org. Chem.* **2008**, *73*, 652−660.

(12) Whitby, L. R.; Ando, Y.; Setola, V.; Vogt, P. K.; Roth, B. L.; Boger, D. L. Design, Synthesis, and Validation of a beta-Turn Mimetic Library Targeting Protein−Protein and Peptide−Receptor Interactions. *J. Am. Chem. Soc.* **2011**, *133*, 10184−10194.

(13) Chugh, J. K.; Wallace, B. A. Peptaibols: models for ion channels. *Biochem. Soc. Trans.* **2001**, *29*, 565−570.

(14) Whitmore, L.; Wallace, B. A. The Peptaibol Database: a database for sequences and structures of naturally occurring peptaibols. *Nucleic Acids Res.* **2004**, *32*, D593−D594.

(15) Marr, A. K.; Gooderham, W. J.; Hancock, R. E. W. Antibacterial peptides for therapeutic use: obstacles and realistic outlook. *Curr. Opin. Pharmacol.* **2006**, *6*, 468−472.

(16) Shai, Y. Mode of action of membrane active antimicrobial peptides. *Biopolymers* **2002**, *66*, 236−248.

(17) Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **2002**, *415*, 389−395.

(18) Castro, T. G.; Micaêlo, N. M. Modeling of Peptaibol Analogues Incorporating Nonpolar α,α-Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation. *J. Phys. Chem. B* **2014**, *118*, 649−658.

(19) Marshall, G. R.; Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrocki, J.; Leplawy, M. T. Factors Governing Helical Preference of Peptides Containing Multiple alpha,alpha-Dialkyl Amino-Acids. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 487−491.

(20) Improta, R.; Barone, V.; Kudin, K. N.; Scuseria, G. E. Structure and Conformational Behavior of Biopolymers by Density Functional Calculations Employing Periodic Boundary Conditions. I. The Case of Polyglycine, Polyalanine, and Poly-α-aminoisobutyric Acid in Vacuo. *J. Am. Chem. Soc.* **2001**, *123*, 3311−3322.

(21) Karle, I. L.; Balaram, P. Structural Characteristics of Alpha-Helical Peptide Molecules Containing Aib Residues. *Biochemistry* **1990**, *29*, 6747−6756.

(22) Marshall, G. R.; Bosshard, H. E. Angiotensin II. Studies on the biologically active conformation. *Circ. Res.* **1972**, *31* (Suppl 2), 143−50.

(23) Schweitzer-Stenner, R.; Gonzales, W.; Bourne, G. T.; Feng, J. A.; Marshall, G. R. Conformational manifold of alpha-aminoisobutyric acid (Aib) containing alanine-based tripeptides in aqueous solution explored by vibrational spectroscopy, electronic circular dichroism spectroscopy, and molecular dynamics simulations. *J. Am. Chem. Soc.* **2007**, *129*, 13095−13109.

(24) Rodrigues, L. M.; Fonseca, J. I.; Maia, H. L. S. Synthesis and conformational investigation of tetrapeptide analogues of the fragment B23-B26 of insulin. *Tetrahedron* **2004**, *60*, 8929−8936.

(25) Fu, Y.; Etienne, M. A.; Hammer, R. P. Facile Synthesis of α,α-Diisobutylglycine and Anchoring Its Derivatives onto PAL-PEG-PS Resin. *J. Org. Chem.* **2003**, *68*, 9854−9857.

(26) Valle, G.; Crisma, M.; Bonora, G. M.; Toniolo, C.; Lelj, F.; Barone, V.; Fraternali, F.; Hardy, P. M.; Langrangoldsmith, A.; Maia, H. L. S. Structural versatility of peptides from C-alpha,alpha disubstituted glycines − preferred conformational of the C-alpha,alpha-dibenzylglycine residue. *J. Chem. Soc., Perkin Trans. 2* **1990**, 1481−1487.

(27) Casanovas, J.; Nussinov, R.; Alemàn, C. Intrinsic Conformational Preferences of Cα,α-Dibenzylglycine. *J. Org. Chem.* **2008**, *73*, 4205−4211.

(28) Casanovas, J.; Zanuy, D.; Nussinov, R.; Alemán, C. Intrinsic Conformational Characteristics of α,α-Diphenylglycine. *J. Org. Chem.* **2007**, *72*, 2174−2181.

(29) Pavone, V.; Lombardi, A.; Saviano, M.; De Simone, G.; Nastri, F.; Maglio, O.; Omote, Y.; Yamanaka, Y.; Yamada, T. Conformational behavior of C-alpha,C-alpha-diphenyl glycine: Extended conformation in tripeptides containing consecutive D phi G residues. *Biopolymers* **2000**, *53*, 161−168.

(30) Rodriguez-Ropero, F.; Zanuy, D.; Casanovas, J.; Nussinov, R.; Aleman, C. Application of 1-aminocyclohexane carboxylic acid to protein nanostructure computer design. *J. Chem. Inf. Model* **2008**, *48*, 333−343.

(31) Alemán, C. Conformational Properties of α-Amino Acids Disubstituted at the α-Carbon. *J. Phys. Chem. B* **1997**, *101*, 5046−5050.

(32) Fox, R. O.; Richards, F. M. A voltage-gated ion channel model inferred from the crystal-structure of Alamethicin at 1.5-A resolution. *Nature* **1982**, *300*, 325−330.

(33) Pandey, R. C.; Cook, J. C.; Rinehart, K. L. High-resolution and field desorption mass spectrometry studies and revised structures of Alamethicin I and Alamethicin II. *J. Am. Chem. Soc.* **1977**, *99*, 8469−8483.

(34) Bechinger, B. The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. *Biochim. Biophys. Acta, Biomembr.* **1999**, *1462*, 157−183.

(35) Bechinger, B.; Salnikov, E. S. The membrane interactions of antimicrobial peptides revealed by solid-state NMR spectroscopy. *Chem. Phys. Lipids* **2012**, *165*, 282−301.

101

Tarsila G. Castro and Nuno M. Micaêlo.:dx.doi.org/10.1021/jp505400q | J. Phys. Chem. B 2014, 118, 9861−9870.

Article

(36) Bak, M.; Bywater, R. P.; Hohwy, M.; Thomsen, J. K.; Adelhorst, K.; Jakobsen, H. J.; Sørensen, O. W.; Nielsen, N. C. Conformation of Alamethicin in Oriented Phospholipid Bilayers Determined by 15N Solid-State Nuclear Magnetic Resonance. *Biophys. J.* **2001**, *81*, 1684−1698.

(37) Salnikov, E. S.; De Zotti, M.; Formaggio, F.; Li, X.; Toniolo, C.; O'Neil, J. D. J.; Raap, J.; Dzuba, S. A.; Bechinger, B. Alamethicin Topology in Phospholipid Membranes by Oriented Solid-state NMR and EPR Spectroscopies: a Comparison. *J. Phys. Chem. B* **2009**, *113*, 3034−3042.

(38) Tieleman, D. P.; Berendsen, H. J. C.; Sansom, M. S. P. Surface binding of alamethicin stabilizes its helical structure: Molecular dynamics simulations. *Biophys. J.* **1999**, *76*, 3186−3191.

(39) Tieleman, D. P.; Sansom, M. S. P.; Berendsen, H. J. C. Alamethicin helices in a bilayer and in solution: Molecular dynamics simulations. *Biophys. J.* **1999**, *76*, 40−49.

(40) Tieleman, D. P.; Hess, B.; Sansom, M. S. P. Analysis and evaluation of channel models: Simulations of alamethicin. *Biophys. J.* **2002**, *83*, 2393−2407.

(41) Tieleman, D. P.; Breed, J.; Berendsen, H. J. C.; Sansom, M. S. P. Alamethicin channels in a membrane: molecular dynamics simulations. *Faraday Discuss.* **1998**, *111*, 209−223.

(42) Leitgeb, B.; Szekeres, A.; Manczinger, L.; Vágvölgyi, C.; Kredics, L. The History of Alamethicin: A Review of the Most Extensively Studied Peptaibol. *Chem. Biodivers.* **2007**, *4*, 1027−1051.

(43) Fraternali, F. Restrained and unrestrained molecular dynamics simulations in the NVT ensemble of alamethicin. *Biopolymers* **1990**, *30*, 1083−1099.

(44) Biggin, P. C.; Breed, J.; Son, H. S.; Sansom, M. S. P. Simulation Studies of Alamethicin−Bilayer Interactions. *Biophys. J.* **1997**, *72*, 627−636.

(45) Gibbs, N.; Sessions, R. B.; Williams, P. B.; Dempsey, C. E. Helix bending in alamethicin: molecular dynamics simulations and amide hydrogen exchange in methanol. *Biophys. J.* **1997**, *72*, 2490−2495.

(46) Sessions, R. B.; Gibbs, N.; Dempsey, C. E. Hydrogen Bonding in Helical Polypeptides from Molecular Dynamics Simulations and Amide Hydrogen Exchange Analysis: Alamethicin and Melittin in Methanol. *Biophys. J.* **1998**, *74*, 138−152.

(47) Thogersen, L.; Schiott, B.; Vosegaard, T.; Nielsen, N. C.; Tajkhorshid, E. Peptide Aggregation and Pore Formation in a Lipid Bilayer: A Combined Coarse-Grained and All Atom Molecular Dynamics Study. *Biophys. J.* **2008**, *95*, 4337−4347.

(48) Mihajlovic, M.; Lazaridis, T. Antimicrobial peptides in toroidal and cylindrical pores. *Biochim. Biophys. Acta, Biomembr.* **2010**, *1798*, 1485−1493.

(49) Pan, J. J.; Tristram-Nagle, S.; Nagle, J. F. Alamethicin Aggregation in Lipid Membranes. *J. Membr. Biol.* **2009**, *231*, 11−27.

(50) Qian, S.; Wang, W. C.; Yang, L.; Huang, H. W. Structure of the alamethicin pore reconstructed by X-ray diffraction analysis. *Biophys. J.* **2008**, *94*, 3512−3522.

(51) Constantin, D.; Brotons, G.; Jarre, A.; Li, C.; Saldift, T. Interaction of alamethicin pores in DMPC bilayers. *Biophys. J.* **2007**, *92*, 3978−3987.

(52) He, K.; Ludtke, S. J.; Worcester, D. L.; Huang, H. W. Neutron scattering in the plane of membranes: Structure of alamethicin pores. *Biophys. J.* **1996**, *70*, 2659−2666.

(53) Hall, J. E.; Vodyanoy, I.; Balasubramanian, T. M.; Marshall, G. R. Alamethicin − A Rich Model for Channel Behavior. *Biophys. J.* **1984**, *45*, 233−247.

(54) Sansom, M. S. P.; Breed, J.; Berendsen, H. J. C.; Tieleman, P. Alamethicin channels − Simulation studies. *Biophys. J.* **1998**, *74*, A232−A232.

(55) Pan, J.; Tieleman, D. P.; Nagle, J. F.; Kučerka, N.; Tristram-Nagle, S. Alamethicin in lipid bilayers: Combined use of X-ray scattering and MD simulations. *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788*, 1387−1397.

(56) Mihajlovic, M.; Lazaridis, T. Antimicrobial peptides bind more strongly to membrane pores. *Biochim. Biophys. Acta, Biomembr.* **2010**, *1798*, 1494−1502.

(57) Dittmer, J.; Thøgersen, L.; Underhaug, J.; Bertelsen, K.; Vosegaard, T.; Pedersen, J. M.; Schiøtt, B.; Tajkhorshid, E.; Skrydstrup, T.; Nielsen, N. C. Incorporation of Antimicrobial Peptides into Membranes: A Combined Liquid-State NMR and Molecular Dynamics Study of Alamethicin in DMPC/DHPC Bicelles. *J. Phys. Chem. B* **2009**, *113*, 6928−6937.

(58) *The PyMOL Molecular Graphics System*, version 1.3r1; Schrödinger, LLC: New York, 2010.

(59) Huang, W.; Lin, Z. X.; van Gunsteren, W. F. Validation of the GROMOS 54A7 Force Field with Respect to beta-Peptide Folding. *J. Chem. Theory Comput.* **2011**, *7*, 1237−1243.

(60) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.

(61) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. Protein Data Bank − Computer-based archival file for macro-molecular structures. *J. Mol. Biol.* **1977**, *112*, 535−542.

(62) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(63) Tieleman, D. P. Palmitoyl-linoleylphosphatidylcholine bilayer with 128 lipids. http://wcm.ucalgary.ca/tieleman/downloads.

(64) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.

(65) Spoel, D. v. d.; Lindahl, E.; Hess, B.; Buuren, A. R. v.; Apol, E.; Meulenhoff, P. J.; Tieleman, P.; Sjibers, A. L. T. M.; Feenstra, K. A.; Drunen, R. v.; Berendsen, H. J. C. Gromacs user manual version 4.5. 2010. http://www.gromacs.org.

(66) Smith, P. E.; Vangunsteren, W. F. Consistent dielectric-properties of the simple point-charge and extended simple point-charge water models at 227 and 300K. *J. Chem. Phys.* **1994**, *100*, 3169−3174.

(67) Akerlof, G. Dielectric Constants of Some Organic Solvent−Water Mixture at Various Temperatures. *J. Am. Chem. Soc.* **1932**, *54*, 4125−4139.

(68) Dwyer, D. S. Molecular simulation of the effects of alcohols on peptide structure. *Biopolymers* **1999**, *49*, 635−645.

(69) Patel, S.; Brooks, C. L., Structure, thermodynamics, and liquid-vapor equilibrium of ethanol from molecular-dynamics simulations using nonadditive interactions. *J. Chem. Phys.* **2005**, *123*, -.

(70) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory. Comput.* **2008**, *4*, 116−122.

(71) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(72) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* **1998**, *108*, 10220−10230.

(73) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(74) Kollman, P. Free Energy Calculations − Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.

(75) Oostenbrink, C.; Van Gunsteren, W. F. Single-step perturbations to calculate free energy differences from unphysical reference states: Limits on size, flexibility, and character. *J. Comput. Chem.* **2003**, *24*, 1730−1739.

(76) Riniker, S.; Christ, C. D.; Hansen, N.; Mark, A. E.; Nair, P. C.; van Gunsteren, W. F., Comparison of enveloping distribution sampling and thermodynamic integration to calculate binding free energies of phenylethanolamine *N*-methyltransferase inhibitors. *J. Chem. Phys.* **2011**, *135*.

(77) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **1963**, *7*, 95−99.

The Secondary Structure of Antiamoebin I and Zervamicin II Peptaibols

Incorporating D-Amino Acids and Proline Analogues. A Modeling Study

The Secondary Structure of Antiamoebin I and Zervamicin II Peptaibols Incorporating D-Amino Acids and Proline Analogues.

A Modeling Study

Tarsila G. Castro[1,2], Nuno M. Micaêlo[1] and Manuel Melle-Franco[2]*

[1]Departamento de Química, Escola de Ciências, Universidade do Minho
Campus de Gualtar 4710 – 057 Braga Portugal

[2]Departamento de Informática, Centro ALGORITMI, Universidade do Minho,
4710-057 Braga, Portugal

*Correspondence: manuelmelle.research@gmail.com

Abstract

Antiamoebin I (AAM-I) and Zervamicin II (Zrv-IIB) are peptaibols with antibiotic activity that perform their function through insertion/disruption of cell membranes. In this study we investigate the folding properties of two classes of non-canonical amino acids inserted in these peptaibols: proline analogues and asymmetrical D-α,α-dialkyl glycines. Systematic substitution of native Aib, Pro, Hyp and Iva residues were done to understand the folding properties of the peptides incorporating non-canonical residues. The peptaibols secondary structure is related to their ability to incorporate in membranes and therefore to their function. Our findings revealed that native Zrv-IIB suffers considerable unfold in water. The non-canonical proline analogue, cis-3-amino-L-proline (ALP) and Iva induce helical structures in both peptaibols. Asymmetric glycines, such as α-methyl-D-leucine (MDL) and α-methyl-D-phenylalanine (MDP) are folding inducers for the two peptaibols. This pre-organization in water may help to overcome the energy barrier required for peptide insertion into the membrane, as well as to facilitate the formation of transmembrane channels.
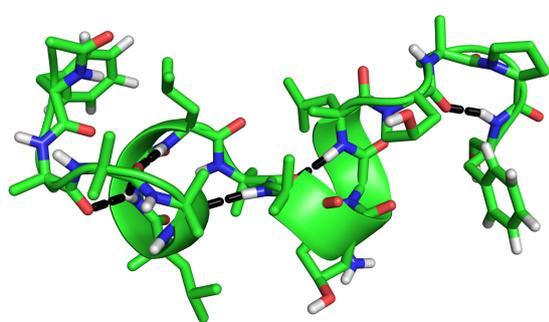
AAM-I peptidomimetic with MDL
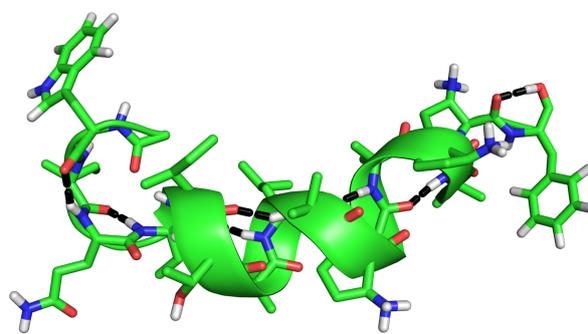
Zrv-IIB peptidomimetic with ALP

Introduction

Peptaibols are antibiotic peptides originated by fungi that present Aib (α-aminoisobutyric acid) in their composition, an amino alcohol C-terminal and have a length of 5 to 20 residues.[1-2] They are normally organized in amphiphilic helices due the presence of non-canonical helix-promoting residues like Aib or Iva (isovaline; α-ethylalanine). In this family of peptides, Alamethicin has been largely investigated in the past decades.[3-4] Antiamoebins and Zervamicins are representative examples of the peptaibol family.[1,5-6]

One of the major health problems today is resistance to conventional antibiotics.[7-8] That is why investigating agents like peptaibols is important, in order to understand the relationship between their structure and mechanism of action. Several studies indicate that the interaction with cell membranes and the formation of pores/channels correspond to the mechanisms of antibiotic activity, and this function is directly related to the structure of the peptides.[9-22] However, for shorter or intermediate Peptaibols, the mechanisms of interaction with the membrane are not well understood and seems to vary according to the nature of the constituent amino acids as well as with the length of the peptide.[1]

We address in this study two 16-amino acid length peptaibols: Antiamoebin I (AAM-I) and Zervamicin-IIB (Zrv-IIB). Antiamoebins are produced by fungi of the species *Emericellopsis poonensis* and have this name due to the antiamoebic properties.[23] AAM-I is one of the most representative members of the Antiamoebins family and has the sequence: Ac-Phe-Aib-Aib-Aib-Iva-Gly-Leu-Aib-Aib-Hyp-Gln-Iva-Hyp-Aib-Pro-Phl.[6,24] Zervamicins were isolated from cultures of *Emericellopsis salmosynnemata*.[25] Zrv-IIB primary structure is: Ac-Trp-Ile-Gln-Iva-Ile-Thr-Aib-Leu-Aib-Hyp-Gln-Aib-Hyp-Aib-Pro-Phl. This peptide is active against Gram-positive bacteria and nontoxic for eukaryotic cells.[25]

The primary structures of AAM-I and Zrv-IIB are very similar at the C-terminal segment, they share the same amino acids residues at the segments 9-11 and 13-16. (see Figure 1). In addition, the position 12 is an asymmetric D-α,α-dialkyl glycine on AAM-I (Iva) and a symmetric/achiral α,α-dialkyl glycine on Zrv-IIB (Aib). Both peptaibols present a high content of three non-canonical amino acids: Aib, Iva and Hyp. There are 6 and 4 Aib residues in AAM-I and in Zrv-IIB respectively. The N-terminal segment in Zrv-IIB is more polar.

The secondary structure (SS) of AAM-I can be classified as a right-handed helix, but the helix has three different secondary structures: α-helix for the residues 1-9, 3_{10}-helix for the short segment 10-12 and an overlapping series of β-turns for the residues 12-16.[6,24] The entire

structure of Zrv-IIB represents an amphiphilic helix with ~26Å of length, 1Å more than AAM-I.[5-6]
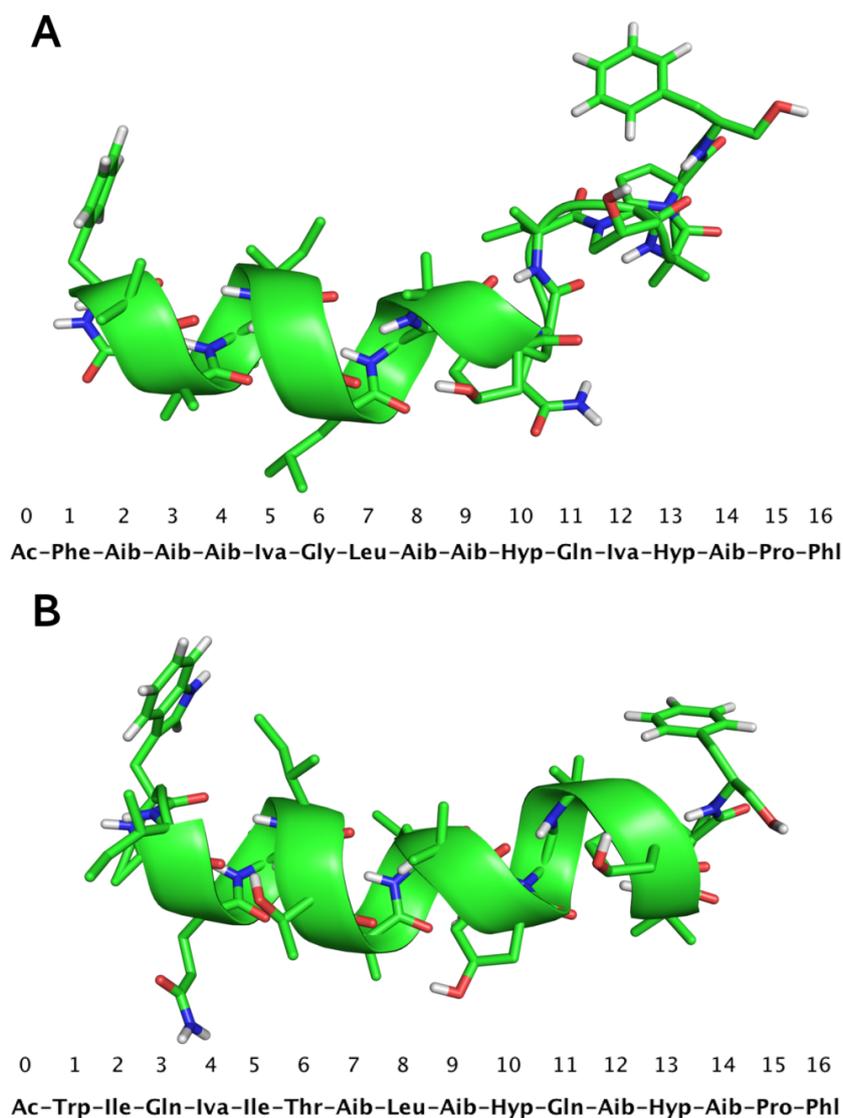
Figure 1. (A) X-ray structure of Antiamoebin I monomer (PDB: 1JOH) and (B) NMR structure of Zervamicin II-B (PDB: 1IH9) with respective primary structures.

The N-terminal segment forms an α-helix with residues 1-8, the C-terminal part (residues 9-16) is organized into a β-ribbon, presenting three $i \rightarrow i+3$ (3_{10}-helix portion) hydrogen bonds and two $i \rightarrow i+4$ hydrogen bonds.[5] Both peptaibols are bent due the presence of Hyp[10], yet the bend angle on AAM-I is higher than in Zrv-IIB, contributing for the slightly short length of AAM.[5-6, 26]

The Zrv-IIB mechanism of action suggests that this peptaibol adopts a helical conformation when approaching the bilayer/water interface. Also, experimental data indicates that the Barrel Stave (BS) model is the preferable pore organization.[5, 19, 21] In contrast, there is

evidence that AAM-I uses a different mechanism to promote ion movement, sometimes forming a complex with the ions (carrier) and sometimes as BS ion channel. [6, 9, 18, 27] This difference is probably related to the bend angles (at Hyp[10]), shape and folded length of both peptides. In addition to the bend angle being lower on Zrv-IIB, the shorter α-helical portion in this peptaibol generates a slightly longer peptide, which in turn, enables Zrv-IIB to span the entire membrane bilayer.[6, 28-29]

Our main goal on this study is to see if new non-canonical amino acids could stabilize AAM-I and Zrv-IIB analogs in a helical form (α-helix or 3_{10}-helix preferentially), generating consequently, longer peptides able to optimize the peptide function as ion-channel forming. We evaluate this by inserting two classes of non-canonical amino acids: asymmetrical D-α,α-dialkyl glycines (similar to the natural D-Iva found on the two peptaibols under study; see Figures 2 and 3) and proline analogs. The asymmetrical D-α,α-dialkyl glycines under investigation are: MCP (2-amino-2-cyclopentylpropanoic acid), MDC (2-amino-2-(2-cyclopentenyl)propanoic acid), MDL (α-methyl-D-leucine), MDP (α-methyl-D-phenylalanine) and MPR (2-amino-2-methyl-4-pentenoic acid). The proline analogs are: ALP (cis-3-amino-L-proline), HLP (trans-3-hydroxy-L-proline) and MLP (cis-4-methyl-L-proline). Note that we have assigned a new three-letter code for the residues that did not receive a prior terminology in other works, namely: ALP, HLP, MLP MCP, MDC, MDL, MDP and MPR

In our previous studies we focused on symmetrical α,α-dialkyl glycines inserted on Peptaibolin and Alamethicin peptaibols and both works suggest the foldamer potential towards α-helical SS of Dhg and Ac_6c.[30-31] Since AAM-I and Zrv-IIB are peptaibols carrying other types of non-canonical amino acids, we chose to study these two new classes (asymmetrical D-α,α-dialkyl glycines and proline analogs) to investigate if and how the secondary structure differs from natural AAM-I and Zrv-IIB (with D-Iva and Hyp).

There are a number of experimental studies addressing several of the non-canonical amino acids explored in this work. Ross and co-workers[32] reported in 1993 the synthesis of α−amino acids, including three asymmetrical α,α-dialkyl glycines under investigation: MPR, MCP and MDC. Mendel and co-workers[33] reported the protein biosynthesis with conformational restricted residues, addressing different classes of amino acids, which included Aib, Iva, MPR and MDL, and they considered this method a powerful approach to generate peptides with well-defined secondary structures. Amino acids analogs to proline have been studied more frequently over the years.[34-38] However, only few studies address

peptide secondary structure.[34-35] This study investigates the variation of the structural features induced by the following non-canonical amino acids: Aib, Iva, Hyp, ALP, HLP, MLP, MCP, MDC, MDL, MDP and MPR, in AAM-I and Zrv-IIB. The objective is to determine which amino acids have greater tendency to induce α-helical or 3_{10}-helical secondary structures (SS) in these antibiotic peptaibols.

Materials and Methods

Non-canonical amino acid FF parameters

The non-canonical amino acids investigated on this work were built with the program PyMOL.[39] The GROMOS topologies, that include the bonded and non-bonded parameters, were based on the similar encoded residues within the GROMOS 54a7 force field (FF).[40] The new bonded and non-bonded parameters are listed in the Supporting Information (SI) using the G54a7 FF syntax (Table 1S).

System Preparation

The X-ray structure of AAM-I and the NMR structure of Zrv-IIB used in this study are available on the Protein Data Bank, with the codes 1JOH and 1IH9 respectively (see Figure1).[5-6]

We created several peptides analogs by inserting proline analogs and asymmetrical D-α,α-dialkyl glycines. First, we replace all Aibs positions in AAM-I and Zrv-IIB for amino acids of a different nature, i.e., we exchanged Aib, that is a symmetric α,α-dialkyl glycine, by the asymmetric D-α,α-dialkyl glycine, Iva, or by Hyp, naturally found residues in these peptides. The second type of modification performed, was replacing all D-Iva positions by analogous residues (MCP, MDC, MDL, MDP and MPR). Similarly, we have substituted the positions originally occupied by Hyp or Pro by proline analogs (ALP, HLP and MLP). The third type of change was similar to the second, but in this case we replaced only one or two possible positions, to explore the effect of the insertion of a single residue in the final structure. We named the generated peptides using the three letters of the substituting amino acid followed by its position in the peptide (see Figure 2 and 3, and Tables 1 and 2).
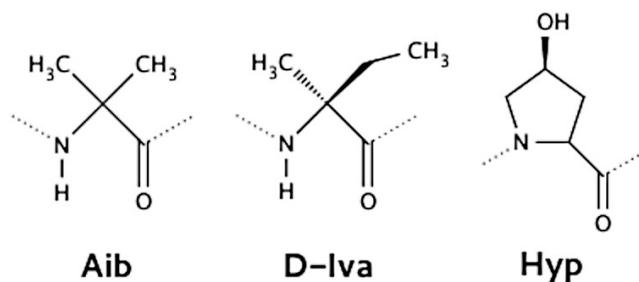
Figure 2. Two-dimensional structures of the non-canonical amino acids naturally found on AAM-I and Zrv-IIB peptaibols: Aib (α-amino isobutyric acid), D-Iva (isovaline; α-ethylalanine) and Hyp (4-hydroxyproline).
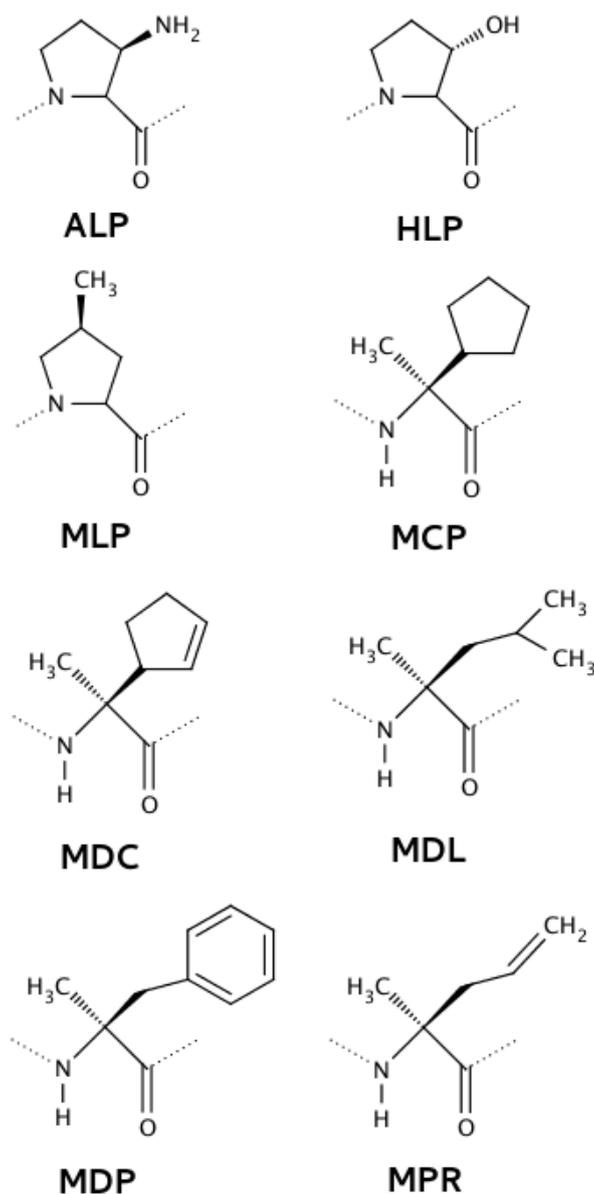
Figure 3. Two-dimensional structures of the non-canonical amino acids under investigation: proline analogs and asymmetrical D-amino acids. The proline analogs are: ALP (cis-3-amino-L-proline), HLP (trans-3-hydroxy-L-proline) and MLP (cis-4-methyl-L-proline). The asymmetrical D-α,α-dialkyl glycines under investigation are: MCP (2-amino-2-cyclopentylpropanoic acid), MDC (2-amino-2-(2-cyclopentenyl)propanoic acid), MDL (α-methyl-D-leucine), MDP (α-methyl-D-phenylalanine) and MPR (2-amino-2-methyl-4-pentenoic acid).

The new AAM-I and Zrv-IIB peptidomimetics were modeled in water with the simple point charge (SPC) water model in an octahedral box with a hydration layer of at least 1.5 nm between the peptide and the walls. Thus, the AAM-I and analogs systems had about 5200-5600 water molecules, and the Zrv-IIB system had about 4000-4500 water molecules. In the case of analogs carrying the residue ALP, which is protonated at physiological conditions, we add Cl- atoms until the box had zero net charge.

Table 1. Sequences generated through the incorporation of the non-canonical amino acids under study to the AAM-I peptaibol.

Residues	Peptide Sequence
native (Aib, Iva, Hyp)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-Hyp-Aib**-Pro-Phl
Aib → Iva (2, 3, 4, 8, 9, 14)	Ac-Phe-**Iva-Iva-Iva-Iva**-Gly-Leu-**Iva-Iva-Hyp**-Gln-**Iva-Hyp-Iva**-Pro-Phl
Aib → Hyp (2, 3, 4, 8, 9, 14)	Ac-Phe-**Hyp-Hyp-Hyp-Iva**-Gly-Leu-**Hyp-Hyp-Hyp**-Gln-**Iva-Hyp-Hyp**-Pro-Phl
Hyp/Pro → ALP (10, 13, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-ALP**-Gln-**Iva-ALP-Aib-ALP**-Phl
Hyp/Pro → HLP (10, 13, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-HLP**-Gln-**Iva-HLP-Aib-HLP**-Phl
Hyp/Pro → MLP (10, 13, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-MLP**-Gln-**Iva-MLP-Aib-MLP**-Phl
Iva → MCP (5, 12)	Ac-Phe-**Aib-Aib-Aib-MCP**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MCP-Hyp-Aib**-Pro-Phl
Iva → MDC (5, 12)	Ac-Phe-**Aib-Aib-Aib-MDC**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MDC-Hyp-Aib**-Pro-Phl
Iva → MDL (5, 12)	Ac-Phe-**Aib-Aib-Aib-MDL**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MDL-Hyp-Aib**-Pro-Phl
Iva → MDP (5, 12)	Ac-Phe-**Aib-Aib-Aib-MDP**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MDP-Hyp-Aib**-Pro-Phl
Iva → MPR (5, 12)	Ac-Phe-**Aib-Aib-Aib-MPR**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MPR-Hyp-Aib**-Pro-Phl
Aib → Iva (8, 14)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Iva-Aib-Hyp**-Gln-**Iva-Hyp-Iva**-Pro-Phl
Aib → Iva (9, 14)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Iva-Hyp**-Gln-**Iva-Hyp-Iva**-Pro-Phl
Aib → Hyp (8, 14)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Hyp-Aib-Hyp**-Gln-**Iva-Hyp-Hyp**-Pro-Phl
Aib → Hyp (9, 14)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Hyp-Hyp**-Gln-**Iva-Hyp-Hyp**-Pro-Phl
Hyp/Pro → ALP (10, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-ALP**-Gln-**Iva-Hyp-Aib-ALP**-Phl
Hyp/Pro → ALP (13, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-ALP-Aib-ALP**-Phl
Hyp/Pro → HLP (10, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-HLP**-Gln-**Iva-Hyp-Aib-HLP**-Phl
Hyp/Pro → HLP (13, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-HLP-Aib-HLP**-Phl
Hyp/Pro → MLP (10, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-MLP**-Gln-**Iva-Hyp-Aib-MLP**-Phl
Hyp/Pro → MLP (13, 15)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-MLP-Aib-MLP**-Phl
Iva → MCP (5)	Ac-Phe-**Aib-Aib-Aib-MCP**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-Hyp-Aib**-Pro-Phl
Iva → MCP (12)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MCP-Hyp-Aib**-Pro-Phl
Iva → MDC (5)	Ac-Phe-**Aib-Aib-Aib-MDC**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-Hyp-Aib**-Pro-Phl
Iva → MDC (12)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MDC-Hyp-Aib**-Pro-Phl
Iva → MDL (5)	Ac-Phe-**Aib-Aib-Aib-MDL**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-Hyp-Aib**-Pro-Phl
Iva → MDL (12)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MDL-Hyp-Aib**-Pro-Phl
Iva → MDP (5)	Ac-Phe-**Aib-Aib-Aib-MDP**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-Hyp-Aib**-Pro-Phl
Iva → MDP (12)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MDP-Hyp-Aib**-Pro-Phl
Iva → MPR (5)	Ac-Phe-**Aib-Aib-Aib-MPR**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**Iva-Hyp-Aib**-Pro-Phl
Iva → MPR (12)	Ac-Phe-**Aib-Aib-Aib-Iva**-Gly-Leu-**Aib-Aib-Hyp**-Gln-**MPR-Hyp-Aib**-Pro-Phl

Table 2. Sequences generated through the incorporation of the non-canonical amino acids under study to the Zrv-IIB peptaibol.

Residues	Sequence
native (Aib, Iva, Hyp)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Aib → Iva (7, 9, 12, 14)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Iva**-Leu-**Iva**-**Hyp**-Gln-**Iva**-**Hyp**-**Iva**-Pro-Phl
Aib → Hyp (7, 9, 12, 14)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Hyp**-Leu-**Hyp**-**Hyp**-Gln-**Hyp**-**Hyp**-**Hyp**-Pro-Phl
Hyp/Pro → ALP (10, 13, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**ALP**-Gln-**Aib**-**ALP**-**Aib**-**ALP**-Phl
Hyp/Pro → HLP (10, 13, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**HLP**-Gln-**Aib**-**HLP**-**Aib**-**HLP**-Phl
Hyp/Pro → MLP (10, 13, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**MLP**-Gln-**Aib**-**MLP**-**Aib**-**MLP**-Phl
Iva → MCP (4)	Ac-Trp-Ile-Gln-**MCP**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Iva → MDC (4)	Ac-Trp-Ile-Gln-**MDC**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Iva → MDL (4)	Ac-Trp-Ile-Gln-**MDL**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Iva → MDP (4)	Ac-Trp-Ile-Gln-**MDP**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Iva → MPR (4)	Ac-Trp-Ile-Gln-**MPR**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Aib → Iva (7, 12)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Iva**-Leu-**Aib**-**Hyp**-Gln-**Iva**-**Hyp**-**Aib**-Pro-Phl
Aib → Iva (7, 14)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Iva**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Iva**-Pro-Phl
Aib → Iva (9, 12)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Iva**-**Hyp**-Gln-**Iva**-**Hyp**-**Aib**-Pro-Phl
Aib → Iva (9, 14)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Iva**-**Hyp**-Gln-**Aib**-**Hyp**-**Iva**-Pro-Phl
Aib → Iva (12, 14)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Iva**-**Hyp**-**Iva**-Pro-Phl
Aib → Hyp (7)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Hyp**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Aib → Hyp (9)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Hyp**-**Hyp**-Gln-**Aib**-**Hyp**-**Aib**-Pro-Phl
Aib → Hyp (12)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Hyp**-**Hyp**-**Aib**-Pro-Phl
Aib → Hyp (14)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**Hyp**-**Hyp**-Pro-Phl
Hyp/Pro → ALP (10, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**ALP**-Gln-**Aib**-**Hyp**-**Aib**-**ALP**-Phl
Hyp/Pro → ALP (13, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**ALP**-**Aib**-**ALP**-Phl
Hyp/Pro → HLP (10, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**HLP**-Gln-**Aib**-**Hyp**-**Aib**-**HLP**-Phl
Hyp/Pro → HLP (13, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**HLP**-**Aib**-**HLP**-Phl
Hyp/Pro → MLP (10, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**MLP**-Gln-**Aib**-**Hyp**-**Aib**-**MLP**-Phl
Hyp/Pro → MLP (13, 15)	Ac-Trp-Ile-Gln-**Iva**-Ile-Thr-**Aib**-Leu-**Aib**-**Hyp**-Gln-**Aib**-**MLP**-**Aib**-**MLP**-Phl

Molecular Dynamics Simulations

All simulations were performed using the GROMACS 4.5.4 version.[41-42] For the treatment of long-range interactions, we used the Reaction Field method, with a cut-off of 1.4 nm and for consistency a dielectric constant of 54 for water.[43-44] The van der Waals interactions were also truncated with a twin-range cut-off of 0.8 and 1.4nm. The algorithm LINCS[45-46] was used to constrain the chemical bonds of the peptides and the algorithm SETTLE[47] in the case of water. The pressure and temperature Berendsen algorithms[48] were used to control the temperature and pressure at 310K and 1 atm, respectively. We used the following coupling constants: $\tau_T = 0.10$ ps and $\tau_P = 1.0$ ps.

Three steps of energy minimization were performed. In the first two steps, position restraints (with force constant of 1000 kJ·mol^{-1}·nm^{-2}) were applied to all heavy atoms of the peptide and afterwards on the main chain. In the third step of energy minimization no

position restrains were applied. One molecular dynamics simulation of 100 ps was done with position restraints (force constant of 1000 kJ·mol^{-1}·nm^{-2}) on the heavy atoms and afterwards a 200 ps simulation was done with position restraints (force constant of 1000 kJ·mol^{-1}·nm^{-2}) on the main chain. The systems were equilibrated and sampled using 100 ns molecular dynamics simulations with an integration interval of 2 fs. To ensure a better sampling of the conformational states of these peptides in water, at least three replicates of each system were performed.

Analysis

Typically, in order to analyze the structural stability of peptides and proteins in a MD simulation, the root-mean-square deviation (rmsd) computed against the starting experimental structure is monitored with time. However, in order to follow subtle changes in secondary structure (specifically in α-helix and 3_{10}-helix content), we computed the Secondary Structure (SS), by the DSSP (Dictionary of Secondary Structure in Proteins) method implemented in GROMACS.[41] MD was run for 100 ns, with 40 ns of equilibration followed by 60 ns of production in which we selected the replicate with the highest percentage of helical secondary structure conformations (α-helix + 3_{10}-helix). Tables 3 and 4 report the percentage of conformations with at least 3 or 4 residues forming a 3_{10}-helix or α-helix, respectively. Also, the average number of residues in α-helix and 3_{10}-helix and the presence of Intramolecular Hydrogen Bonds were computed.

Central conformations, shown on Figures 4 and 5, are the ones that minimize the RMSD variance when fitted against all other conformations of the trajectory, these conformations correspond to the most populated conformation of the simulation. In addition, we also calculate the number of residues in different SS, through a DSSP analysis, and the number of intramolecular hydrogen bonds, for the most likely conformations. Additionally data from DSSP analysis (Figures S2 and S3) is presented as SI.

Results and Discussion

The Maintenance or Formation of Helical Secondary Structure

All Aib substitutions: changing the amino acid nature

The 2-dimensional structures of all non-canonical amino acids investigated in this work are shown on Figures 2 and 3. Aib, D-Iva and Hyp are the naturally found non-canonical residues present in AAM-I and Zrv-IIB (Figure 2). Aib and Iva are symmetrical and

asymmetrical derivatives of α,α-dialkyl glycines and Hyp is a proline analog. ALP, HLP and MLP are the new proline analogs that we investigated of which ALP is charged at physiological pH, HLP residue is polar due to the hydroxyl group, and MLP is apolar. MCP, MDC, MDL, MDP and MPR are D-asymmetrical disubstituted glycines. Figure 1 shows the primary and secondary structures of AAM-I (panel A) and Zrv-IIB (panel B).

The initial conformations for MD simulations were generated by replacing the residues of interest on the experimental structures for the non-canonical amino acids under study. The SS analysis was used to quantify the percentage of conformations presenting helical structure in the last 60 ns of the simulation in aqueous medium (Figure 1S on SI shows that after 40ns, the peptides reach equilibrium). This can tell us conformational preferences of the peptides under study, and indirectly, if these peptides have moved away from the initial structure, predominantly helical.

Table 3 and Table 4 present the percentage of conformations with helical SS for AAM-I and Zrv-IIB peptaibols and analogs, respectively. Note that, due to the length of these peptides, one configuration might contain α-helix and 3_{10}-helix at the same time.

AAM-I has six Aib residues that are known to be mostly α-helical inducers.[49-51] In fact, when we simulated AAM-I in water, the percentage of conformations with α-helical and 3_{10}-helix portions are 77% and 20% respectively.

The replacement of all Aib amino acids for D-Iva or Hyp increases the total percentage of conformations with helical SS in comparison with the native AAM-I in water (Table 3), suggesting that this is a favorable change for helix formation. Typically, the imino group present on Hyp causes the breakdown of α-helical structures, since it is geometrically incompatible with the spiral towards the right of the α-helix.[5-6] Thus, generally, this group inserts a bend on the chain, which interrupts the helical structure. This is not the case here, probably due to the presence of the polar six-hydroxyl groups. In fact, although less probable, β-turn structures can be accommodated into helical backbones (with a hydrogen pattern $i \rightarrow i+3$), and this is the preferable SS conformation of Pro and analogues.[34] One example of this is the structure of collagen, where consecutive Pro and Hyp, generate a helix.[52]

Table 3. Percentage of conformations with helical secondary structure (α-helix and 3_{10}-helix and) obtained for the wild type AAM-I and each peptide analog under study, considering the last 60ns of simulation time.

Peptides	% of conformations with helical SS	
	α-helix	3_{10}-helix
AAM-I	76.7	19.6
Aib $\longrightarrow$ Iva $_{(2, 3, 4, 8, 9, 14)}$	93.4	6.6
Aib $\longrightarrow$ Hyp $_{(2, 3, 4, 8, 9, 14)}$	99.5	2.3
Hyp/Pro $\longrightarrow$ ALP $_{(10, 13, 15)}$	92.6	53.2
Hyp/Pro $\longrightarrow$ HLP $_{(10, 13, 15)}$	78.3	45.2
Hyp/Pro $\longrightarrow$ MLP $_{(10, 13, 15)}$	24.4	57.1
Iva $\longrightarrow$ MCP $_{(5, 12)}$	45.6	47.9
Iva $\longrightarrow$ MDC $_{(5, 12)}$	53.2	25.0
Iva $\longrightarrow$ MDL $_{(5, 12)}$	84.2	28.1
Iva $\longrightarrow$ MDP $_{(5, 12)}$	91.2	34.0
Iva $\longrightarrow$ MPR $_{(5, 12)}$	0.2	48.8
Aib $\longrightarrow$ Iva $_{(8, 14)}$	23.2	55.0
Aib $\longrightarrow$ Iva $_{(9, 14)}$	--	72.8
Aib $\longrightarrow$ Hyp $_{(8, 14)}$	3.5	89.7
Aib $\longrightarrow$ Hyp $_{(9, 14)}$	26.6	41.9
Hyp/Pro $\longrightarrow$ ALP $_{(10, 15)}$	52.1	53.4
Hyp/Pro $\longrightarrow$ ALP $_{(13, 15)}$	77.6	49.5
Hyp/Pro $\longrightarrow$ HLP $_{(10, 15)}$	53.5	33.6
Hyp/Pro $\longrightarrow$ HLP $_{(13, 15)}$	46.3	29.1
Hyp/Pro $\longrightarrow$ MLP $_{(10, 15)}$	99.6	0.1
Hyp/Pro $\longrightarrow$ MLP $_{(13, 15)}$	50.7	47.8
Iva $\longrightarrow$ MCP $_{(5)}$	30.0	18.3
Iva $\longrightarrow$ MCP $_{(12)}$	35.0	40.4
Iva $\longrightarrow$ MDC $_{(5)}$	--	49.6
Iva $\longrightarrow$ MDC $_{(12)}$	18.3	44.1
Iva $\longrightarrow$ MDL $_{(5)}$	69.1	32.2
Iva $\longrightarrow$ MDL $_{(12)}$	66.7	35.9
Iva $\longrightarrow$ MDP $_{(5)}$	9.6	82.8
Iva $\longrightarrow$ MDP $_{(12)}$	70.7	18.6
Iva $\longrightarrow$ MPR $_{(5)}$	67.0	41.4
Iva $\longrightarrow$ MPR $_{(12)}$	68.0	21.9

Zrv-IIB unfolds considerably when simulated in water, as show by the low percentage of conformations with helical SS (Table 4). Note that the experimental structure was obtained by NMR in DPC micelles, a membrane-like environment that promotes folding; and in water, it is probable that the peptide undergoes a hydrophobic unfolding as the nonpolar residues tend to aggregate to protect themselves from water. Furthermore, this peptide contains fewer Aib residues than AAM-I, only four, which may not be sufficient to impose the proper constraint to the structure.

Similarly to the observed for AAM-I, the substitution of all the Aib on Zrv-IIB favors the helical content. D-Iva increases the population of α-helical conformations and Hyp induces both α-helix and 3_{10}-helix SS.

Table 4. Percentage of conformations with helical secondary structure (α-helix and 3_{10}-helix) obtained for the wild type Zrv-IIB and each peptide analog under study, considering the last 60ns of simulation time.

Peptides	% of conformations with helical SS	
	α-helix	3_{10}-helix
Zrv-IIB	20.1	7.6
Aib → Iva $_{(7, 9, 12, 14)}$	77.3	11.3
Aib → Hyp $_{(7, 9, 12, 14)}$	20.4	55.5
Hyp/Pro → ALP $_{(10, 13, 15)}$	66.4	32.6
Hyp/Pro → HLP $_{(10, 13, 15)}$	24.2	15.1
Hyp/Pro → MLP $_{(10, 13, 15)}$	38.6	27.3
Iva → MCP $_{(4)}$	3.3	71.6
Iva → MDC $_{(4)}$	1.2	27.9
Iva → MDL $_{(4)}$	81.9	8.4
Iva → MDP $_{(4)}$	83.1	5.9
Iva → MPR $_{(4)}$	32.0	26.0
Aib → Iva $_{(7, 12)}$	12.7	46.0
Aib → Iva $_{(7, 14)}$	6.8	49.4
Aib → Iva $_{(9, 12)}$	46.2	3.7
Aib → Iva $_{(9, 14)}$	0.2	15.0
Aib → Iva $_{(12, 14)}$	4.4	70.6
Aib → Hyp $_{(7)}$	14.5	29.7
Aib → Hyp $_{(9)}$	63.2	13.1
Aib → Hyp $_{(12)}$	56.9	18.4
Aib → Hyp $_{(14)}$	73.9	21.3
Hyp/Pro → ALP $_{(10, 15)}$	41.0	41.3
Hyp/Pro → ALP $_{(13, 15)}$	72.1	15.4
Hyp/Pro → HLP $_{(10, 15)}$	68.8	9.0
Hyp/Pro → HLP $_{(13, 15)}$	36.2	28.4
Hyp/Pro → MLP $_{(10, 15)}$	71.4	32.7
Hyp/Pro → MLP $_{(13, 15)}$	68.3	26.9

We also analyzed through DSSP, the percentage of conformations presenting turns and bend SS. The algorithm written by Kabsch and Sander[53] establishes that a turn occurs when exists a hydrogen bond between CO(i) to NH(i+n), where n=3, 4 or 5. When the interactions are consecutive and according to dihedral preferences, a helix of types 3_{10}, α or π, take place. In addition, a bend corresponds to a region of high curvature (at least 70°). For both AAM-I and Zrv-IIB peptidomimetics the presence of bends was not observed. However, we obtained a percentage of conformations presenting turns, from 70% to 99%, for all systems under study, indicating that a minimum of one turn is present in most cases.

Iva and Hyp/Pro replacements

The second type of analogs studied corresponds to the substitution of the naturally found Iva by asymmetrical D-α,α-dialkyl glycines or the replacement of the naturally found Hyp/Pro by proline analogs (in all possible positions).

In AAM-I peptides carrying D-Iva and analogues (Table 3), only the residues MDL and MDP induce more conformations with helical SS than native simulated AAM-I. In the case of proline analogs, the change of Hyp and Pro positions for ALP or HLP increases significantly the helical SS content. Note that the substituents NH_2 and OH on these residues are in position 3, while Hyp has an OH on position 4 (Figures 2 and 3). This difference could minimize steric hindrance and improve the global number of hydrogen bonds necessary to the formation of α-helical structures.

For Zrv-IIB analogs, on Table 4, we observe many substitutions that favor the maintenance/formation of helical conformations. ALP stands out with a large proportion of conformations containing α-helices when the substitution occurs in all three possible positions (10, 13 and 15). In the class of D-Iva analogues, MDL and MDP present a significant number of conformations in α-helix, and MCP presents about 72% of conformations with the 3_{10}-helix form. Importantly, Zrv-IIB has only one position on its sequence with a D-Iva amino acid, so it is remarkable that only one residue substitution is able to induce such conformational change in the peptide structure.

Single and Double Substitutions

In this section we evaluate if a minimum number of substitutions is able to increase the number of helical conformations. To do this, we designed analogues where we change either one or two residues combined in different ways.

For the AAM-I analogs (Table 3), we highlight the peptide containing $MLP_{(10,15)}$ that presents 99.6% of conformations with α-helical structure. Remarkably, changing D-Iva for MPR only in position 5 we can induce conformations with a good balance of α-helical and 3_{10}-helical SS (67.0% and 41.4%, respectively).

Considering the peptides based on Zrv-IIB (Table 4), we observe many substitutions that increase the percentages of helical SS, since the Zrv-IIB suffers considerable unfold in water. The peptides carrying Iva $_{(7,12)}$, Iva $_{(7,14)}$, Iva $_{(9,12)}$ and Iva $_{(12,14)}$ exhibit high percentages of helical SS, but the replacement in all four possible positions, Iva $_{(7,9,12,14)}$, induce more conformations with helical SS. Similarly, we observe that the analogs containing $Hyp_{(7)}$,

$Hyp_{(9)}$, $Hyp_{(12)}$ and $Hyp_{(14)}$ all have more conformations with helical SS than Zrv-IIB, but although 3 out of 4 favor α-helix conformations, the combination of four positions, $Hyp_{(7,9,12,14)}$, favors the 3_{10}-helix form.

Interestingly, the residue $MLP_{(10,15)}$ with a methyl (Me) substituent on position 4, similar to Hyp but with an alkyl group, induces a great number of conformations with helical SS with the combination of only two positions. This is an indication that the substituent position is equally or more relevant than the polarity of the group attached to the proline ring.

AAM-I and Zrv-IIB peptidomimetics

We selected the AAM-I and Zrv-IIB analogs with the highest number of helical SS in water for a more detailed analysis. In this sense, we analyzed the percentage of conformations containing types $i \rightarrow i+3$ and $i \rightarrow i+4$ of intramolecular hydrogen bonds (corresponding to 3_{10}-helix, α-helix or turn) and the average number of residues in a specific SS. This is relevant because it reveals the peptides that have a natural tendency for helical structures in aqueous media, which in turn might translate to better channel formation or other forms of membrane disruption.

On Tables 5 and 6 the percentage of conformations with type $i \rightarrow i+3$ and $i \rightarrow i+4$ intramolecular hydrogen bonds show in many peptides suggest that the α-helix, 3_{10}-helix and turn SS coexist, in agreement with experimental results.[5-6] Also, the high values of percentage of conformations with type $i \rightarrow i+3$ and $i \rightarrow i+4$ hydrogen bonds, indicate that there is no tendency for a random coil or extended SS.

Table 5. Percentage of conformations presenting intramolecular hydrogen bonds of types $i \rightarrow i+3$ (3_{10}-helix or turn) and $i \rightarrow i+4$ (α-helix or turn), obtained for AAM-I and analogs, considering the last 60ns of simulation time.

Peptide	% of conformations with hydrogen bond $i \rightarrow i+3$	% of conformations with hydrogen bond $i \rightarrow i+4$
AAM-I	66.0	95.8
Aib → Iva $_{(2,3,4,8,9,14)}$	81.7	98.2
Aib → Hyp $_{(2,3,4,8,9,14)}$	41.3	98.9
Hyp/Pro → ALP $_{(10,13,15)}$	87.0	97.1
Hyp/Pro → HLP $_{(10,13,15)}$	84.0	98.3
Hyp/Pro → MLP $_{(10,15)*}$	2.2	99.8
Iva → MDL $_{(5,12)}$	69.6	99.3
Iva → MDP $_{(5,12)}$	90.1	96.5
Iva → MPR $_{(5)}$	91.6	93.7
* Hyp/Pro → MLP $_{(10,15)}$ has 98.6% of $i \rightarrow i+2$ hydrogen bonds.		

Table 6. Percentage of conformations presenting intramolecular hydrogen bonds of types $i \rightarrow i+3$ (3_{10}-helix or turn) and $i \rightarrow i+4$ (α-helix or turn), obtained for Zrv-IIB and analogs, considering the last 60ns of simulation time.

Peptide	% of conformations with hydrogen bond $i \rightarrow i+3$	% of conformations with hydrogen bond $i \rightarrow i+4$
Zrv-IIB	99.7	48.2
Aib $\rightarrow$ Iva $_{(7,9,12,14)}$	86.8	99.8
Aib $\rightarrow$ Hyp $_{(7,9,12,14)}$	82.3	25.1
Hyp/Pro $\rightarrow$ ALP $_{(10,13,15)}$	81.9	96.0
Hyp/Pro $\rightarrow$ HLP $_{(10,15)}$	59.2	95.4
Hyp/Pro $\rightarrow$ MLP $_{(10,15)}$	82.9	94.7
Hyp/Pro $\rightarrow$ MLP $_{(13,15)}$	96.2	96.5
Iva $\rightarrow$ MCP $_{(4)}$	99.1	36.7
Iva $\rightarrow$ MDL $_{(4)}$	74.8	97.5
Iva $\rightarrow$ MDP $_{(4)}$	91.1	99.5

AAM-I and Zrv-IIB analogs present, in most cases, high percentages of both $i \rightarrow i+3$ and $i \rightarrow i+4$ hydrogen bonds. In the case of Zrv-IIB analogs, we observed that the ones carrying MLP and MDP are well structured, since the two interactions are highly populated.

Comparing directly the wild types AAM-I and Zrv-IIB, we see that the percentage of $i \rightarrow i+3$ hydrogen bonds is higher and dominant on Zrv-IIB. This fact explains why this peptide is longer than AAM-I, as observed experimentally. [5-6] Concerning the AAM-I analogs (Table 5), the peptides carrying, Iva, ALP, HLP, MDP and MPR show a higher number of conformations with type $i \rightarrow i+4$ hydrogen bond, as well as, a higher number of type $i \rightarrow i+3$ hydrogen bond. This indicates that these peptides have a good portion of their structures stabilized on α-helix, 3_{10}-helix or turns. In the case of Zrv-IIB analogs (Table 6), the peptides containing Iva, ALP, MLP and MDP are those that have the highest numbers of conformations with types $i \rightarrow i+3$ and $i \rightarrow i+4$ of intramolecular hydrogen bonds.

Tables 7 and 8 present the average number of residues in α-helix and in 3_{10}-helix type, for the peptaibols and chosen analogs. As a reference, we analyzed the X-ray monomer structure of AAM-I and the NMR structure of Zrv-IIB with the same method used to quantify the SS of the proposed analogs, and we found that only 50% of AAM-I X-ray structure is in α-helix and no 3_{10}-helix was present. In the case of Zrv-IIB 75% of helical SS was observed. Proportionally, the modeled Zrv-IIB deviates more from experimental structure than AAM-I, but this is justified by the fact that the Zrv-IIB structure has been obtained in micelles.

For AAM-I and analogs (Table 7) we highlight that the peptides carrying Iva, ALP, MDL and MPR, have the highest sums of residues in helical SS (9.2, 8.6, 9,2 and 9.2, respectively). This indicates that these amino acids are able to stabilize a large portion of the peptide in a

helical SS, during the simulation on water. On Table 8 we observe that the peptide carrying Iva and ALP are the ones with the highest number of residues in helical SS (9.1 and 9, respectively). Iva and ALP share the same tendency to fold a helical SS in both peptaibols.

Table 7. Average number of residues in α-helical and 3_{10}-helical conformations, with respective standard deviation obtained for AAM-I and chosen analogs, considering the last 60ns of simulation time.

Peptide	Average number of residues in α-helix	Standard deviation	Average number of residues in 3_{10}-helix	Standard deviation
AAM-I X-ray	8	--	--	--
AAM-I	4.4	± 0.9	3.6	± 1.0
Aib → Iva $_{(2,3,4,8,9,14)}$	5.3	± 0.8	3.9	± 1.0
Aib → Hyp $_{(2,3,4,8,9,14)}$	5.1	± 1.2	3.2	± 0.8
Hyp/Pro → ALP $_{(10,13,15)}$	4.9	± 1.8	3.7	± 1.3
Hyp/Pro → HLP $_{(10,13,15)}$	5.1	± 1.9	3.3	±0.8
Hyp/Pro → MLP $_{(10,15)}$	4.1	± 0.6	4.3	± 1.9
Iva → MDL $_{(5,12)}$	4.8	± 1.3	4.4	± 1.8
Iva → MDP $_{(5,12)}$	4.3	± 0.9	3.6	± 1.1
Iva → MPR $_{(5)}$	5.0	± 1.2	4.2	± 1.6

Table 8. Average number of residues in α-helical and 3_{10}-helical conformations, with respective standard deviation obtained for Zrv-IIB and chosen analogs, considering the last 60ns of simulation time.

Peptide	Average number of residues in α-helix	Standard deviation	Average number of residues in 3_{10}-helix	Standard deviation
Zrv-IIB NMR	12	--	--	--
Zrv-IIB	5.1	± 0.9	3.2	± 0.7
Aib → Iva $_{(7,9,12,14)}$	5.6	± 1.4	3.5	± 0.9
Aib → Hyp $_{(7,9,12,14)}$	4.2	± 0.6	3.7	± 1.1
Hyp/Pro → ALP $_{(10,13,15)}$	5.3	± 1.8	3.7	± 1.3
Hyp/Pro → HLP $_{(10,15)}$	5.4	± 1.3	3.5	± 0.9
Hyp/Pro → MLP $_{(10,15)}$	4.5	± 0.8	3.2	± 0.7
Hyp/Pro → MLP $_{(13,15)}$	5.2	± 1.3	3.7	± 1.0
Iva → MCP $_{(4)}$	5.2	± 1.2	3.7	± 1.1
Iva → MDL $_{(4)}$	4.9	± 0.9	3.4	± 0.9
Iva → MDP $_{(4)}$	5.2	± 1.0	3.4	± 0.9

Figure 4 shows the central conformation obtained for AAM-I and analogs, indicating the most representative structure of the simulation. The position 4 in all central conformations (Figure 4) participates in the formation of a helical SS, usually as hydrogen bond acceptor. Aib or Iva generally occupies this position. Also, for AAM-I and analogs we observed that the peptides well structured at the N-terminal are: AAM-I, Iva $_{(2,3,4,8,9,14)}$, ALP $_{(10,13,15)}$, MDL $_{(5,12)}$ and MPR $_{(5)}$. The peptide with MDL $_{(5,12)}$ is also well structured at the middle of the

sequence. The peptide carrying MDP $_{(5, 12)}$ is in 3_{10}-helix type at N-terminal, and adopts an α-helix at the middle.

AAM-I: $i \rightarrow i+3$: CO (Aib^4) --- HN (Leu^7)
$i \rightarrow i+4$: CO (Aib^2) --- HN (Gly^6)
CO (Leu^7) --- HN (Gln^{11})
CO (Aib^8) --- HN (Iva^{12})

Iva $_{(2, 3, 4, 8, 9, 14)}$: $i \rightarrow i+2$: CO (Phe^1) --- HN (Iva^3)
CO (Leu^7) --- HN (Iva^9)
$i \rightarrow i+3$: CO (Gln^{11}) --- HN (Iva^{14})
$i \rightarrow i+4$: CO (Iva^2) --- HN (Gly^6)
CO (Iva^3) --- HN (Leu^7)
CO (Iva^4) --- HN (Iva^8)

ALP $_{(10, 13, 15)}$: $i \rightarrow i+3$: CO (Aib^2) --- HN (Iva^5)
CO (Aib^3) --- HN (Gly^6)
CO (Aib^4) --- HN (Leu^7)
$i \rightarrow i+4$: CO (Aib^8) --- HN (Iva^{12})

MDL $_{(5, 12)}$: $i \rightarrow i+3$: CO (Aib^3) --- HN (Gly^6)
CO (Hyp^{13}) --- HN (Phl^{16})
$i \rightarrow i+4$: CO (Aib^2) --- HN (Gln^{11})
CO (Aib^3) --- HN (Leu^7)
CO (Aib^4) --- HN (Aib^8)
CO (Aib^8) --- HN (MDL^{12})

MDP $_{(5, 12)}$: $i \rightarrow i+3$: CO (Aib^2) --- HN (MDP^5)
CO (Aib^4) --- HN (Leu^7)
$i \rightarrow i+4$: CO (Leu^7) --- HN (Gln^{11})
CO (Aib^8) --- HN (MDP^{12})

MPR $_{(5)}$: $i \rightarrow i+2$: CO (Leu^7) --- HN (Aib^9)
$i \rightarrow i+3$: CO (Gln^{11}) --- HN (Aib^{14})
CO (Hyp^{13}) --- HN (Phl^{16})
$i \rightarrow i+4$: CO (Phe^1) --- HN (MPR^5)
CO (Aib^2) --- HN (Gly^6)
CO (Aib^3) --- HN (Leu^7)
CO (Aib^4) --- HN (Aib^8)

Figure 4. Central structures of AAM-I and analogs carrying Iva$_{(2, 3, 4, 8, 9, 14)}$, ALP$_{(10,13,15)}$, MDL$_{(5,12)}$, MDP$_{(5,12)}$ and MPR$_{(5)}$, from the last 60 ns in water. The coloring of the atoms follows the convention: green for carbon, blue for nitrogen, red for oxygen, white for hydrogen, green for the cartoon that defines the helical SS and black dashed traces to highlight hydrogen bonds. The water molecules were omitted for better visualization.

Zrv-IIB: $i \rightarrow i+2$: CO (Ace^0) --- HN (Ile^2)
CO (Trp^1) --- HN (Gln^3)
CO (Aib^7) --- HN (Aib^9)
$i \rightarrow i+3$: CO (Ile^2) --- HN (Ile^5)
CO (Ile^5) --- HN (Leu^8)
CO (Leu^8) --- HN (Gln^{11})
CO (Gln^{11}) --- HN (Aib^{14})

Iva $_{(7, 9, 12, 14)}$**:** $i \rightarrow i+2$: CO (Thr^6) --- HN (Leu^8)
$i \rightarrow i+3$: CO (Hyp^{13}) --- HN (Phl^{16})
$i \rightarrow i+4$: CO (Ace^0) --- HN (Iva^4)
CO (Trp^1) --- HN (Ile^5)
CO (Ile^2) --- HN (Thr^6)
CO (Gln^3) --- HN (Iva^7)
CO (Iva^7) --- HN (Gln^{11})
CO (Leu^8) ---HN (Iva^{12})

ALP $_{(10, 13, 15)}$**:** $i \rightarrow i+2$: CO (Trp^1) --- HN (Gln^3)
CO (Ile^2) --- HN (Iva^4)
$i \rightarrow i+3$: CO (Gln^{11}) --- HN (Aib^{14})
$i \rightarrow i+4$: CO (Iva^4) --- HN (Leu^8)
CO (Ile^5) --- HN (Aib^9)
CO (Leu^8) --- HN (Aib^{12})
*CO (ALP^{15}) --- HO (Phl^{16})

MDP $_{(4)}$**:** $i \rightarrow i+3$: CO (Gln^{11}) --- HN (Aib^{14})
$i \rightarrow i+4$: CO (Ace^0) --- HN (MDP^4)
CO (Trp^1) --- HN (Ile^5)
CO (Ile^2) --- HN (Thr^6)
CO (Gln^3) --- HN (Aib^7)

MLP $_{(10, 15)}$**:** $i \rightarrow i+2$: CO (Ile^5) --- HN (Aib^7)
$i \rightarrow i+3$: CO (Gln^{11}) --- HN (Aib^{14})
$i \rightarrow i+4$: CO (Ace^0) --- HN (Iva^4)
CO (Trp^1) --- HN (Ile^5)
CO (Ile^2) --- HN (Thr^6)
$i \rightarrow i+5$: CO (Gln^3) --- HN (Leu^8)

MLP $_{(13, 15)}$**:** $i \rightarrow i+3$: CO (Ile^5) --- HN (Leu^8)
CO (Gln^{11}) --- HN (Aib^{14})
CO (MLP^{13}) --- HN (Phl^{16})
$i \rightarrow i+4$: CO (Ace^0) --- HN (Iva^4)
CO (Ile^2) --- HN (Thr^6)
CO (Gln^3) --- HN (Aib^7)
*CO (MLP^{15}) --- HO (Phl^{16})

Figure 5. Central structures of Zrv-IIB and analogs carrying Iva$_{(7, 9, 12, 14)}$, ALP$_{(10,13,15)}$, MDP$_{(4)}$, MLP$_{(10,15)}$ and MLP$_{(13,15)}$, from the last 60 ns in water. The coloring of the atoms follows the convention described in Figure 4.

On Figure 5, it is observed that the wild-type Zrv-IIB central conformation does not have $i \rightarrow i+4$ intramolecular hydrogen bonds (α-helices). The peptides well structured at the N-

terminal are: Iva $_{(7,9,12,14)}$, MDP $_{(4)}$, MLP $_{(10,15)}$ and MLP $_{(13,15)}$. The peptide carrying ALP $_{(10,13,15)}$ is one example of the central conformation is structured at the middle.

The central conformations shown in Figures 4 and 5 show that not always the non-canonical residue inserted directly participates in the helix, as hydrogen bond donor or receptor, but may be able to modify the geometry of the peptide on another portion of the peptide.

In both peptaibols and analogues Aib frequently participates on the formation of helical SS, as hydrogen bond donor or receptor. On the one hand, this agrees with the existing studies on the remarkable ability of this residue to induce helical structures, suggesting that this mechanism is stronger for the Aib than for D-Iva, Hyp and its analogues. On the other hand, we find that the Peptaibols, especially Zrv-IIB, lose much of the original structure when simulated in water, and that replacing some Aibs for other residues improve the folding. This indicates that a combination of the existing Aibs with non-canonical amino acids in substitution of D-Iva or Hyp, should produce better constrained analogs. Examples of this effect can be seen in the central structures of the AAM-I analogs with MDL$_{(5,12)}$ and MPR$_{(5)}$, wherein the Aib in position 4 participates in hydrogen bonding and also the non-canonical residues. In both cases, we see an improved number of hydrogen bonds, 6 in the case of MDL$_{(5,12)}$ and 7 for MPR$_{(5)}$.

Tables 9 and 10 indicate the number of residues on a specific SS, as well as, the number of intramolecular hydrogen bond estimated with PyMOL and GROMACS for the central conformations presented on Figures 4 and 5.

Table 9. Number of residues in a specific SS (α-helix, turn or 3_{10}-helix) and number of intramolecular hydrogen bonds, obtained for the central AAM-I and chosen analogs, considering the last 60ns of simulation time.

Central Conformation	Number of residues in a SS			Number of intramolecular hydrogen bonds
	α-helix	turn	3_{10}-helix	
AAM	4 (25%)	1 (6%)	3 (19%)	4
Aib $\rightarrow$ Iva $_{(2,3,4,8,9,14)}$	5 (31%)	4 (25%)	-	6
Hyp/Pro $\rightarrow$ ALP $_{(10,13,15)}$	4 (25%)	-	4 (25%)	4
Iva $\rightarrow$ MDL $_{(5,12)}$	4 (25%)	8 (50%)	-	6
Iva $\rightarrow$ MDP $_{(5,12)}$	4 (25%)	4 (25%)	-	4
Iva $\rightarrow$ MPR $_{(5)}$	6 (38%)	4 (25%)	-	7

Table 9 shows that the central conformation of native AAM-I simulated in water has 4 residues in α-helix and 3 residues in 3_{10}-helix, agreeing with the previous analysis that indicates that this peptide is capable of maintaining a portion of its structure in helical form. In contrast, we observed that despite the fact that Zrv-IIB presents a good number of hydrogen bonds (Table 10), these interactions do not contribute to the formation of helical SS.

Table 10. Number of residues in a specific SS (α-helix, turn or 3_{10}-helix) and number of intramolecular hydrogen bonds, obtained for the central Zrv-IIB and chosen analogs, considering the last 60ns of simulation time.

Central Conformation	Number of residues in a SS			Number of intramolecular hydrogen bonds
	α-helix	turn	3_{10}-helix	
Zrv-IIB	-	8 (50%)	-	7
Aib $\rightarrow$ Iva $_{(7, 9, 12, 14)}$	9 (56%)	3 (19%)	-	8
Hyp/Pro $\rightarrow$ ALP $_{(10, 13, 15)}$	7 (44%)	-	3 (19%)	7
Hyp/Pro $\rightarrow$ MLP $_{(10, 15)}$	4 (25%)	4 (25%)	-	6
Hyp/Pro $\rightarrow$ MLP $_{(13, 15)}$	4 (25%)	5 (31%)	-	7
Iva $\rightarrow$ MDP $_{(4)}$	5 (31%)	2 (12%)	-	5

The central conformations observed for AAM-I analogs reinforce the fact that new non-canonical amino acids are capable to induce more structured peptides. Similarly, for Zrv-IIB analogues, we also find non-canonical amino acids that induce a greater portion of helical SS that native Aib, Iva and Hyp on Zrv-IIB peptaibol. We highlight Iva $_{(7, 9, 12, 14)}$ and ALP $_{(10, 13, 15)}$ that have 75% and 63% of their chain, respectively, well structured.

Conclusions

Our modeling studies concerning the insertion of asymmetrical D-α,α-dialkyl glycines and proline analogs suggest that some of the new non-canonical amino acids are more capable of inducing helical conformations and pre-organization in water than the native Aib, Iva and Hyp in AAM-I and Zrv-IIB.

Focusing on AAM-I and analogues, the peptides carrying Iva $_{(2, 3, 4, 8, 9, 14)}$, Hyp $_{(2, 3, 4, 8, 9, 14)}$, ALP $_{(10, 13, 15)}$, HLP $_{(10, 13, 15)}$, MLP $_{(10, 15)}$, MDL $_{(5, 12)}$, MDP $_{(5, 12)}$ and MPR $_{(5)}$ present a percentage of conformations containing helical SS superior that the one observed for the wild type AAM-I. Also, these peptides have a preference for type $i \rightarrow i+4$ intramolecular hydrogen

bond. Analogs carrying Iva, ALP, MDL and MPR, have the largest sums of residues in helical SS. This indicates that these amino acids are able to stabilize larger portion of the peptide in a helical SS.

The central conformations for AAM-I and chosen analogs, and data shown on Table 9, indicate that the non-canonical amino acids Iva $_{(2, 3, 4, 8, 9, 14)}$, ALP $_{(10, 13, 15)}$, MDL $_{(5, 12)}$, MDP $_{(5, 12)}$ and MPR $_{(5)}$ on these positions are the most promising residues, since they induce a good pre-organization in aqueous medium.

About Zrv-IIB and analogs, we observed that the peptides carrying Iva $_{(7, 9, 12, 14)}$, Hyp $_{(7, 9, 12, 14)}$, ALP $_{(10, 13, 15)}$, HLP $_{(10, 15)}$, MLP $_{(10, 15)}$, MLP $_{(13, 15)}$, MCP $_{(4)}$, MDL $_{(4)}$ and MDP $_{(4)}$ are the ones with percentage of conformations with helical SS higher than the observed for the native Zrv-IIB. This peptaibol naturally has only one Iva residue, which makes remarkable that the exchange by MCP, MDL or MDP induces so much structure. Analyzing the percentage of conformation with hydrogen bonds, we found that this peptide and its analogs show, in general, a greater number of $i \rightarrow i+3$ interactions than AAM-I and analogs, but also have, in most cases a high percentage of conformations with type $i \rightarrow i+4$ interaction, which is an indication of the presence of two types of SS and more structured analogs. Among these chosen analogs, we observed that the peptides carrying Iva and ALP are the ones with the highest number of residues in helical SS.

Detailed analysis on Zrv-IIB (central conformations) and analogs indicates that the non-canonical residues Iva $_{(7, 9, 12, 14)}$, ALP $_{(10, 13, 15)}$, MLP $_{(10, 15)}$, MLP $_{(13, 15)}$ and MDP $_{(4)}$ are the residues more capable to induce well-defined and helical SS on Zrv-IIB peptaibol.

Comparing both peptaibols, Iva, ALP and MDP show improved foldamer characteristics, however, further experiments using different peptides and environments are required to propose more general conclusions about their foldamer role. We emphasize that the peptaibols incorporating these residues, on the specific positions analyzed, would be the most pre-arranged in water and, may show improved insertion into the membrane.

The modeling of AAM-I, Zrv-IIB and analogues by inserting new asymmetrical α,α-dialkyl glycines and proline analogs suggest that it is possible to optimize the characteristics of native peptaibols and obtain novel peptides that have improved structural stability in water, which might translate into improved antibiotic activity.

ASSOCIATED CONTENT

Supporting Information

The parameterizations for the non-canonical amino acids discussed in this article are available as Supporting Information (Table S1). This section also presents a running average of the number of residues in α-helix, observed on the total simulation time (Figure S1). More detailed DSSP data are shown on Figures S2 and S3. This information is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*Manuel Melle-Franco

DI / Centro ALGORITMI, Universidade do Minho, 4710-057 Braga, Portugal

E-mail: manuelmelle.research@gmail.com

Contact: (+ 351) 253 604 459

References

(1)　　Chugh, J. K.; Wallace, B. A., Peptaibols: Models for Ion Channels. *Biochem. Soc. Trans.* **2001,** *29*, 565-570.

(2)　　Whitmore, L.; Wallace, B. A., The Peptaibol Database: a Database for Sequences and Structures of Naturally Occurring Peptaibols. *Nucleic Acids Res.* **2004,** *32*, D593-D594.

(3)　　Fox, R. O.; Richards, F. M., A voltage-gated ion channel model inferred from the crystal-structure of Alamethicin at 1.5-A resolution. *Nature* **1982,** *300*, 325-330.

(4)　　Leitgeb, B.; Szekeres, A.; Manczinger, L.; Vágvölgyi, C.; Kredics, L., The History of Alamethicin: A Review of the Most Extensively Studied Peptaibol. *Chem. Biodivers.* **2007,** *4*, 1027-1051.

(5)　　Shenkarev, Z. O.; Balashova, T. A.; Efremov, R. G.; Yakimenko, Z. A.; Ovchinnikova, T. V.; Raap, J.; Arseniev, A. S., Spatial Structure of Zervamicin IIB Bound to DPC Micelles: Implications for Voltage-Gating. *Biophys. J.* **2002,** *82*, 762-771.

(6)　　Snook, C. F.; Woolley, G. A.; Oliva, G.; Pattabhi, V.; Wood, S. P.; Blundell, T. L.; Wallace, B. A., The structure and function of antiamoebin I, a proline-rich membrane-active polypeptide. *Structure* **1998,** *6*, 783-792.

(7)　　Bulet, P.; Stocklin, R.; Menin, L., Anti-microbial peptides: from invertebrates to vertebrates. *Immunol. Rev.* **2004,** *198*, 169-184.

(8)　　Zasloff, M., Antimicrobial peptides of multicellular organisms. *Nature* **2002,** *415*, 389-395.

(9)　　Wallace, B. A.; Snook, C. F.; Duclohier, H.; O'Reilly, A., Antiamoebin: A polypeptide ion carrier and channel. In *Peptides for the New Millennium*, Fields, G.; Tam, J.; Barany, G., Eds. Springer Netherlands: 2002; Vol. 6, pp 733-735.

(10)　　Karle, I. L.; Perozzo, M. A.; Mishra, V. K.; Balaram, P., Crystal structure of the channel-forming polypeptide antiamoebin in a membrane-mimetic environment. *Proc. Natl. Acad. Sci. U.S.A.* **1998,** *95*, 5501-5504.

(11)　　Levtsova, O. V.; Antonov, M. Y.; Naumenkova, T. V.; Sokolova, O. S., Interaction of zervamicin IIB with lipid bilayers. Molecular dynamics study. *Comput. Biol. Chem.* **2011,** *35*, 34-39.

(12)　　Sansom, M. P.; Kerr, I.; Mellor, I., Ion channels formed by amphipathic helical peptides. *Eur. Biophys. J.* **1991,** *20*, 229-240.

(13)　　Kropacheva, T. N.; Salnikov, E. S.; Nguyen, H. H.; Reissmann, S.; Yakimenko, Z. A.; Tagaev, A. A.; Ovchinnikova, T. V.; Raap, J., Membrane association and activity of 15/16-membered peptide antibiotics: Zervamicin IIB, ampullosporin A and antiamoebin I. *Biochim. Biophys. Acta (BBA) - Biomembranes* **2005,** *1715*, 6-18.

(14)　　Bechinger, B.; Salnikov, E. S., The membrane interactions of antimicrobial peptides revealed by solid-state NMR spectroscopy. *Chem. Phys. Lipids* **2012,** *165*, 282-301.

(15)　　Weidema, A. F.; Kropacheva, T. N.; Raap, J.; Ypey, D. L., Membrane Permeabilization of a Mammalian Neuroendocrine Cell Type (PC12) by the Channel-Forming Peptides Zervamicin, Alamethicin, and Gramicidin. *Chem. Biodivers.* **2007,** *4*, 1347-1359.

(16)　　Shai, Y., Mode of action of membrane active antimicrobial peptides. *Biopolymers* **2002,** *66*, 236-248.

(17)　　Wilson, M. A.; Wei, C. Y.; Bjelkmar, P.; Wallace, B. A.; Pohorille, A., Molecular Dynamics Simulation of the Antiamoebin Ion Channel: Linking Structure and Conductance. *Biophys. J.* **2011,** *100*, 2394-2402.

(18)　　Shenkarev, Z. O.; Paramonov, A. S.; Lyukmanova, E. N.; Gizatullina, A. K.; Zhuravleva, A. V.; Tagaev, A. A.; Yakimenko, Z. A.; Telezhinskaya, I. N.; Kirpichnikov, M. P.; Ovchinnikova, T. V.; Arseniev, A. S., Peptaibol Antiamoebin I: Spatial Structure, Backbone Dynamics, Interaction with Bicelles and Lipid-Protein Nanodiscs, and Pore Formation in Context of Barrel-Stave Model. *Chem. Biodivers.* **2013,** *10*, 838-863.

(19)   Balaram, P.; Krishna, K.; Sukumar, M.; Mellor, I. R.; Sansom, M. P., The properties of ion channels formed by zervamicins. *Eur. Biophys. J.* **1992,** *21*, 117-128.

(20)   Menestrina, G.; Voges, K.-P.; Jung, G.; Boheim, G., Voltage-Dependent Channel Formation by Rods of Helical Polypeptides. *J. Membrain Biol.* **1986,** *93*, 111-132.

(21)   Kropacheva, T. N.; Raap, J., Voltage-dependent interaction of the peptaibol antibiotic zervamicin II with phospholipid vesicles. *FEBS Letters* **1999,** *460*, 500-504.

(22)   Agarwalla, S.; Mellor, I. R.; Sansom, M. S. P.; Karle, I. L.; Flippen-Anderson, J. L.; Uma, K.; Krishna, K.; Sukumar, M.; Balaram, P., Zervamicins, a structurally characterised peptide model for membrane ion channels. *Biochem. Biophys. Res. Commun.* **1992,** *186*, 8-15.

(23)   Thirumalachar, M. J., Antiamoebin, a new antiprotozoal-anthelmintic antibiotic. I. Production and biological studies. *Hindustan Antibiot. Bull.* **1968,** *10*, 287-9.

(24)   Galbraith, T. P.; Harris, R.; Driscoll, P. C.; Wallace, B. A., Solution NMR Studies of Antiamoebin, a Membrane Channel-Forming Polypeptide. *Biophys. J.* **2003,** *84*, 185-194.

(25)   Argoudel.Ad; Dietz, A.; Johnson, L. E., Zervamicins I and II, Polypeptide Antibiotics Produced by Emericellopsis-Salmosynnemata. *J. Antibiot.* **1974,** *27*, 321-328.

(26)   Karle, I. L.; Flippen-Anderson, J. L.; Agarwalla, S.; Balaram, P., Crystal structure of [Leu1]zervamicin, a membrane ion-channel peptide: implications for gating mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* **1991,** *88*, 5307-11.

(27)   Duclohier, H.; Snook, C. F.; Wallace, B. A., Antiamoebin can function as a carrier or as a pore-forming peptaibol. *Biochim. Biophys. Acta (BBA) - Biomembranes* **1998,** *1415*, 255-260.

(28)   Ballesteros, J. A.; Weinstein, H., The role of Pro/Hyp-Kinks in Determining the Transmembrane Helix Length and Gating Mechanism of a Leu Zervamicin Channel. *Biophys. J.* **1992,** *62*, 110-111.

(29)   Sansom, M. S. P., The Biophysics of Peptide Models of Ion Channels. *Prog. Biophys. Mol. Biol.* **1991,** *55*, 139-235.

(30)   Castro, T. G.; Micaêlo, N. M., Modeling of Peptaibol Analogues Incorporating Nonpolar α, α-Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation. *J. Phys. Chem. B* **2014,** *118*, 649-658.

(31)   Castro, T. G.; Micaelo, N. M., Conformational and Thermodynamic Properties of Non-Canonical alpha,alpha-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies. *J. Phys. Chem. B* **2014,** *118*, 9861-9870.

(32)   Roos, E. C.; Lopez, M. C.; Brook, M. A.; Hiemstra, H.; Speckamp, W. N.; Kaptein, B.; Kamphuis, J.; Schoemaker, H. E., Synthesis of .alpha.-substituted .alpha.-amino acids via cationic intermediates. *J. Org. Chem.* **1993,** *58*, 3259-3268.

(33)   Mendel, D.; Ellman, J.; Schultz, P. G., Protein biosynthesis with conformationally restricted amino acids. *J. Am. Chem. Soc.* **1993,** *115*, 4359-4360.

(34)   Baures, P. W.; Ojala, W. H.; Gleason, W. B.; Johnson, R. L., Conformational analysis of homochiral and heterochiral diprolines as β-turn-forming peptidomimetics: unsubstituted and substituted models. *J. Pept. Res.* **1997,** *50*, 1-13.

(35)   Caumes, C.; Delsuc, N.; Azza, R. B.; Correia, I.; Chemla, F.; Ferreira, F.; Carlier, L.; Luna, A. P.; Moumne, R.; Lequin, O.; Karoyan, P., Homooligomers of substituted prolines and [small beta]-prolines: syntheses and secondary structure investigation. *New J. Chem.* **2013,** *37*, 1312-1319.

(36)   Bach, T.; Takagi, H., Properties, metabolisms, and applications of l-proline analogues. *Appl. Microbiol. Biotechnol.* **2013,** *97*, 6623-6634.

(37)   Tsogoeva, S. B.; Jagtap, S. B.; Ardemasova, Z. A., 4-trans-Amino-proline based di- and tetrapeptides as organic catalysts for asymmetric C–C bond formation reactions. *Tetrahedron: Asymmetry* **2006,** *17*, 989-992.

(38)   Torino, D.; Mollica, A.; Pinnen, F.; Feliciani, F.; Spisani, S.; Lucente, G., Novel chemotactic For-Met-Leu-Phe-OMe (fMLF-OMe) analogues based on Met residue replacement by 4-amino-proline scaffold: Synthesis and bioactivity. *Bioorg. Med. Chem.* **2009**, *17*, 251-259.

(39)   Schrödinger *The PyMOL Molecular Graphics System*, 1.3r1; LLC: 2010.

(40)   Huang, W.; Lin, Z. X.; van Gunsteren, W. F., Validation of the GROMOS 54A7 Force Field with Respect to beta-Peptide Folding. *J. Chem. Theory. Comput.* **2011**, *7*, 1237-1243.

(41)   Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory. Comput.* **2008**, *4*, 435-447.

(42)   Spoel, D. v. d.; Lindahl, E.; Hess, B.; Buuren, A. R. v.; Apol, E.; Meulenhoff, P. J.; Tieleman, P.; Sjibers, A. L. T. M.; Feenstra, K. A.; Drunen, R. v.; Berendsen, H. J. C., Gromacs user manual version 4.5. 2010. http://www.gromacs.org.

(43)   Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The missing term in effective pair potentials. *J. Phys. Chem-US* **1987**, *91*, 6269-6271.

(44)   Smith, P. E.; Vangunsteren, W. F., Consistent dielectric-properties of the simple point-charge and extended simple point-charge water models at 227 and 300K. *J. Chem. Phys.* **1994**, *100*, 3169-3174.

(45)   Hess, B., P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory. Comput.* **2008**, *4*, 116-122.

(46)   Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463-1472.

(47)   van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C., A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* **1998**, *108*, 10220-10230.

(48)   Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.

(49)   Aravinda, S.; Shamala, N.; Balaram, P., Aib Residues in Peptaibiotics and Synthetic Sequences: Analysis of Nonhelical Conformations. *Chem. Biodivers.* **2008**, *5*, 1238-1262.

(50)   Crisma, M.; Andreetto, E.; De Zotti, M.; Moretto, A.; Peggion, C.; Formaggio, F.; Toniolo, C., Crystal-state 3D-structural characterization of novel, Aib-based, turn and helical peptides. *J. Pept. Sci.* **2007**, *13*, 190-205.

(51)   Karle, I. L.; Balaram, P., Structural Characteristics of Alpha-Helical Peptide Molecules Containing Aib Residues. *Biochemistry* **1990**, *29*, 6747-6756.

(52)   Shoulders, M. D.; Raines, R. T., Collagen Structure and Stability. *Annual review of biochemistry* **2009**, *78*, 929-958.

(53)   Kabsch, W.; Sander, C., Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577-2637.

Chapter VI

New Self-Assembled Supramolecular Hydrogels Based on Dehydropeptides

Theoretical part only.

Journal of
Materials Chemistry B

PAPER

New self-assembled supramolecular hydrogels based on dehydropeptides

H. Vilaça,[a] G. Pereira,[a] T. G. Castro,[a] B. F. Hermenegildo,[b] J. Shi,[c] T. Q. Faria,[d] N. Micaêlo,[a] R. M. M. Brito,[d] B. Xu,[c] E. M. S. Castanheira,[b] J. A. Martins*[a] and P. M. T. Ferreira*[a]

Supramolecular hydrogels rely on small molecules that self-assemble in water as a result of the cooperative effect of several relatively weak intermolecular interactions. Peptide-based low molecular weight hydrogelators have attracted enormous interest owing to the simplicity of small molecules combined with the versatility and biocompatibility of peptides. In this work, naproxen, a well known non-steroidal anti-inflammatory drug, was *N*-conjugated with various dehydrodipeptides to give aromatic peptide amphiphiles that resist proteolysis. Molecular dynamics simulations were used to obtain insight into the underlying molecular mechanism of self-assembly and to rationalize the design of this type of hydrogelators. The results obtained were in excellent agreement with the experimental observations. Only dehydrodipeptides having at least one aromatic amino acid gave hydrogels. The characterization of the hydrogels was carried out using transmission electron microscopy (TEM), circular dichroism (CD), fluorescence spectroscopy and also rheological assays.

Introduction

The preparation of biomaterials, such as hydrogels, using a "bottom-up" approach is based on molecular self-assembly through non-covalent interactions such as hydrogen bonding, van der Waals forces and π–π and electrostatic interactions. Small peptides with bulky aromatic moieties can self-assemble into nanostructures that interweave giving three-dimensional (3D) networks that entrap water giving biocompatible and biodegradable hydrogels. These biomaterials have a wide range of applications, from drug delivery to tissue engineering and regenerative medicine.[1–8] One limitation of peptide based hydrogelators is their susceptibility to enzymatic hydrolysis, which shortens their *in vivo* lifetime and narrows the scope of their application. One of the strategies used to circumvent this limitation is to introduce non-proteinogenic amino acids into peptide hydrogelators.[9–13] Recently Xu *et al.* described the synthesis of new hydrogelators based on dipeptides containing unnatural D-amino acids *N*-capped with naproxen.[14] The D-amino acids in the conjugates not only increased the proteolytic stability of the hydrogelators but

also enhanced their selectivity for inhibiting cyclooxygenase-2 (COX-2). The same authors also prepared and studied other peptides *N*-conjugated with other NSAID drugs, namely ibuprofen and flurbiprofen.[15]

In this work, a multidisciplinary approach that combines molecular dynamics simulations with experimental results was devised for developing new efficient dehydropeptide hydrogelators. Five dehydrodipeptides *N*-conjugated with naproxen were prepared and studied. These compounds were designed taking into consideration several factors: dehydro-amino acids[16–21] are known to increase the resistance of peptides against proteolytic enzymes, the naproxen-capped hydrogelators (and hydrogels) are likely to retain the NSAID properties of naproxen;[14,15] the naphthalene moiety of naproxen is prone to engage in intermolecular π–π stacking interactions as described for other peptide hydrogels functionalised with naphthalene moieties.[12,14,15,22]

The goal of this work was to understand the self-assembly behaviour of aromatic dehydrodipeptide amphiphiles and to create a rational basis for the design of new dehydropeptide hydrogelators. Molecular dynamics simulations, together with characterization assays [circular dichroism (CD), fluorescence spectroscopy, transmission electron microscopy (TEM) and rheometry], evidenced the propensity of dehydrodipeptides containing an aromatic amino acid and *N*-conjugated to a polyaromatic moiety to self-assemble into nanostructures that give hydrogels. Furthermore, this new class of hydrogelators are found to resist proteolytic degradation.

[a] *Centre of Chemistry, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. E-mail: pmf@quimica.uminho.pt*
[b] *Centre of Physics, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal*
[c] *Centre for Neuroscience and Cell Biology, University of Coimbra, Coimbra 3004-517, Portugal*
[d] *Department of Chemistry, Brandeis University, Waltham, MA, 02454, USA*

This journal is © The Royal Society of Chemistry 2015

J. Mater. Chem. B, 2015, **3**, 6355–6367 | 6355

Results and discussion

Synthesis

Five new dehydrodipeptides *N*-protected with naproxen (Npx) were prepared from the corresponding methyl esters of *N-tert*-butoxycarbonyl-β-hydroxydipeptides. The strategy employed involved a dehydration reaction followed by the cleavage of the *tert*-butoxycarbonyl group (Boc), reaction with (*S*)-(+)-naproxen chloride and alkaline hydrolysis of the methyl esters (Scheme 1). The dehydroamino acids used were dehydrophenylalanine (ΔPhe) and dehydroaminobutyric acid (ΔAbu). This synthetic methodology was chosen to avoid racemization issues concerning the naproxen moiety. The *N,C*-diprotected dipeptides having a β-hydroxyamino acid (Scheme 1, **1a–e**) were dehydrated in good to high yields by treatment with *tert*-butyldicarbonate (Boc$_2$O) and 4-dimethylaminopyridine (DMAP) followed by *N,N,N′,N′*-tetramethylguanidine (TMG)[17] (Scheme 1, **2a–e**). The Boc group was removed using trifluoroacetic acid (TFA) (Scheme 1, **3a–e**) and the *N*-deprotected dehydrodipeptides were conjugated with (*S*)-(+)-naproxen (Scheme 1, **4a–e**). Finally, the methyl esters were removed by treatment with a solution of NaOH (1 M) affording compounds **5a–e** in good yields (Scheme 1).

The stereochemistry of compounds **2–5** was confirmed by Nuclear Overhauser (NOE) difference experiments by irradiating the α-NH proton of the dehydroamino acid residue and observing NOE enhancements in the signals of the β-methyl or β-phenyl protons. All the ^{1}H NMR spectra of compounds **5a–e** in dimethyl-sulfoxide (DMSO-d_6) show one doublet and two singlets between 8.11 pm and 12.67 ppm due to the NH and CO$_2$H protons. The β-CH proton of the dehydroamino acid residues appears in the aromatic region in the case of dehydrophenylalanine and as a quartet near 6.5 ppm for dehydroaminobutyric acid. In the ^{13}C NMR spectra of these compounds the signals assigned to the β-carbon atoms of the dehydroamino acid residues appear in a narrow zone of a high chemical shift between 131.81 ppm and 133.14 ppm. This is due to deprotection resulting from conjugation of the α,β-double bond with the carbonyl group.
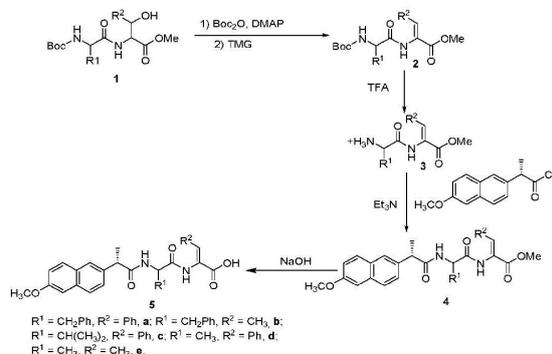
Scheme 1 Synthesis of dehydrodipeptides *N*-conjugated to naproxen (Npx): Npx-L-Phe-*Z*-ΔPhe-OH, **5a**; Npx-L-Phe-*Z*-ΔAbu-OH, **5b**; Npx-L-Val-*Z*-ΔPhe-OH, **5c**; Npx-L-Ala-*Z*-ΔPhe-OH, **5d**; Npx-L-Ala-*Z*-ΔAbu-OH, **5e**.

Molecular dynamics simulations

Molecular dynamics simulations (MD) were carried out for all dehydrodipeptides prepared (**5a–e**) to examine the spontaneity of the self-assembly process. The peptides were placed in cubic boxes of 4.5 × 4.5 × 4.5 mm solvated with the SPC water model.[23] The average number of clusters observed for each peptide is presented in Table 1. This analysis was carried out by calculating the number of peptides that are clustered using a cut-off of 1.4 nm for each simulation frame and averaged over the last 20 ns sampling time. The average number of inter-molecular hydrogen bonds was calculated and averaged over the same time interval. The number of π-stacking interactions was divided by the number of peptides present in the simulation box and normalized by the number of simulated frames (last 20 ns). This average percentage of π–π interaction is read as the number of π-stacking/number of peptides in solution. Different types of π-stacking interactions considered were: sandwich (S), parallel displaced (PD) and T-shaped (T). Experimental and theoretical data[24] were used to define the cut-off distances and angles between aromatic groups that characterize these types of π-stacking interactions. The intermolecular π-stacking interactions analyzed were those between naproxen groups, naproxen and the aromatic moieties of phenylalanine (Phe) and dehydrophenylalanine (ΔPhe) and between the phenyl groups of the two aromatic amino acid residues. All these types of π-stacking interactions could be found solely for peptide **5a**, the other peptides show only some of these types of interactions. In the case of peptide **5e** the only π-stacking interactions observed are those between the naproxen moieties.

Fig. 1 shows the clustering arrangement of a given frame taken from the simulation trajectory of **5a–e**.

Visual analysis of the clustering behaviour within each frame clearly shows the self-aggregation of peptides **5a–d**, and, poor aggregation of peptide **5e**. Hydrogelators **5b** and **5c** show the lowest numbers of clusters (Fig. 1 and Table 1), which means a higher extent of aggregation. The number of inter-molecular hydrogen bonds is very low and comparable for all systems (Table 1) suggesting that this type of molecular inter-action does not explain the aggregation properties observed for peptides **5a–d**. The aggregation phenomenon seems to be better explained by the formation of intermolecular π-stacking interactions between the naproxen moiety and the aromatic amino acids.

Fig. 2 shows the representative π-stacking interactions observed for compound **5b**.

Naproxen π-stacking, preferentially as the S type mode, occurs in all five dehydrodipeptides studied. PD π-stacking is also significant (Fig. 2a). It is interesting to note that naproxen π-stacking alone cannot explain the aggregation phenomenon. Although peptide **5e** presents a high percentage of naproxen π-stacking interactions, the theoretical simulations suggest that this peptide is not able to self-aggregate (Fig. 1). This leads to the conclusion that the sole intermolecular interaction between naproxen groups is not sufficient to promote peptide aggrega-tion. On the other hand, the interactions between naproxen and

6356 | *J. Mater. Chem. B*, 2015, **3**, 6355–6367

This journal is © The Royal Society of Chemistry 2015

134

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo,
R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B,* 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Table 1 Average number of clusters, hydrogen bonds and percentage of intra/intermolecular π-stacking interactions observed for each system. The total number of π-stacking interactions and also the individual contribution of each π-stacking geometry, S, PD and T, are shown. Standard deviation for the first two analyses is shown in parenthesis

		Npx-Phe-ΔPhe-OH, 5a	Npx-Phe-ΔAbu-OH, 5b	Npx-Val-ΔPhe-OH, 5c	Npx-Ala-ΔPhe-OH, 5d	Npx-Ala-ΔAbu-OH, 5e
Average number of clusters		3.2 (±0.10)	2.4 (±0.09)	2.8 (±0.08)	3.8 (±0.10)	4.5 (±0.09)
Average number of hydrogen bonds		3.8 (±0.04)	4.7 (±0.04)	3.5 (±0.04)	3.2 (±0.05)	3.7 (±0.04)
Percentage of π stacking interactions[a]	Npx-Npx	8.5 (S: 5.1; PD: 2.0; T: 1.4)	6.5 (S: 4.3; PD: 1.1; T: 1.1)	7.0 (S: 4.5; PD: 1.5; T: 1.0)	6.3 (S: 4.4; PD: 1.5; T: 0.4)	11.3 (S: 7.7; PD: 2.3; T: 1.3)
	Npx-Phe	3.6 (S: 1.9; PD: 0.7; T: 1.0)	4.5 (S: 2.5; PD: 1.1; T: 0.9)	—	—	—
	Npx-ΔPhe	1.7 (S: 0.8; PD: 0.4; T: 0.5)	—	2.7 (S: 1.3; PD: 0.5; T: 0.9)	2.0 (S: 0.9; PD: 0.4; T: 0.7)	—
	Phe-Phe	0.5 (S: 0.1; PD: 0.1; T: 0.3)	0.6 (S: 0.2; PD: 0.1; T: 0.3)	—	—	—
	ΔPhe-ΔPhe	0.4 (S: 0.1; T: 0.3)	—	0.1 (T: 0.1)	0.2 (PD: 0.1; T: 0.1)	—
	Phe-ΔPhe	0.8 (S: 0.1; PD: 0.2; T: 0.5)	—	—	—	—
	Total	15.5	11.6	9.8	8.5	11.3

[a] Sandwich (S): $R \leq 4.5$ Å and $\theta \leq 15°$ or $\theta \geq 165°$; parallel displaced (PD): $R \leq 4.0$ Å and $15° < \theta < 30°$ or $150° < \theta < 165°$; T-shaped (T): $R \leq 3.5$ Å and $\theta \leq 10°$.
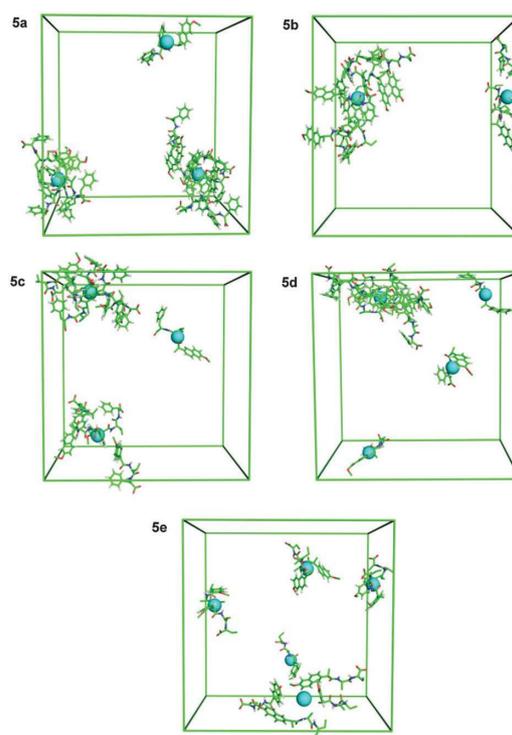
Fig. 1 Snapshot of the MDS of peptides **5a–e** after equilibration. The water molecules have been omitted for better viewing of the peptides. The spheres in cyan represent the geometric center of each cluster. Figures were obtained with Pymol.[25]

other aromatic groups seem to be responsible for the aggregation of the peptides **5a–d**. Analyzing the systems of peptides **5a** or **5b** suggests that the self-aggregation phenomenon is determined by the intermolecular interaction between naproxen and the phenyl group of the aromatic amino acid (Fig. 2b). The system containing peptide **5a** also suggests that the presence of two aromatic amino acid residues does not show any additive effect on cluster formation, since the **5a** system shows less self-aggregation than **5b**. In fact, combining two aromatic residues such as phenylalanine and dehydrophenylalanine seems to have a detrimental effect on self-aggregation. The systems containing peptides **5c** and **5d**, indicate that self-aggregation in dehydro-dipeptides with a single aromatic residue in the C-terminal position is less effective. Systems containing peptides **5a**, **5c** and **5d** also demonstrate that the dehydrophenylalanine residue does not establish significant intermolecular interactions within itself, instead, this amino acid seems to interact preferentially with naproxen. Furthermore, although **5c** and **5d** are structurally similar, **5c** shows a higher propensity to form clusters. This could result from the presence of valine that makes peptide **5c** less polar than peptide **5d**. From these results it is possible to conclude that the aggregation of this type of peptides is dominated by π–π interactions between the N-aromatic component

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B,* 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).
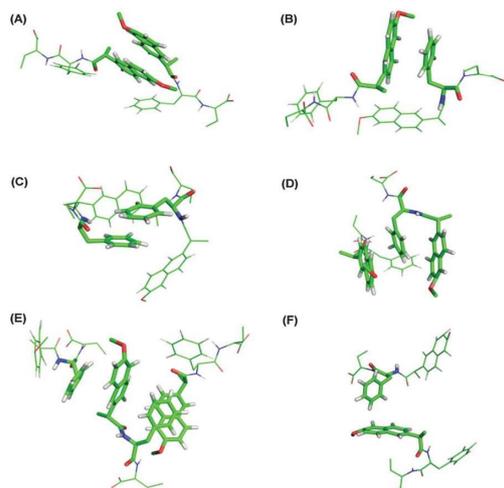
Fig. 2 Illustration of representative π-stacking interactions observed for hydrogelator **5b** during the last 20 ns of the simulation. (A) Intermolecular interaction between the naproxen moieties, (B) intermolecular interaction between Npx and Phe, (C) intermolecular interaction between Phe groups, (D) intra and intermolecular interactions, (E) multiple intermolecular interactions and (F) T-shaped interaction between Phe and Npx.

and other aromatic amino acid moieties. Aromatic amino acid residues in addition to the *N*-aromatic capping group are required for peptide aggregation (as seen for **5e**).

Hydrogelation

Dehydrodipeptides **5a–e** were tested as hydrogelators in order to validate this methodology for the rational design of efficient hydrogelators. Gelation was triggered *via* pH change and/or heating and subsequent cooling. Compounds **5a** and **5b** were solubilized in PBS buffer at 60 °C and gelation occurred upon cooling to room temperature (Fig. 3). Compounds **5c** and **5d** were dissolved in water with the addition of NaOH (1 M) and gelation occurred by pH adjustment with HCl (1 M) (Fig. 3). The results showed that gelation of compounds **5a–d** occurs at low
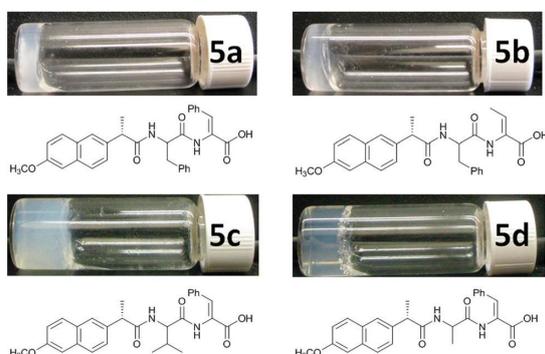
Fig. 3 Optical images of hydrogels **5a–d**.

Table 2 CGC and gel–sol phase transition pH of compounds **5a–d**

Compound	CGC [wt%]	pHgs
Npx-L-Phe-Z-ΔPhe-OH, **5a**	0.4	8.0
Npx-L-Phe-Z-ΔAbu-OH, **5b**	0.4	6.0
Npx-L-Val-Z-ΔPhe-OH, **5c**	0.6	8.0
Npx-L-Ala-Z-ΔPhe-OH, **5d**	0.8	5.0

critical gelation concentrations (CGC), between 0.4 wt% and 0.8 wt% and a gel–sol phase transition pH (pHgs) between 5 and 8 (Table 2).

As predicted by molecular dynamics simulations, compound **5e** (Npx-L-Ala-Z-ΔAbu-OH), lacking an aromatic amino acid residue, failed to give a hydrogel under all conditions tested. The dehydropeptides with a capped *N*-terminal phenylalanine residue (**5a** and **5b**) display lower CGC compared to peptides **5c** and **5d**, bearing an alkyl *N*-terminal amino acid (Val or Ala). These experimental results are in excellent agreement with those obtained by the molecular dynamics simulations. This means that the molecular modelling methodology presented here might be a valuable new tool for the design of efficacious peptide hydrogelators.

Comparing the experimental conditions for hydrogelation of dehydrodipeptide **5a** with the dipeptide phenylalanylphenylalanine *N*-protected with naproxen (0.4 wt% and pH 8.0 *vs.* 0.8 wt% and pH 7, respectively),[14] it is possible to conclude that the presence of the α,β-double bond decreases the CGC value and increases the gel–sol phase transition pH.

Hydrogelators **5a** and **5b**, containing different dehydro-amino acids, dehydrophenylalanine and dehydroaminobutyric acid, respectively, were selected for further characterization, namely fluorescence studies, circular dichroism (CD) studies and proteolitic stability assays, to get insight into the effect of the structure of the dehydroaminoacid on the self-assembly and gelation processes and proteolitic stability. The properties of hydrogelator **5a** can be directly compared to its "natural" dipeptide analogue phenylalanylphenylalanine *N*-protected with naproxen.[14] Moreover, hydrogelator **5a** (bearing two aromatic amino acid residues) was expected to display fluorescence and CD spectra more insightful towards the self-assembly and gelation processes than **5c** and **5d** (containing only one aromatic amino acid residue). Rheological and Transmission Electron Microscopy (TEM) characterisation was carried out with gels **5a–d** (**5e** failed to produce a hydrogel).

Photophysical studies

The critical aggregation concentration (CAC) as well as the influence of pH on the aggregation of peptides **5a** and **5b** were studied by fluorescence spectroscopy. Fig. 4 shows the influence of pH on the fluorescence properties of molecules **5a** and **5b**. The fluorescence emission of peptides **5a** and **5b** is clearly dominated by the emission of the naproxen moiety, λ_{max} = 353 nm (λ_{exc} = 290 nm), similar to the results reported for naproxen in methanol and water.[26] However, it is possible to observe a second fluorescence band, with maximum emission near 440 nm. This band is ascribed to the formation of an

6358 | *J. Mater. Chem. B*, 2015, **3**, 6355–6367

This journal is © The Royal Society of Chemistry 2015

136

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B*, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).
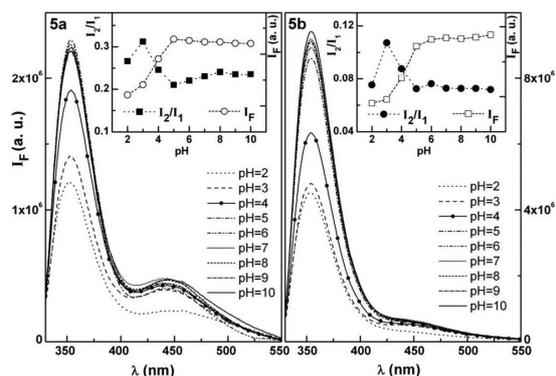
Fig. 4 Fluorescence spectra of compounds **5a** and **5b** (2×10^{-6} M) at different pH values (λ_{exc} = 290 nm). Insets: variation of the maximum fluorescence intensity and intensity ratio I_2/I_1 with pH.

Fig. 5 Excitation spectra (λ_{em} = 360 nm and λ_{em} = 450 nm) of compounds **5a** and **5b** (2×10^{-6} M) at selected representative pH values. Insets: normalized spectra at the peak of minimum energy.

emissive dimmer between naproxen and the phenylalanine residues. At the excitation wavelength used (λ_{exc} = 290 nm), phenylalanine is not electronically excited.

For both compounds **5a** and **5b**, the maximum emission intensity rises with pH, with a tendency to stabilize for pH values above 5 for compound **5a** and 6 for compound **5b**. Considering the ratio between the aggregate band and the naproxen monomer band, I_2/I_1, a different behaviour is observed: the maximum value of I_2/I_1 occurs at pH 3 for both compounds, near the pK_a value of the peptide terminal carboxylic acid group (p$K_a \sim 3$).[27] For both compounds, a minimum is observed at pH 5 with stabilization observed thereafter. A slight rise in the I_2/I_1 ratio is observed at pH 8 for compound **5a** and pH 6 for compound **5b**, identified as the pHgs values. The pH at which a gel is formed is highly dependent on the molecular structure of the hydrogelator and correlates with the apparent pK_a of the peptide.[28] The extent of deprotonation of the carboxylic acid group on the hydrogelators is pH-dependent and determines their hydrophilicity. Accordingly, compound **5b**, bearing only one aromatic amino acid, is more hydrophilic than **5a** (bearing two aromatic amino acid residues) thus requiring lower pH (pH 0.6) for gelation compared to **5a** (pH 8.0).

Excitation spectra provide relevant information about the nature of the aggregate emission band (Fig. 5). It can be observed that upon collection of emission in the naproxen band (360 nm) or in the aggregate band (450 nm), the spectra completely modify showing different excited species, and not dynamic exciplexes were formed at the excited state.

Fig. 6 shows the dependence of fluorescence emission of compounds **5a** and **5b** on concentration.

It can be seen that the ratio of intensities of the aggregate and monomer bands, I_2/I_1 (insets of Fig. 6), is almost constant for concentrations below 0.4 wt%, but increases sharply for higher concentrations. At this concentration, a clear change in the naproxen maximum emission wavelength is also detected, with a red shift that tends to stabilize at higher concentrations (inset of Fig. 6, left). These results point clearly to hydrogel formation at 0.4 wt% for compound **5a**. For compound **5b**, the
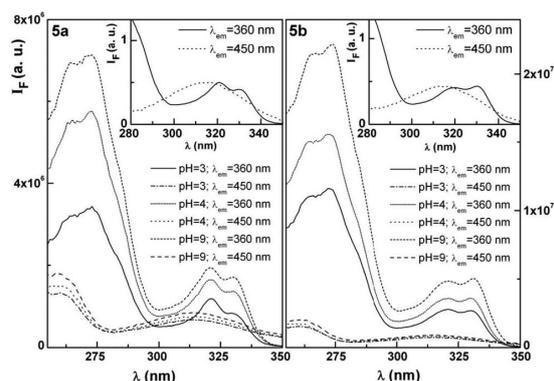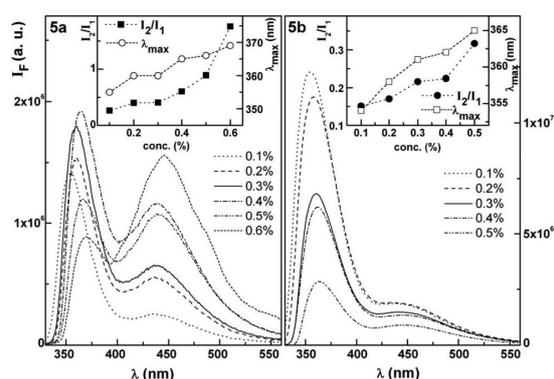
Fig. 6 Fluorescence spectra of compounds **5a** (pH 8) and **5b** (pH 6) at several concentrations (λ_{exc} = 290 nm). Insets: variation with the concentration of the maximum emission wavelength of the first band and intensity ratio I_2/I_1.

ratio I_2/I_1 follows the same trend (presenting lower values) with a noticeable rise above 0.4 wt%, as for compound **5a**. The red shift in naproxen emission with increasing hydrogelator concentration is smaller than that observed for compound **5a**. From these two indicators, a clear change in behaviour is detected above 0.4 wt%, pointing to hydrogel formation (inset of Fig. 6, right). Above 0.4 wt% concentration, the aggregate emission band is clearly higher than what is observed at lower concentrations, which may indicate the formation of intermolecular aggregates that play an important role in gel formation.

These results show that fluorescence spectroscopy is a good methodology to estimate the critical gelation concentration and to get insight into the intramolecular/intermolecular interactions between the aromatic moieties of these molecules.

Hydrogel characterization

The CD spectra of hydrogelators **5a** and **5b** and of the dipeptide Npx-phenylalanylphenylalanine (Npx-L-Phe-L-Phe-OH) were

137

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B,* 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Fig. 7 CD spectra of Npx-L-Phe-L-Phe-OH (A), Npx-Phe-ΔPhe-OH **5a** (B) and Npx-Phe-ΔAbu-OH **5b** (C).

obtained for Npx-L-Phe-L-Phe-OH. This suggests the same type of intermolecular interactions and the β-sheet like structure. However, in Npx-L-Phe-L-Phe-OH the strongest bands are the ones resulting from the peptide backbone (195 nm and 220 nm), while in dehydrodipeptides **5a** and **5b**, the most intense bands originate from the naphthalene interactions (220 nm and 235 nm). This indicates that for the dehydrodipeptide hydrogelators the naphthalene interactions are more important than the peptide backbone arrangement. The variation of the CD spectra of these three compounds with temperature shows a progressive loss of the structure up to 40 °C. For higher temperatures, the absence of CD signals suggests high mobility of the peptide backbone and the lack of an organized structure.[31] Cooling the peptide solutions leads to gelation, shown by the enhancement of the CD signals[31] and the blue shift of the λ_{max} of the bands. This indicates a gradual transition from an isotropic solution to an anisotropic environment in the gel state.[35] In the case of Npx-L-Phe-L-Phe-OH, the structure formation is abrupt and occurs at temperatures below 40 °C, where the signal strength increases rapidly as the temperature decreases. For compound **5a** the appearance of organized structures starts at a slightly higher temperature (around 60 °C), suggesting that the dehydrophenylalanine residue increases the propensity for hydrogelation at higher temperatures. The hydrogelator **5b** showed a behaviour similar to that observed for Npx-L-Phe-L-Phe-OH. According to these results it is possible to conclude that the dehydrophenylalanine residue increases the hydrogel thermal stability.

Morphological analysis of the new hydrogels based on dehydrodipeptides was carried out by transmission electron microscopy (TEM). The TEM images of hydrogels obtained from compounds **5a–d** are shown in Fig. 8.

Hydrogelator **5a** self-assembles into non-uniform nanofibers displaying different widths: a minimum width of 9 nm and a maximum width of 18 nm, similar to those shown by Npx-L-Phe-L-Phe-OH.[14] Hydrogel **5b** exhibits long nanofibers that

recorded in order to get insight into the peptide secondary structure (Fig. 7).

The three peptides show similar CD spectra, displaying Cotton effects only in the far UV wavelength (190–260 nm). The CD spectra of Npx-L-Phe-L-Phe-OH (Fig. 7A) at low temperatures exhibit bands around 196 nm (positive peak), 220 nm (broad negative peak) and 235 nm (positive Cotton effect). The signals at 196 nm and 220 nm indicate a β-sheet like arrangement of the peptide backbone, corresponding to π–π* and n–π* transitions, respectively.[29–34] The CD spectra of compounds **5a** and **5b** (Fig. 7B and C) at low temperatures are similar to those

Fig. 8 TEM images of hydrogels obtained from dehydrodipeptides (A) **5a**; (B) **5b**; (C) **5c**; (D) **5d** (scale bar of 100 nm).

6360 | J. Mater. Chem. B, 2015, **3**, 6355–6367

This journal is © The Royal Society of Chemistry 2015

138

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo,
R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: J. Mater. Chem. B, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Table 3 Rheological properties of hydrogels formed by naproxen-dehydropeptides **5a–d**

Hydrogel[a]	Dynamic strain sweep			Dynamic frequency sweep	
	G_{max}' [Pa]	G_{max}'' [Pa]	Critical strain [%]	G'^b [Pa]	G''^b [Pa]
5a	1.6×10^3	2.2×10^2	5.0	1.7×10^3	2.2×10^2
5b	8.1×10^2	92.7	1.6	7.1×10^2	79.3
5c	5.9×10^2	1.1×10^2	0.3	6.6×10^2	89.1
5d	9.8×10^2	1.0×10^2	8.0	8.0×10^2	1.17×10^2

[a] The concentration of the hydrogel is 0.4 wt% for compounds **5a** and **5b**, 0.6 wt% for compound **5c** and 0.8 wt% for compound **5d**. [b] The value is taken at 6.32 rad s^{-1}.

entangle to form a network, with an average width of 10 nm. The nanofibers of hydrogelator **5c** are short (length between 170–750 nm) and non-uniform displaying a minimum width of 12 nm and a maximum width of 16 nm. Hydrogel **5d** comprises of long and entangled nanofibers with widths ranging between 8 and 16 nm.

The rheological data obtained with hydrogels **5a–d** are presented in Table 3.

All hydrogels presented a storage modulus (G') significantly higher than their loss modulus (G'') and independent of frequency, which indicates a viscoelastic behaviour. The hydrogels of compounds **5d** and **5a** present greater storage modulus and critical strains (8.0% and 5.0%, respectively), suggesting a more resilient network in these two hydrogels. Comparison between the critical strains of hydrogels of Npx-L-Phe-Z-ΔPhe-OH (**5a**), Npx-L-Phe-L-Phe-OH (0.62%)[15] and Npx-D-Phe-D-Phe-OH (1.0%)[14] shows that the α,β-double bond in **5a** increases the resistance of this gel to an external force.

Enzymatic and toxicity assays

The stability of the new dehydrodipeptide hydrogelators **5a** and **5b** against proteolytic degradation with α-chymotrypsin was compared to Npx-L-Phe-L-Phe-OH (Fig. 9). α-Chymotrypsin was chosen for its ability to preferentially cleave peptide amide
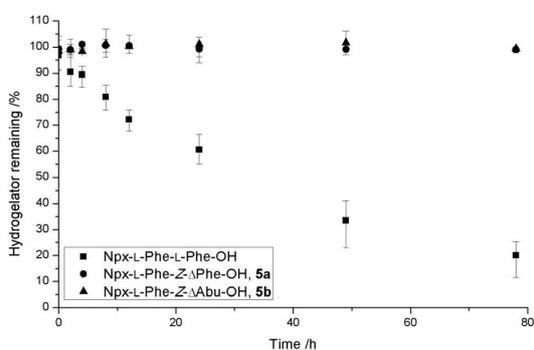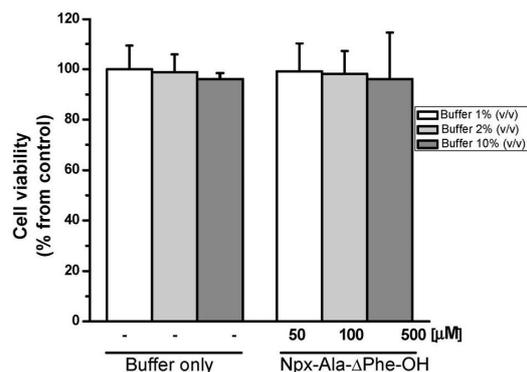
Fig. 10 Cell viability of adult human skin fibroblasts after incubation for 48 hours with 50 μM, 100 μM or 500 μM of Npx-L-Ala-Z-ΔPhe-OH (**5d**), as compared with buffer controls. No significant differences were observed ($P > 0.05$).

bonds where the carboxyl side of the amide bond is an aromatic amino acid. Thus, the peptide bond between the Phe residues on the control substrate Npx-L-Phe-L-Phe-OH and the peptide bond between the Phe residue and the dehydroamino acid residue in peptides **5a** (Npx-L-Phe-Z-ΔPhe-OH) and **5b** (Npx-L-Phe-Z-ΔAbu-OH) are the likely cleavage sites for chymotrypsin. Dehydropeptides **5c** (Npx-L-Val-Z-ΔPhe-OH) and **5d** (Npx-L-Ala-Z-ΔPhe-OH) lacking an aromatic amino acid residue in position P1 are not likely to be recognized by chymotrypsin as substrates. The results show that while the control substrate Npx-L-Phe-L-Phe-OH undergoes fast proteolytic degradation, the dehydrodipeptides **5a** and **5b** are completely stable when treated with α-chymotrypsin for 80 hours (Fig. 9). The capping N-terminal amide bond of naproxen on peptides **5a** and **5b** was also found to be stable towards chymotrypsin-catalysed hydrolysis.

Replacement of the C-terminal Phe residue on the control substrate by a dehydroamino acid renders the peptide bond resistant to hydrolysis.

A preliminary evaluation of the cellular toxicity of the dehydrodipeptide hydrogelators was carried out on adult human skin fibroblasts.

Cell viability after incubation for 48 hours with 50 μM, 100 μM and 500 μM of hydrogelator **5d** is presented in Fig. 10. The dehydrodipeptide did not show toxicity, even at concentrations as high as 500 μM. Hydrogelator **5d** was selected for testing due to its higher solubility at 37 °C in the cell culture medium, compared to hydrogelators **5a–c**, thanks to its relatively high CGC (presumably substantially higher at physiological pH). This allowed the studying of the effect of the hydrogelator on cellular viability for a wide range of concentrations, up to 500 μM, without potential interference from nano/microstructures that form before macroscopic gel formation can be detected. For relevant biological applications the biocompatibility of hydrogels needs also to be assessed. The biocompatibility of the hydrogels as cell culture media will be evaluated in the near future.

Fig. 9 Evaluation of the proteolitic stability of hydrogelators Npx-L-Phe-L-Phe-OH; Npx-L-Phe-Z-ΔPhe-OH, **5a** and Npx-L-Phe-Z-ΔAbu-OH, **5b** in the presence of α-chymotrypsin (pH 7.4, 37 °C) for 80 hours.

139

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B,* 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Experimental

Melting points (°C) were determined in a Gallenkamp apparatus and are uncorrected. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker Avance III at 400 and 100.6 MHz, respectively, or in a Varian Unity Plus 300 at 300 and 75.4 MHz, respectively. ^{1}H–^{1}H spin–spin decoupling and DEPT θ 45° were used. HMQC and HMBC were used to attribute some signals. Chemical shifts (δ) are given in parts per million (ppm) and coupling constants (J) in hertz (Hz). High resolution mass spectrometry (HRMS) data were recorded by the mass spectrometry service of the University of Vigo, Spain. Elemental analysis was performed on a LECO CHNS 932 elemental analyzer. Column chromatography was performed on a Macherey–Nagel silica gel 230–400 mesh. Petroleum ether refers to the boiling range 40–60 °C.

Transmission electron microscopy (TEM)

TEM images were recorded on a Morgagni 268 Transmission Electron Microscope operating at 80 kV. The samples were prepared using the uranyl acetate negative staining method. In a carbon coated copper grid (400 mesh), 4 μL of the hydrogel was placed and was left for 30 seconds. The excess of water from the hydrogel was removed using a filter paper, and washed with water and a solution of uranyl acetate (2%) (3 times each).

Rheometry

Rheological studies were performed on an ARES-G2 rheometer with a parallel plate at 25 °C. Dynamic strain sweep and frequency sweep experiments were carried out. During the strain sweep experiments the hydrogels were under different oscillation strains, constant frequency (6.28 rad s^{-1}) and constant temperature (25.5 °C). In a frequency sweep, the experiments were carried out under different frequencies (0.1 to 200 rad s^{-1}), constant oscillation amplitude and temperature (25.5 °C).

CD spectroscopy

The CD spectra were obtained using an OLIS DSM-20 CD spectropolarimeter operating in the UV-Visible spectral region, equipped with a Peltier temperature control unit. The near UV spectra (500–260 nm) were obtained with 1 second accumulations every 1 nm. The far UV spectra (260–190 nm) were obtained with 5 seconds accumulations every 1 nm. Optical cells with path lengths ranging from 0.05 to 1.00 mm were used. Baselines with the buffer used in each hydrogel were obtained at 20 °C and 80 °C. As no relevant differences in the spectra were observed with the variation of temperature, the hydrogel spectra were corrected with the baselines at 20 °C. Correction in relation to the path length of the optical cell used was also made. The data were smoothed mathematically in Origin 8 software. The optical cells were filled with each hydrogel pre-heated at 80 °C to form a clear solution, and then introduced into the CD spectropolarimeter with the temperature being previously programmed to 80 °C. The spectra were obtained 10–15 minutes after each change in temperature.

Photophysical studies

Fluorescence measurements were performed using a Fluorolog 3 spectrofluorimeter, equipped with double monochromators in both excitation and emission, Glan–Thompson polarizers and a temperature controlled cuvette holder. Fluorescence emission and excitation spectra were corrected for the instrumental response of the system.

Molecular dynamic simulations

The molecular structure of the peptides **5a–e** with unnatural α,β-dehydroamino acids under study (Scheme 1) was designed with the program Pymol.[25] These molecules were parameterized using parameters transferred from the natural amino acids in the GROMOS 54a7[36,37] force field. To validate the proposed parameters, the new amino acids were subjected to 12 000 steps of energy minimization calculations with the steepest descent algorithm and 100 ps MD simulation in a cubic box solvated with Simple Point Charge (SPC) water model.[23] The validation was done by analyzing the convergence of the system's potential energy and the geometry of the amino acids. The naproxen group present in all five peptides under study, **5a–e**, which were also parameterized according to the GROMOS54a7 force field and the molecule was subjected to the same protocol used to validate the α,β-dehydroamino acids. The topologies of Npx, ΔPhe and ΔAbu are available upon request. The five peptides **5a–e** were designed and eleven copies were placed in a cubic box of size 4.5 × 4.5 × 4.5 nm solvated with an SPC water model.[23] Each system was energy minimized with the steepest descent algorithm and 60 ns of MD simulations were run, the first 40 ns were spent for equilibration and the last 20 ns were used for analysis. In these experiments the simulation was made in 30 000 000 steps with an integration interval of 2 fs. All simulations were run with the GROMACS 4.5.4 software package.[38,39] In all MD simulations the system was maintained at a constant temperature and a pressure of 310 K and 1 atm, respectively, using the Berendsen thermostat and barostat methods,[40] with τ_T = 0.20 ps and τ_P = 0.10 ps. The SETTLE algorithm[41] was used to constrain bond lengths and angles of water molecules, while the bond lengths and angles of peptides were constrained with the LINCS algorithm.[42] For the treatment of long-range interactions, we used the reaction field method, with a cut-off of 1.4 nm and a dielectric constant of ε = 54 for water. The van der Waals interactions were also calculated with a cut-off radius of 1.4 nm. The aggregation properties of each peptide system were evaluated by identifying the occurrence of peptide clusters formed in the simulation box. Peptide clusters were detected by clustering peptides using the single-linkage method at a cut-off of 1.4 nm between the center of mass of each peptide. The number of clusters for each system was counted and characterized. The average number of intra and intermolecular hydrogen bonds, and also the average number of intra and intermolecular π-stacking interactions were calculated in order to understand the interactions responsible for the formation of aggregates.

6362 | *J. Mater. Chem. B*, 2015, **3**, 6355–6367

This journal is © The Royal Society of Chemistry 2015

140

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo,
R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B*, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Enzymatic resistance assay

To illustrate the enzymatic resistance of dehydrodipeptides, diluted solutions of compounds **5a** and **5b** and of Npx-L-Phe-L-Phe-OH (0.5 mg mL^{-1}) were prepared in sodium phosphate buffer pH 7.47 0.1 M and divided into three samples of 100 μL. A solution of α-chymotrypsin in the same buffer was also prepared (1.0 mg mL^{-1}; 51.33 U mL^{-1}). All the solutions were incubated at 37 °C and 20 rpm overnight. The enzyme solution (100 μL) was added to each hydrogelator solution. Samples of 10 μL were taken at 0 h, 2 h, 4 h, 8 h, 12 h, 24 h, 49 h and 78 h. and analyzed by HPLC (λ = 276 nm; water/acetonitrile, 1:1 with 0.1% TFA). The percentage of the gelator was determined using the peptide peak area in each sample and comparing it with the area of the same peak in the diluted solutions without the enzyme. To verify whether these solutions were stable at 37 °C, the samples of each peptide were analyzed by HPLC after 78 hours at 37 °C and 20 rpm.

MTT assay

Adult human skin fibroblasts (ASF-2 cells) were maintained at 37 °C in a humidified 5% CO$_2$ atmosphere grown in Dulbecco's Modified Eagle's Medium (DMEM, Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS, Lonza, Verviers, Belgium), 10 mM Hepes and 1% antibiotic/antimycotic solution (Sigma-Aldrich, St. Louis, MO, USA). Prior to culture, cells within a sub-confluent monolayer were trypsinized using trypsin (0.05%)–EDTA·Na (0.53 mM) solution and resuspended in DMEM to obtain a cell concentration of around 50 000 cells per mL. The cells were plated in 96-multiwell culture plates (100 μL per well) 24 hours before incubation with compound **5d**. Cells were then treated with different concentrations of **5d**, and prepared as follows: Npx-L-Ala-Z-ΔPhe-OH (**5d**) was dissolved in phosphate buffer of 0.1 M, pH 8, obtaining a solution of 5.0 mM. The 5 mM solution was used to prepare solutions of 50 μM, 100 μM and 500 μM in DMEM. Solutions of phosphate buffer 0.1 M pH 8 at 1%, 2% and 10% in DMEM were prepared as controls. 100 μL aliquots of buffer controls and **5d** solutions were placed into the wells of the plate with the cell culture, with three replicas of each. The plate was incubated at 37 °C for 48 hours. Cells were then incubated for 60 minutes with MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide, Sigma-Aldrich, St. Louis, MO, USA] to a final concentration of 0.5 mg mL^{-1}. Then, the medium was removed, and the formazan crystals formed by the cell's capacity to reduce MTT were dissolved in a 50:50 (v/v) DMSO: ethanol solution, and the absorbance was measured at 570 nm (with background subtraction at 690 nm), in a SpectroMax Plus384 absorbance microplate reader. The results were expressed as percentage relative to the control (cells with buffer solution).

Synthesis of β-hydroxydipeptides derivatives (1a–e). The synthesis of compound **1a**,[43] **1b**[43] and **1e**[16] was described elsewhere.

Synthesis of Boc-L-Val-D,L-Phe(β-OH)-OMe (1c). Boc-L-Val-OH (4.34 g, 20 mmol) was treated with H-D,L-Phe(β-OH)-OMe,HCl

(4.63 g, 20 mmol) in acetonitrile using the standard N,N'-dicyclohexylcarbodiimide (DCC)/1-hydroxybenzotriazole (HOBt) procedure, giving **1c** (6.89 g, 87%) as an oil; ^{1}H NMR (300 MHz, CDCl$_3$, δ): 0.60–0.67 (dd, J = 6.3 and 7.2 Hz, 6H, γCH$_3$ Val), 0.79–0.89 (dd, J = 6.9 and 9.2 Hz, 6H, γCH$_3$ Val), 1.40 (s, 18H, CH$_3$ Boc), 1.83–2.00 (m, 2H, βCH Val), 3.72 (s, 3H, OCH$_3$), 3.73 (s, 3H, OCH$_3$), 3.77–3.97 (m, 2H, αCH Val), 4.86–4.92 [m, 2H, αCH Phe(β-OH)], 5.01–5.14 (dd, J = 9.0 and 21 Hz, 2H, NH Val), 5.30–5.38 [dd, J = 4.0 and 28.4 Hz, 2H, βCH Phe(β-OH)], 7.09–7.16 [m, 2H, NH Phe(β-OH)], 7.20–7.39 (m, 10H, Ar H); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 17.14 (γCH$_3$ Val), 17.74 (γCH$_3$ Val), 18.99 (γCH$_3$ Val), 28.20 (CH$_3$ Boc), 30.86 (βCH Val), 52.47 (OCH$_3$), 52.58 (OCH$_3$), 58.04 [αCH Phe(β-OH)], 58.22 [αCH Phe(β-OH)], 59.48 (αCH Val), 59.78 (αCH Val), 73.00 [βCH Phe(β-OH)], 73.47 [βCH Phe(β-OH)], 79.91 [(CH$_3$)$_3$C], 125.64 (CH), 125.86 (CH), 127.75 (CH), 127.81 (CH), 128.25 (CH), 128.30 (CH), 139.73 (C), 139.78 (C), 155.89 (C=O), 155.94 (C=O), 170.73 (C=O), 170.99 (C=O), 171.90 (C=O), 172.01 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{20}$H$_{30}$N$_2$NaO$_6$ 417.19961; found, 417.19957.

Synthesis of Boc-L-Ala-D,L-Phe(β-OH)-OMe (1d). Boc-L-Ala-OH (1.89 g, 10 mmol) was treated with H-D,L-Phe(β-OH)-OMe,HCl (2.32 g, 10 mmol) in acetonitrile using the standard DCC/HOBt procedure, giving **1d** (3.50 g, 95%) as an oil; ^{1}H NMR (300 MHz, CDCl$_3$, δ): 1.10–1.25 (m, 6H, βCH$_3$ Ala), 1.40 (2s, 9H, CH$_3$ Boc), 1.42 (2s, 9H, CH$_3$ Boc), 3.35 (brs, 2H, OH), 3.71 (2s, 3H, OCH$_3$), 3.73 (2s, 3H, OCH$_3$), 4.14 (brs, 2H, αCH Ala), 4.84–4.88 [dd, J = 3.3 and 7.2 Hz, 2H, αCH Phe(β-OH)], 5.08 (brs, 1H, NH Ala), 5.18 (brs, 1H, NH Ala), 5.26–5.30 [dd, J = 3.0 and 6.6 Hz, 2H, βCH Phe(β-OH)], 7.12 [brd, J = 8.7 Hz, 2H, NH Phe(β-OH)], 7.24–7.37 (m, 10H, Ar H); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 18.20 (βCH$_3$ Ala), 18.53 (βCH$_3$ Ala), 28.21 (CH$_3$ Boc), 49.84 (αCH Ala), 52.52 (OCH$_3$), 52.59 (OCH$_3$), 58.09 [αCH Phe(β-OH)], 73.33 [βCH Phe(β-OH)], 73.55 [βCH Phe(β-OH)], 80.04 [(CH$_3$)$_3$C], 125.77 (CH), 125.90 (CH), 127.90 (CH), 128.24 (CH), 128.26 (CH), 139.67 (C), 155.39 (C=O), 155.42 (C=O), 170.77 (C=O), 170.96 (C=O), 172.93 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{18}$H$_{26}$N$_2$NaO$_6$ 389.16831; found, 389.16827.

Synthesis of dehydrodipeptides derivatives (2a–e). DMAP (0.1 equiv.) was added to solutions of compounds **1a–e** in dry acetonitrile (1 M) followed by Boc$_2$O (1.0 equiv.) under rapid stirring at room temperature. The reaction was monitored by ^{1}H NMR until all the reactant had been consumed. Then TMG (2% in volume) was added, stirring was continued and the reaction was followed by ^{1}H NMR. When all the reactant had been consumed, evaporation at reduced pressure gave a residue that was partitioned between ethyl acetate (50 mL) and KHSO$_4$ (30 mL, 1 M). The organic phase was thoroughly washed with KHSO$_4$ (1 M) and brine (2 × 30 mL, each), and dried with MgSO$_4$. The removal of the solvent afforded compounds **2a–e**.

The synthesis of compounds **2a**,[43] **2b**[44] and **2e**[45] was described elsewhere.

Synthesis of Boc-L-Val-Z-ΔPhe-OMe (2c). Compound **1c** (1.97 g, 5 mmol) was treated according to the procedure described above to give compound **2c** (1.70, 90%) as a white solid; mp 152.0–153.0 °C; ^{1}H NMR (300 MHz, CDCl$_3$, δ): 1.00 (dd, J = 6.6 Hz, 6H, γCH$_3$ Val), 1.44 (s, 9H, CH$_3$ Boc), 2.18–2.28

141

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B*, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

(m, 1H, βCH Val), 3.81 (s, 3H, OCH$_3$), 4.01–4.16 (m, 1H, αCH Val), 5.14 (d, J = 9.0 Hz, 1H, NH Val), 7.28–7.35 (m, 3H, Ar H), 7.34 (s, 1H, βCH ΔPhe), 7.48 (d, J = 6.9 Hz, 2H, Ar H), 7.81 (s, 1H, NH ΔPhe); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 17.56 (γCH$_3$ Val), 19.26 (γCH$_3$ Val), 28.25 (CH$_3$ Boc), 30.49 (βCH Val), 52.50 (OCH$_3$), 59.99 (αCH Val), 80.04 [(CH$_3$)$_3$$C$], 123.96 (C), 128.53 (CH), 129.43 (C), 129.70 (CH), 132.66 (CH), 133.44 (βCH ΔPhe), 155.96 (C=O), 165.40 (C=O), 170.72 (C=O); anal. calcd for C$_{20}$H$_{28}$N$_2$O$_5$: C 63.81, H 7.50, N 7.44; found: C 63.36, H 7.36, N 7.40.

Synthesis of Boc-ʟ-Ala-Z-ΔPhe-OMe (2d). Compound **1d** (3.50 g, 9.5 mmol) was treated according to the procedure described above to give compound **2d** (3.05 g, 92%,) as a white solid; mp 107.0–108.0 °C; ^{1}H NMR (300 MHz, CDCl$_3$, δ): 1.38–1.44 (m, 12H, βCH$_3$ Ala and CH$_3$ Boc), 3.79 (s, 3H, OCH$_3$), 4.36 (brs, 1H, αCH Ala), 5.27 (brs, 1H, NH Ala), 7.28–7.36 (m, 3H, Ar H), 7.40 (s, 1H, βCH ΔPhe), 7.47 (d, J = 6.6 Hz, 2H, Ar H), 8.02 (s, 1H, NH ΔPhe); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 17.87 (βCH$_3$ Ala), 28.22 (CH$_3$ Boc), 50.22 (αCH Ala), 52.55 (OCH$_3$), 80.19 [(CH$_3$)$_3$$C$], 123.79 (αC), 128.46 (CH), 129.44 (CH), 129.71 (CH), 133.20 (βCH ΔPhe), 133.46 (C), 155.58 (C=O), 165.46 (C=O), 171.54 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{18}$H$_{24}$N$_2$NaO$_5$ 371.15774; found, 371.15774.

Synthesis of dehydrodipeptides 3a–e. TFA (0.3 M) was added to compounds **2a–e**. After 2 hours the mixture was taken to dryness at reduced pressure to afford the corresponding dehydrodipeptide methyl ester.

Synthesis of H-ʟ-Phe-Z-ΔPhe-OMe,TFA (3a). The general procedure described above was followed using compound **2a** (0.86 g, 1.95 mmol) giving compound **3a** (0.79 g, 92%) as a white solid; mp 87.0–88.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 2.94–3.00 (dd, J = 9.2 and 5.2 Hz, 1H, βCH$_2$ Phe), 3.24–3.29 (dd, J = 4.8 and 9.2 Hz, 1H, βCH$_2$ Phe), 3.73 (s, 3H, OCH$_3$), 4.25 (brs, 1H, αCH Phe), 7.30–7.41 (m, 9H, Ar H and βCH ΔPhe), 7.58–7.60 (dd, J = 2.0 and 4.0 Hz, 2H, Ar H), 8.26 (brs, 3H, NH$_3$$^+$), 10.37 (s, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 36.69 (βCH$_2$ Phe), 52.37 (OCH$_3$), 53.61 (αCH Phe), 125.03 (αC), 127.31 (CH), 128.64 (CH), 128.71 (CH), 129.57 (CH), 129.72 (CH), 130.01 (CH), 132.31 (βCH ΔPhe), 132.83 (C), 134.80 (C), 164.87 (C=O), 168.30 (C=O); HRMS (micrOTOF) m/z: [M]$^+$ calcd for C$_{19}$H$_{21}$N$_2$O$_3$$^+$ 325.15467; found, 325.15545.

Synthesis of H-ʟ-Phe-Z-ΔAbu-OMe,TFA (3b). The general procedure described above was followed with compound **2b** (1.74 g, 4.8 mmol) giving compound **3b** (1.60 g, 89%) as an oil; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.60 (d, J = 7.2 Hz, 3H, γCH$_3$ ΔAbu), 2.99–3.05 (dd, J = 8.0 and 6.0 Hz, 1H, βCH$_2$), 3.14–3.19 (dd, J = 6.0 and 8.0 Hz, 1H, βCH$_2$ Phe), 3.66 (s, 3H, OCH$_3$), 4.18–4.20 (m, 1H, αCH Phe), 6.60 (q, J = 7.2 Hz, 1H, βCH ΔAbu), 7.27–7.34 (m, 5H, Ar H), 8.29 (brs, 3H, NH$_3$$^+$), 9.89 (s, 1H, NH ΔAbu); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 13.45 (γCH$_3$ ΔAbu), 37.04 (βCH$_2$ Phe), 52.00 (OCH$_3$), 53.41 (αCH Phe), 126.71 (αC), 127.22 (CH), 128.57 (CH), 129.57 (CH), 133.68 (βCH ΔAbu), 134.74 (C), 164.11 (C=O), 167.22 (C=O); HRMS (micrOTOF) m/z: [M]$^+$ calcd for C$_{14}$H$_{19}$N$_2$O$_3$$^+$ 263.13902; found, 263.13925.

Synthesis of H-ʟ-Val-Z-ΔPhe-OMe,TFA (3c). The general procedure described above was followed with compound **2c**

(0.75 g, 2.0 mmol) giving compound **3c** (0.70 g, 90%) as a white solid; mp 228.5–230.0 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 0.98 (d, J = 6.9 Hz, 3H, γCH$_3$ Val), 1.04 (d, J = 6.9 Hz, 3H, γCH$_3$ Val), 2.19–2.30 (m, 1H, βCH Val), 3.71 (s, 3H, OCH$_3$), 3.82 (brs, 1H, αCH Val), 7.31 (s, 1H, βCH ΔPhe), 7.36–7.48 (m, 3H, Ar H), 7.65–7.69 (m, 2H, Ar H), 8.27 (brs, 3H, NH$_3$$^+$ Val), 10.25 (brs, 1H, NH ΔPhe); ^{13}C NMR (75.4 MHz, DMSO-d_6, δ): 16.98 (γCH$_3$ Val), 18.31 (γCH$_3$ Val), 29.91 (βCH Val), 52.26 (OCH$_3$), 57.35 (αCH Val), 125.30 (αC), 128.68 (CH), 129.74 (CH), 130.17 (CH), 132.43 (βCH ΔPhe), 132.87 (C), 164.88 (C=O), 168.32 (C=O); anal. calcd for C$_{17}$H$_{21}$N$_2$O$_5$F$_3$: C 52.31, H 5.42, N 7.18; found: C 51.83, H 5.47, N 7.19.

Synthesis of H-ʟ-Ala-Z-ΔPhe-OMe,TFA (3d). The general procedure described above was followed with compound **2d** (1.39 g, 4.0 mmol) giving compound **3d** (1.33 g, 95%) as an oil; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 1.46 (d, J = 6.9 Hz, 3H, βCH$_3$ Ala), 3.72 (s, 3H, OCH$_3$), 4.08 (brt, J = 5.7 Hz, 1H, αCH Ala), 7.37–7.42 (m, 4H, Ar H and βCH ΔPhe), 7.63–7.67 (m, 2H, Ar H), 8.28 (brs, 3H, NH$_3$$^+$), 10.17 (s, 1H, NH ΔPhe); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 16.77 (βCH$_3$ Ala), 48.44 (αCH Ala), 52.44 (OCH$_3$), 125.16 (αC), 128.76 (CH), 129.88 (CH), 130.12 (CH), 133.03 (C), 133.29 (βCH ΔPhe), 164.99 (C=O), 169.70 (C=O).

Synthesis of H-ʟ-Ala-Z-ΔAbu-OMe,TFA (3e). The general procedure described above was followed with compound **2e** (0.69 g, 2.4 mmol) giving compound **3e** (0.60 g, 86%) as an oil; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 1.42 (d, J = 6.9 Hz, 3H, βCH$_3$ Ala), 1.69 (d, J = 7.5 Hz, 3H, γCH$_3$ ΔAbu), 3.66 (s, 3H, OCH$_3$), 3.99 (brs, 1H, αCH Ala), 6.66 (q, J = 6.9 Hz, 1H, βCH ΔAbu), 8.19 (brs, 3H, NH$_3$$^+$), 9.74 (s, 1H, NH ΔAbu); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 13.49 (γCH$_3$ ΔAbu), 17.18 (βCH$_3$ Ala), 48.25 (αCH Ala), 52.07 (OCH$_3$), 125.83 (αC), 134.03 (βCH ΔAbu), 164.13 (C=O), 168.81 (C=O).

Synthesis of dehydrodipeptides 4a–e. Triethylamine (2.2 equiv.) was added to a solution of dehydrodipeptide methyl ester hydrochloride (**3a–e**) in dichloromethane (0.1 M), and (S)-(+)-naproxen chloride (1 equiv.) was then slowly added with vigorous stirring and cooling in an ice bath. After stirring at 0 °C for 30 minutes the solution was stirred at room temperature overnight. The reaction mixture was then concentrated and partitioned between ethyl acetate (100 mL) and KHSO$_4$ (1 M, 50 mL) and washed with KHSO$_4$ (1 M), NaHCO$_3$ (1 M) and brine (3 × 30 mL). After drying over MgSO$_4$ the extract was taken to dryness at reduced pressure to afford the corresponding N-protected dehydrodipeptide methyl ester (**4a–e**).

Synthesis of compound Npx-ʟ-Phe-Z-ΔPhe-OMe (4a). The general procedure described above was followed with compound **3a** (0.40 g, 0.91 mmol) giving compound **4a** (0.35 g, 71%) as a white solid; mp 176.0–177.0 °C; ^{1}H NMR (400 MHz, CDCl$_3$, δ): 1.52 (d, J = 7.2 Hz, 3H, CH$_3$), 2.95–3.01 (dd, J = 7.6 Hz, 1H, βCH$_2$ Phe), 3.08–3.12 (dd, J = 6.4 Hz, 1H, βCH$_2$ Phe), 3.64 (q, J = 7.2 Hz, 1H, CH), 3.72 (s, 3H, OCH$_3$), 3.93 (s, 3H, OCH$_3$), 4.82 (q, J = 7.6 Hz, 1H, αCH Phe), 5.96 (d, J = 8.0 Hz, 1H, NH Phe), 7.04 (d, J = 7.8 Hz, 2H, Ar H), 7.09 (d, J = 2.4 Hz, 2H, Ar H), 7.11–7.23 (m, 7H, Ar H), 7.30–7.32 (m, 3H, Ar H), 7.51 (s, 1H, βCH), 7.61 (dd, J = 8.4 and 8.8 Hz, 2H, Ar H), 7.77 (brs, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, CDCl$_3$, δ): 18.10 (CH$_3$), 36.80

142

H. Vilaça, a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo,
R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B*, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

(βCH$_2$ Phe), 46.85 (αCH Phe), 52.50 (OCH$_3$), 54.41 (CH), 55.28 (OCH$_3$), 105.59 (CH), 119.15 (CH), 123.72 (CH), 125.94 (CH), 126.17 (CH), 126.90 (CH), 127.17 (CH), 127.69 (CH), 128.58 (CH), 128.93 (C), 129.21 (CH), 129.27 (C), 129.42 (CH), 129.66 (CH), 132.70 (CH), 133.29 (C), 133.81 (C), 135.21 (C), 136.05 (C), 157.80 (C-O), 165.14 (C=O), 169.70 (C=O), 175.07 (C=O); anal. calcd for C$_{33}$H$_{32}$N$_2$O$_5$: C 73.86, H 6.01, N 5.22; found: C 73.45, H 6.023, N 4.924.

Synthesis of Npx-L-Phe-Z-ΔAbu-OMe (4b). The general procedure described above was followed with compound **3b** (0.438 g, 1 mmol) giving compound **4b** (0.433 g, 91%) as a white solid; mp 164.0–165.0 °C; ^{1}H NMR (400 MHz, CDCl$_3$, δ): 1.52–1.56 (m, 6H, CH$_3$ and γCH$_3$ ΔAbu), 2.96–3.01 (dd, J = 7.6 Hz, 1H, βCH$_2$ Phe), 3.10–3.15 (dd, J = 6.4 Hz, 1H, βCH$_2$ Phe), 3.64–3.70 (m, 4H, CH and OCH$_3$), 3.92 (s, 3H, OCH$_3$), 4.83 (q, J = 7.2 Hz, 1H, αCH Phe), 6.18 (d, J = 6.8 Hz, 1H, NH Phe), 6.67 (q, J = 7.2 Hz, 1H, βCH ΔAbu), 7.04–7.17 (m, 8H, H Ar and NH ΔAbu), 7.25 (dd, J = 2.0 Hz, 1H, Ar H), 7.56 (s, 1H, Ar H), 7.64 (dd, J = 8.4 and 8.8 Hz, 2H, Ar H); ^{13}C NMR (100.6 MHz, CDCl$_3$, δ): 14.15 (γCH$_3$ ΔAbu), 18.08 (CH$_3$), 37.19 (βCH$_2$ Phe), 46.76 (CH), 52.11 (OCH$_3$), 54.27 (αCH Phe), 55.27 (OCH$_3$), 105.53 (CH), 119.06 (CH), 125.85 (C), 126.01 (CH), 126.06 (CH), 126.81 (CH), 127.52 (CH), 128.47 (CH), 128.90 (C), 129.17 (CH), 129.24 (CH), 133.75 (C), 134.28 (CH), 135.32 (C), 136.18 (C), 157.72 (C-O), 164.39 (C=O), 169.36 (C=O), 174.91 (C=O); HRMS (micrOTOF) m/z: [M + Na]$^+$ calcd for C$_{28}$H$_{30}$N$_2$NaO$_5$ 497.20524; found, 497.20604.

Synthesis of Npx-L-Val-Z-ΔPhe-OMe (4c). The general procedure described above was followed with compound **3c** (0.59 g, 1.5 mmol) giving compound **4c** (0.50 g, 68%) as a white solid; mp 213.0–214.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 0.91 (d, J = 6.8 Hz, 3H, γCH$_3$ Val), 0.96 (d, J = 6.4 Hz, 3H, γCH$_3$ Val), 1.43 (d, J = 7.2 Hz, 3H, CH$_3$), 1.97–2.07 (m, 1H, βCH Val), 3.60 (s, 3H, OCH$_3$), 3.83 (s, 3H, OCH$_3$), 3.96 (q, J = 6.8 Hz, 1H, CH), 4.82 (t, J = 8.0 Hz, 1H, αCH Val), 7.04–7.11 (m, 4H, Ar H), 7.18 (s, 1H, βCH ΔPhe), 7.24 (d, J = 2.4 Hz, 1H, Ar H), 7.47 (dd, J = 2.0 and 8.8 Hz, 1H, Ar H), 7.51 (dd, J = 2.0 and 8.4 Hz, 2H, Ar H), 7.66 (d, J = 4.8 Hz, 1H, Ar H), 7.69 (d, J = 4.4 Hz, 1H, Ar H), 7.72 (s, 1H, Ar H), 8.14 (d, J = 8.8 Hz, 1H, NH Val), 9.73 (brs, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 18.34 (γCH$_3$ Val), 19.17 (γCH$_3$ Val), 19.34 (CH$_3$), 30.64 (βCH Val), 40.13 (αCH Val), 52.05 (OCH$_3$), 55.19 (OCH$_3$), 57.70 (CH), 105.72 (CH), 118.49 (CH), 125.42 (CH), 125.85 (C), 126.50 (CH), 126.73 (CH), 128.41 (C), 128.45 (CH), 129.10 (CH), 129.27 (CH), 129.98 (CH), 132.37 (CH), 133.08 (C), 133.17 (C), 137.38 (C), 156.99 (C-O), 165.32 (C=O), 171.37 (C=O), 173.76 (C=O); HRMS (micrOTOF) m/z: [M + Na]$^+$ calcd for C$_{29}$H$_{32}$N$_2$NaO$_5$ 511.22089; found, 511.22136.

Synthesis of Npx-L-Ala-Z-ΔPhe-OMe (4d). The general procedure described above was followed with compound **3d** (0.70 g, 2 mmol) giving compound **4d** (0.62 g, 67%) as a white solid; mp 169.0–170.0 °C; ^{1}H NMR (400 MHz, CDCl$_3$, δ): 1.37 (d, J = 7.2 Hz, 3H, βCH$_3$ Ala), 1.55 (d, J = 7.2 Hz, 3H, CH$_3$), 3.70 (s, 4H, CH and OCH$_3$), 3.90 (s, 3H, OCH$_3$), 4.68–4.76 (m, 1H, αCH Ala), 6.29 (d, J = 6.4 Hz, 1H, NH Ala), 7.05 (d, J = 2.4 Hz, 1H, Ar H), 7.09–7.12 (dd, J = 2.4 and 6.4 Hz, 1H, Ar H), 7.22–7.24 (m, 3H, Ar H), 7.26–7.29 (dd, J = 1.6 and 6.8 Hz, 1H, Ar H), 7.33 (s, 1H, βCH ΔPhe), 7.38–7.40 (m, 2H, Ar H), 7.60 (s, 2H, Ar H), 7.62 (s, 1H, Ar

H), 8.02 (brs, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, CDCl$_3$, δ): 17.58 (βCH$_3$ Ala), 18.37 (CH$_3$), 46.73 (CH), 49.11 (αCH Ala), 52.48 (OCH$_3$), 55.27 (OCH$_3$), 105.57 (CH), 119.05 (CH), 123.66 (αC), 125.94 (CH), 126.09 (CH), 127.57 (CH), 128.46 (CH), 128.91 (C), 129.21 (CH), 129.47 (CH), 129.69 (CH), 133.73 (βCH ΔPhe), 133.33 (C), 133.73 (C), 135.64 (C), 157.68 (C-O), 165.21 (C=O), 171.00 (C=O), 174.76 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{27}$H$_{28}$N$_2$O$_5$ 483.18904; found, 483.18917.

Synthesis of Npx-L-Ala-Z-ΔAbu-OMe (4e). The general procedure described above was followed with compound **3e** (0.69 g, 2.4 mmol) giving compound **4e** (0.89 g, 92%) as a white solid; mp 150.0–151.0 °C; ^{1}H NMR (300 MHz, CDCl$_3$, δ): 1.38 (d, J = 7.2 Hz, 3H, βCH$_3$ Ala), 1.50 (d, J = 7.5 Hz, 3H, γCH$_3$ ΔAbu), 1.58 (d, J = 7.2 Hz, 3H, CH$_3$), 3.63 (s, 3H, OCH$_3$), 3.73 (q, J = 7.2 Hz, 1H, CH), 3.89 (s, 3H, OCH$_3$), 4.66–4.76 (m, 1H, αCH Ala), 6.49 (d, J = 7.5 Hz, 1H, NH Ala), 6.66 (q, J = 7.2 Hz, 1H, βCH ΔAbu), 7.06 (d, J = 2.4 Hz, 1H, Ar H), 7.09–7.13 (dd, J = 2.7 and 6.3 Hz, 1H, Ar H), 7.31–7.35 (dd, J = 1.5 and 6.9 Hz, 1H, Ar H), 7.62–7.66 (m, 3H, Ar H), 7.91 (brs, 1H, NH ΔAbu); ^{13}C NMR (75.4 MHz, CDCl$_3$, δ): 13.92 (γCH$_3$ ΔAbu), 17.86 (βCH$_3$ Ala), 18.27 (CH$_3$), 46.61 (CH), 48.86 (αCH Ala), 52.08 (OCH$_3$), 55.22 (OCH$_3$), 105.48 (CH), 118.97 (CH), 125.98 (CH and αC), 126.44 (C), 127.40 (CH), 128.85 (CH), 129.17 (CH), 133.67 (C), 134.50 (βCH ΔAbu), 135.66 (C), 157.60 (C-O), 164.45 (C=O), 170.83 (C=O), 174.68 (C=O); anal. calcd for C$_{22}$H$_{26}$N$_2$O$_5$: C 66.32, H 6.58, N 7.03; found: C 66.23, H 6.25, N 6.51.

Synthesis of dehydrodipeptides 5a–e. NaOH (1 equiv., 1 M) was added to a solution of N-acyl dehydrodipeptide methyl ester (**4a–e**) in dioxane (3 mL). The solution was stirred at room temperature (the reaction was followed by TLC until no starting material was detected). The dioxane was removed under reduced pressure and the reaction mixture was acidified to pH 2–3 with HCl (1 M) and the solid formed was filtered.

Synthesis of Npx-L-Phe-Z-ΔPhe-OH (5a). The general procedure described above was followed with compound **4a** (0.268 g, 0.5 mmol) giving compound **5a** (0.183 g, 70%) as a white solid; mp 195.0–196.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.21 (d, J = 7.2 Hz, 3H, CH$_3$), 2.80–2.86 (dd, J = 10.8 and 13.6 Hz, 1H, βCH$_2$ Phe), 3.11–3.15 (dd, J = 3.6 and 14.0 Hz, 1H, βCH$_2$ Phe), 3.76 (q, J = 7.2 Hz, 1H, CH), 3.83 (s, 3H, OCH$_3$), 4.72–4.78 (m, 1H, αCH Phe), 7.09 (dd, J = 2.4 and 8.8 Hz, 1H, Ar H), 7.17–7.28 (m, 8H, Ar H), 7.31–7.33 (m, 2H, Ar H), 7.37 (dd, J = 1.6 and 8.4 Hz, 1H, Ar H), 7.51–7.74 (m, 2H, βCH + Ar H), 7.65–7.68 (m, 3H, βCH + Ar H), 8.31 (d, J = 8.8 Hz, 1H, NH Phe), 9.68 (s, 1H, NH ΔPhe), 12.71 (brs, 1H, CO$_2$H); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 18.86 (CH$_3$), 37.33 (βCH$_2$ Phe), 44.64 (CH), 53.80 (αCH Phe), 55.10 (OCH$_3$), 105.62 (CH), 118.37 (CH), 125.34 (CH), 126.23 (CH), 126.35 (C), 126.39 (CH), 126.65 (CH), 127.99 (CH), 128.29 (C), 128.36 (CH), 128.45 (CH), 129.04 (CH), 129.32 (CH), 129.91 (CH), 131.83 (CH), 133.04 (C), 133.50 (C), 137.07 (C), 137.87 (C), 156.89 (C-O), 166.14 (C=O), 170.93 (C=O), 173.33 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{32}$H$_{30}$N$_2$NaO$_5$ 545.20469; found, 545.20483.

Synthesis of Npx-L-Phe-Z-ΔAbu-OH (5b). The general procedure described above was followed with compound **4b** (0.17 g, 0.36 mmol) giving compound **5b** (0.160 g, 97%) as a white solid;

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo,
R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B*, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

mp 186.0–187.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.21 (d, J = 7.2 Hz, 3H, CH$_3$), 1.47 (d, J = 7.2 Hz, 3H, γCH$_3$ ΔAbu), 2.80–2.86 (dd, J = 10.0 and 3.6 Hz, 1H, βCH$_2$ Phe), 3.06–3.11 (dd, J = 4.4 and 9.2 Hz, 1H, βCH$_2$ Phe), 3.77 (q, J = 6.8 Hz, 1H, CH), 3.84 (m, 3H, OCH$_3$), 4.68–4.74 (m, 1H, αCH Phe), 6.50 (q, J = 7.2 Hz, 1H, βCH ΔAbu), 7.11 (dd, J = 2.8 and 6.0 Hz, 1H, Ar H), 7.18–7.31 (m, 6H, Ar H), 7.56 (s, 1H, Ar H), 7.37 (dd, J = 1.6 and 6.8 Hz, 1H, Ar H) 7.64–7.73 (m, 2H, Ar H), 8.28 (d, J = 8.4 Hz, 1H, NH Phe), 9.19 (s, 1 H, NH ΔAbu), 12.48 (brs, 1H, CO$_2$H); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 13.52 (γCH$_3$ ΔAbu), 18.55 (CH$_3$), 37.90 (βCH$_2$ Phe), 44.57 (CH), 53.64 (αCH Phe), 55.10 (OCH$_3$), 105.62 (CH), 118.42 (CH), 125.29 (CH), 126.21 (CH), 126.60 (CH), 127.93 (CH), 127.98 (C), 128.30 (C), 129.04 (CH), 129.16 (CH), 129.32 (CH), 132.11 (CH), 133.05 (C), 137.02 (C), 137.79 (C), 156.91 (C-O), 165.37 (C=O), 169.88 (C=O), 173.17 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{27}$H$_{28}$N$_2$NaO$_5$ 483.18904; found, 483.18921.

Synthesis of Npx-L-Val-Z-ΔPhe-OH (5c). The general procedure described above was followed with compound **4c** (0.43 g, 0.88 mmol) giving compound **5c** (0.35 g, 85%) as a white solid; mp 218.0–219.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 0.90 (d, J = 6.4 Hz, 3H, γCH$_3$ Val), 0.96 (d, J = 6.4 Hz, 3H, γCH$_3$ Val), 1.43 (d, J = 7.2 Hz, 3H, CH$_3$), 2.00–2.09 (m, 1H, βCH Val), 3.84 (s, 3H, OCH$_3$), 3.96 (q, J = 7.2 Hz, 1H, CH), 4.36 (t, J = 8.0 Hz, 1H, αCH Val), 6.99–7.11 (m, 4H, Ar H), 7.20 (s, 1H, βCH ΔPhe), 7.24 (d, J = 2.8 Hz, 1H, Ar H), 7.45–7.52 (m, 3H, Ar H), 7.67–7.72 (dd, J = 2.4 and 6.4 Hz, 2H, Ar H), 7.72 (s, 1H, Ar H), 8.11 (d, J = 9.2 Hz, 1H, NH Val), 9.54 (s, 1H, NH ΔPhe), 12.67 (brs, 1H, CO$_2$H); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 18.27 (γCH$_3$ Val), 19.23 (γCH$_3$ Val), 19.32 (CH$_3$), 30.66 (βCH Val), 44.45 (CH), 55.11 (OCH$_3$), 57.62 (αCH Val), 105.64 (CH), 118.39 (CH), 125.35 (CH), 126.41 (CH), 126.67 (C), 126.69 (CH), 128.24 (CH), 128.35 (C), 128.82 (CH), 129.05 (CH), 129.76 (CH), 131.81 (CH), 133.08 (C), 133.45 (C), 137.34 (C), 156.90 (C-O), 166.13 (C=O), 170.91 (C=O), 173.59 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{28}$H$_{30}$N$_2$NaO$_5$ 497.20469; found, 497.20479.

Synthesis of Npx-L-Ala-Z-ΔPhe-OH (5d). The general procedure described above was followed with compound **4d** (0.46 g, 1 mmol) giving compound **5d** (0.38 g, 85%) as a white solid; mp 185.0–186.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.30 (d, J = 6.8 Hz, 3H, βCH$_3$ Ala), 1.41 (d, J = 6.4 Hz, 3H, CH$_3$), 3.83–3.86 (m, 4H, OCH$_3$ and CH), 4.44–4.52 (m, 1H, αCH Ala), 7.08–7.11 (dd, J = 2.4 and 6.4 Hz, 1H, Ar H), 7.19–7.23 (m, 5H, Ar H), 7.43–7.46 (dd, J = 1.6 and 6.8 Hz, 1H, Ar H), 7.52–7.55 (m, 2H, βCH ΔPhe and Ar H), 7.66–7.71 (m, 3H, Ar H), 8.23 (d, J = 7.6 Hz, 1H, NH Ala), 9.43 (s, 1H, NH ΔPhe), 12.64 (brs, 1H, CO$_2$H); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 17.95 (βCH$_3$ Ala), 18.85 (CH$_3$), 44.41 (CH), 48.11 (αCH Ala), 55.10 (OCH$_3$), 105.63 (CH), 118.38 (CH), 125.34 (CH), 126.28 (C), 126.43 (CH), 126.62 (CH), 128.32 (CH), 129.02 (C), 129.04 (CH), 129.05 (CH), 129.87 (CH), 131.87 (CH), 133.04 (C), 133.49 (C), 137.24 (C), 156.89 (C-O), 166.11 (C=O), 171.78 (C=O), 173.18 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{26}$H$_{26}$N$_2$NaO$_5$ 469.17339; found, 469.17353.

Synthesis of Npx-L-Ala-Z-ΔAbu-OH (5e). The general procedure described above was followed with compound **4e** (0.40 g, 1 mmol) giving compound **5e** (0.31 g, 80%) as a white solid; mp

170.0–171.0 °C; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.27 (d, J = 7.2 Hz, 3H, βCH$_3$ Ala), 1.39 (d, J = 7.2 Hz, 3H, CH$_3$), 1.47 (d, J = 7.2 Hz, 3H, γCH$_3$ ΔAbu), 3.84–3.86 (m, 4H, OCH$_3$ and CH), 4.40–4.47 (m, 1H, αCH Ala), 6.47 (q, J = 7.2 Hz, 1H, βCH ΔAbu), 7.09–7.13 (dd, J = 2.4 and 6.4 Hz, 1H, Ar H), 7.24 (d, J = 2.4 Hz, 1H, Ar H), 7.42–7.45 (dd, J = 2.0 and 6.8 Hz, 1H, Ar H), 7.68–7.74 (m, 3H, H Ar), 8.21 (d, J = 7.6 Hz, 1H, NH Ala), 8.94 (s, 1H, NH ΔAbu), 12.41 (brs, 1H, CO$_2$H); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 13.46 (γCH$_3$ ΔAbu), 18.45 (CH$_3$), 18.53 (CH$_3$), 44.39 (CH), 48.05 (αCH Ala), 55.10 (OCH$_3$), 105.64 (CH), 118.44 (CH), 125.29 (CH), 126.45 (CH), 126.59 (CH), 127.94 (αC), 128.33 (C), 129.04 (CH), 132.05 (CH), 133.06 (C), 137.18 (C), 156.91 (C-O), 165.33 (C=O), 170.88 (C=O), 173.09 (C=O); HRMS (ESI) m/z: [M + Na]$^+$ calcd for C$_{21}$H$_{24}$N$_2$NaO$_5$ 407.15774; found, 407.15778.

Conclusions

Several N-aromatic dehydrodipeptide amphiphiles were prepared and studied as new hydrogelators. Molecular dynamics simulations were used to obtain insights into the underlying molecular mechanism responsible for aggregation. The results obtained were in excellent agreement with the experimental observations. This allowed the rationalization of the structural features that govern the self-assembly of dehydrodipeptide amphiphiles. Thus, compounds with at least one aromatic amino acid gave hydrogels at low critical gelation concentrations. TEM images of the new hydrogels prepared revealed that they comprise of nanofibers with different widths that entangle to give a 3D network. All hydrogels showed a viscoelastic behaviour with a storage modulus higher than the loss modulus and independent of the frequency. The CD spectra of two hydrogelators, **5a** and **5b**, were compared with that obtained from the dipeptide phenylalanylphenylalanine N-protected with naproxen. The CD spectra were similar and point to a structural organization with the characteristics of a β-sheet arrangement. Fluorescence spectroscopy studies showed that this is a good methodology to determine the CGC and the gelation pH. Preliminary toxicity assays were performed using one of the hydrogelators and it was found that this compound was not toxic even at concentrations of 500 μM. The resistance of some of the new hydrogelators towards α-chymotrypsin was tested in an 80 h assay and it was found that the presence of the dehydroamino acid in the conjugates confers proteolytic resistance to the hydrogelator. Given the properties of this new class of hydrogelators it is possible to conclude that they constitute promising candidates for biomedical applications.

Acknowledgements

6366 | *J. Mater. Chem. B*, 2015, **3**, 6355–6367

This journal is © The Royal Society of Chemistry 2015

144

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B*, 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Notes and references

1 J. Kopecek and J. Yang, *Angew. Chem., Int. Ed.*, 2012, **51**, 7396.

2 S. Marchesan, Y. Qu, L. J. Waddington, C. D. Easton, V. Glattauer, T. J. Lithgow, K. M. McLean, J. S. Forsythe and P. G. Hartley, *Biomaterials*, 2013, **34**, 3678.

3 M. Guvendiren, H. D. Lu and J. A. Burdick, *Soft Matter*, 2012, **8**, 260.

4 G. Fichman and E. Gazit, *Acta Biomater.*, 2014, **10**, 1671.

5 A. Raspa, G. A. A. Saracino, R. Pugliese, D. Silva, D. Cigognini, A. Vescovi and F. Gelain, *Adv. Funct. Mater.*, 2014, **24**, 6317.

6 M. Ma, Y. Kuang, Y. Gao, Y. Zhang, P. Gao and B. Xu, *J. Am. Chem. Soc.*, 2010, **132**, 2719.

7 B. Frohm, J. E. DeNizio, D. S. M. Lee, L. Gentile, U. Olsson, J. Malm, K. S. Akerfeldt and S. Linse, *Soft Matter*, 2015, **11**, 414.

8 C. Tang, R. V. Ulijn and A. Saiani, *Langmuir*, 2011, **27**, 14438.

9 Z. Yang, G. Liang, M. Ma, Y. Gao and B. Xu, *Small*, 2007, **3**, 558.

10 G. Liang, Z. Yang, R. Zhang, L. Li, Y. Fan, Y. Kuang, Y. Gao, T. Wang, W. W. Lu and B. Xu, *Langmuir*, 2009, **25**, 8419.

11 X. Li, X. Du, J. Li, Y. Gao, Y. Pan, J. Shi, N. Zhou and B. Xu, *Langmuir*, 2012, **28**, 13512.

12 J. Li, Y. Gao, Y. Kuang, J. Shi, X. Du, J. Zhou, H. Wang, Z. Yang and B. Xu, *J. Am. Chem. Soc.*, 2013, **135**, 9907.

13 M. M. Nguyen, K. M. Eckes and L. J. Suggs, *Soft Matter*, 2014, **10**, 2693.

14 J. Li, Y. Kuang, Y. Gao, X. Du, J. Shi and B. Xu, *J. Am. Chem. Soc.*, 2013, **135**, 542.

15 J. Li, Y. Kuang, J. Shi, Y. Gao, J. Zhou and B. Xu, *Beilstein J. Org. Chem.*, 2013, **9**, 908.

16 P. M. T. Ferreira, H. L. S. Maia, L. S. Monteiro and J. Sacramento, *J. Chem. Soc., Perkin Trans. 1*, 1999, 3697.

17 P. M. T. Ferreira, L. S. Monteiro, G. Pereira, L. Ribeiro, J. Sacramento and L. Silva, *Eur. J. Org. Chem.*, 2007, 5934.

18 A. S. Abreu, P. M. T. Ferreira, L. S. Monteiro, M. J. R. P. Queiroz, I. C. F. R. Ferreira, R. C. Calhelha and L. M. Estevinho, *Tetrahedron*, 2004, **60**, 11821.

19 P. M. T. Ferreira, E. M. S. Castanheira, L. S. Monteiro, G. Pereira and H. Vilaça, *Tetrahedron*, 2010, **66**, 8672.

20 P. M. T. Ferreira, L. S. Monteiro, G. Pereira, E. M. S. Castanheira and C. G. Frost, *Eur. J. Org. Chem.*, 2013, 550.

21 G. Pereira, H. Vilaça and P. M. T. Ferreira, *Amino Acids*, 2013, **44**, 335.

22 (*a*) G. Laverty, A. P. McCloskey, B. F. Gilmore, D. S. Jones, J. Zhou and B. Xu, *Biomacromolecules*, 2014, **15**, 3429; (*b*) Y. Gao, Y. Kuang, X. Du, J. Zhou, P. Chandran, F. Horkay and B. Xu, *Langmuir*, 2013, **29**, 15191; (*c*) J. Raeburn, T. O. McDonald and D. J. Adams, *Chem. Commun.*, 2012, **48**, 9355; (*d*) X. Li, X. Du, J. Li, Y. Gao,

Y. Pan, J. Shi, N. Zhou and B. Xu, *Langmuir*, 2012, **28**, 13512; (*e*) L. Chen, G. Pont, K. Morris, G. Lotze, A. Squires, L. C. Serpell and D. J. Adams, *Chem. Commun.*, 2011, **47**, 12071; (*f*) Y. Zhang, Y. Kuang, Y. Gao and B. Xu, *Langmuir*, 2011, **27**, 529.

23 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269.

24 M. O. Sinnokrot, E. F. Valeev and C. D. Sherrill, *J. Am. Chem. Soc.*, 2002, **124**, 10887.

25 The PyMOL Molecular Graphics System, Version 1.3.1 Schrödinger, LLC.

26 M. M. Velazquez, M. Valero, L. J. Rodríguez, S. M. B. Costa and M. A. Santos, *J. Photochem. Photobiol., B*, 1995, **29**, 23.

27 G. D. Fasman, *Handbook of Biochemistry and Molecular Biology, Proteins, I*, CRC Press, 3rd edn, 1976.

28 L. Chen, K. Morris, A. Laybourn, D. Elias, M. R. Hicks, A. Rodger, L. Serpell and D. J. Adams, *Langmuir*, 2010, **26**, 5232.

29 N. Berova, L. D. Bari and G. Pescitelli, *Chem. Soc. Rev.*, 2007, **36**, 914.

30 G. Impellizzeri, F. D'Alessandro, G. Pappalardo and C. Tringali, *J. Inclusion Phenom.*, 2005, **51**, 173.

31 K. Nakanishi, N. Berova and R. W. Woody, *Circular Dichroism Principles and Applications*, VCH Publishers Inc, New York, 1994.

32 X. Li, K. Yi, J. Shi, Y. Gao, H.-C. Lin and B. Xu, *J. Am. Chem. Soc.*, 2011, **133**, 17513.

33 S. Debnath, A. Shome, D. Das and P. K. Das, *J. Phys. Chem. B*, 2010, **114**, 4407.

34 A. M. Smith, R. J. Williams, C. Tang, P. Coppo, R. F. Collins, M. L. Turner, A. Saiani and R. V. Ulijn, *Adv. Mater.*, 2008, **20**, 37.

35 H. F. Chow and J. Zhang, *Tetrahedron*, 2005, **61**, 11279.

36 W. Huang, Z. X. Lin and W. F. van Gunsteren, *J. Chem. Theory Comput.*, 2011, **7**, 1237.

37 N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark and W. F. van Gunsteren, *Eur. Biophys. J. Biophys. Lett.*, 2011, **40**, 843.

38 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435.

39 D. van der Spoel, E. Lindahl, B. Hess, A. R. Buuren, E. Apol, P. J. Meulenhoff, P. Tieleman, A. L. T. M. Sjibers, K. A. Feenstra, R. Drunen and H. J. C. Berendsen, *Gromacs user manual version*, 4.5, 2010, www.gromacs.org.

40 H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.

41 D. van der Spoel, P. J. van Maaren and H. J. C. Berendsen, *J. Chem. Phys.*, 1998, **108**, 10220.

42 B. Hess, H. Bekker, H. J. C. Berendsen and J. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463.

43 P. M. T. Ferreira, L. S. Monteiro and G. Pereira, *Amino Acids*, 2010, **39**, 499.

44 A. S. Abreu, E. M. S. Castanheira, P. M. T. Ferreira, L. S. Monteiro, G. Pereira and M. J. R. P. Queiroz, *Eur. J. Org. Chem.*, 2008, 5697.

45 R. Ramesh, K. De and S. Chandrasekaran, *Tetrahedron*, 2007, **63**, 10534.

145

H. Vilaça,a G. Pereira, **T. G. Castro**, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaêlo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins and P. M. T. Ferreira: *J. Mater. Chem. B,* 2015, 3, 6355 (DOI: 10.1039/c5tb00501a).

Chapter VII

Self-healing RGD dehydropeptide hydrogel

Theoretical part only.

FULL PAPER   **WILEY-VCH**

Self-healing RGD dehydropeptide hydrogel

Helena Vilaça,[a] Tarsila G. Castro,[a,f] Loly Torres Pérez,[b] Ashkan Dehsorkhi,[c] Cristóvão F. Lima,[d] Catarina Gonçalves,[e] Manuel Melle-Franco,[f] Loic Hilliou,[b] Miguel Gama,[e] Ian W. Hamley,[c] José A. Martins,[a] Paula M. T. Ferreira*[a]

Abstract: A new dehydropeptide supramolecular hydrogelator with the tripeptide cell adhesion motif arginine-glycine-aspartic acid (RGD) was designed, prepared and characterized. The dehydrodipeptide naproxen-alanyl-dehydrophenylalanine was conjugated with a pentapeptide GRGDG using conventional peptide synthesis protocols. This compound self-assembles in water to form nanofibres and produce a hydrogel at the concentration of 0.3 wt% and a pH of 6.0. The hydrogel has strong viscoelastic properties and presents self-healing and thermoreversible properties.

Introduction

Compounds with the right balance between hydrophobicity and hydrophilicity are able to trap the solvent in a three-dimensional (3D) network, resulting in a gel. If the solvent is water, they are known as hydrogelators. In physical gels non-covalent interactions like electrostatic, dipole-dipole, van der Waals, π-π stacking and hydrogen bonding are the drive for the gelator self-assembly into nanostructures.[1] Recently, the field of hydrogels has been focused on a new class of materials made from small molecules and known as low molecular weight gelators (LMWG). Small peptides with bulky aromatic motifs can form hydrogels with biomedical applications such as drug delivery, biossensing, tissue engineering and wound healing.[2] The major disadvantage of this type of materials is their susceptibility to enzymatic hydrolysis which diminishes their potential applications. One of the strategies used to circumvent this difficulty is to use non-proteinogenic amino acids. Recently, we reported hydrogelators resistant to proteolysis made of

dehydrodipeptides N-conjugated with naproxen.[3]

A major challenge in the biomaterials area is to create systems that can mimic the extracellular matrix (ECM).[4] Most of the hydrogels for 3D cell culture reported until now are synthetic or natural polymer networks.[4] However, several problems, such as biodegradability and biocompatibility, in the case of synthetic polymers, and differences between batches, in the case of natural polymers may be a drawback to the use of this gelators.[4] Therefore, peptide hydrogels present a good alternative to polymer hydrogels for 3D cell culture. They are easy to synthesize, there is an uniformity between batches, are easy to functionalize (e.g., with adhesion motifs) and have low cost. Other characteristics, such as gel stiffness and porosity, can be regulated by peptide sequence, concentration and by the method of gelation. These gels have to be easy to handle at 37 °C and physiological pH; have rapid and reproducible gelation under mild conditions and mechanical properties that resemble those of natural tissue (0.1-100 kPa); present uniformity at the macro, micro and nanoscopic levels; have optical transparency for straight forward analysis of results and be compatible with long term culture.[5] The possibility to formulate gels that match cell type is a great advantage.[5] The introduction of non-proteinogenic amino acids can modulate the rate of hydrolysis of the peptides, which is another great advantage of these systems.[6]

The insertion of bioactive ligands in the hydrogelator structure is a common strategy used in the design of new biomaterials that mimic the ECM,[7] as they induce biological responses in the cells. The sequence arginine-glycine-aspartic acid (RGD), in particular, is one of the most studied natural ligands and has been applied in the synthesis of these materials.[7a,8] The RGD sequence is responsible for the interactions of proteins of the extracellular matrix with a group of cell-surface receptors called integrins,[9] particularly $\alpha_v\beta_3$ and $\alpha_5\beta_1$ integrins, located in the cell membranes.[10] RGD has been used in small to ultra-small peptide hydrogels for both 2D and 3D cell culture,[8a,10-11] although the results have sometimes presented some problems, due to gel contraction,[8a] presence of organic solvents,[11] or formation of aggregates.[10] Thus, the development of new hydrogels capable of mimicking the ECM is still in demand.

In this work, it was decided to conjugate an RGD peptide [GR(Pbf)GD(O^{t}Bu)G] with the dehydrodipeptide Npx-L-Ala-Z-ΔPhe-OH to increase its resistance to enzymatic hydrolysis. This dehydrodipeptide has previously shown to form hydrogels with good rheological properties and did not presented toxicity towards fibroblasts.[3]

Results and Discussion

A new dehydropeptide containing the RGD sequence was prepared using a combination of solution and solid phase peptide synthesis. The naproxen-alanyl-dehydrophenylalanine

[a] Department of Chemistry
University of Minho
Campus de Gualtar, 4710-057 Braga, Portugal
E-mail: pmf@quimica.uminho.pt
[b] Institute for Polymers and Composites/I3N
Department of Polymer Engineering
University of Minho
Campus de Azurém, 4800-058, Guimarães, Portugal
[c] Chemistry Department
University of Reading
Chemistry Building, Whiteknights , PO Box 224, Reading, RG6 6AD, UK
[d] CITAB
Department of Biology
University of Minho
Campus de Gualtar, 4710-057 Braga, Portugal
[e] IBB, Institute for Biotechnology and Bioengineering
Center of Biological Engineering
University of Minho
Campus de Gualtar, 4710-057 Braga, Portugal
[f] Center ALGORITMI
Universidade do Minho
Campus de Gualtar, 4710-057 Braga, Portugal

Supporting information for this article is given via a link at the end of the document.

(Npx-L-Ala-Z-ΔPhe-OH) (**1**, Scheme 1) was prepared from the methyl ester of *tert*-butoxycarbonyl-alanyl-β-hydroxyphenylalanine (Boc-L-Ala-Phe(β-OH)-OMe) according to the procedure already described.[3] The pentapeptide with the RGD sequence was synthesized by solid phase peptide synthesis using a fluorenyl-9-methoxycarbonyl (Fmoc) strategy and a 2-chlorotrityl chloride resin. For side-chain protection the 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl group (Pbf) for arginine and the *tert*-butyl ester group for aspartic acid were used. The peptide was elongated using Fmoc-amino acids and diisopropylcarbodiimide (DIC) and 1-hydroxybenzotriazole (HOBt). The coupling of the naproxen dehydrodipeptide **1** was carried out in solid phase using the same methodology (Scheme 1). The RGD dehydropeptide was cleaved from the resin using a mixture of 2,2,2-trifluoroethanol (TFE) and acetic acid (AcOH). The side-chain protecting groups were removed by treatment with trifluoroacetic acid (TFA). The RGD dehydropeptide **3** (Scheme 1) was obtained in 45% overall yield.

In the ^{1}H NMR spectrum of peptide **3** it is possible to observe the signals corresponding to the β-proton and to the α-NH of the dehydrophenylalanine residue at 7.11 ppm and 9.71 ppm, respectively. The signals corresponding to CH_2 protons of the glycine residues appear as a multiplet between 3.69 and 3.76 ppm. The Z-stereochemistry was confirmed by NOE difference experiments.
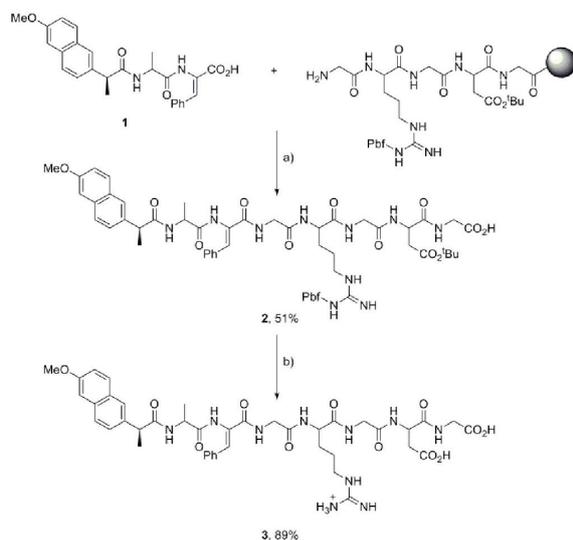
Scheme 1. Synthesis of the RGD dehydropeptide **3**; a) i) HOBt, DIC, DMF, rt, 18 h, ii) AcOH/TFE/DCM (1:1:3), rt, 4 h; b) TFA, rt, 5 h.

The hydrogelation capacity of peptide **3** was tested and it was found that this peptide formed consistent and stable hydrogels with a critical gelation concentration (CGC) of 0.32 wt% in phosphate buffer pH 6.0 (0.1 M), after a heat/cool cycle (Figure 1). Comparing these results with those obtained with the naproxen-alanyl-dehydrophenylalanine hydrogelator **1**, namely a

CGC of 0.8 wt% and a pH of 5.0,[3] it is possible to observe an increase in the gelation pH and a decrease in the CGC.
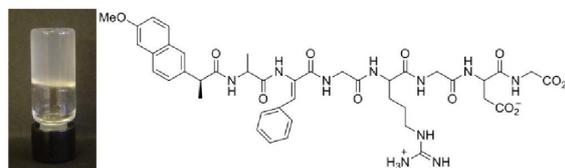
Figure 1. Hydrogel from RGD dehydropeptide **3** (0.32 wt%) in phosphate buffer pH 6.0 (0.1 M).

Peptide **3** gave consistent hydrogels in a few minutes between concentrations of 0.32 and 0.50 wt%. However, the turbidity of the gels increased with concentration (Figure S1). At concentrations above 0.50 wt% peptide **3** was not totally soluble giving suspensions and at concentrations below 0.32 wt% no consistent hydrogels were formed. The most concentrated gels (0.46 wt% - 0.50 wt%) showed syneresis after a few days and eventually settled into suspensions. The gels with the lowest concentrations (0.32 wt% - 0.40 wt%) proved to be stable for more than a year at room temperature. All the aged gels were re-heated at 80 ºC, giving colourless solutions. When cooled to room temperature, the solutions formed gels again. This cycle could be repeated several times. The gel-sol transition temperature (T_{GS}) for peptide **3** determined using the inverted tube test was 53 ºC. Raising the temperature causes disassembly of the structure and dissolution of the gel. The hydrogel organization is restored on standing at lower temperatures. Although the gel could be considered thermoreversible, after three cycles of heating and cooling, the T_{GS} changed from 53 ºC to 39 ºC suggesting that the gel strength decreases during this process. Applying mechanical forces to this hydrogel also destroys the 3D structure giving a clear solution that re-assembles to give a hydrogel upon standing a few minutes.

The structure of RGD dehydropeptide **3** hydrogel was studied using electron microscopy techniques SEM (Figure S2) and TEM. Figure 2 shows a transmission electron microscopy (TEM) image of **3**.
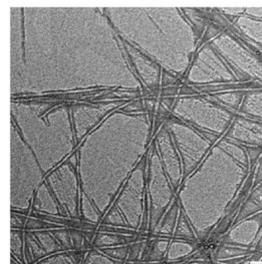
Figure 2. TEM image of peptide **3** obtained from a stained (uranyl acetate) and dried sample of a solution of 0.060 wt% in phosphate buffer pH 6 (0.1 M); Scale bar 100 nm.

The dehydropeptide self-assembles in a dense network of long nanofibres with some polydispersity in diameter and a mean width of 23 nm. The fibres showed some flexibility.

The circular dichroism (CD) spectrum of the hydrogel of peptide **3** at 0.32 wt% presented four bands (Figure 3 - gel) The positive Cotton effect around 240 nm and the negative band around 202 nm correspond to interactions between the aromatic moieties.[12] The negative Cotton effect around 200-205 nm suggest the presence of an unordered structure, which is common in peptides with ionized side chains.[13] It has been previously observed that a hydrogel with the RGD sequence has a less ordered structure than an aromatic dipeptide hydrogel.[8a] Unordered structures usually present a strong band just below 200 nm, while peptides with β-sheets show a positive band near 196 nm and a strong negative band around 216 nm.[13] The presence of both would give a spectrum with overlay of bands, thus presenting a more complicated picture. The band around 300 nm is due to π-π* transitions in the naphthalene moiety.[14] After breaking the gel, the CD spectrum (Figure 3 - solution) showed a decrease in the Cotton effect at 300 nm and an increase in the negative band at 200 nm, suggesting less affective π-π stacking interactions and an unordered structure.
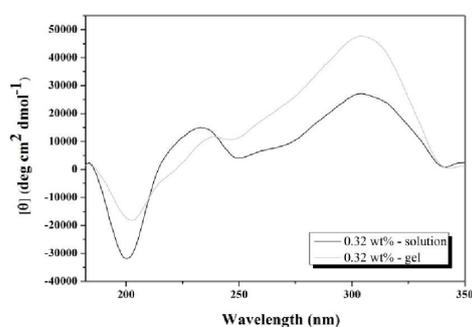
Figure 3. CD spectra of peptide **3** at 0.32 wt% (solution and hydrogel) in phosphate buffer pH 6 (0.1 M) at rt.

The secondary structure of peptide **3** was also investigated using X-ray diffraction (XRD) of a gel (0.32 wt%) (Figure 4). The 2D XRD (Figure 4A) shows a partially aligned pattern, with meridional 5.5 Å reflections and main equatorial reflections at 42.9 Å and 10.4 Å. Less aligned patterns revealed reflections at 4.5 Å, 4.0 Å, 3.8 Å and 3.5 Å. The reflections at 10.4 Å, 5.5 Å and 4.5 Å arise from the stacking of β-sheets.[15] The 5.5 Å, 4.5 Å and 3.8 Å peaks have previously been observed for the heptapeptide A$_6$K, that self-assembles into nanotubes in aqueous solution.[15b] The origin of these peaks was attributed to the helical wrapping of peptide dimmers. The broad reflections at 4.5 Å in the meridional axis were attributed to the separation of the β-strands,[15b] while the 10.4 Å reflection was attributed to the spacing of the β-sheets.[16] The sharp peaks at 3.8 Å correspond to the distance between the α-carbon atoms in the peptide backbone.[15b] The peaks between 4.0 Å and 3.5 Å are

typical of van der Waals packed peptide side-chains.[17] The peak at 42 Å is consistent with the fibre width measured in TEM.
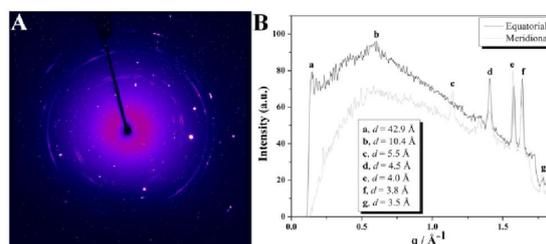
Figure 4. A) Fibre X-ray diffraction pattern obtained from a dried stalk of RGD gel (0.32 wt%, pH 6); B) Fibre X-ray diffraction one-dimensional radial averages with indicated d-spacings.

The viscoelasticity of a hydrogel is an essential characteristic to be considered in biomedical applications. The sol-gel transition temperature (T$_{SG}$) of peptide **3** determined by the crossover between the storage modulus G' and the loss modulus G" was 24 °C (Table 1) (Figure S3A). When compared with the T$_{GS}$ measured using the inverted tube test, the T$_{SG}$ is much lower than the T$_{GS}$ suggesting that the gel is kinetic dependent and after being formed undergoes structural changes that lead to a stronger gel. It is possible that the gel is formed by aggregates that establish crosslinks after a certain concentration and with time. Thus it is expected to obtain a stronger hydrogel with an increase in the gelation time. The mechanical spectrum (Figure S3C), obtained after 30 minutes of gel structural built-up at 20 °C, showed a very weak frequency dependence of G' which is 10 times larger than G" (Table 1), indicating a strong physical gel. However, the slow increase in gel elasticity with time may contribute to the apparent weak frequency dependence. In the dynamic strain sweep (Figure S3D), the gel showed a critical strain value of 3% for the onset of yielding. G" becomes larger than G' only for strain values in excess of 100%, suggesting that a large strain is needed to reversibly fluidize the gel. When the gel is allowed to re-build for 30 minutes the new mechanical spectrum (Figure S3F) shows a storage modulus 10 times larger than the loss modulus, indicating that this gel is self-healing. Nevertheless, both G' and G" are much smaller (Table 1), suggesting that the new gel has a different structure. When re-heated at 65 °C, the re-built time was increased to 60 minutes at 20 °C (Figure S4). The new T$_{SG}$ obtained was of 47 °C (Table 1, Figure S4A) which is higher than the first determined. This is in agreement with the explanation above: if heating the gel at 65 °C does not destroy all of the aggregates, it is necessary less time and a lower drop in temperature to restore the gel. After 30 minutes of structural build-up, the gel recovered 21% of its storage modulus, and after 60 minutes, 31% (Figure S4B). Both in the frequency and in the strain sweeps, the new gel presented values much smaller than the previous one (Table 1, Figure S4C-D). This new gel broke at strains of 1% and G" become larger than G' for strain values of only 10% (Figure S4D).

However, after breaking the gel and 60 minutes of structural re-build, the gel presented values for both storage and loss moduli higher than the gel obtained in the first cycle, with a G' 10 times larger than G'' (Table 1, Figure S4F), indicating a stronger physical gel. These values are to be taken with caution as experimental errors such as wall slip and non-homogeneous stresses and strains upon gel break up might affect the values of the moduli. However, from the rheological data is clear the structural healing of the hydrogel of peptide **3**.

Table 1. Rheological properties of the hydrogel (0.50 wt% in phosphate buffer pH 6, 0.1 M) of peptide **3**.

	T_{SG}[a] [ºC]	Dynamic Frequency Sweep[b]		Dynamic Strain Sweep[c]			Dynamic Frequency Sweep[d]	
		G' [Pa]	G'' [Pa]	G' [Pa]	G'' [Pa]	Critical Strain [%]	G' [Pa]	G'' [Pa]
1st cycle	24	581	32	1240	56	3.3	38	4
2nd cycle	47	171	11	192	13	1.0	247	18

[a] The values are taken at f = 1 Hz and γ = 0.5%; [b] The values are taken at f = 1 Hz, γ = 0.5%, T = 20 ºC, after 30 minutes (1st cycle) or 60 minutes (2nd cycle) of structural build-up (the test took 20 minutes); [c] The values are taken at f = 1 Hz and T = 20 ºC (the test took 7 minutes); [d] The values are taken at f = 1 Hz, γ = 0.5%, T = 20 ºC, after shear breaking and 30 minutes (1st cycle) or 60 minutes (2nd cycle) of structural build-up (the test took 20 minutes).

The toxicity of peptide **3** was tested in human skin dermal fibroblasts and the results were compared with those obtained with Npx-L-Ala-Z-ΔPhe-OH **1**. According to the results obtained peptide **3** did not show any toxicity even at concentrations of 500 µM (Figure 5). These results are similar to those reported with the dehydrodipeptide **1**. The effect of peptide **3** on cell adhesion was also tested in human skin fibroblasts and compared with the results obtained with the dehydrodipeptide **1**. It was found that cell adhesion 1 hour after seeding was remarkably delayed in the presence of peptide **3**, but not in the presence of dehydrodipeptide **1** (Figure S5). This indicates that the RGD sequence may be blocking the adhesion sites of integrins.
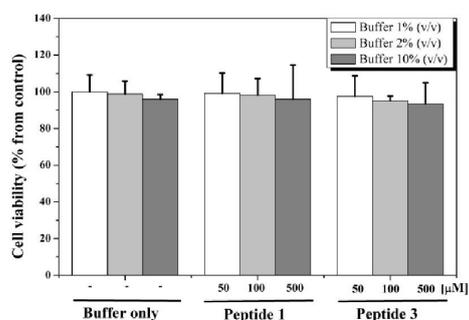
Figure 5. Cell viability of adult human skin fibroblasts for 48 h, after incubation with 50 µM, 100 µM or 500 µM of peptides **3** and **1**. No significant differences were observed (P > 0.05).

Taking in consideration the results obtained with the RGD dehydropeptide **3** it was decided to test this hydrogelator as a 3D cell culture medium. Thus, it was necessary to prepare a hydrogel of peptide **3** in a medium that was suitable for cell growth, i.e. a medium with a pH and an osmotic concentration similar to those of the physiological medium. The hydrogel will have to be formed at a higher pH and should be stable at 37 ºC. Since compound **3** proved to drop the pH of buffers when in higher concentrations, it was necessary to previously neutralize the solution of peptide **3** with NaOH. The self-assembly capacity of the neutralized peptide **3** at 37 ºC and pH 7-8 was tested using several buffer solutions. The hydrogelation of peptide **3** in a phosphate-buffered saline (PBS) solution at pH 7.39 was tested, as this is the most common buffer used in biological studies, as it is isotonic. Despite forming turbid gels between 1.6 and 3.2 wt%, the pH dropped to 6.53-6.40, indicating that the PBS buffering capability is not enough in this case. Several phosphate buffers were also tested and were used to form gels of RGD peptide **3**, without losing its buffer capacity. Figure 6 shows the relation between peptide **3** concentration and temperature in phosphate buffers with pH from 7.11 to 8.33. It is possible to conclude that at 60 ºC solutions are obtained independent from the pH used and the peptide concentration. Gels are obtained at 37 ºC between 2.0 and 3.5 wt% in phosphate buffers between pH 7.11 and 8.33. At 1.7 wt% (pH 7.36) however, the gel that forms at 20 ºC turns into a solution at 37 ºC, indicating that gels must have a higher concentration of peptide **3**.
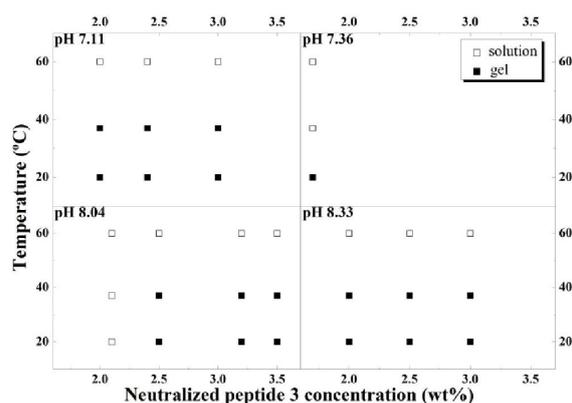
Figure 6. Phase diagram of neutralized peptide **3** at different concentrations and temperatures (60, 37 and 20 ºC) in phosphate buffers (0.1 M) with different pHs (pH 7.11, 7.36, 8.04 and 8.33).

As the gels for cell culture would require not only stability at 37 ºC, but also resistance to changes due to the cell culture medium (DMEM) these characteristics were also tested (Figure 7). All of the gels absorbed some or all of the culture medium added. Only gels obtained at pH 8.04 in a buffer solution and in concentrations between 3.2 and 3.5 wt% were stable for 24 hours after the addition of DMEM at 37 ºC. Daily changes of the

medium revealed that these gels were stable for more than four days.
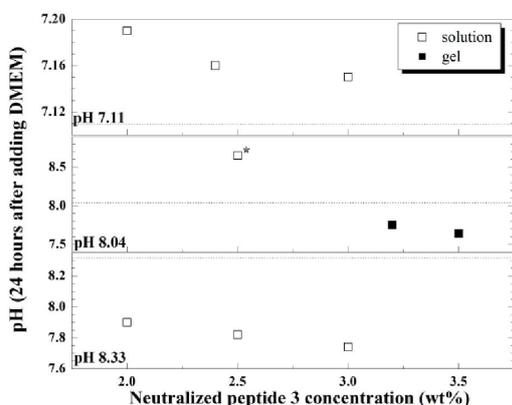
Figure 7. Phase diagram of neutralized peptide **3** at different concentrations in phosphate buffers (0.1 M) with different pHs (pH 7.11, 8.04 and 8.33), obtained at 37 ºC, 24 hours after adding DMEM (pH 8.07) to the gels of this peptide; *After 24 hours was a mixture of gel and solution; turned a solution only 72 hours after adding DMEM.

Given the results, it was decided to use a buffer solution (pH 8.04) of neutralized peptide **3** with CGC of 3.2 w% to give gels to be tested as 3D cell medium. After preparing these hydrogels of neutralized peptide **3** in phosphate buffer and having culture media it was necessary to evaluate their capability to mimic the extracellular matrix. Therefore it was decided to prepare another dehydropeptide having the sequence arginine-glycine-glutamic acid (RGE) (Scheme 2, compound **4**) instead of the well known RGD sequence. The preparation of this compound was carried out using a strategy similar to that described for peptide **3**. Peptide **4** was obtained in a 69% overall yield.
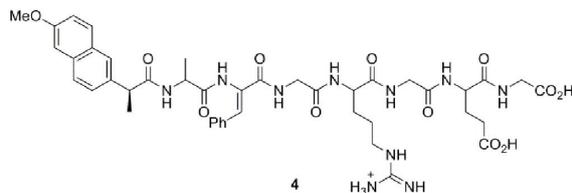
Scheme 2. Structure of peptide **4**.

Peptide **4** was also tested in the gelation conditions established for peptide **3** and proved to form almost clear gels at 3.2 wt% that resisted to the addition of DMEM at 37 ºC, absorbing part of it. At 3.6 wt%, however, like what was observed with **3** (result not shown), the gel turned solution 24 hours after adding DMEM. The viscoelasticity of the gels of neutralized dehydropeptides **3** and **4** was determined by rheometry. Phosphate buffer (pH 8, 0.1 M) was added to peptides **3** and **4** and the mixture heated to 80 ºC, in order to obtain solutions/gels at 3.2 wt%. The gels were not obtained in the cooling ramp (Figure S7A), but only after structural build-up for 1.5-4.5 hours (Figure S7B) at 20 ºC, indicating that the self-assembly of the neutralized peptides **3** and **4** in these conditions is mainly driven by kinetics, although they are temperature dependent. The gel of peptide **3** was impossible to measure, as the gel was too brittle and could not sustain the strains and stresses used during rheological testing. This is evident from the strain sweep (Figure S6A and C), since no linear behaviour is obtained at 0.1%. However, the gel did showed instantaneous re-building of a stiff gel (Figure S6B). The gel of **4** was less brittle and, at the applied strain (0.1%), it was possible to obtain rheological data (Figure S7). The mechanical spectra (Figure S7C), obtained after ~8 h of gel formation and structural build-up, showed a weak frequency dependence of the elastic storage modulus *G'* which is 3-5 times larger than the loss modulus *G''* (Table 2), indicating a strong physical gel, that broke at strain of 0.3% (Figure S7D).

Table 2. Rheological properties of the hydrogels of neutralized peptides **3** and **4** (3.2 wt% in phosphate buffer pH 8, 0.1 M).

	Time to form gel[b] [h]	Dynamic Frequency Sweep[c]		Dynamic Strain Sweep[d]		Dynamic Frequency Sweep[e]		T_{GS}[f] [ºC]
		G' [kPa]	*G''* [kPa]	Critical Strain [%]	*G''* crosses *G'* [%]	*G'* [kPa]	*G''* [kPa]	
3	4h30	---	---	---	---	---	---	71.6
3[a]	5h30	---	---	---	---	---	---	---
4	0h43	2.9	0.94	0.3	7.1	1.2	0.38	68.0
4[a]	1h30	0.98	0.29	0.2	10.4	---	---	---

[a] After a heat/cool cycle; [b] The values are taken at *f* = 1 Hz, *γ* = 0.1%, T = 20 ºC, and corresponds to the point where *G'* crosses *G''*; [c] The values are taken at *f* = 1 Hz, *γ* = 0.1%, T = 20 ºC, after 8h20 at 20 ºC (the test took 22 minutes); [d] The values are taken at *f* = 1 Hz, T = 20 ºC (the test took 7 minutes); [e] The values are taken at *f* = 1 Hz, *γ* = 0.1%, T = 20 ºC, after 2 hours of structural build-up; [f] The values are taken at *f* = 1 Hz, *γ* = 0.1% (heating at 6.0 ºC min⁻¹).

The gel was then allowed to re-build for 2 hours and a new mechanical spectrum was obtained (Figure S7F). The self-recovered gel still presented storage moduli 3-4 times larger than the loss moduli, indicating that the gel is self-healing. However, the recovered **4** gel presented lower *G'* and *G''* (Table 2), which suggests that the new gel has different structure when

compared with the initial gel. Again, experimental errors inherent to rheological testing under larger deformation (see above) may contribute to such differences. Both gels were then heated again to 80 ºC and the protocol repeated. In the heating sweep it was possible to determine the gel-solution temperature (T_{GS}), which is similar for both gels. **3** suffers a transition gel-solution at 71.6 ºC and **4** only just below, at 68 ºC (Table 2). In the second cooling sweep, similar results as in the first cycle were produced, with the gels only forming after some time at 20 ºC. For both samples, the time for gel formation was around one more hour than initially (Table 2). After ~8 h at 20 ºC, a new mechanical spectrum for the gel of **4** was obtained (Figure S7J), showing that the thermo-recovered gel has 34% of the initial storage moduli (Table 2). In the strain sweep of **4**, the thermo-recovered gel showed a lower strain of yielding (Table 2, Figure S7K). The results indicated that the gels of **3** and **4** are not completely thermo-reversible, as reported for the gel of **3** before neutralization.

Transmission electron microscopy was used to analyze the type of fibres formed by the neutralized peptides **3** and **4** at pH 8. Solutions 5 times more diluted than the gels were used, as also a solution of **3** in pH 6 buffer, in order to compare the differences, if any, between the fibres formed in both pHs. In all cases, long (length above 4 μm) and thin fibres were formed (Figure 8). Some of the fibres are straighter, while others bent. Peptide **3** in pH 6 buffer formed fibres with non uniform thickness, ranging from 50 to 110 nm. Some bifurcations and crosslinks are observed (Figure 8A). In pH 8, more fibres are observed than at pH 6, with much more bifurcations and crosslinks, and the fibres present more uniform thickness, ranging between 65-80 nm (Figure 8B). The bifurcations in the fibres do not change their thickness. Peptide **4** at the same pH and concentration as **3** presented more fibres, with more bifurcations and crosslinks (Figure 8C). Some fibres are twisted around each other (Figure S8A). The fibres, however, are slightly thinner than in **3**, with diameters between 50 and 60 nm.
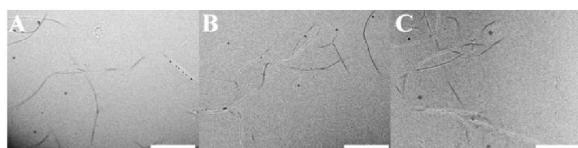
Figure 8. TEM images of; A) Peptide **3** [solution at 0.064 wt% in phosphate buffer (pH 6, 0.1 M)]; B) Peptide **3** [solution at 0.64 wt% in phosphate buffer (pH 8, 0.1 M)]; C) Peptide **4** [solution at 0.64 wt% in phosphate buffer (pH 8, 0.1 M)]; Scale bars 5000 nm.

The cell viability in the presence of neutralized peptides **3** and **4** was evaluated using the MTS assay. The cell viability was tested in phosphate buffer (its non-toxicity is shown in Figure S9) and at 37 ºC. Thus, the neutralized peptides **3** and **4** were incubated with 3T3 fibroblasts in 17% phosphate buffer pH 7.4, for up to 48 hours (Figure 9). Neutralized peptide **3**, contrary to peptide **3**, showed to be toxic at concentrations of 0.5 mM and higher toxicity at 1.2 mM. Unexpectedly, giving the structural

similarities between both peptides, the neutralized peptide **4** was not toxic until a concentration of 1.2 mM.
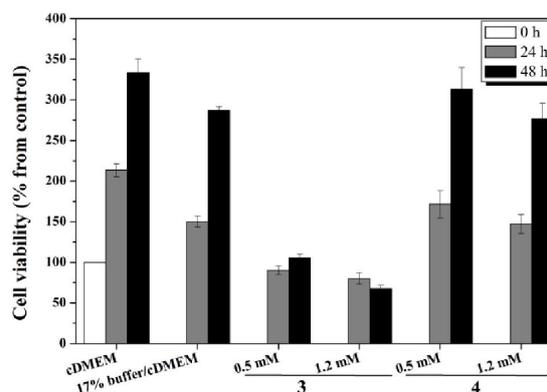
Figure 9. Cell viability of 3T3 fibroblasts after incubation for 24 and 48 hours with 0.5 mM and 1.2 mM of neutralized peptides **3** or **4**, as compared with control (cDMEM) at 0 hours. Cells with just cell culture medium and 17%(v/v) buffer (phosphate buffer, pH 7.36) in cDMEM were used as controls. Results determined through the MTS assay. Shown are mean ± SD values (n = 3).

Observation at the microscope of the cells with neutralized peptides **3** and **4** revealed that the solutions of **3**, at both concentrations, formed precipitates/crystals, which could lead to cell death, even though the compound itself is not toxic. The crystals and precipitates are not visible in cell cultures of peptide **4**. In order to try to understand if this result could be in any way related to the presence of the cells or cDMEM, solutions of peptides **3** and **4** were prepared at 7.06 mM in buffer and at 1.2 and 0.5 mM in 17% buffer/cDMEM, and observed using microscopy. In the more concentrated solution of **3** was observed an increase in the turbidity and some precipitates could be observed. The 1.2 and 0.5 mM solutions were clear, but also presented some precipitates at the microscope. Peptide **4** formed clear solutions and presented uniformity at the microscope. It has been previously reported that the formation of aggregates is prejudicial to cells.[10,18]

Despite the toxicity found for peptide **3** it was decided to test the gels of compounds **3** and **4** in cell culture. As the UV light could influence the gelation process, solutions of **3** and **4** at 3.2 wt% in pH 8.04 (phosphate buffer 0.1 M) were prepared and let to gel with or without the presence of UV light, and were then incubated for 18 hours at 37 ºC. A suspension of 3T3 fibroblasts in cDMEM was added on top of each gel and these were incubated at 37 ºC for 48 hours. A Live/Dead assay was then performed to visualize the cells to access their viability. Fluorescent images of this Live/Dead assay are presented in Figure 10.

WILEY-VCH
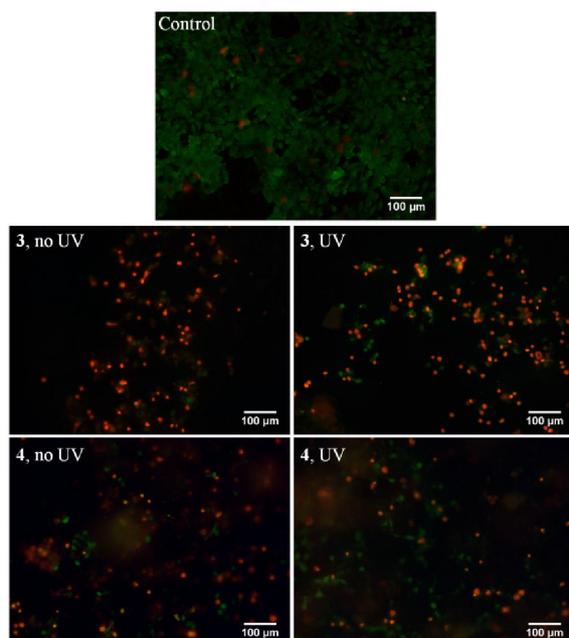
Figure 10. Fluorescence photographs of 3T3 cells incubated with the gels of **3** and **4** (3.2 wt%, pH 8.04 phosphate buffer) gelled with or without the presence of UV light, stained with LIVE/DEAD® Viability/Cytotoxicity Kit for mammalian cells. Live cells are stained in green and dead cells stained in red.

In wells with 3T3 fibroblasts in cDMEM (control wells) the number of living cells (stained in green) is superior to the number of dead cells (stained in red). When cells are seeded on top of **3** gels, the number of cells stained in red augments, outnumbering the number of cells stained in green, some presenting the typical spread morphology of adhered fibroblasts. In the gel of **3** formed in the presence of UV light, a similar result was obtained. In both cases, the cells were found mainly on the borders of the wells, indicating that the gels did not had a uniform surface and did not allowed cell penetration. In the gel of **4**, despite also showing an increase in the number of cells stained in red, the number of these is comparable to the number of cells stained in green. The cells stained in green all presented a typical morphology of the fibroblasts indicating that these have adhered to the gel. It was also observed that the cells are in different planes in the gel, which indicates that despite the seeding of the cells 2D over the gels, they have penetrated in the gels and grew in a 3D environment.

The great elasticity exhibited by the gel of neutralized peptide **3** (larger than 100 kPa) is high above the natural tissues stiffness (0.1-100 kPa). This explains in part why, in this case, the cells did not penetrate in the gel. In the case of **4**, its elastic modulus (2.9 kPa) is more in harmony with the natural tissues stiffness and, specifically, the ideal values for fibroblasts environment,[4] a reason for which the cells easily penetrated through the gel and spread in three dimensions.

These results indicate that the gel of peptide **4** can be a good matrix for 3D cell culture. However, further studies are required to better access this. In order to understand these results molecular dynamics simulations were carried out. A conformation cluster analysis was carried out through a single-linkage method with a rmsd cut-off of 0.12 nm[19] to analyze the conformational variation of the sequences GRGDG and GRGED in the RGD (**3**) and RGE (**4**) dehydropeptides from MD simulations (Figure 11).
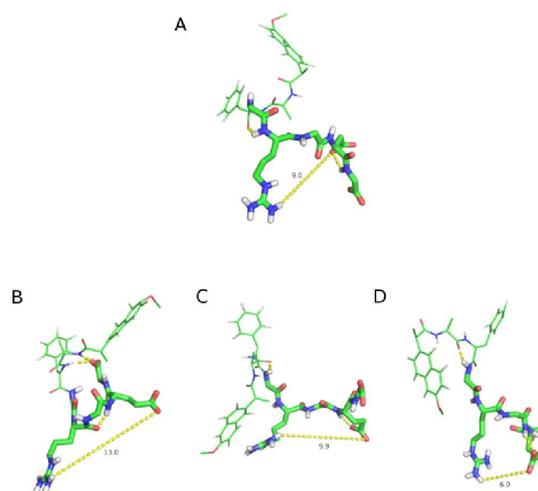
Figure 11. Most likely conformations obtained for; A) RGD dehydropeptide **3**; B) 1st RGE dehydropeptide **4** cluster; C) 2nd RGE dehydropeptide **4** cluster; D) 3th RGE dehydropeptide **4** cluster. The colouring of the atoms follows the convention: green for carbon, blue for nitrogen, red for oxygen, white for hydrogen. The dashed points highlight hydrogen bond interactions, or, when labelled, show electrostatic interaction between Arg and Asp (RGD) or between Arg and Glu (RGE). Water molecules and counter-ions were omitted for better visualization.

The conformation analysis shows that during the simulation RGD dehydropeptide **3** populates only one conformation 85% of the time, (Figure 11A). In contrast, RGE dehydropeptide **4** presents three distinct conformations with populations of 49%, 15% and 11% (Figure 11B-D). The electrostatic interactions in these peptides are relevant as there are two opposites charged amino acids and a deprotonated C-terminus, that probably direct the crystallization of RGD, and are highlighted in Figure 11. The large conformational stability of RGD is probably related with the observed propensity of peptide **3** to aggregation in solution. The existence of different RGE conformers hinders the kinetics of RGE molecules clustering in solution and consequently hinders their aggregation.

The biostability of neutralized peptides **3** and **4** was studied through incubation of peptide solutions with α-chymotrypsin at 37 ºC. The solutions were then followed by analytical HPLC-ESI-MS. The peptides have similar rate of hydrolysis and undergo fast proteolytic degradation, remaining 50% only 6 hours after

155

incubation. After that, the rate of cleavage decreased and only after 48 hours of incubation did all the peptide disappeared (Figure 12). This is a much lower rate of hydrolysis when compared to other RGD peptides with a phenylalanyl-phenylalanine sequence in the presence of proteinase K, but higher than a similar peptide, with only a phenylalanine residue.[20] As the peptides suffered enzymatic breakdown, the chromatogram changed and two new peaks in the MS spectrum appeared, with a rate similar to the hydrolysis of the peptide. Interestingly, the new peaks resulting from the cleavage in **4** had exactly the same mass as the peaks that appeared for the cleavage of **3** (m/z = 660 and 682) (Figure S10). α-Chymotrypsin was chosen for its ability to preferentially cleave peptide amide bonds where the carboxyl side of the amide bond belongs to an aromatic amino acid. Previously, it was shown the resistance of naproxen-dehydrodipeptides to α-chymotrypsin.[3] Given that, in here, we have the dehydrodipeptide conjugated with a pentapeptide, we expected that the proteolytic cleavage would occur in the amide bond between the ΔPhe and Gly residues. However, the peaks observed in the mass spectrum do not correspond to this type of cleavage.
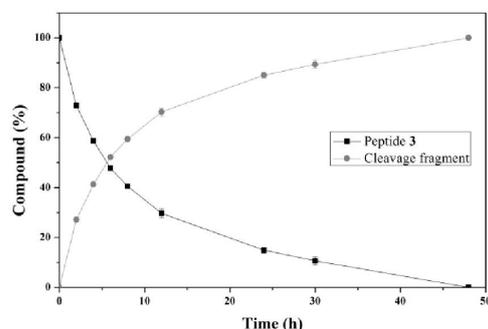
Figure 12. Evaluation of the proteolitic stability of peptide **3** (peptide **4** showed a similar result) in the presence of α-chymotrypsin (pH 7.4, 37 °C) for 48 hours.

Conclusions

A new dehydropeptide hydrogelator with the RGD sequence was designed, synthesized and characterized. The new hydrogels prepared could be useful in several biomedical applications such as drug delivery or as 3D surrogates of the extracellular matrix (ECM). Herein the use of this hydrogelator as a 3D ECM surrogate was evaluated and although the RGD hydrogel is viscoelastic, thermoreversible and self-healing, the fact that it gave hydrogels with precipitates affected the cell growth. When compared with a similar hydrogelator that has a glutamic acid residue instead of an aspartic acid it is possible to conclude that the RGE dehydropeptide is a much more promising surrogate of ECM than the RGD dehydropeptide. A possible explanation to these results was obtained by molecular dynamics simulations. These studies revealed that the RGD dehydropeptide have a preferential conformation which could be

responsible for the formation of aggregates. The conformational stability of RGE peptide is much smaller and thus this peptide do not aggregates.

Both RGD and RGE dehydropeptides are susceptible to enzymatic hydrolysis.

These new hydrogelators, due to their excellent viscoelastic properties, can be used in drug delivery and the peptide with the RGE sequence can mimic the ECM and be used in 3D cell cultures.

Experimental Section

General methods

^{1}H and ^{13}C NMR spectra were recorded on a Bruker Avance III at 400 and 100.6 MHz, respectively. [1]H–[1]H spin–spin decoupling and DEPT θ 45° were used. HMQC and HMBC were used to attribute some signals. Chemical shifts (δ) are given in parts per million (ppm) and coupling constants (J) in hertz (Hz). High resolution mass spectrometry (HRMS) data were recorded by the mass spectrometry service of the University of Vigo, Spain. DCM was dried over calcium chloride (CaCl) and calcium hydride (CaH$_2$) and then distilled and stored under molecular sieves. In solid phase peptide synthesis (SPPS) was used the resin 2-chlorotrytil chloride (100-200 mesh) 1% DVB, with a loading capacity of 1.4 mmol g^{-1}. All solutions were made up with ultra filtered (18 MΩ) water from a Barnstead Nanopure system. Phosphate buffers were prepared from NaH$_2$PO$_4$ and Na$_2$HPO$_4$ with a final concentration of 0.1 M and pH 6.00, 7.11, 7.36, 7.49, 8.04 or 8.33. Phosphate-buffered saline (PBS) was prepared from 2.7 mM KCl, 137 mM NaCl, 10 mM Na$_2$HPO$_4$.2H$_2$O and 1.8 mM KH$_2$PO$_4$ with pH 7.39.

Self-assembly in buffer: Briefly, compounds were weighted into a sample vial, the buffer was added and the mixture was heated to 80 °C and left to cool at room temperature or at 37 °C.

Circular dichroism: The CD spectra were recorded at 20 °C on a Chirascan spectropolarimeter (AppliedPhotophysics, UK). Peptide hydrogels were loaded into 0.1 mm quartz cells. Spectra display absorbance <2 at any measured point with 0.5 nm step, 1 nm bandwidth and 1 second collection time per step, taking three averages. The post-acquisition smoothing tool from Chirascan software was used to remove random noise elements from the averaged spectra. A residual plot was generated for each curve in order to verify whether or not the spectrum has been distorted during the smoothing process. Following background (buffer) correction, the CD data were normalized to molarmean residue ellipticity.

Scanning transmission electron microscopy: STEM experiments were performed using an ultra-high resolution field emission gun scanning electron microscopy (FEG-SEM), NOVA 200 Nano SEM, FEI Company (SEMAT/UM), operated at 15 and 18.5 kV, using a STEM detector. Cu-C grids (S160-4 AGAR) were immersed in the peptide hydrogels. The grid was then allowed to dry at room temperature.

Transmission electron microscopy: TEM experiments with peptide **3** were performed using a Philips CM20 transmission electron microscope operated at 200 kV. The shiny side of 300 mesh Cu grids coated with a carbon film (Agar Scientific, UK) was placed over one drop of the peptide solution (0.060 wt% in phosphate buffer pH 6, 0.1 M) for 1 minute. The excess at the sides of the grid was cleaned very carefully. The shiny side of the grid was placed over a drop of aqueous uranyl acetate (1 wt%)

WILEY-VCH

(Agar Scientific, UK) for 1 minute. The excess at the sides of the grid was cleaned very carefully. The grid was then allowed to dry at room temperature. TEM experiments with neutralized peptides **3** and **4** were performed using a FEI-Tecnai G2 Spirit Biotwin transmission electron microscope (IBILI, Faculty of Medicine, University of Coimbra) operated at 100 kV. The samples were prepared as follows: 5 µL of the peptide solution was placed over the shiny side of a 300 mesh carbon coated copper grid (TAAB) for 1 minute. The excess at the sides of the grid was cleaned very carefully. The shiny side of the grid was placed over a drop of uranyl acetate aqueous solution (2%) (Agar Scientific Ltd.) for 1 minute. The excess at the sides of the grid was cleaned very carefully. The grids were then let to dry. The solutions of the peptides prepared for TEM were 5 times more diluted than the gel concentrations (0.064 or 0.64 wt%), in phosphate buffer (pH 8.04 or 6.00).

X-ray diffraction: XRD measurements were performed on a stalk prepared by drying samples of the hydrogel. The hydrogel was suspended between the ends of wax-coated capillaries and dried. The stalk was mounted (vertically) onto the four axis goniometer of a RAXIS IV++ X-ray diffractometer (Rigaku) equipped with a rotating anode generator. The XRD data was collected using a Saturn 992 CCD camera. One-dimensional profiles in the equatorial and meridional reflections (with appropriate re-alignment of images to allow for fibril orientation) were obtained using the software CLEARER[21] which was also used to fit peak positions.

Rheology: Rheological experiments were performed on a PaarPhysica MCR300 stress-controlled rheometer, equipped with a temperature controlled Couette geometry (diameter 10 mm). The hydrogel of peptide **3** (0.5 wt%, phosphate buffer pH 6, 0.1 M) was heated to 65 °C and transferred to the rheometer, pre-programmed to 65 °C. During the temperature cooling ramp (5 °C min^{-1}), the solution/gel was tested at 1 Hz and 0.5% strain. In the kinetic studies, the gel was sheared at 1 Hz, 0.5% strain and 20 °C. Dynamic frequency sweeps were performed at 0.5 % strain and 20 °C. During the strain sweep experiments the gel was submitted to different strains (0.1 to 500%), at constant frequency (1 Hz) and temperature (20 °C). For the second cycle, the rheometer was heated to 65 °C and the cycle repeated. Neutralized peptides **3** and **4** were dissolved at 80 °C in phosphate buffer (pH 8.04, 0.1 M) to a final concentration of 3.2 wt%, and the solutions were transferred to the rheometer, pre-programmed to 80 °C. During the temperature cooling ramp (1.17 °C min^{-1}), the solution/gel was sheared at 1 Hz and 0.1% strain. In the kinetic studies, the gel was sheared at 1 Hz, 0.1% strain and 20 °C. Mechanical spectra were recorded using constant strain (0.1%) and temperature (20 °C). During the strain sweep experiments the gel was under different strains (0.1 to 500%), constant frequency (1 Hz) and temperature (20 °C). During the temperature heating ramp (6 °C min^{-1}), the gel/solution was submitted to a 0.1% strain cycled at 1 Hz.

HPLC-MS / Enzymatic resistance assay: Diluted solutions of neutralized peptides **3** and **4** (0.5 mg mL^{-1}) were prepared in sodium phosphate buffer (pH 7.49, 0.1 M). Samples of these solutions were filtered (PES 0.2 µm), 20% MeOH was added (100 µL / 500 µL of buffer) and the solutions were used in direct injection in the MS, to define the parameters for the measurements during the assay. A solution of α-chymotrypsin in the same buffer was also prepared (0.5 mg mL^{-1}; 51.33 U mL^{-1}). All solutions were incubated at 37 °C and 20 rpm overnight. The solutions of the peptides were divided into 9×3 vials of 300 µL each. The enzyme solution (300 µL) was added to each vial of peptide solution. Samples of 10 µL were taken at 0 h, 2 h, 4 h, 8 h, 12 h, 24 h, 49 h and 78 h and analyzed by HPLC (λ = 276 nm; water/acetonitrile, 1:1 with 0.1% TFA). The percentage of gelator was determined using the peptide peak area in each sample and comparing it with the area of the same peak in the diluted solutions without the enzyme. To verify that these solutions were

stable at 37 °C, the samples of each peptide were analyzed by HPLC after 78 hours at 37 °C and 20 rpm. Also, at 0 h, 2 h, 4 h, 6 h, 8 h, 12 h, 24 h, 30 h and 48 h, 3 vials of each peptide/enzyme solutions were removed from the incubator, filtered (PES 0.2 µm) and 20% MeOH added. The solutions (V_{inj} 25 µL) were then monitored by analytical LC-ESI-MS (water/acetonitrile, 1:1 with 0.1% formic acid; flow 0.4 mL min^{-1}).

*MTT assay with peptides **1** and **3**:* Adult human skin fibroblasts (ASF-2 cells) were maintained at 37 °C in a humidified 5% CO$_2$ atmosphere grown in Dulbecco´s Modified Eagle's Medium (DMEM, Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS, Lonza, Verviers, Belgium), 10 mM Hepes and 1% antibiotic/antimycotic solution (Sigma-Aldrich, St. Louis, MO, USA). Prior to culture, cells within a sub-confluent monolayer were trypsinized using trypsin (0.05%)-EDTA.4Na (0.53 mM) solution and resuspended in DMEM to obtain a cell concentration of around 50 000 cells per mL. The cells were plated in 96-multiwell culture plates (100 µL per well) 24 hours before incubation with compounds **1** and **3**. Cells were then treated with different concentrations of **1** and **3**, prepared as follows: the peptides were dissolved in phosphate buffer (0.1 M, pH 8), obtaining solutions of 5.0 mM. The 5 mM solutions were used to prepare solutions of 50 µM, 100 µM and 500 µM in DMEM. Solutions of phosphate buffer (0.1 M, pH 8) at 1%, 2% and 10% in DMEM were prepared as controls. 100 µL aliquots of buffer controls and peptide solutions were placed into the wells of the plate with the cell culture, with three replicas of each. The plate was incubated at 37 °C for 48 hours. Cells were then incubated for 60 minutes with MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Sigma-Aldrich, St. Louis, MO, USA] to a final concentration of 0.5 mg mL^{-1}. Then, the medium was removed, and the formazan crystals formed by the cell's capacity to reduce MTT were dissolved with a 50:50 (v/v) DMSO:ethanol solution, and absorbance measured at 570 nm (with background subtraction at 690 nm), in a SpectroMax Plus384 absorbance microplate reader. The results were expressed as percentage relative to the control (cells with buffer solution).

*Biological assays with neutralized peptides **3** and **4**:*

Cell cultivation: Mouse embryo fibroblasts 3T3 (ATCC CCL-164) were grown in DMEM (Biochrom GmbH, Berlin, Germany) supplemented with 10% newborn calf serum (Invitrogen, CA) and 1 mg mL^{-1} penicillin/streptavidin (DMEM complete medium [cDMEM]) at 37 °C in a 95% humidified air containing 5% CO$_2$. At 80% confluency, 3T3 fibroblasts were harvested with 0.05 %(w/v) trypsin-EDTA and subcultivated in the same medium.

Preparation of buffer and peptide solutions for MTS assay: Buffer solutions were prepared with 5, 10 and 17 %(v/v) of PBS (pH 7.39) and phosphate buffers (pH 7.36 and 8.04) in cDMEM. The peptide solutions were prepared as follows: neutralized peptides **3** and **4** were weighted and left under UV light (UV-C) overnight. The required volume of phosphate buffer pH 7.36 to form solutions at 7.06 mM was added and the samples heated to 80 °C to help solubilisation. cDMEM was added to form solutions with 1.2 mM concentration [17 %(v/v) of phosphate buffer] and these were kept at 37 °C. These solutions were used to prepare the 0.5 mM solutions [17 %(v/v) of phosphate buffer].

Evaluation of buffers/peptides cellular cytotoxicity: Cellular cytotoxicity was assessed using the MTS assay. The 3T3 fibroblasts were seeded (75 000 cells per mL, 100 µL per well, in a 96-well polystyrene plate) and incubated at 37 °C, 5% CO$_2$ for 18 hours before incubation with the buffer/peptide solutions. Then, the culture medium was removed and replaced with cDMEM containing a different concentration of buffer/peptide solution (200 µL per well) and cells further incubated for 24 or 48 hours. Control cells were incubated with fresh medium and wells

WILEY-VCH

containing only growth medium were used as blanks. All assays were made with triplicate cell incubations.

MTS assay: Cellular viability was assessed by measuring cell concentration via mitochondrial reduction of the tetrazolium salt MTS (3-(4,5-dimethyl-2-yl)-5-(3-carboxy-methoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) in the presence of 5% phenazinemethosulfate over a 2 hours incubation period. The coloured reaction product formazan is soluble in the culture medium and can be measured spectrophotometrically at a wavelength of 490 nm with a reference wavelength of 570 nm. After the incubation time, the culture medium of each well was replaced with 100 mL of fresh culture medium and 20 mL of CellTiter 961 AQueous One Solution Reagent (Promega, CA) was added and the plate further incubated for 2 hours at 37 °C, 5% CO_2, as indicated by the manufacturer. The amount of soluble formazan produced by cellular reduction of MTS was measured at 490 nm. An MTS assay on cells incubated with cDMEM before applying the buffer/peptide solutions was also preformed (t = 0 h). Data are presented as means ± standard deviation (SD) of the indicated number (n) of determinations.

Cell culture on gel: Neutralized peptides **3** and **4** were weighted into sterile eppendorfs and put under UV light (UV-C) for 18 hours. Phosphate buffer pH 8.04 was added to a sample of **3** and another of **4**, to obtain solutions at 3.2 wt%. The samples were heated to ~60 °C and transferred to μ-slide angiogenesis ibiTreat (Ibidi, Germany) 10 μL per well, and let to form gels at room temperature under UV light (UV-C) for 1h30. Phosphate buffer pH 8.04 was added to a sample of **3** and another of **4**, to obtain solutions at 3.2 wt%. The samples were heated to ~60 °C and transferred to the μ-slide, 10 μL per well, and let to form gels at room temperature for 1h30. The plate was incubated at 37 °C in a 95% humidified air containing 5% CO_2 for 24 hours. The 3T3 fibroblasts were seeded (200 000 cells per mL, 50 μL per well) in the μ-slide with the gels, and incubated at 37 °C, 5% CO_2 for 48 hours. Control cells were incubated with fresh medium, and wells containing each gel with only growth medium were used as controls. All assays were made with duplicate cell incubations.

Live-dead assay: The LIVE/DEAD® Viability/Cytotoxicity Kit for mammalian cells (Invitrogen, CA) was used to determine cell viability and observe the cells on the peptide gels. This kit provides two-colour fluorescence cell viability assay, based on the simultaneous determination of live and dead cells with two probes that measure intracellular esterase activity and plasma membrane integrity. Briefly, after the incubation time, the growth medium was removed and the gels/seeded cells were washed with sterile PBS (2×50 μL per well), 25 μL per well of a 4 μM calcein AM and 5 μM ethidium homodimer-1 solution in sterile PBS were added to the wells, incubated for 30 minutes at 37 °C and 5% CO_2 (as indicated by the manufacturer), the solution was removed and the gels/seeded cells were washed again with sterile PBS (50 μL per well), Mounting Medium (Ibidi, Germany) was added and the cells visualized in a fluorescence microscope.

Molecular dynamics simulations: The molecular structure of the RGD and RGE dehydropeptides were designed with the program PyMOL.[22] The α,β-dehydroamino acid, ΔPhe, was parameterized and validated in previous work by the authors,[3] and the topology (bonded and non bonded parameters) was based on the equivalent encoded amino acid present in the GROMOS 54a7 force field.[23] The peptides were designed in extended conformation and placed in dodecahedral boxes of water considering a hydration layer of at least 1.5 nm between the peptide and the walls in all directions. Thus, the systems have about 3250-3300 water molecules. We used the Simple Point Charge (SPC) water model.[24] Boxes were made neutral with the addition of one Na^+ ion. Each peptide was energy minimized with a steepest descent algorithm, and submitted

to an equilibration step of 1 ns. After that, a production run of 10 ns of NPT MD was performed at 310K and 1 atm with a Berendsen bath[25] with τ = 0.10 ps. The SETTLE algorithm[26] was used to constrain bond lengths and angles of water molecules, while the bond lengths and angles of peptides were constrained with the LINCS algorithm[27] which allowed the use of a 2 fs timestep. For the treatment of long-range interactions, we used the reaction field method, with a cut-off of 1.4 nm and a dielectric constant of ε = 54 (corresponding to SPC water). The van der Waals interactions were truncated with a twin-range cut-off of 0.8 and 1.4 nm. All simulations were run with the GROMACS 4.5.4 software package.[28]

Synthesis

2,4,6-Trinitrobenzenesulfonic acid (TNBS) test:[29] A sample of the resin was washed with dimethylformamide (DMF) 2 times. A few drops of DMF, two drops of a solution of 10% N,N-diisopropylethylamine (DIPEA) in DMF and two drops of 1% TNBS solution in DMF were added. After 5 minutes, the colour of the resin was observed (red is sign of the presence of free NH_2 groups).

Synthesis of dehydrodipeptide (1): The synthesis of Npx-L-Ala-Z-ΔPhe-OH (**1**) was described elsewhere.[3]

Synthesis of Npx-L-Ala-Z-ΔPhe-Gly-L-Arg(Pbf)-Gly-L-Asp(O^tBu)-Gly-OH (2): Fmoc-Gly-OH (1.20 *equiv* (resin), 0.50 g, 1.68 mmol) was dissolved in dry DCM (10 mL). DIPEA (4.00 *equiv* (Fmoc-Gly-OH), 1.16 mL, 6.72 mmol) and the resin (1.00 g) were added. The mixture was left stirring at room temperature for 6 hours. The resin was filtered and washed successively with a mixture of DCM/MeOH/DIPEA (17:2:1, 3×10 mL), DCM (3×10 mL), DMF (3×10 mL) and DCM (3×10 mL). The resin was left drying under reduced pressure overnight. The loading was measured by the absorbance of the dibenzofulvene-piperidine adducts at 290 nm (0.89 mmol g^{-1}). After washing the resin with DMF (2×10 mL), a solution of 20% piperidine in DMF (10 mL) was added. The mixture was left stirring at rt for 2 hours. The resin was filtered and washed successively with DMF (2×10 mL), 2-propanol (2×10 mL), DMF (2×10 mL) and 2-propanol (2×10 mL). The TNBS test was used to verify the cleavage. Fmoc-L-Asp(O^tBu)-OH (3.00 *equiv*) 1-hydroxybenzotriazole (HOBt) (3.00 *equiv*) and N,N'-diisopropylcarbodiimide (DIC) (3.00 *equiv*) were dissolved in DMF (10 mL). The solution was added to the resin and the mixture was left stirring at room temperature overnight. The resin was filtered and washed successively with DMF (3×10 mL) and DCM (3×10 mL). The coupling was verified by the TNBS test. The cleavage of the Fmoc group and coupling of the amino acids was repeated for Fmoc-Gly-OH (3.00 *equiv*), Fmoc-L-Arg(Pbf)-OH (3.00 *equiv*), Fmoc-Gly-OH (3.00 *equiv*) and peptide **1** (2.00 *equiv*). A mixture of AcOH/TFE/DCM (1:1:3, 20 mL) was added to the resin and it was left stirring at room temperature for 4 hours. The solution was filtered and the solvent removed under reduced pressure. Precipitation with diethyl ether afforded compound **2** (0.55 g, 51%) as a pale pink solid; ^1H NMR (400 MHz, DMSO-d_6, δ): 1.29 (d, J = 6.8 Hz, 3H, CH_3 Ala), 1.34 (s, 9H, 3×CH_3), 1.34-1.42 (m, 2H, γCH_2 Arg), 1.38 (s, 6H, 2×CH_3), 1.41 (d, J = 7.2 Hz, 3H, CH_3 Npx), 1.42-1.56 (m, 1H, βCH Arg), 1.61-1.70 (m, 1H, βCH Arg), 1.98 (s, 3H, CH_3 Pbf), 2.40-2.47 (m, 1H, βCH Asp), 2.41 (s, 3H, CH_3 Pbf), 2.47 (s, 3H, CH_3 Pbf), 2.65 (dd, J = 5.4 and 15.8 Hz, 1H, βCH Asp), 2.93 (s, 2H, CH_2 Pbf), 2.98-3.01 (m, 2H, δCH_2 Arg), 3.64-3.77 (m, 6H, 3×CH_2 Gly), 3.78-3.86 (m, 1H, CH Npx), 3.84 (s, 3H, OCH_3), 4.24-4.30 (m, 1H, αCH Arg), 4.37-4.41 (m, 1H, αCH Ala), 4.61-4.67 (m, 1H, αCH Asp), 6.59 (vbs, 2H, 2×NH Arg), 7.03 (brs, 1H, εNH Arg), 7.08-7.11 (m, 1H, Ar H Npx), 7.10 (s, 1H, βCH), 7.21-7.27 (m, 4H, Ar H Npx, 3×Ar H ΔPhe), 7.43 (dd, J = 1.6 and 8.8 Hz, 1H, Ar H Npx), 7.49-7.51 (m, 2H, Ar H ΔPhe), 7.68-7.74 (m, 3H, Ar H Npx), 7.87 (d, J = 7.6 Hz, 1H, αNH Arg), 8.00-8.02 (m, 1H, NH RGD*G*), 8.14-8.20 (m, 3H, NH Asp, NH *G*RGD, NH R*G*D), 8.36 (d, J = 6.0 Hz, 1H, NH

Ala), 9.72 (s, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 12.26 (CH$_3$ Pbf), 17.04 (CH$_3$ Ala), 17.59 (CH$_3$ Pbf), 18.94 (CH$_3$ Npx, CH$_3$ Pbf), 25.26 (γCH$_2$ Arg), 27.63 (3×CH$_3$), 28.29 (2×CH$_3$), 29.19 (βCH$_2$ Arg), 37.53 (βCH$_2$ Asp), 41.18 (CH$_2$ Gly), 41.98 (CH$_2$ Gly), 42.45 (CH$_2$ Pbf), 42.72 (CH$_2$ Gly), 43.74 (δCH$_2$ Arg), 44.41 (CH Npx), 48.87 (αCH Ala), 49.25 (αCH Asp), 52.24 (αCH Arg), 55.11 (OCH$_3$ Npx), 80.16 ($\underline{C}$(CH$_3$)$_3$), 86.27 (2-C Pbf), 105.65 (CH Npx), 116.25 (C Pbf), 118.44 (CH Npx), 124.31 (C Pbf), 125.42 (CH Npx), 126.49 (CH Npx), 126.60 (CH Npx), 128.32 (C Npx), 128.36 (αC), 128.41 (2×CH ΔPhe), 128.66 (CH ΔPhe), 128.77 (βCH ΔPhe), 129.06 (CH Npx), 129.54 (2×CH ΔPhe), 131.42 (C Pbf), 133.08 (C Npx), 133.72 (C$_i$ ΔPhe), 134.20 (C Pbf), 137.09 (C-2 Npx), 137.25 (C Pbf), 156.16 (C=N Arg), 156.93 (C-6 Npx), 157.42 (C Pbf), 164.93 (C=O ΔPhe), 168.60 (C=O R$\underline{G}$D), 168.80 (C=O $\underline{G}$RGD), 169.24 (C=O Asp), 170.39 (C=O Asp), 171.06 (C=O RGD$\underline{G}$), 171.72 (C=O Arg), 172.51 (C=O Ala), 173.98 (C=O Npx); HRMS (ESI) *m/z*: [M+H]$^+$ calcd for C$_{59}$H$_{77}$N$_{10}$O$_{15}$S$^+$ 1197.52851; found, 1197.52727.

Synthesis of Npx-L-Ala-Z-ΔPhe-Gly-L-Arg(Pbf)-Gly-L-Glu(O^{t}Bu)-Gly-OH: The synthesis was carried out as described for peptide **2**, substituting Fmoc-L-Asp(O^{t}Bu)-OH for Fmoc-L-Glu(O^{t}Bu)-OH. The loading of the resin was 0.72 mmol g^{-1}. The peptide (0.63 g, 72%) was obtained as a pale cream solid; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.29 (d, *J* = 6.8 Hz, 3H, CH$_3$ Ala), 1.35-1.42 (m, 2H, γCH$_2$ Arg), 1.36 (s, 9H, 3×CH$_3$), 1.38 (s, 6H, 2×CH$_3$), 1.41 (d, *J* = 7.2 Hz, 3H, CH$_3$ Npx), 1.43-1.51 (m, 1H, βCH Arg), 1.62-1.77 (m, 2H, βCH Arg and CH Glu), 1.82-1.94 (m, 1H, CH Glu), 1.98 (s, 3H, CH$_3$ Pbf), 2.20-2.25 (m, 2H, CH$_2$ Glu), 2.41 (s, 3H, CH$_3$ Pbf), 2.47 (s, 3H, CH$_3$ Pbf), 2.93 (s, 2H, CH$_2$ Pbf), 2.99-3.01 (m, 2H, δCH$_2$ Arg), 3.66-3.79 (m, 6H, 3×CH$_2$ Gly), 3.83-3.90 (m, 1H, CH Npx), 3.84 (s, 3H, OCH$_3$), 4.20-4.32 (m, 2H, αCH Arg and αCH Glu), 4.36-4.42 (m, 1H, αCH Ala), 6.47 (vbs, 2H, 2×NH), 6.83 (brs, 1H, εNH Arg), 7.08-7.11 (m, 1H, Ar H Npx), 7.10 (s, 1H, βCH), 7.21-7.27 (m, 4H, Ar H Npx, 3×Ar H ΔPhe), 7.43 (dd, *J* = 1.6 and 8.4 Hz, 1H, Ar H Npx), 7.50 (dd, *J* = 1.6 and 7.6 Hz, 2H, 2×Ar H ΔPhe), 7.67-7.70 (m, 1H, Ar H Npx), 7.68 (s, 1H, Ar H Npx), 7.70 (d, *J* = 2.4 Hz, 1H, Ar H Npx), 7.86 (d, *J* = 7.6 Hz, 1H, NH), 7.98 (d, *J* = 8.0 Hz, 1H, NH), 8.13-8.19 (m, 3H, 3×NH), 8.35 (d, *J* = 6.0 Hz, 1H, NH Arg), 9.72 (s, 1H, NH ΔPhe), 12.25 (brs, 1H, CO$_2$H); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 12.27 (CH$_3$ Pbf), 17.04 (CH$_3$ Ala), 17.60 (CH$_3$ Pbf), 18.95 (CH$_3$ Npx, CH$_3$ Pbf), 25.27 (γCH$_2$ Arg), 27.40 (CH$_2$ Glu), 27.73 (3×CH$_3$), 28.29 (2×CH$_3$), 29.16 (βCH$_2$ Arg), 31.06 (CH$_2$ Glu), 40.80 (CH$_2$ Gly), 41.90 (CH$_2$ Gly), 42.46 (CH$_2$ Pbf), 42.73 (CH$_2$ Gly), 43.76 (δCH$_2$ Arg), 44.42 (CH Npx), 48.88 (αCH Ala), 51.54 (αCH Glu), 52.26 (αCH Arg), 55.11 (OCH$_3$ Npx), 79.64 ($\underline{C}$(CH$_3$)$_3$), 86.28 (2-C Pbf), 105.65 (CH Npx), 116.26 (C Pbf), 118.45 (CH Npx), 124.32 (C Pbf), 125.44 (CH Npx), 126.50 (CH Npx), 126.60 (CH Npx), 128.33 (C Npx), 128.42 (3×CH ΔPhe), 128.68 (βCH ΔPhe), 128.79 (αC ΔPhe), 129.06 (CH Npx), 129.55 (2×CH ΔPhe), 131.44 (C Pbf), 133.09 (C Npx), 133.72 (C$_i$ ΔPhe), 134.18 (C Pbf), 137.08 (C Npx), 137.27 (C Pbf), 156.10 (C=N Arg), 156.94 (C Npx), 157.44 (C Pbf), 164.94 (C=O ΔPhe), 168.60 (C=O R$\underline{G}$E), 168.82 (C=O $\underline{G}$RGE), 171.12 (C=O RGE$\underline{G}$), 171.24 (C=O Glu), 171.68 (C=O Glu or Arg), 171.72 (C=O Glu or Arg), 172.54 (C=O Ala), 173.99 (C=O Npx); HRMS (ESI) *m/z*: [M+H]$^+$ calcd for C$_{60}$H$_{79}$N$_{10}$O$_{15}$S$^+$ 1211.54416; found, 1211.54322.

Synthesis of Npx-L-Ala-Z-ΔPhe-Gly-L-Arg-Gly-L-Asp-Gly-OH,TFA (3): TFA (6 mL mmol^{-1}) was added to Npx-L-Ala-Z-ΔPhe-Gly-L-Arg(Pbf)-Gly-L-Asp(O^{t}Bu)-Gly-OH (**2**) (0.54 g, 0.45 mmol) and the mixture was left stirring at room temperature for 5 hours. The solvent was removed under reduced pressure. Diethyl ether was added and the solvent removed again under reduced pressure. Precipitation with diethyl ether afforded compound **3** (0.40 g, 89%) as a beige solid; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.30 (d, *J* = 6.8 Hz, 3H, CH$_3$ Ala), 1.42 (d, *J* = 6.8 Hz, 3H, CH$_3$ Npx), 1.49-1.55 (m, 3H, CH$_2$ Arg), 1.68-1.77 (m, 1H, CH Arg), 2.48-2.49 (m, 1H, βCH Asp), 2.65-2.70 (m, 1H, βCH Asp), 3.06-3.08 (m, 2H, δCH$_2$ Arg), 3.69-3.76 (m, 6H, 3×CH$_2$ Gly), 3.82-3.85 (m, 1H, CH Npx), 3.84 (s,

3H, OCH$_3$), 3.97 (brs, 3H, NH$_3^+$), 4.29-4.32 (m, 1H, αCH Arg), 4.37-4.40 (m, 1H, αCH Ala), 4.59-4.64 (m, 1H, αCH Asp), 6.83 (brs, 1H, NH), 7.09-7.12 (m, 1H, Ar H Npx), 7.11 (s, 1H, βCH), 7.22-7.27 (m, 4H, Ar H Npx, 3×Ar H ΔPhe), 7.42-7.45 (m, 2H, Ar H Npx, εNH Arg), 7.49-7.51 (m, 2H, 2×Ar H ΔPhe), 7.68-7.72 (m, 3H, 3×Ar H Npx), 7.90 (d, *J* = 8.0 Hz, 1H, αNH Arg), 8.12 (t, *J* = 5.8 Hz, 1H, NH RGD$\underline{G}$), 8.16-8.22 (m, 3H, NH Asp, NH $\underline{G}$RGD, NH GR$\underline{G}$D), 8.38 (d, *J* = 6.4 Hz, 1H, NH Ala), 9.71 (s, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 17.03 (CH$_3$ Ala), 18.95 (CH$_3$ Npx), 24.87 (CH$_2$ Arg), 29.06 (CH$_2$ Arg), 36.33 (βCH$_2$ Asp), 40.40 (δCH$_2$ Arg), 40.81 (CH$_2$ Gly), 41.82 (CH$_2$ Gly), 42.73 (CH$_2$ Gly), 44.41 (CH Npx), 48.92 (αCH Ala), 49.30 (αCH Asp), 52.07 (αCH Arg), 55.13 (OCH$_3$ Npx), 105.66 (CH Npx), 118.47 (CH Npx), 125.44 (CH Npx), 126.49 (CH Npx), 126.63 (CH Npx), 127.38 (αC), 128.33 (C Npx), 128.45 (2×CH ΔPhe), 128.74 and 128.84 (CH ΔPhe and βCH ΔPhe), 129.08 (CH Npx), 129.55 (2×CH ΔPhe), 133.09 (C Npx), 133.68 (C$_i$ ΔPhe), 137.09 (C-2 Npx), 156.61 (C=N Arg), 156.95 (C-6 Npx), 164.95 (C=O ΔPhe), 168.62 (C=O R$\underline{G}$D), 168.83 (C=O $\underline{G}$RGD), 170.93 (C=O Asp and C=O RGD$\underline{G}$), 171.59 and 171.61 (C=O Arg and C=O Asp), 172.58 (C=O Ala), 174.08 (C=O Npx); HRMS (ESI) *m/z*: [M]$^+$ calcd for C$_{42}$H$_{53}$N$_{10}$O$_{12}^+$ 889.38389; found, 889.38563.

Synthesis of Npx-L-Ala-Z-ΔPhe-Gly-L-Arg-Gly-L-Glu-Gly-OH,TFA (4): The synthesis was carried out as described for peptide **3**. Npx-L-Ala-Z-ΔPhe-Gly-L-Arg(Pbf)-Gly-L-Glu(O^{t}Bu)-Gly-OH (0.54 g, 0.45 mmol) afforded peptide **4** (0.43 g, 96%) as a beije solid; ^{1}H NMR (400 MHz, DMSO-d_6, δ): 1.30 (d, *J* = 7.2 Hz, 3H, CH$_3$ Ala), 1.42 (d, *J* = 7.2 Hz, 3H, CH$_3$ Npx), 1.46-1.58 (m, 3H, γCH$_2$ Arg and βCH Arg), 1.69-1.79 (m, 2H, βCH Arg and βCH Glu), 1.86-1.96 (m, 1H, βCH Glu), 2.25 (t, *J* = 8.0 Hz, 2H, γCH$_2$ Glu), 3.05-3.09 (m, 2H, δCH$_2$ Arg), 3.62-3.81 (m, 6H, 3×CH$_2$ Gly), 3.73 (brs, 3H, NH$_3^+$), 3.82-3.88 (m, 1H, CH Npx), 3.84 (s, 3H, OCH$_3$ Npx), 4.26-4.32 (m, 2H, αCH Arg and αCH Glu), 4.39 (quint, *J* = 6.8 Hz, 1H, αCH Ala), 6.83 (brs, 1H, NH), 7.09-7.12 (m, 1H, Ar H Npx), 7.10 (s, 1H, βCH ΔPhe), 7.21-7.28 (m, 4H, Ar H Npx, 2×H$_m$ ΔPhe, H$_p$ ΔPhe), 7.41-7.45 (m, 2H, Ar H Npx and εNH Arg), 7.49-7.51 (m, 2H, 2×H$_o$ ΔPhe), 7.68-7.71 (m, 3H, 3×Ar H Npx), 7.91 (d, *J* = 8.0 Hz, 1H, αNH Arg), 7.99 (d, *J* = 8.0 Hz, 1H, NH Glu), 8.16-8.19 (m, 2H, NH $\underline{G}$RGE, NH R$\underline{G}$E), 8.25 (t, *J* = 6.0 Hz, 1H, NH RGE$\underline{G}$), 8.38 (d, *J* = 6.0 Hz, 1H, NH Ala), 9.72 (s, 1H, NH ΔPhe); ^{13}C NMR (100.6 MHz, DMSO-d_6, δ): 17.04 (CH$_3$ Ala), 18.95 (CH$_3$ Npx), 24.89 (γCH$_2$ Arg), 27.49 (βCH$_2$ Glu), 29.01 (βCH$_2$ Arg), 29.89 (γCH$_2$ Glu), 40.40 (δCH$_2$ Arg), 40.62 (CH$_2$ RGE$\underline{G}$), 41.84 (CH$_2$ Gly), 42.74 (CH$_2$ Gly), 44.42 (CH Npx), 48.92 (αCH Glu), 51.66 (αCH Glu), 52.13 (αCH Arg), 55.13 (OCH$_3$ Npx), 105.63 (CH Npx), 118.47 (CH Npx), 125.45 (CH Npx), 126.50 (CH Npx), 126.63 (CH Npx), 128.29 (αC ΔPhe), 128.34 (C Npx), 128.45 (C$_m$ ΔPhe), 128.74 (C$_m$ ΔPhe or βCH ΔPhe), 128.79 (C$_m$ ΔPhe or βCH ΔPhe), 129.08 (CH Npx), 129.56 (2×C$_o$ ΔPhe), 133.10 (C Npx), 133.69 (C$_i$ ΔPhe), 137.09 (C Npx), 156.61 (C=N Arg), 156.96 (C Npx), 164.97 (C=O ΔPhe), 168.58 (C=O R$\underline{G}$E), 168.86 (C=O $\underline{G}$RGE), 171.04 (C=O RGE$\underline{G}$), 171.43 (C=O Glu), 171.58 (C=O Arg), 172.60 (C=O Ala), 174.01 (C=O Glu or C=O Npx), 174.08 (C=O Glu or C=O Npx); HRMS (ESI) *m/z*: [M]$^+$ calcd for C$_{43}$H$_{55}$N$_{10}$O$_{12}^+$ 903.39954; found, 903.39822.

Neutralization of the peptides 3 and 4: Distilled water (16-21 mM) was added to peptides 3 and 4 and NaOH (1 M) was added till a suspension with pH ~7 was obtained. Sonication and heating to 40 °C was used to obtain a more uniform suspension. The solvent was removed under reduced pressure. Diethyl ether was added and the mixture taken to dryness under reduced pressure.

Npx-L-Ala-Z-ΔPhe-Gly-L-Arg-Gly-L-Asp-Gly-OH,CF$_3$CO$_2$Na: Npx-L-Ala-Z-ΔPhe-Gly-L-Arg-Gly-L-Asp-Gly-OH,TFA (**3**) (0.40 g, 0.40 mmol), water (25 mL) and NaOH (1 M) (2.76 *equiv*, 1.1 mL) gave neutralized peptide **3** (0.40 g, 98%) as a beije solid. MS (ESI) *m/z*: [M+Na]$^+$ calcd for C$_{42}$H$_{51}$N$_{10}$Na$_2$O$_{12}^+$ 933.35; found, 933.50.

Npx-L-Ala-Z-ΔPhe-Gly-L-Arg-Gly-L-Glu-Gly-OH,CF₃CO₂Na: Npx-L-Ala-Z-ΔPhe-Gly-L-Arg-Gly-L-Glu-Gly-OH,TFA (**4**) (0.43 g, 0.42 mmol), water (20 mL) and NaOH (1 M) (2.37 *equiv*, 1.0 mL) gave neutralized peptide **4** (0.35 g, 81%) as a beige solid. MS (ESI) *m/z*: [M+Na]$^+$ calcd for $C_{43}H_{53}N_{10}Na_2O_{12}^+$ 947.36; found, 947.50.

Acknowledgements

Keywords: RGD • hydrogel • cell culture • dehydropeptide

[1]   L. A. Estroff, A. D. Hamilton, *Chem. Rev.* **2004**, *104*, 1201-1218.
[2]   a) X. Zhao, F. Pan, H. Xu, M. Yaseen, H. Shan, C. A. E. Hauser, S. Zhang, J. R. Lu, *Chem. Soc. Rev.* **2010**, *39*, 3480-3498; b) Y. Zhang, H. Gu, Z. Yang, B. Xu, *J. Am. Chem. Soc.* **2003**, *125*, 13680-13681; c) D. M. Ryan, B. L. Nilsson, *Polymer Chemistry* **2012**, *3*, 18-33.
[3]   H. Vilaça, G. Pereira, T. G. Castro, B. F. Hermenegildo, J. Shi, T. Q. Faria, N. Micaelo, R. M. M. Brito, B. Xu, E. M. S. Castanheira, J. A. Martins, P. M. T. Ferreira, *J. Mater. Chem. B* **2015**, *3*, 6355-6367.
[4]   J. Thiele, Y. Ma, S. M. C. Bruekers, S. Ma, W. T. S. Huck, *Adv. Mater.* **2014**, *26*, 125-148.
[5]   V. Jayawarna, S. M. Richardson, A. R. Hirst, N. W. Hodson, A. Saiani, J. E. Gough, R. V. Ulijn, *Acta Biomater.* **2009**, *5*, 934-943.
[6]   a) T. C. Holmes, S. de Lacalle, X. Su, G. Liu, A. Rich, S. Zhang, *Proc. Natl. Acad. Sci.* **2000**, *97*, 6728-6733; b) J. E. Gough, A. Saiani, A. F. Miller, *Bioinspired, Biomimetic Nanobiomater.* **2011**, *1*, 4-12.
[7]   a) S. Maude, E. Ingham, A. Aggeli, *Nanomedicine* **2013**, *8*, 823-847; b) E. C. Wu, S. Zhang, C. A. E. Hauser, *Adv. Funct. Mater.* **2012**, *22*, 456-468; c) T. G. Kim, H. Shin, D. W. Lim, *Adv. Funct. Mater.* **2012**, *22*, 2446-2468.
[8]   a) M. Zhou, A. M. Smith, A. K. Das, N. W. Hodson, R. F. Collins, R. V. Ulijn, J. E. Gough, *Biomaterials* **2009**, *30*, 2523-2530; b) R. Orbach, I.

Mironi-Harpaz, L. Adler-Abramovich, E. Mossou, E. P. Mitchell, V. T. Forsyth, E. Gazit, D. Seliktar, *Langmuir* **2012**, *28*, 2015-2022.
[9]   L. Marinelli, A. Lavecchia, K.-E. Gottschalk, E. Novellino, H. Kessler, *J. Med. Chem.* **2003**, *46*, 4393-4404.
[10]   G. Cheng, V. Castelletto, R. R. Jones, C. J. Connon, I. W. Hamley, *Soft Matter* **2011**, *7*, 1326-1333.
[11]   a) R. Orbach, L. Adler-Abramovich, S. Zigerson, I. Mironi-Harpaz, D. Seliktar, E. Gazit, *Biomacromolecules* **2009**, *10*, 2646-2651; b) Y. Wang, Z. Zhang, L. Xu, X. Li, H. Chen, *Colloids Surf., B* **2013**, *104*, 163-168.
[12]   S. Wenzel, V. Buss, *J. Phys. Org. Chem.* **1992**, *5*, 748-754.
[13]   K. Nakanishi, N. Berova, R. W. Woody in *Circular Dichroism Principles and Applications*, VCH Publishers, Inc., New York, **1994**.
[14]   Z. Yang, G. Liang, M. Ma, Y. Gao, B. Xu, *J. Mater. Chem.* **2007**, *17*, 850-854.
[15]   a) I. W. Hamley, G. D. Brown, V. Castelletto, G. Cheng, M. Venanzi, M. Caruso, E. Placidi, C. Aleman, G. Revilla-López, D. Zanuy, *J. Phys. Chem. B* **2010**, *114*, 10674-10683; b) V. Castelletto, D. R. Nutt, I. W. Hamley, S. Bucak, C. Cenker, U. Olsson, *Chem. Commun.* **2010**, *46*, 6270-6272; c) O. Rathore, D. Y. Sogah, *J. Am. Chem. Soc.* **2001**, *123*, 5231-5239.
[16]   I. W. Hamley, *Angew. Chem. Int. Ed.* **2007**, *46*, 8128-8147.
[17]   F. Gelain, D. Silva, A. Caprini, F. Taraballi, A. Natalello, O. Villa, K. T. Nam, R. N. Zuckermann, S. M. Doglia, A. Vescovi, *ACS Nano* **2011**, *5*, 1845-1859.
[18]   A. D. Martin, A. B. Robinson, A. F. Mason, J. P. Wojciechowski, P. Thordarson, *Chem. Commun.* **2014**, *50*, 15541-15544.
[19]   A. E. Torda, W. F. van Gunsteren, *J. Comput. Chem.* **1994**, *15*, 1331-1340.
[20]   X. Li, X. Du, Y. Gao, J. Shi, Y. Kuang, B. Xu, *Soft Matter* **2012**, *8*, 7402-7407.
[21]   O. Sumner Makin, P. Sikorski, L. C. Serpell, *J. Appl. Crystallogr.* **2007**, *40*, 966-972.
[22]   Schrödinger, in *The PyMOL Molecular Graphics System*, 1.3r1 ed. ed., LLC, **2010**.
[23]   a) W. Huang, Z. X. Lin, W. F. van Gunsteren, *J. Chem. Theory Comput.* **2011**, *7*, 1237-1243; b) N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, W. F. van Gunsteren, *Eur. Biophys. J. Biophys. Lett.* **2011**, *40*, 843-856.
[24]   H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* **1987**, *91*, 6269-6271.
[25]   H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola, J. R. Haak, *The Journal of Chemical Physics* **1984**, *81*, 3684-3690.
[26]   D. van der Spoel, P. J. van Maaren, H. J. C. Berendsen, *J. Chem. Phys.* **1998**, *108*, 10220-10230.
[27]   B. Hess, H. Bekker, H. J. C. Berendsen, J. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463-1472.
[28]   a) B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *Journal of Chemical Theory and Computation* **2008**, *4*, 435-447; b) D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, P. Tieleman, A. L. T. M. Sjibers, K. A. Feenstra, R. van Drunen, H. J. C. Berendsen, **2010**.
[29]   W. S. Hancock, J. E. Battersby, *Anal. Biochem.* **1976**, *71*, 260-264.

WILEY-VCH

Entry for the Table of Contents

FULL PAPER
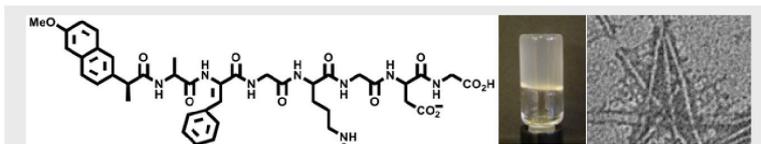
A dehydropeptide containing the RGD sequence is obtained with good yield using a combined solid phase and solution phase synthesis. The peptide gives hydrogels under physiological conditions and low critical gelation concentrations, thus presenting possible biomedical applications.

Dehydropeptide hydrogels

Helena Vilaça, Tarsila G. Castro, Loly Torres Pérez, Ashkan Dehsorkhi, Cristóvão F. Lima, Catarina Gonçalves, Manuel Melle-Franco, Loic Hilliou, Miguel Gama, Ian W. Hamley, José A. Martins, Paula M. T. Ferreira

Page No. – Page No.

Self-healing RGD dehydropeptide hydrogel

CONCLUSIONS

Chapter VIII

Final Remarks

Conclusions and Final Remarks

The Molecular Dynamics Simulations performed in the systems addressed in this thesis, answered many questions placed as hypotheses at the beginning of this work. Our findings demonstrated that only a few representatives of each class of non-canonical amino acids studied correspond to well-constrained structures able to act as foldamers.

In this work we focused on five classes of unnatural amino acids, namely: symmetrical α,α-dialkyl glycines, asymmetrical α,α-dialkyl glycines, proline analogues, $C\alpha$ to $C\alpha$ cyclized amino acids and α,β-dehydroamino acids. We opted for introducing and following these classes of amino acids in peptides with well-known secondary structure and biological function, or by suggesting novel constrained peptides with potential applicability in medicine and other fields.

The symmetrical α,α-dialkyl glycines were studied in two peptaibols, one shorter and one longer, with antimicrobial properties: Peptaibolin and Alamethicin. The residues Dhg (α,α-dihexyl glycine) and Ac_6c (α,α-cyclohexyl glycine) proved to be α-helical inducers in these systems, promoting pre-organization in water. In the case of Peptaibolin the presence of the new amino acids were found to explicitly help the insertion in POPC membranes.

The asymmetrical α,α-dialkyl glycines such as Iva (isovaline), and the proline analogs like Hyp (hydroxyproline) were studied in Antiamoebin and Zervamicin peptaibols. In these studies, two amino acid analogs of Iva were found to induce improved helical secondary structure, namely MDL (α-methyl-D-leucine) and MDP (α-methyl-D-phenylalanine), which may be linked to the antibiotic activity of these peptaibols. Regarding proline analogs, the analog ALP (cis-3-amino-L-proline) proved to generate improved helical content in both peptaibols, which was unexpected as proline and many proline derivatives induce bends in helical secondary structures.

Simulations on peptides carrying the α,β-dehydroamino acids Δ^{Z}Phe and Δ^{Z}Abu were performed to understand experimental results obtained by co-workers. These peptides self-assemble in hydrogels which improves their range of application, for instance, in drug delivery.

$C\alpha$ to $C\alpha$ cyclized amino acids (Ac_nc residues) can also be considered α,α-dialkyl glycines, due the double substitution at the $C\alpha$ atom. As previously mentioned, the Ac_6c amino acid stands as a good helical folder. In addition, this residue together with other residues of this class, namely: Ac_3c, Ac_4c and Ac_5c, were studied in hexa and nonapeptides in three different solvents. Preliminary results indicate that these residues behave very differently according to the environment, yet more simulations are necessary to fully confirm this finding (Appendix VI).

167

This thesis addresses systematically how the insertion of one of more unnatural, non-canonical, amino acids may affect the structure and, through the structure, the function of peptides relevant for medicinal chemistry. The findings obtained show that certain unnatural amino acids stabilize conformations with well-defined secondary structure. More generally, this work shows how in-silico experiments can be a valuable tool for the rational design of biomolecules with improved properties.

APPENDIX

APPENDIX I

Chapter III – Supporting Information

Modeling of Peptaibol Analogs Incorporating Non-Polar α,α-Dialkyl Glycines Shows Improved α-Helical Pre-Organization and Spontaneous Membrane Permeation

Tarsila Gabriel Castro and Nuno Miguel Micaêlo

This section presents the topologies for the new α,α-dialkylglycines under study. These topologies were developed based on the natural amino acids parameterized in de GROMOS 54a7 force field.

Aib

Non-bonded parameters

Atom name	Atom type	Charge (q)	
N	N	-0.31	
H	H	0.31	
CA	C	0.00	
CB1	CH_3	0.00	
CB2	CH_3	0.00	
C	C	0.45	
O	O	-0.45	

Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21
	CA	CB1	gb_27
	CA	CB2	gb_27
	CA	C	gb_27
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	CB1	ga_13
	N	CA	C	ga_19
	CB1	CA	C	ga_13
	N	CA	CB2	ga_13
	CB1	CA	CB2	ga_13
	CB2	CA	C	ga_13
	CA	C	O	ga_30
	CA	C	+N	ga_19
	O	C	+N	ga_33

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45

Impropers dihedrals	ai	aj	ak	al	Gromos improper type
	N	-C	CA	H	gi_1
	CA	N	C	CB1	gi_2
	C	CA	+N	O	gi_1
	CA	N	CB2	C	gi_2
	CA	N	CB1	CB2	gi_2

Deg

Non-bonded parameters

Atom name	Atom type	Charge (q)	
N	N	-0.31	
H	H	0.31	
CA	C	0.00	
C1	CH_2	0.00	
C2	CH_3	0.00	
C3	CH_2	0.00	
C4	CH_3	0.00	
C	C	0.45	
O	O	-0.45	
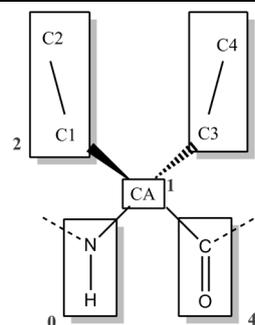
Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21
	CA	C1	gb_27
	CA	C3	gb_27
	C1	C2	gb_27
	C3	C4	gb_27
	CA	C	gb_27
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	C1	ga_13
	N	CA	C	ga_19
	N	CA	C3	ga_13
	C1	CA	C	ga_13
	C	CA	C3	ga_13
	CA	C1	C2	ga_13
	CA	C3	C4	ga_13
	C1	CA	C3	ga_13
	CA	C	O	ga_30
	CA	C	+N	ga_19
	O	C	+N	ga_33

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	C	CA	C1	C2	gd_34
	C	CA	C3	C4	gd_34

174

	ai	aj	ak	al	Gromos improper type
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
Impropers dihedrals	ai	aj	ak	al	Gromos improper type
	N	-C	CA	H	gi_1
	CA	N	C	C1	gi_2
	C	CA	+N	O	gi_1
	CA	N	C3	C	gi_2
	CA	N	C1	C3	gi_2

Dpg

Non-bonded parameters

Atom name	Atom type	Charge (q)	
N	N	-0.31	
H	H	0.31	
CA	C	0.00	
C1	CH$_2$	0.00	
C2	CH$_2$	0.00	
C3	CH$_3$	0.00	
C4	CH$_2$	0.00	
C5	CH$_2$	0.00	
C6	CH$_3$	0.00	
C	C	0.45	
O	O	-0.45	
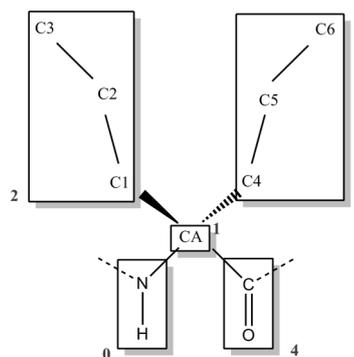
Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21
	CA	C1	gb_27
	CA	C4	gb_27
	C1	C2	gb_27
	C2	C3	gb_27
	C4	C5	gb_27
	C5	C6	gb_27
	CA	C	gb_27
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angles type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	C	CA	C4	ga_13
	C	CA	C1	ga_13
	C1	CA	C4	ga_13
	N	CA	C1	ga_13
	N	CA	C4	ga_13
	N	CA	C	ga_19
	CA	C1	C2	ga_13
	CA	C4	C5	ga_13
	C1	C2	C3	ga_13
	C4	C5	C6	ga_13
	CA	C	O	ga_30
	CA	C	+N	ga_19

175

	ai	aj	ak	al	
	O	C	+N		ga_33
Propers dihedrals	**ai**	**aj**	**ak**	**al**	**Gromos dihedral type**
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	C	CA	C1	C2	gd_34
	C	CA	C4	C5	gd_34
	CA	C4	C5	C6	gd_34
	CA	C1	C2	C3	gd_34
Impropers dihedrals	**ai**	**aj**	**ak**	**al**	**Gromos improper type**
	N	-C	CA	H	gi_1
	CA	N	C	C4	gi_2
	C	CA	+N	O	gi_1
	CA	N	C1	C	gi_2
	CA	N	C4	C1	gi_2

Dibg
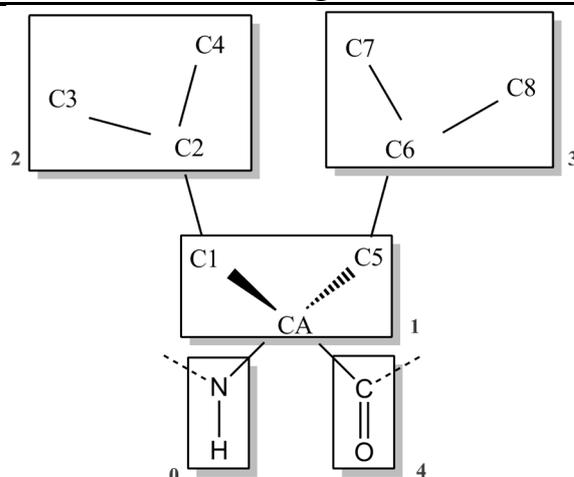
Non-bonded parameters

Atom name	Atom type	Charge (q)
N	N	-0.31
H	H	0.31
CA	C	0.00
C1	CH_2	0.00
C2	CH	0.00
C3	CH_3	0.00
C4	CH_3	0.00
C5	CH_2	0.00
C6	CH	0.00
C7	CH_3	0.00
C8	CH_3	0.00
C	C	0.45
O	O	-0.45

Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21

176

	CA	C1	gb_27
	CA	C5	gb_27
	C1	C2	gb_27
	C2	C3	gb_27
	C2	C4	gb_27
	C5	C6	gb_27
	C6	C7	gb_27
	C6	C8	gb_27
	CA	C	gb_27
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	C1	ga_13
	C	CA	C1	ga_13
	CA	C1	C2	ga_13
	CA	C5	C6	ga_13
	C1	C2	C3	ga_15
	C1	C2	C4	ga_15
	C3	C2	C4	ga_15
	C5	C6	C7	ga_15
	C5	C6	C8	ga_15
	C7	C6	C8	ga_15
	C5	CA	C	ga_13
	C1	CA	C5	ga_13
	N	CA	C5	ga_13
	N	CA	C	ga_19
	CA	C	O	ga_30
	CA	C	+N	ga_19
	O	C	+N	ga_33

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	N	CA	C1	C2	gd_34
	N	CA	C5	C6	gd_34
	CA	C5	C6	C7	gd_34
	CA	C1	C2	C3	gd_34

Impropers dihedrals	ai	aj	ak	al	Gromos improper type
	N	-C	CA	H	gi_1
	CA	N	C5	C1	gi_2
	CA	N	C1	C	gi_2
	CA	N	C	C5	gi_2
	C6	C7	C5	C8	gi_2
	C2	C4	C3	C1	gi_2
	C	CA	+N	O	gi_1

Dhg
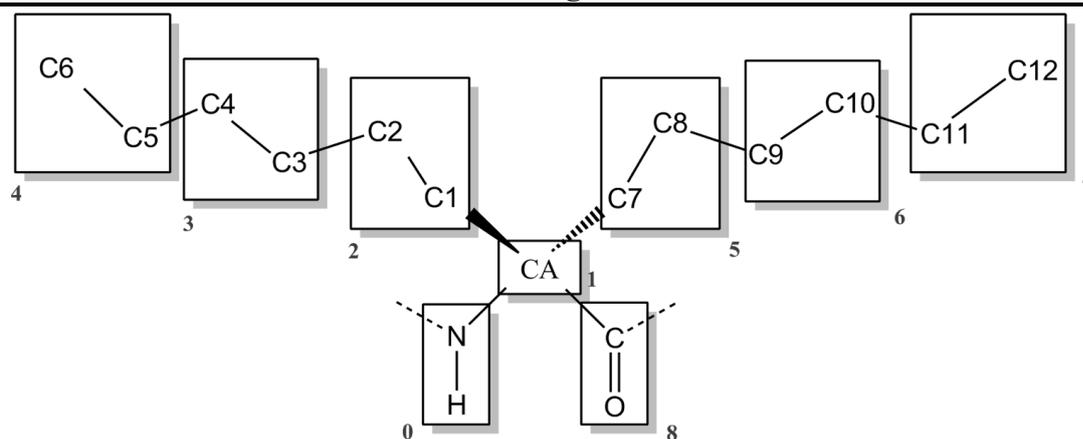
Non-bonded parameters

Atom name	Atom type	Charge (q)
N	N	-0.31
H	H	0.31
CA	C	0.00
C1	CH_2	0.00
C2	CH_2	0.00
C3	CH_2	0.00
C4	CH_2	0.00
C5	CH_2	0.00
C6	CH_3	0.00
C7	CH_2	0.00
C8	CH_2	0.00
C9	CH_2	0.00
C10	CH_2	0.00
C11	CH_2	0.00
C12	CH_3	0.00
C	C	0.45
O	O	-0.45

Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21
	CA	C1	gb_27
	CA	C7	gb_27
	C1	C2	gb_27
	C2	C3	gb_27
	C3	C4	gb_27
	C4	C5	gb_27
	C5	C6	gb_27
	C7	C8	gb_27
	C8	C9	gb_27
	C9	C10	gb_27
	C10	C11	gb_27
	C11	C12	gb_27
	CA	C	gb_27
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32

-C	N	CA		ga_31
H	N	CA		ga_18
C	CA	C1		ga_13
C	CA	C7		ga_13
C1	CA	C7		ga_13
CA	C1	C2		ga_13
CA	C7	C8		ga_13
C1	C2	C3		ga_13
C2	C3	C4		ga_13
C3	C4	C5		ga_13
C4	C5	C6		ga_13
C7	C8	C9		ga_13
C8	C9	C10		ga_13
C9	C10	C11		ga_13
C10	C11	C12		ga_13
N	CA	C1		ga_13
N	CA	C7		ga_13
N	CA	C		ga_19
CA	C	O		ga_30
CA	C	+N		ga_19
O	C	+N		ga_33

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	C	CA	C1	C2	gd_34
	C	CA	C7	C8	gd_34
	CA	C1	C2	C3	gd_34
	C1	C2	C3	C4	gd_34
	C2	C3	C4	C5	gd_34
	C3	C4	C5	C6	gd_34
	CA	C7	C8	C9	gd_34
	C7	C8	C9	C10	gd_34
	C8	C9	C10	C11	gd_34
	C9	C10	C11	C12	gd_34

Impropers dihedrals	ai	aj	ak	al	Gromos improper type
	N	-C	CA	H	gi_1
	CA	N	C	C7	gi_2
	CA	N	C7	C1	gi_2
	CA	N	C1	C	gi_2
	C	CA	+N	O	gi_1

Dϕg
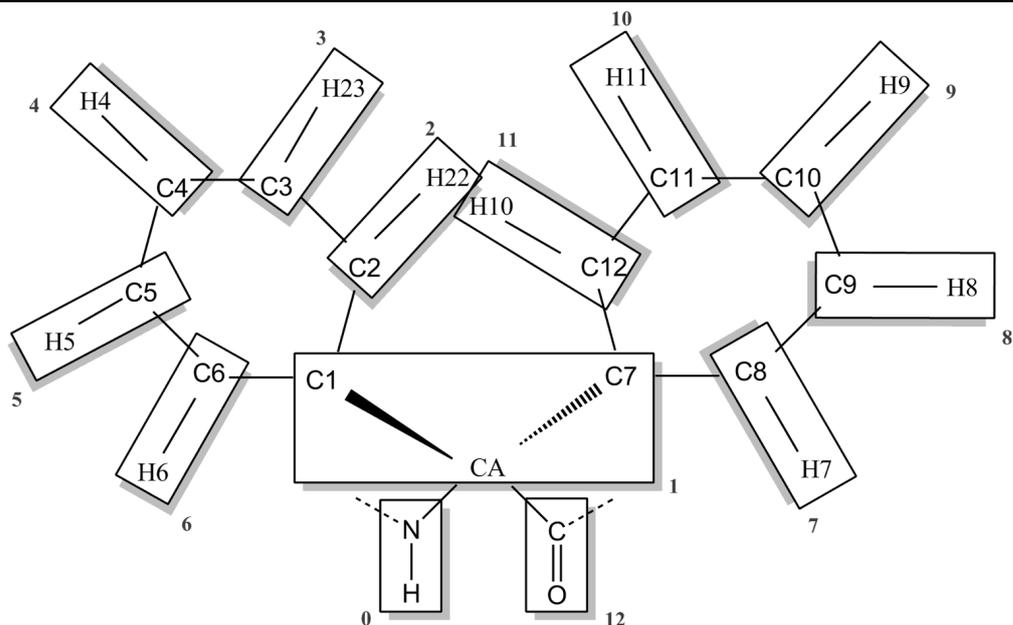
Non-bonded parameters

Atom name	Atom type	Charge (q)
N	N	-0.31
H	H	0.31
CA	C	0.00
C1	C	0.00
C7	C	0.00
C2	C	-0.14
H22	HC	0.14
C3	C	-0.14
H23	HC	0.14
C4	C	-0.14
H4	HC	0.14
C5	C	-0.14
H5	HC	0.14
C6	C	-0.14
H6	HC	0.14
C8	C	-0.14
H7	HC	0.14
C9	C	-0.14
H8	HC	0.14
C10	C	-0.14
H9	HC	0.14
C11	C	-0.14
H11	HC	0.14
C12	C	-0.14
H10	HC	0.14
C	C	0.45
O	O	-0.45

Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2

N	CA	gb_21
CA	C1	gb_27
CA	C7	gb_27
CA	C	gb_27
C1	C2	gb_16
C2	C3	gb_16
C2	H22	gb_3
C3	C4	gb_16
C3	H23	gb_3
C4	C5	gb_16
C4	H4	gb_3
C5	C6	gb_16
C5	H5	gb_3
C6	C1	gb_16
C6	H6	gb_3
C7	C8	gb_16
C8	C9	gb_16
C8	H7	gb_3
C9	C10	gb_16
C9	H8	gb_3
C10	C11	gb_16
C10	H9	gb_3
C11	C12	gb_16
C11	H11	gb_3
C12	C7	gb_16
C12	H10	gb_3
C	O	gb_5
C	+N	gb_10

Exclusions	ai	aj
	C1	H23
	C1	H5
	C1	C4
	C2	C5
	C2	H4
	C2	H6
	C3	C6
	C3	H5
	C4	H22
	C4	H6
	C5	H23
	C6	H22
	C6	H4
	H22	H23
	H23	H4
	H4	H5
	H5	H6
	C7	H8
	C7	H11
	C7	C10
	C8	H10
	C8	H9
	C8	C11
	C9	C12
	C9	H11
	C10	H10
	C10	H7
	C11	H8

	ai	aj	ak	Gromos angle type
	C12			H9
	C12			H7
	H7			H8
	H8			H9
	H9			H11
	H10			H11

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	C1	ga_13
	N	CA	C7	ga_13
	N	CA	C	ga_19
	C7	CA	C	ga_13
	C1	CA	C	ga_13
	C1	CA	C7	ga_13
	CA	C7	C8	ga_15
	CA	C7	C12	ga_15
	CA	C1	C2	ga_15
	CA	C1	C6	ga_15
	C1	C2	H22	ga_25
	C1	C6	H6	ga_25
	C1	C2	C3	ga_27
	C1	C6	C5	ga_27
	C2	C1	C6	ga_27
	H22	C2	C3	ga_25
	H6	C6	C5	ga_25
	C2	C3	C4	ga_27
	C2	C3	H23	ga_25
	C6	C5	C4	ga_27
	C6	C5	H5	ga_25
	C3	C4	C5	ga_27
	H23	C3	C4	ga_25
	H5	C5	C4	ga_25
	H4	C4	C5	ga_25
	H4	C4	C3	ga_25
	C7	C8	H7	ga_25
	C7	C12	H10	ga_25
	C7	C6	C9	ga_27
	C7	C12	C11	ga_27
	C7	C8	C9	ga_27
	C8	C7	C12	ga_27
	H7	C8	C9	ga_25
	H10	C12	C11	ga_25
	C8	C9	C10	ga_27
	C8	C9	H8	ga_25
	C12	C11	C10	ga_27
	C12	C11	H11	ga_25
	C9	C10	C11	ga_27
	H8	C9	C10	ga_25
	H11	C11	C10	ga_25
	H9	C10	C9	ga_25
	H9	C10	C11	ga_25
	CA	C	O	ga_30
	CA	C	+N	ga_19
	O	C	+N	ga_33

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	C	CA	C1	C6	gd_34
	C	CA	C7	C8	gd_34
	CA	C1	C2	C3	gd_34
	CA	C7	C8	C9	gd_34
	C3	C4	C5	C6	gd_34
	C9	C10	C11	C12	gd_34
Impropers dihedrals	**ai**	**aj**	**ak**	**al**	**Gromos improper type**
	N	-C	CA	H	gi_1
	CA	N	C	C7	gi_2
	CA	N	C7	C1	gi_2
	CA	N	C1	C	gi_2
	C1	C2	C6	CA	gi_1
	C7	C8	C12	CA	gi_1
	C1	C2	C3	C4	gi_1
	C1	C6	C5	C4	gi_1
	C7	C8	C9	C10	gi_1
	C7	C12	C11	C10	gi_1
	C2	C1	C6	C5	gi_1
	C2	C3	C4	C5	gi_1
	C8	C7	C12	C11	gi_1
	C8	C9	C10	C11	gi_1
	H22	C1	C3	C2	gi_1
	H23	C2	C4	C3	gi_1
	H4	C3	C5	C4	gi_1
	H5	C4	C6	C5	gi_1
	H6	C1	C5	C6	gi_1
	H7	C7	C9	C8	gi_1
	H8	C8	C10	C9	gi_1
	H9	C9	C11	C10	gi_1
	H11	C10	C12	C11	gi_1
	H10	C11	C7	C12	gi_1
	C	CA	+N	O	gi_1

Db$_z$g
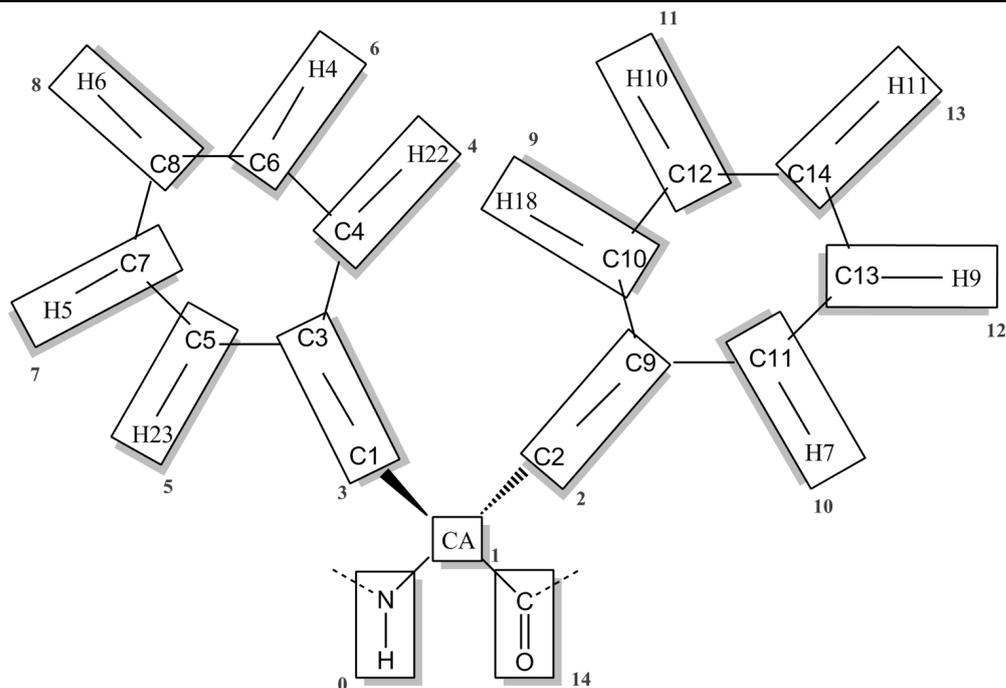
Non-bonded parameters

Atom name	Atom type	Charge (q)
N	N	-0.31
H	H	0.31
CA	C	0.00
C2	CH2	0.00
C9	C	0.00
C1	CH2	0.00
C3	C	0.00
C4	C	-0.14
H22	HC	0.14
C5	C	-0.14
H23	HC	0.14
C6	C	-0.14
H4	HC	0.14
C7	C	-0.14
H5	HC	0.14
C8	C	-0.14
H6	HC	0.14
C10	C	-0.14
H8	HC	0.14
C11	C	-0.14
H7	HC	0.14
C12	C	-0.14
H10	HC	0.14
C13	C	-0.14
H9	HC	0.14
C14	C	-0.14
H11	HC	0.14
C	C	0.45
O	O	-0.45

184

Bonded parameters			
Bonds	**ai**	**aj**	**Gromos bond type**
	N	H	gb_2
	N	CA	gb_21
	CA	C1	gb_27
	CA	C2	gb_27
	CA	C	gb_27
	C1	C3	gb_27
	C3	C4	gb_16
	C3	C5	gb_16
	C4	H22	gb_3
	C4	C6	gb_16
	C5	H23	gb_3
	C5	C7	gb_16
	C6	H4	gb_3
	C6	C8	gb_16
	C7	H5	gb_3
	C7	C8	gb_16
	C8	H6	gb_3
	C2	C9	gb_27
	C9	C10	gb_16
	C9	C11	gb_16
	C10	H8	gb_3
	C10	C12	gb_16
	C11	H7	gb_3
	C11	C13	gb_16
	C12	H10	gb_3
	C12	C14	gb_16
	C13	H9	gb_3
	C13	C14	gb_16
	C14	H11	gb_3
	C	O	gb_5
	C	+N	gb_10
Exclusions	**ai**		**aj**
	C1		H22
	C1		H23
	C1		C6
	C1		C7
	C3		C8
	C3		H4
	C3		H5
	C4		C7
	C4		H6
	C4		H23
	C5		C6
	C5		H6
	C5		H22
	C6		H5
	C8		H22
	C8		H23
	C7		H4
	H6		H4
	H6		H5
	H5		H23
	C2		H7
	C2		H8

C2	C13
C2	C12
C9	C14
C9	H10
C9	H9
C10	C13
C10	H11
C10	H7
C11	C12
C11	H11
C11	H8
C12	H9
C13	H10
H10	H8
H9	H7
C14	H7
C14	H8
H11	H9
H11	H10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	C1	ga_13
	N	CA	C2	ga_13
	N	CA	C	ga_19
	C2	CA	C	ga_13
	C1	CA	C	ga_13
	C1	CA	C2	ga_13
	CA	C2	C9	ga_15
	CA	C1	C3	ga_15
	C2	C9	C11	ga_27
	C2	C9	C10	ga_27
	C1	C3	C4	ga_27
	C1	C3	C5	ga_27
	C11	C9	C10	ga_27
	C4	C3	C5	ga_27
	C3	C5	H23	ga_25
	C9	C11	H7	ga_25
	C3	C5	C7	ga_27
	C9	C11	C13	ga_27
	H23	C5	C7	ga_25
	H7	C11	C13	ga_25
	C3	C4	H22	ga_25
	C9	C10	H8	ga_25
	C3	C4	C6	ga_27
	C9	C10	C12	ga_27
	H22	C4	C6	ga_25
	H8	C10	C12	ga_25
	C5	C7	H5	ga_25
	C11	C13	H9	ga_25
	C5	C7	C8	ga_27
	C11	C13	C14	ga_27
	H5	C7	C8	ga_25
	H9	C13	C14	ga_25
	C4	C6	H4	ga_25
	C10	C12	H10	ga_25

186

C4	C6	C8	ga_27	
C10	C12	C14	ga_27	
H4	C6	C8	ga_25	
H10	C12	C14	ga_25	
C7	C8	C6	ga_27	
C13	C14	C12	ga_27	
C7	C8	H6	ga_25	
C13	C14	H11	ga_25	
C6	C8	H6	ga_25	
C12	C14	H11	ga_25	
CA	C	O	ga_30	
CA	C	+N	ga_19	
O	C	+N	ga_33	

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	C	CA	C1	C3	gd_34
	C	CA	C2	C9	gd_34
	CA	C1	C3	C5	gd_40
	CA	C2	C9	C11	gd_40

Impropers dihedrals	ai	aj	ak	al	Gromos improper type
	N	-C	CA	H	gi_1
	CA	N	C	C2	gi_2
	CA	N	C2	C1	gi_2
	CA	N	C1	C	gi_2
	C3	C4	C5	C1	gi_1
	C9	C11	C10	C2	gi_1
	C3	C5	C7	C8	gi_1
	C9	C11	C13	C14	gi_1
	C3	C5	C6	C8	gi_1
	C9	C11	C12	C14	gi_1
	C5	C3	C4	C6	gi_1
	C11	C9	C10	C12	gi_1
	C5	C3	C7	H23	gi_1
	C11	C9	C13	H7	gi_1
	C5	C7	C8	C6	gi_1
	C11	C13	C14	C12	gi_1
	C4	C3	C5	C7	gi_1
	C10	C9	C11	C13	gi_1
	C4	C3	C6	H22	gi_1
	C10	C9	C12	H8	gi_1
	C4	C6	C8	C7	gi_1
	C10	C12	C14	C13	gi_1
	H5	C5	C8	C7	gi_1
	H9	C11	C14	C13	gi_1
	H4	C4	C8	C6	gi_1
	H10	C10	C14	C12	gi_1
	C8	C7	C6	H6	gi_1
	C14	C13	C12	H11	gi_1
	C	CA	+N	O	gi_1

Ac$_6$c

Non-bonded parameters

Atom name	Atom type	Charge (q)
N	N	-0.31
H	H	0.31
CA	C	0.00
C1	CH2R	0.00
C2	CH2R	0.00
C3	CH2R	0.00
C4	CH2R	0.00
C5	CH2R	0.00
C	C	0.45
O	O	-0.45
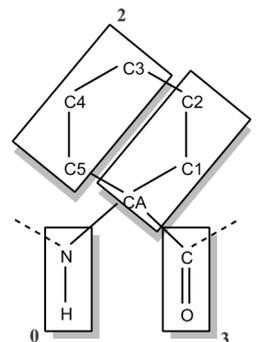
Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21
	CA	C1	gb_27
	CA	C5	gb_27
	C1	C2	gb_27
	C2	C3	gb_27
	C3	C4	gb_27
	C4	C5	gb_27
	CA	C	gb_27
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	C1	ga_13
	N	CA	C5	ga_13
	N	CA	C	ga_13
	C1	CA	C	ga_13
	C5	CA	C	ga_13
	C1	CA	C5	ga_13
	CA	C1	C2	ga_13
	CA	C5	C4	ga_13
	C1	C2	C3	ga_13
	C2	C3	C4	ga_13
	C3	C4	C5	ga_13
	CA	C	O	ga_30
	CA	C	+N	ga_19
	O	C	+N	ga_33

Propers dihedrals	ai	aj	ak	al	Gromos dihedral type
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	N	CA	C5	C4	gd_34
	N	CA	C1	C2	gd_34
	H	N	CA	C1	gd_34

	ai	aj	ak	al	Gromos improper type
	CA	C1	C2	C3	gd_34
	CA	C5	C4	C3	gd_34
	C1	C2	C3	C4	gd_34
	C2	C3	C4	C5	gd_34
	CA	C1	C2	C3	gd_34
Impropers dihedrals	**ai**	**aj**	**ak**	**al**	**Gromos improper type**
	N	-C	CA	H	gi_1
	C	CA	+N	O	gi_1
	CA	N	C1	C	gi_2
	CA	N	C5	C1	gi_2
	CA	N	C	C5	gi_2

Dmg

Non-bonded parameters

Atom name	Atom type	Charge (q)	
N	N	-0.310	
H	H	0.310	
CA	C	0.000	
CB	CH2	0.266	
OB	OA	-0.674	
HB	H	0.408	
CG	CH2	0.266	
OG	OA	-0.674	
HG	H	0.408	
C	C	0.45	
O	O	-0.45	
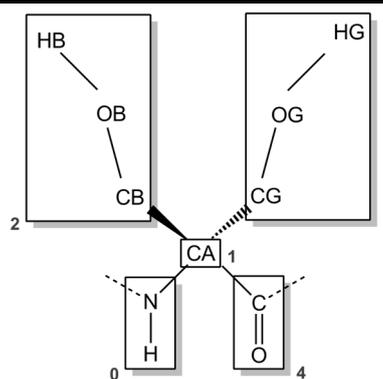
Bonded parameters

Bonds	ai	aj	Gromos bond type
	N	H	gb_2
	N	CA	gb_21
	CA	CB	gb_27
	CA	CG	gb_27
	CA	C	gb_27
	CB	OB	gb_18
	CG	OG	gb_18
	OB	HB	gb_1
	OG	HG	gb_1
	C	O	gb_5
	C	+N	gb_10

Angles	ai	aj	ak	Gromos angle type
	-C	N	H	ga_32
	-C	N	CA	ga_31
	H	N	CA	ga_18
	N	CA	CB	ga_13
	N	CA	CG	ga_13
	N	CA	C	ga_13
	CG	CA	C	ga_13
	CB	CA	C	ga_13
	CA	CB	OB	ga_13
	CB	OB	HB	ga_12
	CA	CG	OG	ga_13
	CG	OG	HG	ga_12
	CB	CA	CG	ga_13
	CA	C	O	ga_30
	CA	C	+N	ga_19

	ai	aj	ak	al	
	O	C	+N		ga_33
Propers dihedrals	**ai**	**aj**	**ak**	**al**	**Gromos dihedral type**
	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_42
	-C	N	CA	C	gd_43
	N	CA	C	+N	gd_44
	N	CA	C	+N	gd_45
	N	CA	CB	OB	gd_34
	N	CA	CG	OG	gd_34
	CA	CB	OB	HB	gd_23
	CA	CG	OG	HG	gd_23
Impropers dihedrals	**ai**	**aj**	**ak**	**al**	**Gromos improper type**
	N	-C	CA	H	gi_1
	CA	N	C	CB	gi_2
	C	CA	+N	O	gi_1
	CA	N	C	CG	gi_2

Figure S1. Central Structures

Peptaibolin at A: t=0-10ns, B: t=40-50ns, C: t=50-60ns and D: t=90-100ns.
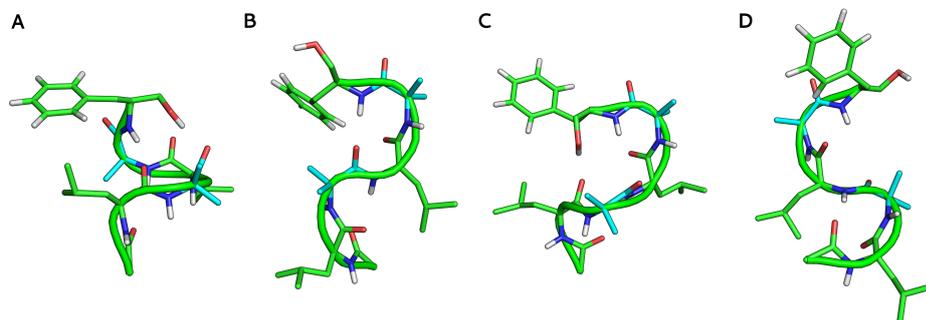
Analogue with ALA at A: t=0-10ns, B: t=40-50ns, C: t=50-60ns and D: t=90-100ns.
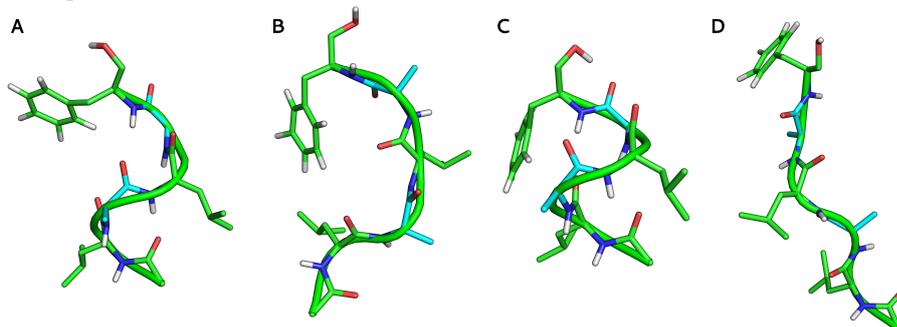
Analogue with Dhg at A: t=0-10ns, B: t=40-50ns, C: t=50-60ns and D: t=90-100ns.
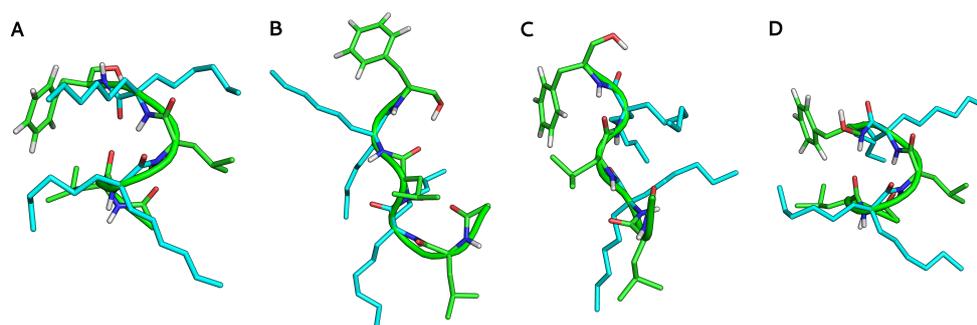
Figure S2. Ramachandran Plots

This section presents the probability contours (φ and ψ) superimposed on the Ramachandran diagram,

for the non-canonical amino acids Deg, Dpg, Dibg, Dϕg, Db$_z$g and Dmg.
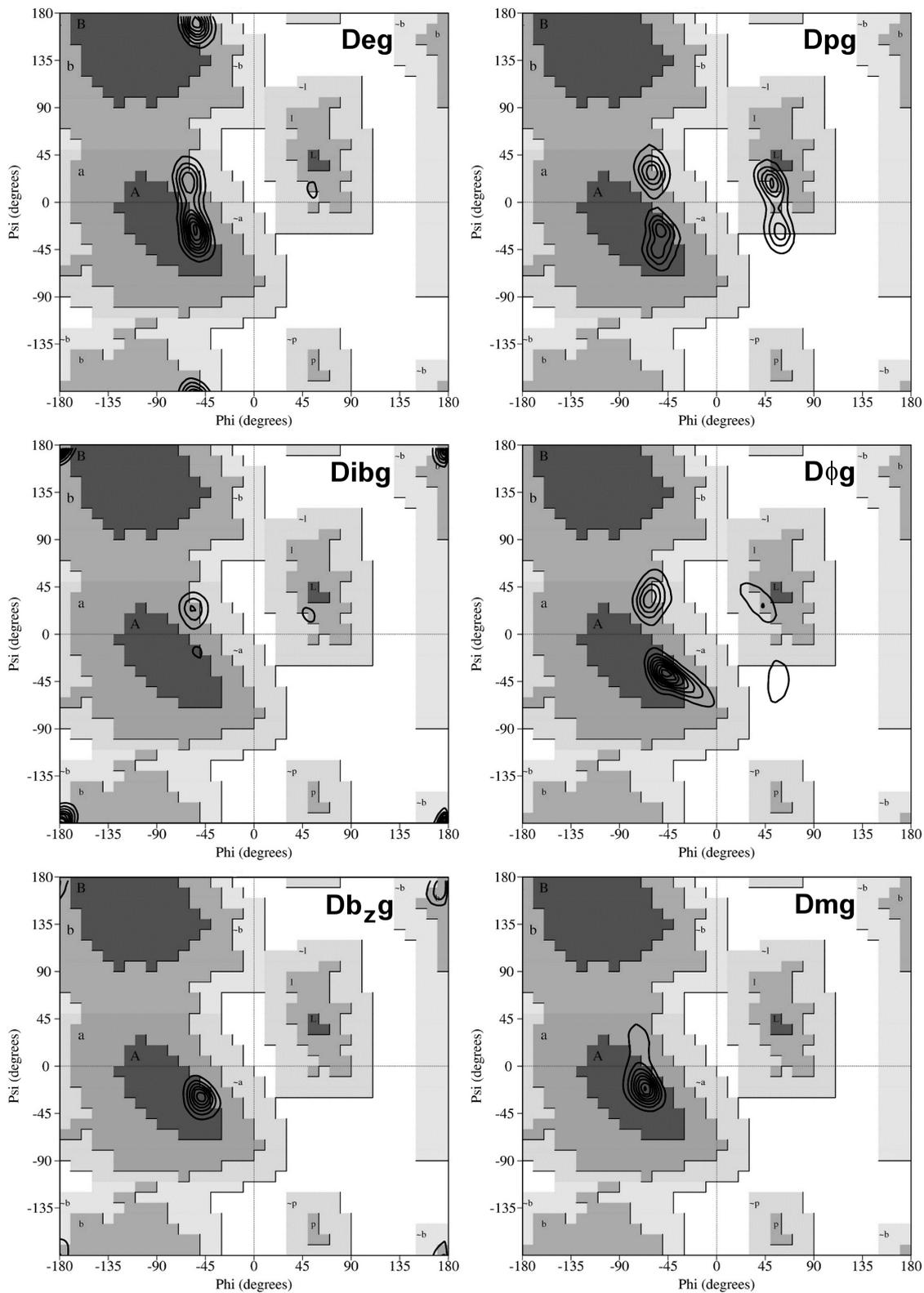
APPENDIX II

Chapter IV – Supporting Information

Conformational and Thermodynamic Properties of Non-Canonical α,α-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies

Tarsila Gabriel Castro and Nuno Miguel Micaêlo

Supporting Information

G54A7 FF Parameters

The Force Field parameters (bonded and non-bonded terms) for the α,α-dialkylglycines under study were developed based on the natural amino acids parameterized in de GROMOS 54a7 force field. They are the same showed in Appendix I, as these two works used the same residues.

Ramachandran Plots

This section presents the probability contours (φ and ψ) superimposed on the Ramachandran diagram, for the non-canonical amino acids Deg, Dibg, Dϕg, Db$_z$g, Ac$_6$c and Dmg.
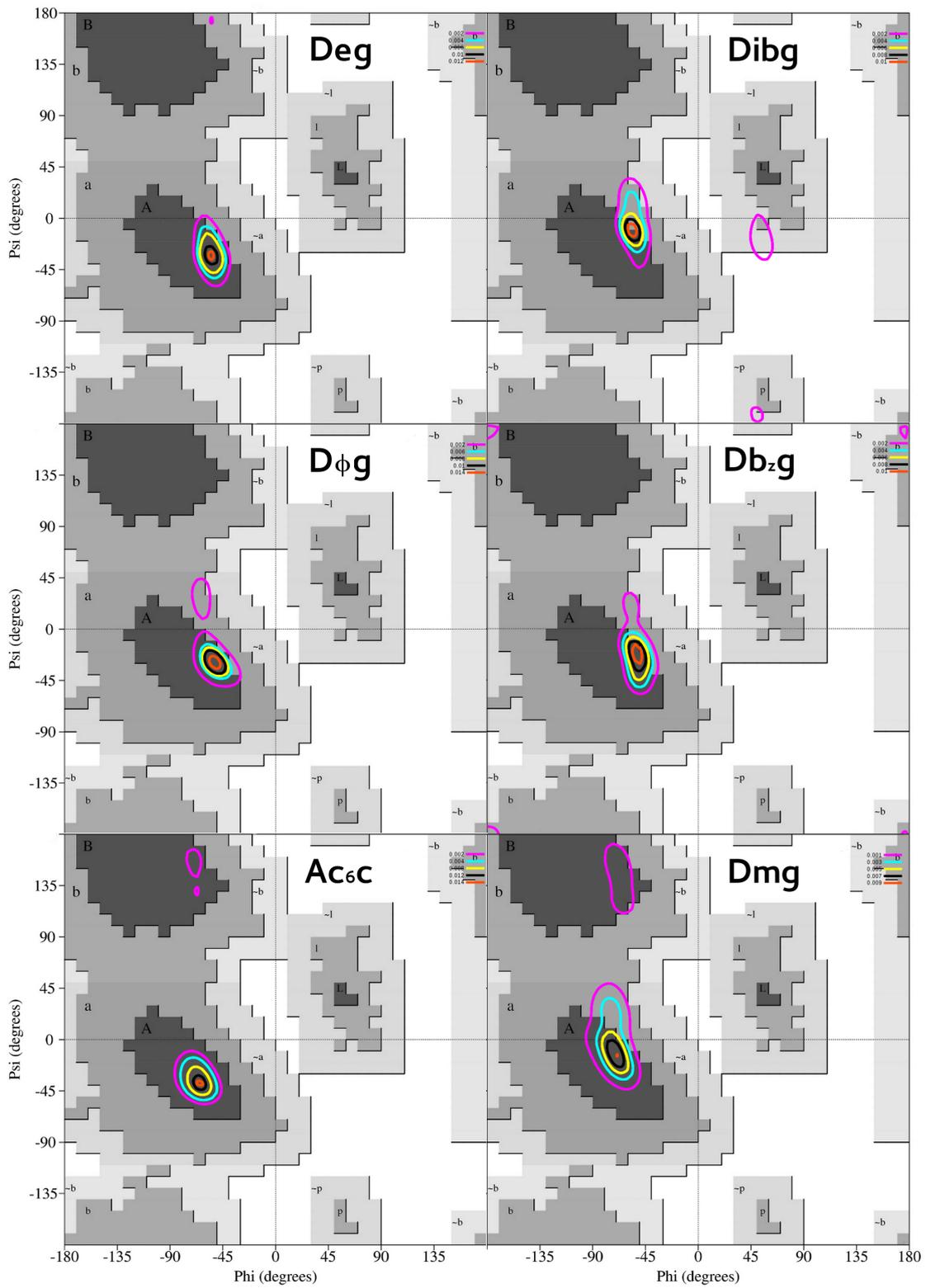
APPENDIX III

Chapter V – Supplementary Material and G54a7 FF parameters

The Secondary Structure Properties of Antiamoebin I and Zervamicin II Peptaibols Incorporating D-Amino Acids and Proline Analogues. A Modelling Study

Tarsila G. Castro, Nuno M. Micaêlo and Manuel Melle-Franco

Supporting Information

Table S1. Gromos 54a7 topologies

This section presents the topologies for the new asymmetrical α,α-dialkylglycines and proline analogs under study. These topologies were developed based on the natural amino acids parameterized in de GROMOS 54a7 force field.

Asymmetrical α,α-dialkyl glycines

Iva (isovaline):

```
[ DIV ]
 [ atoms ]
  N    N   -0.31000    0
  H    H    0.31000    0
  CA   C    0.00000    1
  CB2  CH3  0.00000    1
  CB1  CH2  0.00000    2
  CG1  CH3  0.00000    2
  C    C    0.450      3
  O    O   -0.450      3
 [ bonds ]
  N    H    gb_2
  N    CA   gb_21
  CA   CB1  gb_27
  CA   CB2  gb_27
  CA   C    gb_27
  CB1  CG1  gb_27
  C    O    gb_5
  C    +N   gb_10
 [ angles ]
; ai   aj   ak   gromos type
 -C    N    H    ga_32
 -C    N    CA   ga_31
  H    N    CA   ga_18
  N    CA   CB1  ga_13
  N    CA   C    ga_19
  CB1  CA   C    ga_13
  N    CA   CB2  ga_13
  CB1  CA   CB2  ga_13
  CB2  CA   C    ga_13
  CA   CB1  CG1  ga_13
```

```
  CA   C   O    ga_30
  CA   C   +N   ga_19
  O    C   +N   ga_33
[ impropers ]
; ai  aj  ak  al  gromos type
  N   -C   CA   H    gi_1
  CA   N   C    CB2  gi_2
  C    CA  +N   O    gi_1
  CA   N   CB1  C    gi_2
  CA   N   CB2  CB1  gi_2
[ dihedrals ]
; ai  aj  ak  al  gromos type
 -CA  -C   N    CA   gd_14
 -C    N   CA   C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C    N   CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N    CA  C    +N   gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N    CA  C    +N   gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  C    CA  CB1  CG1  gd_34
```

alpha-Methyl-D-Leucine

[MDL]
```
[ atoms ]
  N    N    -0.31000   0
  H    H     0.31000   0
  CA   C     0.00000   1
 CB2   CH3   0.00000   1
 CB1   CH2   0.00000   1
 CG1   CH1   0.00000   2
 CG2   CH3   0.00000   2
 CG3   CH3   0.00000   2
  C    C     0.450     3
  O    O    -0.450     3
[ bonds ]
  N    H    gb_2
  N    CA   gb_21
  CA   CB1  gb_27
  CA   CB2  gb_27
  CA   C    gb_27
 CB1   CG1  gb_27
 CG1   CG2  gb_27
 CG1   CG3  gb_27
  C    O    gb_5
  C    +N   gb_10
[ angles ]
; ai  aj  ak  gromos type
 -C    N    H    ga_32
 -C    N    CA   ga_31
  H    N    CA   ga_18
  N    CA   CB1  ga_13
  N    CA   CB2  ga_13
  N    CA   C    ga_13
 CB1   CA   C    ga_13
 CB2   CA   C    ga_13
 CB1   CA   CB2  ga_13
  CA   CB1  CG1  ga_15
 CB1   CG1  CG2  ga_15
 CB1   CG1  CG3  ga_15
```

```
 CG2  CG1  CG3    ga_15
  C   CA  CG2    ga_13
  CA   C   O    ga_30
  CA   C   +N    ga_19
  O   C   +N    ga_33
[ impropers ]
; ai  aj   ak   al  gromos type
  N  -C   CA   H    gi_1
  C   CA   +N   O    gi_1
  CA   N   C   CB2   gi_2
  CA   N   CB1   C    gi_2
  CA   N   CB2   CB1   gi_2
 CG1  CB1  CG2  CG3   gi_2
[ dihedrals ]
; ai  aj   ak   al  gromos type
 -CA  -C   N   CA    gd_14
 -C   N   CA   C     gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N   CA   C     gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA  CB1  CG1   gd_34
  N   CA   C   +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   C   +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 CA  CB1  CG1  CG2   gd_34
```

alpha-Methyl-D-phenylalanine

[MDP]
```
[ atoms ]
  N   N   -0.31000    0
  H   H    0.31000    0
  CA   CH1   0.00000    1
 CB1   CH2   0.00000    1
 CB2   CH3   0.00000    1
  CG   C    0.00000    1
 CD1   C   -0.14000    2
 HD1   HC   0.14000    2
 CD2   C   -0.14000    3
 HD2   HC   0.14000    3
 CE1   C   -0.14000    4
 HE1   HC   0.14000    4
 CE2   C   -0.14000    5
 HE2   HC   0.14000    5
  CZ   C   -0.14000    6
  HZ   HC   0.14000    6
  C   C    0.450    7
  O   O   -0.450    7
[ bonds ]
  N   H   gb_2
  N   CA   gb_21
  CA  CB1   gb_27
  CA  CB2   gb_27
  CA   C   gb_27
 CB1   CG   gb_27
  CG  CD1   gb_16
  CG  CD2   gb_16
 CD1  HD1   gb_3
 CD1  CE1   gb_16
 CD2  HD2   gb_3
 CD2  CE2   gb_16
```

```
CE1  HE1   gb_3
CE1  CZ    gb_16
CE2  HE2   gb_3
CE2  CZ    gb_16
CZ   HZ    gb_3
C    O     gb_5
C    +N    gb_10
```

[exclusions]
```
;  ai   aj
CB1  HD1
CB1  HD2
CB1  CE1
CB1  CE2
CG   HE1
CG   HE2
CG   CZ
CD1  HD2
CD1  CE2
CD1  HZ
HD1  CD2
HD1  HE1
HD1  CZ
CD2  CE1
CD2  HZ
HD2  HE2
HD2  CZ
CE1  HE2
HE1  CE2
HE1  HZ
HE2  HZ
```

[angles]
```
;  ai   aj   ak   gromos type
-C   N    H    ga_32
-C   N    CA   ga_31
H    N    CA   ga_18
N    CA   CB1  ga_13
N    CA   C    ga_13
CB1  CA   C    ga_13
CA   CB1  CG   ga_15
CB1  CG   CD1  ga_27
CB1  CG   CD2  ga_27
CD1  CG   CD2  ga_27
CG   CD1  HD1  ga_25
CG   CD1  CE1  ga_27
HD1  CD1  CE1  ga_25
CG   CD2  HD2  ga_25
CG   CD2  CE2  ga_27
HD2  CD2  CE2  ga_25
CD1  CE1  HE1  ga_25
CD1  CE1  CZ   ga_27
HE1  CE1  CZ   ga_25
CD2  CE2  HE2  ga_25
CD2  CE2  CZ   ga_27
HE2  CE2  CZ   ga_25
CE1  CZ   CE2  ga_27
CE1  CZ   HZ   ga_25
CE2  CZ   HZ   ga_25
CA   C    O    ga_30
```

```
   CA   C   +N   ga_19
   O    C   +N   ga_33
   N   CA   CB2   ga_13
  CB1  CA   CB2   ga_13
  CB2  CA   C    ga_13
[ impropers ]
; ai   aj   ak   al   gromos type
   N   -C   CA   H    gi_1
   CA   N    C   CB2   gi_2
  CG  CD1  CD2  CB1   gi_1
  CG  CD1  CE1  CZ    gi_1
  CG  CD2  CE2  CZ    gi_1
  CD1  CG  CD2  CE2   gi_1
  CD1  CG  CE1  HD1   gi_1
  CD1 CE1  CZ  CE2    gi_1
  CD2  CG  CD1  CE1   gi_1
  CD2  CG  CE2  HD2   gi_1
  CD2 CE2  CZ  CE1    gi_1
  HE1 CD1  CZ  CE1    gi_1
  HE2 CD2  CZ  CE2    gi_1
  CZ  CE1  CE2  HZ    gi_1
   C   CA   +N   O    gi_1
   CA   N   CB1   C    gi_2
   CA   N   CB2  CB1   gi_2
[ dihedrals ]
; ai   aj   ak   al   gromos type
 -CA  -C   N    CA    gd_14
 -C    N    CA   C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C    N    CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   CB1   CG    gd_34
  N   CA   C    +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   C    +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA  CB1  CG  CD1     gd_40
```

Methyl-2-cyclopentyl-2-(formylamino)propanoate
alpha-methyl-D-cyclopentyl

```
[ MCP ]
[ atoms ]
  N    N   -0.31000   0
  H    H    0.31000   0
  CA   C    0.00000   1
  CB2 CH3   0.00000   1
  CB1 CH1   0.00000   1
  CG1 CH2r  0.00000   2
  CG2 CH2r  0.00000   2
  CG3 CH2r  0.00000   3
  CG4 CH2r  0.00000   3
  C    C    0.450     4
  O    O   -0.450     4
[ bonds ]
  N    H   gb_2
  N    CA  gb_21
  CA   C   gb_27
  CA   CB2 gb_27
  CA   CB1 gb_27
  C    O   gb_5
  C    +N  gb_10
```

```
 CB1  CG1   gb_27
 CB1  CG4   gb_27
 CG1  CG2   gb_27
 CG2  CG3   gb_27
 CG3  CG4   gb_27
[ angles ]
; ai  aj   ak  gromos type
 -C   N    H    ga_32
 -C   N    CA   ga_31
  H   N    CA   ga_18
  N   CA   C    ga_13
 CA   C    O    ga_30
 CA   C    +N   ga_19
  O   C    +N   ga_33
  N   CA   CB2  ga_13
 CB1  CA   CB2   ga_13
 CB2  CA   C    ga_13
  N   CA   CB1   ga_13
  C   CA   CB1   ga_13
 CA   CB1  CG1   ga_13
 CA   CB1  CG4   ga_13
 CG1  CB1  CG4    ga_7
 CB1  CG4  CG3    ga_7
 CG4  CG3  CG2    ga_7
 CG3  CG2  CG1    ga_7
 CG2  CG1  CB1    ga_7
[ impropers ]
; ai  aj   ak  al  gromos type
  N   -C   CA   H    gi_1
 CA   N    C   CB2    gi_2
 CA   N    CB1  C    gi_2
 CA   N    CB2  CB1   gi_2
  C   CA   +N   O    gi_1
 CB1  CA   CG4  CG1    gi_2
[ dihedrals ]
; ai  aj   ak  al  gromos type
 -CA  -C   N    CA    gd_14
 -C   N    CA   C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N    CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   CB1  CG4    gd_34
 CA   CB1  CG4  CG3    gd_34
 CA   CB1  CG1  CG2    gd_34
 CB1  CG1  CG2  CG3    gd_34
 CG2  CG3  CG4  CB1    gd_34
 CG1  CG2  CG3  CG4    gd_1
  N   CA   C    +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   C    +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

2-amino-2-(2-cyclopentenyl)propanoic acid
alpha-methyl-D-cyclopentenyl (MDC)

[MDC]
```
[ atoms ]
  N    N   -0.31000   0
  H    H    0.31000   0
 CA    C    0.00000   1
 CB2  CH3   0.00000   1
 CB1  CH1   0.00000   1
```

```
  CG1 CH2r   0.00000   2
  CG2 CH2r   0.00000   2
  CG3 CH1    0.00000   3
  CG4 CH1    0.00000   3
  C   C      0.450     4
  O   O     -0.450     4
[ bonds ]
  N   H   gb_2
  N   CA  gb_21
  CA  C   gb_27
  CA  CB2 gb_27
  CA  CB1 gb_27
  C   O   gb_5
  C   +N  gb_10
  CB1 CG1 gb_27
  CB1 CG4 gb_27
  CG1 CG2 gb_27
  CG2 CG3 gb_27
  CG3 CG4 gb_27
[ angles ]
; ai  aj   ak   gromos type
 -C    N    H    ga_32
 -C    N    CA   ga_31
  H    N    CA   ga_18
  N    CA   C    ga_13
  CA   C    O    ga_30
  CA   C    +N   ga_19
  O    C    +N   ga_33
  N    CA   CB2  ga_13
  CB1  CA   CB2  ga_13
  CB2  CA   C    ga_13
  N    CA   CB1  ga_13
  C    CA   CB1  ga_13
  CA   CB1  CG1  ga_13
  CA   CB1  CG4  ga_13
  CG1  CB1  CG4  ga_7
  CB1  CG4  CG3  ga_27
  CG4  CG3  CG2  ga_27
  CG3  CG2  CG1  ga_7
  CG2  CG1  CB1  ga_7
[ impropers ]
; ai  aj   ak   al  gromos type
  N   -C   CA   H    gi_1
  CA   N   C    CB2  gi_2
  CA   N   CB1  C    gi_2
  CA   N   CB2  CB1  gi_2
  C   CA   +N   O    gi_1
  CB1 CA   CG4  CG1  gi_2
  CB1 CG4  CG3  CG2  gi_1
[ dihedrals ]
; ai  aj   ak   al  gromos type
 -CA  -C    N   CA   gd_14
 -C    N    CA   C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C    N    CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   CB1  CG4  gd_34
  CA  CB1  CG4  CG3  gd_34
  CA  CB1  CG1  CG2  gd_34
  CB1 CG1  CG2  CG3  gd_34
```

```
  CG2  CG3  CG4  CB1    gd_5
 CG1  CG2  CG3  CG4    gd_1
   N   CA   C   +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   N   CA   C   +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

2-amino-2methyl-4-pentenoic acid
alpha-methyl-D-2-propeno

[MPR]
```
[ atoms ]
   N    N    -0.31000    0
   H    H     0.31000    0
  CA    C     0.00000    1
  CB2  CH3    0.00000    1
  CB1  CH2    0.00000    2
  CG1  CH1    0.00000    2
  CG2  CH2    0.00000    2
   C    C     0.450      3
   O    O    -0.450      3
[ bonds ]
   N    H    gb_2
   N    CA   gb_21
  CA    CB1  gb_27
  CA    CB2  gb_27
  CA    C    gb_27
  CB1   CG1  gb_27
  CG1   CG2  gb_27
   C    O    gb_5
   C   +N    gb_10
[ angles ]
; ai  aj   ak   gromos type
  -C   N    H    ga_32
  -C   N    CA   ga_31
   H   N    CA   ga_18
   N   CA   CB1  ga_13
   N   CA   C    ga_19
  CB1  CA   C    ga_13
   N   CA   CB2  ga_13
  CB1  CA   CB2  ga_13
  CB2  CA   C    ga_13
  CA   CB1  CG1  ga_13
  CB1  CG1  CG2  ga_28
  CA   C    O    ga_30
  CA   C   +N    ga_19
   O   C   +N    ga_33
[ impropers ]
; ai  aj   ak   al  gromos type
   N   -C   CA   H    gi_1
  CA   N    C    CB2  gi_2
   C   CA  +N    O    gi_1
  CA   N    CB1  C    gi_2
  CA   N    CB2  CB1  gi_2
[ dihedrals ]
; ai  aj   ak   al  gromos type
 -CA  -C   N    CA   gd_14
  -C   N    CA   C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  -C   N    CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   N   CA   C   +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

```
   N   CA   C   +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   C   CA   CB1  CG1    gd_34
  CA   CB1  CG1  CG2    gd_34
```

Proline analogs

4-hydroxiproline

```
[ HYP ]
 [ atoms ]
   N    N    0.00000    0
  CA   CH1    0.00000    1
  CB  CH2r   0.00000    1
  CG   CH1   0.26600    2
  OD1   OA   -0.67400    2
  HD1   H    0.40800    2
  CD  CH2r   0.00000    3
   C    C    0.450    4
   O    O    -0.450    4
 [ bonds ]
   N   CA    gb_21
   N   CD    gb_21
  CA   CB    gb_27
  CA    C    gb_27
  CB   CG    gb_27
  CG   OD1    gb_18
  CG   CD    gb_27
  OD1   HD1    gb_1
   C    O    gb_5
   C   +N    gb_10
 [ angles ]
; ai   aj   ak   gromos type
  -C    N   CA    ga_31
  -C    N   CD    ga_31
  CA    N   CD    ga_21
   N   CA   CB    ga_13
   N   CA    C    ga_13
  CB   CA    C    ga_13
  CA   CB   CG    ga_13
  CB   CG   OD1    ga_13
  CB   CG   CD    ga_13
  OD1   CG   CD    ga_13
  CG   OD1   HD1    ga_12
   N   CD   CG    ga_13
  CA    C    O    ga_30
  CA    C   +N    ga_19
   O    C   +N    ga_33
 [ impropers ]
; ai   aj   ak   al  gromos type
   N   -C   CA   CD   gi_1
  CA    C   CB    N   gi_2
  CG   OD1   CB   CD   gi_2
   C   CA   +N    O   gi_1
 [ dihedrals ]
; ai   aj   ak   al   gromos type
  -CA  -C    N   CA    gd_14
  -C    N   CA    C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

```
 -C   N   CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA   N   CD   CG    gd_39
  N   CA   CB   CG    gd_34
  N   CA   C   +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   C   +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA   CB   CG   CD    gd_34
  CB   CG   OD1   HD1    gd_11
  CB   CG   CD   N    gd_34
```

cis-3-amino-L-proline

[ALP]
```
 [ atoms ]
  N   N   0.00000   0
  CA  CH1   0.00000   1
  CG  CH2r   0.00000   2
  CD  CH2r   0.00000   2
  CB  CH1   0.12700   3
  N01   NZ   0.12900   4
  H01   H   0.24800   4
  H02   H   0.24800   4
  H03   H   0.24800   4
  C   C   0.450   5
  O   O   -0.450   5
 [ bonds ]
  N   CA   gb_21
  N   CD   gb_21
  CA   CB   gb_27
  CA   C   gb_27
  CB   CG   gb_27
  CB   N01   gb_9
  N01   H01   gb_2
  N01   H02   gb_2
  N01   H03   gb_2
  CG   CD   gb_27
  C   O   gb_5
  C   +N   gb_10
 [ angles ]
 ; ai   aj   ak   gromos type
  -C   N   CA   ga_31
  -C   N   CD   ga_31
  CA   N   CD   ga_21
  N   CA   CB   ga_13
  N   CA   C   ga_13
  CB   CA   C   ga_13
  CA   CB   CG   ga_13
  CA   CB   N01   ga_19
  CB   CG   CD   ga_13
  N01   CB   CG   ga_19
  CB   N01   H01   ga_23
  CB   N01   H02   ga_23
  CB   N01   H03   ga_23
  H02   N01   H01   ga_10
  H02   N01   H03   ga_10
  H01   N01   H03   ga_10
  N   CD   CG   ga_13
  CA   C   O   ga_30
  CA   C   +N   ga_19
```

```
     O   C   +N    ga_33
[ impropers ]
;  ai   aj   ak   al   gromos type
   N   -C   CA   CD   gi_1
   CA   C   CB   N   gi_2
   C   CA   +N   O   gi_1
   CB   N01   CA   CG   gi_2
[ dihedrals ]
;  ai   aj   ak   al   gromos type
  -CA   -C   N   CA   gd_14
  -C   N   CA   C   gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  -C   N   CA   C   gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   CA   N   CD   CG   gd_39
   CB   CA   N   CD   gd_39
   N   CA   CB   CG   gd_34
   N   CA   C   +N   gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   N   CA   C   +N   gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   CA   CB   CG   CD   gd_34
   CA   CB   N01   H01   gd_14
   CB   CG   CD   N   gd_34
```

cis-4-methyl-L-proline

```
[ MLP ]
[ atoms ]
   N   N   0.00000   0
   CA   CH1   0.00000   1
   CB   CH2r   0.00000   1
   CG   CH1   0.00000   2
   C01   CH3   0.00000   2
   CD   CH2r   0.00000   3
   C   C   0.450   4
   O   O   -0.450   4
[ bonds ]
   N   CA   gb_21
   N   CD   gb_21
   CA   CB   gb_27
   CA   C   gb_27
   CB   CG   gb_27
   CG   C01   gb_27
   CG   CD   gb_27
   C   O   gb_5
   C   +N   gb_10
[ angles ]
;  ai   aj   ak   gromos type
  -C   N   CA   ga_31
  -C   N   CD   ga_31
   CA   N   CD   ga_21
   N   CA   CB   ga_13
   N   CA   C   ga_13
   CB   CA   C   ga_13
   CA   CB   CG   ga_13
   CB   CG   C01   ga_13
   CB   CG   CD   ga_13
   C01   CG   CD   ga_13
   N   CD   CG   ga_13
   CA   C   O   ga_30
   CA   C   +N   ga_19
```

```
   O    C    +N     ga_33
[ impropers ]
; ai   aj   ak   al   gromos type
  N    -C   CA   CD   gi_1
  CA   C    CB   N    gi_2
  CG   C01  CB   CD   gi_2
  C    CA   +N   O    gi_1
[ dihedrals ]
; ai   aj   ak   al   gromos type
 -CA  -C    N    CA   gd_14
 -C    N    CA   C    gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C    N    CA   C    gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA   N    CD   CG   gd_39
  N    CA   CB   CG   gd_34
  N    CA   C    +N   gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N    CA   C    +N   gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA   CB   CG   CD   gd_34
  CB   CG   CD   N    gd_34
```

trans-3-hydroxy-L-proline

[HLP]
```
[ atoms ]
  N    N     0.00000    0
  CA   CH1   0.00000    1
  CG   CH2r  0.00000    2
  CD   CH2r  0.00000    2
  CB   CH1   0.26600    3
  O01  OA   -0.67400    3
  H01  H     0.40800    3
  C    C     0.450      4
  O    O    -0.450      4
[ bonds ]
  N    CA   gb_21
  N    CD   gb_21
  CA   CB   gb_27
  CA   C    gb_27
  CB   CG   gb_27
  CB   O01  gb_18
  O01  H01  gb_1
  CG   CD   gb_27
  C    O    gb_5
  C    +N   gb_10
[ angles ]
; ai   aj   ak   gromos type
 -C    N    CA   ga_31
 -C    N    CD   ga_31
  CA   N    CD   ga_21
  N    CA   CB   ga_13
  N    CA   C    ga_13
  CB   CA   C    ga_13
  CA   CB   CG   ga_13
  CA   CB   O01  ga_13
  CG   CB   O01  ga_13
  CB   O01  H01  ga_12
  CB   CG   CD   ga_13
  N    CD   CG   ga_13
  CA   C    O    ga_30
```

```
  CA    C    +N    ga_19
   O    C    +N    ga_33
 [ impropers ]
; ai   aj   ak    al   gromos type
  N   -C   CA   CD    gi_1
  CA    C   CB    N    gi_2
   C   CA   +N    O    gi_1
  CB   O01   CA   CG    gi_2
 [ dihedrals ]
; ai   aj   ak    al   gromos type
 -CA   -C    N    CA    gd_14
 -C    N    CA    C     gd_42 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C    N    CA    C     gd_43 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA    N   CD    CG    gd_39
  CB   CA    N    CD    gd_39
  N    CA   CB    CG    gd_34
  N    CA    C    +N    gd_44 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N    CA    C    +N    gd_45 ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA   CB   CG    CD    gd_34
  CA   CB   O01   H01    gd_11
  CB   CG   CD    N     gd_34
```

Figures

Figure S1
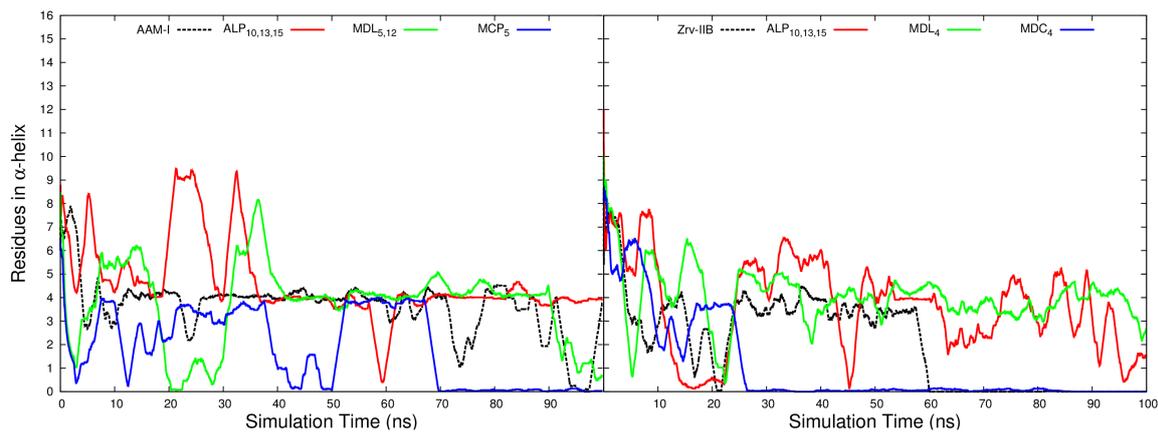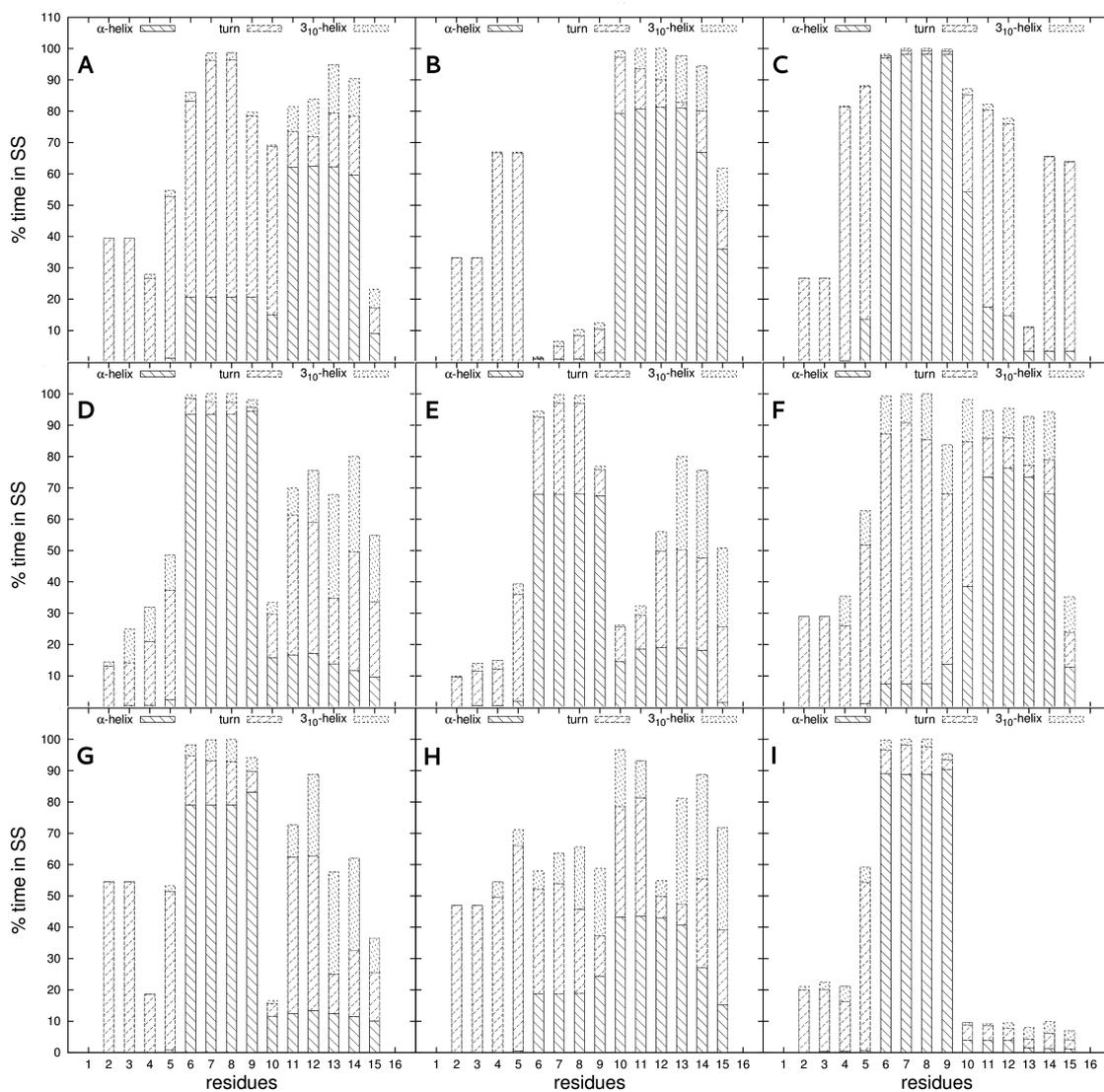
Figure S2. Percentage of time in a-helix, turn and 3_{10}-helix conformations for each residue of the AAM-I peptide (A) and analogs carrying (B) Iva$_{(2, 3, 4, 8, 9, 14)}$, (C) Hyp$_{(2, 3, 4, 8, 9, 14)}$, (D) ALP$_{(10,13,15)}$, (E) HLP$_{(10,13,15)}$, (F) MDL$_{(5,12)}$, (G) MDP$_{(5,12)}$, (H) MPR$_{(5)}$, (I) MLP$_{(10,15)}$, considering the last 60ns of simulation time.
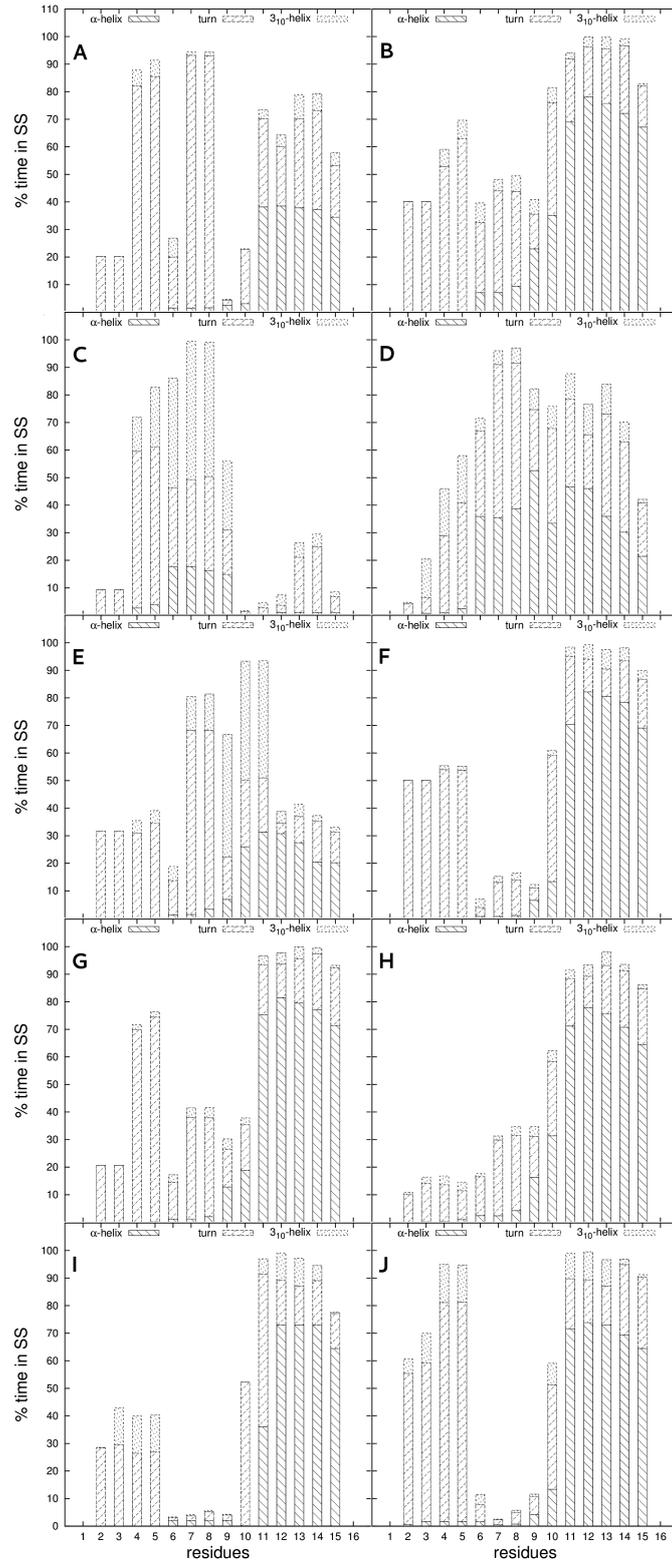
Figure S3. Percentage of time in α-helix, turn and 3_{10}-helix conformations for each residue of the Zrv-IIB peptide (A) and analogs carrying (B) Iva$_{(7, 9, 12, 14)}$, (C) Hyp$_{(7, 9, 12, 14)}$, (D) ALP$_{(10,13,15)}$, (E) MCP$_{(4)}$, (F) MDL$_{(4)}$, (G) MDP$_{(4)}$, (H) HLP$_{(10,15)}$, (I) MLP$_{(10,15)}$ and (J) MLP$_{(13,15)}$, considering the last 60ns of simulation time.

APPENDIX IV

Chapters VI and VII– G54a7 FF Parameters

Dehydroamino Acids FF parameters

; This file has non-natural AA developed by Tarsila and Micaelo

; Dehydro amino acids

[bondedtypes]
; bonds angles dihedrals impropers
     2   2   1   2

; alpha,beta-dehydro amino acids

[DPH]
[atoms]
```
  N    N    -0.31000    0
  H    H     0.31000    0
  CA   C     0.00000    1
  CB   CH1   0.00000    1
  CG   C     0.00000    1
  CD1  C    -0.14000    2
  HD1  HC    0.14000    2
  CD2  C    -0.14000    3
  HD2  HC    0.14000    3
  CE1  C    -0.14000    4
  HE1  HC    0.14000    4
  CE2  C    -0.14000    5
  HE2  HC    0.14000    5
  CZ   C    -0.14000    6
  HZ   HC    0.14000    6
  C    C     0.450      7
  O    O    -0.450      7
```
[bonds]
```
  N    H    gb_2
  N    CA   gb_21
  CA   CB   gb_3
  CA   C    gb_27
  CB   CG   gb_27
  CG   CD1  gb_16
  CG   CD2  gb_16
  CD1  HD1  gb_3
  CD1  CE1  gb_16
  CD2  HD2  gb_3
  CD2  CE2  gb_16
  CE1  HE1  gb_3
  CE1  CZ   gb_16
  CE2  HE2  gb_3
  CE2  CZ   gb_16
  CZ   HZ   gb_3
  C    O    gb_5
  C    +N   gb_10
```
[exclusions]
; ai   aj
```
  CB   HD1
  CB   HD2
  CB   CE1
  CB   CE2
  CG   HE1
  CG   HE2
  CG   CZ
  CD1  HD2
  CD1  CE2
  CD1  HZ
  HD1  CD2
  HD1  HE1
  HD1  CZ
  CD2  CE1
  CD2  HZ
  HD2  HE2
  HD2  CZ
  CE1  HE2
  HE1  CE2
  HE1  HZ
  HE2  HZ
```
[angles]
```
; ai   aj   ak   gromos type
  -C   N    H    ga_32
  -C   N    CA   ga_31
  H    N    CA   ga_18
  N    CA   CB   ga_26
  N    CA   C    ga_26
  CB   CA   C    ga_26
  CA   CB   CG   ga_15
  CB   CG   CD1  ga_27
  CB   CG   CD2  ga_27
  CD1  CG   CD2  ga_27
  CG   CD1  HD1  ga_25
  CG   CD1  CE1  ga_27
  HD1  CD1  CE1  ga_25
  CG   CD2  HD2  ga_25
  CG   CD2  CE2  ga_27
  HD2  CD2  CE2  ga_25
  CD1  CE1  HE1  ga_25
  CD1  CE1  CZ   ga_27
  HE1  CE1  CZ   ga_25
  CD2  CE2  HE2  ga_25
  CD2  CE2  CZ   ga_27
  HE2  CE2  CZ   ga_25
  CE1  CZ   CE2  ga_27
```

```
 CE1   CZ   HZ    ga_25                        H    N    CA    ga_18
 CE2   CZ   HZ    ga_25                        N    CA   CB    ga_26
 CA    C    O     ga_30                        N    CA   C     ga_26
 CA    C    +N    ga_19                        CB   CA   C     ga_26
 O     C    +N    ga_33                        CA   CB   CG    ga_15
[ impropers ]                                  CA   C    O     ga_30
; ai  aj   ak   al  gromos type               CA   C    +N    ga_19
 N    -C   CA   H    gi_1                       O    C    +N    ga_33
 CA   N    C    CB   gi_1                      [ impropers ]
 CA   CB   N    C    gi_1                      ; ai  aj   ak   al  gromos type
 CG   CD1  CD2  CB   gi_1                       N    -C   CA   H    gi_1
 CG   CD1  CE1  CZ   gi_1                       CA   N    C    CB   gi_1
 CG   CD2  CE2  CZ   gi_1                       C    CA   +N   O    gi_1
 CD1  CG   CD2  CE2  gi_1                      [ dihedrals ]
 CD1  CG   CE1  HD1  gi_1                      ; ai  aj   ak   al  gromos type
 CD1  CE1  CZ   CE2  gi_1                       -CA  -C   N    CA   gd_14
 CD2  CG   CD1  CE1  gi_1                       -C   N    CA   C    gd_42
 CD2  CG   CE2  HD2  gi_1                      ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 CD2  CE2  CZ   CE1  gi_1                       -C   N    CA   C    gd_43
 HE1  CD1  CZ   CE1  gi_1                      ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 HE2  CD2  CZ   CE2  gi_1                       C    CA   CB   CG   gd_47
 CZ   CE1  CE2  HZ   gi_1                       N    CA   C    +N   gd_44
 C    CA   +N   O    gi_1                      ;backbone dihedral, changed by Ying Xue Sep 29. 2009
[ dihedrals ]                                  N    CA   C    +N   gd_45
; ai  aj   ak   al  gromos type               ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -CA  -C   N    CA   gd_14
 -C   N    CA   C    gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N    CA   C    gd_43                     [ DLE ]
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 C    CA   CB   CG   gd_47                     [ atoms ]
 N    CA   C    +N   gd_44                      N    N    -0.31000    0
;backbone dihedral, changed by Ying Xue Sep 29. 2009  H    H    0.31000     0
 N    CA   C    +N   gd_45                      CA   C    0.00000     1
;backbone dihedral, changed by Ying Xue Sep 29. 2009  CB   CH1  0.00000     1
 CA   CB   CG   CD1  gd_40                      C1   CH1  0.00000     2
                                               C2   CH3  0.00000     2
                                               C3   CH3  0.00000     2
[ ABU ]                                        C    C    0.450      3
 [ atoms ]                                     O    O    -0.450     3
  N    N    -0.31000    0                      [ bonds ]
  H    H    0.31000     0                       N    H    gb_2
  CA   C    0.00000     1                       N    CA   gb_21
  CB   CH1  0.00000     1                       CA   CB   gb_27
  CG   CH3  0.00000     1                       CA   C    gb_27
  C    C    0.450      2                        CB   C1   gb_27
  O    O    -0.450     2                        C1   C2   gb_27
 [ bonds ]                                      C1   C3   gb_27
  N    H    gb_2                                C    O    gb_5
  N    CA   gb_21                               C    +N   gb_10
  CA   CB   gb_27                              [ angles ]
  CA   C    gb_27                              ; ai  aj   ak  gromos type
  CB   CG   gb_27                               -C   N    H    ga_32
  C    O    gb_5                                -C   N    CA   ga_31
  C    +N   gb_10                               H    N    CA   ga_18
 [ angles ]                                     N    CA   CB   ga_26
; ai  aj   ak  gromos type                      N    CA   C    ga_26
  -C   N    H    ga_32                          CB   CA   C    ga_26
  -C   N    CA   ga_31                          CA   CB   C1   ga_26
                                                CB   C1   C2   ga_15
                                                CB   C1   C3   ga_15
```

```
C2   C1   C3   ga_15
CA   C    O    ga_30
CA   C    +N   ga_19
O    C    +N   ga_33
[ impropers ]
;  ai   aj   ak   al   gromos type
 N   -C   CA   H    gi_1
CA   N    C    CB   gi_1
C1   CB   C3   C2   gi_2
 C   CA   +N   O    gi_1
[ dihedrals ]
;  ai   aj   ak   al  gromos type
 -CA  -C   N    CA   gd_14
 -C   N    CA   C    gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N    CA   C    gd_43
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 C    CA   CB   C1   gd_47
 N    CA   C    +N   gd_44
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 N    CA   C    +N   gd_45
;backbone dihedral, changed by Ying Xue Sep 29. 2009
CA   CB   C1   C2   gd_34


[ DVA ]
[ atoms ]
 N    N    -0.31000    0
 H    H    0.31000     0
CA    C    0.00000     1
CB    C    0.00000     1
C1   CH3   0.00000     1
C2   CH3   0.00000     1
 C    C    0.450       2
 O    O    -0.450      2
[ bonds ]
 N    H    gb_2
 N    CA   gb_21
CA   CB   gb_27
CA   C    gb_27
CB   C1   gb_27
CB   C2   gb_27
 C    O    gb_5
 C    +N   gb_10
[ angles ]
;  ai   aj   ak  gromos type
 -C   N    H    ga_32
 -C   N    CA   ga_31
 H    N    CA   ga_18
 N    CA   CB   ga_26
 N    CA   C    ga_26
CB   CA   C    ga_26
CA   CB   C1   ga_26
CA   CB   C2   ga_26
C1   CB   C2   ga_26
CA   C    O    ga_30
CA   C    +N   ga_19
 O    C    +N   ga_33
[ impropers ]
```

```
;  ai   aj   ak   al   gromos type
 N   -C   CA   H    gi_1
CA   N    C    CB   gi_1
CB   CA   C2   C1   gi_1
 C   CA   +N   O    gi_1
[ dihedrals ]
;  ai   aj   ak   al  gromos type
 -CA  -C   N    CA   gd_14
 -C   N    CA   C    gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N    CA   C    gd_43
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 C    CA   CB   C2   gd_47
 C    CA   CB   C1   gd_46
 N    CA   C    +N   gd_44
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 N    CA   C    +N   gd_45
;backbone dihedral, changed by Ying Xue Sep 29. 2009


[ DAL ]
[ atoms ]
 N    N    -0.31000    0
 H    H    0.31000     0
CA    C    0.00000     1
CB   CH2   0.00000     1
 C    C    0.450       2
 O    O    -0.450      2
[ bonds ]
 N    H    gb_2
 N    CA   gb_21
CA   CB   gb_27
CA   C    gb_27
 C    O    gb_5
 C    +N   gb_10
[ angles ]
;  ai   aj   ak  gromos type
 -C   N    H    ga_32
 -C   N    CA   ga_31
 H    N    CA   ga_18
 N    CA   CB   ga_26
 N    CA   C    ga_26
CB   CA   C    ga_26
CA   C    O    ga_30
CA   C    +N   ga_19
 O    C    +N   ga_33
[ impropers ]
;  ai   aj   ak   al   gromos type
 N   -C   CA   H    gi_1
CB   N    C    CA   gi_1
 C   CA   +N   O    gi_1
[ dihedrals ]
;  ai   aj   ak   al  gromos type
 -CA  -C   N    CA   gd_14
 -C   N    CA   C    gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N    CA   C    gd_43
```

;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   C   +N   gd_44
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA   C   +N   gd_45
;backbone dihedral, changed by Ying Xue Sep 29. 2009

[DTR]
[atoms]
  N   N   -0.31000   0
  H   H   0.31000   0
  CA   C   0.00000   1
  CB   CH1   0.00000   1
  CG   C   -0.21000   2
  CD1 C   -0.14000   2
  HD1 HC   0.14000   2
  CD2 C   0.00000   2
  NE1 NR   -0.10000   2
  HE1 H   0.31000   2
  CE2 C   0.00000   2
  CE3 C   -0.14000   3
  HE3 HC   0.14000   3
  CZ2 C   -0.14000   4
  HZ2 HC   0.14000   4
  CZ3 C   -0.14000   5
  HZ3 HC   0.14000   5
  CH2 C   -0.14000   6
  HH2 HC   0.14000   6
  C   C   0.450   7
  O   O   -0.450   7
[bonds]
  N   H   gb_2
  N   CA   gb_21
  CA   CB   gb_27
  CA   C   gb_27
  CB   CG   gb_27
  CG   CD1 gb_10
  CG   CD2 gb_16
  CD1 HD1 gb_3
  CD1 NE1 gb_10
  CD2 CE2 gb_16
  CD2 CE3 gb_16
  NE1 HE1 gb_2
  NE1 CE2 gb_10
  CE2 CZ2 gb_16
  CE3 HE3 gb_3
  CE3 CZ3 gb_16
  CZ2 HZ2 gb_3
  CZ2 CH2 gb_16
  CZ3 HZ3 gb_3
  CZ3 CH2 gb_16
  CH2 HH2 gb_3
  C   O   gb_5
  C   +N   gb_10
[exclusions]
; ai   aj

CB   HD1
CB   NE1
CB   CE2
CB   CE3
CG   HE1
CG   HE3
CG   CZ2
CG   CZ3
CD1 CE3
CD1 CZ2
HD1 CD2
HD1 HE1
HD1 CE2
CD2 HE1
CD2 HZ2
CD2 HZ3
CD2 CH2
NE1 CE3
NE1 HZ2
NE1 CH2
HE1 CZ2
CE2 HE3
CE2 CZ3
CE2 HH2
CE3 CZ2
CE3 HH2
HE3 HZ3
HE3 CH2
CZ2 HZ3
HZ2 CZ3
HZ2 HH2
HZ3 HH2
[angles]
; ai   aj   ak   gromos type
 -C   N   H   ga_32
 -C   N   CA   ga_31
  H   N   CA   ga_18
  N   CA   CB   ga_26
  N   CA   C   ga_26
  CB   CA   C   ga_26
  CA   CB   CG   ga_26
  CB   CG   CD1   ga_37
  CB   CG   CD2   ga_37
  CD1 CG   CD2   ga_7
  CG   CD1 HD1   ga_36
  CG   CD1 NE1   ga_7
  HD1 CD1 NE1   ga_36
  CG   CD2 CE2   ga_7
  CG   CD2 CE3   ga_39
  CE2 CD2 CE3   ga_27
  CD1 NE1 HE1   ga_36
  CD1 NE1 CE2   ga_7
  HE1 NE1 CE2   ga_36
  CD2 CE2 NE1   ga_7
  CD2 CE2 CZ2   ga_27
  NE1 CE2 CZ2   ga_39
  CD2 CE3 HE3   ga_25
  CD2 CE3 CZ3   ga_27

220

```
HE3  CE3  CZ3    ga_25              H    H    0.31000    0
CE2  CZ2  HZ2    ga_25              CA   C    0.00000    1
CE2  CZ2  CH2    ga_27              CB   CH1  0.00000    1
HZ2  CZ2  CH2    ga_25              CG   C    0.00000    1
CE3  CZ3  HZ3    ga_25              CD1  C    -0.14000   2
CE3  CZ3  CH2    ga_27              HD1  HC   0.14000    2
HZ3  CZ3  CH2    ga_25              CD2  C    -0.14000   3
CZ2  CH2  CZ3    ga_27              HD2  HC   0.14000    3
CZ2  CH2  HH2    ga_25              CE1  C    -0.14000   4
CZ3  CH2  HH2    ga_25              HE1  HC   0.14000    4
 CA   C    O      ga_30              CE2  C    -0.14000   5
 CA   C    +N     ga_19              HE2  HC   0.14000    5
 O    C    +N     ga_33              CZ   C    -0.14000   6
[ impropers ]                       HZ   HC   0.14000    6
; ai  aj  ak  al  gromos type       C    C    0.450      7
  N   -C   CA   H     gi_1           O    O    -0.450     7
  CA   N    C    CB    gi_1         [ bonds ]
  CG  CD1  CD2   CB    gi_1           N    H    gb_2
  CG  CD1  NE1  CE2    gi_1          N    CA   gb_21
  CG  CD2  CE2  NE1    gi_1          CA   CB   gb_3
 CD1   CG  CD2  CE2    gi_1          CA   C    gb_27
 CD1   CG  NE1  HD1    gi_1          CB   CG   gb_27
 CD1  NE1  CE2  CD2    gi_1          CG   CD1  gb_16
 CD2   CG  CD1  NE1    gi_1          CG   CD2  gb_16
 CD2  CE2  CE3   CG    gi_1         CD1   HD1  gb_3
 CD2  CE2  CZ2  CH2    gi_1         CD1   CE1  gb_16
 CD2  CE3  CZ3  CH2    gi_1         CD2   HD2  gb_3
 NE1  CD1  CE2  HE1    gi_1         CD2   CE2  gb_16
 CE2  CD2  CE3  CZ3    gi_1         CE1   HE1  gb_3
 CE2  CD2  CZ2  NE1    gi_1         CE1   CZ   gb_16
 CE2  CZ2  CH2  CZ3    gi_1         CE2   HE2  gb_3
 CE3  CD2  CE2  CZ2    gi_1         CE2   CZ   gb_16
 CE3  CD2  CZ3  HE3    gi_1          CZ   HZ   gb_3
 CE3  CZ3  CH2  CZ2    gi_1          C    O    gb_5
 CZ2  CE2  CH2  HZ2    gi_1          C    +N   gb_10
 CZ3  CE3  CH2  HZ3    gi_1        [ exclusions ]
 CH2  CZ2  CZ3  HH2    gi_1        ; ai  aj
  C   CA   +N    O     gi_1          CB   HD1
[ dihedrals ]                        CB   HD2
; ai  aj  ak  al  gromos type        CB   CE1
 -CA  -C    N    CA   gd_14          CB   CE2
 -C    N    CA   C    gd_42          CG   HE1
;backbone dihedral, changed by Ying Xue Sep 29. 2009    CG   HE2
 -C    N    CA   C    gd_43          CG   CZ
;backbone dihedral, changed by Ying Xue Sep 29. 2009    CD1  HD2
  C   CA   CB   CG    gd_47         CD1   CE2
  N   CA   C    +N    gd_44         CD1   HZ
;backbone dihedral, changed by Ying Xue Sep 29. 2009    HD1  CD2
  N   CA   C    +N    gd_45          HD1  HE1
;backbone dihedral, changed by Ying Xue Sep 29. 2009    HD1  CZ
  CA  CB   CG   CD2   gd_40         CD2   CE1
                                    CD2   HZ
                                    HD2   HE2
## ; dehydro amino acids E position   HD2   CZ
                                    CE1   HE2
### [ EDP ]                          HE1   CE2
[ atoms ]                            HE1   HZ
  N    N   -0.31000    0             HE2   HZ
```

221

[angles]
; ai aj ak gromos type
 -C   N   H   ga_32
 -C   N   CA   ga_31
  H   N   CA   ga_18
  N   CA CB   ga_26
  N   CA C   ga_26
 CB   CA C   ga_26
 CA   CB CG   ga_15
 CB   CG CD1 ga_27
 CB   CG CD2 ga_27
 CD1 CG CD2 ga_27
 CG   CD1 HD1 ga_25
 CG   CD1 CE1 ga_27
 HD1 CD1 CE1 ga_25
 CG   CD2 HD2 ga_25
 CG   CD2 CE2 ga_27
 HD2 CD2 CE2 ga_25
 CD1 CE1 HE1 ga_25
 CD1 CE1 CZ   ga_27
 HE1 CE1 CZ   ga_25
 CD2 CE2 HE2 ga_25
 CD2 CE2 CZ   ga_27
 HE2 CE2 CZ   ga_25
 CE1 CZ CE2 ga_27
 CE1 CZ HZ   ga_25
 CE2 CZ HZ   ga_25
 CA   C   O   ga_30
 CA   C   +N   ga_19
  O   C   +N   ga_33

[impropers]
; ai aj ak al gromos type
  N   -C CA H   gi_1
 CA   N   C   CB   gi_1
 CA   CB N   C   gi_1
 CG   CD1 CD2 CB   gi_1
 CG   CD1 CE1 CZ   gi_1
 CG   CD2 CE2 CZ   gi_1
 CD1 CG CD2 CE2 gi_1
 CD1 CG CE1 HD1 gi_1
 CD1 CE1 CZ CE2 gi_1
 CD2 CG CD1 CE1 gi_1
 CD2 CG CE2 HD2 gi_1
 CD2 CE2 CZ CE1 gi_1
 HE1 CD1 CZ CE1 gi_1
 HE2 CD2 CZ CE2 gi_1
 CZ   CE1 CE2 HZ   gi_1
  C   CA +N O   gi_1

[dihedrals]
; ai aj ak al gromos type
 -CA -C N   CA   gd_14
 -C   N   CA C   gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N   CA C   gd_43
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  C   CA CB CG   gd_46
  N   CA C   +N   gd_44
;backbone dihedral, changed by Ying Xue Sep 29. 2009

  N   CA C   +N   gd_45
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 CA   CB CG CD1 gd_40

[EDA]
[atoms]
  N   N   -0.31000   0
  H   H   0.31000   0
 CA   C   0.00000   1
 CB   CH1 0.00000   1
 CG   CH3 0.00000   1
  C   C   0.450   2
  O   O   -0.450   2
[bonds]
  N   H   gb_2
  N   CA   gb_21
 CA   CB   gb_27
 CA   C   gb_27
 CB   CG   gb_27
  C   O   gb_5
  C   +N   gb_10
[angles]
; ai aj ak gromos type
 -C   N   H   ga_32
 -C   N   CA   ga_31
  H   N   CA   ga_18
  N   CA CB   ga_26
  N   CA C   ga_26
 CB   CA C   ga_26
 CA   CB CG   ga_15
 CA   C   O   ga_30
 CA   C   +N   ga_19
  O   C   +N   ga_33
[impropers]
; ai aj ak al gromos type
  N   -C CA H   gi_1
 CA   N   C   CB   gi_1
  C   CA +N O   gi_1
[dihedrals]
; ai aj ak al gromos type
 -CA -C N   CA   gd_14
 -C   N   CA C   gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
 -C   N   CA C   gd_43
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  C   CA CB CG   gd_46
  N   CA C   +N   gd_44
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N   CA C   +N   gd_45
;backbone dihedral, changed by Ying Xue Sep 29. 2009

[EDL]
[atoms]
  N   N   -0.31000   0
  H   H   0.31000   0
 CA   C   0.00000   1
 CB   CH1 0.00000   1
 C1   CH1 0.00000   2

```
  C2   CH3    0.00000    2                0    2
  C3   CH3    0.00000    2                CD2   C    0.00000    2
   C    C     0.450     3                 NE1   NR  -0.10000    2
   O    O    -0.450     3                 HE1   H    0.31000    2
[ bonds ]                                 CE2   C    0.00000    2
  N    H    gb_2                          CE3   C   -0.14000    3
  N    CA   gb_21                         HE3   HC   0.14000    3
  CA   CB   gb_27                         CZ2   C   -0.14000    4
  CA   C    gb_27                         HZ2   HC   0.14000    4
  CB   C1   gb_27                         CZ3   C   -0.14000    5
  C1   C2   gb_27                         HZ3   HC   0.14000    5
  C1   C3   gb_27                         CH2   C   -0.14000    6
  C    O    gb_5                          HH2   HC   0.14000    6
  C    +N   gb_10                          C    C    0.450     7
[ angles ]                                 O    O   -0.450     7
; ai  aj   ak  gromos type               [ bonds ]
 -C    N    H     ga_32                    N    H    gb_2
 -C    N    CA    ga_31                    N    CA   gb_21
  H    N    CA    ga_18                    CA   CB   gb_27
  N    CA   CB    ga_26                    CA   C    gb_27
  N    CA   C     ga_26                    CB   CG   gb_27
  CB   CA   C     ga_26                    CG   CD1  gb_10
  CA   CB   C1    ga_26                    CG   CD2  gb_16
  CB   C1   C2    ga_15                    CD1  HD1  gb_3
  CB   C1   C3    ga_15                    CD1  NE1  gb_10
  C2   C1   C3    ga_15                    CD2  CE2  gb_16
  CA   C    O     ga_30                    CD2  CE3  gb_16
  CA   C    +N    ga_19                    NE1  HE1  gb_2
  O    C    +N    ga_33                    NE1  CE2  gb_10
[ impropers ]                             CE2  CZ2  gb_16
; ai  aj   ak   al  gromos type           CE3  HE3  gb_3
  N   -C    CA   H    gi_1                 CE3  CZ3  gb_16
  CA   N    C    CB   gi_1                 CZ2  HZ2  gb_3
  C1   CB   C3   C2   gi_2                 CZ2  CH2  gb_16
  C    CA   +N   O    gi_1                 CZ3  HZ3  gb_3
[ dihedrals ]                             CZ3  CH2  gb_16
; ai  aj   ak   al  gromos type           CH2  HH2  gb_3
 -CA  -C    N    CA   gd_14                 C    O    gb_5
 -C    N    CA   C    gd_42                 C    +N   gb_10
;backbone dihedral, changed by Ying Xue Sep 29. 2009    [ exclusions ]
 -C    N    CA   C    gd_43                ; ai  aj
;backbone dihedral, changed by Ying Xue Sep 29. 2009    CB   HD1
  C    CA   CB   C1   gd_46                CB   NE1
  N    CA   C    +N   gd_44                CB   CE2
;backbone dihedral, changed by Ying Xue Sep 29. 2009    CB   CE3
  N    CA   C    +N   gd_45                CG   HE1
;backbone dihedral, changed by Ying Xue Sep 29. 2009    CG   HE3
  CA   CB   C1   C2   gd_34                CG   CZ2
                                          CG   CZ3
                                          CD1  CE3
[ EDT ]                                   CD1  CZ2
  [ atoms ]                               HD1  CD2
   N    N   -0.31000    0                 HD1  HE1
   H    H    0.31000    0                 HD1  CE2
   CA   C    0.00000    1                 CD2  HE1
   CB   CH1  0.00000    1                 CD2  HZ2
   CG   C   -0.21000    2                 CD2  HZ3
   CD1  C   -0.14000    2                 CD2  CH2
   HD1  HC   0.1400
```

```
NE1  CE3                          HZ3  CZ3  CH2    ga_25
NE1  HZ2                          CZ2  CH2  CZ3    ga_27
NE1  CH2                          CZ2  CH2  HH2    ga_25
HE1  CZ2                          CZ3  CH2  HH2    ga_25
CE2  HE3                          CA   C    O      ga_30
CE2  CZ3                          CA   C    +N     ga_19
CE2  HH2                          O    C    +N     ga_33
CE3  CZ2                       [ impropers ]
CE3  HH2                       ; ai  aj  ak  al gromos type
HE3  HZ3                          N   -C   CA   H     gi_1
HE3  CH2                          CA   N    C    CB    gi_1
CZ2  HZ3                          CG   CD1  CD2  CB    gi_1
HZ2  CZ3                          CG   CD1  NE1  CE2   gi_1
HZ2  HH2                          CG   CD2  CE2  NE1   gi_1
HZ3  HH2                          CD1  CG   CD2  CE2   gi_1
[ angles ]                        CD1  CG   NE1  HD1   gi_1
; ai   aj   ak  gromos type       CD1  NE1  CE2  CD2   gi_1
 -C   N    H    ga_32             CD2  CG   CD1  NE1   gi_1
 -C   N    CA   ga_31             CD2  CE2  CE3  CG    gi_1
  H   N    CA   ga_18             CD2  CE2  CZ2  CH2   gi_1
  N   CA   CB   ga_26             CD2  CE3  CZ3  CH2   gi_1
  N   CA   C    ga_26             NE1  CD1  CE2  HE1   gi_1
 CB   CA   C    ga_26             CE2  CD2  CE3  CZ3   gi_1
 CA   CB   CG   ga_26             CE2  CD2  CZ2  NE1   gi_1
 CB   CG   CD1  ga_37             CE2  CZ2  CH2  CZ3   gi_1
 CB   CG   CD2  ga_37             CE3  CD2  CE2  CZ2   gi_1
 CD1  CG   CD2  ga_7              CE3  CD2  CZ3  HE3   gi_1
 CG   CD1  HD1  ga_36             CE3  CZ3  CH2  CZ2   gi_1
 CG   CD1  NE1  ga_7              CZ2  CE2  CH2  HZ2   gi_1
 HD1  CD1  NE1  ga_36             CZ3  CE3  CH2  HZ3   gi_1
 CG   CD2  CE2  ga_7              CH2  CZ2  CZ3  HH2   gi_1
 CG   CD2  CE3  ga_39             C    CA   +N   O     gi_1
 CE2  CD2  CE3  ga_27          [ dihedrals ]
 CD1  NE1  HE1  ga_36          ; ai  aj  ak  al gromos type
 CD1  NE1  CE2  ga_7            -CA  -C   N    CA   gd_14
 HE1  NE1  CE2  ga_36           -C   N    CA   C    gd_42
 CD2  CE2  NE1  ga_7           ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 CD2  CE2  CZ2  ga_27           -C   N    CA   C    gd_43
 NE1  CE2  CZ2  ga_39          ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 CD2  CE3  HE3  ga_25           C    CA   CB   CG   gd_46
 CD2  CE3  CZ3  ga_27           N    CA   C    +N   gd_44
 HE3  CE3  CZ3  ga_25          ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 CE2  CZ2  HZ2  ga_25           N    CA   C    +N   gd_45
 CE2  CZ2  CH2  ga_27          ;backbone dihedral, changed by Ying Xue Sep 29. 2009
 HZ2  CZ2  CH2  ga_25           CA   CB   CG   CD2  gd_40
 CE3  CZ3  HZ3  ga_25
 CE3  CZ3  CH2  ga_27
```

APPENDIX V

FF G54a7 Parameters

GROMOS bond-stretching parameters

- Bond type code

- Force constant

- Ideal bond length

- Examples of usage in terms of non-bonded atom types

- This file has been change by Ying Xue Sep, 29 2009

- **This file has been changed by Castro and Micâelo, 2012-2015**

ICB(H)[N]   CB[N] B0[N]

#define gb_1   0.1000 1.5700e+07
; H - OA   750
;
#define gb_2   0.1000 1.8700e+07
; H - N (all) 895
;
#define gb_3   0.1090 1.2300e+07
; HC - C   700
;
#define gb_4   0.112 3.7000e+07
; C - O (CO in heme) 2220
;
#define gb_5   0.1230 1.6600e+07
; C - O   1200
;
#define gb_6   0.1250 1.3400e+07
; C - OM   1000
;
#define gb_7   0.1320 1.2000e+07
; CR1 - NR (6-ring)   1000
;
#define gb_8   0.1330 8.8700e+06
; H - S   750
;
#define gb_9   0.1330 1.0600e+07
; C - NT, NL 900
;
#define gb_10   0.1330 1.1800e+07
; C, CR1 - N, NR, CR1, C (peptide, 5-ring)
1000
;

#define gb_11   0.1340 1.0500e+07
; C - N, NZ, NE   900
;
#define gb_12   0.1340 1.1700e+07
; C - NR (no H) (6-ring)   1000
;
#define gb_13   0.1360 1.0200e+07
; C - OA   900
;
#define gb_14   0.1380 1.1000e+07
; C - NR (heme)   1000
;
#define gb_15   0.1390 8.6600e+06
; CH2 - C, CR1 (6-ring)   800
;
#define gb_16   0.1390 1.0800e+07
; C, CR1 - CH2, C, CR1 (6-ring)   1000
;
#define gb_17   0.1400 8.5400e+06
; C, CR1, CH2 - NR (6-ring)   800
;
#define gb_18   0.1430 8.1800e+06
; CHn - OA   800
;
#define gb_19   0.1430 9.2100e+06
; CHn - OM   900
;
#define gb_20   0.1435 6.1000e+06
; CHn - OA (sugar)   600
;

```
#define gb_21     0.1470 8.7100e+06
; CHn - N, NT, NL, NZ, NE    900
;
#define gb_22     0.1480 5.7300e+06
; CHn - NR (5-ring)   600
;
#define gb_23     0.1480 7.6400e+06
; CHn - NR (6-ring)  800
;
#define gb_24     0.1480 8.6000e+06
; O, OM - P   900
;
#define gb_25     0.1500 8.3700e+06
; O - S    900
;
#define gb_26     0.1520 5.4300e+06
; CHn - CHn (sugar)  600
;
#define gb_27     0.1530 7.1500e+06
; C, CHn - C, CHn   800
;
#define gb_28     0.1610 4.8400e+06
; OA - P   600
;
#define gb_29     0.1630 4.7200e+06
; OA - SI   600
;
#define gb_30     0.1780 2.7200e+06
; FE - C (Heme)
;
#define gb_31     0.1780 5.9400e+06
; CH3 - S   900
;
#define gb_32     0.1830 5.6200e+06
; CH2 - S   900
;
#define gb_33     0.1870 3.5900e+06
; CH1 - SI   600
;
#define gb_34     0.198 0.6400e+06
; NR - FE   120
;
#define gb_35     0.200 0.6280e+06
; NR (heme) - FE   120
;
#define gb_36     0.2040 5.0300e+06
; S - S    1000
;
```

```
#define gb_37     0.221 0.5400e+06
; NR - FE    126
;
#define gb_38     0.1000 2.3200e+07
; HWat - OWat    1110
;
#define gb_39     0.1100 1.2100e+07
; HChl - CChl    700
;
#define gb_40     0.1758 8.1200e+06
; CChl - CLChl    1200
;
#define gb_41     0.1530 8.0400e+06
; ODmso - SDmso    900
;
#define gb_42   0.193799 4.9500e+06
; SDmso - CDmso    890
;
#define gb_43     0.1760 8.1000e+06
; CCl4 - CLCl4    1200
;
#define gb_44     0.1265 1.3100e+07
; CUrea - OUrea     1000
;
#define gb_45      0.135 1.0300e+07
; CUrea - NUrea    900
;
#define gb_46   0.163299 8.7100e+06
; HWat - HWat    1110
;
#define gb_47   0.233839 2.6800e+06
; HChl - CLChl    700
;
#define gb_48   0.290283 2.9800e+06
; CLChl - CLChl    1200
;
#define gb_49   0.279388 2.3900e+06
; ODmso - CDmso    890
;
#define gb_50   0.291189 2.1900e+06
; CDmso - CDmso    890
;
#define gb_51     0.2077 3.9700e+06
; HMet - CMet    820
;
#define gb_52   0.287407 3.0400e+06
; CLCl4 - CLCl4    1200
;
```

#define gb_53   0.1430 8.1800e+06
;parameter ATB PEG2 - N-C1
;
#define gb_54   0.1520 5.4300e+06
;parameter ATB PEG2 - C1-C2 or C3-C4
;

#define gb_55   0.1435 6.1000e+06
;parameter ATB PEG2 - C2-O2 or C4-C5
;
#define gb_56   0.1000 2.3200e+07
;parameter ATB PEG2 - O2-C3

GROMOS bond-angle bending parameters

- Bond-angle type code

- Force constant

- Ideal bond angle

- Example of usage in terms of non-bonded atom types

ICT(H)[N] CT[N] (TO[N])

#define ga_1   90.00   380.00
; NR(heme) - FE - C   90
;
#define ga_2   90.00   420.00
; NR(heme) - FE - NR(heme) 100
;
#define ga_3   96.00   405.00
; H - S - CH2   95
;
#define ga_4   100.00   475.00
; CH2 - S - CH3   110
;
#define ga_5   103.00   420.00
; OA - P - OA   95
;
#define ga_6   104.00   490.00
; CH2 - S - S   110
;
#define ga_7   108.00   465.00
; NR, C, CR1(5-ring)   100
;
#define ga_8   109.50   285.00
; CHn - CHn - CHn, NR(6-ring) (sugar) 60
;
#define ga_9   109.50   320.00
; CHn, OA - CHn - OA, NR(ring) (sugar)
68
;

#define ga_10   109.50   380.00
; H - NL, NT - H, CHn - OA - CHn(sugar)
80
;
#define ga_11   109.50   425.00
; H - NL - C, CHn   H - NT - CHn 90
;
#define ga_12   109.50   450.00
; X - OA, SI - X   95
;
#define ga_13   109.50   520.00
; CHn,C - CHn - C, CHn, OA, OM, N, NE
110
;
#define ga_14   109.60   450.00
; OM - P - OA   95
;
#define ga_15   111.00   530.00
; CHn - CHn - C, CHn, OA, NR, NT, NL
110
;
#define ga_16   113.00   545.00
; CHn - CH2 - S   110
;
#define ga_17   115.00   50.00
; NR(heme) - FE - NR 10
;

```
#define ga_18      115.00    460.00
; H - N - CHn     90
;
#define ga_19      115.00    610.00
; CHn, C - C - OA, N, NT, NL      120
;
#define ga_20      116.00    465.00
; H - NE - CH2    90
;
#define ga_21      116.00    620.00
; CH2 - N - CH1    120
;
#define ga_22      117.00    635.00
; CH3 - N - C, CHn - C - OM      120
;
#define ga_23      120.00    390.00
; H - NT, NZ, NE - C      70
;
#define ga_24      120.00    445.00
; H - NT, NZ - H    80
;
#define ga_25      120.00    505.00
; H - N - CH3, H, HC - 6-ring, H - NT - CHn
90
;
#define ga_26      120.00    530.00
; P, SI - OA - CHn, P      95
;
#define ga_27      120.00    560.00
; N, C, CR1 (6-ring, no H)      100
;
#define ga_28      120.00    670.00
; NZ - C - NZ, NE   120
;
#define ga_29      120.00    780.00
; OM - P - OM      140
;
#define ga_30      121.00    685.00
; O - C - CHn, C       CH3 - N - CHn 120
;
#define ga_31      122.00    700.00
; CH1, CH2 - N - C  120
;
#define ga_32      123.00    415.00
; H - N - C  70
;
#define ga_33      124.00    730.00
```

```
; O  - C  - OA, N, NT, NL   C - NE - CH2
120
;
#define ga_34      125.00    375.00
; FE - NR - CR1 (5-ring)      60
;
#define ga_35      125.00    750.00
; -    120
;
#define ga_36      126.00    575.00
; H, HC  - 5-ring      90
;
#define ga_37      126.00    640.00
; X(noH)  - 5-ring      100
;
#define ga_38      126.00    770.00
; OM  - C - OM 120
;
#define ga_39      132.00    760.00
; 5, 6 ring connnection 100
;
#define ga_40      155.00    2215.00
; SI - OA - SI      95
;
#define ga_41      180.00    91350.00
; Fe - C - O (heme) 57
;
#define ga_42      109.50    434.00
; HWat - OWat - HWat  92
;
#define ga_43      107.57    484.00
; HChl - CChl - CLChl 105
;
#define ga_44      111.30    632.00
; CLChl - CChl - CLChl      131
;
#define ga_45      97.40     469.00
; CDmso - SDmso - CDmso      110
;
#define ga_46      106.75    503.00
; CDmso - SDmso - ODmso      110
;
#define ga_47      108.53    443.00
; HMet - OMet - CMet  95
;
#define ga_48      109.50    618.00
; CLCl4 - CCl4 - CLCl4      131
;
```

```
#define ga_49     107.60   507.00        #define ga_53     117.2   636.00
; FTFE - CTFE - FTFE     100             ; NUrea - CUrea - NUrea    120
;                                        ;
#define ga_50     109.50   448.00        #define ga_54     121.4   690.00
; HTFE - OTFE - CHTFE     85             ; OUrea - CUrea - NUrea    120
;                                        ;
#define ga_51     110.3    524.00        #define ga_55     60.00   520.00
; OTFE - CHTFE - CTFE     97             ; cyclopropane-ring    100
;                                        ; Tarsila
#define ga_52     111.4    532.00        #define ga_56     88.00   520.00
; CHTFE - CTFE - FTFE     95             ; cyclobutane-ring     100
;                                        ; Tarsila
;
```

GROMOS improper (harmonic) dihedral angle parameters

- Improper dihedral-angle type code

- Force constant

- Ideal improper dihedral angle

- Example of usage

ICQ(H)[N] CQ[N] (Q0[N])


```
#define gi_1      0.0   167.42309         #define gi_4      180.0   167.42309
; planar groups 40                       ; Planar Groups (Alan Mark -ref- bvictor 29
;                                        November 2010
#define gi_2    35.26439   334.84617      ;
; tetrahedral centres   80               #define gi_5    -35.26439   334.84617
;                                        ; Tetrahedral Groups (Alan Mark -ref- bvictor 29
#define gi_3      0.0   669.69235         November 2010
; heme iron    160                       ;
;
```

GROMOS (trigonometric) dihedral torsional angle parameters

- Dihedral-angle type code

- Force constant

- Phase shift

- Multiplicity

- Example of usage in terms of non-bonded atom types

ICP(H)[N] CP[N] PD[N] NP[N]

```
#define gd_1    180.000    2.67       1          ; -CH1(sugar)-NR(base)  0.0
; CHn-CHn-CHn-OA (sugar) 0.6                      ;
;                                                 #define gd_17   0.000    0.418      2
#define gd_2    180.000    3.41       1          ; O-CH1-CHn-no O      0.1
; OA-CHn-OA-CHn,H (beta sugar) 0.8                ;
;                                                 #define gd_18   0.000    2.09       2
#define gd_3    180.000    4.97       1          ; O-CH1-CHn-O  0.5
; OA-CHn-CHn-OA (sugar)      1.2                   ;
;                                                 #define gd_19   0.000    3.14       2
#define gd_4    180.000    5.86       1          ; -OA-P- 0.75
; N-CHn-CHn-OA (lipid)   1.4                       ;
;                                                 #define gd_20   0.000    5.09       2
#define gd_5    180.000    9.35       1          ; O-P-O- (dna, lipids)    1.2
; OA-CHn-CHn-OA (sugar)      2.2                   ;
;                                                 #define gd_21   0.000    16.7       2
#define gd_6    180.000    9.45       1          ; -S-S-   4.0
; OA-CHn-OA-CHn,H (alpha sugar) 2.3                ;
;                                                 #define gd_22   0.000    1.05       3
#define gd_7    0.000      2.79       1          ; -OA-P- 0.25
; P-O5*-C5*-C4* (dna)   0.7                        ;
;                                                 #define gd_23   0.000    1.26       3
#define gd_8    0.000      5.35       1          ; -CHn-OA(no sugar)- 0.3
; O5*-C5*-C4*-O4* (dna)1.3                         ;
;                                                 #define gd_24   0.000    1.30       3
#define gd_9    180.000    1.53       2          ; HTFE-OTFE-CHTFE-CTFE      0.3
; C1-C2-CAB-CBB       (heme) 0.4                   ;
;                                                 #define gd_25   0.000    2.53       3
#define gd_10   180.000    5.86       2          ; O5*-C5*-C4*-O4* (dna)0.6
; -C-C-   1.4                                      ;
;                                                 #define gd_26   0.000    2.93       3
#define gd_11   180.000    7.11       2          ; -CH2-S-        0.7
; -C-OA,OE- (at ring)     1.7                      ;
;                                                 #define gd_27   0.000    3.19       3
#define gd_12   180.000    16.7       2          ; O-P-O- (dna, lipids)    0.8
; -C-OA,OE- (carboxyl)   4.0                       ;
;                                                 #define gd_28   0.000    3.65       3
#define gd_13   180.000    24.0       2          ; OA-CHn-OA-CHn,H (alpha sugar)      0.9
; CHn-OE-C-CHn (ester lipid)   5.7                 ;
;                                                 #define gd_29   0.000    3.77       3
#define gd_14   180.000    33.5       2          ; -C,CHn,SI-     0.9
; -C-N,NT,NE,NZ,NR-    8.0                          ;
;                                                 #define gd_30   0.000    3.90       3
#define gd_15   180.000    41.8       2          ; CHn-CHn-OA-H (sugar) 0.9
; -C-CR1- (6-ring)          10.0                   ;
;                                                 #define gd_31   0.000    4.18       3
#define gd_16   0.000      0.0        2          ; HC-C-S-        1.0
```

232

```
;
#define gd_32      0.000        4.69         3
; AO-CHn-OA-CHn,H (beta sugar)
;
#define gd_33      0.000        5.44         3
; HC-C-C-         1.3
;
#define gd_34      0.000        5.92         3
; -CHn,SI-CHn-  1.4
;
#define gd_35      0.000        7.69         3
; OA-CHn-CHn-OA (sugar)         1.8
;
#define gd_36      0.000        8.62         3
; N-CHn-CHn-OA (lipid)   2.1
;
#define gd_37      0.000        9.50         3
; OA-CHn-CHn-OA (sugar)         2.3
;
#define gd_38      0.000         0.0         4
; -NR-FE-          0.0
;
#define gd_39   180.000         1.0          6
; -CHn-N,NE-     0.24
;
#define gd_40      0.000        1.0          6
; -CHn-C,NR(ring), CR1- 0.24
;
#define gd_41      0.000        3.77         6
; -CHn-NT-        0.9
;
```

; Below are the changes made by Ying Xue, Sep 29, 2009

#define gd_42   0.000   2.8   3
; Backbone dihedral angle -C-N-CA-C-   0.67
;
#define gd_43   180.000   0.7   6
; Backbone dihedral angle -C-N-CA-C-   0.17
;
#define gd_44   180.000   3.5   2
; Backbone dihedral angle -N-CA-C-N-   0.84
;
#define gd_45   0.000   0.4   6
; Backbone dihedral angle -N-CA-C-N-   0.096
;
; Dihedrals for dehydro amino acids (double bond)
#define gd_46   180.000   53.50   1
; C-CA-CB-CG   E:CG cis C   1.3
;
#define gd_47   0.000   53.50   1
; C-CA-CB-CG   Z:CG trans C   1.3
;
;
#define gd_49   0.00   5.92   3
; O1-C1-C2-O2

; get the constraint distances for dummy atom constructions

#include "ff_dum.itp"

[constrainttypes]

; now the constraints for the rigid NH3 groups
 MNH3   C   2   DC_MNC1
 MNH3 CH1   2   DC_MNC2
 MNH3 CH2   2   DC_MNC2
 MNH3 MNH3   2   DC_MNMN
; and the angle-constraints for OH and SH groups in proteins:
 CH2   H   2   DC_CO
 CH1   H   2   DC_CO
  C   H   2   DC_CO
  P   H   2   DC_PO

; bond-, angle- and dihedraltypes for specbonds:
[bondtypes]
S   S   2   gb_36
NR   FE   2   gb_34

```
[ angletypes ]
CH1   CH2   S    2   ga_16
CH2   S     S    2   ga_6
CR1   NR    FE   2    ga_34
NR    FE    NR   2   ga_17

[ dihedraltypes ]
S    S    1   gd_21
NR    FE    1   gd_38
CH2   S    1   gd_26
```

APPENDIX VI

Conformational Properties of the Non-canonical Cyclic Ac_nc Amino Acids:

A Molecular Modeling Study

Abstract

The α-helix and 3_{10}-helix folding properties of a series of non-canonical cyclic amino acids, Ac_3c, Ac_4c, Ac_5c, Ac_6c, (S,S)-Ac_5c^{dOM} and (R,R)-Ac_5c^{dOM}, were studied using molecular modeling methodologies. The helical propensity of these residues was evaluated using leucine-based, hexa and nonapeptides. The secondary structure properties of the peptides incorporating cyclic and non-cyclic α,α-disubstituted amino acids were investigated in water, chloroform and in trifluoroethanol/water mixture. We show that, in water, leucine nonapeptides carrying Ac_5c and (R,R)-Ac_5c^{dOM} residues show a high tendency to form α-helical secondary structures. The number of residues in α-helix was found also to change as a function of the solvent. In chloroform, residues Ac_5c, Ac_6c, (S,S)-Ac_5c^{dOM} and (R,R)-Ac_5c^{dOM} induced the formation of 3_{10}-helices, in agreement with previous experimental reports. The TFE/H_2O (50/50 v/v) mixture increases the population of α-helical secondary structure for the hexapeptides, relative to the aqueous media. In summary, we show that some of the non-canonical amino acids under study are strong helical inducers of our model peptides and, this effect is also dependent on the peptide size and solvent environment.

(R,R) Ac_5c^{dOM} α-helical nonapeptide in H_2O

(R,R) Ac_5c^{dOM} 3_{10}-helical nonapeptide in $CHCL_3$
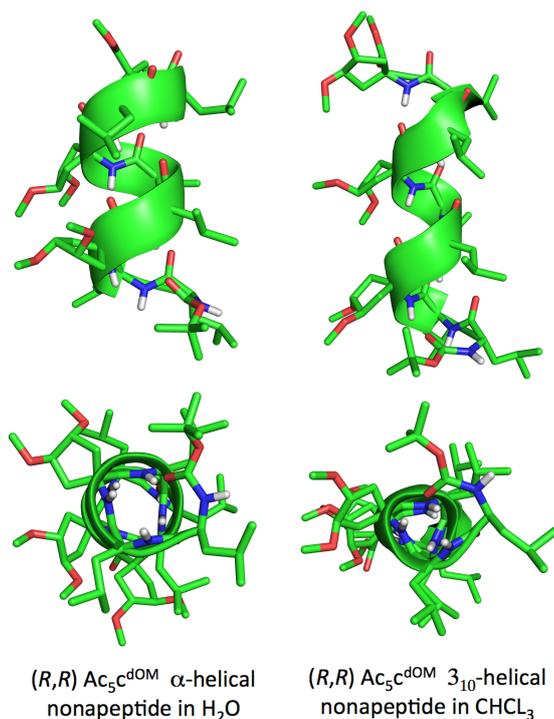
1.Introduction

Non-canonical constrained amino acids are being used for the design of novel peptidomimetics in drug discovery (Gentilucci et al. 2006; Grauer and Konig 2009; Giannis 1993; Vagner et al. 2008). The incorporation of constrained amino acids into peptides is a promising approach to induce well-defined and stable secondary structure (SS) (Toniolo et al. 2001; Hill et al. 2001; Goodman et al. 2007). In fact, constrained residues have been used as building blocks, with the goal to improve the global structural stability and to optimize peptide function (Ballet et al. 2011; Whitby et al. 2011; Mallareddy et al. 2011; Feytens et al. 2007; Ressurreicao et al. 2008; Oh and Lee 1999). Other advantages of using constrained amino acids are the improvement on the bioavailability and stability in physiological conditions (Balaram 1992; Toniolo et al. 2001). Also, this type of amino acids has been extensively used on the synthesis of therapeutic peptides to prevent proteolytic degradation in vivo (Balaram 1992; Karle et al. 1990; Oh and Lee 1999).

An important class of non-canonical constrained amino acids, the α,α-disubstituted amino acids (dAAs), has been designed and incorporated into known peptides and proteins (Bürgi et al. 2001; Prasad et al. 2006). The α-amino isobutyric acid (Aib) is a well-known residue, largely investigated, and the prototype of this class (Marshall and Bosshard 1972; Marshall et al. 1990). Aib induces well defined different SS in peptides, namely β-bend (Rose et al. 1985; Venkatachalam 1968) and $3_{10}/\alpha$-helix (Marshall and Bosshard 1972; Marshall et al. 1990; Toniolo and Benedetti 1991), according to the chain length (Venkatraman et al. 2001; Mendel et al. 1993; Toniolo et al. 2001). The lack of chirality and the geometrical constrain around the Cα atom as a result of the double substitution at this position, are ultimately responsible for these observations. Using this rational, we address in this study non-canonical constrained amino acids that are also highly constrained at the Cα position and present consequently similar folding properties as the Aib residue: the cyclic Ac$_n$c (1-aminocycloalkane-1-carboxylic acids) residues, where n refers to the size of the cycle.

The Ac$_n$c amino acids are the result of the Cα to Cα cyclization of symmetrical α,α-disubstituted amino acids (Benedetti et al. 1997; Toniolo 1990). The cyclization process generates residues with even more restricted conformational flexibility than Aib or its analogs (Alemán 1997; Zanuy et al. 2009). Previous experimental and theoretical results indicate that the Ac$_n$c with cycles with more than 3 atoms (n = 4–12) explore, mostly, a main chain geometry similar to Aib (ϕ , ϕ $\approx \pm 60°$, $\pm 30°$) which is typical of 3_{10}-helix or α-helix SS (Ballano et al. 2008; Benedetti et al. 1997; Gatos et al. 1997a;

Gatos et al. 1997b; Moretto et al. 2001; Santini et al. 1996; Saviano et al. 2000a; Saviano et al. 2000b). The residues Ac$_5$c (1-aminocyclopentane-1-carboxylic acid) and Ac$_6$c (1-aminocyclohexane-1-carboxylic acid) have been found to originate γ-turn conformations in small peptides (Aschi et al. 2003; Paradisi et al. 1995). On the other hand, Ac$_3$c (1-aminocyclopropane-1-carboxylic acid) is the only member of Ac$_n$c family that prefers molecular geometries on the bridge region (ϕ, ψ $\approx$ $\pm90°$, $0°$) and this particularity (Zimmerman et al. 1977; Aschi et al. 2003; Rodriguez-Ropero et al. 2008; Alemán 1997; Zanuy et al. 2009) has been the subject of experimental and theoretical studies over the past decades (Ballano et al. 2008; Crisma et al. 1989; Headley et al. 2003; Jiménez et al. 2011; Zimmerman et al. 1977; Gomez-Catalan et al. 2000).
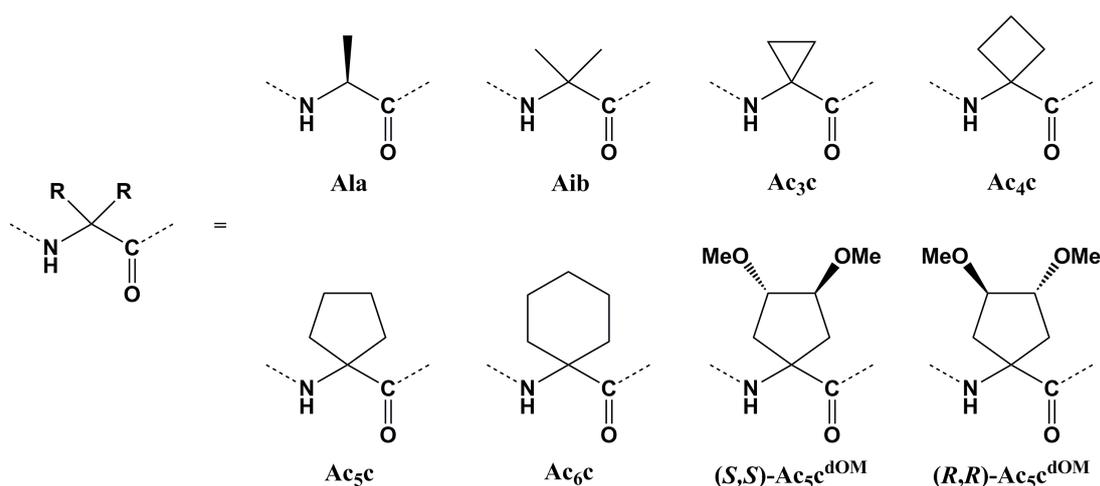
Figure 1. Two-dimensional structures of Ala and the non-canonical dAAs under study: Aib, Ac$_3$c, Ac$_4$c, Ac$_5$c, Ac$_6$c, (*S,S*)-Ac$_5$c^{dOM} and (*R,R*)-Ac$_5$c^{dOM}.

Mendel and co-workers (Mendel et al. 1993) reported in 1993 the protein biosynthesis with conformationally restricted amino acids, including the dAAs: Aib, Ac$_3$c, Ac$_4$c (1-aminocyclobutanecarboxylic acid), Ac$_5$c and Ac$_6$c. Recently, Demizu and his group performed experimental conformational studies on peptides containing Ac$_5$c, (Demizu et al. 2011; Demizu et al. 2010) and the chiral disubstituted forms (*S,S*)-Ac$_5$c^{dOM} and (R,R)-Ac$_5$c^{dOM}, and reported their capability to induce α-helices and 3$_{10}$-helices. The aim of this work is to study peptides incorporating these cyclic dAAs in aqueous and non-aqueous media, to compute their intrinsic folding properties in order to asses how these dAAs can be used in the design of peptides with a specific SS. In this sense we studied two sets of peptides: the eight peptides investigated by Demizu (Demizu et al. 2011; Demizu et al. 2010) and eight new peptides analogues incorporating a new series of cycloaliphatic residues.

2. Materials and Methods

2.1 Non-canonical amino acid force field parameters

The molecular structure of the α,α-disubstituted amino acids investigated in this study was designed with PyMol (Schrödinger 2010). The dAAs are not parameterized in the GROMOS force field. The parameters for the new, non-canonical dAAs (bonded and non-bonded terms) were based on the equivalent encoded amino acids present in the GROMOS 54a7 force field (FF) (Huang et al. 2011; Schmid et al. 2011).

For some cycloalkanes, the angle parameters were adjusted to reproduce the geometry of these cyclic structures. In addition, the N-terminal of the hexa and nonapeptides, the protecting groups benzyloxycarbonyl (Cbz) and tert-butyl carbamate (Boc), respectively, were also parameterized. Topology files and further detail for the new parameters can be found in the Supporting Information (SI).

2.2 System preparation

The initial geometry for all peptides (Figure 1) corresponded to a fully extended (φ,ψ = 180°, 180°), non-helical conformation. An extended conformation was adopted for all solvents to avoid any bias in the SS populations. For this we built extended conformations of the hexa and nonapeptides synthesized by Demizu et al.,(Demizu et al. 2011) Cbz-(L-Leu-L-Leu-dAA)$_2$-OMe and Boc-(L-Leu-L-Leu-dAA)$_3$-OMe, respectively, where dAA is a α , α -disubstituted amino acid: Aib, Ac$_5$c, (S,S)-Ac$_5$c^{dOM} or (R,R)-Ac$_5$cdOM. In addition, we created a new set of peptidomimetics, by replacing the dAAs positions for: Ala, Ac$_3$c, Ac$_4$c and Ac$_6$c. In total, we studied 8 hexapeptides and 8 nonapeptides with two control peptides: peptides with the Ala (canonical amino acid) and peptides with the Aib residue (non-canonical non-cyclic amino acid). All peptides were modeled in three solvents, which were also studied experimentally with some of these systems: water, a mixture (50/50 v/v) of trifluoroethanol (TFE) and water, and chloroform (CHCL$_3$) (Demizu et al. 2011). Hexapeptides were simulated in 4x4x4 (nm) cubic boxes of solvent. These boxes contained 2000-2200 water molecules, 220-230 TFE molecules and 990 water molecules, and 450-500 molecules of CHCL$_3$. The nonapeptides were solvated in water using octahedral boxes with 2000-2300 water molecules while 5x5x5 (nm) cubic boxes were used for the other solvents. The boxes contained 400-450 molecules of TFE and 2000-2200 water molecules, and 700-800 molecules of CHCL$_3$. The solvated boxes of CHCL$_3$ and TFE/H2O were made with PACKMOL (Martinez et al. 2009). The peptides were modeled in water with the simple point charge (SPC) water model.

2.3 Molecular Dynamics Simulations

All simulations were performed using GROMACS 4.5.4 (Lindahl et al. 2001; Bekker et al. 1993; Spoel et al. 2010). For the treatment of long-range interactions, we used the Reaction Field method, with 1.4 nm cut-off and, for consistency, a dielectric constant of 54 for water (Smith and Vangunsteren 1994; Berendsen et al. 1987), 52 to TFE/H2O and 4.81 for $CHCL_3$. Van der Waals interactions were also truncated with a twin-range cut-off of 0.8 and 1.4 nm. The algorithm LINCS (Hess et al. 1997; Hess 2008) was used to constrain the chemical bonds of the peptides and the algorithm SETTLE (van der Spoel et al. 1998) in the case of water. The pressure and temperature Berendsen algorithms were used to control the temperature and pressure at 310K and 1 atm, respectively (Berendsen et al. 1984). In all solvents $\tau_T = 0.2$ ps and $\tau_P = 1.0$ ps were used for the Berendsen temperature and pressure coupling parameter respectively. One stage of energy minimization was performed using a maximum of 12000 steps with a steepest descent algorithm. All the systems (peptide in water, TFE/H2O and $CHCL_3$) were sampled using 200 ns molecular dynamics simulations with an integration interval of 2 fs.

3. Results and Discussion

We investigated six cyclic non-canonical amino acids: Ac_3c, Ac_4c, Ac_5c, Ac_6c, (S,S)-Ac_5c^{dOM} and (R,R)-Ac_5c^{dOM}. In addition, we also studied Ala, as a reference for canonical amino acids, and Aib, as non-canonical non-cyclic reference of amino acids. The structural formula of all investigated amino acids is presented in Figure 1. Figure 2A-B shows the 2-dimensional structures and sequence of the hexa and nona peptides studied. On the hexapeptide (Figure 2A), positions 3 and 6 were replaced by the amino acids under study whereas in the nonapeptide (Figure 2B), the replaced positions were 3, 6 and 9.

A

1   2   3   4   5   6
Cbz-L-Leu-L-Leu-dAA-L-Leu-L-Leu-dAA-O-Me

B

1   2   3   4   5   6   7   8   9
Boc-L-Leu-L-Leu-dAA-L-Leu-L-Leu-dAA-L-Leu-L-Leu-dAA-O-Me
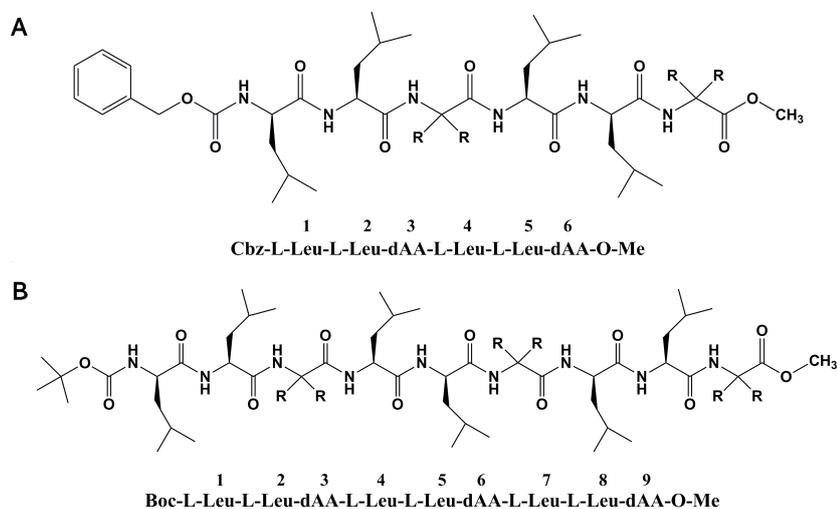
Figure 2. Two-dimensional sructure and sequence of the (A) hexa and (B) nonapeptide studied in this work.

3.1 Cα,α-disubstituted glycines that induce helical SS in water

Figure 3 shows the percentage of conformations with helical SS observed for the hexa and nonapeptide in the three solvents through the simulation computed using the DSSP (Dictionary of Secondary Structure in Proteins) method (Hess et al. 2008) .Two type of SS statistics were computed: in the first one (Figure 3), we count the number of conformations involving a minimum number of residues for each SS type, namely: 3, 4 or 5 for 3_{10}-helix, α-helix or π-helix (5-helix), respectively. Then we normalize this value by the total number of frames analyzed. For the second percentage presented (Figure 5), we count how many times the same residue had a specific type of conformation, and we normalize this value by the total number of frames in the simulation.

First of all, there are negligible conformations for the π-helix (< 1.7 %) in our peptides, indicating that is not a typical SS for the cyclized amino acids under investigation.

Hexapeptides have fewer conformations with helical SS in water (<30 %) than nonapeptides, Figure 3A. This suggests that this hexapeptide is likely too short to fold into a stable helical structure in water, regardless of the substitutions incorporated on his sequence. The hexapeptides incorporating Ala and Aib show a similar low percentage of helical conformations. The experimental data about the Aib residue on the leucine based hexapeptide (Demizu et al. 2011) indicates that this amino acid induce a 3_{10}-helical conformation, while, our results show a small number of conformations presenting this SS type and a more significant contribution of the α-helical form. This difference is justified by the fact that, in solution, the peptide under study can populate different conformations that can be distinct from the ones present in crystal structure. In addition the α-helix is a common SS for peptides carrying Aib residues.

The hexapeptides containing Ac_5c and the chiral forms of this residue were also reported as having 310-helix SS for Ac_5c and 3_{10}-helix/α-helix for the chiral residues in water. These types of SS are also present in our simulations, although with low percentages: 10% of the Ac_5c peptide conformations present helical SS and 15% of helical conformations for the peptide containing (R,R)-Ac_5c^{dOM}, this confirms the observation that short peptides are less able to fold in helical structures.

Interestingly, increasing the size of the peptide bearing the Ala residue resulted in a complete loss of helical conformations, while in the case of Aib we observe an increase of helical SS of the α-helix and 3_{10}-helix types. In addition, the others nonapeptides containing non-canonical amino acids are more prone to adopt helical conformations (Figure 3A). The nonapeptide in water with Aib shows an increase in 3_{10}-helix SS percentage, compared to the hexapeptide in water.
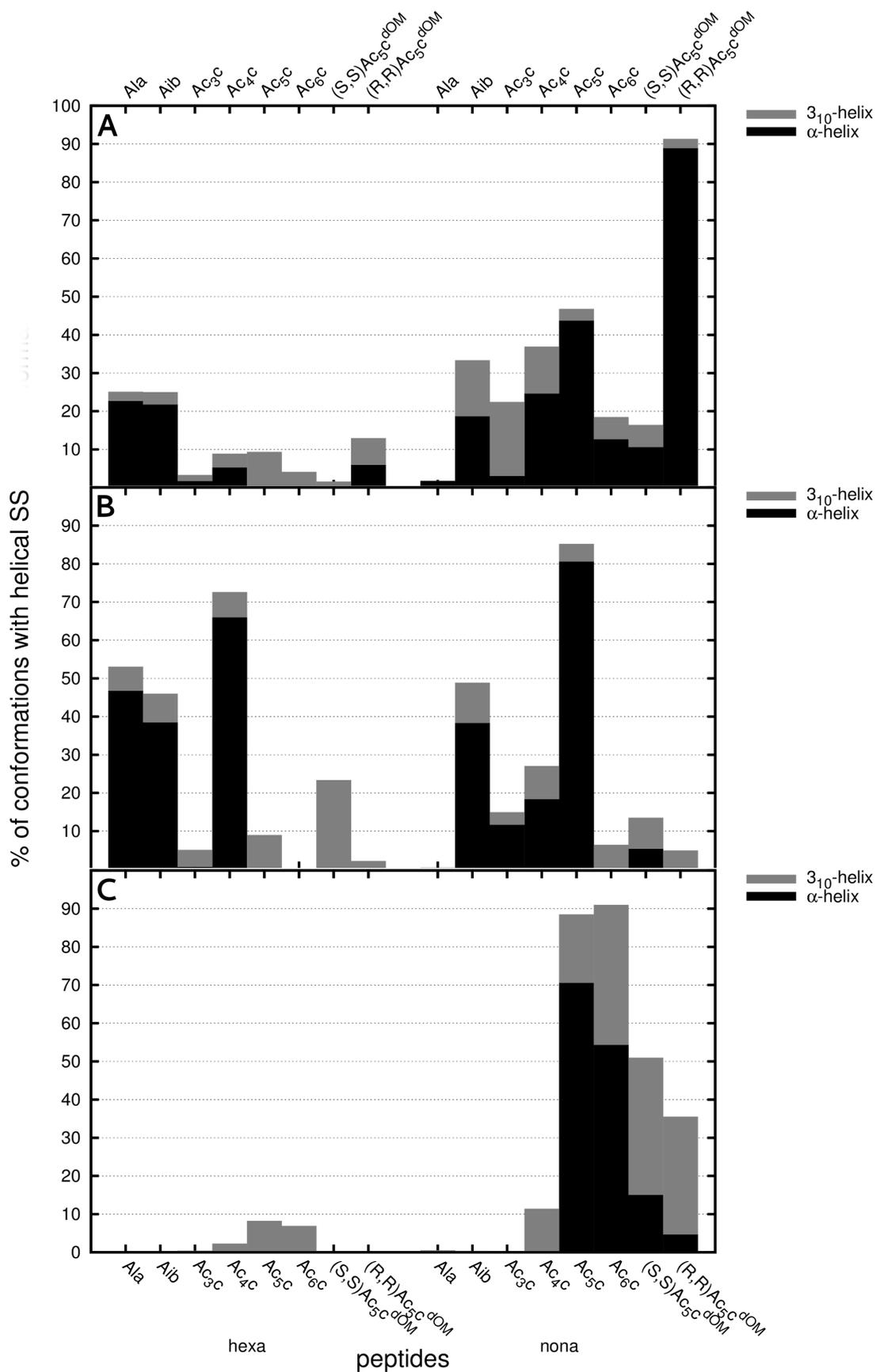
Figure 3. Percentage of conformations with helical SS (α-helix and 3$_{10}$-helix) observed for hexa and nonapeptides in (a) H$_2$O, (b) TFE/H$_2$O and (c) CHCL$_3$.

Furthermore, the small residues Ac_3c and Ac_4c also seem to induce an important percentage of 3_{10}-helix conformations. Remarkably, the highest numbers of conformations with helical SS on the nonapeptide in water are for Ac_5c and (R,R)-Ac_5c^{dOM}, ≈40% and 90% respectively. Note that in comparison the chiral image of (R,R)-Ac_5c^{dOM}, (S,S)-Ac_5c^{dOM} has only 10% of conformations presenting the α-helix form. Figure 4A depicts the intramolecular hydrogen bonds (3.5Å cut-off) involved in this helical structure.

The Ac_5c based residues generally stabilize the helical form of peptides in water. The cyclized side-chain imposes geometrical constrains around de Ca carbon so that the phi and psi angles of the peptide main chain populate the α-helical space of the Ramachandran Plot, thus promoting the formation of helical structures (see next in discussion). However, we also observe that, for Ac_5c, the peptide helicity can be improved if these residues are functionalized with two methoxy groups with a specific chirality, such as (R,R)-Ac_5c^{dOM} as opposed to its mirror image: (S,S)-Ac_5c^{dOM}.
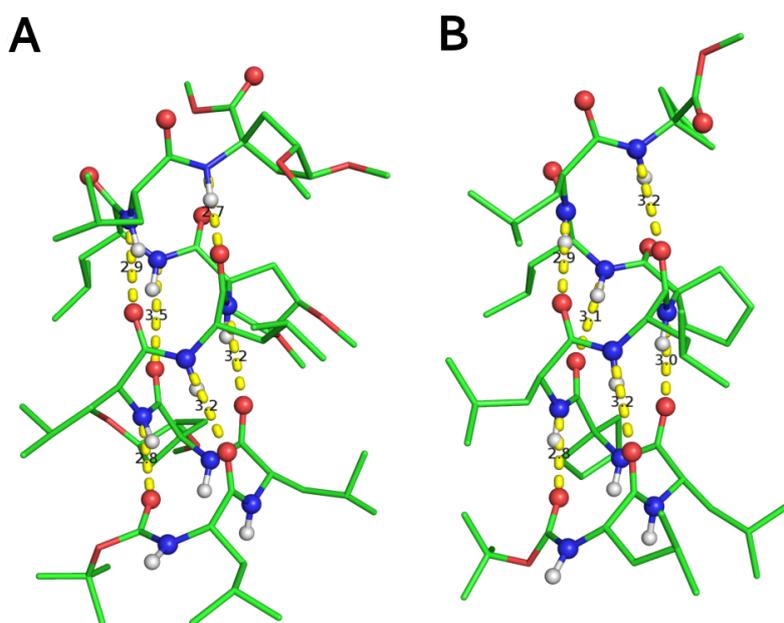
Figure 4. Representative α-helical structures of the nonapeptides carrying (A) (R,R)-Ac_5c^{dOM} in H_2O and (B) Ac_5c in TFE/H_2O. The coloring of atoms follows the convention: green for carbon, blue for nitrogen, red for oxygen, white for hydrogen and yellow to highlight the hydrogen bonds under 3.5Å cut-off and angle (donor-hydrogen-acceptor) less than 30°. The solvent molecules were omitted for better visualization.

This suggests that increasing slightly the polar character of this particular residue enhances the helicity of this model peptide. However, there is not a clear reason why the (R,R)-Ac_5c^{dOM} outperforms its equivalent chiral counterpart, (S,S)-Ac_5c^{dOM}. The substitution in the S,S or R,R affects differently the main chain dihedrals and τ angle (N-Cα-C), which in consequence influences the residues on the neighborhood (see the difference of τ among the residues on Supporting Information).

Figure 5 shows the SS helicoidal population for each residue during the simulation for selected nonapeptide cases. Residues from position 2 to 8 of the nonapeptides with Ac_5c, and (R,R)-Ac_5c^{dOM} are all involved in a helical SS during most of the simulation, Figure 5A-B. Interestingly, the functionalization of Ac_5c to give (R,R)-Ac_5c^{dOM} has a dramatic effect on stabilizing the α helical SS in this peptide doubling the number of populations in most residues.
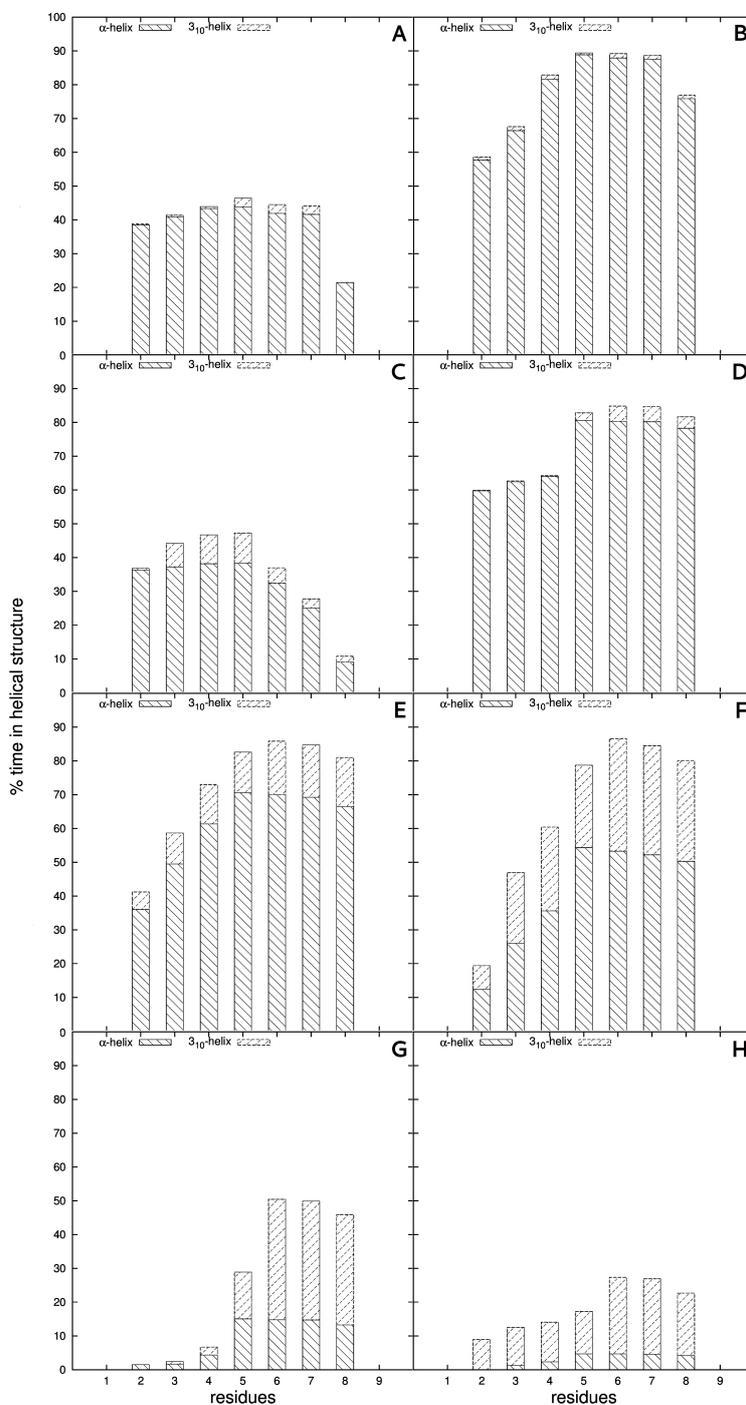
Figure 5. Percentage of simulation time in a-helix and 3_{10}-helix conformations for each residue in the sequence order of the following nonapeptides: (A) Ac_5c in H_2O, (B) (R,R)-Ac_5c^{dOM} in H_2O, (C) Aib in TFE/H_2O mixture, (D) Ac_5c in TFE/H_2O mixture, (E) Ac_5c in $CHCL_3$, (F) Ac_6c in $CHCL_3$, (G) (S,S)-Ac_5c^{dOM} in $CHCL_3$ and (H) (R,R)-Ac_5c^{dOM} in $CHCL_3$. The dAAs are in positions 3, 6, and 9.

3.2 Solvent effect on the SS of the peptides

The stability of peptide conformations is determined by the sequence of residues that form the primary structure, but also by the interaction with the solvent. TFE (2,2,2-trifluoroethanol) was used as a cosolvent (TFE/H$_2$O mixture) for the study of our peptides in solution. Demizu and co-workers used this mixture to perform Circular Dichroism (CD) spectra analysis (Demizu et al. 2011). Also, this organic solvent was chosen because it is known to protect the peptides of water molecules promoting conformational stability of hydrophobic residues (Hong et al. 1999; Reiersen and Rees 2000; Roccatano et al. 2002; Luo and Baldwin 1997). The TFE molecules can reduce the intermolecular interactions between the peptide and water molecules and the reduction of the hydrophobic effect enables effective formation or maintenance of intramolecular hydrogen bonds (Hong et al. 1999; Reiersen and Rees 2000; Roccatano et al. 2002; Luo and Baldwin 1997).

Figure 3B shows that hexapeptides incorporating Ala, Aib and Ac$_4$c solvated in TFE/H2O have higher percentages of conformations with helical SS compared to water (Figure 3A). Remarkably, we observe that the nonapeptide bearing the Ac$_5$c residue has significantly increased the helical SS content in TFE/H$_2$O, which is the expected effect of this solvent. Figure 4B shows a conformation of this peptide in TFE/H$_2$O, highlighting the intramolecular hydrogen bonds involved in the α-helical conformation. On the other hand, the peptides carrying Ac$_6$c and (*S,S*)-Ac$_5$c^{dOM} show a diminution of the total percentage of helical structure and, importantly, the residue (*R,R*)-Ac$_5$c^{dOM} does not promote helical SS in the TFE/H$_2$O. This fact can be attributed to the molecular properties of the solvent mixture and the polar properties of (*R,R*)-Ac$_5$c^{dOM} residue. In addition, the steric hindrance of the substituents at Ac$_5$c rings and the total volume of these residues, may predominate over the protection effect that the TFE molecules can offer.

Figure 5C and D show that all internal residues of the nonapeptides carrying Aib or Ac$_5$c in TFE/H$_2$O mixture, participate in the formation of helical SS. For the nonapeptide with Aib (Figure 5C), we observe a slight increase of a population of 3$_{10}$-helix for the residues in positions 3, 4, 5 and 6, indicating that these residues are not exclusively in α-helix SS. The Ac$_5$c nonapeptide in in TFE/H$_2$O mixture (Figure 5D) was highly structured and stable as observed in water with at least 80% conformers with helical SS (Figure 5A).

3.3 The 3$_{10}$-helix SS type in chloroform

Some authors reported in the last two decades that 3$_{10}$-helical structures are favored in chloroform (Karle et al. 1990; Formaggio et al. 2012; Lettieri et al. 2013; Awasthi et al. 2001). The hexapeptides

under study show little or none helicoidal SS in chloroform (<10%), being the largest Ac_4c, Ac_5c and Ac_6c (Figure 3C).

Nonapeptides in $CHCL_3$ (Figure 3C) carrying Ac_5c, Ac_6c, (S,S)-Ac_5c^{dOM} and (R,R)-Ac_5c^{dOM} have significant percentages of conformations with helical SS in this solvent. Ac_6c is the one that induces highest percentage of structures with helical SS of the α-helix and 3_{10}-helix type. Previous computational studies done by some of the authors also showed that Ac_6c has a good tendency to induce helical SS in peptaibols of different sizes and sequences (Castro and Micaelo 2014; Castro and Micaêlo 2014). This indicates that the foldamer behavior of Ac_6c is a feature that might be present in more peptides.

The nonapeptides with Ac_5c, Ac_6c, (S,S)-Ac_5c^{dOM} and (R,R)-Ac_5c^{dOM} show in chloroform the largest populations of 3_{10}-helix conformations. This agrees with experimental studies, which suggest that chloroform generally induces this type of secondary structure or promote the transition between the most stable conformations (Karle et al. 1990; Formaggio et al. 2012; Lettieri et al. 2013; Awasthi et al. 2001).
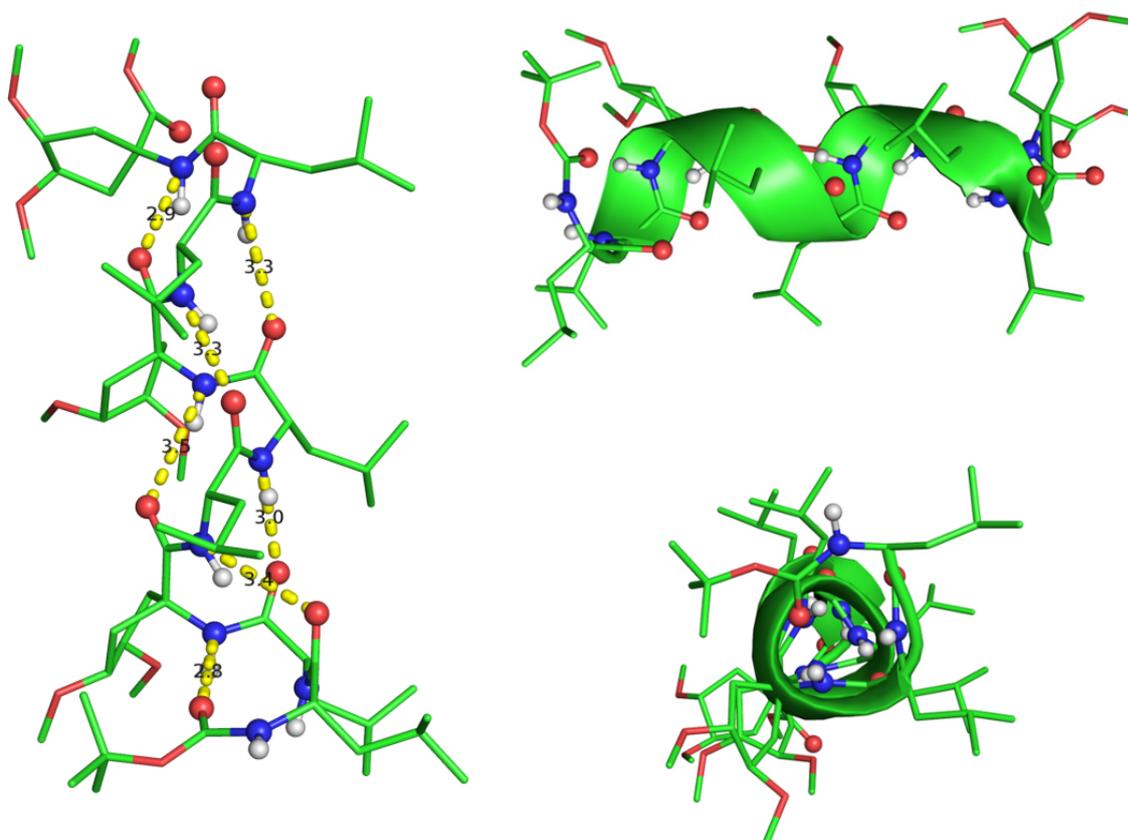
Figure 6. Three views of 3_{10}-helix of nonapeptide with (R,R)-Ac_5c^{dOM} in $CHCL_3$. The coloring of atoms was defined as in figure 4. The hydrogen bonds are highlighted in yellow and the peptide shows the cartoon that defines its SS.

Nonapeptides with (S,S)-Ac$_5$c^{dOM} and (R,R)-Ac$_5$c^{dOM} induce preferentially α-helix conformations in water, while in chloroform we observe an increase of the 3_{10}-helix population. In other words it is possible to shift the helical conformation of a peptide incorporating non-canonical amino acids, from α-helix to 3_{10}-helix, by changing the media from aqueous to chloroform. Solvents that have a lower tendency to interact with backbone peptide groups, as $CHCl_3$, induce mostly helical structures (Awasthi et al. 2001). This fact explains why the 3_{10}-helix is favored in this medium since it allows the greatest number possible of intramolecular hydrogen bonds. Particularly, $CHCL_3$ favor the folding in peptides sequences above 7 residues (Awasthi et al. 2001).

Figure 5E-H show the distribution of 3_{10}-helix conformations for each residue of the nonapeptides containing Ac$_5$c, Ac$_6$c, (S,S)-Ac$_5$cdOM and (R,R)-Ac$_5$c^{dOM}, respectively. It is evident that most of the residues are involved in 3_{10}-helix conformations and α-helix. Remarkably, for the nonapeptide with the (R,R)-Ac5c^{dOM} residue (Figure 5H), most of the peptide residues are arranged in 3_{10}-helix conformation. Figure 6 illustrates one conformation of the nonapeptide with (R,R)-Ac$_5$c^{dOM} solvated in $CHCL_3$ with 3_{10}-helix SS.

The structural preference of Ac$_5$c (Figure 5E), (S,S)-Ac$_5$c^{dOM} (Figure 5G) and (R,R)-Ac$_5$c^{dOM} (Figure 5H) towards helical SS is also evident from the observation of the distribution of dihedrals pairs on the Ramachandran space shown on Figure 7. The plots were generated with the φ and ψ information of the dAAs at the positions of the non-canonical residues (3, 6 and 9).

In figure 7 we observed that, as expected, Ala explores regions assigned to β-sheets, right α-helix, left α-helix and extended conformations, with highest density of φ and ψ pairs in β-sheet region. On the other hand, it is expect that the dAAs under study explore a more constrained region in Ramachandran space due the double substitution at the Cα. Ac$_5$c (Figure 7B) is found in conformations mostly in right α-helix, with a small population of dihedral pairs in the left a-helix region. (S,S)-Ac$_5$c^{dOM} (Figure 6C) also is found in conformations mostly in right a -helix, but also in the β-sheet region, left α-helix region and fully extended conformations at 180°, revealing a flexible arrangement of this residue despite the constrained imposed by the double substitution at the Cα and the bulky side chain. (R,R)-Ac$_5$c^{dOM} (Figure 7D) sampled similar conformational space as (S,S)-Ac$_5$c^{dOM}, but mostly concentrated in the right α-helix and in the region of ψ ≈ -180° to 45° and φ ≈ 45° to 90°. Ramachandran plots for the others dAAs under study (Aib, Ac$_3$c, Ac$_4$c, and Ac$_6$c), in chloroform, are included as supporting information.
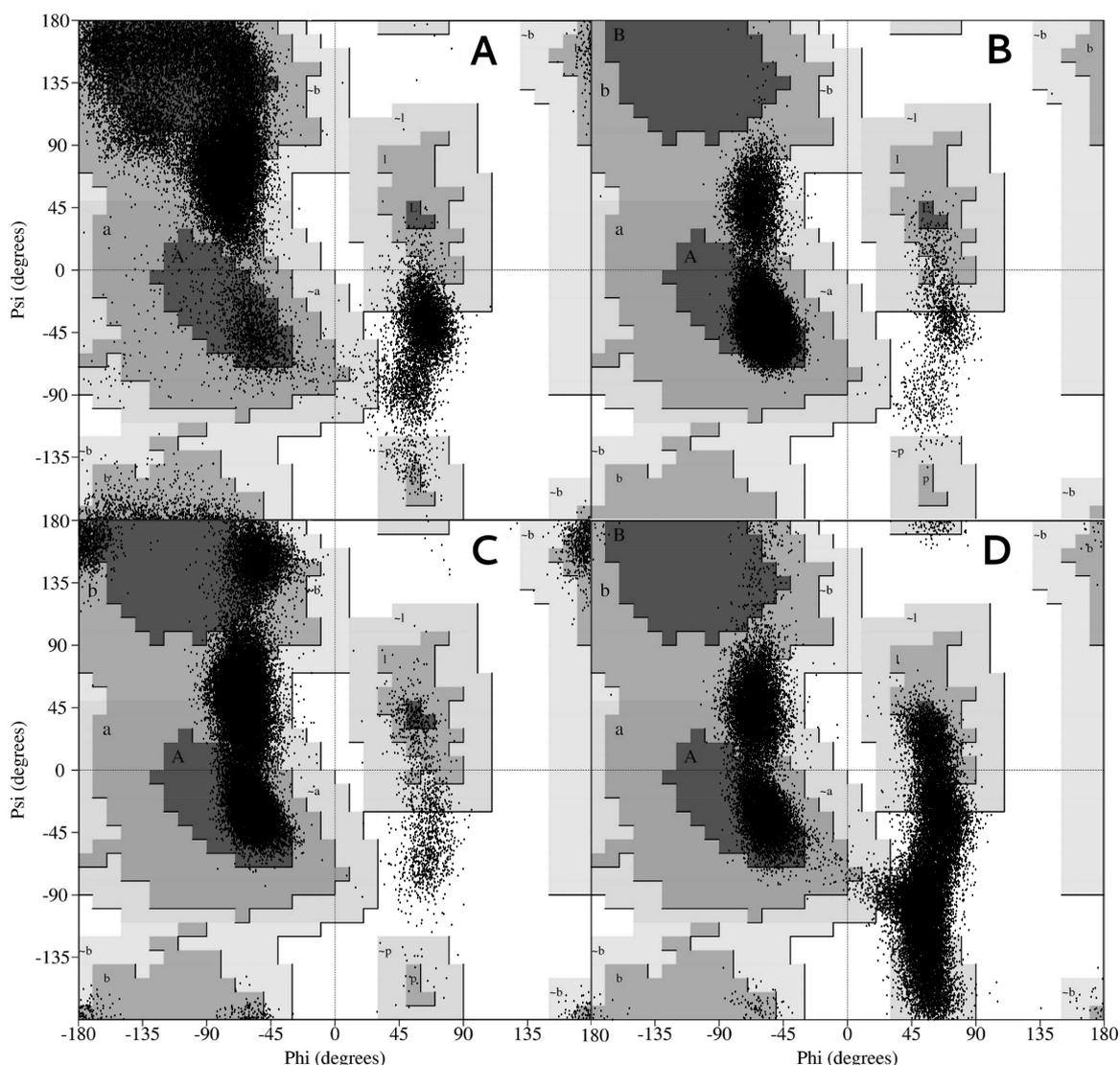
Figure 7. Nonapeptides φ and ψ dihedrals pair distribution for the amino acids (A) Ala, (B) Ac_5c, (C) (S,S)-Ac_5c^{dOM} and (D) (R,R)-Ac_5c^{dOM} in chloroform, superimposed on the Ramachandran diagram. In Ramachandran the region (a) corresponds to typical dihedrals or right a-helix, (b) corresponds to b-sheets space and (l) to left a-helix region.

4. Conclusions

We investigated the folding properties of different non-canonical dAAs towards the formation of α-helix and 3_{10}-helix SS in different solvents. We observed that some non-canonical residues have significant propensity to induce helical SS.

In water, Ac_5c and (R,R)-Ac_5c^{dOM} are the most capable to induce α-helical SS but only if inserted in nonapeptides as the same residues do not induce structure in equivalent hexapeptides. On the other hand, TFE/H_2O mixture induces an increase of α-helix conformations for the hexa and nonapeptides bearing apolar and less bulky dAAs, as Aib, Ac_4c and Ac_5c. This confirms that the presence of TFE in

the solvent helps the formation of helical SS, as previously suggested (Hong et al. 1999; Reiersen and Rees 2000; Roccatano et al. 2002; Luo and Baldwin 1997).

In CHCL$_3$, a significant shift of α-helix to 3$_{10}$-helix SS content was observed in several nonapeptides, especially the one incorporating the (*R,R*)-Ac$_5$c^{dOM} residue. In general, these results fit available experimental data (Demizu et al. 2011), as Ac$_5$c, (*S,S*)-Ac$_5$c^{dOM} and (*R,R*)-Ac$_5$c^{dOM} induce a mixture of α-helix and 3$_{10}$-helix conformations. However, the hexapeptides have no helical structure on this environment, indicating that a minimum length might be important to fold in this medium.

Summing up, we found that the presence of Ac$_5$c based amino acids have a strong tendency to induce nonapeptides helical conformations on the three solvents studied compared to Ala and Aib. The knowledge about the α-helical and 3$_{10}$-helical SS inducer potential of these non-canonical amino acids could be useful in the design of peptides with ad-hoc helical SS.

5. Acknowledgments

6. Supporting Information

The parameters for the non-canonical amino acids discussed in this article are available as supporting information (Table 1S). This section also presents Ramachandran plots for the non-canonical amino acids Aib, Ac$_3$c, Ac$_4$c and Ac$_6$c, in chloroform (Figure 1S) and the τ angles (N-Cα-C') for all the non-canonical amino acids under investigation (Table 2S).

References

Alemán C (1997) Conformational Properties of α-Amino Acids Disubstituted at the α-Carbon. J Phys Chem B 101 (25):5046-5050. doi:10.1021/jp963339+

Aschi M, Lucente G, Mazza F, Mollica A, Morera E, Nalli M, Paradisi MP (2003) Peptide backbone folding induced by the C-alpha-tetrasubstituted cyclic alpha-amino acids 4-amino-1,2-dithiolane-4-carboxylic acid (Adt) and 1-aminocyclopentane-1-carboxylic acid (Ac(5)c). A joint computational and experimental study. Org Biomol Chem 1 (11):1980-1988. doi:10.1039/b212247b

Awasthi SK, Shankaramma SC, Raghothama S, Balaram P (2001) Solvent-induced β-hairpin to helix conformational transition in a designed peptide. Biopolymers 58 (5):465-476. doi:10.1002/1097-0282(20010415)58:5<465::aid-bip1022>3.0.co;2-t

Balaram P (1992) Non-standard amino acids in peptide design and protein engineering. Curr Opin Chem Biol 2 (6):845-851. doi:http://dx.doi.org/10.1016/0959-440X(92)90110-S

Ballano G, Zanuy D, Jiménez AI, Cativiela C, Nussinov R, Alemán C (2008) Structural Analysis of a β-Helical Protein Motif Stabilized by Targeted Replacements with Conformationally Constrained Amino Acids. J Phys Chem B 112 (41):13101-13115. doi:10.1021/jp8032116

Ballet S, Feytens D, Buysse K, Chung NN, Lemieux C, Tumati S, Keresztes A, Van Duppen J, Lai J, Varga E, Porreca F, Schiller PW, Broeck JV, Tourwe D (2011) Design of Novel Neurokinin 1 Receptor Antagonists Based on Conformationally Constrained Aromatic Amino Acids and Discovery of a Potent Chimeric Opioid Agonist-Neurokinin 1 Receptor Antagonist. J Med Chem 54 (7):2467-2476. doi:10.1021/jm1016285

Bekker H, Berendsen HJC, Dijkstra EJ, Achterop S, Vondrumen R, Vanderspoel D, Sijbers A, Keegstra H, Reitsma B, Renardus MKR (1993) GROMACS - A Parallel Computer for Molecular-Dynamics Simulations Physics Computing '92. World Scientific Publ Co Pte Ltd, Singapore

Benedetti E, Di Blasio B, Iacovino R, Menchise V, Saviano M, Pedone C, Maria Bonora G, Ettorre A, Graci L, Formaggio F, Crisma M, Valle G, Toniolo C (1997) Conformational restriction through Calpha i - Calpha i cyclization: 1-aminocycloheptane-1-carboxylic acid (Ac7c). J Chem Soc, Perkin Trans 2 (10):2023-2032. doi:10.1039/a701473b

Berendsen HJC, Grigera JR, Straatsma TP (1987) The missing term in effective pair potentials J Phys Chem 91 (24):6269-6271. doi:10.1021/j100308a038

Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR (1984) Molecular-dynamics with coupling to an external bath. J Chem Phys 81 (8):3684-3690. doi:10.1063/1.448118

Bürgi R, Daura X, Mark A, Van Gunsteren W, Bellanda M, Mammi S, Peggion E (2001) Folding study of an Aib-rich peptide in DMSO by molecular dynamics simulations. J Pept Res 57 (2):107-118. doi:10.1034/j.1399-3011.2001.00793.x

Castro TG, Micaelo NM (2014) Conformational and Thermodynamic Properties of Non-Canonical alpha,alpha-Dialkyl Glycines in the Peptaibol Alamethicin: Molecular Dynamics Studies. J Phys Chem B 118 (33):9861-9870. doi:10.1021/jp505400q

Castro TG, Micaêlo NM (2014) Modeling of Peptaibol Analogues Incorporating Nonpolar α,α-Dialkyl Glycines Shows Improved α-Helical Preorganization and Spontaneous Membrane Permeation. J Phys Chem B 118 (3):649-658. doi:10.1021/jp4074587

Crisma M, Bonora GM, Toniolo C, Barone V, Benedetti E, Di Blasio B, Pavone V, Pedone C, Santini A, Fraternali F, Bavoso A, Lelj F (1989) Structural versatility of peptides containing Cα, α-dialkylated glycines: conformational energy computations, i.r. absorption and 1H n.m.r. analysis of 1-aminocyclopropane-1-carboxylic acid homopeptides. Int J Biol Macromolec 11 (6):345-352. doi:http://dx.doi.org/10.1016/0141-8130(89)90006-8

Demizu Y, Doi M, Kurihara M, Okuda H, Nagano M, Suemune H, Tanaka M (2011) Conformational studies on peptides containing alpha,alpha-disubstituted alpha-amino acids: chiral cyclic alpha,alpha-disubstituted alpha-amino acid as an alpha-helical inducer. Org Biomol Chem 9 (9):3303-3312. doi:10.1039/c0ob01146k

Demizu Y, Tanaka M, Doi M, Kurihara M, Okuda H, Suemune H (2010) Conformations of peptides containing a chiral cyclic alpha, alpha-disubstituted alpha-amino acid within the sequence of Aib residues. J Pept Sci 16 (11):621-626. doi:10.1002/psc.1273

Feytens D, Cescato R, Reubi JC, Tourwe D (2007) New sst(4/5)-selective somatostatin peptidomimetics based on a constrained tryptophan scaffold. J Med Chem 50 (14):3397-3401. doi:10.1021/jm070246f

Formaggio F, Crisma M, Ballano G, Peggion C, Venanzi M, Toniolo C (2012) Novel peptide foldameric motifs: a step forward in our understanding of the fully-extended conformation/310-helix coexistence. Org Biomol Chem 10 (12):2413-2421. doi:10.1039/c2ob06863j

Gatos M, Formaggio F, Crisma M, Toniolo C, Bonora GM, Benedetti Z, Di Blasio B, Iacovino R, Santini A, Saviano M, Kamphuis J (1997a) Conformational Characterization of the 1-Aminocyclobutane-1-carboxylic Acid Residue in Model Peptides. J Pept Sci 3 (2):110-122. doi:10.1002/(sici)1099-1387(199703)3:2<110::aid-psc88>3.0.co;2-6

Gatos M, Formaggio F, Crisma M, Valle G, Toniolo C, Bonora GM, Saviano M, Iacovino R, Menchise V, Galdiero S, Pedone C, Benedetti E (1997b) Conformational characterization of peptides rich in the cycloaliphatic Cα, α-disubstituted glycine 1-amino-cyclononane-1-carboxylic acid. J Pept Sci 3 (5):367-382. doi:10.1002/(sici)1099-1387(199709)3:5<367::aid-psc116>3.0.co;2-v

Gentilucci L, Tolomelli A, Squassabia F (2006) Peptides and peptidomimetics in medicine, surgery and biotechnology. Curr Med Chem 13 (20):2449-2466. doi:10.2174/092986706777935041

Giannis A (1993) Peptidomimetics for Receptor Ligands Discovery, Development, and Medical Perpectives. Angew Chem-Int Edit Engl 32 (9):1244-1267. doi:10.1002/anie.199312441

Gomez-Catalan J, Aleman C, Perez JJ (2000) Conformational profile of 1-aminocyclopropanecarboxylic acid. Theor Chem Acc 103 (5):380-389. doi:10.1007/s002149900066

Goodman CM, Choi S, Shandler S, DeGrado WF (2007) Foldamers as versatile frameworks for the design and evolution of function. Nat Chem Biol 3 (5):252-262. doi:10.1038/nchembio876

Grauer A, Konig B (2009) Peptidomimetics - A Versatile Route to Biologically Active Compounds. Eur J Org Chem (30):5099-5111. doi:10.1002/ejoc.200900599

Headley AD, Ganesan R, Nam J (2003) The effect of the cyclopropyl group on the conformation of chemotactic formyl tripeptides. Bioorg Chem 31 (2):99-108. doi:http://dx.doi.org/10.1016/S0045-2068(02)00522-9

Hess B (2008) P-LINCS: A parallel linear constraint solver for molecular simulation. J Chem Theory Comput 4 (1):116-122. doi:10.1021/ct700200b

Hess B, Bekker H, Berendsen HJC, Fraaije J (1997) LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18 (12):1463-1472. doi:10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.3.co;2-I

Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comput 4 (3):435-447. doi:10.1021/ct700301q

Hill DJ, Mio MJ, Prince RB, Hughes TS, Moore JS (2001) A field guide to foldamers. Chem Rev 101 (12):3893-4011. doi:10.1021/cr990120t

Hong D-P, Hoshino M, Kuboi R, Goto Y (1999) Clustering of Fluorine-Substituted Alcohols as a Factor Responsible for Their Marked Effects on Proteins and Peptides. J Am Chem Soc 121 (37):8427-8433. doi:10.1021/ja990833t

Huang W, Lin ZX, van Gunsteren WF (2011) Validation of the GROMOS 54A7 Force Field with Respect to beta-Peptide Folding. J Chem Theory Comput 7 (5):1237-1243. doi:10.1021/ct100747y

Jiménez AI, Vaquero V, Cabezas C, López JC, Cativiela C, Alonso JL (2011) The Singular Gas-Phase Structure of 1-Aminocyclopropanecarboxylic Acid (Ac3c). J Am Chem Soc 133 (27):10621-10628. doi:10.1021/ja2033603

Karle IL, Flippen-Anderson JL, Uma K, Balaram H, Balaram P (1990) Peptide design: Influence of a guest Aib-Pro segment on the stereochemistry of an Oligo-Val sequence—solution conformations and crystal structure of Boc-(Val)2-Aib-Pro-(Val)3-OMe. Biopolymers 29 (10-11):1433-1442. doi:10.1002/bip.360291010

Lettieri R, Bischetti M, Gatto E, Palleschi A, Ricci E, Formaggio F, Crisma M, Toniolo C, Venanzi M (2013) Looking for the peptide 2.05-helix: A solvent- and main-chain length-dependent conformational switch probed by electron transfer across c α , α -diethylglycine homo-oligomers. Pept Sci 100 (1):51-63. doi:10.1002/bip.22190

Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model 7 (8):306-317

Luo P, Baldwin RL (1997) Mechanism of Helix Induction by Trifluoroethanol: A Framework for Extrapolating the Helix-Forming Properties of Peptides from Trifluoroethanol/Water Mixtures Back to Water†. Biochemistry 36 (27):8413-8421. doi:10.1021/bi9707133

Mallareddy JR, Borics A, Keresztes A, Kover KE, Tourwe D, Toth G (2011) Design, Synthesis, Pharmacological Evaluation, and Structure-Activity Study of Novel Endomorphin Analogues with Multiple Structural Modifications. J Med Chem 54 (5):1462-1472. doi:10.1021/jm101515v

Marshall GR, Bosshard HE (1972) Angiotensin II. Studies on the biologically active conformation. Circ Res 31 (9):Suppl 2:143-150

Marshall GR, Hodgkin EE, Langs DA, Smith GD, Zabrocki J, Leplawy MT (1990) Factors Governing Helical Preference of Peptides Containing Multiple alpha,alpha-dialkyl Amino-Acids. Proc Natl Acad Sci U S A 87 (1):487-491. doi:10.1073/pnas.87.1.487

Martinez L, Andrade R, Birgin EG, Martinez JM (2009) PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. J Comput Chem 30 (13):2157-2164. doi:10.1002/jcc.21224

Mendel D, Ellman J, Schultz PG (1993) Protein-Biosynthesis with Conformationally Restricted Amino-Acids. J Am Chem Soc 115 (10):4359-4360. doi:10.1021/ja00063a063

Moretto A, Formaggio F, Crisma M, Toniolo C, Saviano M, Benedetti E, Iacovino R, Vitale RM (2001) Ac10c: a medium-ring, cycloaliphatic C α, α -disubstituted glycine. Incorporation into model peptides and preferred conformation. J Pept Res 57 (4):307-315. doi:10.1046/j.1397-002X.2000.00834.x

Oh JE, Lee KH (1999) Synthesis of novel unnatural amino acid as a building block and its incorporation into an antimicrobial peptide. Bioorg Med Chem 7 (12):2985-2990

Paradisi MP, Torrini I, Zecchini GP, Lucente G, Gavuzzo E, Mazza F, Pochetti G (1995) Gamma-turn conformation induced by alpha,alpha-disubstituted amino acids with a cyclic 6-membered side-chain. Tetrahedron 51 (8):2379-2386

Prasad S, Mathur A, Sharma R, Gupta N, Ahuja R, Jaggi M, Singh A, Mukherjee R (2006) Octapeptide Analogs of Somatostatin Containing α, α -Dialkylated Amino Acids with Potent Anticancer Activity. Int J Pept Res Ther 12 (2):179-185. doi:10.1007/s10989-005-9005-0

Reiersen H, Rees AR (2000) Trifluoroethanol may form a solvent matrix for assisted hydrophobic interactions between peptide side chains. Protein Eng 13 (11):739-743. doi:10.1093/protein/13.11.739

Ressurreicao ASM, Bordessa A, Civera M, Belvisi L, Gennari C, Piarulli U (2008) Synthesis and conformational studies of peptidomimetics containing a new bifunctional diketopiperazine scaffold acting as a beta-hairpin inducer. J Org Chem 73 (2):652-660. doi:10.1021/jo702072z

Roccatano D, Colombo G, Fioroni M, Mark AE (2002) Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. Proc Natl Acad Sci U S A 99 (19):12179-12184. doi:10.1073/pnas.182199699

Rodriguez-Ropero F, Zanuy D, Casanovas J, Nussinov R, Aleman C (2008) Application of 1-aminocyclohexane carboxylic acid to protein nanostructure computer design. J Chem Inf Model 48 (2):333-343. doi:10.1021/ci700291x

Rose GD, Glerasch LM, Smith JA (1985) Turns in Peptides and Proteins. In: C.B. Anfinsen JTE, Frederic MR (eds) Advances in Protein Chemistry, vol Volume 37. Academic Press, pp 1-109. doi:http://dx.doi.org/10.1016/S0065-3233(08)60063-7

Santini A, Di BB, Galdiero S, R. I, Pedone C, Benedetti E, Crisma M, Toniolo C (1996) Molecular and crystal structures of two terminally-blocked Ac5c homo-oligopeptides. Zeitschrift für Kristallographie, vol 211. doi:10.1524/zkri.1996.211.9.616

Saviano M, Iacovino R, Benedetti E, Moretto V, Banzato A, Formaggio F, Crisma M, Toniolo C (2000a) Preferred conformation of peptides based on cycloaliphatic C α, α -disubstituted glycines: 1-amino-cycloundecane-1-carboxylic acid (Ac11c). J Pept Sci 6 (11):571-583. doi:10.1002/1099-1387(200011)6:11<571::aid-psc290>3.0.co;2-r

Saviano M, Iacovino R, Menchise V, Benedetti E, Bonora GM, Gatos M, Graci L, Formaggio F, Crisma M, Toniolo C (2000b) Conformational restriction through C α i   C α i cyclization: Ac12c, the largest cycloaliphatic C α, α - disubstituted glycine known. Biopolymers 53 (2):200-212. doi:10.1002/(sici)1097-0282(200002)53:2<200::aid-bip10>3.0.co;2-l

Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, van Gunsteren WF (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur Biophys J Biophys Lett 40 (7):843-856. doi:10.1007/s00249-011-0700-9

Schrödinger (2010) The PyMOL Molecular Graphics System. 1.3r1 edn. LLC,

Smith PE, Vangunsteren WF (1994) Consistent dielectric-properties of the simple pointe-charge and extended simple poit-charge water models at 277 and 300K J Chem Phys 100 (4):3169-3174. doi:10.1063/1.466407

Spoel Dvd, Linddahl E, Hess B, Buuren ARv, Apol E, Meulenhoff PJ, Tieleman DP, Sijbers ALTM, Feenstra KA, Drunen Rv, Berendsen HJC (2010) Gromacs User Manual version 4.5.4.

Toniolo C (1990) Conformationally Restricted Peptides Through Short-Range Cyclizations. Int J Pept Protein Res 35 (4):287-300

Toniolo C, Benedetti E (1991) The Polypeptide-3(10)-helix. Trends Biochem Sci 16 (9):350-353

Toniolo C, Crisma M, Formaggio F, Peggion C (2001) Control of peptide conformation by the Thorpe-Ingold effect (C(alpha)-tetrasubstitution). Biopolymers 60 (6):396-419. doi:10.1002/bip.10184

Vagner J, Qu HC, Hruby VJ (2008) Peptidomimetics, a synthetic tool of drug discovery. Curr Opin Chem Biol 12 (3):292-296. doi:10.1016/j.cbpa.2008.03.009

van der Spoel D, van Maaren PJ, Berendsen HJC (1998) A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. J Chem Phys 108 (24):10220-10230. doi:10.1063/1.476482

Venkatachalam CM (1968) Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. Biopolymers 6 (10):1425-1436. doi:10.1002/bip.1968.360061006

Venkatraman J, Shankaramma SC, Balaram P (2001) Design of folded peptides. Chem Rev 101 (10):3131-3152. doi:10.1021/cr000053z

Whitby LR, Ando Y, Setola V, Vogt PK, Roth BL, Boger DL (2011) Design, Synthesis, and Validation of a beta-Turn Mimetic Library Targeting Protein-Protein and Peptide-Receptor Interactions. J Am Chem Soc 133 (26):10184-10194. doi:10.1021/ja201878v

Zanuy D, Ballano G, Jimenez AI, Casanovas J, Haspel N, Cativiela C, Curco D, Nussinov R, Aleman C (2009) Protein Segments with Conformationally Restricted Amino Acids Can Control Supramolecular Organization at the Nanoscale. J Chem Inf Model 49 (7):1623-1629. doi:10.1021/ci9001487

Zimmerman SS, Pottle MS, Némethy G, Scheraga HA (1977) Conformational Analysis of the 20 Naturally Occurring Amino Acid Residues Using ECEPP. Macromolecules 10 (1):1-9. doi:10.1021/ma60055a001

Supplementary Material of Appendix VII

Table 1S. GROMOS 54a7 Force Field topologies: Bonded and non-bonded parameters.

[Aib]

```
[ atoms ]
; atom label, atom type, charge, energy group
  N    N   -0.31000    0
  H    H    0.31000    0
  CA   C    0.00000    1
  CB1  CH3  0.00000    1
  CB2  CH3  0.00000    1
  C    C    0.450      2
  O    O    -0.450     2
[ bonds ]
  N    H    gb_2
  N    CA   gb_21
  CA   CB1  gb_27
  CA   CB2  gb_27
  CA   C    gb_27
  C    O    gb_5
  C    +N   gb_10
[ angles ]
; ai  aj  ak  gromos type
  -C   N    H    ga_32
  -C   N    CA   ga_31
  H    N    CA   ga_18
  N    CA   CB1  ga_13
  N    CA   C    ga_19
  CB1  CA   C    ga_13
  N    CA   CB2  ga_13
  CB1  CA   CB2  ga_13
  CB2  CA   C    ga_13
  CA   C    O    ga_30
  CA   C    +N   ga_19
  O    C    +N   ga_33
[ impropers ]
; ai  aj  ak  al  gromos type
  N    -C   CA   H    gi_1
  CA   N    C    CB1  gi_2
  C    CA   +N   O    gi_1
  CA   N    CB2  C    gi_2
  CA   N    CB1  CB2  gi_2
[ dihedrals ]
; ai  aj  ak  al  gromos type
  -CA  -C   N    CA   gd_14
  -C   N    CA   C    gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  -C   N    CA   C    gd_43
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N    CA   C    +N   gd_44
;backbone dihedral, changed by Ying Xue Sep 29. 2009
  N    CA   C    +N   gd_45
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

[Ac₃c]

```
; atom label, atom type, charge, energy group
[ atoms ]
  N    N   -0.31000    0
  H    H    0.31000    0
  CA   C    0.00000    1
  C1   CH2r  0.00000   1
  C2   CH2r  0.00000   1
  C    C    0.450      2
  O    O    -0.450     2
[ bonds ]
  N    H    gb_2
  N    CA   gb_21
  CA   C1   gb_27
  CA   C2   gb_27
  C1   C2   gb_27
  CA   C    gb_27
  C    O    gb_5
  C    +N   gb_10
[ angles ]
; ai  aj  ak  gromos type
  -C   N    H    ga_32
  -C   N    CA   ga_31
  H    N    CA   ga_18
  N    CA   C    ga_13
  N    CA   C1   ga_13
  N    CA   C2   ga_13
  C2   CA   C    ga_13
  C1   CA   C    ga_13
  C1   CA   C2   ga_55
; (cyclopropane ring ga_55 = 60°; Force constant = 520)
  CA   C2   C1   ga_55
  CA   C1   C2   ga_55
  CA   C    O    ga_30
  CA   C    +N   ga_19
  O    C    +N   ga_33
[ impropers ]
; ai  aj  ak  al  gromos type
  N    -C   CA   H    gi_1
  CA   N    C1   C    gi_2
  CA   N    C    C2   gi_2
  CA   N    C2   C1   gi_2
  C    CA   +N   O    gi_1
[ dihedrals ]
; ai  aj  ak  al  gromos type
  -CA  -C   N    CA   gd_14
  -C   N    CA   C    gd_42
;backbone dihedral, changed by Ying Xue Sep 29. 2009
-C   N    CA   C    gd_43
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

```
  N   CA   C1   C2    gd_34
  C   CA   C2   C1    gd_34
  N   CA   C    +N    gd_44
```

[Ac₄c]

; atom label, atom type, charge, energy group

[atoms]
```
  N    N    -0.31000   0
  H    H     0.31000   0
  CA   C     0.00000   1
  C1   CH2r  0.00000   1
  C2   CH2r  0.00000   1
  C3   CH2r  0.00000   1
  C    C     0.450     2
  O    O    -0.450     2
```

[bonds]
```
  N    H    gb_2
  N    CA   gb_21
  CA   C1   gb_27
  CA   C3   gb_27
  C1   C2   gb_27
  C2   C3   gb_27
  CA   C    gb_27
  C    O    gb_5
  C    +N   gb_10
```

[angles]
; ai aj ak gromos type
```
  -C   N    H     ga_32
  -C   N    CA    ga_31
  H    N    CA    ga_18
  N    CA   C     ga_13
  N    CA   C1    ga_13
  N    CA   C3    ga_13
  C1   CA   C3    ga_56
```
; (cyclobutane ring ga_56 = 88° ; Force constant = 520)
```
  C1   C2   C3    ga_56
  CA   C1   C2    ga_56
  CA   C3   C2    ga_56
  C3   CA   C     ga_13
  C    CA   C1    ga_13
  CA   C    O     ga_30
  CA   C    +N    ga_19
  O    C    +N    ga_33
```

[impropers]
; ai aj ak al gromos type
```
  N   -C   CA   H    gi_1
  CA   N    C    C1   gi_2
  CA   N    C3   C    gi_2
  CA   N    C1   C3   gi_2
  C    CA   +N   O    gi_1
```

[dihedrals]
; ai aj ak al gromos type
```
  -CA  -C   N    CA   gd_14
  -C   N    CA   C    gd_42
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```
  -C   N    CA   C    gd_43
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009

;backbone dihedral, changed by Ying Xue Sep 29. 2009
```
  N    CA   C    +N    gd_45
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009

```
  N    CA   C1   C2    gd_34
  C    CA   C3   C2    gd_34
  CA   C3   C2   C1    gd_34
  CA   C1   C2   C3    gd_34
  N    CA   C    +N    gd_44
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```
  N    CA   C    +N    gd_45
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009

[Ac₅c]

; atom label, atom type, charge, energy group

[atoms]
```
  N    N    -0.31000   0
  H    H     0.31000   0
  CA   C     0.00000   1
  C1   CH2r  0.00000   1
  C2   CH2r  0.00000   1
  C3   CH2r  0.00000   2
  C4   CH2r  0.00000   2
  C    C     0.450     3
  O    O    -0.450     3
```

[bonds]
```
  N    H    gb_2
  N    CA   gb_21
  CA   C1   gb_27
  CA   C4   gb_27
  C1   C2   gb_27
  C2   C3   gb_27
  C3   C4   gb_27
  CA   C    gb_27
  C    O    gb_5
  C    +N   gb_10
```

[angles]
; ai aj ak gromos type
```
  -C   N    H     ga_32
  -C   N    CA    ga_31
  H    N    CA    ga_18
  N    CA   C     ga_13
  N    CA   C1    ga_13
  N    CA   C4    ga_13
  C    CA   C4    ga_13
  C    CA   C1    ga_13
  C1   CA   C4    ga_7
  CA   C1   C2    ga_7
  CA   C4   C3    ga_7
  C1   C2   C3    ga_7
  C2   C3   C4    ga_7
  CA   C    O     ga_30
  CA   C    +N    ga_19
  O    C    +N    ga_33
```

[impropers]
; ai aj ak al gromos type
```
  N   -C   CA   H    gi_1
```

```
  CA   N   C   C1    gi_2              N   CA   C    ga_13
  CA   N   C4  C     gi_2              N   CA   C1   ga_13
  CA   N   C1  C4    gi_2              N   CA   C4   ga_13
   C   CA  +N  O     gi_1              C   CA   C4   ga_13
[ dihedrals ]                         C   CA   C1   ga_13
; ai  aj  ak  al  gromos type         C1  CA   C4   ga_7
 -CA  -C   N   CA   gd_14             CA   C1   C2   ga_7
 -C   N   CA   C    gd_42             CA   C4   C3   ga_7
;backbone dihedral, changed by Ying Xue Sep 29. 2009   C1  C2  C3  ga_7
 -C   N   CA   C    gd_43             C2   C3   C4   ga_7
;backbone dihedral, changed by Ying Xue Sep 29. 2009   C2  O01  C5  ga_12
  N   CA  C1   C2   gd_34             C3   O02  C6   ga_12
  C   CA  C4   C3   gd_34             O01  C2   C1   ga_13
  CA  C4  C3   C2   gd_34             O02  C3   C4   ga_13
  CA  C1  C2   C3   gd_34             O01  C2   C3   ga_13
  C1  C2  C3   C4   gd_1              O02  C3   C2   ga_13
  N   CA  C   +N    gd_44             CA   C    O    ga_30
;backbone dihedral, changed by Ying Xue Sep 29. 2009   CA  C  +N  ga_19
  N   CA  C   +N    gd_45              O   C   +N   ga_33
;backbone dihedral, changed by Ying Xue Sep 29. 2009  [ impropers ]
                                      ; ai  aj  ak  al  gromos type
```

[(S,S)-Ac$_5$c^{dOM}]

; atom label, atom type, charge, energy group

```
[ atoms ]                              N  -C   CA   H    gi_1
  N    N   -0.31000    0               CA   N   C   C1   gi_2
  H    H    0.31000    0               CA   N   C1  C4   gi_2
  CA   C    0.00000    1               CA   N   C4  C    gi_2
  C1   CH2r  0.00000   1               C3   O02  C2  C4   gi_2
  C2   CH2r  0.00000   1               C2   O01  C3  C1   gi_2
  C3   CH2r  0.00000   2                C   CA  +N   O    gi_1
  C4   CH2r  0.00000   2              [ dihedrals ]
  O01  OE   -0.450     3              ; ai  aj  ak  al  gromos type
  C5   CH3   0.450     3               -CA  -C   N   CA   gd_14
  O02  OE   -0.450     4               -C   N   CA   C    gd_42
  C6   CH3   0.450     4              ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  C    C     0.450     5               -C   N   CA   C    gd_43
  O    O    -0.450     5              ;backbone dihedral, changed by Ying Xue Sep 29. 2009
[ bonds ]                              N   CA  C1   C2   gd_34
  N    H    gb_2                       C   CA  C4   C3   gd_34
  N    CA   gb_21                      CA  C4  C3   C2   gd_34
  CA   C1   gb_27                      CA  C1  C2   C3   gd_34
  CA   C4   gb_27                      C4  C3  O02  C6   gd_13
  C1   C2   gb_27                      C1  C2  O01  C5   gd_13
  C2   C3   gb_27                      C1  C2  C3   C4   gd_1
  C3   C4   gb_27                      N   CA  C   +N    gd_44
  CA   C    gb_27                     ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  C2   O01  gb_13                      N   CA  C   +N    gd_45
  O01  C5   gb_18                     ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  C3   O02  gb_13
  O02  C6   gb_18
  C    O    gb_5
  C    +N   gb_10
[ angles ]
; ai  aj  ak  gromos type
 -C   N   H    ga_32
 -C   N   CA   ga_31
  H   N   CA   ga_18
```

[(R,R)-Ac$_5$c^{dOM}]

; atom label, atom type, charge, energy group

```
[ atoms ]
  N    N   -0.31000    0
  H    H    0.31000    0
  CA   C    0.00000    1
  C1   CH2r  0.00000   1
  C2   CH2r  0.00000   1
  C3   CH2r  0.00000   2
  C4   CH2r  0.00000   2
```

```
O01   OE    -0.450    3
C5   CH3    0.450    3
O02   OE    -0.450    4
C6   CH3    0.450    4
 C    C     0.450    5
 O    O    -0.450    5
[ bonds ]
 N    H    gb_2
 N    CA   gb_21
CA    C1   gb_27
CA    C4   gb_27
C1    C2   gb_27
C2    C3   gb_27
C3    C4   gb_27
CA    C    gb_27
C2    O01  gb_13
O01   C5   gb_18
C3    O02  gb_13
O02   C6   gb_18
 C    O    gb_5
 C    +N   gb_10
[ angles ]
;  ai   aj   ak   gromos type
 -C    N    H    ga_32
 -C    N    CA   ga_31
  H    N    CA   ga_18
  N    CA   C    ga_13
  N    CA   C1   ga_13
  N    CA   C4   ga_13
  C    CA   C4   ga_13
  C    CA   C1   ga_13
 C1    CA   C4   ga_7
 CA    C1   C2   ga_7
 CA    C4   C3   ga_7
 C1    C2   C3   ga_7
 C2    C3   C4   ga_7
 C2    O01  C5   ga_12
 C3    O02  C6   ga_12
 O01   C2   C1   ga_13
 O02   C3   C4   ga_13
 O01   C2   C3   ga_13
 O02   C3   C2   ga_13
 CA    C    O    ga_30
 CA    C    +N   ga_19
  O    C    +N   ga_33
[ impropers ]
; ai   aj   ak   al   gromos type
  N   -C    CA   H    gi_1
 CA    N    C    C1   gi_2
 CA    N    C4   C    gi_2
 CA    N    C1   C4   gi_2
 C3    O02  C4   C2   gi_2
 C2    O01  C1   C3   gi_2
  C    CA   +N   O    gi_1
[ dihedrals ]
; ai   aj   ak   al   gromos type
 -CA  -C    N    CA   gd_14
 -C    N    CA   C    gd_42
```

;backbone dihedral, changed by Ying Xue Sep 29. 2009
```
 -C    N    CA   C    gd_43
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```
  N    CA   C1   C2   gd_34
  C    CA   C4   C3   gd_34
 CA    C4   C3   C2   gd_34
 CA    C1   C2   C3   gd_34
 C4    C3   O02  C6   gd_13
 C1    C2   O01  C5   gd_13
 C1    C2   C3   C4   gd_1
  N    CA   C    +N   gd_44
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```
  N    CA   C    +N   gd_45
```
;backbone dihedral, changed by Ying Xue Sep 29. 2009

[Ac$_6$c]
```
 ; atom label, atom type, charge, energy group
[ atoms ]
  N    N    -0.31000    0
  H    H     0.31000    0
 CA    C     0.00000    1
 C1   CH2r   0.00000    1
 C2   CH2r   0.00000    1
 C3   CH2r   0.00000    2
 C4   CH2r   0.00000    2
 C5   CH2r   0.00000    2
  C    C     0.450    3
  O    O    -0.450    3
[ bonds ]
  N    H    gb_2
  N    CA   gb_21
 CA    C1   gb_27
 CA    C5   gb_27
 CA    C    gb_27
 C1    C2   gb_27
 C2    C3   gb_27
 C3    C4   gb_27
 C4    C5   gb_27
  C    O    gb_5
  C    +N   gb_10
[ angles ]
;  ai   aj   ak   gromos type
 -C    N    H    ga_32
 -C    N    CA   ga_31
  H    N    CA   ga_18
  N    CA   C    ga_13
  N    CA   C1   ga_13
  N    CA   C5   ga_13
 C1    CA   C    ga_13
 C5    CA   C    ga_13
 C1    CA   C5   ga_13
 CA    C1   C2   ga_13
 CA    C5   C4   ga_13
 C1    C2   C3   ga_13
 C2    C3   C4   ga_13
 C3    C4   C5   ga_13
 CA    C    O    ga_30
```

```
  CA    C   +N    ga_19                    -C    N    CA    C    gd_43
   O    C   +N    ga_33               ;backbone dihedral, changed by Ying Xue Sep 29. 2009
[ impropers ]                              N    CA   C5    C4   gd_34
; ai   aj   ak   al  gromos type           C    CA   C1    C2   gd_34
   N   -C   CA    H   gi_1                  N    CA    C   +N    gd_44
  CA    N   C1    C   gi_2               ;backbone dihedral, changed by Ying Xue Sep 29. 2009
  CA    N    C   C5   gi_2                  N    CA    C   +N    gd_45
  CA    N   C5   C1   gi_2               ;backbone dihedral, changed by Ying Xue Sep 29. 2009
   C   CA   +N    O   gi_1                 CA   C1    C2   C3   gd_34
[ dihedrals ]                             CA   C5    C4   C3   gd_34
; ai   aj   ak   al  gromos type          C1   C2    C3   C4   gd_34
 -CA   -C    N   CA   gd_14               C2   C3    C4   C5   gd_34
  -C    N   CA    C   gd_42                N   CA    C1   C2   gd_34
;backbone dihedral, changed by Ying Xue Sep 29. 2009
```

Figure 1S. Ramachandran Plots in CHCL$_3$

This section presents the dihedrals pair distribution (φ and ψ) superimposed on the Ramachandran diagram, for the non-canonical amino acids Aib (A), Ac$_3$c (B), Ac$_4$c (C) and Ac$_6$c (D).

Tarsila G. Castro, Nuno M. Micaêlo and Manuel Melle-Franco, 2015.
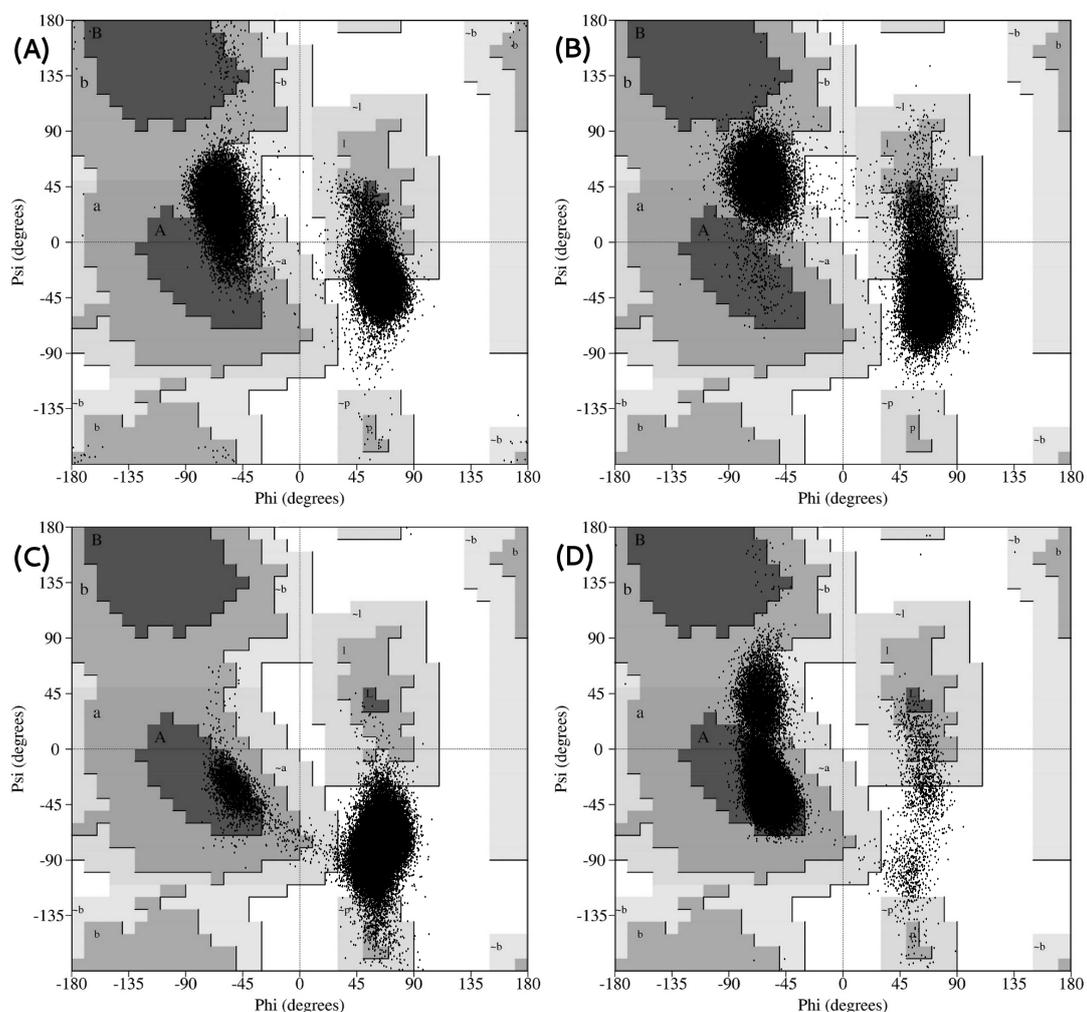
Table 2S. τ angle (degrees) (N-Cα-C') for the non-canonical amino acids under study, in all three solvent environments.

τ angle	hexapeptides			nonapeptides		
	H_2O	TFE/H_2O	$CHCL_3$	H_2O	TFE/H_2O	$CHCL_3$
Aib	116.7	116.4	116.9	117.2	117.1	117.7
Ala	114.3	114.4	114.4	115.1	114.1	114.6
Ac_3c	136.3	136.8	137.2	118.9	119.3	119.5
Ac_4c	114.0	113.6	113.8	114.1	114.4	114.2
Ac_5c	114.0	114.2	114.7	113.1	112.6	112.9
Ac_6c	111.9	111.0	111.5	112.4	112.8	112.8
[(*S,S*)-Ac_5c^{dOM}]	114.3	114.3	114.3	114.3	113.8	113.8
[(*R,R*)-Ac_5c^{dOM}]	113.7	115.4	113.9	112.6	114.2	113.9