INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

CONFORMATIONAL STUDIES OF BIOLOGICALLY ACTIVE MOLECULES

by

Octavian Antohi

A dissertation submitted to the Graduate Faculty in Chemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1998

This manuscript has been read and accepted for the Graduate Faculty in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

6/29/1998
Date

Chair of the Examining Committee

8/17/1998
Date

Executive Officer

Ruth E. Stark

Charlotte S. Russell

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

CONFORMATIONAL STUDIES OF BIOLOGICALLY ACTIVE MOLECULES

by

Octavian Antohi

Adviser: Professor Anne-Marie Sapse

Computational techniques(molecular modeling and *ab initio* calculations) are used in conjunction with nuclear magnetic resonance spectroscopy in order to probe and determine the conformation of a family of peptides that are related to the *Saccharomyces Cerevisiae* α-Factor pheromone.

The *ab initio* calculations provide quantitative values for the energy barriers introduced by the usage of certain side chain constraints. These quantitative results are found to correlate well with the conformational preferences of peptides that are built around cores containing such amino acids.

The strengths and weaknesses of force field molecular modeling techniques are analyzed in detail and their applicability to the design and interpretation of experimental setups is discussed.

Both the physical chemical analysis of analogs and the calculations seem to support the currently accepted model in which a turn in the 7 to 10 region of this peptide molecule is considered to play a decisive role in its pheromonal activity. Besides shedding light on the possible

biologically active conformation, the analysis of cyclized analogs clarifies the effects of lactam cycle constraints which are found to be consistently independent of the size of the peptide chain but strongly dependent on the size of the lactamization ring.

The limitations of solid state nmr dipolar dephasing experiments are discussed and a first order approximation theoretical basis for the computation of natural abundance corrections is developed. New multiple spin experiments are hypothetically considered.

Motto

"We can offer you little, but we offer all we have." (a human in love)

"Evil does not come to you as evil. Evil will smile at you, will give you money, and will tell you that you are wonderful." (approximate quotation from the Honorable Judge Guy DePhillips)

DEDICATION

To the Old

To Stefan and Maria,

my parents

in whose strength and weakness

I recognize my own

To naked shamelessness and to the cold

To brainless cats

And to other cats turned rats

To litmus Jims who would like to be bats

To those who lie, kill, steal, and claw

Yet, above all, to the lovely and generous,

The ones that are sold

To Marjorie's eyes

To Daniel's smile

To Cherryl's heart

To Stefan's soul - the best of us all

To all the little ones who were not allowed to grow / And to all the little ones who were forced to grow too fast / By wolves, cats, would be bats, and rats,

My thoughts

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Professor Anne - Marie Sapse, my mentor, whose help and encouragement were essential for the completion of this thesis.

I would like to thank Professors Ruth E. Stark and Charlotte Russell for their interest in my project and for serving on my Thesis Committee.

Thanks to all the members of the Peptide Chemistry group at the College of Staten Island, specifically, to Drs. F. R. Naider, E. Krainer, M. Breslav, J. Gounarides, and B. Arshava from whom I learned more than I could acknowledge.

A special Thank You to Dr. Hanumantha Rao Marepalli, who has proven to be a true friend.

Although not directly connected to this dissertation the following persons influenced my thought and personality profoundly enough to grant mentioning them in connection with any of my achievements.

I would like to acknowledge two scientists on whose attitudes, habits, and knowledge I have built my understanding of science: my father, Dr. Stefan Antohi and Dr. Tudor Aurel Marian from the University of Bucharest.

Cherryl Ann, my wife, taught me the meaning of love, kindness, and patience while lifting me from my darkest fall. My son, Stefan, showed me that courage, and grace under hardship are not just words.

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Receptor - ligand interaction is a central process to biochemistry and pharmacology. By understanding the ways in which biologically active molecules interact with their receptors, not only profound insights in the cell physiological responses (involved in such diverse processes as nerve impulse propagation, plant tropisms, differentiation, immune response, etc) can be gained, but also, the stage when specific responses may be artificially induced in order to elicit desired effects like growth, or to cure diseases, can eventually be reached. Among ligands, the molecules that transmit signals between cells are a particularly exciting subject since hormonal activity is based on them and since cell - cell recognition seems to be a beautiful model (and maybe at the bottom of all explanations) for such intriguing processes as socialization and relationships between living beings.

When the signaling molecule is a peptide hormone/pheromone one of the subtle points is its flexibility. Due to the relative freedom of the φ and ψ angles in most amino acids, a peptide with, for

instance, 12 or 13 residues has a very large conformational space available. One of the major goals of current studies on such biologically active molecules is to determine the shape that they assume when triggering receptor response. [1, 2]

In many cases neither the exact shape of the receptor, nor the exact region to which the ligand binds are known. Thus, not only is it hard to mimic the conditions that the membrane environment is offering to the incoming ligand, but it is virtually impossible to guess the presence of hydrogen bonding regions, or of pockets that would likely accept aromatic side chains and thereby force the ligand into given conformations. In the absence of direct information concerning the bound peptide, an alternative is to determine as much information as possible on the wild type pheromone and/or judiciously chosen related molecules. A reasonable assumption is that the wild type pheromone molecular structure and solution or solid state conformations may offer hints towards what its favored shape might be at the receptor. Based on this assumption, several converging lines of investigation, both experimental and computational, can be used. The leading role is taken by the synthesis of analogs with appropriately designed constraints (initially suggested by the observations on the wild type pheromone) that can

only assume certain conformations in key regions of their structure. By studying the biological activity (or inactivity) of such analogs, precious clues on the importance of the conformations realized by the used constraints are gained. In parallel, it is possible to perform physical chemical studies of these analogs under different conditions (solution, powder, micelles) and then apply molecular modeling methods based on the results from these physical chemical experiments. The underlying hypotheses are that an active analog probably assumes a conformation that is similar to that of the pheromone (while an inactive one should hint towards properties that the pheromone does not have) and that the receptor cannot have such a drastic action as to change the conformations of the constrained regions.

Furthermore, one can study the physical chemistry of highly constrained model compounds that are not active but that comprise a region of critical interest of the analogs. The importance of such studies should not be underestimated: the simplified picture offered, for example, by a tetrapeptide compared to a 13-mer peptide may help elucidate properties not simply understood on the larger peptides. Moreover, as it will be shown in the Chapters on *ab initio* calculations as well as in the treatment of cyclized tetrapeptides, the

smaller molecules sometimes exhibit interesting traits of their own.

The *Saccharomyces cerevisiae* pheromones (the 13-mer peptide α-factor and the methylated and farnesylated 12-mer peptide a-factor) offer an excellent paradigm for the analysis of peptide ligands and their interactions with receptors. Besides the obvious advantages of being easy to grow and genetically well characterized, the yeast offers an accessible system that resembles mammalian hormones and receptors in many ways [1, 2]. Combined biological, synthetic, and spectroscopic studies [1, 2, 3, 4, 5] of α-factor pheromone analogs suggest that their activity is highly dependent on the conformation of the region between the 7th and 10th residues Specifically, all the active analogs seem to favor the formation of a type II β-turn in this region.

The purposes of this thesis are:

i) to quantitatively analyze, by *ab initio* calculations, the energetic basis of the physical chemical features exhibited [3] by biologically active analogs of the α-factor pheromone from *Saccharomyces cerevisiae* obtained by means of introducing side chain, local conformational hindrances [4, 5];

ii) to contribute to the understanding of the validity of using molecular modeling techniques in the investigation of peptide molecules (previously discussed

in [6]) and show practical applications in which such usage contributes to the comprehension or design of physical chemical or biochemical experiments;

iii) to blend computational techniques and physical chemical techniques such as high resolution solution NMR and solid state NMR in order to elucidate the solution conformation of biologically active analogs of the α-factor that had been previously synthesized and biologically assayed [7] and of some model compounds designed to mimic the putative turn region [8] as well as to qualitatively analyze by molecular modeling and to provide putative working models for the actions of biologically active molecules [9, 10, 11];

v) to consider the theoretical background of solid state NMR dipolar dephasing experiments, review the validity of natural abundance corrections, and eventually propose possible new variants for these experiments.

BACKGROUND

The life cycle of the yeast *Saccharomyces cerevisiae* comprises haploid (one DNA chromosome set) as well as diploid (two DNA chromosome sets) cell development. As depicted in Figure 1, the haploid cells are of two kinds, **a** and α which could also be considered two genders since besides asexual reproduction (entailing vegetative growth, mitosis, and separation into mother and daughter cells by budding), they can also mate. The mating is strictly heterosexual and it involves the cell fusion of an **a** cell with an α cell, resulting in the formation of a zygote, which at its turn can propagate into a diploid line by budding or undergo meiosis and sporulation ensuing - after spore release - a new haploid phase of the cycle [1, 2]. The way that the two kinds of cells recognize each other is biochemical: the **a** cells release the S-Farnesylated and carboxy-terminus methylated dodecapeptide [2]:

NH_2-Tyr-Ile-Ile-Lys-Gly-Val-Phe-Trp-Asp-Pro-Ala-

-Cys(Farnesyl)-COOMe

known as **a**-factor, while the α cells secrete the tridecapeptide [2]:

NH_2-Trp-His-Trp-Leu-Gln-Leu-Lys-Pro-Gly-Gln-Pro-Met-Tyr-COOH

known as α-factor.

While the α-factor has no noticeable action on α cells, it has dramatic effects on the **a** cells: it arrests cell growth in the G-1 phase of the cell cycle by a mechanism that involves specific inhibition of DNA replication, it increases cell surface agglutinability towards the cells of opposite mating type, and it initiates an aberrant elongation of the **a** haploids - termed shmooing. Similar results are obtained when assaying α cells with the **a**-factor.

The α-factor receptor, i.e., the molecule in the **a** cell membrane to which it binds in order to trigger the cascade of changes described above, is a 431 residue protein, coded by a gene known as STE-2 [12, 13], with 7 folds comprising 4 extra cellular domains, 4 intracellular domains and 7 membrane domains [2]. After ligand binding, in order for the signal transduction to proceed, the receptor must be coupled to the heterotrimeric guanine nucleotide-binding protein (known as the G-protein), as was indicated by studies on mutant yeast strains [14]. The present model of the signal transduction pathway is that after the pheromone

binding, a conformational change of the pheromone occurs and the G-protein is activated resulting in the dissociation of the G_α subunit from the $G_{\beta\gamma}$ subunit. After being released, the $G_{\beta\gamma}$ subunit of the G-protein interacts with a number of other factors to prompt the biochemical cascade underlying the physiological responses described above (Fig. 2 adapted from [2]).

The same kind of membrane folding structure and signal transduction pathway was found for other eukaryotic cells receptors such as rhodopsin [15], the β_2-adrenergic receptors [16], and the nicotinic acetylcholine receptors [17]. This emphasizes the importance of understanding the α-factor STE-2 receptor system, a challenging and meaningful problem in itself, and also as a part of the wider picture of the G-protein coupled receptor family.

The approach of using constraints either in the form of bulky side chains or by covalent cyclization for the purpose of restricting the flexibility of biologically active peptides is widely used [18,] and it proved instrumental in gaining conformational information about such diverse systems as:

- The mammalian gonadotropin releasing hormone (GnRH), a hypothalamic decapeptide that stimulates the release of luteinizing hormone and follicle stimulating hormone by the pituitary gland and thereby regulates

ovogenesis and spermatogenesis. Active analogs that helped elucidating the relevance of a type II' β-turn centered around positions 6 and 7 were prepared by N-methylation of residue 6 [19] and by S-γ-lactam linking of residues 6 and 7 [20];

- Bombesin, a 14-mer peptide that stimulates amylase release by rat pancreatic acini cells. Active analogs were obtained by replacing Gly^{11} with *D*-Ala and by cyclizing the molecule by means of a disulfide bond connecting residues 6 and 14 [21];

- α-Melanotropin, the tridecamer that regulates secretion of melanin in the vertebrate skin by α-melanocytes. A turn spanning the region between residues 5 and 9 was hypothesized on the basis of the bioactivity of an analog with inverted chirality at position seven (*D*-Phe^7 instead of *L*-Phe^7) [22]. The hypothesis was confirmed by synthesizing the extremely active analog in which the residues 4 and 10 are linked by a Cys-Cys disulfide bond [23];

- Oxytocin, a 1 to 6 disulfide cyclic nonapeptide neurophyseal hormone that mediates uterine contraction and milk ejection. A potent antagonist was synthesized with the introduction of a second cycle by substituting Gln^4 with a Glu and Leu^8 with a Lys and linking their side chains by means of a lactam bond [24].

The presence of proline and glycine at positions 8 and 9 of the α-factor, a sequence that was associated with the formation of a type II β-turn [25,] suggested the synthesis of analogs and model compounds that would test both the biological effects and conformational consequences of different constraints around this region [1]. It was found that replacing the Gly^9 with a *D*-alanine resulted in an analog that had equal activity and equal binding affinity compared with the wild type pheromone [5, 26] and its nuclear magnetic resonance analysis confirmed the presence of a transient type II β -turn [4]. To stabilize the turn, cyclized 13-mer analogs with a lactam bond between residue *i* and residue *i+3* and cyclized tetramer model peptides that spanned only the turn region were synthesized and analyzed [7, 8]. To facilitate the cyclization, the Gln^{10} of the original α-factor sequence was replaced with aspartic acid or glutamic acid. In order to further investigate the effects of size of the cycle, beside lysine, diamino propionic acid (Dpr), diamino butyric acid (Dab), and ornithine were introduced at position 7 to give a series of 8 linear analogs (with no lactam bond between positions 7 and 10) 8 cyclic analogs and 8 cyclic model tetramers. The generic formulas for these molecules are illustrated in Figures 3a and 3b.

The study produced a surprisingly diverse series of molecules having very different biological activities and unexpected properties such as conformational isomerism observable on the observational NMR time scale. Because of their different biological activities, understanding the conformations of these molecules is an important step in the elucidation of the α-factor activity. Since they have very similar structures (the 13-mer analogs differ from each other by the presence or absence of some methylene groups) and since the cyclized peptides have a quite restricted conformational space, explaining the conformational differences between active and inactive compounds is a bridge towards understanding what may make the pheromone work or not work. Not only is it important to analyze the properties of the actual molecules, but, since the turn (residue 7 to residue 10) region seems to play such an important role in determining the activity of pheromone analogs, the energetics of this portion of the molecule which is ultimately resposible for its shape is a very interesting subject on its own. In this thesis the conformations of the tetramers will be analyzed by molecular modeling and constrained molecular modeling, some of the 13-mer analogs will be studied by a combined approach based on high resolution NMR and molecular modeling, the REDOR [27] (solid state NMR) measurement

of some significant distances in the wild type α-factor will be discussed, and *ab initio* calculations of some peptides relevant to the turn region will be carried out.

These systems were studied using theoretical and experimental methods. The theoretical methods consist of *ab initio* calculations (the theory as well as the techniques underlying this method are described in Chapter 2) and of molecular modeling studies (this approach is described in Chapter 3). The experimental methods used in the NMR approach are described in Chapter 4.

CHAPTER 2

Computational Techniques

The purpose of computational chemistry is to reproduce and eventually predict (to as high a degree of accuracy as possible) the structure, shape, and behavior of atoms and molecules by means of computer programs.

The first step in any kind of computational modeling of molecules is adopting a method for calculating the energy of a given conformation and providing a way for finding energy minima (a minimization method). The atoms are borderline particles, they are small enough so that a consistent and rigorous treatment of their motion should be based on quantum mechanics, but yet they are large enough to allow reasonable results to be obtained from classical mechanical calculations. Thus, from the onset, there are two very different ways of approaching computational chemistry based on the theoretical framework and the definition of particle state that are adopted:
i) *Ab initio* calculations start with the one-particle Hilbert space states given by wave functions

$$\Psi(\text{observable quantities})$$

on which Hermitian operators representing other observables act and are thus established as a quantum mechanical approach. The kinetic energy of a particle of mass m, for instance, is represented by the operator:

$$-\frac{\hbar^2}{2m}\Delta$$

which is the quantum mechanical form of the classical expression

$$\frac{p^2}{2m}$$

ii) the so called molecular mechanics/molecular dynamics methods are based on the one-particle classical states described by continuous, simultaneously measurable parameters

$$(x,\ y,\ z,\ v_x,\ v_y,\ v_z)$$

and, consequently, they are based on a classical approach. If the molecule is fixed (i.e., all the velocities are zero and the kinetic energy is zero) and only the variation of the potential energy with spatial configuration is considered, the study is termed **molecular mechanics.** If the atoms are allowed to move, and thus the kinetic energy and the temperature of the

system are not zero, the computation is considered a **molecular dynamics** study.

In both the classical and quantum mechanical based approaches, the introduction of the interactions between particles brings in the common and unavoidable need for approximations. Although their electron shell structure and mutual interactions are based exclusively on electromagnetic forces which are very well understood (the contributions from other interactions can be ignored without losing anything from the accuracy of the calculation because their effects are beyond the experimental limits as far as chemistry and biochemistry are concerned), the atoms, especially when aggregated in molecules, which is the case of interest in these studies, are and form complex objects for which no exact solutions of the equations of motion can be written. It is here that the terms force field, approximation schemes, and perturbation methods make their appearance. A force field is an alternative, more mathematically tracTable, potential energy expression. Using the full electromagnetic interaction potential ($e_{i,j}$ are charges, $d_{k,l}$ are dipole moments, and r are distances)

$$\sum_{i,j}\frac{e_i e_j}{r_{ij}}+\sum_{i,j}\frac{\mathbf{d}_k\cdot\mathbf{d}_l r^2-3(\mathbf{d}_k\cdot\mathbf{r}_{kl})(\mathbf{d}_l\cdot\mathbf{r}_{kl})}{r_{kl}^{5}}+\ldots$$

leads, in the cases of many-particle systems, to intracTable expressions. Thus, instead of starting from it, one considers various other ways of expressing the forces between the particles involved. For instance, the basis of the Hartree-Fock (HF) method which underlies most *ab initio* calculations is approximating part or the whole potential energy component of the Hamiltonian operator by regarding all the particles in the system as interacting with a single central field depending only on the distances between the particles and the center of the potential:

$$U_{HF}(r_1, r_2, r_3, \ldots)$$

The method is iterative and self consistent in that, at every step of the calculation, it first evaluates the particle states and energies and then, based on the new values, it reevaluates the interaction. Convergence is only considered to be reached when the particle states and the potential do not change from one step to the next. [28] These methods were first applied in atomic and nuclear physics. For the purpose of quantum chemistry, a very important step was the application of the self consistent field (SCF) methods to diatomic and polyatomic molecules.

In general terms, the Hartree-Fock method looks for a set of molecular orbitals that minimize the variational integral

$$E_{HF} = 2\sum_{i=1}^{n/2} h_i + \sum_{i=1}^{n/2}\sum_{j=1}^{n/2}(2J_{ij} - K_{ij}) + V_{NN}$$

with the sums being carried out over all the n/2 occupied spatial orbitals φ_i of the n-electron molecule (it is assumed here that a "usual" molecule has an even number of electrons; special cases like NO or NO_2 must be considered separately). The one particle operators are defined by the following equation:

$$h_i \equiv \left\langle \varphi_i(1) \middle| -\frac{1}{2}\nabla_1^2 - \sum_\alpha \frac{Z_\alpha}{r_{1\alpha}} \middle| \varphi_i(1) \right\rangle$$

and the two particle operators defined by the following two equations:

$$J_{ij} \equiv \left\langle \varphi_i(1)\varphi_j(2) \middle| \frac{1}{r_{12}} \middle| \varphi_i(1)\varphi_j(2) \right\rangle$$

$$K_{ij} \equiv \left\langle \varphi_i(1)\varphi_j(2) \middle| \frac{1}{r_{12}} \middle| \varphi_j(1)\varphi_i(2) \right\rangle$$

In technical (numerical) terms, the whole procedure is reduced to a eigenvalue problem:

$$\hat{F}(1)\varphi_i(1) \equiv \varepsilon_i \varphi_i(1)$$

with

$$\hat{F}(1) = -\frac{1}{2}\nabla_1^2 - \sum_\alpha \frac{Z_\alpha}{r_{1\alpha}} + \sum_{j=1}^{n/2}\left[2\hat{J}_j(1) - 2\hat{K}_j(1)\right]$$

where the two one-particle operators, known as the *Coulomb operator* ($\hat{J}_i$) and the exchange operator ($\hat{K}_i$), are defined by:

$$\hat{J}_j(1)\varphi_i(1) \equiv \varphi_i(1)\int dv_2 \frac{|\varphi_j(2)|^2}{r_{12}}$$

and

$$\hat{K}_j(1)\varphi_i(1) \equiv \varphi_j(1)\int dv_2 \frac{\varphi_j^*(2)\varphi_i(2)}{r_{12}}$$

Obviously, $\hat{F}$ is a one particle operator and thus the many mututally interacting particles problem has been reduced to one of many particles evolving in an external field provided by the two interaction terms in $\hat{F}$. The self consistent feature of the problem is uncovered by analyzing $\hat{J}_i$ and $\hat{K}_i$, which are one particle

operators that change if the set of orbital functions (φ_i) is changed. Thus, at every step of a successive approximations procedure the form of $\hat{F}$ changes (because both $\hat{J}_i$ and $\hat{K}_i$ change) and the ideal solutions of the Hartree-Fock system of equations are supposed to be eigenvectors of this continuously changing $\hat{F}$ operator.

The eigenvalues (ε_i) of this system of equations are the orbital energies of the molecular electrons and their expression is:

$$\varepsilon_i = h_i + \sum_{j=1}^{n/2} (2J_{ij} - K_{ij})$$

which, after summation over the doubly occupied orbitals and substitution in the total energy expression, gives:

$$E_{HF} = 2\sum_{i=1}^{n/2} \varepsilon_i - \sum_{i=1}^{n/2}\sum_{j=1}^{n/2} (2J_{ij} - K_{ij}) + V_{NN}$$

From a practical point of view, considerable simplification of these equations was achieved by introducing standard complete sets of basis functions, χ_i. Although formally identical to the original problem, the matriceal set of equations resulting from the application of this procedure is better organized and it provides the framework for a systematic and homogeneous

approach to virtually all of these types of calculations. The actual system of equations reads:

$$\sum_k c_{ki}\left(F_{jk} - \varepsilon_i S_{jk}\right) = 0$$

with

$$c_{ki} = \langle \chi_k | \varphi_i \rangle$$

the coefficients of the linear (Fourier) expansion of the molecular orbital functions (φ_i) in terms of the standard complete set functions (χ_k),

$$F_{ki} = \langle \chi_k | \hat{F} | \chi_i \rangle$$

the matrix coefficients of the Hartree-Fock operator, $\hat{F}$ as expressed in the basis of the χ_i vectors, and

$$S_{ki} = \langle \chi_k | \chi_i \rangle$$

the normalization coefficients of the system (that may or may not be convenient to be brought to othonormal form).

With these notations, the problem is brought to the solution of a set of homogeneous linear equations that gives rise to the secular equation:

$$\det(F_{ij}-\varepsilon_i S_{ij}) = 0$$

This system of equations (known as the Hartree-Fock- Roothaan equations) together with the entailed secular determinantal equation constitute the basis of the modern computational approach to quantum chemistry.

This approach provides good energy values in the sense that comparing the minimized energies obtained for two different conformers of a structure gives a good indication of which one of them is actually more sTable [29, 30, 31]. The drawback of the method is that due to the extreme complexity of the equations involved, only relatively small molecules and virtually no dynamics can be considered.

As opposed to the HF approximation discussed above, the force fields used in molecular mechanics and molecular dynamics are of a completely empirical nature in that they consist mostly of potential wells around experimentally determined parameters (for instance X-ray determined bond lengths, bond angles and even torsions). Although it is essentially also a self consistent method, a crucial difference lies within the fact that

the equations are classical and thus, much less complex. The main consequence is that the size of the molecules that are tracTable in this system is considerably larger, standard packages, like Sybyl or Insight being able to treat molecules with ~10^4 atoms [32]. Because of the very design of the computational system, which is meant to reproduce well empirically determined geometries rather than to account for actual interactions, the energy values obtained from molecular mechanics (or molecular dynamics) calculations are much less reliable than their *ab initio* counterparts. Consequently, only very cautious use of the numbers obtained from such calculations should be made. In many cases, direct experimental evidence will throw off many conclusions about the relative stability of different conformations as deduced from this quasi-classical approach. In general, having two well (molecular mechanics) minimized conformations of a molecule means only that neither of the two structures exhibits unaccepTable features (like absurd bond angles or bond lengths). Comparing the two minimization energies can be misleading. Therefore, these methods are seldom used as stand alone procedures to produce/analyze new conformations, but rather, they are employed in the interpretation of physical chemical data. A typical example is provided by nuclear magnetic resonance

Overhauser spectroscopy data [34] (experimentally determined proximities) that can be input as constraints so that the molecular modeling results in a structure with a conformation that is not only accepTable (i.e., it has good bond lengths, good bond angles and it does not violate van der Waals radii), but also close to the one adopted by the actual molecule [35, 36, 37].

Although using various techniques (Steepest Descents, Conjugate Gradients and Powell to quote only a few of the most popular), all the methods for finding the conformational energy minima starting from a given conformation (once the way of calculating the energy is selected, as described above), are gradient based [32, 33]. Specifically, the potential energy and its derivatives with respect to the structural parameters are calculated and new conformations are considered by giving small variations to the conformation in the direction pointed to by the largest negative gradient. The procedure is repeated until a point of minimum (where all derivatives are meeting a convergence requirement) is found.

The key expression in the above description of minimization is *small variations*. Because of this feature of the technique and because the potential energy surface is a complex geometrical object involving many local extrema, there is no guarantee (and in fact,

it is highly unlikely) that, when starting with a given input conformation, the minimization procedure is going to take the conformation to the global minimum. Moreover, as discussed above, in the case of the classical approach, it may be that the global minimum does not carry much significance and it is argued [39] that even if the computed global minimum energy conformation had relevance in determining the most sTable shape of a ligand, the actual biologically active structure may be different (a minor conformer or one that is induced by the interaction with the receptor). Thus, due to the inherent flexibility of most interesting ligands (implying highly complex potential energy surfaces) and due to biochemical arguments, any molecular modeling study must involve a "searching" procedure [18, 22, 39].

The simplest, most efficient and, in general, the perfect way to conduct a search is a small angle window grid scan search, that is considering all possible conformations generated when the torsion angles are varied by a step of e.g., 60^{o} and thoroughly minimizing them. This insures a homogeneous probing of the conformational space and precise identification of the global and local minima. Like most perfect things, in the case of interesting molecules, this procedure is not applicable due to the huge amount (increasing by a

factor of 6 with each rotaTable bond) of computational effort involved.

A second very powerful method of searching for interesting conformations is random sampling [32, 33, 40, 41]. In essence, it consists of randomly selecting various conformers and minimizing them. This method is at the basis of such popular schemes as Monte Carlo simulations and Diana. It may be efficient in probing constrained molecules [39, 41], but in the case of highly flexible, larger molecules, it simply produces too many structures for any kind of analysis to be feasible and good physical results are mandatory both for performing the actual searches and for analyzing their outcome.

The third type of search methods (known as annealing or Boltzmann searches) is based on high temperature dynamics [23, 36]. Since during a dynamics simulation the molecule is moving, it can change conformation and if the motion is ample, barriers between potential energy minima can be overcome. A dynamics search at temperatures between 500 and 1,000K is considered an accepTable procedure for generating a reasonable number of conformational candidates [35]. Technically, the methods consist of consecutively "heating" (i.e., increasing the average kinetic energy of the atoms) and "cooling" (i.e., decreasing the

average kinetic energy) of the atoms in the model and generating a structure at the end of each cooling cycle [33]. The generated structures can be further minimized and searched until the desired degree of compatibility with the experimental data is attained.

Once a good candidate for the molecular conformation is found, dynamics in vacuum and in solvent at 300K can be run and the quantitative attributes exhibited in the simulation (hydrogen bonds, average dihedral angles, average distances) can be compared with the experimental results (temperature coefficients, coupling constants, NOE distances) [35]. If the comparison is satisfactory, the dynamics simulation is considered a good representation of the actual conformational behavior of the studied molecule and, based on it, inferences about such characteristics as the presence of turn, sheet, or helical regions can be made or strengthened. The conformational features discovered can be further correlated with the biological action of the molecules and, eventually considered in the design of future physical chemical experiments and/or syntheses. Figure 4 provides a flow chart that summarizes the heuristic process going back and forth between experiment and computation.

CHAPTER 3

Ab initio calculations on Pro-Ala and Pro-Gly dipeptides - a quantitative evaluation of the effects of side chains on conformational preferences

As mentioned in the Introduction, an important working conjecture in the study of the biologically active conformations of yeast pheromones is that the region spanned by amino acids 7 to 10 of the α-factor is prone to adopt a turn conformation and that this shape plays an important role in the activity of the pheromone. In order to test this hypothesis analogs in which Gly^9 was replaced with the less flexible amino acids *D*-Ala and *L*-Ala were synthesized. It was found that the [*D*-Ala^9] α-factor analog was a considerably more potent agonist than the [*L*-Ala^9] α-factor analog, the former being equally active with the wild type pheromone while the latter was ten-fold less active [5]. This difference has been correlated with the relative tendencies of the *D*-Ala and the *L*-Ala containing analogs to assume transient β-turns as judged by solution NMR spectroscopy [41, 42] and vibrational circular dichroism [3]. While the *D*-Ala containing 13mer exhibited a transient type II β-turn observable by both NMR and VCD,

the *L*-Ala containing 13mer possessed a short lived type I β-turn conformation observable only by the short time scale technique VCD [3].

The secondary structure of a peptide chain is the result of a delicate interplay of local steric and electronic effects, hydrogen bonding, and interactions with the medium [43]. In this case, it is the local effect of the methyl side chain that is responsible for the dramatic physical chemical and biological differences between the *D*-Ala^9 and *L*-Ala^9 13mer analogs [4, 5, 41].

In this Chapter we investigate the local effect of different chiralities of an alanyl residue in a Pro-X sequence by applying *ab initio* calculations to determine the preferred conformations of *L*-Pro-Gly, *L*-Pro-*L*-Ala and *L*-Pro-*D*-Ala dipeptides. This furnishes a direct quantum mechanical assessment of the energetic consequences of a methyl side chain in the *i+2* position of a turn. It should be emphasized that none of these dipeptides spans a complete turn region. Nevertheless, we suggest here that, by comparing the favored conformations of these dipeptides we can gain insight into the energetic basis of the behavior of the α-factor analogs since the structural difference in the center of these larger molecules is the same as the difference between the dipeptides considered in these calculations.

The amino acids proline, glycine and alanine [29], the Gly-Ala dipeptide [29] as well as other dipeptides [30] have been examined previously with similar calculational methods. Here, we attempted to find the geometry-optimized conformations associated with all possible local minima of the unblocked dipeptides, their energies and geometric parameters.

Footnote: **The work discussed in this Chapter was published in reference 47.**

METHODS

The Gaussian 90 computer program [44] is used to solve the Hartree-Fock equations, discussed in detail in Chapter 2, for the *L*-Pro-Gly, *L*-Pro-*L*-Ala and *L*-Pro-*D*-Ala dipeptides, shown in Figures 5, 6, and 7. The calculations were performed by using the *ab initio* method with the 6-31G basis set. The purpose of the calculations is to obtain the optimum geometries of the dipeptide conformers and the geometry optimization is carried out by using the gradient method [45]. The starting conformations for the dipeptides included the *cis* and *trans* peptide bond isomers (ω=0° and 180° respectively), but since the *cis* peptide bond containing conformers had substantially higher energies (at least 12.5kcal/mol higher than the lowest *trans* conformer) we will only discuss the *trans* peptide bond conformers. The geometry of the proline ring was optimized keeping the $N_1C_2C_3C_4$ dihedral angle at 0° and optimizing the $C_2C_3C_4C_5$ angle, which accounts for the puckering of the proline ring. The different optimized conformers were obtained by a grid scan procedure, i.e., by giving the ψ of proline (the dihedral $N_8C_6C_5N_1$) and the φ of the X (i.e., Gly or Ala) amino acid (the dihedral $C_{10}C_9N_8C_6$) angles the initial values 0°, 60°, 120°, 180°, 240°, 300°. To illustrate, a starting conformation might

feature $\psi=0^{o}$ and $\varphi=0^{o}$, or $\psi=0^{o}$ and $\varphi=60^{o}$. A total of 36 initial conformations were examined for each dipeptide. The optimization of all molecular parameters was then carried out on each resulting initial structure. This procedure insures the proper identification of all local minima and of the global minimum for each compound.

Tables 1a and 1b show the total energies of the different conformers, as well as their relative energies correlated with the φ and ψ angle values. Tables 2a, 2b, and 2c show the optimized bond lengths; Tables 3a, 3b, and 3c exhibit the optimized bond angles; Tables 4a, 4b, and 4c, tabulate the dihedral angles of *L*-Pro-*L*-Ala, *L*-Pro-*D*-Ala and *L*-Pro-Gly, respectively. The Tables 5a, 5b, and 5c show the net atomic charges, as calculated via Mulliken Population Analysis. The optimized conformations of the lowest energy 4 conformers for each of the alanine containing structures and of the lowest *L*-Pro-Gly conformer are shown as follows: *L*-Pro-*L*-Ala in Figures 5a, b, c and d; *L*-Pro-*D*-Ala in Figures 6a, b, c and d; and, in Figure 7, *L*-Pro-Gly.

RESULTS AND DISCUSSION

As mentioned above, the *L*-Pro-Gly, *L*-Pro-*L*-Ala and *L*-Pro-*D*-Ala dipeptides all were found to have energetically preferred conformations with $\omega=180^{o}$, that is, they all adopt a *trans* peptide bond. The study of the *L*-Pro-Gly dipeptide produced only five energy minimized conformations. In contrast, *ab initio* minimization of the two alanine containing dipeptides resulted in 9 and 12 different optimized conformations, respectively. Thus, a first observable effect of the methyl side chain is an increase in the number of local minima available to the alanyl containing molecules. Each of these exhibits two *gauche* local minima around φ (Ala) for every *trans* local minimum, while in *L*-Pro-Gly for two of the three values of ψ(Pro) for which *trans* minima were observed no *gauche* optimized conformations were found.

As illustrated in Table 1, for both *L*-Pro-*L*-Ala and *L*-Pro-*D*-Ala, the lowest energy conformations are those with φ(Ala) close to 180^{o} (-158.4 and 162.8 respectively). However, whereas *L*-Pro-*L*-Ala features ψ(Pro) near -20^{o} for three of the lowest 4 conformations, the lowest 5 conformations of *L*-Pro-D-Ala have positive ψ(Pro). It is important to note that the

second lowest energy conformation computed for *L*-Pro-*L*-Ala exhibits almost perfect type I β-turn values: ψ (Pro)=-18.6° and φ(Ala)=-85.0° *vs* the ideal values of ψ(*i+1*)=-30° and φ(*i+2*)=-90°. In contrast, *L*-Pro-*D*-Ala has a low energy conformation (#4, 2.45 kcal/mol above the lowest energy structure) which has ψ(Pro)=+149.3°, φ(Ala)=+85.3°, values close to those for an idealized type II β-turn (ψ(*i+1*)=+120°, φ(*i+2*)=+80°). Furthermore, the only conformation of *L*-Pro-*L*-Ala remotely resembling a type II β-turn (#12, ψ=+83°, φ=+69°) is higher than the lowest energy structure by 13.8 kcal/mol, i.e., approximately as high as a *cis* peptide bond containing conformation. It appears then, that the quantum mechanical approach shows that a type II β-turn would be accessible to *L*-Pro-*D*-Ala but inaccessible to *L*-Pro-*L*-Ala.

A puzzling characteristic of the results is the fact that the study of the *L*-Ala containing dipeptide resulted in 12 optimized conformations while the *L*-Pro-*D*-Ala produced only 9. It is difficult to escape the temptation to associate this larger number of quantum levels with the observed higher flexibility of the corresponding *L*-Ala^9 α-factor analog 13mer. Specifically, evidence for any structure in the *L*-Ala^9 α-factor is found only using the very short time scale technique VCD [3] whereas, the *D*-Ala containing 13mer

exhibits a solution structure sTable enough to be probed by the slower NMR NOESY experiments.

As shown in Tables 4a, b and c, the puckering of the proline ring is substantial, with a dihedral angle $C_2C_3C_4C_5$ of approximately 20°. The calculated bond lengths and bond angles are within less than 2% of the experimental values [46] with some exceptions that are discussed below (see Tables 2a, b, c and 3a, b, c). The calculated C_{10} to O_{13} double bond in Ala (both *L* and *D*) is shorter than its experimental value by 0.04Å to 0.05Å and the C_9-C_{10}-O_{12} calculated bond angle is around 8° smaller than the X-ray result in both alanine containing dipeptides. These discrepancies together with the ones associated with the C_9-C_{10} bond length may be due to the fact that the experimental data are from blocked peptides in which the C_{10} is part of an amide bond (see legend of Tables 2a and 2b), while in our study the end group is a carboxyl. For the *L*-Pro-Gly case, where the experimental values are determined on a dipeptide with a free carboxyl, the X-ray measured C_{10}-O_{12} double bond length is very close to the one estimated here. The N_8-H_{22} bond length that is found to be 1.46Å-1.47Å in the X-ray experiments on both Ala containing dipeptides proves that H_{22} must be hydrogen bonded in those crystal structures and also offers a possible explanation for the differences between the experimental and calculated

bond angle values for virtually all angles involving the peptide H. We do not have a rationalization for the following: a) the C_9-C_{10} bond is overestimated by 0.04-0.05Å in the *L*-Pro-*D*-Ala while it is very slightly underestimated in the other two dipeptides (~0.007Å in *L*-Pro-*L*-Ala and ~0.03Å in *L*-Pro-Gly); b) the N_8-C_9-C_{10} bond angle is underestimated by 6° in *L*-Pro-*L*-Ala, 2° to 5° in *L*-Pro-*D*-Ala and both underestimated and overestimated in the *L*-Pro-Gly calculated conformations.

It may be remarked that some trends are in almost perfect accord with the experimental data, the most noticeable being the values of the N_1-C_5-C_6 (N(Pro)-Cα (Pro)-C(Pro)) and C_5-C_6-O_7 (Cα(Pro)-C(Pro)-O(Pro)) bond angles, which, although centered 4 bonds and respectively 3 bonds away from the methyl side chain are substantially affected by the chirality of alanine. As seen from the X-ray crystallography results, the first one of these two angles is smaller in the presence of *D*-Ala and larger in the presence of *L*-Ala while the second exhibits an opposite behavior - both trends being well reproduced by the calculation results (Tables 3a and 3b). The experimentally determined dihedral angles for the X-ray studied dipeptides are closest to the ones calculated in structure #6 for *L*-Pro-*L*-Ala (Table 1), structure #4 for *L*-Pro-*D*-Ala and none of the calculated *L*-Pro-Gly conformations. Thus, due to either long range

intramolecular H-bonding, or to the intermolecular interactions responsible for crystallization, the structure with *L*-Pro-*L*-Ala moiety described in reference [46] is overcoming a local intramolecular barrier of approximately 5.4kcal/mol and the structure with *L*-Pro-*D*-Ala moiety is overcoming about 2.45kcal/mol to adopt a conformation that is somehow favorable to the other interactions involved in the crystal. Interestingly, the *L*-Pro-*D*-Ala X-ray studied dipeptide exhibits a conformation very close to that corresponding to a type II β-turn.

In conclusion: 1) we show that meaningful inferences about the conformational preferences of peptides can be made on the basis of the *ab initio* study of the local minima; 2) due to the presence of an *L*-Ala at position *i+2* the formation of a type II β-turn is heavily disfavored by comparison to an extended or type I turn conformation (over 10kcal/mol) and, thereby, we give a quantitative computational confirmation of the modeling based observation in [25]; 3) the comparison with X-ray crystallography reveals some interesting features of the influences of the side chain and gives an estimate of how high an energy difference due to local hindrance can be overcome by other interactions; 4) we support the conclusion [5] that, since the *D*-Ala^9 α-factor analog is 10 times more active than its *L*-Ala

counterpart, a type II β-turn may be important for the biological activity of that 13mer.

Of course, the physical chemical experiments on the pheromone analogs were carried out in solvents (water, DMSO, chloroform) and the bioactivity assays involve both a medium and the presence of the receptor. Our *ab initio* calculations do not account for solvent effects. Nevertheless, the change in chirality at position 9 (replacing *D*-Ala^9 with *L*-Ala^9) should not alter considerably the ability of a 13mer to interact with a surrounding solvent and if it modifies its shape this should be due to local internal alterations of the potential energy of the molecule. This is why we suggest that we can rationalize the physical chemical behavior of the α-factor analogs by studying the favored conformations of the dipeptides.

CHAPTER 4

Ab initio calculations on Ac-*L*-Pro-*D*-Ala-NH_2 and Ac-*L*-Pro-*L*-Ala-NH_2 dipeptides - a quantitative quantum mechanical evaluation of the effects of side chain chirality on the conformational preferences of a complete turn region peptide model

The work considered in this Chapter is a direct extension of the problem considered in the previous Chapter. It was found [47] that *ab initio* calculations on the nonblocked dipeptides *L*-Pro-*L*-Ala and *L*-Pro-*D*-Ala support the fact that the presence of *D*-Ala favors type II β-turn like backbone angles (as opposed to *L*-Ala). Subtle, but nevertheless significant geometrical differences between the *L*-Pro-*L*-Ala and *L*-Pro-*D*-Ala diastereomers determined by X-ray analysis [46] were also well reproduced by the previous calculation [47].

Here we are considering the blocked peptides Ac-*L*-Pro-*L*-Ala-NH_2 and Ac-*L*-Pro-*D*-Ala-NH_2 that model a complete β-turn region in the sense that they can exhibit the characteristic 10-membered ring hydrogen bond closure which is thought to stabilize β-turn conformations. By computing the *ab initio* energies and geometric parameters of the conformers of these

structures we furnish a direct quantum mechanical evaluation of the role of the chirality of the alanyl side chain in the stabilization (or destabilization) of given conformations.

METHODS

The Gaussian 90 computer program [44] is used, as described in Chapter 3, with the same 6-31G basis set, to solve the Hartree-Fock equations for the Ac-*L*-Pro-*D*-Ala-NH_2 and Ac-*L*-Pro-*L*-Ala-NH_2 dipeptides, shown in Figures 8 and 9. Based on the experience gained by studying the nonblocked dipeptides, we considered only *trans* peptide bond conformations ($\omega=180^o$). The geometry of the proline ring was optimized keeping the $N_1C_2C_3C_4$ dihedral angle at 0^o and optimizing the $C_2C_3C_4C_5$ angle. The different optimized conformers were obtained by a grid scan procedure, i.e., by giving the ψ of proline (the dihedral $N_8C_6C_5N_1$), the φ of the alanine (the dihedral $C_{10}C_9N_8C_6$) and the ψ of the alanine torsional angles the initial values 0^o, 90^o, 180^o, and 270^o. For example, a starting conformation might feature ψ(Pro)$=0^o$, φ(Ala)$=0^o$, and ψ(Ala)$=0^o$; or ψ(Pro)$=0^o$, φ(Ala)$=0^o$, and ψ(Ala)$=90^o$. A total of 64 initial conformations were examined for each dipeptide. The optimization of all molecular parameters was then

carried out on each resulting initial structure. As discussed in the previous Chapters, this procedure insures the proper identification of all local minima and of the global minimum for each compound.

Tables 6a and 6b show the total energies of the four lowest energy conformers for each of the studied molecules, as well as their relative energies correlated with the ψ(Pro), φ(Ala), and ψ(Ala) angle values. Tables 7b, 7b, 8a, 8b, 9a, and 9b show the optimized bond lengths, bond angles and dihedral angles, respectively, for Ac-*L*-Pro-*L*-Ala-NH_2 and Ac-*L*-Pro-*D*-Ala-NH_2. Tables 10a and 10b display the net atomic charges, as calculated via Mulliken Population Analysis. The optimized conformations of the two lowest energy conformers for each molecule are shown as follows: Ac-*L*-Pro-*D*-Ala-NH_2 in Figures 8a and 8b; Ac-*L*-Pro-*L*-Ala-NH_2 in Figures 9a and 9b.

It should be emphasized that the two molecules analyzed in this work are diastereomers and that the results found are also valid for the two molecules that are their respective mirror images. To be specific, the mirror images of the conformations reported here for Ac-*L*-Pro-*L*-Ala-NH_2 are the lowest energy 4 conformations available for the Ac-*D*-Pro-*D*-Ala-NH_2 and their absolute and relative energies should be identical to those found

in this calculation; the same relationship exists between Ac-*L*-Pro-*D*-Ala-NH_2 and Ac-*D*-Pro-*L*-Ala-NH_2.

RESULTS AND DISCUSSION

In contrast to the nonblocked case [47], where the lowest energy conformations for both the *L*-Pro-*L*-Ala and the *L*-Pro-*D*-Ala dipeptides had the alanine in an extended structure, the presence of the acetyl amino terminus and amide carboxy terminus groups seems to accentuate the differences between the two diastereomers. With the exception of conformation #4 of Ac-*L*-Pro-*D*-Ala-NH_2 (Table 6a), in all the other conformations presented, the backbone dihedral angles of the second amino acid, φ(Ala) and especially ψ(Ala), take values that are far from the $\pm 180^{o}$ which would correspond to an extended structure. This may be due to the hydrogen bonds that can be formed in the presence of the blocking groups.

The lowest energy structure obtained for the *D*-Ala containing blocked dipeptide exhibits a good type II β-turn conformation with $\psi_2 = \psi(\text{Pro}) = 103.0^{o}$, $\varphi_3 = \varphi(\text{Ala}) = 126.2^{o}$, and $\psi_3 = \psi(\text{Ala}) = -9.8^{o}$ (Table 6a), close to the ideal values of 120^{o}, 80^{o}, and 0^{o}. The distance between oxygen 15 (O15) and hydrogen 29 (H29) in this molecule, displayed in Fig. 8a, is 2.2066Å (Table 7a) and suggests a strong hydrogen bond between these two atoms. The presence of this hydrogen bond is further supported by the longer covalent bond, of approximately

1Å between H29 and nitrogen 12 to which it is attached, the relatively large charges on O15 (-0.68 units, Table 10a) and on H29 (0.45 units, Table 10a) and by the near collinearity of atoms C14, O15, N12, and H29 which form a dihedral angle of 165.9°. This hydrogen bond is the typical 10 member ring closure that occurs in a β-turn. As can be seen from both Table 6a and Figure 8, oxygen 15 is also close (2.61Å) to hydrogen 24 that is attached to the peptide bond nitrogen. Although this hydrogen bond does not seem to be as strong as the former, and although it is not a classical characteristic of a type II β-turn, its presence singles out this conformation among all the minimum energy structures found in this work. Due to this ability to form two hydrogen bonds, one that closes a 10 member ring and another that closes a 7 member ring, the type II β-turn like conformation exhibited by Ac-*L*-Pro-*D*-Ala-NH_2 is not only the lowest energy structure found for this molecule, but it is also 1.82 kcal/mol lower in energy than the lowest minimum found for Ac-*L*-Pro-*L*-Ala-NH_2 (as it can be calculated from the data in Tables 6a and 6b) and at a considerable gap (5.77 kcal/mol) below the next available minimum for a Ac-*L*-Pro-*D*-Ala-NH_2 conformation. Both features reverse the trends observed in the nonblocked dipeptides where the lowest overall energy conformation belonged to *L*-Pro-*L*-Ala and there was a smaller gap between the

lowest energy conformations available to the *D*-Ala containing nonblocked dipeptide than those corresponding to the *L*-Ala nonblocked dipeptide. The potential presence of this bifurcated hydrogen bond, also observed in the modeling of some tetrapeptides cyclized around a Pro-Gly dipeptide [48, 49], and the energetic characteristics that it imparts to the molecule may account for the fact that of the two linear tridecapeptide α-factor analogs obtained by replacing Gly^9 with *D*-alanine and *L*-alanine, respectively, only the *D*-Ala^9 analog exhibits a type II β-turn like conformation sTable enough to be observed by high resolution solution NMR [4, 42]. The second lowest energy conformation found for the *L-D* dipeptide exhibits, unexpectedly, a distorted type I β-turn conformation with $\psi_2 = \psi(Pro) = -13.6^o$, $\varphi_3 = \varphi(Ala) = -74.8^o$, and $\psi_3 = \psi(Ala) = 54.6^o$ (structure #2, Table 6a), close enough to the defining values of -30^o, -90^o, and 0^o to allow for the characteristic 10 member ring formation. The small distortions at ψ_2 and φ_3 probably add up to favor a less axial position for the methyl side chain, while the large distortion exhibited by the ψ angle of *D*-alanine accommodates the formation of a hydrogen bond between O15 and H29 in the presence of the alanyl side chain. The distance between O15 and H29 in this structure is 2.397Å, the distance between H29 and

the N12 to which H29 is covalently bonded is 0.99Å, and the dihedral angle formed by the atoms C14, O15, N12, and H29 is 114.7°. All these values suggest that the hydrogen bond in Ac-*L*-Pro-*D*-Ala-NH_2 structure #2 (depicted in Fig. 8b) is not not very strong. However, this hydrogen bond offers the only conceivable explanation for the very existence of this optimized minimum energy conformation. Indeed, this structure does not correspond to any of the conformations found for the nonblocked dipeptides [47,] and its relatively low energy can only be accounted for on the basis of the relatively sTable 10 member ring closure made possible by the presence of the blocking groups. Since these two lowest energy conformations obtained for Ac-*L*-Pro-*D*-Ala-NH_2 (#1 and #2 in Table 6a) correspond to the two possible types of β-turns realizable when the *i+1* amino acid is proline, their examination provides a direct quantitative insight into the known experimental and theoretical inference [4, 25] that a *D* amino acid at postition *i+2* favors a type II like β-turn.

In contrast, the Ac-*L*-Pro-*L*-Ala-NH_2 exhibits 3 minimum energy conformations within an approximately 4 kcal/mol energy gap. This structure of the energy spectrum may be a result of the more numerous minima that the nonblocked dipeptide exhibits [47] and is consistent with the higher flexibility observed in the

physical chemical study of the linear 13-mer containing this dipeptide [3]. Interestingly, just as in the nonblocked case, the lowest energy conformation for the *L-L* blocked dipeptide does not correspond to a type I β-turn. Rather, the lowest energy conformation of Ac-*L*-Pro-*L*-Ala-NH_2 exhibits a structure in which a very strong hydrogen bond closes a 7 membered ring involving the acetyl carbonyl oxygen and the Pro-Ala peptide bond hydrogen. As judged by the small distance between the oxygen 15 and hydrogen 24 involved (2.11Å) and the covalent bond of 1.004Å between H24 and N8 (see structure 1 in Table 6b, Table 7b and Fig 9a) as well as by the large O15 and H24 charges from Table 10b, this bond seems to be the strongest of all the H-bonds exhibited by the computed structures. However, its presence seems to be mutually exclusive with the possibility of 10 membered H-bonded ring formation. In the type I β-turn conformation (structure 2 in Table 6b, Fig. 9b), the typical *i* to *i+3* hydrogen bond between O15 and H29 is present (2.47Å distance from O15 to H29 and 0.9932 distance between H29 and N12) but not as strong as the O15 to H24 H-bond in structure 1. Although O15 and H24 are relatively close (2.93Å), they appear unable to come close enough to form a 7 member ring hydrogen bond and the conformation does not exhibit the marked

energy minimum exhibited by the type II β-turn in the Ac-*L*-Pro-*D*-Ala-NH_2.

While the second lowest energy conformation for the Ac-*L*-Pro-*D*-Ala-NH_2 is a distorted type I β-turn like conformation, as discussed above, no minima corresponding to a type II like conformation were found for Ac-*L*-Pro-*L*-Ala-NH_2. This shows that although the blocked dipeptide allows for the possibility of a hydrogen bond closure, the *L* methyl group steric hindrance is strong enough to prevent the 10 member ring of the type II β-turn from forming. Based on fixed conformation calculations we estimate that the hindrance energy by which a type II β-turn is disfavored in Ac-*L*-Pro-*L*-Ala-NH_2 is approximately 13kcal/mol, *i.e.*, the same value found in the nonblocked dipeptides [47].

The calculated bond lengths and bond angles are very close to the X-ray crystallography experimental values [46], in ways quite similar to those of the nonblocked dipeptide calculations [47], with some exceptions that are discussed below (see Tables 7a, b and 8a, b). The calculated values involving H24, the peptide bond hydrogen, are very far (approximately 0.5Å and 10^{o}) off from the crystallographic data for both molecules . As discussed before [47], it is possible that this hydrogen atom is involved in some way in the intermolecular bonding responsible for crystallization.

This hypothesis would account for the unusually long NH bonds found in those experiments. The C_α to carbonyl carbon bonds are both substantially larger than the experimental ones in the *L*-Ala containing blocked dipeptide while they are very close to the X-ray values in the *D*-Ala containing blocked dipeptide.
In contrast to the nonblocked study [47], there are no problems in matching the calculated values involving the angles C9-C10-N12 and C9-C10-O13. This shows that blocking the carboxy end of the dipeptide to an amide solves the problems that arise from studying a free carboxyl (as assumed in [47]).

The same trends discussed in [47] involving the bond angles N1-C5-C6 (N(Pro)-Cα(Pro)-C(Pro)) and C5-C6-O7(Cα(Pro)-C(Pro)-O(Pro)) hold here too. The crystallographic results show the first one of these two angles to be smaller in the presence of *D*-Ala and larger in the presence of *L*-Ala, while the second exhibits an opposite behavior - both trends being well reproduced by the calculation results (Tables 8a and 8b). The dihedral angles determined for the X-ray studied dipeptides are very close to the ones calculated for the lowest energy structure of Ac-*L*-Pro-*D*-Ala-NH_2 (Table 9a), but they do not correspond to any of the calculated conformations in the case of Ac-*L*-Pro-*L*-Ala-NH_2. This may be due to the intermolecular interactions

responsible for crystallization. As shown in Tables 9a and b, the puckering of the proline ring seems to be more substantial for the *D*-Ala containing blocked dipeptides than for the *L*-Ala ones.

As mentioned in our methods section the results reported here for the two molecules Ac-*L*-Pro-*D*-Ala-NH_2 and Ac-*L*-Pro-*L*-Ala-NH_2 carry verbatim to the two mirror image structures Ac-*D*-Pro-*L*-Ala-NH_2 and Ac-*D*-Pro-*D*-Ala-NH_2 because there are exactly two chiral centers in the molecule and the quantum mechanical calculation has inversion symmetry.

CONCLUSIONS

1) Since the lowest energy structure for the Ac-*L*-Pro-*D*-Ala-NH_2 is a type II β-turn, the study of the blocked dipeptides supports in an unambiguous way the known [8] tendency of Pro-*D*-Ala moieties in peptides and proteins to adopt type II β-turn conformations.

2) The study provides further computational support for both the observed type II β-turn conformation in linear peptides containing the *L*-Pro-*D*-Ala moiety and for the conformational flexibility observed in linear peptides containing the *L*-Pro-*L*-Ala moiety.

3) The quantitative computational evaluation of the barrier introduced by the presence of an *L* amino acid at position *i+3* is consistent with the one calculated in [9] (around 13kcal/mol);

4) the comparison with X-ray crystallography, when considered in parall to the nonblocked dipeptide study, shows that the computational effects of blocking the carboxy terminus are in good agreement with the experimental data.

Footnote: **The work discussed in this Chapter is being published. The paper is in press with *Journal of Molecular Structure*.**

CHAPTER 5

Stand-Alone Molecular Modeling on α-factor Related Peptides; the Importance of Initial Conditions

INTRODUCTION

This Chapter describes some results obtained by performing stand-alone molecular modeling in vacuum on the α-factor pheromone and on the eight side chain cyclized tetramer molecules modeling its turn region. The term stand-alone is used here to designate a molecular modeling study that was not based on quantitative physical chemical data. The reason for performing these studies was utilitarian: at the time of the actual calculations the only data available concerning the α-factor pheromone were qualitative inferences about its propensity to form a turn centered around the residues Pro^8 and Gly^9. Nevertheless, working models of the peptide pheromone, as well as of its tetrapeptide turn model analogs were necessary in order to gain at least a qualitative insight on the shape of this family of molecules and to plan further experiments or, eventually design other analogs.

Although unsophisticated and to a large extent superseded by subsequent careful usage of quantitative NMR data [46, 47], these studies provided the impetus for many subsequent experimental and modeling efforts and they brought to light some simple, but yet nontrivial features of these peptides.

A first result worth mentioning was that the stand-alone computations seemed to support a significant conformational diversity of the cyclic tetrapeptides. For instance, a marked propensity to form an almost perfect type II β-turn was observed in the case of the T22 molecule (cyclo 1,4 Ac-Dab-Pro-Gly-Glu-NH_2), while other members of the family seemed to have very different conformational behavior. A second set of results that had a significant impact on ensuing experimental design was the observation that, if the 13-mer molecules were to be bent, specific distances were to be expected between the two ends of the polypeptide chain as well as between certain atoms within the molecule. Moreover, this study seemed to indicate the possibility of distinguishing between different turned/bent/extended conformations of the pheromone by measuring judiciously selected distances between given atoms of the molecule. To be specific, the study indicated that by determining a few carbon to nitrogen distances within the chain, one might discern, with a

reasonable degree of certainty, whether the peptide assumes one of some available conformations. Since such measurements are the exact object of new solid state NMR techniques, such as REDOR [27, 50, 51], these models were the at the origin of the idea to determine the conformation of the pheromone in powder and it was based on these models that the planning and design of the labeling that led to the results described in reference [50] took place.

Last, but not least, this study points to the importance that the initial conditions play, especially in the case of larger molecules as well as to the inherent shortcomings of computational methods and, in particular, of the classical force field evaluations of energy values.

METHODS

Cyclic Tetrapeptides

The molecules were built using the peptide builder module of the QUANTA software. The cyclization bonds were constructed by writing different specific scripts for each molecule considered. The resulting structures, exhibiting the correct atom connectivities, were minimized using the CHARMm force field. The minimum energy conformation obtained for each of the eight

molecules considered was then extensively searched by using the following procedure:

1) A random sampling (essentially Monte Carlo) search was conducted on all the backbone dihedral angles. The most relevant parameters were window=60° and samples=20.

2) The 20 molecules resulting from the search were minimized and analyzed for redundancy (*i.e.*, elimination of identical conformations).

3) The remaining conformations were submitted to steps 1) and 2) described above.

4) The computational cycle was considered complete when no new meaninful conformations were produced after going through steps 1) and 2) on all the available conformations.

5) Of all the minimum energy conformations identified for a given molecule, the lowest energy one was chosen and considered to be the likely conformation of the molecule.

The procedure involved writing numerous automation routines that enabled the computer to continuously go from one molecule to the next and provide at the end of each cycle a list of interpreTable results. Due to the lengthy and intricate nature of the code involved in such a routine we are not giving here any details. The author will provide them to any interested reader.

The α-factor

Because of its sheer size, and to the fact that there are no interresidue hindrances, a study similar to the one described above for the cyclic tetrapeptides was not possible for this linear 13-mer molecule. However, since the conventional wisdom, supported by the biological activity of cyclized analogs [4] seemed to point out that a turn in the residue 7 to residue 10 region of the molecule could have played an important role in its shape, the following three initial conditions were considered:

1) Linear peptide. All the backbone dihedral angles with the exception of the two proline residues were given the extended value (180°).

2) Type II β-turn. The Pro^8 and Gly^9 residues were put in a type II β-turn conformation (-60°, 120°, 80°, 0°) while the rest of the torsional angles were all given the extended value (180°).

3) Type I β-turn. The Pro^8 and Gly^9 residues were put in a type I β-turn conformation (-60°, -30°, -90°, 0°) while the rest of the torsional angles were all given the extended value (180°).

These three conformations were minimized and submitted to a search procedure that was similar to the one described above, with the exception that a Boltzmann

search with a mild temperature (300° to 500°) was used instead of the random sampling. For each conformation, those conformations that resulted from the search but had lost their initial shape in the critical (residue 7 to residue 10 region) were discarded.

The lowest minimum energy structure that resulted for each of these conformations was considered a good representative for the given conformation of the peptide.

RESULTS AND DISCUSSION

Cyclic Tetrapeptides

In very few words, the random sampling study of cyclic tetrapeptides indicated that a Glu residue at position *i+3* favors the formation of a type II like β-turn, while an Asp residue at the same position has no definite effect other than precluding the formation of a type II like β-turn.

These results are summarized in Tables 11a and 11b and the models of the peptides are shown in Figures 10 and 11.

From the point of view of the interpretation of biological effects, these results have the following very elegant connotation: since, invariably, all the

cyclized analogs having glutamine substituting the amino acid at position 10 of the pheromone - which is the *i+3* amino acid of the propsed turn region, one could as well simply state that the type II β-turn analogs are more active.

The problem with this simple and elegant interpretation is that, as it will be explained in subsequent Chapters and as already published [7, 8, 48, 49] neither circular dichroism, nor high resolution nuclear magnetic resonance seem to support such a simple conclusion for the solution behavior of these peptides. The only case in which all studies converge is the Dab-Glu (the T_{22} member of the family - so named because it has 2 methylenic carbons on each side of the cyclization lactam peptide bond.) This molecule is found by all the studies - should they be experimentally based or not - in the same, almost perfect type II β-turn conformation.

In the case of the other members of the family, the physical chemical experiments, and especially the solution NMR based constrained dynamics simulations seem to support a much richer conformational behavior than the stand alone computational approach, which points to the weaknesses exhibited by molecular modeling studiescarried out without experimnetal support.

The α-factor

The results of this study are summarized in Figure 12 and Table 12.

The most dramatic observations are the end to end distance of approximately 12.5Å and the delineation of critical carbon to nitrogen distances that could distinguish between different conformations adopted by the crucially important turn region. These calculated distances constituted the theoretical data underlying the design of the labeled analogs on which the solid state dipolar dephasing experiments described in subsequent Chapters were performed, as well as the basis on which the interpretation of these experiments was carried out.

Fluorescence transfer measurements (private communication from Dr. Haas *via* Dr. Naider) seemed to confirm that, in solution, the distance between the Trp^3 and Tyr^{13} aromatic side chains is the one expected for a bent molecule.

CHAPTER 6

Experimental Techniques Used for the Conformational Analysis of Peptides

PHYSICAL CHEMICAL TECHNIQUES

High Resolution Solution NMR

The high resolution solution 2D NMR techniques used in order to assign the proton resonances and subsequently obtain distance constraints that are the basis of conformational analyses were COSY, TOCSY, NOESY, and ROESY.

COSY

Correlated spectroscopy is the classical technique that provides connectivities between J-coupled vicinal protons [53]. The pulse sequence of the COSY experiment is drawn in Figure 13. It consists of two 90° pulses enclosing an evolution time t_1 and followed by an acquisition time t_2. The evolution time is gradually increased for each successive 1D acquisition. The density matrix derivation of the effect of the COSY pulse sequence on the spin system is given in many standard texts [54, 55.] The significance of the result

is that the transverse components of the spin quantum operators (of two vicinal J coupled spins I and S) I_x, S_x, I_y, and S_y are multiplied in the final form of the density matrix (which describes the spin motion during acquisition) by oscillating factors depending both on $\omega_X t_1$ and $\omega_Y t_2$. (with ω_X and ω_Y , the two resonance pulsations). Thus, the spectrum generated after Fourier transformation of the data acquired by applying this pulse sequence to a sample contains cross peaks linking the resonance frequencies of vicinal J-coupled spins [43]. Its usefulness is that all the pairs of resonances of vicinal protons are identified. The more sophisticated versions of the experiment (for example, DQ-COSY) offer very good resolution and, even with the advent of the very powerful technique TOCSY (see below), it still pays to analyze the simpler COSY spectra and confirm the assignments based on the more complex techniques. The reason for this is that, sometimes, crowding problems that may arise in the TOCSY spectra may be resolved by looking at a COSY type spectrum which has many fewer cross peaks.

TOCSY

The name is an acronym for TOtal Correlation SpectroscopY [54] and the purpose of this sequence is to connect all the resonances of all the spins (or, at

least, many spin resonances) in a J-system. Because of the broad range of connectivities that can be established by analyzing a TOCSY spectrum, it usually plays the key role in the process of any resonance assignment. In our work we will use it for the proton assignment of peptides, but it can also be utilized in the study of ^{13}C labeled proteins [55].

In the TOCSY experiment (Fig. 13), the mixing 90° pulse of the COSY sequence is replaced with a mixing sequence whose main part is a spin lock turned on for a time that inpeptides usually lasts for 30ms to 80ms. The main purpose of this "pulse" (in effect a sequence of pulses with very short delays between them) is to allow J-coupled spins to interact with each other by keeping them in the transverse plane for a long period of time. The net result is that a term of the density matrix containing a component of the spin I before the spin lock is applied, is turned into a sum of terms containing all the spins that are J-coupled to I (let them be named S_1, S_2, S_3, ...):

$$I_x \rightarrow S_{1x} + S_{2x} + S_{3x} + \ldots$$

The ensuing evolution gives rise to observable terms that upon Fourier transformation exhibit cross peaks at (ω_I, ω_{S1}), (ω_I, ω_{S2}), (ω_I, ω_{S3}), ... Longer mixing times (around 80ms) allow one to see many bond correlations. For instance, with a long enough mixing time a cross

peak between the NH of an amino acid like Nle and the protons on the ε-carbon can be observed. However, many times a shorter mixing time may be preferable (around 65ms) because not too many connectivities are lost and the spectrum is "cleaner", that is, a better baseline and fewer artifacts are obtained.

Because the protons in peptide molecules are not J-coupled across the peptide bonds, in order to completely and unambiguously assign all resonances, it is useful to complement the J connectivities obtained from COSY/TOCSY experiments with the through space proximity information from NOESY and/or ROESY experiments.

NOESY and ROESY

In order to perform any kind of experiment-based molecular modeling, one needs distance constraints. These constraints are obtained from experiments that measure the dipole-dipole (through space) interactions between spins. The basis of dipole-dipole interaction measurements is the nuclear overhauser effect in which one spin is selectively irradiated with the effect that the resonance amplitude of a spatially near spin is affected. The effect is distance dependent because, unlike the J-coupling term, the dipole-dipole interaction Hamiltonian depends on the inverse cube of the internuclear separation. Due to the fact that the

NOE is a second order perturbation effect, an overall inverse sixth power dependence is observed. Thus, by using an internal reference, for instance, the distance between two geminal protons, one can obtain distance constraints for the studied molecule according to the formula [34]:

$$r_1 = r_0\left(\frac{I_0}{I_1}\right)^{1/6}$$

with r_0 the internal reference (fixed distance), r_1 the "unknown" distance to be determined and $I_{0,1}$ the intensities of the effect for the two spin pairs.

With the advent of 2D NMR, bidimensional variants of NOE experiments were designed which have the great virtue of allowing one to study all the through space connectivities for a whole molecule without worrying about selective irradiation and on the same spectrum. Technically, there are currently many variants of pulse sequences that use cross relaxation to obtain spatial information about the studied molecules [56], but we will concentrate on just two of them.

The NOESY sequence (Fig. 13) consists of a 90° pulse that creates transverse coherences, an evolution time, during which the "first" (t_1) axis resonances are created, a 180° pulse followed by a mixing time during which the spins interact through space and a final 90°

"reading " pulse followed by acquisition, on which the "second" (t_2) axis evolution is recorded. The intensities of the cross peaks can be introduced directly in the formula presented above in order to obtain the distance constraints, provided that care is taken to keep unwanted effects to a minimum.

The main source of unwanted effects is spin diffusion, which is another name for multistep magnetization transfer arising from cross-relaxation. While the desired effect is generated by magnetization transfer between two spins I↔S, a multistep process is due to magnetization transfer between three or more spins, some of which act as intermediaries: I↔S↔X. There are two undesirable consequences of such a transfer: a) the results of the interactions of S with both I and X are distorted, b) the spin I may seem to interact with spin X although, in fact they are not directly exchanging magnetization and may be spatially far apart. In order to avoid such artifacts, the NOE build-up curve, i.e., the dependence of the NOE intensities on the mixing time must be studied and the mixing times for the actual constraints determining experiments must be chosen in the initial (linear) region of this curve.

Another problem that may occur with NOESY is that the effect is dependent on the molecular motion

(expressed by the correlation time, i.e., the time needed by a molecule to rotate by approximately 1 radian) and on spectrometer frequency. For any given spectrometer frequency, there is a motional region, corresponding to medium size molecules in which there is no observable cross relaxation (Fig. 14). One way around this problem is to work with a different frequency spectrometer. However, this is seldom practical for two reasons: a) the "bad" NOESY region is wide and consequently in many cases little improvement is obtained by changing spectrometer frequency, b) using a considerably lower frequency spectrometer obliterates all the resolution advantages introduced by a higher field spectrometer. Because of that, there are whole classes of molecules (for instance, tetrapeptides) for which NOESY cannot provide any conformationally relevant information.

Although the same care must be taken in eliminating spin diffusion distorted results, the rotating frame nuclear Overhauser effect, which is the basis of the 2D ROESY experiments, exhibits a different dependence on molecular size/motion and allows the study of molecules that cannot be approached by NOESY. The physical basis of the ROESY experiment [37, 57] is the through-space exchange of magnetization by spins which are parallel in the rotating frame. Just as in the TOCSY experiment,

constraining the spins to be parallel in the rotating frame is realized by means of a spin lock field. As a matter of fact, the pulse sequence for the ROESY experiment is almost identical to the TOCSY (Fig. 13) sequence and one of the major concerns when performing ROESY is to prevent the TOCSY cross peaks from distorting the cross peaks resulting from true through space interaction. This is realized by choosing spin lock field values and offsets that favor ROESY interactions and disfavor TOCSY interactions [58, 59]. It is known that setting the frequency of the spin lock field in the middle of the spectrum and using spin lock field values that tilt the magnetization by 90^{o} introduce strong J-coupling mediated (TOCSY) interactions, while setting the spin lock frequency on one side of the spectrum and using a spin lock strength that tilts the magnetization by 18^{o} to 32^{o} favors the through space (ROESY) interactions.

To obtain ROESY distances the same formula, that provides the NOESY distance constraints described abcve, is used.

The ROESY experiments also have the remarkable feature that they exhibit conformational exchange peaks [48]. To be more explicit, in the case of molecules that exhibit conformational isomerism on a time scale long enough to be observable by NMR, cross peaks between

proton resonances corresponding to the same proton on two different conformers can be observed. While the regular, through space, magnetization exchange cross peaks have signs opposite to those of the diagonal peaks, the conformational exchange cross peaks have the same sign as the diagonal. This feature proves invaluable in the assignment of resonances of minor conformers which are otherwise very difficult to correlate [8]. It may be emphasized that, while for carrying out the conformational anlysis of molecules that are large enough to be studied by NOESY, the ROESY technique may seem cumbersome, in the case of molecules that exhibit conformational isomerism the rotating frame NOE is irreplaceable. Although it is a more demanding experiment than NOESY, if properly obtained (*i.e.*, paying attention at the technical details described above) the ROESY spectra provide reliable information about the conformational features of the molecules studied [8, 58].

CHAPTER 7

Solution Structures of *i* to *i+3* Cyclized Model Peptides: Building Blocks Mimicking Specific Conformations

INTRODUCTION

The cyclization of peptides is one of the techniques of choice both for the conformational analysis and in the design of potent analogs. The usefulness of this method is based on the reduction of conformational freedom that occurs when the peptide backbone is constrained. The analytical objective is to obtain molecules that are less conformationally flexible and thus more prone to analysis, both by experimental and computational methods. The biochemical objective is to synthesize molecules that have a conformation close to the actual active conformation locked in. In academic terms, the cyclized molecules should present higher conformational homogeneity which, if indeed centered on a highly active molecular shape, might lead

to increased receptor specificity, increased agonist or antagonist potency, or extended resistance to enzymatic or proteolytic degradation [60-67]. Constraints attempted so far include chain end to chain end, side chain to chain end and side chain to side chain cyclization. Cyclization involving backbone atoms has only been attempted infrequently [68, 69] due to the synthesis difficulties involved and to the low likelihood of obtaining active molecules in this way. Chain end to chain end cyclization often results in biologically inactive analogs [70]. This seems to point to the conclusion that the C- and N-termini of the peptides are essential for an efficient interaction with the corresponding receptor since their elimination has such drastic effects on the peptide activity. On the other hand, both side chain to chain end and side chain to side chain cyclization seem to have consistently produced interesting analogs, as the following examples show. Side chain cyclization using cystine, resulting in the synthesis of an α-melanotropin analog constrained *via* a disulfide bridge [71,] gave a super potent

agonist. This success made the disulfide technique of conformational design exceptionally popular with many peptide chemists [70, 71]. Similarly, side chain to side chain *i* to *i+2* cyclization using amide bond formation was utilized to develop highly potent and selective opioid peptide analogs [72-74]. A longer range cyclization, *i* to *i+4* side chain lactamization, has been employed to stabilize helical segments and yielded highly potent and metabolically sTable growth hormone releasing factor analogs [75, 76], a human calcitonin analog [77], and a potent parathyroid hormone-related protein antagonist [78]. Furthermore, *i* to *i+3* lactam-bridging gave high affinity and selectivity for cholecystokinin analogs [79-81], increased anticoagulant activity in a thrombin binding hirudin analog [82], and enhanced antagonist potency in luteinizing hormone-releasing hormone [83, 84]. Despite these successes, prior to this work, the exact relationship between the size of the lactam ring and the conformation of the constrained region for *i* to *i+3* side

chain lactamization of turn regions in peptides had not been systematically examined.

The molecule whose analogs are studied in this thesis, the α-factor pheromone that mediates mate recognition in the MATa haploid *Saccharomyces cerevisiae* cells, offers an excellent framework for the investigation of the effects of cyclization on pheromone analogs. Furthermore, it is as well a good example to consider within the wider context of peptide hormone - receptor interactions [1, 2, 7, 13, 20]. To initiate signal transduction, the α-factor molecule binds to a receptor that is a member of the seven transmembrane G-protein coupled receptor family [2, 20]. As part of the efforts to elucidate the biologically active conformation of this peptide, eight stereochemically restricted tetrapeptide models mimicking the critical region spanning residues 7 to 10 are studied here. The conformational restriction is introduced by means of an amide bond between the side chains of the first and fourth amino acids. Biological studies, performed on the corresponding tridecamer peptides, revealed that the

activity and receptor binding affinity of those analogs were related to the size and composition of the lactam ring in the center of the pheromone [7]. In order to relate those biological results to the conformation of the lactam ring and to study the conformational freedom in peptides containing side chain to side chain amide bonds, N^{α}-acetyl carboxyl-amide terminal cyclic tetrapeptides with an XxxProGlyZzz sequence where Xxx = Lys, Orn, Dab, & Dpr; Zzz = Glu & Asp were synthesized [8]. Initial characterization efforts [7, 8] revealed the presence of multiple conformers in these constrained model compounds. CD studies in water, trifluoroethanol and methanol suggest that the cyclic peptides exist in β -turn like conformations, likely involving the central Pro-Gly dipeptide [8]. Cyclo^{1,4}[Ac-Dab-Pro-Gly-Glu-NH_2] appears to assume a type II β-turn based on comparison with CD spectra calculated for idealized β-turns. The other peptides in the series probably exist as distorted versions of this conformation. The analysis of the CD curves was complicated by the fact that CD spectra are additive in nature and these peptides exist as mixtures

of different conformers. To provide definitive information on the structures assumed by individual conformers these model compounds were subjected to biophysical analyses using 2D NMR and molecular modeling. In this Chapter a detailed picture of different solution structures of the major conformers of these *i* to *i+3* cyclized model tetrapeptides is presented.

Materials and Experimental Methods

Peptide Molecules. The peptides examined in this study were all more than 99% homogeneous and were characterized by HPLC, FABMS, and NMR spectroscopy by the authors of reference [8]. These peptides are all cyclic tetrapeptides and have *L*-amino acids. They are named as $Tetra_{xx}$ where the suffix xx indicates the number of CH_2 groups on the side chains of first and fourth residues, respectively. For example cyclo^{1,4}[Ac-Lys-Pro-Gly-Glu-NH_2] is designated $Tetra_{42}$ whereas cyclo^{1,4}[Ac-Dpr-Pro-Gly-Asp-NH_2] is designated $Tetra_{11}$ (Fig. 3a).

NMR procedures. NMR spectra were acquired in DMSO-d_6 on a Varian Unityplus 600 MHz NMR spectrometer. Peptides dried in an Abderhalden drying pistol over refluxing methanol under high vacuum for 72 hr were used for the NMR studies. DMSO-d_6 (100%) from Aldrich, Milwaukee, WI was used to make solutions for NMR studies. TMS was used as the internal reference for proton resonances. Carbon chemical shifts were

referenced against DMSO-d_6 at 39.5 ppm. Solutions of 0.5-20 mM peptides were used for assessing aggregation and 5 mM solutions for all conformational analyses. One dimensional ^{1}H-NMR Spectra for determining temperature coefficients were obtained at 295-325 K in increments of 5 K. Sample temperatures were controlled with the variable temperature unit of the instrument.

Complete proton resonance assignments were made [8] using COSY, RELAYED-COSY and TOCSY experiments. HMQC and HMQCTOCSY were used to assign unambiguously proton resonances of the conformationally heterogeneous cyclic tetrapeptides. ROESY [37, 57] experiments were used to extract interproton distances. ROESY spectra were acquired at 300 K in phase sensitive mode using the method of States *et. al.* [85]. Spectra were collected with 2K data blocks for 512 t_1-increments with a relaxation delay of 1.5 s and 64 transients for each t_1-increment. The spectral width in both dimensions was 6000 Hz. ROESY spectra were obtained using 32° flip angle spin lock pulses and resonance offset compensation to suppress TOCSY contributions [57, 58]. The effective

field strength for spin lock is 1.9 KHz. The ROESY spectra were recorded with five different mixing times *viz.* 50, 100, 150, 200 and 250 ms to generate ROE buildup curves. The ROE intensities obtained at 100 ms mixing time were found to be in the middle of the linear curves. Therefore, interproton distances were deduced from ROESY experiments carried out with 100 ms mixing time.

NMR data acquired on a Varian spectrometer, were processed on a SUN Sparc station *IPX* using VNMR software. Prior to Fourier transformation, ROESY time domain data were apodized using shifted sine-bell or Gaussian window functions in both dimensions and zero filled to 2K x 2K real points. Multiplication of the first t_1-increment by 0.5 prior to the second Fourier transformation was carried out to reduce the t_1 ridges in all ROESY spectra.

Modeling Procedures. *Vacuum structures*: The N^{α}-acetyl, carboxyl-amide, 1-4 side chain cyclized tetrapeptides structures were built with the BIOPOLYMER module of the SYBYL software on an INDIGO2 SGI computer

and then minimized using the AMBER force field. The amino acid models Lys, Orn, Dab, Dpr, Asp and Glu, used in cyclizations were custom made to be recognized by the AMBER force field using the dictionary options of the SYBYL package. Constrained searches and minimizations were conducted using the annealing procedure in SYBYL (consecutive dynamics at T=500K and T=20K) starting with the minimized structures. The constraints were input as ranges centered on the experimentally determined distance (±10% of the distance value). The results were analyzed and the conformations compatible with the experimental data were identified and further searched and minimized.

Solvent box and solvated tetrapeptides: The DMSO molecule was optimized by *ab initio* calculations using the 6-31G* [44] basis set. The resulting parameters *i*.e., charges: C, -0.703; S, 0.944; O, -0.784 and H, 0.221, 0.213, 0.188; bond lengths and bond angles: SC, 1.80Å; SO, 1.49Å; CH, 1.10Å; CSC, 97.7°; CSO, 106.7°; SCH, 109.5° were introduced into the Kollman All-Atom parameter set of the SYBYL software. A DMSO box of

molecules having these parameters was built and minimized using the AMBER (Kollman All-Atom) force field. This box was subsequently used to generate the solvent environment for the cyclic tetrapeptide computations. We used parallelepipedic (almost cubic) boxes. The typical length of a solvent box side was between 32 and 36 Å and contained around 325 molecules of solvent. The best (lowest energy) vacuum minimized structures resulting from the searches were solvated and minimized with periodic boundary conditions.

Dynamics: Constrained Molecular Dynamics (MD) simulations in vacuum and DMSO were performed on the eight tetrapeptides using the AMBER force field and the Verlet algorithm. The simulations were carried out at constant temperature under canonical ensemble conditions and were started from well minimized structures obtained as explained above. The final temperature of 300 K was reached in three steps: heating at 1000 K for 1 ps, cooling at 500 K for 5 ps, equilibration at 300 K for 14 ps and simulations at 300 K for 60 ps [35]. The equations of motion integration step was 1 and snapshots

were taken every 60 fs. For $Tetra_{42}$, which is the largest molecule among our model peptides, simulations were carried out with the same algorithm with an averaging period of 120 ps. There were no significant differences from the 60 ps results. Averages for the quantities of interest (distances and dihedral angles) were computed over the simulation (60 ps) periods and compared to the experimental data and are presented in the text and in Tables 13 to 20.

Results

Conformational heterogeneity in *i* to *i+3* monomeric lactams. Initial NMR analyses of these cyclic peptides showed one predominant resonance for each amide *NH* in $Tetra_{42}$, $Tetra_{41}$, $Tetra_{32}$ and $Tetra_{31}$ suggesting either one dominant isomer or fast conformational averaging on the NMR time scale [8]. In contrast, the coexistence of slowly interconverting multiple conformers in $Tetra_{22}$, $Tetra_{21}$, $Tetra_{12}$ and $Tetra_{11}$ peptides was established by *NH/NH* and $C^{\alpha}H/C^{\alpha}H$ exchange cross-peaks (same sign as the diagonal) in their corresponding ROESY spectra. The *NH/NH* exchange cross-peaks of major isomers A and B for $Tetra_{21}$ are shown in Figure 15. Heating the samples resulted in collapse of the multiple peaks to single resonances for each *NH* and $^{\alpha}CH$ proton. This effect was reversible upon cooling the samples and unequivocally establishes the occurrence of slowly interconverting conformers in these peptides.

$Tetra_{11}$ exhibited two isomers (A and B) in a 6:1 ratio. $Tetra_{22}$ and $Tetra_{21}$ showed the presence of three

isomers (A, B and C) in 16:3:1 and 32:23:1 ratios, respectively [8]. The NMR spectrum of $Tetra_{12}$ revealed the existence of four isomers (A, B, C and D) in a 20:6:4:1 ratio [8]. Isomer A of $Tetra_{11}$ and $Tetra_{21}$ and both isomers A and B of $Tetra_{12}$ and $Tetra_{22}$ adopt a *trans* configuration with respect to the Xxx-Pro peptide bond as indicated by characteristic ROE cross peaks from the $XxxC^{\alpha}H$ to the $ProC^{\delta}H$ protons. A *cis* configuration in isomer B of $Tetra_{21}$ and both the isomers C and D of $Tetra_{12}$ was evidenced by a $XxxC^{\alpha}H$ and $ProC^{\alpha}H$ cross peaks in their ROESY spectra. This is illustrated for $Tetra_{12}$ in Figure 16. Nothing can be definitely said regarding isomers C in both the $Tetra_{21}$ and $Tetra_{22}$ due to a lack of cross peaks. All of the conformationally homogeneous peptides ($Tetra_{42}$, $Tetra_{41}$, $Tetra_{32}$ and $Tetra_{31}$) exhibited a *trans* Xxx-Pro peptide bond. In the present investigation, solution structures were elucidated only for the *trans* isomers in each of the cyclic peptides. In the case of the conformationally heterogeneous peptides ($Tetra_{22}$, $Tetra_{21}$, $Tetra_{12}$ and $Tetra_{11}$) the major *trans* isomer was examined.

Temperature gradients of amide *NHs*. Temperature coefficients offer a means for assessing the presence of hydrogen bonds. The temperature coefficients for amide protons were calculated by a least-squares analysis of the temperature dependence of their resonances (Table 21). The *cis/trans* isomers for the tetrapeptides ($Tetra_{22}$, $Tetra_{21}$, $Tetra_{12}$ and $Tetra_{11}$) are conformationally distinct over the temperature range which is used for calculating temperature coefficients. Temperature coefficients for the major *trans* isomer of each peptide are presented in Table 21. The α_{NH} of the *i* residue is outside the cyclic system and is expected to be solvent exposed. As anticipated its temperature coefficient has the highest value (4.15 - 7.44 ppb/K) in each peptide. The temperature coefficient of the *i+3* residue α_{NH} is very low in seven of the eight cyclic peptides indicating its probable involvement in intramolecular H-bonding. The exception is the eighteen membered cyclic peptide, $Tetra_{42}$, which has a coefficient of 4.8 ppb/K. Each of these protons except

in $Tetra_{42}$ resonate up field (6.95 - 7.32 ppm) compared to other *NH*s in their corresponding peptides [8]. This suggests that these α*NH*s are not solvated by DMSO [86, 87]. The temperature coefficient of the Asp α*NH* in the Asp-containing tetrapeptides was found to be highest for $Tetra_{31}$ followed by $Tetra_{41}$ and $Tetra_{11}$ while it is smallest for $Tetra_{21}$ (Table 21). We noticed a corresponding upfield shift of the Asp α*NH* in these peptides (7.32, 7.26, 7.03 and 6.83 ppm, respectively) [8]. A similar trend was observed in the Glu-containing tetrapeptides. With the exception of $Tetra_{42}$ these observations further support the involvement of the *i+3* α*NH* in an intramolecular hydrogen bond. We also observed that the temperature coefficient of the *i+2* residue α*NH* is exceptionally low in $Tetra_{12}$ (0.82 ppb/K) implying that it is either strongly shielded from solvent or that it participates in a relatively sTable H-bond. In $Tetra_{42}$, $Tetra_{31}$ and $Tetra_{11}$ the moderately low coefficient of the *i+2* α*NH* (~ 3 ppb/K), also suggests moderate solvent shielding or participation in H-bonding. Similarly, the lactam *NH* displayed a

relatively low ($-\Delta\delta/\Delta T$) coefficient in only one cyclic peptide ($Tetra_{32}$, 2.28 ppb/K) and moderately low values in another four peptides suggesting that these amide protons were moderately solvent shielded.

$^3J_{NH-C^{\alpha}H}$ coupling constants and backbone dihedral angles derived from coupling constants and modeling. The $^3J_{NH-C^{\alpha}H}$ coupling constants were extracted from well digitized 1D proton spectra. The resolution of the spectra was enhanced by applying a Gaussian window function prior to Fourier transformation. Bystrov's Karplus equations [88, 89] were used to calculate the dihedral angles (ϕ) from these coupling constants (see the supplementary material, Table 21). The average dihedral angles computed from the 60 ps simulations in DMSO are generally within +/- 30° of at least one of the values derived from the coupling constants.

Ramachandran Plots. The dihedral angles of the turn region (ϕ_2, ψ_2, ϕ_3 and ψ_3) obtained by averaging over the 60 ps dynamics simulations are tabulated along with the temperature coefficients of $\alpha_{NH(i+3)}$ and two key interproton distances in Table 22. The ϕ_2 dihedral

angle is constrained to be around -60° by the pyrrolidine ring of the proline residue at position *i+1*. For $Tetra_{42}$ and $Tetra_{31}$ the ψ_2 dihedral angle is more than 20° lower than the value corresponding to a type II β-turn (120°). The average value of the backbone dihedral angles for Pro of $Tetra_{42}$ corresponds to an inverse γ-turn around this residue (*i+1* position, Table 22). The dihedral angles for 1000 discrete steps of the molecular dynamics simulation are shown in a Ramachandran map (Fig. 17a) and cluster near the values expected for the inverse γ-turn. The Ramachandran plot of the Gly residue in $Tetra_{32}$ is consistent with the presence of a γ-turn around this residue (Fig. 17b; Table 22). Interestingly, $Tetra_{31}$ has ϕ, ψ angles of Pro similar to those of $Tetra_{42}$ and ϕ, ψ angles of Gly similar to those of $Tetra_{32}$, implying the presence of both conformational features in this molecule (Fig 17c; Table 22). On the other hand, the Ramachandran map of Pro(*i+1*) and Gly(*i+2*) dihedral angles in $Tetra_{22}$ agree with a very good type II β-turn (Fig. 17d, Table 22.) In $Tetra_{12}$ the ϕ, ψ of both the Pro and the Gly are spread over the

type II β-turn and both the γ- and γ'-turn regions (Fig. 17e) suggesting the coexistence of several important conformations in the major isomer of this tetrapeptide. $Tetra_{21}$ exhibits the largest distortion for the Pro ϕ (-95° average on the solvent simulation). This indicates significant strain for the prolyl residue and may correlate with the fact that this molecule exhibits the largest percentage of *cis* isomer for any tetrapeptide (41%).

Interproton distances deduced from ROE intensities and from modeling. Low spin-lock field strength and small (32°) flip angle pulses for the mixing sequence were used while acquiring the ROESY spectra to minimize TOCSY contributions [58, 59]. The position of the offset was also chosen in such a way so that the *J*-coupling contributions are minimal. ROESY spectra with mixing time 50, 100, 150, 200, and 250 ms were used to obtain ROE buildup curves (Fig. 18). Average ROE intensities from both sides of the diagonal were used. The dependence of the ROE integrals on mixing time was seen to be approximately linear up till 150 ms.

Assuming isotropic motion and the isolated two spin approximation (ISPA) [34], the observed average integrals from 100 ms ROESY spectra were converted to interproton distances. The obtained distances were calibrated with respect to the distance between Gly methylenic protons (1.77 Å). Both intra- and interresidue ROESY cross peaks were obtained for all of the cyclic tetrapeptides. The *NH-NH(i+2; i+3)* interproton distance is the crucial distance which indicates the occurrence of a β-turn or a γ-turn. This is expected to be around 2.4 Å for both type I and type II β-turns and around 3.70 Å for a γ-turn. $Tetra_{42}$ is unique in this series of cyclic peptides in that its spectrum was devoid of this ROE cross peak (Table 22). The *NH-NH(i+2; i+3*) distance was found to be between 3.40 and 3.30 Å for $Tetra_{32}$ and $Tetra_{31}$, respectively, and 2.70 - 2.85 Å for the remaining peptides. An important interproton distance that discriminates a type I β-turn from type II β- and γ-turns is α_{CH}-$\alpha_{NH(i+1;\ i+2)}$. This is expected to be around 3.40 Å for a type I β-turn and between 2.20 - 2.30 Å for both a type II β-

turn and a γ-turn. ROE measurements for all peptides yielded a shorter distance (around 2.20 - 2.65 Å) between these two atoms, indicating the presence of type II β-turns and/or γ-turns. These measurements allow us to rule out the existence of type I β-turn-like conformations in any of the cyclic peptides.

Among the vacuum structures generated for these cyclic peptides, only those which satisfied NMR derived interproton distances (root mean square deviation being less than 0.055 Å/distance, Table 23) and temperature coefficients are selected for further studies in solvent. Table 24 exhibits the results of using the Karplus equation to calculate dihedral angles from coupling constants and comparing them with conformational averages. These structures were solvated with DMSO, further minimized and dynamics simulations were carried out (*vide supra*). The modeling distances and dihedral angles are presented in Table 22 where the NMR distances and modeling distances are also compared. The values from the modeling represent average values extracted from a 60 ps molecular dynamics simulation in

DMSO. In general, the average solvent dynamics values were in good agreement with the ROE distances (a *rmsd* of less than 0.05 Å/distance, Table 23). The relevance of the distance and dihedral angles to the conformations of the *i* to *i+3* lactams is given in the discussion section.

Discussion

Beta turns are a commonly reported conformational feature that have been associated with the biologically active state of many peptides [90-95]. In particular, a Pro-Gly sequence often occurs in type II β-turns [96, 97], and cyclization is attempted in many laboratories to eliminate conformational heterogeneity of linear peptides. The peptides involved in the present study are both cyclic and possess a central Pro-Gly dipeptide unit. Hence all of these cyclic tetrapeptides are expected to manifest a type II β-turn as their conformational feature if the side chain length variation, and/or ring size has no influence on backbone conformation. The NMR and modeling studies reported

herein suggest that the stereochemical situation is considerably more complex for *i to i + 3* cyclization around a central Pro-Gly unit.

Peptide conformations. Turns occur in peptides and proteins when a growing peptide chain folds back on its backbone. A 10-membered H-bonded ring comprising the carbonyl of the *i* th residue and the amide *NH* of the *i+3* residue is defined as a β-turn [25, 98]. An interprotonic distance of 2.40 Å between the α_{NH}s of the *i+2* and *i+3* residues is expected for both type I and II β-turns. The proximity of the $C^{\alpha}H$ *(i+1)* to the α_{NH} *(i+2)* 3.40 Å and 2.20 Å, respectively, discriminates between type I and II β-turns. Type I and II β-turns also differ in their torsion angles ψ_2 and ϕ_3 which are -30°, -90° and 120°, 80°, respectively. The differences reflect opposite orientations of the *i+1* carbonyl group with respect to the plane of the 10-membered ring. On the other hand, a 7 membered H-bonded ring between the carbonyl of *i* th residue and amide *NH* of *i+2* residue results in a γ-turn around the *i+1* residue. The interproton distance between α_{NH}s of the *i+1* and *i+2*

residues is 3.40 Å and the backbone dihedral angles ϕ and ψ of the *i+1* residue about which the γ-turn is located are 70° to 85° and -60° to -70° respectively [99, 100].

We have used the above considerations to model solution structures for a series of cyclic tetrapeptides using NMR derived constraints. The Ramachandran plots of the critical dihedral angles give a clear representation of the variations found in an ensemble of 1000 structures. The rms deviations for the average distance and the torsional angle in these ensembles are presented in Table 23. The consideration of Table 23 in conjunction with Tables 13 to 21 shows the average interproton distances obtained from the analysis of the vacuum and DMSO simulations of these peptides to be in very good agreement with their corresponding ROE derived distances. H-bonding patterns observed in the dynamics simulations of these molecules support the temperature coefficients obtained from NMR for the corresponding amide *NH* protons (*vide infra* and also Tables 21 and 22). However, careful scrutiny of the dynamics simulations shows that several of these

cyclic peptides exist as ensembles of coexisting structures involving both β- and γ-turns. The conformation of each peptide is discussed separately below.

Cyclo^{1,4}[Ac-Lys-Pro-Gly-Glu-NH$_2$] (Tetra$_{42}$) : This structure is unique in the series in that no ROESY cross peak is observed between the *i+2* (Gly) and *i+3* (Glu) *NH*'s (Table 22). The large (4.80 ppb/K) temperature coefficient of the *i+3 NH* suggests that it points outside the cyclization ring and consequently the *i+2* residue of this tetrapeptide is far from adopting a type II β-turn conformation. On the other hand, the moderately low temperature coefficient of the Gly (*i+2*) *NH* (3.14 ppb/K) suggests that this atom is somewhat solvent shielded or inside the lactam cycle. The MD simulation in both vacuum and DMSO revealed a 7-membered ring centered around Pro and closed by a hydrogen bond between the Gly *NH* and Lys CO (Figure 19a). The average ϕ and ψ angles of Pro, -77° and 77°, respectively in the vacuum simulation and -77° and 82° in DMSO, correspond to a γ'-turn conformation. As seen in the simulation,

except for this part of the molecule, the structure seems to be very flexible and few constraints were found in the ROESY analysis.

Cyclo^{1,4}[Ac-Orn-Pro-Gly-Glu-NH_2] (Tetra$_{32}$) : The distinctive features of Tetra$_{32}$ are the large ROE distance (3.40 Å) between the *NH*'s at the *i+2* and *i+3* positions, the low temperature coefficient of the Glu *NH* (0.90 ppb/K) and the relatively low temperature coefficient of the lactam *NH* (2.30 ppb/K, the lowest of all side chain *NH*'s in the series). The vacuum and DMSO simulations based on the ROE constraints seem to accommodate these facts very well. The simulation revealed a very sTable H-bond between the Glu *NH* and the CO of Pro that closes a 7-membered ring around Gly in a γ -turn like conformation (Figure 19b). The average dihedral angles ϕ and ψ of Gly, 84° and -68° in the vacuum simulation and 81° and -68° in the DMSO simulation, correspond to a C_7 γ-turn conformation. The Glu *NH* also exhibited some ability to H-bond to the side chain CO of Glu, forming a transient intraresidue 7-membered ring. The side chain *NH* of ornithine formed an

H-bond with the acetyl CO. Although such interaction gives rise to a 9-membered ring that is outside the cyclization ring, this H-bond was sTable throughout the dynamics, and it provides the only possible rationale for the low temperature coefficient of this side chain amide bond.

Cyclo^{1,4}[Ac-Dab-Pro-Gly-Glu-NH_2] (Tetra$_{22}$) : The Glu *NH* of this molecule exhibited the lowest temperature coefficient (-1.80 ppb/K) of the series while the remaining *NH*'s had high temperature coefficients (between 5.7 and 7.6 ppb/K; Table 21). The Pro αCH to Gly *NH* and Gly *NH* to Glu *NH* distances (2.19 and 2.70 Å, respectively) together with this very low *NH* temperature coefficient suggest that the backbone of this molecule adopts a relatively sTable type II β-turn conformation. This conclusion agrees well with the CD results which were run in different solvents and reflect the average for all conformers of this peptide [8]. The average interproton distances obtained from the analysis of the vacuum and DMSO dynamics simulations of this molecule are in very good agreement with its ROE derived

distances (Table 23 shows the lowest *rms* deviations for this structure.) The dynamics showed a very sTable H bond between the Glu *NH* and the Dab carbonyl. The average backbone dihedral angles ϕ and ψ of Pro and Gly are -56°, 122°, 82° and 19°, respectively, in vacuum and -71°, 126°, 87° and 5°, respectively, in DMSO. They correspond to an almost perfect type II β-turn around the Pro and Gly residues (Figure 19d).

Cyclo^{1,4}[Ac-Dpr-Pro-Gly-Glu-NH$_2$] (Tetra$_{12}$) : The distance between the Gly *NH* and the Pro $^{\alpha}$CH in the major isomer of this molecule is larger than in any other tetrapeptide (2.65 Å). The Gly *NH* also has the lowest temperature coefficient (0.82 ppb/K) of all Gly *NH*'s in the series (Table 21). Together these two facts suggest that in this conformer the Gly *NH* is oriented towards the lactam ring. The vacuum and solvent simulations show that this *NH* is able to H-bond to the CO of Dpr, forming a 7-membered ring that is highly populated during the dynamics evolutions. The temperature coefficient of the Glu *NH* is found to be 0.22 ppb/K, also indicating strong H-bonding. Indeed, both

simulations show this *NH* forms hydrogen bonds with the carbonyls of Dpr (corresponding to a typical type II β-turn 10-membered ring) and Pro (a 7-membered ring around Gly) that are very sTable throughout the dynamics. On the average, the molecule seems to adopt a type II β-turn and at the same time oscillates between two extreme conformations, one corresponding to a γ'-turn around Pro and the other to a γ-turn around Gly (Figure 19e). This oscillation among three structures results in average dihedral angles that are close to those expected for a type II β-turn and a somewhat lengthened Pro$^{\alpha}$CH-Gly$^{\alpha}$*NH* distance. The averaging over several structures is observed as a large distribution of ϕ, ψ angles in the Ramachandran plot for this tetrapeptide (Fig. 17e).

Cyclo^{1,4}[Ac-Lys-Pro-Gly-Asp-NH$_2$] (Tetra$_{41}$) : The short interproton ROESY distances from the *i+2 NH* to the *i+1* $^{\alpha}$CH (2.29 Å) and to the *i+3 NH* (2.80 Å) as well as the low temperature coefficient of the *i+3 NH* (2.75 ppb/K) point to a type II β-turn conformation in this molecule. The constrained dynamics in both vacuum and DMSO exhibited a strong intramolecular hydrogen bond,

the characteristic 10-member ring type II β-turn closure, between the Asp *NH* and the carbonyl of Lys. The Asp *NH* also showed a less populated hydrogen bond with the carbonyl of Pro. The $^{\varepsilon}NH$ of Lys was inside the ring during the entire simulation and showed some ability to H-bond to the Lys carbonyl, consistent with its temperature coefficient (3.33 ppb/K). On the average, the backbone dihedral angles and the H bond patterns from the simulations performed on this structure correspond to a type II β-turn.

Cyclo^{1,4}[Ac-Orn-Pro-Gly-Glu-NH$_2$] (Tetra$_{31}$) : The relatively large (3.30 Å) ROESY distance observed between the *NH*'s of Gly and Asp excludes the possibility of a type II β-turn with Asp at the *i+3* position. Together with the low temperature coefficient exhibited by the Asp *NH* (2.65 ppb/K) this distance suggests the presence of a γ-turn like ring around Gly. This conformational feature is observed in the dynamics simulations concurrently with a transient γ'-turn centered around Pro. The hydrogen bond between the Gly *NH* and the Orn carbonyl associated with the γ'-like ring

repeatedly forms and breaks during the simulations which is consistent with the measured temperature coefficient of the Gly *NH* (3.45 ppb/K). A 10-membered H-bonded ring involving the Gly *NH*(*i+3*) and the Orn CO(*i*) is never observed in the simulations of this tetrapeptide. The average dihedral angles found in the simulations (Table 22) are consistent with two C_7 rings around Pro and Gly (Figure 19c).

Cyclo^{1,4}[Ac-Dab-Pro-Gly-Asp-NH_2] (Tetra$_{21}$) : The simulations show that the *i+3* residue *NH* (Asp *NH*) forms H-bonds to 3 different CO's belonging to the *i*(Dab), *i+1* (Pro) and *i+3* (Asp) residues. This extensive H-bonding is consistent with the very low temperature coefficient determined for the Asp *NH* (-1.40 ppb/K). Although the average values of the backbone dihedral angles of the *i+1* and *i+2* residues are close to a type II β-turn conformation (Table 22), the 10-membered ring was the least populated in the MD simulation. The 7-membered ring involving the Pro CO appeared to be most sTable. During the simulation a transient γ'-turn like 7-membered ring closed by an H bond between the *NH* of Gly and the

CO of Dab was also observed. Given the relatively high temperature coefficient of this *NH* (4.04 ppb/K) it is likely that this γ'-turn is not highly populated.

Cyclo^{1,4}[Ac-Dpr-Pro-Gly-Asp-NH$_2$] (Tetra$_{11}$) : The MD simulations of this molecule exhibited a highly populated bifurcated hydrogen bond from the Asp *NH* to the α carbonyl of Dpr and to the β carbonyl of Asp. This is consistent with the very low temperature coefficient observed for this proton (-0.90 ppb/K). The Gly *NH* and the side chain *NH* both show temperature coefficients around 3 ppb/K. However, we note that all temperature coefficients for Tetra$_{11}$ are relatively low (Table 21). This molecule is the most constrained of the series and consequently its atoms experience less motion and perhaps less overall interaction with the solvent. The constrained dynamics indicated that the Gly *NH* and the lactam *NH* had a very low tendency to form an intramolecular H-bond. The absence of such H-bonding for the lactam *NH* is supported by 6 experimental ROESY-derived distance constraints for this proton which result in its orientation to the outside of the cyclic

portion of the molecule. The ROESY interproton distances from the Gly *NH* to the Pro αCH and to the Asp *NH*, 2.48 Å and 2.84 Å, respectively, and the 60 ps MD averaged dihedral angles (Table 22) support a type II β-turn conformation for this tetrapeptide.

Conformational correlation. Previous investigations have attempted to correlate the effects of *i* to *i +3* and *i* to *i + 4* lactamization with helix formation [101-104]. These studies concluded that both ring size and the orientation of the side chains involved in lactam formation influenced the stability of the helical segment under investigation. In particular an *i* to *i +3* lactam bridge with an 18-membered covalent ring was found to be helix destabilizing [105].

When trying to correlate the backbone conformation to the ring size in this series of closely related peptides, one is struck by the rich conformational diversity offered by what would be expected to be similar molecules. In the sixteen-membered cyclic peptides, $Tetra_{22}$ appears to form a sTable type II β-turn while $Tetra_{31}$ assumes conformations involving γ-turn

and γ'-turn-like C_7 rings. Similarly, in seventeen membered cyclic peptides $Tetra_{41}$ seems to adopt a type II β-turn while the characteristic feature exhibited by $Tetra_{32}$ is that of a γ-turn like C_7 ring. It was also found that cyclic peptides ($Tetra_{21}$ and $Tetra_{12}$) with fifteen-membered lactam rings exhibited entirely different structures. It was somewhat surprising that a β-turn like structure was absent in the relatively flexible 18-membered lactam containing tetrapeptide ($Tetra_{42}$), while it was observed in $Tetra_{41}$. This finding together with the structural conclusions on the tetrapeptides containing smaller rings leads us to conclude that the exact conformation depends on the specific composition of the ring and does not correlate simply with the ring size. This conclusion could be further tested by preparing lactam rings in which the basic and acidic residues have been inverted in the primary sequence.

From the perspective of our analysis of the biologically active structure of α-factor, it is important to note that tridecapeptides containing

lactams corresponding to $Tetra_{42}$, $Tetra_{32}$, $Tetra_{12}$, and $Tetra_{11}$ maintained the same preferred conformations as the model peptides (see next Chapter). Hence, it appears that one can control the architecture of a peptide by using the cyclized tetrapeptides discussed here as conformation fixing building blocks. This finding is important since many biologically active peptides contain β-turns and it is clear that cyclization through lactam bridges could result in the stabilization of this conformation. However, our findings indicate that the choice of the residues used for such lactamization is critical as different side chains can induce quite different structures.

Footnote: **The work discussed in this Chapter was published in reference 48.**

CHAPTER 8

Conformational Analysis of Cyclic Analogs of the *Saccharomyces cerevisiae* α-factor Pheromone

Introduction

As explained in previous Chapters, the sexual conjugation by means of which *S. cerevisiae* goes from its haploid level to its diploid level has been extensively studied because it offers an excellent model of cell - cell signaling. In particular, the size and and the nature of the α-factor pheromone (13 amino acid residues) and of the receptor (a seven-loop transmembrane molecule coupled with with a G protein) make understanding this biological mechanism an exceptionally interesting problem with a high potential for generalization. A typical example of a related problem is, for instance, the gonadotropin releasing hormone, a mammalian peptide secreted by the hypothalamus that triggers the release of follicle-stimulating hormone and luteinizing hormone by the

pituitary gland. Indeed, the similarity of the two molecules is underscored by the fact that the yeast pheromone stimulates (at a low level) mammalian gonadotrophs [106]. The structure-activity relationship in α-factor has been the object of numerous studies [1, 2]. Solution physical chemical investigations suggest that this pheromone molecule has a substantial tendency to fold into a β-turn around the lysine-proline-glycine-glutamine sequence, which constitutes the 7 to 10 region of the its amino acid sequence [3, 4]. Furthermore, the studies on the biological activity of the peptide [41, 107, 108, 109] suggest this region to be extremely important for pheromone-receptor interaction.

The control of the stereochemistry and/or topology of regions of molecules is an important part of drug design and of understanding the interactions between recognition molecules and their cognate ligands. The consideration of the 7 to 10 lactamized tridecapeptides is directly linked to the study described in the previous Chapter. Indeed, the model tetrapeptides considered before are the building blocks around which

the design of the pheromone analogs is centered. The critical proline-glycine moiety was found to exhibit a variety of conformations related to type II β-turns and γ -turns but never type I β-turns in those cyclic tetrapeptides. When incorporating the cyclic tetrapeptide building blocks into the center of tridecamer sequence to obtain constrained α-factor analogs, biochemical investigations showed them to preserve high biological potency despite the fact that they did not bind to Ste2p, the α-factor receptor molecule, with high affinity [7].

Clarifying the connection between overall activity, receptor affinity, and peptide molecular shape is a major goal in the study of any hormone, pheromone, or substance mediating cell-cell signaling. The main purpose of the work described in this Chapter is to shed light on this question in the case of the α-factor. It consists of a detailed conformational analysis of four of the constrained tridecapeptides in which the center of the peptide contained an *i to i + 3* lactam with the sequence (Xxx-Pro-Gly-Glu; Xxx = Lys, Orn, Dab, Dpr)

using high resolution ^{1}H nmr spectroscopy and computational modeling. The results of this study also allowed us to ascertain whether the cyclized tetrapeptides are indeed conformational building blocks, *i.e.*, whether they retain their conformation when inserted into a longer peptide.

Experimental Procedures

Materials

The peptides investigated in the present study are residue 7 to residue 10 side chain to side chain cyclized analogs of the *Saccharomyces cerevisiae* α-factor analog that has the sequence WHWLQLKPGQPNleY. *L*-norleucine (Nle), which is isosteric to *L*-methionine, was incorporated at position 12 to replace Met^{12} of the naturally occuring α-factor. This replacement improves the synthesis and stability of the resulting peptide without affecting the activity or receptor binding as noted by the authors of reference [7] who synthesized these peptides. The different analogs were obtained by replacing the 10^{th} residue with Glu and the 7^{th} residue with Lys, Orn, Dab, and Dpr respectively. The acronyms that we will subsequently use for these peptides were published previously and reflect the number of methylene groups contained in the side chains of the 7^{th} and 10^{th} residues. For example, the $Lys^{7}Glu^{10}$ cyclized analog was named **C42**, the $Orn^{7}Glu^{10}$ cyclized analog was named

C32, and so on (Fig 3b). The syntheses of these peptides and their biological properties were reported previously [7].

NMR Procedures

The techniques used in the characterization of the cyclic tridecamer analogs of the α-factor were the same as the ones described for the cyclized tetrapeptides, in the previous Chapter, with the following exceptions: a series of TOCSY [110] and one dimensional ^{1}H-NMR spectra for determining temperature coefficients of the cyclic tridecapeptides were obtained at 295-325 K in increments of 5 K. One dimensional NMR experiments alone were not enough to measure the temperature coefficients due to the considerable crowding exhibited by the spectra of 13mer molecules. Sample temperatures were controlled with the variable temperature unit of the instrument that was thoroughly calibrated before each experiment.

The complete proton resonance assignments were made using DQFCOSY [111], TOCSY, ROESY [37, 57] and NOESY [112] experiments. The DQFCOSY experiments were

necessary, in this case, again due to the crowding of the larger molecules' spectra. The NOESY (as opposed to the ROESY) spectra utilized here, both in confirming the assignments and to extract interproton distances were the appropriate versions of Overhauser effect experiments to use with pheromone (approximately 2,000 Daltons) sized molecules. All 2D NMR spectra were acquired at 298 K in phase sensitive mode using the method of States *et al* [85]. The NOESY spectra were collected for 512 t_1-increments with 2K data blocks and 32 transients for each t_1-increment. The spectral width in both dimensions was 7200 Hz, which covered a region approximately 2ppm to 3ppm wider than the one in which all the meaningful resonances appeared. The NOESY spectra were recorded with a mixing time of 250 ms which, according to our analyses was the most appropriate for collecting quantitative conformational data. It may be emphasized that this time scale is significantly larger than the one used for tetrapeptides, which, again has to do with the larger size of the molecules. The NOE intensities were

converted to interproton distances. The spin lock time for the TOCSY spectra was 60ms. The NMR data, acquired on a Varian Unityplus 600 spectrometer, were processed on a SILICON GRAPHICS workstationstation using the SYBYL TRIAD software. Prior to Fourier transformation, time domain data from the 2D experiments were apodized using shifted sine-bell or Gaussian window functions in both dimensions and zero filled to 2K x 2K real points.

Modeling Procedures

The tridecamer α-factor analogs models were built using the BIOPOLYMER module of the SYBYL software on an INDIGO 2 SGI computer. Special templates for the residues used in the cyclizations (Lys, Orn, Dab, Dpr, and Glu) were created using the dictionary options of the SYBYL package. Constrained DIANA searches were used to generate 200 initial conformations for each studied molecule. The constraints were derived from NOESY spectra. The constraints were input as the experimentally determined internuclear distance in Å ± 10%. Pseudo atom corrections were applied to those

methylene protons that were indistinguishable. The generated structures were minimized using the AMBER force field and their deviations from the physical chemical data were evaluated by computing RMSD values of the constrained distances. For each molecule, 10 conformations with the lowest RMSD values were selected and further searched and minimized (using repeated annealing and minimization routines). The lowest energy conformation for each molecule was solvated in a DMSO box [48] and thoroughly minimized. The dynamics simulations were run on both the vacuum and solvated structures in five steps: heating at 1000K (10 ps), cooling to 500K (10 ps), cooling to 300K (10 ps), equilibration at 300K (20 ps), and simulation at 300K (200 ps). The equation of motion integration step was 1 and snapshots were taken every 100 fs. Dihedral angles deduced from coupling constants were not used as constraints due to degeneracy, but rather used for evaluating the model obtained from NMR derived distance constraints.

Conformational Heterogeneity in Cyclic Lactams

Each of these peptides showed a single sharp homogeneous peak in analytical HPLC under two different conditions [7]. One-dimensional proton NMR spectra of 18- and 17-membered ring containing α-factor analogs **C42** and **C32** exhibited one resonance for each amide NH, suggesting either one dominant isomer on the NMR time scale or fast conformational averaging. However, both of the smaller ring peptides **C22** (16-membered) and **C12** (15-membered) were heterogeneous on the NMR time scale and showed the presence of two isomers in 30:1 and 19:1 ratio, respectively. The coexistence of two conformers was established from the exchange cross-peaks (same sign as the diagonal) in their corresponding ROESY spectra (data not shown). The isomers were most likely a result of slow Xxx-Pro peptide bond isomerization as compared to the NMR time scale. The major isomer adopts a *trans* configuration with respect to the Xxx-Pro peptide bond as evidenced from the characteristic ROE cross-peak from the Xxx$C^{\alpha}H$ to the Pro$C^{\delta}H$, while the minor isomer adopts a

cis configuration, as revealed by the XxxC$^{\alpha}$H to the ProC $^{\alpha}$H ROE cross-peak. In the case of the conformationally heterogeneous peptides (**C22** and **C12**) the conformational analysis was carried out only for the major *trans* isomer.

Characterization of peptide pheromones

The complete proton resonance assignment of all the molecules studied in DMSO-d_6 was accomplished by analyzing standard DQF-COSY, TOCSY and NOESY experiments. Peak resolution and typical TOCSY connectivities are shown for **C22** (Fig. 20). All protons for each residue are labeled in the TOCSY spectrum and the connectivities are easily followed for the β-NH proton of Dab^{7} (δ= 7.23). The characteristic downfield resonances of C$^{\beta}$Hs (2.80-3.20 ppm) were used to differentiate the Trp, His and Tyr residues from the non-aromatic residues, while the upfield resonances (C$^{\delta}$Hs and C$^{\varepsilon}$Hs; 0.80-0.90 ppm) were used to identify the Leu and Nle residues. The side chain ring protons of Trp^{1},

His^2, and Trp^3, which form independent spin systems, were identified by inter- and intra-residue NOE connectivities. Identical spin systems such as Trp(1 and 3), Leu(4 and 6) and Pro(8 and 11) were distinguished by means of sequential NOE connectivities as illustrated for **C32** (Fig. 21). The composition of the peptides was confirmed by the complete NMR resonance assignments. Chemical shift values of the cyclic tridecapeptides are summarized in Table 25, which presents complete assignments for **C22**, and in Table 26 which presents the assignments for the four residues involved in the ring cyclized region of **C42**, **C32**, and **C12**.

Temperature coefficients of amide NHs

The temperature coefficients for amide protons were calculated by a least-squares analysis of the temperature dependence of their resonances (Table 27). When considering the *i* to *i+3* cyclized residues (amino acids 7 to 10), the αNH of the *i*-th residue is outside

the cyclic system and is expected to be solvent exposed. The temperature coefficient exhibited by residue Xxx^7 was between 6.0 and 6.7 ppb/K and was consistently the highest of the three NH's involved in the constrained region of the cyclized tridecapeptides. It was also consistently comparable to the temperature coefficients of the NH's from nonconstrained residues of these α-factor analogs, suggesting that this NH experienced a solvent exposure similar to the amide NH's that were not in the cyclic region in each of these peptides. The α_{NH} of the *i+3* residue exhibited very small temperature coefficients (0.06 to 1.78 ppb/K) in the **C32**, **C22**, and **C12** peptides, indicating its probable involvement in intramolecular H-bonding. As in the study of the cyclic tetrapeptides, the NH coefficient of the *i+3* residue in **C42** (6.3 ppb/K) is an exception and suggests solvent exposure of this proton. In addition, the α_{NH} proton of the 10^{th} residue in **C32**, **C22**, and **C12** resonated upfield (6.47 - 7.50 ppm) compared to all other amide NHs in each of these peptides, further indicating that in these molecules this α_{NH} was not solvent exposed [86, 87]

Moderately low temperature coefficients, 3.1 ppb/K and 3.4 ppb/K were observed for the glycine amide proton in **C42** and **C12**, respectively, implying that in these two pheromone analogs the $^{\alpha}$NH of the *i+2* residue is partially shielded from the DMSO solvent and/or involved in some intra- or intermolecular hydrogen bond.

Interproton distances calculated from NOE intensities and from molecular modeling.

NOESY spectra were recorded with mixing times of 150, 250, 300 and 350 ms. The intensities of corresponding peaks from both sides of the diagonal were averaged to obtain NOE buildup curves. Based on the isotropic motion assumption and the isolated two spin approximation (ISPA) [34], the observed average integrals from 250 ms NOESY spectra (which fell within the linear range of the NOE build up curve) were converted to interproton distances. In cases where both methylene protons resonated at the same chemical shift the observed NOE cross-peak volume (I) has been

corrected using the formula $I_{cor} = 2/3\ I_{obs}$. The distance between the Gly methylenic protons (1.77 Å) was used as an internal standard for calculation of all other distances. Of utmost importance in the understanding of the molecular geometry are the αCH-NH(*i+1, i+2*) and NH-NH(*i+2; i+3*) interproton distances because they differentiate between distinct types of β- and γ-turns. The former distance takes a value of approximately 2.2-2.3Å in type II β-turn and γ-turn conformations and a value of approximately 3.4Å for a type I β-turn, while the latter distance is around 2.4Å in type I and II β-turns and around 3.7Å in γ-turns. **C42** is the only member of this series of cyclized 13-mer pheromone analogs which does not exhibit the NH-NH(*i+2; i+3*) NOE cross peak, indicating an almost extended conformation for the Gly residue (Table 28). NOESY spectra of all remaining peptides exhibited this cross peak and the corresponding inter atomic distance was found to be between 2.58 and 2.78 Å. This is consistent with a bent conformation around the glycyl residue in each of these molecules. The $^{\alpha}CH-^{\alpha}NH(i+1; i+2)$

interproton distance extracted from NOE crosspeak volumes was between 1.97 Å and 2.40 Å for all constrained tridecapeptides indicating the presence of type II β-turns and/or γ-turns and excluding the possibility of a type- I β-turn. The average interproton distances obtained from the analysis of molecular dynamics simulations in vacuum (data not shown) and DMSO of these cyclic peptide pheromones were in very good agreement with their corresponding NOE derived distances (Table 29, and supplementary material, Tables 30 - 32). The *rms* deviations for each average distance in the ensemble of two thousand structures are also presented in these Tables.

$^{3}J_{NH-C^{\alpha}H}$ coupling constants and backbone dihedral angles derived from coupling constants and modeling

The $^{3}J_{NH-C^{\alpha}H}$ coupling constants were extracted from well digitized DQFCOSY and one dimensional ^{1}H NMR spectra. Bystrov's Karplus equations [88, 89] were used to calculate the dihedral angles (ϕ) from these coupling

constants. The observed coupling constants displayed a well defined pattern, with the residues from the noncyclic region exhibiting $^{3}J_{NH-C\alpha H}$ values between 6.7 Hz to ~8.0 Hz (there was only one value of 8.3 Hz, all the others being below 7.9), while the cyclic region residues exhibited values as low as 5.7 Hz and as high as 12.9 Hz. The $^{3}J_{NH-C\alpha H}$ values for the constrained region amino acids as well as the dihedral angles calculated from Bystrov's Karplus equations[19] together with average dihedral angles computed from 200 ps MD trajectories in DMSO are given in the supplementary material, Table 33. The dynamics average dihedral angles are generally within +/- 25° of at least one of the values derived from the coupling constants.

Ramachandran Plots

The dihedral angles of the Pro-Gly region (ϕ_{i+1}, ψ_{i+1}, ϕ_{i+2} and ψ_{i+2}) for the 2000 discrete steps of the 200 ps molecular dynamics simulations in DMSO are shown in Ramachandran maps for the four molecules studied (Fig. 22). Table 28 presents the dynamics averages of

these dihedral angles and the averages of the two important interproton distances from the Gly^9 αNH to the Pro^8 αCH and the Glu^{10} αNH. The average values exhibited by ϕ_{i+1} are confined to the range between -47° and -80°, due to the presence of the pyrrolidine ring of the proline residue at position *i+1* which constrains this angle to values around -60°. The other three backbone torsional angles which determine the shape of the central lactam region exhibited considerable heterogeneity implying substantial conformational divergence in this otherwise closely related series of cyclic peptides. The dihedral angle ψ_{i+1} average was found to vary between 76° in **C42** and 147° in **C32**, the ϕ_{i+2} average varied between 63° in **C12** and 165° in **C42**, and the ψ_{i+2} average varied between -164° in **C42** to +25° in **C22**.

The ψ_{i+1} dihedral angle average of **C42** is 44° lower than the standard type II β-turn value (120°). The Ramachandran plot as well as the dynamics simulation averages of the torsional angles of Pro^8 (ϕ_{i+1}, ψ_{i+1}) in **C42** correspond to an inverse γ-turn around this residue

(Fig. 22a, Table 28). The Gly^9 residue of **C42** adopts an almost extended conformation (Table 28). In the **C32** dynamics simulation, the Gly^9 dihedral angles are consistent with the presence of a γ-turn around this residue (Fig. 22b; Table 28). The average backbone torsional angles in **C22** reflect a reasonably good type II β-turn (Fig. 22c, Table 28) as do the angles of **C12**. However, significant spread is observed in the Ramachandran maps of the molecular dynamics simulations (Fig. 22d) indicating the coexistence of several additional turn structures (see Discussion). It is very interesting that the smallest lactam ring exhibits higher conformational diversity than some of the larger homologs.

The conformational features exhibited by the residue 7 to residue 10 regions of the cyclized tridecapeptides were very similar to those exhibited by the corresponding model cyclic tetrapeptides [48]. The only noTable difference was the wider spread of the glycine torsional angles in the Ramachandran map of **C22** to accommodate a transient γ-turn, as opposed to the

corresponding tetrapeptide which seemed to exhibit a very sTable type II β-turn.

DISCUSSION

The region encompassed by residues 7 and 10 of α-factor has been shown to play a very important role in the biological activity of this pheromone [90-95]. In the native pheromone this region is centered on a Pro-Gly dipeptide unit which induces a transient type II β-turn conformation [41.] Although side chain to side chain cyclization is utilized in many laboratories with the purpose of eliminating the conformational diversity of linear peptides, our previous studies on side chain lactamized tetramers that model this critical region of the α-factor revealed a quite heterogeneous family of structures [48]. Depending on the methylene composition of the side chains used for lactamization, the cyclized tetrapeptides exhibited various combinations of conformational features including the type II β-turn, the γ-turn and the γ'-turn. In order to determine whether these conformational preferences were modified

when the same constraint was placed in a longer peptide, we used ^{1}H-NMR to study a family of analogs of the tridecapeptide α-factor that incorporated the model cyclic tetrapeptides in the critical region discussed above. The conformational behavior of these cyclic pheromone analogs as revealed by NMR and constrained dynamics simulations is discussed below for each of these peptides.

C42: The large temperature coefficient (6.30 ppb/K) of the *i+3* NH, in conjunction with the absence of a NOESY cross peak between the *i+2* (Gly) and *i+3* (Glu) NH's (Table 28,) suggest that the *i+3* residue amide proton points outside the lactam ring and consequently that the *i+2* residue of this tridecapeptide is not part of a type II β-turn. The simulation averages of the backbone ϕ, ψ dihedral angles of the *i+2* residue (165° and -164°) substantiate this conclusion. In this molecule the Gly NH displayed a temperature coefficient (3.10 ppb/K,) which was the lowest among all the amide NH's in **C42** and among all the glycine NH coefficients in this series of cylic peptides. This implies that this atom is

moderately solvent shielded and that this NH is likely oriented towards the inside of the lactam ring and/or involved in hydrogen bonding. A 7-membered ring centered around Pro and closed by a hydrogen bond between the Gly NH and Lys CO is indeed observed in the molecular dynamics simulation performed on this molecule in both vacuum and DMSO (Figure 23a). The major conformational feature of **C42** corresponds to a γ'-turn around Pro^{8} and a relatively extended Gly^{9} residue which is consistent with the average angles found in constrained molecular dynamics simulations in both vacuum and DMSO.

C32: The *i+3* residue NH exhibited the lowest temperature coefficient (0.06 ppb/K) of all the amide protons in this series of molecules suggesting its involvement in a strong intramolecular hydrogen bond. However, in this analog a relatively large NOE distance (2.78 Å) between the NH's at the *i+2* and *i+3* residues was found which appears to contradict the presence of a type II β-turn. These facts were well reproduced by the dynamics simulations which revealed a sTable H-bond

between the Glu NH and the CO of Pro closing a 7-membered ring around Gly in a γ-turn like conformation (Fig. 23b). Indeed, the average backbone torsional angles calculated for Gly, 89° and -51° in the vacuum simulation and 87° and -63° in the DMSO simulation, correspond to a C_7 γ-turn conformation.

C22: The Pro $^{\alpha}$CH to Gly NH and Gly NH to Glu NH internuclear distances derived from the NOE experiments were found to be 2.07 and 2.58 Å, respectively, for this molecule. In combination with these data, the low temperature coefficient of the Glu residue NH (1.78 ppb/K) suggests that the backbone of this molecule adopts a type II β-turn like conformation. Besides a quite sTable H bond between the Glu NH and the Dab carbonyl, the dynamics simulations also exhibited a transient γ turn centered around the glycine residue (Fig. 23c). This accounts for the larger scattering of the points observed in the Ramachandran plot of this peptide (Fig. 22c) as opposed to the very narrow range displayed by the corresponding model tetrapeptide.[9]

C12: Similarly to the corresponding model tetrapeptide, in the molecular dynamics simulation of this molecule three conformational features (type II β-turn, γ- and γ'-like C7 turns) coexisted (Fig. 23d). The very low temperature coefficient (0.5 ppb/K) of the Glu NH which is extensively H-bonded in the dynamics simulation, and the moderately low temperature coefficient (3.40 ppb/K) of the Gly which seems to spend a considerable time inside the lactam cyclization ring, both support this picture. The averaging over several conformations is also reflected by the extensive spread of the points in the Ramachandran plot (Fig. 22d).

Biological and biochemical ramifications of the effects of the tetrapeptide building blocks*

It is relevant to discuss the above conformational conclusions in light of the biological activities of the **CXX** α-factor analogs. A growth arrest assay showed that all of these analogs except for **C32** had virtually the same activity which was about ten-fold lower than that

of the linear tridecapeptide. The equivalency of the activities for **C42, C22** and **C12** coupled with the different conformational preferences of the cyclic central region of these constrained molecules makes it rather unlikely that any one cyclic structure is identical to the biologically active structure of α-factor. This conclusion is supported by the fact that the cyclic peptides all have much lower receptor affinities than linear α-factor and implies that the receptor selects or induces the bound conformation of the α-factor. However, the reasonably high potency of these covalently constrained pheromones strongly suggests that the linear pheromone does assume a bent structure in its biologically active state. The exact nature of this structure will only be revealed when studies on the receptor-peptide complex can be carried out.

Footnote: **The work discussed in this Chapter was published in reference 49.*The discussion of the biological effects is presented here for completion. The biologically relevant experiments were performed in Dr. J. M. Becker's lab, University of Tennessee.**

motto

"Look around and see, this is my land, these are my people"

CHAPTER 9

Considerations on the Usage of Solid State NMR Dipolar Dephasing Experiments in the Conformational Analysis of Peptides

The REDOR effect [27, 113] is a remarkable phenomenon that has been shown to be particularly useful for the investigation of the conformation of active peptides and proteins [114-117]. Although the interpretation of the 2-spin heteronuclear interaction is quite clear, and very fast algorithms have been developed for the numerical simulation of the data [118], problems arise from the effect of natural abundance dephasing spins [119]. In a typical experiment, a REDOR spectrum is measured for the interaction of two labeled nuclei such as ^{13}C and ^{15}N. The full echo and the REDOR spectrum are subtracted and the difference spectrum is corrected for interactions between the observe nucleus and natural abundance nuclei within 1 or 2 bonds. After proper correction the resultant dephasing between the labeled nuclei can be

related to their internuclear distance. Empirical methods to correct for natural abundance dephasing have been successfully applied to a variety of molecules [114, 115]. However, when the observed nucleus is within approximately the same distance from many similar dephasing nuclei a more careful theoretical treatment may be needed, given that the effect is not linear.

With that in mind, in this work we analyze the expressions resulting from the equations of evolution due to the dipolar interactions of three or more spins and we find that the heteronuclear 3-spin case is a direct generalization of two spins problem. The theory underlying 3-spin REDOR experiments and REDOR experiments on multiple - labeled substances is worked out.

THE TRUNCATED DIPOLAR INTERACTION HAMILTONIAN FOR THREE DIFFERENT SPINS *I*, *S*, AND *X*

We consider the problem of the evolution of a spin *I* that has heteronuclear dipolar couplings with spins *S* and *X* in a constant magnetic field, under MASS, *i.e.*, magic angle spinning, conditions. *I*, *S* and *X* are different spin ½ nuclei, i.e., $\omega_{0I} \neq \omega_{0S} \neq \omega_{0X}$. The problem setting is pictorially described in the Figures 24a and 24b, with the (x_{rotor}, y_{rotor}, z_{rotor}) reference frame (*RF*) parallel to the rotor *RF* and originating at *I* (z_{rotor} is parallel to the MAS rotation axis). In this *RF*

(α, β) are the polar angles of the line connecting *I* and *S* (same convention as in [27, 113]). Because there are three spins involved, we must introduce an extra pair of angles that we will denote φ and γ (Figure 24b) and which give the relative spatial position of the spin *X* with respect to the *S* spin. They are, respectively, the azimuth and the colatitude of the *IX* line in the *RF* (not drawn) in which the *IS* line is the "z" axis. We will note here that while γ is a molecular conformation dependent angle that is found in the final formulas, φ is a molecular orientation dependent angle (just like α and β), that in a powder takes, with equal probability, all possible values (between 0 and 2π or $-\pi$ and π which

is the convention used in this work). Thus, the powder average that gives the formula of the measurable intensity will be, in this case, a triple integral over α, β and φ.

The reference frames relevant to our problem are the laboratory (denoted *lab*) frame, the rotor (denoted *ro*) frame and a frame (denoted *spin*) with *IS* as the "z" axis. The *ro* frame is made parallel to the *lab* frame by a rotation of Euler angles $(-\omega_r t, \theta_m, 0)$ with θ_m the magic angle and ω_r the MAS angular frequency [119, 120] (we use $-\omega_r$ instead of $+\omega_r$ in order to obtain the same relative signs for α and $\omega_r t$ in the dipolar dephasing frequencies as in reference [27], which is just a matter of convention and does not affect the results on either the dephasing phase, or the total signal, since the effect is clearly independent of the sense of the rotor rotation). The *spin* frame is made parallel to the *ro* frame by a rotation of Euler angles $(\pi, \beta, \pi-\alpha)$ which is the inverse of $R(\alpha, \beta, 0)$ [121]. Obviously, there is a degree of arbitrariness in the choice of the *spin* frame that is resolved in the final formulas by the integration over φ. The first purpose of our problem is then to express the truncated dipolar interaction part of the Hamiltonian, governing the lab frame evolution of *I*, in the variables $(\alpha, \beta, \gamma, \varphi)$ and this is done

according to [119, 120] using the properties of spherical tensors.

$$H_D^{(trunc,lab)} = \omega_{(0),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_z S_z + \omega_{(1),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_z X_z + \omega_{(S,X),20}^{(lab)}(t)\sqrt{\frac{2}{3}}S_z X_z \qquad [1]$$

The $\omega_{(0),20}^{(lab)}(t)$ and $\omega_{(1),20}^{(lab)}(t)$ are the (20) spherical components of the dipolar interaction tensors of spin *I* with *S* and of *I* with *X* in the *lab* frame. They are given by

$$\begin{aligned}\omega_{(0),20}^{(lab)}(t) &= \sum_{q=-2}^{+2}\omega_{(0),2q}^{(ro)}(\alpha,\beta)D_{q0}^{(2)}(-\omega_r t,\theta_m,0)\\ \omega_{(1),20}^{(lab)}(t) &= \sum_{p,q=-2}^{+2}\omega_{(1),2p}^{(spin)}(\varphi,\gamma)D_{pq}^{(2)}(\pi,\beta,\pi-\alpha)D_{q0}^{(2)}(-\omega_r t,\theta_m,0)\end{aligned} \qquad [2]$$

with $\omega_{(0),2q}^{(ro)}(\alpha,\beta)$, the standard [119] spherical components of the dipolar tensor of *I* and *S* in the rotor frame and $\omega_{(1),2p}^{(spin)}(\varphi,\gamma)$, the standard spherical components of the dipolar tensor of *I* and *X* in the *spin* frame.

The problem of the evolution of spin *I* governed by the Hamiltonian [27] from an initial state described by a density matrix $Const \times I_X$ (that is obtained for instance, by cross polarization with the proton reservoir) has the analytical solution

$$\rho(t) = Const \times \left\{ I_x \cos\left[\int_0^t d\tau\omega_{(0),20}^{(lab)}(\tau)S_z + \int_0^t d\tau\omega_{(1),20}^{(lab)}(\tau)X_z\right] + I_y \sin[same]\right\} \qquad [3]$$

where the sin function has the same argument as the cos. The derivation of eq. [3] is discussed in Appendix 1.

THE 3-SPIN EXPERIMENT

We consider an experiment with the pulse sequence described in Fig. 25 performed on a substance of which molecules contain the three spins system described in Figures 24a and 24b. The dephasing angular frequency of spin *I* is calculated in Appendix 2. The term originating in $\omega^{(rot)}_{(0),2q}(\alpha,\beta)$ gives the standard [27] REDOR dephasing and the term originating in $\omega^{(spin)}_{(1),2p}(\varphi,\gamma)$ gives the contribution due to the second spin. The choice of the dephasing spin that is called *S* or *X* is entirely arbitrary. A natural choice is to take as *S* the spin that is more strongly coupled to *I*, so that the contribution to the dephasing due to *X* is the smaller one.

If the dephasing pulses on the third and fourth channels (the *S* and *X* spins channels) are given at the middle of the rotation period (as Figure 25 suggests), upon time integration over an integral number of periods, all the terms in the angular dephasing frequency expression exhibiting a $(2\alpha+2\omega_r t)$ time dependence vanish, leaving the following expression for the dipolar dephasing of spin *I* due to the combined effects of spins *S* and *X*:

$$\frac{\Delta\Phi(\alpha,\beta,\varphi,\gamma,\lambda(0),\lambda(1))}{2\sqrt{2}} = \sin\alpha\left\{\lambda(0)\sin(2\beta)\mp\lambda(1)\sin(2\beta)(\cos^2\varphi\sin^2\gamma-\cos^2\gamma)\pm\lambda(1)\cos(2\beta)\cos\varphi\sin(2\gamma)\right\}$$
$$\pm\lambda(1)\cos\alpha\left[\cos\beta\sin\varphi\sin(2\gamma)-\sin\beta\sin(2\varphi)\sin^2\gamma\right] \qquad [4]$$

with the same λ notation as in *[27]*, i.e.,

$$\lambda(0) = N_C D^{(0)} T_r \qquad D^{(0)} = \frac{\mu_0}{4\pi}\frac{\gamma_I\gamma_S\hbar}{2\pi r_0^3} \text{ in Hz}$$
$$\lambda(1) = N_C D^{(1)} T_r \qquad D^{(1)} = \frac{\mu_0}{4\pi}\frac{\gamma_I\gamma_X\hbar}{2\pi r_1^3} \text{ in Hz} \qquad [4']$$

and r_0, r_1 are the distances from *I* to *S* and *X*, respectively. The different (±) signs correspond to the cases when the two S spins are parallel (top sign) or antiparallel (bottom sign, see Appendix 1 for the analytical basis of this formula). The normalized signal is obtained by a trace and powder average (eq. [16] in Appendix 1) which, in this case, becomes a triple integration (over α, β and φ) and an average of the two cases arising from the (anti)parallelism of the *S* and *X* spins which are, to a very good approximation equiprobable.

$$S_f = \frac{1}{2}\times\frac{1}{8\pi^2}\int_{-\pi}^{\pi}d\alpha\int_{-\pi}^{\pi}d\varphi\int_{0}^{\pi}d\beta\sin\beta\left[\cos\left(\Delta\Phi^{parallel}\right)+\cos\left(\Delta\Phi^{anti}\right)\right] \qquad [5]$$

MANY SPINS

The reasoning presented above for the computation of the 3-spin Hamiltonian is immediately generalized to a many-spins case (more than 3 spins) with the observation that for each additional dephasing spin *Y* we will have to consider an extra γ(Y) conformation dependent angle (between the *IS* and *IY* lines) and an extra azimuth φ(Y) that, in principle, must be averaged. However, a number of complications arise from the following facts: 1) the azimuths are not independent (since giving the three angles α, β and φ entirely determines a molecular orientation), 2) 2^{m-1} (with *m*=the number of dephasing spins) terms arise from the different parallel-antiparallel possibilities, 3) all spins must be different, which obviously places stringent limits on *m*. Still, the reasoning presented in Appendix 1 holds, because, as long as no two spins are identical, the evolution of the initial cross polarized state can be solved analytically. For completeness, we present here the formulas for the Hamiltonian corresponding to 3 dephasing spins, *S*, *X* and *Y*:

$$H_D^{(trunc,lab)} = \omega_{(0),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_zS_z + \omega_{(1),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_zX_z + \omega_{(2),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_zY_z \quad [6]$$

$$\omega_{(0),20}^{(lab)}(t) = \sum_{q=-2}^{+2} \omega_{(0),2q}^{(rot)}(\alpha,\beta) D_{q0}^{(2)}(-\omega_r t, \theta_m, 0)$$

$$\omega_{(1),20}^{(lab)}(t) = \sum_{p,q=-2}^{+2} \omega_{(1),2p}^{(spin)}(\varphi_1,\gamma_1) D_{pq}^{(2)}(\pi,\beta,\pi-\alpha) D_{q0}^{(2)}(-\omega_r t, \theta_m, 0)$$

$$\omega_{(2),20}^{(lab)}(t) = \sum_{p,q=-2}^{+2} \omega_{(2),2p}^{(spin)}(\varphi_2,\gamma_2) D_{pq}^{(2)}(\pi,\beta,\pi-\alpha) D_{q0}^{(2)}(-\omega_r t, \theta_m, 0)$$

NUMERICAL SIMULATION OF THE HETERONUCLEAR 3-SPIN EXPERIMENT SIGNAL

The numerical simulation for the $\frac{\Delta S}{S_0} = 1 - S_f$ value amounts to the numerical calculation of the S_f triple integral given above. This is somewhat facilitated by the formula [122]

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\alpha \cos(u\cos\alpha + v\sin\alpha) = J_0\left(\sqrt{u^2+v^2}\right) \qquad [7]$$

with J_0 the standard zeroth order Bessel function, by means of which the integration over α is analytically performed. Thus, using this formula and the discussion in Appendices 1 and 2, the expression for S_f becomes:

$$S_f = \frac{1}{2} \times \frac{1}{2\pi}\int_0^{\pi} d\varphi \int_0^{\pi} d\beta \sin\beta \left[J_0\left(\sqrt{u^{parallel^2} + v^2}\right) + J_0\left(\sqrt{u^{anti^2} + v^2}\right)\right] \qquad [8]$$

with *u* and *v* defined in Appendix 2 and the *parallel* and *anti* choices corresponding to the upper and lower sign choices in eq. [20]. This integral was computed numerically for the cases illustrated in Fig. 26 and 27 using a two dimensional adaptation of the quad3d routine [123].

MANY HETERONUCLEAR SPIN EXPERIMENTS: STRATEGY AND SCOPE

An important observation based on the results of Appendix 1 and in particular of eqs. [3 and 6] above is that one may label e.g., 4 positions in a molecule with nuclei *I*, *S*, *X* and *Y* (for example ^{13}C, ^{15}N, ^{19}F and ^{31}P) and, in principle, in an ideal case, perform twelve possible standard 2-spin experiments and twelve 3-spin experiments of the kind described in Fig. 25 on this same sample. Indeed, the time dependence of all dipolar dephasing frequencies is cyclic with the same pulsation ω_r, thus, if no pulses are shot on the *X* and *Y* channels, one may perform *I*-observe, *S*-dephase standard REDOR, or, *S*-observe, *I*-dephase standard REDOR on the same molecule as if *X* and *Y* were not present since their effects refocus at the end of each period. The same goes for an *I*-observe, *S*, *X*-dephase 3-spin REDOR experiment to which *Y* is not playing any part, as long as no pulses are given on the *Y* channel. Since all four spins play

equivalent roles any of them can be chosen to be *I*, or *S*, or *X*, giving rise hypothetically to the 24 experiments mentioned above from which a considerable amount of information about the conformation of the molecule could be extracted.

HOMONUCLEAR DEPHASING SPINS

A very important case, both for potential 3-spin experiments and for the interpretation of standard 2-spins REDOR data is $\gamma_I \neq \gamma_S = \gamma_X$. As shown previously [118, 119, 124, 125] the high field truncated dipolar Hamiltonian for three spins, two of which are from identical nuclei, is:

$$H_{D,X=S}^{(trunc,lab)} = \omega_{(0),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_z S_z + \omega_{(1),20}^{(lab)}(t)\sqrt{\frac{2}{3}}I_z X_z + \omega_{(S'),20}^{(lab)}(t)\sqrt{\frac{1}{6}}\left[2S_z X_z - \frac{1}{2}(S_+ X_- + S_- X_+)\right] \qquad [9]$$

with the same notations as in eq. [1] for the pulsations. As shown in Appendix 3, there is no general analytical solution of the evolution governed by this operator, however, our heteronuclear calculations could provide a good approximation for the case when the *SX* coupling is much smaller than the *IS* and *IX* couplings. A typical case would be $I=^{19}F$ and $S=X=^{15}N$. Since $\gamma^{15}_N/\gamma^{19}_F = 0.108$, for most angles γ ($\gamma \geq 60^o$), the problem term in the commutator calculated in eq. [21] would be much smaller than the "nice" one hence the validity of our approximation.

For general simulations, one should adapt to this problem the techniques discussed in references [124, 125].

There is one singular situation in which the homonuclear S and X case is exactly soluble. As noted in Appendix 2, for S, I and X collinear ($\gamma=\pi$) and equidistant ($r_0=r_1$, see Fig. 24a, b)

$$\omega_{(0),20}^{(lab)}(t)-\omega_{(1),20}^{(lab)}(t)=0$$

thus, the second term of the commutator [21] vanishes and the analysis in Appendix 1 holds exactly.

DISCUSSION

Our computations support the experimentally based conclusion [115] that when one of the two couplings is much stronger (e.g., $D^{(0)}$ of the order of $8\times D^{(1)}$ or larger) the more strongly coupled of the two dephasing spins (S within our convention) dominates the observable effect, and the errors introduced by approximating the true 3-spin effect with the sum of two 2-spins effects are no more than 10%. However, the effect is not linear and for $D^{(0)} < 8\times D^{(1)}$ substantial errors are introduced by this approximation.

The dependence on γ for the case $D^{(0)} \approx D^{(1)}$ is illustrated in Fig. 27. For $\lambda(0)\approx\lambda(1)\approx1.0$, the maximum possible effect is twice as large as the minimum possible effect. As explained in the appendices, the measurable signal (eq. [5]) has a $\gamma\leftrightarrow\pi-\gamma$ symmetry, which is somewhat unfortunate, because it prevents unequivocal

determination of the angle from a signal measurement. For γ=0 (which is the same as γ=π, given the just mentioned symmetry) eq. [4] becomes:

$$\Delta\Phi(\alpha,\beta,\varphi,=0\gamma,\lambda(0),\lambda(1))=2\sqrt{2}\sin\alpha\sin(2\beta)\{\lambda(0)\pm\lambda(1)\} \qquad [10]$$

Consequently, at equal couplings, λ(0)=λ(1)=λ, which is the only when we are also able to solve analytically the homonuclear case:

$$\Delta\Phi^{parallel}(\alpha,\beta,\gamma=\pi,\lambda(0)=\lambda(0)=\lambda)=2\sqrt{2}\sin\alpha\sin(2\beta)\times 2\lambda$$
$$\Delta\Phi^{anti}(\alpha,\beta,\gamma=\pi,\lambda(0)=\lambda(0)=\lambda)=0 \qquad [11]$$

and the equation for $\frac{\Delta S}{S_0}$ becomes, in this case:

$$\frac{\Delta S}{S_0}=1-S_f=1-0.5-\frac{1}{2}\times\frac{1}{8\pi^2}\int_{-\pi}^{\pi}d\alpha\int_{-\pi}^{\pi}d\varphi\int_{0}^{\pi}d\beta\sin\beta\left[\cos\left(2\sqrt{2}\sin\alpha\sin(2\beta)\times 2\lambda\right)\right]$$

Clearly, there is no φ dependence and remembering the definition of the 2-spins effect,

$$\frac{\Delta S}{S_0}(\gamma=\pi,\lambda(0)=\lambda(1)=\lambda)=0.5\times\frac{\Delta S}{S_0}^{2-SPIN}(2\times\lambda) \qquad [12]$$

where, $\frac{\Delta S^{2\text{-}SPIN}}{S_0}(2\times\lambda)$ is the expression of the 2-spin REDOR as defined in eqs. 6 to 8 in reference *[27]*. Thus, the effect of 2 homonuclear dephasing spins on an observed (dephased) heteronuclear spin located exactly at the middle of the line segment linking them is equal to half the effect of one of them, located a little closer ($\sqrt[3]{2}$ times closer, to be specific) to the I spin (Fig. 27). Although simple, eq. [12] and Fig. 26 together constitute, probably, the clearest way to exhibit the very different nature of the 3-spin interaction *vs* the 2-spin one. Their significance is that, due to the fact that the two dephasing spins may be oriented oppositely, in many instances (i.e., for many geometries and at many values of the N_C parameter) they interfere destructively and they give a total dephasing which is less than either of the separate dephasings that any single one of them would produce.

These results emphasize the importance of data processing for the deeper understanding of REDOR experiments. In a complex case, quite typical for biologically interesting experiments, a labeled ^{15}N may be within 5.5Å from e.g., 20 carbon atoms in the same molecule (this number was picked from the molecular modeling of molecules similar to those studied in [114]), only one of which is artificially labeled. The

probability of having two ^{13}C nuclei in the same molecule affecting the same observed labeled ^{15}N is

$$P_2 = C_{19}^1 \times (0.01108) \times (1-0.01108)^{18} \approx 0.172 \qquad [13]$$

thus, approximately 17.2% of the molecules experience 3-spin instead of 2-spin effects. Clearly the presence of natural abundance ^{13}C may account for more than 10% of the effect, and our discussion makes a better starting point for the calculation of the correction than the linear approximation. Due to the presence of the γ angles these corrections depend on the working molecular model, thus a consistent processing of REDOR experiments should involve iterative back and forth data refinement/exchange between the fitting of the experimental results and structural modeling.

Another outcome of our theoretical considerations is the interpretation of a putative experiment in which the 3-spin effect is pursued in itself. Some compounds that might test our calculations are depicted in Fig. 28. A substance containing the functional group a) should give an N detected curve very similar to the 3-spins curve shown in Fig. 26, since for this group γ is close to 180^o and the two labeled ^{13}C atoms are equidistant from the observe ^{15}N. The groups b) and c) should give similar results upon N detection since

$$\gamma(b) \approx 180^o - \gamma(c) \qquad [14]$$

Footnote: The work discussed in the above sections of this Chapter constitutes the theoretical basis for a manuscript that is currently under review. Also, to a large extent it parallels the experimental design considered in reference 130. Nevertheless, based on the above discussion, it is this thesis author's feeling that the approximations used in that article are less than ideal for the interpretation of a ^{13}C detected, ^{19}F dephased experiment.

APPLICATIONS OF THE *REDOR* EXPERIMENT TO THE STUDY OF THE α-FACTOR PHEROMONE

Four labeled analogs of the α-factor pheromone were synthesized [by Dr. Michael Breslav]:

1) WHWLQLK-[1-^{13}C]PG-[α-^{15}N]QPNleY

2) WHWLQLK-[^{15}N]P-[1-^{13}C]GQPNleY

3) WHWLQLK-[^{15}N]P-[2-^{13}C]GQPNleY

4) WHWLQLKP-[2-^{13}C]G-[α^{15}N]QPNleY

Figure 29 shows the design of the REDOR experiment and Table 34 shows the results [114] compared to the molecular models obtained as described in Chapter 5 (see also Figure 12.)

The experiment seems to support the fact that, in powder, the α-factor molecule adopts a distorted type I β-turn like conformation.

The analysis presented in this Chapter shows the accuracy of the distances determined from this REDOR experiment to be less reliable than initially considered. Nevertheless the computer assisted process of determining how to design labeled molecules on which to carry out meaningful physical experiments bears considerable promise.

Footnote: **The work discussed in this section was published in reference 114.**

CHAPTER 10

Ab Initio Calculations on Glycine. Comparing Experimental and Theoretical Results.

Introduction

The history of the study of glycine, the simplest amino acid, embodies all the features of the considerably insightful gains that can be obtained from systematic computational studies as well as the challenges of correlating the theoretical results with the experimental data collected under various conditions. Particularly, organic compounds and, especially, biologically relevant molecules like amino acids always bring forth controversial issues concerning the pertinence of quantum mechnical computational results, which are performed in vacuum, in gas phase, or in a dielectrically modeled solvent, to the interpretation of physical chemical experiments that are performed in a variety of real solvent environments

(like DMSO, chloroform, methanol, micelle suspensions), and which, in their turn are invariably less than ideal approximations of the actual biological conditions. Nevertheless, as unexpected as it might be, *ab initio* calculations are one of the few available sources of information for the determination of good conformations for the neutral forms of amino acids [31, 40]. This is due to the fact that in polar media and in the solid state, amino acids form zwitterions. Thus, while the charged forms of amino acids are accessible by X-ray spectroscopy, the determination of the conformation of the neutral form is limited to matrix isolation spectroscopy because almost all amino acids decompose before melting [126, 127].

One of the purposes of the study of glycine is to reproduce theoretically the microwave absorption lines observed in vapor experiments. This endeavor resulted in a controversy because it is not the structure predicted by computational optimization to be the most sTable that seems to be responsible for the main

absorption lines that are due to a higher energy conformation [1126, 128].

Also, the conformational predilections of glycine, i.e., the amount of Ramachandran space that it tends to occupy, have considerable importance because this amino acid is present in many activity-determining regions of biologically important molecules. This is clearly evident from many other published works [25, 36, 126, 127, 128] as well as from the previous nine Chapters centered around the conformation of glycine as a part of larger peptides, replacing glycine, and cyclizing around a glycine containing moiety to obtain interesting analogs.

Our purpose in this Chapter is to use high level computational methods (MP2, MP3, and MP4) in order to investigate possible optimized conformations of glycine and both contribute to the elucidation of the microwave spectrum controversy and lend support to the rest of the computations carried out in our work.

METHODS

The basis set used is 6-311+G**. The 6-311+G** is a triple-zeta basis set which uses three Slater orbitals for the description of valence electrons, one expanded in a series of three gamma functions and the other two approximated by one Gaussian each. The other electrons are described by one Slater orbital expanded in a series of six Gaussians. In addition, diffuse functions and d functions are set in non-hydrogen atoms and p functions are set on the hydrogens. The correlation energy effects are taken into consideration using the Moller-Plesset [129] method, also known as MP up to the fourth perturbation correction term.

The main difference between the *ab initio* calculations described in this Chapter and the similar computations carried out on dipeptides is that here, the program that we run takes into account higher order perturbation terms in the Hamiltonian. The highest order Moller-Plesset correction term used in the calculation is the basis for naming each energy level obtained. The four conformations studied were obtained

by inputing different initial conditions in the Gaussian program [44] and by using four different Hartree-Fock approximations: regular self consistent field (SCF), MP2, MP3 and MP4 with the 6-311+G** basis set. The numeric results of the calculations are exhibited in Tables 35, 36, 37, 38, and 39. The numbers 1, 2, 3, and 4 were arbitrarily assigned to name the optimized molecular conformations found. Figure 30 shows the numbering of the atoms within the glycine molecule that is to be followed throughout the Tables of this Chapter.

RESULTS

The 4 conformations analyzed here are exhibited in Figure 31. The analysis of the energy values exhibited in Table 35 shows the conformations 3 and 4 to be the most sTable. The importance of higher order correction terms is apparent from the fact that in the higher order correction energy levels MP3 and MP4, the conformation 4, which is exhibited to be the best within the framework of the regular SCF and MP2 approximations, is replaced as the lowest energy structure by conformation

3, as shown in Tables 35 and 36. Conformation 2 is consistently, irrespective of approximation level, around 2kcal/mol higher than the lowest energy available and conformation 1 is exhibiting consistently very high energy levels, between 7.0 and 8.3 kcal/mol above the best conformation found in each case.

The comparison of our glycine bond lengths (shown in Table 37) to other models of glycine [126-128] as well as to the results of the dipeptide discussed in Chapter 3, shows the models considered here to have a consistently shorter double bond between the carbon and the oxygen of the carbonyl group. This shortening, consistently present in all our models, of approximately 0.02Å may be attribuTable to our more sophisticated basis set that allows for higher order energy corrections and, consequently, for subtle geometrical adjustments. Except for this bond, all the other bond lengths are in good agreement with the ones from other models. The Mulliken population analysis charges show carbonyl atoms to be much less charged in the case of this glycine model (structure #4, Table 40) than in the

case of the glycyl residue of the Pro-Gly dipeptide considered in Chapter 3.

DISCUSSION

By employing different computational methods our study sheds some light both on the question of the sTable conformers of glycine and on the importance of the sophistication of the basis set used.

Our two most sTable conformations (structures 3 and 4 in Tables 35, 36, and 39) are both stabilized by hydrogen bonds. Structure 3 forms a five membered ring closed by a hydrogen bond between the carboxyl hydrogen atom and the amino nitrogen. In structure 4, the 5-atom ring is closed by a hydrogen bond between the amino hydrogen atoms and the oxygen of the carboxyl.

The lower order perturbation calculations (shown in Table 36) indicate structure 4 to be substantially lower (by 1 kcal/mol in the regular HF approximation and by almost 2kcal/mol in MP2) than the structure 3. Thus at this level of the energy calculation approximation, the bifurcated hydrogen bond due to the presence of the two

hydrogen atoms on the nitrogen that can form bonds to the oxygen seems to be responsible for the determining effect. The higher order perturbation calculation seems to be more sensitive to the high charge on the carboxyl hydrogen (Table 40). In this calculation, the single carboxyl hydrogen seems to bond strongly enough to the nitrogen to offset the energy dip due to the two hydrogen atoms in the other confromation. However, the difference in energy between the two conformations is low, 0.32 kcal/mol at MP3 level and even lower, 0.14 kcal/mol at MP4 level. Thus, the two conformations seem to be indicated by higher order perturbation calculations to be very close which suggests that in experimental setups they may both be equally populated.

This may be the reason why actual experiments have found different conformations for glycine.

CONCLUSIONS

This work attempts to rationalize, on a quantum mechanical basis, the behavior of some peptide molecules containing the proline moiety and to organize a thorough conformational study of four constrained analogs of the *Saccharomyces cerevisiae* α-factor pheromone (Chapter 8) as well as of eight model tetrapeptides (Chapter 7.) Besides being interesting organic compounds and peptide moieties by themselves, all these molecules were considered for the purpose of understanding the residue 7 to residue 10 region of the α-factor that seems to play a key role in the recognition of the pheromone by the STE2 receptor of the MATa yeast haploid cells.

The *ab initio* study can only be carried out for small molecules, but its results are meaningful for the understanding of the physical chemical behavior of the larger peptides. Since the lowest energy structure for the Ac-*L*-Pro-*D*-Ala-NH_2 is a type II β-turn, the study of the blocked dipeptides supports in an unambiguous way the known [8] tendency of Pro-*D*-Ala moieties in peptides and proteins to adopt type II β-turn conformations. The study provides further computational support for both the observed type II β-turn conformation in linear peptides containing the *L*-Pro-*D*-Ala moiety and for the

conformational flexibility observed in linear peptides containing the *L*-Pro-*L*-Ala moiety. The quantitative computational evaluation of the energetic barrier preventing the formation of a type II β-turn, introduced by the presence of an *L* amino acid at position *i+3*, is consistent with the one calculated in the nonblocked dipeptides (around 13kcal/mol.) These quantitative results show unequivocally that the presence of a *D*-Ala amino acid at the position *i+3* in a turn region strongly favors a type II β-turn like conformation. Correlated with the high biological activity of the ^{9}*D*-Ala analog of the α-factor, this result lends convincing theoretical support to the hypothesis that a type II β-turn conformation is important for pheromone recognition. The comparison with X-ray crystallography, when considered in parallel with the nonblocked dipeptide study, shows that the computational effects of blocking the carboxy terminus are in good agreement with the experimental data.

The *ab initio* calculations seem to provide a simple quantitative explanation of the fact that the physical chemistry of the ^{9}*L*-Ala analog is different even at a qualitative level from that of the ^{9}*D*-Ala analog: the ^{9}*L*-Ala analog is probably more flexible because the

energetic spectrum of the *L*-Pro-*D*-Ala dipeptide has many levels within a low energy gap above the ground state.

The detailed conformational analysis of the major *trans* isomers of eight *i* to *i+3* side chain lactamized cyclic tetrapeptides was carried out in order to investigate the effect of a Pro-Gly dipeptide unit at the *i+1* and *i+2* positions of a peptide ring spanning four amino acids. The systematic variation in the size of the constrained region, containing from 14 to 18 member atoms, showed an unexpectedly rich conformational diversity of these peptides. In contrast to what might have beeen expected for Pro-Gly centered peptides, several of these compunds exhibited preferences for γ-turn or γ'-turn like conformations. Other peptides assumed a near perfect type II β-turn and still others were relatively flexible and underwent transitions between structures which had close AMBER force field computed energies. The Ramachandran maps (Fig. 17) clearly indicated that for several structures the range of dihedral angles was highly limited while for others a greater range was allowed, signifying substantial

flexibility. Based on this work, it seems that these plots incorporate all the conformations available to the Pro-Gly region corresponding to the low energy (and thus reasonably accepTable) structures of these molecules. In all the peptides that were studied, the α_{NH} of the Gly is experimentally determined to be close to the α_{CH} of Pro. Consequently, the ψ dihedral angle of the prolyl residue is always positive and all the turn conformations are far from a type I β-like shape.

The fact that the use of different, but seemingly similar, amino acids such as ornithine and lysine, or aspartic acid and glutamic acid (viewed by many analysts as almost identical from the point of view of their conformational effect) resulted in very distinct molecular shapes conveys a serious cautionary warning: careful planning should be undertaken while using such lactam constraints in the design of biologically active molecules with reverse turns and especially, extreme attention should be given not to draw precipitous conclusions or make abrupt assumptions about the shape of the "constrained" molecule. An extra methylene group may make a huge difference both in the shape and in the general flexibility of the structure. However, it is the conclusion of this work that it appears to be

reasonable to try to control the turn topology of a Pro-Gly peptide backbone by side chain lactamization. It is nevertheless critical to use appropriate buiding blocks and realize that simply having the same character side chains (*e.g.*, amino or carboxyl) does not necessarily insure that different amino acids will impose identical, or even similar conformational constraints.

The NMR characterization and conformational analysis of four side chain lactamized analogs of the yeast α-factor which differ from each other only in the composition of the non backbone segment of the cyclized region supports the conclusion of the tetrapeptide study that these rings exhibit considerable conformational diversity. The major structures assumed by these lactams involve type II β-turns and γ-turns. It appears that the cyclic tetramer building blocks (Xxx-Pro-Gly-Glu) preserve their conformational characteristics when incorporated into medium size peptides and, therefore, that they can be used as tools to affect the architecture of such molecules. This is an important result from two different points of view. Firstly, it shows that constrained peptides regions can be analyzed

separately to reach inferences that can be generalized with good accuracy to the larger parent molecules. Secondly, it implies that given constrained building blocks can be consistently used to design larger molecules with desired shapes.

The nuclear magnetic resonance spectroscopy Overhauser effect based molecular modeling study of the model tetrapeptides carried out in comparison with the similar study of some of the cyclized tridecamers also gives credence and rationalization to the biological results that show small differences of the cyclized molecules to have dramatic effects on their bioactivities [7]. Indeed, although apparently similar the different 7 to 10 lactamized peptides were found to have substantially different shapes, and thus, their different activities are not that surprising.

Even in the absence of physical chemical data, a systematic modeling effort may provide good working models, ideas, as well as quantitative data for experimental design and interpretation. The best support for this statement is provided by the $Tetra_{22}$ molecule that has high propensity for forming an almost perfect type II β-turn. This predilection, together with the low likelihood of this molecule deviating far from the favored turn during its dynamic evolution had

been inferred long before any ROESY data supporting its conformational analysis had been collected. However, caution should always be exercised in analyzing the results of stand alone molecular modeling because, as emphasized in this work, the energy differences calculated within a classical framework are far from being as trustworthy as those obtained within a quantum mechanical framework. The structures obtained from stand alone modeling should be viewed as working models which should be supported with further physical chemical data.

REDOR studies can prove very useful in future understanding of biologically active conformations. Nevertheless, good interpretation of the data may require thorough understanding of all the factors involved and very careful application of the software tools. Such factors as the background natural abundance spins may produce significant effects that cannot simply be eliminated by subtraction since the technique measures variations of amplitude which do not depend linearly on the separate spins.

The main molecule considered in these studies, the *Saccharomyces cerevisiae* α-factor, seems to have a high propensity to adopt a turned conformation. All the studies discussed in this work as well as in previous experimental (solid state NMR, high resolution solution

NMR, VCD, or fluorescence transfer measurements) or theoretical (*ab initio* calculations, stand alone modeling, or physical chemistry based modeling) work seem to concur in this respect.

It seems that a type II β-turn conformation centered around Pro^8 and Gly^9 is likely to be a good candidate for the biologically active conformation of this pheromone.

APPENDIX 1

THE VALIDITY OF AVERAGE HAMILTONIAN THEORY IN THE MANY SPIN HETERONUCLEAR CASE

As far as the truncated dipolar Hamiltonian approximation is valid (which is the assumption under which most of the theory of solid state NMR [119, 120] and, in particular, that of the REDOR experiment is written) there is no difference between the 2-spin case *[27]* and the many-spin cases discussed here because as long as all the spins are different, the many-pin truncated dipolar Hamiltonian has fundamentally the same commutation rules with the state prepared by cross polarization of the observed nuclei as the 2-spin one. To be specific, we consider the problem described by the Hamiltonian in eq. [1]. The crucial point in the generalization is that an initial state described by the density matrix:

$$\rho_0 = Const \times I_x \qquad [15]$$

with *Const* a normalization constant (or any time independent function of S_Z and X_Z, for that matter) which is the form of the density matrix describing the

observe nucleus after the cross polarization with the proton "reservoir", evolves under the action of [1] to give, (using the Liouville operator notation [117]):

$$|\rho(t)) = T\exp\left[-i\int_0^t d\tau \hat{H}_D^{(trunc,lab)}\right]|\rho(0))$$

The expression on the right hand side of this equation can be calculated analytically to yield (back to regular operator notation) eq. [3] in the text and the signal is obtained from:

$$S_f = \left\langle \frac{Tr[I_x\rho(t)]}{Tr[I_x\rho(0)]} \right\rangle_{POWDER} \qquad [16]$$

with the angular brackets denoting the powder average. Upon taking the trace, the I_Y term vanishes and so do all the terms containing odd powers of S_Z and X_Z i.e., the terms assuming the form (k, l integers)

$$f(t)S_z^{2k+1}X_z^{2l+1}$$

in the power expansion of the cosine. Clearly, keeping only the even power terms is the same as taking the arithmetic average between the cosine of the difference and the cosine of the sum which is the mathematical basis for the definitions of the dephasing (eq. [4]) and signal (eq. [5]) given there with the intuitive physical

argument based on the relative orientation of the *I* and *X* spins.

APPENDIX 2
CALCULATION AND PROPERTIES OF THE DIPOLAR DEPHASING ANGULAR FREQUENCY AND DIPOLAR DEPHASING IN THE HETERONUCLEAR 3-SPIN CASE

From eq. [1] and the considerations above, we obtain the angular dipolar dephasing frequency of a spin *I*:

$$\omega_D(t) = \frac{1}{2} \times \sqrt{\frac{2}{3}} \left[\omega^{(lab)}_{(0)20}(t) \pm \omega^{(lab)}_{(1)20}(t) \right]$$

with the $\frac{1}{2}$ factor coming from the dephasing spins (mathematically speaking from the fact that each "z" component is half a Pauli matrix) and the ± sign taking into account the *S-X* parallelism effect discussed before. We need not be concerned by the overall sign of the expression, because in the formula for the experimentally observable signal it shows up under a cosine function (eq. [4]) which is even. To be perfectly rigorous, another ± sign should be placed in front of the whole expression *[27]*. Using in this formula the eq. [2], the expressions for the spherical components of the

dipolar tensor (reference [119] pages 11 and 292) and of the rotation matrices [121], we obtain:

$$\begin{aligned}\frac{\omega_D(\alpha,\beta,\varphi,\gamma,t)}{2\pi} = &-\frac{1}{2}D^{(0)}\left[\sin^2\beta\cos(2\alpha+2\omega t)+\sqrt{2}\sin(2\beta)\cos(\alpha+\omega t)\right]\pm\\ &\pm\frac{1}{2\sqrt{2}}D^{(1)}[\cos(2\varphi+\alpha+\omega_r t)\sin^2\gamma\sin\beta(1+\cos\beta)-\\ &-\cos(\varphi+\alpha+\omega_r t)\sin(2\gamma)(2\cos\beta-1)(1+\cos\beta)+\\ &+\cos(\alpha+\omega_r t)(1-3\cos^2\gamma)\sin(2\beta)+\\ &+\cos(-\varphi+\alpha+\omega_r t)\sin(2\gamma)(2\cos\beta+1)(1-\cos\beta)+\\ &+\cos(-2\varphi+\alpha+\omega_r t)\sin^2\gamma\sin\beta(\cos\beta-1)]\pm\\ &\pm\frac{1}{2}D^{(1)}[\cos(2\varphi+2\alpha+2\omega_r t)\sin^2\gamma\cos^4\frac{\beta}{2}-\\ &-\cos(\varphi+2\alpha+2\omega_r t)\sin\gamma\cos\gamma\sin\beta(1+\cos\beta)+\\ &+\frac{1}{2}\cos(2\alpha+2\omega_r t)\sin^2\beta(1-3\cos^2\gamma)+\\ &+\cos(-\varphi+2\alpha+2\omega_r t)\sin\gamma\cos\gamma\sin\beta(\cos\beta-1)+\\ &+\cos(-2\varphi+2\alpha+2\omega_r t)\sin^2\gamma\sin^4\frac{\beta}{2}] \qquad [17]\end{aligned}$$

The 2π denominator on the left side of the equation insures that we use consistently the definitions of the dipolar constants in Hz as given in eq. [3'].
The term containing $D^{(0)}$ is the standard REDOR angular dephasing frequency *[27]*. The relative sign of the two terms in the square bracket of this standard term is different here. This is a matter of how the β angle is defined. Obviously, changing $\beta\leftrightarrow\pi-\beta$ changes the relative sign between the two terms and since a β integration is

to be performed as a part of the powder averaging, this sign bears no significance on the final result.

An important property of the dipolar pulsations that can be seen on the above equation is that for $\gamma=\pi$ (i.e., *I*, *S* and *X* collinear) and for $D^{(1)}=D^{(0)}$, which is, for instance, the case of *S* and *X* homonuclear and equidistant from *I*,

$$\omega^{(lab)}_{(0)20}(t) = \omega^{(lab)}_{(1)20}(t) \qquad [18]$$

thus, in this exceptional case, *S* and *X* contribute identically to the effect and give a zero total antiparallel dephasing.

The 3-spin experiment dephasing over N_C rotational periods is given by *[27]*:

$$\Delta\Phi(\alpha,\beta,\varphi,\gamma,\lambda(0),\lambda(1)) = N_C\left[\int_0^{t_P} dt\omega_D(t) - \int_{t_P}^{T_r} dt\omega_D(t)\right]$$

with t_P the time delay within the rotor period after which the dephasing pulse is given *[27]*. If $t_P = \frac{T_r}{2}$, the last five terms in eq. [17] above (oscillating with $2\times\omega_r t$) vanish upon this integration, and, the expression of $\Delta\Phi$ becomes:

$$\begin{aligned}\Delta\Phi(\alpha,\beta,\varphi;\gamma,\lambda(0),\lambda(1)) = 2\sqrt{2}\lambda(0)\sin\alpha\sin(2\beta)\mp \\ \mp\sqrt{2}\lambda(1)[\sin(2\varphi+\alpha)\sin^2\gamma\sin\beta(1+\cos\beta)- \\ -\sin(\varphi+\alpha)\sin(2\gamma)(2\cos\beta-1)(1+\cos\beta)+ \\ +\sin\alpha(1-3\cos^2\gamma)\sin(2\beta)+ \\ +\sin(-\varphi+\alpha)\sin(2\gamma)(2\cos\beta+1)(1-\cos\beta)+ \\ +\sin(-2\varphi+\alpha)\sin^2\gamma\sin\beta(\cos\beta-1)]\end{aligned} \qquad [19]$$

from which eq. [4] in the text is deduced by means of simple properties of trigonometric functions. We can note that: 1) the $\lambda(0)$ containing term is the regular *[27]* dephasing which is studied at length in [118], 2) one reassuring quick check is that for $\gamma=0$, i.e., for *S* and *X* on the same line, the term in $\lambda(1)$ assumes a form identical to the one in $\lambda(0)$ which is true for the terms in $D^{(1)}$ and $D^{(0)}$ in eq. [17] too.

As mentioned, the dephasing depends on 4 polar angles, 2 azimuths α, φ and two colatitudes β, γ, the first three of these angles being orientation dependent, but the last one being conformation dependent. Some useful properties of the signal function S_f defined in eq. [4] are deduced from
eqs. [4 and 5] as follows. From eq. [4] one realizes that the *u* and *v* parameters in eq. [5] are in fact:

$$u(\beta,\varphi,\gamma) = \lambda(0)\sin(2\beta) \mp \lambda(1)\sin(2\beta)\left(\cos^2\varphi\sin^2\gamma - \cos^2\gamma\right) \pm \lambda(1)\cos(2\beta)\cos\varphi\sin(2\gamma)$$

[20]

$$v(\beta,\varphi,\gamma) = \pm\lambda(1)\left[\cos\beta\sin\varphi\sin(2\gamma) - \sin\beta\sin(2\varphi)\sin^2\gamma\right]$$

Clearly,

$$u(\beta,-\varphi,\gamma) = u(\beta,\varphi,\gamma)$$
$$v(\beta,-\varphi,\gamma) = -v(\beta,\varphi,\gamma)$$

and since from eq. [5] it is clear that the signs of both u and v bear no significance after α averaging, the φ averaging interval can be reduced from $(-\pi,\pi]$ to $[0,\pi]$. This φ integration interval in conjunction with

$$u(\beta,\pi-\varphi,\pi-\gamma) = u(\beta,\varphi,\gamma)$$
$$v(\beta,\pi-\varphi,\pi-\gamma) = -v(\beta,\varphi,\gamma)$$

imply that the γ dependence of the signal (eq.[5]) function S_f exhibits the following symmetry:

$$S_f(\gamma) = S_f(\pi-\gamma)$$

which is seen in all the numerical simulations.

APPENDIX 3

COMMUTATION PROPERTIES OF THE 3-SPIN HAMILTONIAN FOR THE HOMONUCLEAR DEPHASING SPINS CASE

With the notation from eqs. [1 and 2], the first commutator involved in the solution of the equations of motion assumes the usual form:

$$-i\left[H_{D,X=S}^{(trunc,lab)},I_x\right]=\sqrt{\frac{2}{3}}\left[\omega_{(0),20}^{(lab)}(t)S_z+\omega_{(1),20}^{(lab)}(t)X_z\right]I_y$$

but, once the "z" components of the dephasing spins are introduced, due to the presence of the $S_{\pm}$ and $X_{\pm}$ operators in the Hamiltonian, the results become complex:

$$\begin{aligned}(-i)^2\left[H_{D,X=S}^{(trunc,lab)},\left[H_{D,X=S}^{(trunc,lab)},I_x\right]\right]=&-\frac{2}{3}\left[\omega_{(0),20}^{(lab)}(t)S_z+\omega_{(1),20}^{(lab)}(t)X_z\right]^2I_x+\\&+\frac{1}{6}i\left[\omega_{(1),20}^{(lab)}(t)-\omega_{(0),20}^{(lab)}(t)\right]\omega_{(S1),20}^{(lab)}(t)(S_+X_--S_-X_+)I_y\end{aligned}\qquad[21]$$

Obviously, the term containing the raising and lowering operators spoils the hope for a general analytical solution. The exception is the one discussed at eq. [18]. In this singular case the $S_{\pm}$ and $X_{\pm}$ containing terms are zero in all orders of commutators (due to the fact that the first two terms of the Hamiltonian combine to give a multiple of $(S+X)_z$) and an analytical solution can then be written.

TABLE 1a

Minimized energies and backbone dihedral angles of all L-Pro-L-Ala and L-Pro-D-Ala optimized conformations found. The numbering of the conformations given here is respected in all tables and throughout the paper. For example, In table 2a, the bond lengths referring to conformation #2 are the optimized bond lengths obtained for the Pro-L-Ala conformation that is 1.4935 kcal/mol higher than the lowest energy one.

	Pro-L-Ala					Pro-D-Ala			
	ψ(Pro)	φ(Ala)	E(Hartrees)	ΔE(kcal/mol)		ψ(Pro)	φ(Ala)	E(Hartrees)	ΔE(kcal/mol)
1	-17.6	-158.4	-644.32607		1	155.6	162.8	-644.32003	
2	-18.6	-85.0	-644.32369	1.49345	2	33.4	160.3	-644.31936	0.42043
3	160.0	-161.0	-644.32008	3.75872	3	37.5	85.1	-644.31895	0.6777
4	-21.4	66.4	-644.31855	4.7188	4	149.3	85.3	-644.31613	2.44725
5	-150.6	-161.2	-644.31791	5.1204	5	158.8	-68.9	-644.31262	4.64977
6	159.3	-84.1	-644.31744	5.41532	6	-39.8	159.3	-644.31171	5.2208
7	78.4	-161.3	-644.31719	5.5722	7	38.0	-66.0	-644.31114	5.57847
8	-156.7	-87.8	-644.31440	7.32292	8	-31.9	85.0	-644.30920	6.79582
9	76.7	-84.1	-644.31380	7.69942	9	-42.5	-63.0	-644.30129	11.75935
10	-172.4	69.0	-644.30923	10.5671					
11	-153.1	66.0	-644.30866	10.92477					
12	82.7	68.7	-644.30402	13.83637					

Table 1b.
Minimized energies and backbone dihedral angles for the L-Pro-Gly dipeptide.

Pro-Gly			
ψ(Pro)	φ(Gly)	E(Hartrees)	ΔE(kcal/mol)
-18.5	-175.5	-605.304415	
-21.2	93.0	-605.300570	2.412737
-19.0	-84.7	-605.300450	2.488037
-141.4	179.7	-605.296493	4.971055
74.7	179.4	-605.295609	5.525765

TABLES 2a, b and c
Comparison between the calculated and experimental bond lengths for the three dipeptides studied in this work. The first 4 columns in each table give the 6-31G optimized values for the lowest energy 4 conformers. The last column in each table contains the experimental values (from reference [10]) for N-iBu-L-Pro-L-Ala-NH-iPr (Table 2a); N-iBu-L-Pro-D-Ala-NH-iPr (Table 2b); tBoc-L-Pro-Gly-OH. All distances are in Å. (see Table 1a for the corresponding energies)

TABLE 2a. L-Pro-L-Ala

	1	2	3	4	X-RAY
N1-C2	1.4671	1.4671	1.4587	1.4671	-
N1-H14	0.9970	0.9970	0.9954	0.9969	-
N1-C5	1.4541	1.4556	1.4459	1.4552	1.48
C2-C3	1.5470	1.5470	1.5556	1.5470	-
C2-H15	1.0824	1.0824	1.0826	1.0824	-
C3-C4	1.5450	1.5470	1.5447	1.5450	-
C4-C5	1.5408	1.5450	1.5524	1.5408	-
C5-C6	1.5119	1.5408	1.5127	1.5119	1.50
C5-H21	1.0867	1.0867	1.0860	1.0867	-
C6-O7	1.2339	1.2339	1.2323	1.2339	1.23
C6-N8	1.3437	1.3437	1.3487	1.3437	1.33
N8-H22	0.9961	0.9961	0.9941	0.9961	1.46
N8-C9	1.4453	1.4453	1.4493	1.4453	1.44
C9-C10	1.5031	1.5031	1.5041	1.5031	1.51
C9-C11	1.5384	1.5384	1.5379	1.5384	-
C9-H23	1.0824	1.0824	1.0826	1.0824	-
C10-O12	1.3468	1.3468	1.3441	1.3468	-
C10-O13	1.2124	1.2124	1.2132	1.2124	1.25
C11-H24	1.0817	1.0817	1.0814	1.0817	-
C11-H25	1.0825	1.0824	1.0826	1.0824	-
O12-H27	0.9545	0.9545	0.9547	0.9545	-

TABLE 2b. Pro-D-Ala

	1	2	3	4	X-RAY
N1-C2	1.4611	1.4724	1.4712	1.4597	-
N1-H14	0.9964	0.9999	1.0007	0.9959	-
N1-C5	1.4489	1.4680	1.4692	1.4470	1.45
C2-C3	1.5571	1.5607	1.5608	1.5567	-
C2-H15	1.0826	1.0819	1.0816	1.0822	-
C3-C4	1.5445	1.5469	1.5470	1.5444	-
C4-C5	1.5514	1.5317	1.5321	1.5527	-
C5-C6	1.5120	1.5085	1.5079	1.5109	1.49
C5-H21	1.0851	1.0871	1.0869	1.0855	-
C6-O7	1.2326	1.2318	1.2288	1.2297	1.23
C6-N8	1.3478	1.3487	1.3540	1.3542	1.34
N8-H22	0.9942	0.9961	0.9934	0.9917	1.47
N8-C9	1.4494	1.4480	1.4397	1.4422	1.46
C9-C10	1.5038	1.5040	1.5128	1.5120	1.46
C9-C11	1.5378	1.5381	1.5338	1.5342	-
C9-H23	1.0825	1.0819	1.0816	1.0822	-
C10-O12	1.3438	1.3458	1.3513	1.3521	-
C10-O13	1.2132	1.2123	1.2092	1.2077	1.26
C11-H24	1.0814	1.0815	1.0831	1.0833	-
C11-H25	1.0825	1.0819	1.0816	1.0822	-
O12-H27	0.9546	0.9546	0.9543	0.9543	-

TABLE 2c. L-Pro-Gly.

	1	2`	3	4	X-RAY
N1-C2	1.4683	1.4702	1.4696	1.4557	
N1-H13	0.9972	0.9977	0.9979	0.9935	
N1-C5	1.4554	1.4576	1.4574	1.4457	1.469
C2-C3	1.5466	1.5464	1.5469	1.5517	
C2-H14	1.0826	1.0825	1.0826	1.0834	
C3-C4	1.5450	1.5450	1.5450	1.5468	
C4-C5	1.5408	1.5408	1.5398	1.5341	
C5-C6	1.5117	1.5116	1.5115	1.5118	1.520
C5-H20	1.0865	1.0859	1.0861	1.0955	
C6-O7	1.2328	1.2304	1.2302	1.2311	1.223
C6-N8	1.3435	1.3485	1.3482	1.3497	1.346
N8-H21	0.9959	0.9932	0.9927	0.9928	1.0
N8-C9	1.4382	1.4316	1.4323	1.4413	1.447
C9-C10	1.4939	1.5072	1.5079	1.4942	1.520
C9-H22	1.0830	1.0821	1.0805	1.0826	
C9-H23	1.0837	1.0810	1.0822	1.0829	
C10-O11	1.3460	1.3458	1.3455	1.3436	
C10-O12	1.2116	1.2124	1.2122	1.2126	1.228
O11-H24	0.9544	0.9548	0.9548	0.9544	

TABLES 3a, b and c

Comparison between the calculated and experimental bond angles for the three dipeptides studied in this work. The first 4 columns in each table give the 6-31G optimized values for the lowest energy 4 conformers. The last column in each table contains the experimental values (from reference [10]) for N-iBu-L-Pro-L-Ala-NH-iPr (Table 2a); N-iBu-L-Pro-D-Ala-NH-iPr (Table 2b); tBoc-L-Pro-Gly-OH. All angles are in degrees. (see table 1a for the corresponding energies)

Table 3a Pro-L-Ala

	1	2	3	4	X-RAY
N1-C2-C3	104.66	104.66	106.29	104.66	-
N1-C5-C6	113.11	113.03	110.97	112.96	112.6
C2-C3-C4	105.66	105.66	105.67	105.66	-
C3-C4-C5	104.69	104.69	104.30	104.69	-
C4-C5-C6	111.56	111.56	112.02	111.56	-
C5-C6-O7	120.58	120.58	122.09	120.58	119.1
C5-C6-N8	116.46	116.46	115.99	116.46	116.6
O7-C6-N8	122.96	122.96	121.91	122.96	124.3
C6-N8-C9	121.96	121.96	122.10	121.96	124.9
N8-C9-C10	108.08	108.09	107.82	108.09	114.4
N8-C9-C11	112.73	112.73	112.60	112.73	-
C10-C9-C11	110.58	110.44	110.59	114.16	-
C9-C10-O12	112.07	112.07	112.30	112.07	120.7
C9-C10-O13	125.87	125.87	125.49	125.87	
O12-C10-O13	122.06	122.06	122.21	122.06	-
C2-N1-H14	115.79	115.65	117.88	115.06	-
N1-C2-H15	110.96	110.96	110.52	110.96	-
C3-C2-H15	111.23	111.25	110.81	111.23	-
N1-C2-H16	110.96	110.96	110.51	110.96	-
H15-C2-H16	107.84	107.81	107.93	107.84	-
C2-C3-H17	111.35	111.35	111.24	111.35	-
C4-C3-H17	110.55	110.57	110.79	110.55	-
H17-C3-H18	107.43	107.39	107.17	107.43	-
C3-C4-H19	111.35	111.35	111.25	111.35	-
C5-C4-H19	111.04	111.06	111.47	111.04	-
H19-C4-H20	107.43	107.39	107.16	107.43	-
C4-C5-H21	109.90	110.11	109.57	110.13	-
C6-C5-H21	106.29	106.34	108.64	106.32	-
C6-N8-H22	118.44	118.44	120.66	118.44	129.5
C9-N8-H22	119.60	119.60	117.23	119.60	111.7
N8-C9-H23	108.85	108.85	108.98	108.85	-
C10-C9-H23	107.93	106.90	108.01	105.37	-
C11-C9-H23	108.54	109.64	108.73	107.28	-
C9-C11-H24	109.62	109.62	109.69	109.62	-
C9-C11-H25	110.68	110.68	110.53	110.68	-
H24-C11-H25	108.78	108.78	108.83	108.78	-
C9-C11-H26	109.62	109.62	109.69	109.62	-
H24-C11-H26	109.32	109.32	109.25	109.32	-
H25-C11-H26	108.78	108.78	108.83	108.78	-
C10-O12-H27	114.13	114.13	114.37	114.13	-

TABLE 3b Pro-D-Ala

	1	2	3	4	X-RAY
N1-C2-C3	106.41	106.47	106.37	106.35	-
N1-C5-C6	110.76	110.92	110.71	110.82	110.8
C2-C3-C4	105.51	105.64	105.570	105.58	-
C3-C4-C5	104.05	104.28	104.246	104.23	-
C4-C5-C6	112.04	111.56	111.970	111.97	-
C5-C6-O7	121.84	121.89	122.128	122.22	122.9
C5-C6-N8	116.15	115.72	116.117	116.28	116.6
O7-C6-N8	121.99	122.31	121.733	121.57	120.7
C6-N8-C9	122.16	123.63	121.944	121.91	120.3
N8-C9-C10	107.56	109.86	110.619	110.78	112.9
N8-C9-C11	112.91	113.88	110.838	110.79	-
C10-C9-C11	110.53	113.17	110.143	110.12	-
C9-C10-O12	112.23	112.01	110.817	110.82	119.1
C9-C10-O13	125.43	126.09	126.769	126.78	-
O12-C10-O13	122.30	121.90	122.343	122.33	-
C2-N1-H14	117.04	116.97	117.644	117.65	-
N1-C2-H15	110.38	110.37	110.428	110.44	-
C3-C2-H15	110.79	110.81	110.803	110.81	-
N1-C2-H16	110.32	110.36	110.451	110.41	-
H15-C2-H16	108.11	108.04	108.024	108.01	-
C2-C3-H17	111.26	111.32	111.296	111.23	-
C4-C3-H17	110.84	110.79	110.838	110.83	-
C2-C3-H18	111.27	111.32	111.234	111.23	-
C4-C3-H18	110.82	110.74	110.829	110.86	-
H17-C3-H18	107.18	107.05	107.187	107.19	-
C3-C4-H19	111.27	111.32	111.234	111.23	-
C5-C4-H19	111.51	111.46	111.578	111.58	-
C5-C4-H20	111.56	111.49	111.528	111.51	-
H19-C4-H20	107.19	107.04	107.152	107.14	-
C4-C5-H21	109.56	109.72	109.697	109.66	-
C6-C5-H21	109.10	109.06	108.834	108.82	-
C6-N8-H22	120.69	118.62	119.573	119.59	128.0
C9-N8-H22	117.19	117.76	118.507	118.56	111.7
N8-C9-H23	108.94	106.22	108.586	108.64	-
C10-C9-H23	107.98	105.25	106.648	106.61	-
C11-C9-H23	108.80	107.86	109.836	109.80	-
C9-C11-H24	109.68	109.95	110.315	110.31	-
C9-C11-H25	110.51	110.07	110.425	110.47	-
H24-C11-H25	108.84	108.96	108.527	108.54	-
H24-C11-H26	109.22	108.97	108.672	108.67	-
C10-O12-H27	114.34	113.70	114.096	114.11	-

TABLE 3c ProGly

	1	2	3	4	X-RAY
N1-C2-C3	104.74	104.85	104.77	103.81	-
N1-C5-C6	112.84	113.11	111.91	112.42	111.3
C2-C3-C4	105.69	105.70	105.65	105.34	-
C3-C4-C5	104.77	104.83	104.63	103.33	-
C4-C5-C6	111.55	111.14	111.88	118.07	-
C5-C6-O7	120.87	120.87	120.87	120.94	122.9
C5-C6-N8	116.59	115.93	116.14	117.93	112.6
O7-C6-N8	122.54	123.19	122.98	121.13	123.4
C6-N8-C9	121.26	122.25	122.08	120.86	124.9
N8-C9-C10	109.77	115.41	115.37	109.31	110.7
C9-C10-O11	111.61	114.98	115.09	111.82	-
C9-C10-O12	125.85	122.96	122.86	125.46	120.7
O11-C10-O12	122.54	122.06	122.05	122.72	-
C2-N1-H13	115.51	115.16	115.42	120.21	-
N1-C2-H14	110.90	110.85	110.90	111.56	-
C3-C2-H14	111.24	111.21	111.23	111.22	-
H14-C2-H15	107.86	107.91	107.86	107.51	-
C2-C3-H16	111.34	111.36	111.35	111.17	-
C4-C3-H16	110.56	110.55	110.58	110.63	-
H16-C3-H17	107.40	107.37	107.38	107.93	-
C5-C4-H18	111.03	110.99	111.35	111.63	-
C4-C5-H20	109.95	110.26	109.97	108.62	-
C6-C5-H20	106.35	106.49	106.31	106.12	-
C6-N8-H21	118.97	116.85	117.02	121.72	129.5
C9-N8-H21	119.77	120.90	120.91	117.42	114.5
N8-C9-H22	111.25	111.15	109.58	111.44	-
C10-C9-H22	108.94	107.11	106.65	109.10	-
N8-C9-H23	111.90	108.40	110.36	111.51	-
C10-C9-H23	108.67	106.90	107.01	108.95	-
H22-C9-H23	106.19	107.52	106.47	106.45	-
C10-O11-H24	114.28	113.35	113.19	114.39	-

TABLE 4
Calculated and experimental dihedral angles (in degrees).

TABLE 4a Pro-L-Ala

	1	2	3	4	X ray
C2-C3-C4-C5	19.617	19.731	16.161	19.700	
C6-C5-C4-N1	-121.342	-121.085	-120.564	-121.081	
O7-C6-C5-N1	162.397	161.410	-19.996	158.556	
ψ(Pro)=N8-C6-C5-N1	-17.603	-18.59	160.004	-21.444	145
N8-C6-C5-O7	180.000	180.000	180.000	180.000	
φ(Ala)=C10-C9-N8-C6	-158.436	-85.009	-160.978	66.426	-89
C11-C9-N8-C6	79.034	152.631	76.749	-60.705	
O12-C10-C9-N8	169.659	166.724	169.659	-164.854	
O13-C10-C9-O12	180.000	180.000	180.000	180.000	
H14-N1-C5-C2	135.466	135.194	-136.031	134.309	
H15-C2-N1-C3	120.066	120.087	120.299	120.065	
H21-C5-N1-C4	-118.483	-118.694	-118.196	-118.726	
H22-N8-C6-O7	180.000	180.000	180.000	180.000	
H23-C9-N8-C10	116.982	115.764	117.007	113.964	
H24-C11-C9-N8	60.032	59.022	60.163	65.163	

Footnote to Table 4a.
The structure #4 is very close in energy (within 0.8kcal/mol) to the next three structures which (numbers 5, 6 and 7 in Table. 1). Among these, the structure #6 is the one that has the same conformation as the X-ray studied peptide.

TABLE 4b Pro-D-Ala

	1	2	3	4	X ray
C2-C3-C4-C5	17.109	21.096	20.932	16.622	
C6-C5-C4-N1	-120.343	-125.085	-123.867	-120.402	
O7-C6-C5-N1	-24.426	-146.636	-142.458	-30.734	
ψ(Pro)=N8-C6-C5-N1	155.574	33.364	37.542	149.266	137
N8-C6-C5-O7	180.000	180.000	180.000	180.000	
φ(Ala)=C10-C9-N8-C6	162.832	160.293	85.083	85.325	96
C11-C9-N8-C6	285.040	-77.121	-152.383	-152.161	
O12-C10-C9-N8	190.390	-170.88	-168.983	-165.967	
O13-C10-C9-O12	180.000	180.000	180.000	180.000	
H14-N1-C5-C2	-133.403	-124.664	-124.44	-135.087	
H15-C2-N1-C3	120.277	120.317	120.317	120.295	
H21-C5-N1-C4	-118.008	-115.820	-116.185	-118.244	
H22-N8-C6-O7	180.000	180.000	180.000	180.000	
H23-C9-N8-C10	-116.801	-116.835	-116.472	-116.753	
H24-C11-C9-N8	60.132	59.565	60.275	59.965	

TABLE 4c ProGly

	1	2	3	4	X-ray
N1-C2-C3-C4	0	0	0	0	
C2-C3-C4-C5	19.405	19.326	19.852	22.497	
C6-C5-C4-N1	-121.323	-121.046	-121.394	120.239	
O7-C6-C5-N1	161.483	158.551	160.951	38.579	
ψ(Pro)=N8-C6-C5-N1	-18.517	-21.449	-19.049	-141.421	157
N8-C6-C5-O7	180.000	180.000	180.000	180.000	
φ(Gly)=C10-C9-N8-C6	-175.489	99.110	-84.671	179.744	-76
O11-C10-C9-N8	179.485	-10.717	-6.935	179.030	
O12-C10-C9-O12	180.000	180.000	180.000	180.000	
H13-N1-C5-C2	134.699	133.494	133.536	143.501	
H14-C2-N1-C3	120.092	120.097	120.097	119.867	
H20-C5-N1-C4	-118.545	-118.949	-118.423	-116.043	
H21-N8-C6-O7	180.000	180.000	180.000	180.000	
H22-C9-N8-C10	120.653	122.212	120.424	120.678	
H23-C9-N8-C6	63.771	-20.711	154.089	59.220	

Table 5a
Total atomic charges in L-Pro-L-Ala:

	1	2	3	4
1 N	-0.813019	-0.815873	-0.776892	-0.810843
2 C	-0.056169	-0.057772	-0.070516	-0.060753
3 C	-0.352572	-0.352406	-0.349432	-0.352712
4 C	-0.300255	-0.301194	-0.319306	-0.305621
5 C	-0.024998	-0.030888	-0.024003	-0.035822
6 C	0.781408	0.792795	0.798107	0.807402
7 O	-0.646827	-0.641785	-0.644577	-0.647487
8 N	-0.903980	-0.868139	-0.917439	-0.871850
9 C	0.008713	-0.029115	0.006809	-0.038728
10 C	0.770732	0.793040	0.773922	0.806727
11 C	-0.447182	-0.459406	-0.447937	-0.440160
12 O	-0.732669	-0.736615	-0.730402	-0.737758
13 O	-0.560246	-0.555825	-0.563563	-0.553216
14 H	0.339148	0.344722	0.374346	0.331824
15 H	0.151771	0.153996	0.166363	0.151583
16 H	0.173927	0.169375	0.162187	0.169768
17 H	0.173094	0.174253	0.168668	0.171991
18 H	0.176928	0.175311	0.171992	0.177325
19 H	0.188703	0.190151	0.167581	0.189121
20 H	0.182881	0.180808	0.183605	0.208456
21 H	0.195023	0.200225	0.187799	0.194348
22 H	0.441125	0.416483	0.413115	0.415721
23 H	0.253734	0.262331	0.255280	0.223483
24 H	0.159540	0.166627	0.162276	0.227182
25 H	0.184901	0.201915	0.186338	0.180382
26 H	0.211897	0.184916	0.218538	0.157039
27 H	0.444392	0.442070	0.447142	0.442598

Table 5b
Total atomic charges in L-Pro-D-Ala:

	1	2	3	4
1 N	-0.758422	-0.740873	-0.760497	-0.760497
2 C	-0.080600	-0.097253	-0.077390	-0.077390
3 C	-0.348875	-0.361357	-0.348628	-0.348628
4 C	-0.318226	-0.287438	-0.319161	-0.319161
5 C	-0.038638	-0.076934	-0.038272	-0.038272
6 C	0.804184	0.780755	0.806334	0.806334
7 O	-0.647260	-0.636450	-0.635046	-0.635046
8 N	-0.921498	-0.916837	-0.878271	-0.878271
9 C	0.012533	0.012457	-0.022764	-0.022764
10 C	0.774274	0.769279	0.784024	0.784024
11 C	-0.452190	-0.449952	-0.465668	-0.465668
12 O	-0.730416	-0.732417	-0.737399	-0.737399
13 O	-0.563417	-0.557989	-0.540946	-0.540946
14 H	0.370483	0.328063	0.373107	0.373107
15 H	0.169146	0.179712	0.169175	0.169175
16 H	0.163453	0.168049	0.163462	0.163462
17 H	0.168603	0.170003	0.167996	0.167996
18 H	0.171553	0.169689	0.171915	0.171915
19 H	0.164975	0.182505	0.158928	0.158928
20 H	0.184134	0.181896	0.184022	0.184022
21 H	0.190469	0.215473	0.192658	0.192658
22 H	0.414874	0.437834	0.385124	0.385124
23 H	0.255593	0.250265	0.266273	0.266273
24 H	0.219814	0.217823	0.185374	0.185374
25 H	0.187595	0.185953	0.205839	0.205839
26 H	0.160410	0.161621	0.166464	0.166464
27 H	0.447449	0.446125	0.443348	0.443348

Table 5c
Total atomic charges in Pro-Gly:

	ProGly			
	1	2	3	4
1 N	-0.807182	-0.809801	-0.811572	-0.794879
2 C	-0.059516	-0.062214	-0.061506	-0.040454
3 C	-0.352320	-0.351990	-0.352405	-0.352059
4 C	-0.301782	-0.307203	-0.300986	-0.310554
5 C	-0.028911	-0.040839	-0.037866	-0.027490
6 C	0.779652	0.783964	0.782951	0.825041
7 O	-0.641434	-0.632788	-0.632658	-0.637602
8 N	-0.917470	-0.865830	-0.864137	-0.942359
9 C	-0.081042	-0.161255	-0.162023	-0.083360
10 C	0.739973	0.758123	0.762087	0.743265
11 O	-0.729492	-0.715240	-0.719468	-0.726855
12 O	-0.554625	-0.556371	-0.554440	-0.560313
13 H	0.336229	0.333822	0.340711	0.362075
14 H	0.152145	0.154246	0.154803	0.144347
15 H	0.175390	0.172413	0.171947	0.169205
16 H	0.173323	0.174195	0.175335	0.168527
17 H	0.177842	0.177878	0.177441	0.176613
18 H	0.189565	0.190214	0.191777	0.162378
19 H	0.184676	0.200524	0.183006	0.185840
20 H	0.195515	0.198052	0.199743	0.178411
21 H	0.443976	0.417969	0.418862	0.422642
22 H	0.243016	0.226429	0.272109	0.245377
23 H	0.236904	0.275068	0.225960	0.244622
24 H	0.445568	0.440633	0.440329	0.447582

Tables 6a and 6b.

Minimized energies and backbone dihedral angles of the four lowest energy conformers found for Ac-*L*-Pro-*L*-Ala-NH_2 and Ac-*L*-Pro-*D*-Ala-NH_2. The numbering of the conformations given here is followed in all tables and throughout the paper. For example, In table 2a, the bond lengths referring to conformation #2 are the optimized bond lengths obtained for the Ac-*L*-Pro-*D*-Ala-NH_2 conformation that is 5.7730 kcal/mol higher than the lowest energy one.

Table 6a. Ac-*L*-Pro-*D*-Ala-NH_2.

	ψ(Pro)	φ(Ala)	ψ(Ala)	E(Hartree)	ΔE(kcal/mol)
1	103.2	126.2	-9.8	-776.23328	
2	-13.6	-74.8	-9.7	-776.22408	5.77300
3	116.2	117.0	-88.2	-776.22331	6.25617
4	118.2	136.2	-155.6	-776.22005	8.30182

Table 6b. Ac-*L*-Pro-*L*-Ala-NH_2.

	ψ(Pro)	φ(Ala)	ψ(Ala)	E(Hartree)	
1	70.7	-117.7	19.5	-776.23037	
2	-4.0	-117.1	18.7	-776.22811	1.41815
3	140.5	-146.5	56.6	-776.22405	3.96580
4	106.9	-173.6	-17.3	-776.21710	8.32692

TABLES 7a and 7b

Calculated distances in the *ab initio* models reported in this work compared with the X-ray experimental bond lengths measured on some related molecules. The first 4 columns in each table give the 6-31G optimized values for the lowest energy 4 conformers found. The last column in each table contains the experimental values (from reference [10]) for N-iBu-*L*-Pro-*D*-Ala-NH-iPr (Table 2a); N-iBu-*L*-Pro-*L*-Ala-NH-iPr (Table 2b). All distances are in Å (see Tables 1a and 1b for the corresponding energies).

TABLE 7a. Distances in Ac-*L*-Pro-*D*-Ala-NH_2 (Å).

	1	2	3	4	X-RAY
N1-C2	1.4682	1.4657	1.4659	1.4653	-
N1-C14	1.3466	1.3539	1.2240	1.3526	1.34
N1-C5	1.4670	1.4599	1.4542	1.4625	1.45
C2-C3	1.5392	1.5406	1.5429	1.5399	-
C2-H17	1.0811	1.0814	1.0814	1.0815	-
C3-C4	1.5450	1.5451	1.5439	1.5451	-
C4-C5	1.5331	1.5378	1.5403	1.5369	-
C5-C6	1.5232	1.5213	1.5195	1.5270	1.49
C5-H23	1.0792	1.0782	1.0770	1.078	-
C6-O7	1.2312	1.2343	1.2358	1.2333	1.23
C6-N8	1.3454	1.3454	1.3454	1.3380	1.36
N8-H24	0.9948	0.9951	0.9946	0.9973	1.47
N8-C9	1.4516	1.4593	1.4480	1.4416	1.46
C9-C10	1.5292	1.5233	1.5196	1.5329	1.46
C9-C11	1.5401	1.5348	1.5456	1.5424	-
C10-N12	1.3456	1.3482	1.3552	1.3487	1.33
C9-H25	1.0812	1.0787	1.0801	1.0867	-
C10-O13	1.2307	1.2278	1.2241	1.2288	1.26
O15-C14	1.2428	1.2351	1.2327	1.2356	1.27
C16-C14	1.5023	1.5032	1.5040	1.5037	-
C11-H26	1.0815	1.0815	1.0826	1.0804	-
N12-H29	0.9993	0.9908	0.9904	0.9898	-
N12-H30	0.9925	0.9925	0.9925	0.9926	-
C16-H31	1.0813	1.0815	1.0815	1.0815	-
O15-H29	2.2066	2.3970	6.4848	6.7819	-
O15-H24	2.6133	3.1198	3.2545	3.1859	-

TABLE 7b. Distances in Ac-*L*-Pro-*L*-Ala-NH_2 (Å).

	1	2	3	4	X-RAY
N1-C2	1.4693	1.4682	1.4675	1.4694	-
N1-C14	1.3493	1.3561	1.3463	1.3461	1.32
N1-C5	1.4681	1.4599	1.4640	1.4657	1.48
C2-C3	1.5438	1.5406	1.5400	1.5401	-
C2-H17	1.0811	1.0815	1.0818	1.0810	-
C3-C4	1.5449	1.5471	1.5438	1.5442	-
C4-C5	1.5269	1.5395	1.5361	1.5322	-
C5-C6	1.5294	1.5209	1.5322	1.5308	1.50
C5-H23	1.0803	1.0799	1.0805	1.0787	-
C6-O7	1.2300	1.2299	1.2297	1.2310	1.27
C6-N8	1.3513	1.3476	1.3464	1.3505	1.32
N8-H24	1.0004	0.9917	0.9948	0.9945	1.46
N8-C9	1.4547	1.4531	1.4687	1.4693	1.44
C9-C10	1.5291	1.5312	1.5318	1.5299	1.51
C9-C11	1.5402	1.5380	1.5382	1.5512	-
C10-N12	1.3482	1.3458	1.3322	1.3441	1.33
C9-H25	1.0819	1.0816	1.0840	1.0831	-
C10-O13	1.2283	1.2301	1.2291	1.2305	1.25
C14-O15	1.2404	1.2357	1.2349	1.2410	1.23
C16-C14	1.5026	1.5035	1.5027	1.5027	-
C11-H26	1.0812	1.0816	1.0811	1.0781	-
N12-H29	0.9880	0.9932	0.9930	0.9927	-
N12-H30	0.9921	0.9924	0.9926	0.9925	-
C16-H31	1.0813	1.0815	1.0813	1.0813	-
O15-H29	4.1231	2.4655	6.5402	2.6275	-
O15-H24	2.1097	2.9251	3.8052	2.8012	-

TABLE 8a. Bond angles in Ac-*L*-Pro-*D*-Ala-NH_2 (degrees).

	1	2	3	4	X-RAY
N1-C2-C3	105.25	105.30	104.97	104.98	-
N1-C5-C6	108.34	114.93	109.44	109.18	110.8
C2-C3-C4	106.83	107.05	106.48	106.16	-
C3-C4-C5	106.78	107.06	106.01	105.41	-
C4-C5-C6	112.91	112.38	111.29	110.86	-
C5-C6-O7	121.89	119.63	120.44	120.26	122.9
C5-C6-N8	115.39	118.35	116.74	116.86	116.6
O7-C6-N8	123.72	123.02	122.82	122.89	120.8
C6-N8-C9	124.26	124.32	123.40	122.71	123.6
N8-C9-C10	110.95	113.22	107.14	109.04	112.9
N8-C9-C11	112.52	114.67	112.21	112.47	-
C10-C9-C11	111.15	111.48	111.52	110.41	-
C9-C10-N12	118.25	118.29	115.54	115.29	120.1
C9-C10-O13	119.54	118.94	121.46	122.32	119.1
N12-C10-O13	122.21	122.77	122.84	122.39	120.7
C2-N1-C14	126.37	125.77	127.14	127.49	-
N1-C14-O15	120.83	120.60	120.90	121.29	119.7
N1-C14-C16	118.06	117.99	117.71	117.41	-
O15-C14-C16	121.10	121.40	121.39	121.30	-
N1-C2-H17	110.24	110.61	110.37	110.38	-
C3-C2-H17	111.40	111.50	111.41	111.30	-
N1-C2-H18	110.24	110.61	110.37	110.38	-
H17-C2-H18	108.32	107.35	108.31	108.49	-
C2-C3-H19	110.68	110.93	110.839	110.93	-
C4-C3-H19	110.41	110.47	110.44	110.46	-
C2-C3-H20	110.68	110.93	110.84	110.93	-
H19-C3-H20	107.86	107.03	107.83	107.92	-
C5-C4-H21	110.43	110.46	110.68	110.84	-
C4-C5-H23	111.93	111.12	112.60	113.34	-
C6-C5-H23	110.04	104.39	110.62	110.78	-
C6-N8-H24	118.28	118.81	119.505	120.09	128.0
C9-N8-H24	117.45	116.87	117.10	117.21	111.7
N8-C9-H25	106.67	104.56	106.752	106.98	-
C10-C9-H25	105.88	104.53	108.73	108.72	-
C11-C9-H25	109.33	107.42	110.29	109.10	-
C9-C11-H26	110.36	110.69	110.89	110.856	-
C9-C11-H27	110.75	109.65	110.28	110.687	-
H26-C11-H27	108.54	108.11	109.14	108.14	-
H26-C11-H28	108.40	108.87	107.21	107.46	-
C10-N12-H29	124.53	124.17	122.752	123.226	-
C10-N12-H30	117.20	117.58	118.511	118.290	-
H29-N12-H30	118.27	118.24	118.74	118.48	-
C14-C16-H31	110.21	110.30	110.207	110.193	-
H31-C16-H32	108.72	108.63	108.73	108.74	-

TABLE 8b. Bond angles in Ac-*L*-Pro-*L*-Ala-NH_2 (degrees).

	1	2	3	4	X-RAY
N1-C2-C3	104.88	105.18	105.60	105.02	-
N1-C5-C6	110.36	115.5	112.25	110.36	112.6
C2-C3-C4	106.30	106.78	103.05	106.50	-
C3-C4-C5	105.98	106.66	102.85	106.47	-
C4-C5-C6	113.00	112.37	111.60	112.12	-
C5-C6-O7	120.03	118.61	120.92	119.03	119.1
C5-C6-N8	115.00	118.38	117.09	114.84	116.6
O7-C6-N8	123.17	123.01	121.96	123.10	124.3
C6-N8-C9	123.38	123.67	124.40	131.67	124.9
N8-C9-C10	111.93	110.71	112.75	111.70	114.4
N8-C9-C11	111.89	112.35	108.50	116.37	-
C10-C9-C11	111.21	111.59	113.06	111.25	-
C9-C10-N12	117.66	117.57	118.12	119.89	116.7
C9-C10-O13	120.53	120.19	122.24	118.01	120.7
N12-C10-O13	121.80	122.24	119.61	120.93	122.3
C2-N1-C14	126.76	125.98	125.30	127.02	-
N1-C14-O15	120.81	120.65	121.98	120.82	122.3
N1-C14-C16	118.06	117.97	118.05	118.16	-
O15-C14-C16	121.12	121.38	119.96	119.88	-
N1-C2-H17	110.43	110.59	112.25	110.26	-
C3-C2-H17	111.37	111.47	110.31	111.17	-
N1-C2-H18	110.43	110.59	109.21	110.33	-
H17-C2-H18	108.36	107.57	109.84	107.79	-
C2-C3-H19	110.82	111.02	109.50	110.74	-
C4-C3-H19	110.48	110.46	109.58	109.84	-
C2-C3-H20	110.82	111.02	112.32	111.58	-
H19-C3-H20	107.95	107.13	109.85	107.47	-
C5-C4-H21	110.64	110.52	113.63	111.26	-
C4-C5-H23	112.42	111.14	108.85	109.31	-
C6-C5-H23	108.64	104.63	109.33	106.72	-
C6-N8-H24	118.31	119.13	117.47	117.25	129.5
C9-N8-H24	118.31	117.20	118.05	118.43	111.7
N8-C9-H25	106.51	106.32	109.46	108.68	-
C10-C9-H25	105.65	105.91	105.56	105.72	-
C11-C9-H25	109.32	109.62	107.32	108.09	-
C9-C11-H26	110.50	110.72	110.47	110.37	-
C9-C11-H27	110.49	110.24	110.33	110.33	-
H26-C11-H27	108.29	108.29	108.85	108.62	-
H26-C11-H28	108.57	108.34	108.36	108.41	-
C10-N12-H29	122.56	124.02	120.71	122.46	-
C10-N12-H30	118.25	117.68	116.67	117.45	-
H29-N12-H30	119.19	118.30	120.11	118.93	-
C14-C16-H31	110.19	110.29	109.81	110.50	-
H31-C16-H32	108.74	108.63	108.56	108.14	-

TABLE 9a. Calculated and experimental [46] dihedral angles (in degrees) for Ac-*L*-Pro-*D*-Ala-NH_2 conformers.

	1	2	3	4	X ray
C2-C3-C4-C5	-11.0	-7.3	-14.5	-16.5	-
C6-C5-C4-N1	-117.6	-125.5	-117.6	-117.0	-
O7-C6-C5-N1	-76.8	166.4	-63.8	-61.8	-
ψ(Pro)=N8-C6-C5-N1	103.2	-13.6	116.2	118.2	137
N8-C6-C5-O7	180.0	180.0	180.0	180.0	-
φ(Ala)=C10-C9-N8-C6	126.2	-74.8	117.0	136.2	96
C11-C9-N8-C6	-108.5	54.6	120.3	-101.0	-
ψ(Ala)=N12-C10-C9-N8	-9.8	-9.7	-88.2	-155.6	3
O13-C10-C9-N12	180.0	180.0	180.0	180.0	-
C14-N1-C5-C2	180.0	180.0	180.0	180.0	-
H17-C2-N1-C3	-120.2	-120.6	-120.2	-120.1	-
H23-C5-N1-C4	-119.7	-119.2	120.1	-120.8	-
H25-C9-N8-C10	-114.9	-113.2	-116.3	-117.4	-

TABLE 9b. Calculated and experimental [46] dihedral angles (in degrees) for Ac-*L*-Pro-*L*-Ala-NH_2 conformers.

	1	2	3	4	X ray
C2-C3-C4-C5	-16.0	-10.8	-36.0	-13.9	-
C6-C5-C4-N1	-119.4	-125.5	-121.6	-116.0	-
O7-C6-C5-N1	-109.3	176.0	-41.3	-73.1	-
ψ(Pro)=N8-C6-C5-N1	70.7	-4.0	140.5	106.9	156
N8-C6-C5-O7	180.0	180.0	180.0	180.0	-
φ(Ala)=C10-C9-N8-C6	-117.7	-117.1	-146.5	-173.6	66
C11-C9-N8-C6	116.7	117.4	87.5	15.0	-
ψ(Ala)=N12-C10-C9-N8	19.5	18.7	56.6	-17.3	14
O13-C10-C9-N12	180.0	180.0	180.0	180.0	-
C14-N1-C5-C2	180.0	180.0	180.0	180.0	-
H17-C2-N1-C3	-120.1	-120.5	-120.2	-120.1	-
H23-C5-N1-C4	-119.7	-119.0	-116.8	-121.2	-
H25-C9-N8-C10	115.0	114.6	117.2	115.7	-

Table 10a.
Net atomic charges in Ac-*L*-Pro-*D*-Ala-NH_2.

	1	2	3	4
1 N	-0.858756	-0.905420	-0.872732	-0.893356
2 C	-0.014747	-0.017607	-0.016505	-0.011220
3 C	-0.345468	-0.356423	-0.337988	-0.328927
4 C	-0.317665	-0.334753	-0.317465	-0.310725
5 C	-0.027318	0.040307	0.032232	-0.022728
6 C	0.823291	0.788358	0.773873	0.803158
7 O	-0.635780	-0.614820	-0.647591	-0.603249
8 N	-0.907493	-0.921850	-0.918355	-0.902155
9 C	-0.016883	-0.008048	-0.011871	-0.009322
10 C	0.746497	0.776822	0.732549	0.765098
11 C	-0.459590	-0.417405	-0.422385	-0.453782
12 N	-0.967080	-0.951761	-0.943878	-0.962291
13 O	-0.629799	-0.630702	-0.623745	-0.613148
14 C	0.786080	0.752202	0.785603	0.762376
15 O	-0.680389	-0.634017	-0.644511	-0.679122
16 C	-0.537069	-0.534020	-0.533154	-0.541285
17 H	0.178659	0.188482	0.177998	0.184920
18 H	0.199854	0.173880	0.190274	0.185517
19 H	0.176480	0.190168	0.187734	0.182381
20 H	0.205528	0.186807	0.173592	0.192463
21 H	0.179980	0.198232	0.197167	0.181324
22 H	0.214602	0.212403	0.209819	0.213373
23 H	0.222891	0.261514	0.243397	0.252935
24 H	0.400314	0.387049	0.391235	0.389919
25 H	0.273326	0.209781	0.242763	0.253572
26 H	0.152995	0.151856	0.156217	0.155836
27 H	0.173172	0.170527	0.170436	0.172638
28 H	0.222512	0.214943	0.209826	0.223637
29 H	0.450454	0.436191	0.437295	0.437781
30 H	0.382576	0.384019	0.382661	0.379221
31 H	0.218909	0.225207	0.221292	0.217983
32 H	0.196086	0.180501	0.182424	0.183220
33 H	0.193829	0.197575	0.191791	0.193956

Table 10b.
Net atomic charges in Ac-*L*-Pro-*L*-Ala-NH_2.

	1	2	3	4
1 N	-0.873291	-0.891327	-0.887216	-0.859163
2 C	-0.014536	-0.017205	-0.013158	-0.015201
3 C	-0.346358	-0.352378	-0.316952	-0.314924
4 C	-0.325197	-0.323347	-0.279410	-0.303952
5 C	-0.023275	0.030740	0.019537	-0.028395
6 C	0.809214	0.793582	0.784752	0.823740
7 O	-0.647618	-0.650201	-0.621879	-0.617228
8 N	-0.925123	-0.918254	-0.893514	-0.912431
9 C	-0.010344	-0.009482	-0.006918	-0.008917
10 C	0.753292	0.762213	0.779516	0.745716
11 C	-0.461728	-0.417174	-0.445187	-0.427318
12 N	-0.942438	-0.950612	-0.942714	-0.971264
13 O	-0.640517	-0.631278	-0.609381	-0.643927
14 C	0.788165	0.749291	0.721292	0.735941
15 O	-0.684271	-0.631355	-0.675173	-0.651294
16 C	-0.531934	-0.540330	-0.544279	-0.539127
17 H	0.173728	0.184817	0.192017	0.179365
18 H	0.186283	0.178264	0.179218	0.194316
19 H	0.180225	0.189921	0.190551	0.178063
20 H	0.193929	0.188653	0.179310	0.176028
21 H	0.179331	0.194725	0.193072	0.183195
22 H	0.208941	0.214337	0.211937	0.209726
23 H	0.237458	0.258086	0.226823	0.249147
24 H	0.461826	0.384874	0.408291	0.392731
25 H	0.257613	0.210892	0.229148	0.264193
26 H	0.164284	0.152363	0.158027	0.150293
27 H	0.175610	0.171038	0.171560	0.179268
28 H	0.232790	0.224026	0.200591	0.225306
29 H	0.434925	0.448122	0.436129	0.438125
30 H	0.372419	0.378293	0.359405	0.364198
31 H	0.237548	0.232371	0.227332	0.223719
32 H	0.180026	0.189784	0.179248	0.192861
33 H	0.196254	0.196549	0.188023	0.187208

Tables 11a and 11b.

The backbone conformations of the model tetramers as found by a vacuum stand alone random sampling search/minimization procedure.

Table 11a. The cyclic tetramers containing Glu at position 4 (*i+3*)

position *i*	φ_{i+1}	ψ_{i+1}	φ_{i+2}	ψ_{i+2}
Dap	-68.5	99.9	137.9	-40.9
Dab	-41.1	124.2	105.2	-23.0
Orn	-49.0	117.1	82.2	-65.4
Lys	-42.8	122.0	72.7	8.8

Table 11b. The cyclic tetramers containing Asp at position 4 (*i+3*)

position *i*	φ_{i+1}	ψ_{i+1}	φ_{i+2}	ψ_{i+2}
Dap	-65.2	-7.8	-136.8	33.2
Dab	-63.6	-23.2	-95.0	63.9
Orn	-51.4	123.5	85.3	-83.3
Lys	-66.1	117.8	82.0	-73.9

Table 12

Some distances for a type II β-turn conformation of the residue 7 to residue 10 region of the α-Factor are given. The distances are measured from a given Nitrogen atom (for instance in the first column, the N atom is the proline Nitrogen) and different carbon atoms. All the distances are in Å.

To understand to which carbon atom the distance is measured, view the enclosed figure that contains a key. Example: the first distance in the table (5.770Å in the top left corner) is between the Pro^{8} Nitrogen (N73) and carbon # 60 (the β carbon of Leu^{6}).

Carbon Atom	Pro N (N73)	Gly N (N80)	Glu N (N84)	Lys N (N64)
60	5.770	6.842	6.182	3.291
61	4.815	5.460	5.128	3.268
62	5.874	5.755	5.492	4.740
63	3.992	5.219	5.593	3.419
65	2.483	5.106	4.930	1.454
66	1.357	3.605	3.893	2.356
68	3.257	5.725	5.276	2.467
69	4.555	7.250	6.705	3.148
70	5.690	8.025	7.002	3.291
71	6.763	9.332	8.392	4.206
74	1.462	2.432	4.350	4.764
75	2.528	1.345	3.307	5.233
77	2.360	3.585	5.736	5.949
78	2.358	4.826	6.644	5.669
79	1.462	4.745	5.974	4.295
81	4.749	1.460	2.432	6.477
82	5.178	2.492	1.343	6.062
85	5.554	4.124	1.456	4.856
86	6.496	4.761	2.410	5.546
88	5.136	4.750	2.429	3.611
89	6.469	6.270	3.822	4.332
90	6.552	7.091	4.936	3.740

Annex to Table 12.
The critical region of the α-factor with the atom labels used in Table 12.

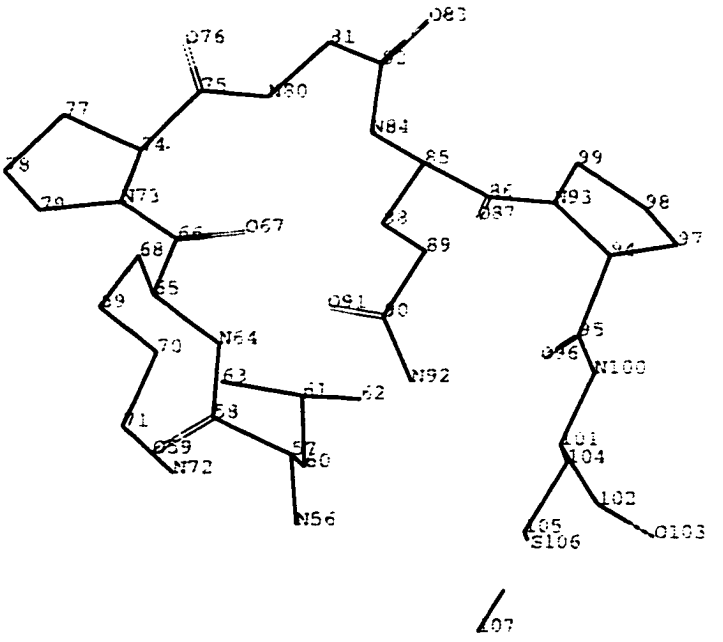


Table 13. Comparison between the experimental and calculated interproton distances for $Tetra_{42}$.[a,b]

Protons	ROESY-distance (Å)[c]	MD in vacuum average (Å)[d]	MD in DMSO average (Å)[e]
$Pro^{\alpha}CH$-GlyNH	2.20	2.23 (0.05)	2.22 (0.06)
GlyNH-$Gly^{\alpha}CH^{proR}$	2.80	2.64 (0.06)	2.65 (0.07)
GlyNH-$Gly^{\alpha}CH^{proS}$	3.10	3.02 (0.06)	3.02 (0.07)
GluNH-$Gly^{\alpha}CH^{proR}$	2.80	2.81 (0.12)	2.78 (0.07)
GluNH-$Gly^{\alpha}CH^{proS}$	2.90	2.81 (0.08)	2.79 (0.10)
GlyNH-GluNH	no cross peak	4.68 (0.16)	4.74 (0.18)
GluNH-$Glu^{\alpha}CH$	2.80	2.86 (0.08)	2.84 (0.07)
$Lys^{\alpha}CH$-$Lys^{\beta}CH1$	2.50	2.44 (0.19)	2.50 (0.14)
$Lys^{\alpha}CH$-$Lys^{\beta}CH2$	2.62	2.98 (0.06)	3.03 (0.09)
$Lys^{\alpha}CH$-$Lys^{\alpha}NH$	2.90	2.90 (0.13)	2.96 (0.09)
$Lys^{\alpha}CH$-$Pro^{\delta}CH1$	2.18	2.40 (0.21)	2.29 (0.20)
$Lys^{\alpha}CH$-$Pro^{\delta}CH2$	2.23	2.46 (0.16)	2.37 (0.17)

[a] Diastereotopic assignments are given as superscripts.
[b] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{42}$.
[c] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{42}$ in DMSO at 300 K.
[d] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{42}$.
[e] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{42}$.

Table 14. Comparison between the experimental and calculated interproton distances for $Tetra_{32}$.[a,b]

Protons	ROESY-distance (Å)[c]	MD in vacuum average (Å)[d]	MD in DMSO average (Å)[e]
Pro$^{\alpha}$CH-GlyNH	2.20	2.31 (0.07)	2.31 (0.07)
GlyNH-Gly$^{\alpha}$CH^{proR}	2.70	2.53 (0.06)	2.52 (0.06)
GlyNH-Gly$^{\alpha}$CH^{proS}	2.90	2.99 (0.07)	2.98 (0.08)
GluNH-Gly$^{\alpha}$CH^{proR}	3.10	3.19 (0.05)	3.19 (0.05)
GlyNH-GluNH	3.40	3.48 (0.10)	3.49 (0.10)
GluNH-Glu$^{\alpha}$CH	2.80	2.88 (0.07)	2.86 (0.08)
GluNH-Glu$^{\beta}$CH1	2.90	2.77 (0.08)	2.74 (0.08)
Glu$^{\alpha}$CH-Glu$^{\beta}$CH1	2.70	2.90 (0.05)	2.91 (0.05)
Glu$^{\alpha}$CH-Glu$^{\gamma}$CH1	2.54	2.46 (0.10)	2.43 (0.09)
Glu$^{\alpha}$CH-Glu$^{\gamma}$CH2	3.00	3.12 (0.10)	3.08 (0.12)
Orn$^{\delta}$NH-Glu$^{\beta}$CH1	2.90	2.81 (0.10)	2.83 (0.10)
Orn$^{\alpha}$CH-Orn$^{\beta}$CH1	3.20	2.96 (0.07)	2.94 (0.07)
Orn$^{\alpha}$CH-Orn$^{\gamma}$CH1	2.70	2.68 (0.05)	2.68 (0.09)
Orn$^{\alpha}$CH-Orn$^{\gamma}$CH2	2.90	3.02 (0.08)	3.02 (0.08)
Orn$^{\alpha}$CH-Pro$^{\delta}$CH1	2.20	2.27 (0.09)	2.27 (0.09)
Orn$^{\alpha}$CH-Pro$^{\delta}$CH2	2.40	2.49 (0.09)	2.47 (0.10)
Orn$^{\alpha}$CH-Orn$^{\alpha}$NH	2.90	2.66 (0.05)	2.64 (0.06)

[a] Diastereotopic assignments are given as superscripts.
[b] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{32}$.
[c] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{32}$ in DMSO at 300 K.
[d] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{32}$.
[e] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{32}$.

Table 15. Comparison between the experimental and calculated interproton distances for $Tetra_{22}$.[a,b,c]

Protons	ROESY-distance (Å)[d]	MD in vacuum average (Å)[e]	MD in DMSO average (Å)[f]
Pro$^{\alpha}$CH-GlyNH	2.19	2.19 (0.08)	2.18 (0.08)
GlyNH-Gly$^{\alpha}$CH^{proR}	2.44	2.22 (0.08)	2.20 (0.08)
GlyNH-Gly$^{\alpha}$CH^{proS}	3.31	2.96 (0.07)	2.99 (0.06)
GluNH-Gly$^{\alpha}$CH^{proR}	3.06	3.24 (0.06)	3.25 (0.07)
GluNH-Gly$^{\alpha}$CH^{proS}	4.26	3.95 (0.04)	3.90 (0.05)
GlyNH-GluNH	2.69	2.76 (0.10)	2.77 (0.10)
GluNH-Glu$^{\alpha}$CH	2.97	2.99 (0.05)	3.00 (0.06)
GluNH-Glu$^{\gamma}$CH2	3.49	3.34 (0.07)	3.35 (0.08)
Glu$^{\alpha}$CH-Glu$^{\gamma}$CH1	2.38	2.53 (0.07)	2.53 (0.07)
Glu$^{\alpha}$CH-Glu$^{\gamma}$CH2	2.78	2.68 (0.08)	2.78 (0.10)
Dab$^{\gamma}$NH-Dab$^{\gamma}$CH1	2.54	2.47 (0.08)	2.46 (0.08)
Dab$^{\gamma}$NH-Dab$^{\beta}$CH1	2.28	2.29 (0.07)	2.31 (0.08)
Dab$^{\alpha}$CH-Dab$^{\beta}$CH2	2.62	2.88 (0.04)	2.88 (0.04)
Dab$^{\alpha}$CH-Dab$^{\alpha}$NH	3.26	3.03 (0.04)	3.57 (0.12)
Dab$^{\alpha}$CH-Pro$^{\delta}$CH1	2.18	2.19 (0.13)	2.25 (0.14)
Dab$^{\alpha}$CH-Pro$^{\delta}$CH2	2.23	2.40 (0.16)	2.37 (0.07)

[a] Major conformer

[b] Diastereotopic assignments are given as superscripts.

[c] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{12}$.

[d] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{12}$ in DMSO at 300 K.

[e] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{12}$.

[f] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{12}$.

Table 16. Comparison between the experimental and calculated interproton distances for $Tetra_{12}$.[a,b,c]

Protons	ROESY-distance (Å)[d]	MD in vacuum average (Å)[e]	MD in DMSO average (Å)[f]
Pro$^{\alpha}$CH-GlyNH	2.65	2.54 (0.04)	2.54 (0.05)
GlyNH-Gly$^{\alpha}$CH^{proR}	2.72	2.56 (0.04)	2.57 (0.05)
GlyNH-Gly$^{\alpha}$CH^{proS}	2.78	2.87 (0.04)	2.86 (0.05)
GluNH-Gly$^{\alpha}$CH^{proR}	3.17	3.23 (0.05)	3.18 (0.07)
GluNH-Gly$^{\alpha}$CH^{proS}	3.28	3.26 (0.06)	3.29 (0.07)
GlyNH-GluNH	2.70	2.69 (0.07)	2.69 (0.08)
GluNH-Glu$^{\alpha}$CH	2.91	2.87 (0.05)	2.86 (0.05)
GluNH-Glu$^{\gamma}$CH2	2.76	2.82 (0.06)	2.97 (0.08)
Glu$^{\alpha}$CH-Glu$^{\beta}$CH1	2.39	2.70 (0.04)	2.73 (0.05)
Glu$^{\alpha}$CH-Glu$^{\beta}$CH2	3.41	3.13 (0.04)	3.11 (0.06)
Glu$^{\alpha}$CH-Glu$^{\gamma}$CH1	2.57	2.71 (0.04)	2.69 (0.06)
Glu$^{\alpha}$CH-Glu$^{\gamma}$CH2	2.94	2.73 (0.04)	2.75 (0.06)
Dpr$^{\beta}$NH-Glu$^{\gamma}$CH1	2.73	2.80 (0.05)	2.86 (0.12)
Dpr$^{\beta}$NH-Glu$^{\gamma}$CH2	3.14	3.11 (0.04)	3.20 (0.12)
Dpr$^{\beta}$NH-Dpr$^{\beta}$CH^{proR}	2.93	2.77 (0.04)	2.80 (0.05)
Dpr$^{\alpha}$CH-Dpr$^{\beta}$CH^{proS}	2.38	2.40 (0.06)	2.37 (0.07)

[a] Major conformer

[b] Diastereotopic assignments are given as superscripts.

[c] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{12}$.

[d] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{12}$ in DMSO at 300 K.

[e] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{12}$.

[f] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{12}$.

Table 17. Comparison between the experimental and calculated interproton distances for $Tetra_{41}$.[a,b]

Protons	ROESY-distance (Å)[c]	MD in vacuum average (Å)[d]	MD in DMSO average (Å)[e]
Pro$^{\alpha}$CH-GlyNH	2.29	2.29 (0.10)	2.26 (0.05)
GlyNH-Gly$^{\alpha}$CH^{proR}	2.35	2.26 (0.07)	2.28 (0.06)
GlyNH-Gly$^{\alpha}$CH^{proS}	2.45	2.63 (0.06)	2.61 (0.05)
AspNH-Gly$^{\alpha}$CH^{proR}	2.79	2.96 (0.09)	2.91 (0.07)
AspNH-Gly$^{\alpha}$CH^{proS}	2.82	3.00 (0.06)	3.03 (0.05)
GlyNH-AspNH	2.80	2.79 (0.10)	2.83 (0.09)
AspNH-Asp$^{\alpha}$CH	2.72	2.81 (0.16)	2.67 (0.08)
Lys$^{\varepsilon}$NH-Lys$^{\gamma}$CH1	2.57	2.42 (0.21)	2.62 (0.12)
Lys$^{\varepsilon}$NH-Lys$^{\gamma}$CH2	2.89	3.17 (0.06)	3.13 (0.06)
Lys$^{\alpha}$CH-Lys$^{\gamma}$CH1	2.82	2.81 (0.16)	2.91 (0.16)
Lys$^{\alpha}$CH-Lys$^{\gamma}$CH2	2.10	2.44 (0.09)	2.43 (0.13)
Lys$^{\alpha}$CH-Pro$^{\delta}$CH1	2.22	2.46 (0.20)	2.35 (0.16)
Lys$^{\alpha}$CH-Pro$^{\delta}$CH2	2.19	2.42 (0.21)	2.31 (0.18)

[a] Diastereotopic assignments are given as superscripts.
[b] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{41}$.
[c] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{41}$ in DMSO at 300 K.
[d] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{41}$.
[e] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{41}$.

Table 18. Comparison between the experimental and calculated interproton distances for $Tetra_{31}$.[a,b]

Protons	ROESY-distance (Å)[c]	MD in vacuum average (Å)[d]	MD in DMSO average (Å)[e]
$Pro^{\alpha}CH$-GlyNH	2.41	2.37 (0.07)	2.32 (0.08)
GlyNH-$Gly^{\alpha}CH^{proR}$	2.56	2.33 (0.09)	2.29 (0.11)
GlyNH-$Gly^{\alpha}CH^{proS}$	2.62	2.94 (0.06)	2.93 (0.07)
AspNH-$Gly^{\alpha}CH^{proS}$	2.62	2.67 (0.16)	2.69 (0.19)
AspNH-$Gly^{\alpha}CH^{proR}$	3.06	3.50 (0.09)	3.54 (0.09)
GlyNH-AspNH	3.30	3.36 (0.07)	3.42 (0.07)
AspNH-$Asp^{\alpha}CH$	2.47	2.60 (0.05)	2.63 (0.06)
AspNH-$Asp^{\beta}CH^{proR}$	2.61	2.43 (0.05)	2.44 (0.06)
AspNH-$Asp^{\beta}CH^{proS}$	3.20	3.41 (0.05)	3.43 (0.05)
$Orn^{\delta}NH$-$Asp^{\alpha}CH$	3.47	3.59 (0.07)	3.57 (0.08)
$Orn^{\delta}NH$-$Asp^{\beta}CH^{proR}$	2.43	2.40 (0.09)	2.44 (0.10)
$Orn^{\delta}NH$-$Asp^{\beta}CH^{proS}$	2.92	3.22 (0.06)	3.21 (0.05)
$Orn^{\delta}NH$-$Orn^{\delta}CH1$	2.56	2.47 (0.08)	2.45 (0.07)
$Orn^{\delta}NH$-$Orn^{\delta}CH2$	2.79	2.94 (0.07)	2.92 (0.06)
$Orn^{\delta}NH$-$Orn^{\gamma}CH1$	3.65	3.52 (0.08)	3.54 (0.07)
$Orn^{\delta}NH$-$Orn^{\gamma}CH2$	3.90	4.04 (0.05)	4.06 (0.05)
$Orn^{\alpha}CH$-$Orn^{\gamma}CH1$	2.78	3.13 (0.04)	3.16 (0.04)
$Orn^{\alpha}CH$-$Orn^{\gamma}CH2$	2.28	2.17 (0.06)	2.18 (0.07)
$Orn^{\alpha}CH$-$Orn^{\delta}CH2$	3.13	3.26 (0.06)	3.27 (0.06)
$Orn^{\alpha}CH$-$Orn^{\alpha}NH$	2.99	2.93 (0.08)	2.99 (0.10)

[a] Diastereotopic assignments are given as superscripts.
[b] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{31}$.
[c] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{31}$ in DMSO at 300 K.
[d] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{31}$.
[e] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{31}$.

Table 19. Comparison between the experimental and calculated interproton distances for $Tetra_{21}$.[a,b,c]

Protons	ROESY-distance (Å)[d]	MD in vacuum average (Å)[e]	MD in DMSO average (Å)[f]
Pro$^{\alpha}$CH-GlyNH	2.34	2.23 (0.06)	2.24 (0.06)
GlyNH-Gly$^{\alpha}$CH^{proS}	2.52	2.88 (0.08)	2.81 (0.08)
AspNH-Gly$^{\alpha}$CH^{proS}	3.53	3.26 (0.13)	3.52 (0.09)
AspNH-Gly$^{\alpha}$CH^{proR}	3.10	3.23 (0.18)	2.96 (0.18)
GlyNH-AspNH	2.85	2.87 (0.09)	2.84 (0.09)
AspNH-Dab$^{\gamma}$NH	3.08	3.15 (0.08)	3.08 (0.10)
Asp$^{\alpha}$CH-Dab$^{\gamma}$NH	3.07	3.35 (0.04)	3.35 (0.04)
Dab$^{\gamma}$NH-Dab$^{\gamma}$CH1	2.57	2.41 (0.04)	2.43 (0.04)
Dab$^{\gamma}$NH-Dab$^{\gamma}$CH2	2.62	2.75 (0.05)	2.74 (0.05)
Dab$^{\gamma}$NH-Dab$^{\beta}$CH1	2.95	3.14 (0.05)	3.16 (0.04)
Dab$^{\gamma}$NH-Dab$^{\beta}$CH2	3.19	3.48 (0.05)	3.42 (0.05)
Dab$^{\alpha}$CH-Dab$^{\beta}$CH1	2.27	2.21 (0.07)	2.15 (0.05)
Dab$^{\alpha}$CH-Dab$^{\beta}$CH2	2.32	2.48 (0.18)	2.51 (0.04)
Dab$^{\alpha}$CH-Dab$^{\gamma}$CH1	4.14	3.92 (0.05)	3.90 (0.04)
Dab$^{\alpha}$CH-Dab$^{\gamma}$CH2	2.91	3.21 (0.04)	3.19 (0.04)
Dab$^{\alpha}$CH-Pro$^{\delta}$CH1	2.25	2.48 (0.09)	2.39 (0.09)
Dab$^{\alpha}$CH-Pro$^{\delta}$CH2	2.40	2.52 (0.21)	2.79 (0.13)
Dab$^{\alpha}$CH-Dab$^{\alpha}$NH	2.89	2.84 (0.06)	2.86 (0.06)
Dab$^{\alpha}$NH-Dab$^{\beta}$CH1	2.40	2.58 (0.14)	2.41 (0.15)
Dab$^{\alpha}$NH-Dab$^{\beta}$CH2	3.23	3.56 (0.09)	3.59 (0.09)

[a] Major conformer

[b] Diastereotopic assignments are given as superscripts.

[c] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{21}$.

[d] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{21}$ in DMSO at 300 K.

[e] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{21}$.

[f] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{21}$.

Table 20. Comparison between the experimental and calculated interproton distances for $Tetra_{11}$.[a,b,c]

Protons	ROESY-distance (Å)[d]	MD in vacuum average (Å)[e]	MD in DMSO average (Å)[f]
Pro$^{\alpha}$CH-GlyNH	2.48	2.37 (0.06)	2.34 (0.07)
AspNH-Gly$^{\alpha}$CH^{proR}	3.10	3.30 (0.05)	3.32 (0.07)
AspNH-Gly$^{\alpha}$CH^{proS}	3.47	3.71 (0.05)	3.71 (0.06)
GlyNH-AspNH	2.84	2.81 (0.10)	2.88 (0.11)
AspNH-Asp$^{\alpha}$CH	2.88	2.95 (0.04)	2.96 (0.05)
AspNH-Asp$^{\beta}$CH^{proR}	3.18	3.02 (0.05)	3.03 (0.08)
AspNH-Asp$^{\beta}$CH^{proS}	3.76	3.61 (0.05)	3.65 (0.07)
Asp$^{\alpha}$CH-Asp$^{\beta}$CH^{proR}	2.32	2.34 (0.05)	2.33 (0.08)
Asp$^{\alpha}$CH-Asp$^{\beta}$CH^{proS}	2.39	2.53 (0.05)	2.51 (0.06)
Dpr$^{\beta}$NH-Asp$^{\beta}$CH^{proR}	2.50	2.36 (0.06)	2.33 (0.06)
Dpr$^{\beta}$NH-Asp$^{\beta}$CH^{proS}	2.60	2.70 (0.05)	2.69 (0.07)
Dpr$^{\beta}$NH-Dpr$^{\beta}$CH^{proS}	2.62	2.50 (0.05)	2.50 (0.06)
Dpr$^{\beta}$NH-Dpr$^{\beta}$CH^{proR}	2.70	2.90 (0.04)	2.93 (0.05)
Dpr$^{\beta}$NH-Dpr$^{\alpha}$CH	2.79	2.94 (0.07)	2.85 (0.10)
Dpr$^{\alpha}$CH-Pro$^{\delta}$CH1	2.21	2.32 (0.16)	2.27 (0.16)
Dpr$^{\alpha}$CH-Pro$^{\delta}$CH2	2.24	2.39 (0.13)	2.40 (0.13)

[a] Major conformer

[b] Diastereotopic assignments are given as superscripts.

[c] *rmsd* value is given in parenthesis for each interproton distance obtained from modeling for $Tetra_{11}$.

[d] Experimental data are extracted from a ROESY spectrum (100 ms) for $Tetra_{11}$ in DMSO at 300 K.

[e] Calculated interproton distances are computed from 60 ps MD trajectories in vacuum for $Tetra_{11}$.

[f] Calculated interproton distances are computed from 60 ps MD trajectories in DMSO for $Tetra_{11}$.

Table 21. Temperature coefficients of *NH* protons (defined as $-\Delta\delta/\Delta T$ and measured in ppb/K) of cyclic peptides **Tetra$_{42}$** to **Tetra$_{11}$** in DMSO-d_6.

Amide Proton	αNH(*i*)	ωNH(*i*)	αNH(*i+2*)	αNH(*i+3*)
Tetra$_{42}$	7.02	5.89	3.14	4.80
Tetra$_{32}$	7.44	2.28	6.76	0.90
Tetra$_{22}$	7.44	5.71	7.57	-1.80
Tetra$_{12}$	4.40	3.99	0.82	0.22
Tetra$_{41}$[a]	5.91	3.33	5.63	0.72
Tetra$_{31}$[a]	5.30	5.40	3.45	2.65
Tetra$_{21}$	4.54	3.72	4.04	-1.40
Tetra$_{11}$	4.15	3.10	3.40	-0.90

[a] *major conformer*

Table 22. Conformational characteristics of cyclic tetrapeptides from NMR and modeling[a]

	Glyα*NH* - Zzzα*NH*		Proα*CH* - Glyα*NH*		temp. coef.	H-bonding	Dihedral angles (modeling)[c]			
Peptides & Ideal turns	ROESY distance (Å)	Modeling distance (Å)	ROESY distance (Å)	Modeling distance (Å)	$\Delta_{zz\alpha NH}$ (ppb/K)	H-bond distance modeling (Å)	ϕ_2 (°)	ψ_2 (°)	ϕ_3 (°)	ψ_3 (°)
$Tetra_{42}$	---	4.74 (0.18)	2.20	2.22 (0.06)	4.80	---	-77 (9)	82 (10)	108 (21)	178 (9)
$Tetra_{32}$	3.40	3.49 (0.10)	2.20	2.31 (0.07)	0.90	2.06[c]	-68 (11)	160 (7)	81 (8)	-68 (8)
$Tetra_{22}$b	2.69	2.77 (0.10)	2.19	2.18 (0.08)	-1.80	2.47[d]	-71 (7)	126 (5)	87 (9)	5 (3)
$Tetra_{12}$b	2.70	2.69 (0.08)	2.65	2.54 (0.05)	0.22	2.46[c,d]	-65 (22)	128 (29)	68 (10)	19 (9)
$Tetra_{41}$	2.80	2.83 (0.09)	2.29	2.26 (0.05)	0.72	2.75[d]	-77 (12)	142 (18)	53 (9)	19 (10)
$Tetra_{31}$	3.30	3.42 (0.07)	2.41	2.32 (0.08)	2.65	1.95[c]	-46 (10)	99 (9)	80 (9)	-52 (9)
$Tetra_{21}$b	2.85	2.84 (0.09)	2.34	2.24 (0.06)	-1.40	2.50[c]	-95 (10)	133 (10)	67 (9)	30 (12)
$Tetra_{11}$b	2.84	2.88 (0.11)	2.48	2.34 (0.07)	-0.90	2.64[d]	-73 (8)	112 (10)	72 (10)	21 (10)
Type II-β turn	2.40		2.20		< 2.5	2.5	-60	120	80	0
Type I-β turn	2.40		3.40		< 2.5	2.5	-60	-30	-90	0
γ-turn	3.70		2.30		< 2.5	2.5			70 to 85	-60 to -70
γ'-turn					< 2.5	2.5	-60 to -70	70 to 85		

a *rmsd* value is given in parenthesis for each interproton distance and torsion angle obtained from modeling.
b Major conformer
c 1← 3 H-bonding
d 1← 4 H-bonding
e Dihedral angels extracted from 60 ps MD trajectories in DMSO.

Table 23. Comparison between the ROE distance values and the simulated average distance values.

	root mean-square deviation, (Å/distance)	
tetrapeptide	in vacuum	in DMSO
$Tetra_{42}$	0.050	0.047
$Tetra_{32}$	0.031	0.033
$Tetra_{22}$[a]	0.024	0.026
$Tetra_{12}$[a]	0.033	0.037
$Tetra_{41}$	0.052	0.043
$Tetra_{31}$	0.029	0.029
$Tetra_{21}$[a]	0.036	0.039
$Tetra_{11}$[a]	0.029	0.027

[a] *major conformer*

Table 24. $^3J_{NH\text{-}C\alpha H}$ coupling constants (Hz), backbone dihedral angles (°) calculated from coupling constants and dihedral angles (°) extracted from 60 ps MD trajectories in DMSO for $Tetra_{42}$ to $Tetra_{11}$

Peptide	Residue	$^3J_{HNC^\alpha H}$[a]	ϕ calc[b]	ϕ model[c]
$Tetra_{42}$	Gly (*NH*, $C^\alpha H$)	11	54, 135, -54, -135	108 (21)
	Lys ($N^\varepsilon H$, $C^\varepsilon H$)	10.4	50, 139, -50, -139	-32 (7)
	Lys (*NH*, $C^\alpha H$)	7.9	43, 77, -87, -153	-106 (12)
	Glu (*NH*, $C^\alpha H$)	9.2	-95, -145	-153 (19)
$Tetra_{32}$	Gly (*NH*, $C^\alpha H$)	11.4	56, 132, -56, -132	81 (8)
	Orn ($N^\delta H$, $C^\delta H$)	10.1	49, 141, -49, -141	101 (10)
	Orn (*NH*, $C^\alpha H$)	6.9	33, 87, -80, -160	92 (14)
	Glu (*NH*, $C^\alpha H$)	9.4	-97, -143	-52 (7)
$Tetra_{22}$[e]	Gly (*NH*, $C^\alpha H$)	10.8	53, 136, -53, -136	87 (9)
	Dab ($N^\gamma H$, $C^\gamma H$)	11.85	59, 129, -59, -129	-89 (10)
	Dab (*NH*, $C^\alpha H$)	7.6	40, 80, -85, -155	-96 (13)
	Glu (*NH*, $C^\alpha H$)	9.9	-101, -139	-106 (9)
$Tetra_{12}$[e]	Gly (*NH*, $C^\alpha H$)	11.9	59, 129, -59, -129	68 (10)
	Dpr ($N^\beta H$, $C^\beta H$)	12.4	63, 125, -63, -125	96 (16)
	Dpr (*NH*, $C^\alpha H$)	NA[d]		
	Glu (*NH*, $C^\alpha H$)	8.6	54, 66, -91, -149	-102 (24)
$Tetra_{41}$	Gly (*NH*, $C^\alpha H$)	11.1	54, 134, -54, -134	53 (9)
	Lys ($N^\varepsilon H$, $C^\varepsilon H$)	11	54, 134, -54, -134	-160 (17)
	Lys (*NH*, $C^\alpha H$)	7.9	42, 78, -87, -153	-74 (24)
	Asp (*NH*, $C^\alpha H$)	8.1	44, 76, -88, -152	-44 (14)
$Tetra_{31}$	Gly (*NH*, $C^\alpha H$)	10.9	53, 136, -53, -136	80 (9)
	Orn ($N^\delta H$, $C^\delta H$)	11.6	57, 131, -57, -131	-86 (11)
	Orn (*NH*, $C^\alpha H$)	7.7	41, 79, -86, -154	60 (8)
	Asp (*NH*, $C^\alpha H$)	7.7	41, 79, -86, -154	-66 (7)
$Tetra_{21}$[e]	Gly (*NH*, $C^\alpha H$)	11.9	59, 129, -59, -129	67 (9)
	Dab ($N^\gamma H$, $C^\gamma H$)	10.8	53, 136, -53, -136	180 (3)
	Dab (*NH*, $C^\alpha H$)	8.4	49, 71, -90, -150	-71 (10)
	Asp (*NH*, $C^\alpha H$)	7.3	37, 83, -83, -157	-136 (15)
$Tetra_{11}$[e]	Gly (*NH*, $C^\alpha H$)	11.5	57, 132, -57, -132	72 (10)
	Dpr ($N^\beta H$, $C^\beta H$)	11.3	56, 133, -56, -133	-76 (6)
	Dpr (*NH*, $C^\alpha H$)	8.3	48, 72, -90, -150	-100 (11)
	Asp (*NH*, $C^\alpha H$)	8.2	46, 74, -88, -152	-95 (9)

[a] Experimental value extracted from 1D spectrum in DMSO-d_6 at 300K.
[b] Possible values for the peptide backbone angles ϕ (°) from Bystrov's Karplus equations; for Glycyl and side chain *NHs* $\Sigma(^3J_{NH\text{-}C^\alpha H}) = -9.4\cos^2\theta - 1.1\cos\theta + 14.9$ and for others $(^3J_{NH\text{-}C^\alpha H}) = 9.4\cos^2\theta - 1.1\cos\theta + 0.4$.
[c] Average dihedral angles from 60 ps constrained MD in DMSO & *rmsd* values are given in parentheses.
[d] No experimental value due to signal overlap. [e] Major conformer

Table 25 ^{1}H Assignments for the cyclic [Dab^7, Glu^{10}, Nle^{12}]α-factor (C22) in DMSO-d_6 (25 °C)

	αNH (ppm)	αCH (ppm)	βCH (ppm)	γCH (ppm)	δCH (ppm)	εCH (ppm)	γNH (ppm)	others (ppm)
Trp^1		4.05	3.14, 2.98					N1 10.92
								C2 7.15
								C4 7.57
								C5 6.81
								C6 7.00
								C7 7.35
His^2	8.88	4.67	3.06					C2 8.68
								C4 7.29
Trp^3	8.27	4.62	3.19, 2.96					N1 10.82
								C2 7.17
								C4 7.66
								C5 6.98
								C6 7.05
								C7 7.32
Leu^4	8.44	4.36	1.49	1.61	0.85			
Gln^5	8.00	4.28	1.88, 1.76	2.10				NH_2 6.76 7.23
Leu^6	7.80	4.26	1.39 1.32	1.56	0.83			
Dab^7	8.41	4.60	1.88 1.66	3.62 2.64			7.23	
Pro^8		4.35	2.12, 1.93	1.84	3.99 3.74			
Gly^9	9.02	3.69 3.59						
Glu^{10}	6.47	4.55	1.46	1.95				
Pro^{11}		4.33	1.98 1.86	1.76	3.52			
Nle^{12}	7.99	4.11	1.61	1.44	1.24	0.85		
Tyr^{13}	7.86	4.33	2.92 2.77					C2,6 7.66
								C3,5 6.97
								OH 9.16

Table 26 ^{1}H Assignments of constrained region for the cyclo$^{7\text{-}10}$ [Xxx^{7}, Glu^{10}, Nle^{12}]α-factor analogs in DMSO-d_6 (25 °C)

	αNH (ppm)	αCH (ppm)	βCH (ppm)	γCH (ppm)	δCH (ppm)	εCH (ppm)	ωNH (ppm)
C42							
Lys^{7}	7.98	4.52	1.71 1.55	1.40	1.31	3.22, 2.89	7.58
Pro^{8}		4.30	2.00	1.85	3.69 3.56		
Gly^{9}	7.77	3.67					
Glu^{10}	7.94	4.55	1.52	2.09, 1.98			
C32							
Orn^{7}	8.23	4.41	1.76 1.50	1.32	2.96		7.04
Pro^{8}		4.32	2.03	1.84	3.80 3.54		
Gly^{9}	7.96	3.73 3.60					
Glu^{10}	7.50	4.62	1.99 1.57	2.15			
C12							
Dpr^{7}	7.87	4.59	3.44 3.59				7.40
Pro^{8}		4.29	2.09 1.90	1.74	3.62		
Gly^{9}	8.15	3.86 3.56					
Glu^{10}	6.69	4.56	1.77	2.35 1.99			

Table 27. Temperature coefficients of *NH* protons (-Δδ/ΔT ppb/K) of cyclic peptides **C42** to **C12** in DMSO-d_6.

Amide Proton	C42	C32	C22	C12
His^2 *αNH*	7.2	4.7	4.5	4.7
Trp^3 *αNH*	6.1	3.9	3.8	4.0
Leu^4 *αNH*	11.4	7.0	6.5	7.3
Gln^5 *αNH*	6.5	4.0	*na*	4.0
Leu^6 *αNH*	6.9	*na*	4.0	6.4
Xxx^7 *αNH*	6.3	6.7	6.0	*na*
Gly^9 *αNH*	3.1	5.9	5.8	3.4
Glu^{10} *αNH*	6.3	0.06	1.78	0.5
Nle^{12} *αNH*	6.2	*na*	*na*	*na*
Tyr^{13} *αNH*	9.6	*na*	5.96	5.8

na not available due to signal overlap
Xxx in **C42, C32, C22, & C12** are Lys, Orn, Dab, & Dpr respectively

Table 28. Conformational characteristics of cyclic lactam containing α-factor analogs from NMR and modeling[a]

	Gly^{9}αNH - Glu^{10}αNH		Pro^{8}αCH - Gly^{9}αNH		temp. coef.	H-bonding	Dihedral angles (modeling)[d]			
Peptides & Ideal turns	NOESY distance (Å)	Modeling distance (Å)	NOESY distance (Å)	Modeling distance (Å)	Glu^{10}α NH (ppb/K)	H-bond modeling (Å)	ϕ_{i+1} (°)	ψ_{i+1} (°)	ϕ_{i+2} (°)	ψ_{i+2} (°)
C42	----	4.59 (0.11)	2.21	2.28 (0.05)	6.30	----	-80 (8)	76 (11)	165 (9)	-164 (12)
C32	2.78	2.55 (0.05)	1.97	2.04 (0.06)	0.06	2.16[b]	-72 (11)	147 (8)	87 (10)	-63 (13)
C22	2.58	2.51 (0.04)	2.07	2.08 (0.09)	1.78	2.47[b,c]	-47 (10)	112 (11)	66 (11)	25 (9)
C12	2.73	2.59 (0.05)	2.40	2.46 (0.06)	0.50	2.44[b,c]	-51 (11)	135 (9)	63 (10)	22 (12)
Type II-β turn	2.40		2.20		< 2.5	2.5	-60	120	80	0
Type I-β turn	2.40		3.40		< 2.5	2.5	-60	-30	-90	0
γ-turn	3.70		2.30		< 2.5	2.5			70 to 85	-60 to -70
γ'-turn					< 2.5	2.5	-60 to -70	70 to 85		

a *rmsd* value is given in parenthesis for each interproton distance and torsion angle obtained from modeling.
b 1← 3 H-bonding
c 1← 4 H-bonding
d Average dihedral angels extracted from 200 ps MD trajectories in DMSO.

Table 29. Comparison between the experimental and calculated interproton distances for **C32**

Protons		NOESY-distance (Å)[a]	MD in DMSO average (Å)[b]
$Trp^1C^{\alpha}H$	His^2NH	2.02	2.09 (0.09)
$Trp^1C^{\beta}H_1$	His^2NH	2.38	2.53 (0.11)
$Trp^1C^{\beta}H_2$	His^2NH	2.21	2.20 (0.06)
Trp^1C_4H	His^2NH	3.36	3.63 (0.14)
His^2NH	$His^2C^{\alpha}H$	2.64	2.59 (0.05)
His^2NH	Trp^3NH	3.01	2.97 (0.06)
$His^2C^{\alpha}H$	Trp^3NH	1.96	2.12 (0.05)
Trp^3NH	Leu^4NH	2.50	2.72 (0.06)
$Trp^3C^{\alpha}H$	Leu^4NH	1.98	2.03 (0.09)
$Trp^3C^{\beta}H_1$	Leu^4NH	2.60	2.76 (0.04)
$Trp^3C^{\beta}H_2$	Leu^4NH	2.79	2.89 (0.07)
$Leu^4C^{\alpha}H$	Gln^5NH	1.83	2.01 (0.09)
$Leu^4C^{\alpha}H$	Leu^4NH	2.60	2.70 (0.04)
Leu^4NH	Gln^5NH	2.64	2.47 (0.06)
Gln^5NH	Leu^6NH	2.14	2.23 (0.09)
$Gln^5C^{\beta}H_1$	Leu^6NH	2.62	2.77 (0.05)
$Gln^5C^{\gamma}H^c$	Leu^6NH	3.60	3.83 (0.11)
Leu^6NH	Orn^7NH	2.76	2.62 (0.06)
$Leu^6C^{\alpha}H$	Orn^7NH	1.98	1.97 (0.05)
Orn^7NH	$Orn^7C^{\alpha}H$	2.43	2.40 (0.08)
Orn^7NH	$Orn^7C^{\beta}H_1$	2.44	2.29 (0.06)
Orn^7NH	$Orn^7C^{\beta}H_2$	2.68	2.56 (0.12)
Orn^7NH	$Orn^7C^{\gamma}H^c$	2.54	2.62 (0.07)
$Orn^7C^{\alpha}H$	$Orn^7C^{\beta}H_1$	2.11	2.10 (0.10)
$Orn^7C^{\alpha}H$	$Orn^7C^{\beta}H_2$	2.35	2.43 (0.04)
$Orn^7C^{\alpha}H$	$Pro^8C^{\delta}H_1$	2.12	2.20 (0.05)
$Orn^7C^{\alpha}H$	$Pro^8C^{\delta}H_2$	2.18	2.26 (0.08)
$Orn^7N^{\delta}H$	$Orn^7C^{\beta}H_1$	3.11	3.23 (0.08)
$Orn^7N^{\delta}H$	$Orn^7C^{\gamma}H^c$	2.70	2.59 (0.07)
$Orn^7N^{\delta}H$	$Orn^7C^{\delta}H^c$	2.32	2.40 (0.06)
$Orn^7N^{\delta}H$	$Glu^{10}C^{\gamma}H^c$	2.30	2.18 (0.11)

Continuation of Table 29

$Orn^{7}N^{\delta}H$	$Glu^{10}C^{\beta}H_1$	2.17	2.26 (0.04)
$Orn^{7}N^{\delta}H$	$Glu^{10}C^{\beta}H_2$	2.44	2.32 (0.08)
$Pro^{8}C^{\alpha}H$	$Gly^{9}NH$	1.97	2.04 (0.06)
$Gly^{9}NH$	$Glu^{10}NH$	2.78	2.55 (0.05)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_1$	2.54	2.47 (0.12)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_2$	2.75	2.84 (0.13)
$Gly^{9}C^{\alpha}H_1$	$Glu^{10}NH$	2.44	2.33 (0.07)
$Gly^{9}C^{\alpha}H_2$	$Glu^{10}NH$	2.33	2.36 (0.04)
$Glu^{10}C^{\alpha}H$	$Glu^{10}NH$	2.80	2.92 (0.17)
$Glu^{10}NH$	$Glu^{10}C^{\beta}H_1$	2.63	2.51 (0.09)
$Glu^{10}NH$	$Glu^{10}C^{\beta}H_2$	2.66	2.74 (0.04)
$Glu^{10}NH$	$Glu^{10}C^{\gamma}H^{c}$	3.13	3.30 (0.16)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\gamma}H^{c}$	2.71	2.70 (0.05)
$Nle^{12}C^{\alpha}H$	$Nle^{12}C^{\beta}H_1$	3.00	2.88 (0.07)
$Nle^{12}C^{\alpha}H$	$Nle^{12}C^{\beta}H_2$	3.48	3.36 (0.11)
$Tyr^{13}C^{\alpha}H$	$Tyr^{13}C^{\beta}H_1$	2.54	2.65 (0.06)
$Tyr^{13}C^{\alpha}H$	$Tyr^{13}C^{\beta}H_2$	2.84	2.77 (0.07)

[a] Experimental data are extracted from a NOESY spectrum (250 ms) in DMSO-d_6.

[b] Calculated interproton distances are computed from 200 ps MD trajectories in DMSO &

rmsd values of the two thousand structures obtained during dynamics are given in parentheses.

[c] Distance involving a pseudoatom.

Table 30. Comparison between the experimental and calculated interproton distances for **C42**

Protons		NOESY-distance (Å)[a]	MD in DMSO average (Å)[b]
$Trp^1C^{\alpha}H$	His^2NH	2.09	2.06 (0.04)
$Trp^1C^{\beta}H_1$	His^2NH	2.47	2.44 (0.08)
$Trp^1C^{\beta}H_2$	His^2NH	2.30	2.39 (0.09)
Trp^1C_4H	His^2NH	3.22	3.34 (0.12)
His^2NH	$His^2C^{\alpha}H$	2.83	2.79 (0.05)
His^2NH	Trp^3NH	2.44	2.47 (0.04)
$His^2C^{\alpha}H$	Trp^3NH	2.02	2.08 (0.06)
$His^2C^{\alpha}H$	Trp^3C_4H	3.06	2.99 (0.08)
Trp^3NH	Leu^4NH	2.67	2.63 (0.05)
$Trp^3C^{\alpha}H$	Leu^4NH	1.99	2.02 (0.04)
$Trp^3C^{\beta}H_1$	Leu^4NH	2.71	2.76 (0.09)
$Trp^3C^{\beta}H_1$	Leu^4NH	2.74	2.83 (0.07)
$Leu^4C^{\alpha}H$	Gln^5NH	2.13	2.28 (0.05)
$Leu^4C^{\alpha}H$	Leu^4NH	2.82	2.79 (0.11)
Leu^4NH	Gln^5NH	2.69	2.81 (0.10)
Gln^5NH	Leu^6NH	2.37	2.36 (0.08)
$Gln^5C^{\beta}H_1$	Leu^6NH	2.47	2.40 (0.05)
$Gln^5C^{\beta}H_2$	Leu^6NH	2.61	2.70 (0.09)
Leu^6NH	Lys^7NH	2.74	2.68 (0.07)
$Leu^6C^{\alpha}H$	Lys^7NH	2.20	2.11 (0.07)
Lys^7NH	$Lys^7C^{\alpha}H$	2.76	2.84 (0.09)
Lys^7NH	$Lys^7C^{\beta}H_1$	2.42	2.31 (0.08)
Lys^7NH	$Lys^7C^{\beta}H_2$	2.88	2.69 (0.15)
$Lys^7C^{\alpha}H$	$Lys^7C^{\beta}H_1$	2.53	2.49 (0.09)
$Lys^7C^{\alpha}H$	$Lys^7C^{\beta}H_2$	2.58	2.71 (0.11)
$Lys^7C^{\alpha}H$	$Lys^7C^{\gamma}H^c$	2.81	2.72 (0.08)
$Lys^7C^{\alpha}H$	$Lys^7C^{\delta}H^c$	2.92	3.03 (0.14)
$Lys^7N^{\varepsilon}H$	$Lys^7C^{\gamma}H^c$	2.53	2.68 (0.07)

Continuation of Table 30

$Lys^{7}N^{\varepsilon}H$	$Lys^{7}C^{\delta}H^{c}$	2.56	2.45 (0.04)
$Lys^{7}N^{\varepsilon}H$	$Lys^{7}C^{\varepsilon}H_1$	2.31	2.26 (0.06)
$Lys^{7}N^{\varepsilon}H$	$Glu^{10}C^{\beta}H^{c}$	2.59	2.67 (0.08)
$Lys^{7}N^{\varepsilon}H$	$Glu^{10}C^{\gamma}H_1$	2.17	2.21 (0.04)
$Lys^{7}N^{\varepsilon}H$	$Glu^{10}C^{\gamma}H_2$	2.28	2.34 (0.06)
$Lys^{7}C^{\alpha}H$	$Pro^{8}C^{\delta}H_1$	2.10	2.09 (0.05)
$Lys^{7}C^{\alpha}H$	$Pro^{8}C^{\delta}H_2$	2.26	2.37 (0.06)
$Pro^{8}C^{\alpha}H$	$Gly^{9}NH$	2.21	2.28 (0.05)
$Gly^{9}NH$	$Glu^{10}NH$	no cross peak	4.59 (0.11)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_1$	2.28	2.33 (0.09)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_2$	2.43	2.39 (0.08)
$Gly^{9}C^{\alpha}H_1$	$Glu^{10}NH$	2.39	2.34 (0.04)
$Gly^{9}C^{\alpha}H_2$	$Glu^{10}NH$	2.72	2.78 (0.06)
$Glu^{10}C^{\alpha}H$	$Glu^{10}NH$	2.93	2.86 (0.11)
$Glu^{10}NH$	$Glu^{10}C^{\beta}H^{c}$	2.64	2.57 (0.07)
$Glu^{10}NH$	$Glu^{10}C^{\gamma}H_1$	2.84	2.70 (0.10)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\beta}H^{c}$	2.74	2.86 (0.09)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\gamma}H_1$	2.44	2.58 (0.05)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\gamma}H_2$	2.52	2.58 (0.08)
$Nle^{12}C^{\alpha}H$	$Nle^{12}NH$	2.80	2.72 (0.06)
$Nle^{12}C^{\alpha}H$	$Nle^{12}C^{\beta}H_1$	2.32	2.36 (0.05)
$Nle^{12}C^{\alpha}H$	$Nle^{12}C^{\beta}H_2$	2.59	3.52 (0.10)
$Tyr^{13}C^{\alpha}H$	$Tyr^{13}C^{\beta}H_1$	2.49	2.71 (0.07)

[a] Experimental data are extracted from a NOESY spectrum (250 ms) in DMSO-d_6.
[b] Calculated interproton distances are computed from 200 ps MD trajectories in DMSO & *rmsd* values of the two thousand structures obtained during dynamics are given in parentheses.
[c] Distance involving a pseudoatom.

Table 31. Comparison between the experimental and calculated interproton distances for **C22**

Protons		NOESY-distance (Å)[a]	MD in DMSO average (Å)[b]
$Trp^{1}C^{\alpha}H$	$His^{2}NH$	2.07	2.11 (0.04)
$Trp^{1}C^{\alpha}H$	$Trp^{1}C_{4}H$	3.06	3.19 (0.11)
$Trp^{1}C_{4}H$	$His^{2}NH$	3.98	3.79 (0.08)
$His^{2}NH$	$His^{2}C^{\alpha}H$	2.81	2.64 (0.09)
$His^{2}NH$	$Trp^{3}NH$	3.25	3.36 (0.09)
$His^{2}NH$	$His^{2}C^{\beta}H^{c}$	2.71	2.88 (0.04)
$His^{2}C^{\alpha}H$	$His^{2}C^{\beta}H^{c}$	2.22	2.08 (0.08)
$His^{2}C^{\alpha}H$	$Trp^{3}NH$	2.03	2.14 (0.06)
$Trp^{3}NH$	$Trp^{3}C^{\alpha}H$	2.55	2.46 (0.06)
$Trp^{3}NH$	$His^{2}C^{\beta}H^{c}$	2.77	2.82 (0.05)
$Trp^{3}C^{\alpha}H$	$Trp^{3}C_{4}H$	3.13	3.01 (0.09)
$Trp^{3}NH$	$Leu^{4}NH$	2.91	2.88 (0.06)
$Trp^{3}C^{\alpha}H$	$Leu^{4}NH$	2.05	2.06 (0.08)
$Trp^{3}C^{\alpha}H$	$Leu^{4}C^{\alpha}H$	2.93	2.77 (0.07)
$Leu^{4}C^{\alpha}H$	$Gln^{5}NH$	2.06	2.01 (0.04)
$Leu^{4}C^{\alpha}H$	$Leu^{4}NH$	2.80	2.69 (0.10)
$Leu^{4}NH$	$Gln^{5}NH$	2.90	2.76 (0.07)
$Leu^{4}NH$	$Leu^{4}C^{\gamma}H$	2.98	2.84 (0.11)
$Leu^{4}NH$	$Leu^{4}C^{\beta}H^{c}$	2.51	2.61 (0.06)
$Leu^{4}C^{\alpha}H$	$Leu^{4}C^{\beta}H^{c}$	2.45	2.41 (0.06)
$Gln^{5}NH$	$Gln^{5}C^{\alpha}H$	2.27	2.20 (0.07)
$Gln^{5}NH$	$Gln^{5}C^{\gamma}H^{c}$	3.38	3.32 (0.15)
$Gln^{5}C^{\alpha}H$	$Gln^{5}C^{\gamma}H^{c}$	2.83	2.77 (0.04)
$Gln^{5}C^{\gamma}H^{c}$	$Leu^{6}NH$	3.40	3.53 (0.10)
$Leu^{6}NH$	$Leu^{6}C^{\gamma}H$	2.81	2.70 (0.08)
$Leu^{6}NH$	$Dab^{7}NH$	3.34	3.10 (0.08)
$Leu^{6}C^{\alpha}H$	$Leu^{6}C^{\gamma}H$	3.02	2.97 (0.07)
$Leu^{6}C^{\alpha}H$	$Dab^{7}NH$	2.10	1.98 (0.04)

Continuation of Table 31

$Leu^{6}C^{\alpha}H$	$Dab^{7}C^{\alpha}H$	3.24	3.39 (0.09)
$Dab^{7}NH$	$Dab^{7}C^{\alpha}H$	2.70	2.63 (0.05)
$Dab^{7}N^{\gamma}H$	$Dab^{7}C^{\alpha}H$	3.75	3.61 (0.11)
$Dab^{7}N^{\gamma}H$	$Glu^{10}NH$	3.82	3.90 (0.05)
$Dab^{7}N^{\gamma}H$	$Glu^{10}C^{\alpha}H$	4.17	4.33 (0.11)
$Dab^{7}N^{\gamma}H$	$Glu^{10}C^{\gamma}H^{c}$	2.12	2.19 (0.04)
$Dab^{7}N^{\gamma}H$	$Glu^{10}C^{\beta}H^{c}$	3.24	3.11 (0.08)
$Pro^{8}C^{\alpha}H$	$Gly^{9}NH$	2.07	2.08 (0.09)
$Gly^{9}NH$	$Glu^{10}NH$	2.58	2.51 (0.04)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_1$	2.18	2.29 (0.07)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_2$	2.42	2.56 (0.09)
$Gly^{9}C^{\alpha}H_1$	$Glu^{10}NH$	2.90	2.79 (0.11)
$Gly^{9}C^{\alpha}H_2$	$Glu^{10}NH$	3.14	3.20 (0.14)
$Glu^{10}C^{\alpha}H$	$Glu^{10}NH$	2.72	2.89 (0.12)
$Glu^{10}NH$	$Glu^{10}C^{\beta}H^{c}$	2.57	2.51 (0.09)
$Glu^{10}NH$	$Glu^{10}C^{\gamma}H^{c}$	2.64	2.53 (0.09)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\beta}H^{c}$	3.07	2.98 (0.08)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\gamma}H^{c}$	2.20	2.34 (0.10)
$Glu^{10}C^{\alpha}H$	$Pro^{11}C^{\delta}H^{c}$	2.14	2.23 (0.06)
$Pro^{11}C^{\alpha}H$	$Nle^{12}NH$	2.04	1.99 (0.08)
$Nle^{12}C^{\alpha}H$	$Nle^{12}NH$	2.86	2.79 (0.05)
$Nle^{12}NH$	$Nle^{12}C^{\beta}H^{c}$	3.24	3.11 (0.07)
$Nle^{12}NH$	$Nle^{12}C^{\gamma}H^{c}$	3.73	3.89 (0.18)
$Nle^{12}C^{\alpha}H$	$Nle^{12}C^{\beta}H^{c}$	3.23	3.41 (0.10)
$Nle^{12}C^{\alpha}H$	$Nle^{12}C^{\gamma}H^{c}$	3.99	3.78 (0.12)
$Nle^{12}NH$	$Tyr^{13}NH$	2.85	2.71 (0.05)
$Nle^{12}C^{\alpha}H$	$Tyr^{13}NH$	2.27	2.18 (0.07)
$Nle^{12}C^{\beta}H^{c}$	$Tyr^{13}NH$	3.48	3.34 (0.09)
$Tyr^{13}NH$	$Tyr^{13}C^{\alpha}H$	2.66	2.57 (0.04)

[a] Experimental data are extracted from a NOESY spectrum (250 ms) in DMSO-d_6.
[b] Calculated interproton distances are computed from 200 ps MD trajectories in DMSO & *rmsd* values of the two thousand structures obtained during dynamics are given in parentheses.
[c] Distance involving a pseudoatom.

Table 32. Comparison between the experimental and calculated interproton distances for **C12**

Protons		NOESY-distance (Å)[a]	MD in DMSO average (Å)[b]
$Trp^1C^{\alpha}H$	$Trp^1C^{\beta}H_1$	2.55	2.49 (0.04)
$Trp^1C^{\alpha}H$	$Trp^1C^{\beta}H_2$	2.14	2.27 (0.06)
$Trp^1C^{\alpha}H$	His^2NH	2.00	2.01 (0.06)
His^2NH	Trp^3NH	3.19	3.11 (0.08)
$His^2C^{\alpha}H$	$His^2C^{\beta}H^c$	2.18	2.23 (0.04)
Trp^3NH	$Trp^3C^{\beta}H_1$	2.07	2.21 (0.05)
Trp^3NH	$Trp^3C^{\beta}H_2$	2.56	2.48 (0.07)
$Trp^3C^{\alpha}H$	$Trp^3C^{\beta}H_1$	2.30	2.30 (0.09)
$Trp^3C^{\alpha}H$	$Trp^3C^{\beta}H_2$	2.30	2.36 (0.06)
Trp^3NH	Leu^4NH	2.54	2.62 (0.08)
$Trp^3C^{\alpha}H$	Leu^4NH	1.92	2.03 (0.06)
$Trp^3C^{\beta}H_1$	Leu^4NH	2.74	2.81 (0.05)
$Trp^3C^{\beta}H_2$	Leu^4NH	2.62	2.68 (0.10)
$Leu^4C^{\alpha}H$	Leu^4NH	2.67	2.83 (0.09)
Leu^4NH	Gln^5NH	2.72	2.74 (0.06)
Leu^4NH	$Leu^4C^{\gamma}H$	2.65	2.70 (0.09)
Leu^4NH	$Leu^4C^{\beta}H^c$	2.27	2.35 (0.05)
$Leu^4C^{\alpha}H$	$Leu^4C^{\beta}H^c$	2.49	2.37 (0.07)
$Leu^4C^{\beta}H^c$	Gln^5NH	2.50	2.46 (0.08)
Gln^5NH	$Gln^5C^{\beta}H_1$	2.75	2.79 (0.09)
Gln^5NH	$Gln^5C^{\beta}H_2$	3.00	3.07 (0.05)
Gln^5NH	$Gln^5C^{\gamma}H^c$	2.90	2.81 (0.09)
$Gln^5C^{\beta}H_1$	Leu^6NH	2.75	2.88 (0.11)
$Gln^5C^{\beta}H_2$	Leu^6NH	3.00	2.99 (0.08)
$Gln^5C^{\gamma}H^c$	Leu^6NH	2.92	3.07 (0.06)
Leu^6NH	$Leu^6C^{\beta}H^c$	2.24	2.42 (007)
Leu^6NH	$Leu^6C^{\gamma}H$	2.79	2.66 (0.05)
Leu^6NH	Dpr^7NH	2.28	2.29 (0.05)

Continuation of Table 32

$Leu^{6}C^{\alpha}H$	$Leu^{6}C^{\beta}H^{c}$	2.81	2.78 (0.04)
$Leu^{6}C^{\beta}H^{c}$	$Dpr^{7}NH$	2.79	2.92 (0.11)
$Dpr^{7}NH$	$Dpr^{7}C^{\alpha}H$	3.00	2.83 (0.011)
$Dpr^{7}N^{\gamma}H$	$Dpr^{7}C^{\alpha}H$	3.43	3.57 (0.05)
$Dpr^{7}N^{\gamma}H$	$Dpr^{7}NH$	2.97	2.88 (0.10)
$Dpr^{7}C^{\alpha}H$	$Pro^{8}C^{\delta}H^{c}$	2.00	2.05 (0.04)
$Dpr^{7}N^{\gamma}H$	$Glu^{10}NH$	3.87	3.92 (0.08)
$Dpr^{7}N^{\gamma}H$	$Glu^{10}C^{\beta}H^{c}$	2.56	2.46 (0.04)
$Pro^{8}C^{\alpha}H$	$Gly^{9}NH$	2.40	2.46 (0.06)
$Gly^{9}NH$	$Glu^{10}NH$	2.73	2.59 (0.05)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_1$	2.46	2.35 (0.09)
$Gly^{9}NH$	$Gly^{9}C^{\alpha}H_2$	2.46	2.51 (0.08)
$Glu^{10}NH$	$Glu^{10}C^{\alpha}H$	3.18	3.30 (0.07)
$Glu^{10}NH$	$Glu^{10}C^{\beta}H^{c}$	2.70	2.69 (0.09)
$Glu^{10}NH$	$Glu^{10}C^{\gamma}H_1$	3.70	3.49 (0.15)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\beta}H^{c}$	2.70	2.84 (0.07)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\gamma}H_1$	3.18	3.04 (0.13)
$Glu^{10}C^{\alpha}H$	$Glu^{10}C^{\gamma}H_2$	2.50	2.62 (0.07)
$Nle^{12}C^{\alpha}H$	$Tyr^{13}NH$	2.43	2.55 (0.08)
$Tyr^{13}NH$	$Tyr^{13}C^{\alpha}H$	2.27	2.40 (0.09)
$Nle^{12}C^{\alpha}H$	$Tyr^{13}C^{\alpha}H$	2.33	2.41 (0.07)
$Tyr^{13}C^{\alpha}H$	$Tyr^{13}C^{\beta}H_1$	2.79	2.83 (0.05)
$Tyr^{13}NH$	$Tyr^{13}C^{\beta}H_1$	3.04	3.16 (0.06)

[a] Experimental data are extracted from a NOESY spectrum (250 ms) in DMSO-d_6.

[b] Calculated interproton distances are computed from 200 ps MD trajectories in DMSO &

rmsd values of the two thousand structures obtained during dynamics are given in parentheses.

[c] Distance involving a pseudoatom.

Table 33. $^3J_{NH\text{-}C^{\alpha}H}$ coupling constants (Hz), backbone dihedral angles (°) calculated from coupling constants and dihedral angles (°) extracted from 200 ps MD trajectories in DMSO for for cyclic lactam containing α-factor analogs.

Peptide	Residue	$^3J_{HNC^{\alpha}H}$[a]	ϕ calc[b]	ϕ model[c]
C42	Gly (*NH*, $C^{\alpha}H$)	9.1	43, 148, -43, -148	165 (9)
	Lys ($N^{\varepsilon}H$, $C^{\varepsilon}H$)	10.8	53, 136, -53, -136	-129 (13)
	Lys (*NH*, $C^{\alpha}H$)	7.4	38, 82, -84, -156	-144 (18)
	Glu (*NH*, $C^{\alpha}H$)	5.9	26, 94, -75, -164	47 (8)
C32	Gly (*NH*, $C^{\alpha}H$)	NA[d]	-	87 (10)
	Orn ($N^{\delta}H$, $C^{\delta}H$)	10.4	50, 139, -50, -139	160 (17)
	Orn (*NH*, $C^{\alpha}H$)	5.7	24, 96, -74, -166	-53 (11)
	Glu (*NH*, $C^{\alpha}H$)	8.9	-94, -147	-71 (7)
C22	Gly (*NH*, $C^{\alpha}H$)	11.0	54, 135, -54, -135	66 (11)
	Dab ($N^{\gamma}H$, $C^{\gamma}H$)	11.2	55, 134, -55, -134	-144 (15)
	Dab (*NH*, $C^{\alpha}H$)	7.2	36, 84, -83, -157	-158 (19)
	Glu (*NH*, $C^{\alpha}H$)	8.5	52, 68, -91, -149	-83 (11)
C12	Gly (*NH*, $C^{\alpha}H$)	12.9	66, 122, -66, -122	63 (10)
	Dpr ($N^{\beta}H$, $C^{\beta}H$)	11.9	59, 129, -59, -129	112 (14)
	Dpr (*NH*, $C^{\alpha}H$)	NA[d]		-141 (17)
	Glu (*NH*, $C^{\alpha}H$)	5.9	26, 94, -75, -164	-67 (8)

[a] Experimental value extracted from DQFCOSY and one dimensional proton spectra in DMSO-d_6 at 300K.
[b] Possible values for the peptide backbone angles ϕ (°) from Bystrov's Karplus equations; for Glycyl and side chain *NH*s $\Sigma(^3J_{NH\text{-}C^{\alpha}H}) = -9.4\cos^2\theta - 1.1\cos\theta + 14.9$ and for others $(^3J_{NH\text{-}C^{\alpha}H}) = 9.4\cos^2\theta - 1.1\cos\theta + 0.4$.
[c] Average dihedral angles from 200 ps constrained MD in DMSO & *rmsd* values are given in parentheses.
[d] No experimental value due to signal overlap or unresolvable signal.

Table 34. Comparison between the REDOR experimental distance values, the ideal distance values, and the simulated distance values (all the distances are in Å).

	minimized conformations			ideal conformations			
	type I β-turn	type I β-turn	extended	type I β-turn	type I β-turn	extended	REDOR
1	3.56	3.31	4.69	3.08	2.92	4.77	3.61
2	4.70	5.18	5.82	4.64	4.93	6.10	4.70
3	4.21	4.75	4.76	4.10	4.67	4.87	4.35
4	2.40	2.40	2.40	2.40	2.40	2.40	2.40

Table 35.

Energies of the four glycine conformations considered optimized with different computational methods (in Hartrees.)

Structure	HF/6-311+G**	MP2/6-311+G**//HF/ 6-311+G**	MP3/6-311+G**//HF/ 6-311+G**	MP4/6-311+G**//HF/ 6-311+G**
1	-282.90856	-283.77127	-283.78335	-283.79826
2	-282.91891	-283.77927	-283.79136	-283.80637
3	-282.92010	-283.778313	-283.79480	-283.80961
4	-282.92180	-283.78236	-283.79429	-283.80939

Table 36

Relative energies (kcal/mol) of the glycine conformations as calculated in first order of the perturbation theory.

	HF/6-311+G**	MP2/6-311+G**//HF/ 6-311+G**
4	0	0
3	1.06675	1.938975
2	1.813475	2.539492
1	8.3081	6.958975

Relative energies (kcal/mol) of the glycine conformations as calculated in first order of the perturbation theory.

	MP3/6-311+G**//HF/ 6-311+G**	MP4/6-311+G**//HF/ 6-311+G**
3	0	0
4	0.320025	0.13805
2	2.1586	2.0331
1	7.184875	7.122125

TABLE 37

Calculated bond lengths in the *ab initio* optimized models of glycine reported.

	1	2	3	4
C_1-C_2	1.507	1.511	1.527	1.520
C_1-N_5	1.441	1.446	1.455	1.440
C_1-O_4	1.181	1.183	1.180	1.184
C_1-O_3	1.330	1.326	1.316	1.327
C_1-H_6	1.088	1.087	1.084	1.084
O_3-H_8	0.946	0.946	0.950	0.946
N_5-H_7	0.998	0.998	0.997	0.999

TABLE 38

Calculated bond angles in the *ab initio* optimized models of glycine reported.

	1	2	3	4
H_6-C_1-H_7	106.575	106.733	106.687	105.967
H_6-C_1-C_2	106.644	105.582	106.433	105.920
H_6-C_1-N_5	112.592	112.199	112.017	109.838
H_7-C_1-C_2	106.642	105.583	106.432	105.922
H_7-C_1-N_5	112.591	112.205	112.016	109.835
C_2-C_1-N_5	111.386	113.952	112.819	118.586
H_9-N_5-H_{10}	109.821	108.702	108.341	107.574
H_9-N_5-C_1	111.742	111.290	112.821	111.830
H_{10}-N_5-C_1	111.736	111.294	112.824	111.831
O_3-C_2-O_4	122.977	122.637	122.886	122.517
O_3-C_2-C_1	111.059	114.102	115.582	114.207
O_4-C_2-C_1	125.964	123.261	121.532	123.277
H_8-O_3-C_2	109.023	108.749	108.580	108.584

TABLE 39

Calculated torsion angles in the *ab initio* optimized models of glycine reported.

	1	2	3	4
H_6-C_1-N_5-H_9	168.054	175.212	-121.083	61.557
H_6-C_1-N_5-H_{10}	-68.442	-63.373	2.112	-177.723
H_7-C_1-N_5-H_9	-71.429	-64.626	-1.241	177.726
H_7-C_1-N_5-H_{10}	52.074	56.789	121.954	-61.554
C_2-C_1-N_5-H_9	48.311	55.296	118.837	-60.359
C_2-C_1-N_5-H_{10}	171.815	176.711	-117.968	60.362
H_6-C_1-C_2-O_3	47.445	-139.543	-123.541	-123.857
H_6-C_1-C_2-O_4	-132.558	40.463	56.483	56.145
H_7-C_1-C_2-O_3	-66.141	107.622	122.936	123.881
H_7-C_1-C_2-O_4	113.856	-72.372	-57.040	-56.118
N_5-C_1-C_2-O_3	170.653	-15.964	-0.302	0.013
N_5-C_1-C_2-O_4	-9.351	164.041	179.723	-179.986
O_4-C_2-O_3-H_8	171.278	8.489	-179.860	-0.012
C_1-C_2-O_3-H_8	-8.725	-171.505	0.165	179.989

Table 40

The charges of the minimized structure #4 as calculated by the Mulliken population analysis. The meaning of the atom numbers is given in Figure 30.

Atom	Charge
C_1	-0.091102
C_2	0.193416
O_3	-0.296634
O_4	-0.401490
N_5	-0.532882
H_6	0.162734
H_7	0.164745
H_8	0.299508
H_9	0.251445
H_{10}	0.250260

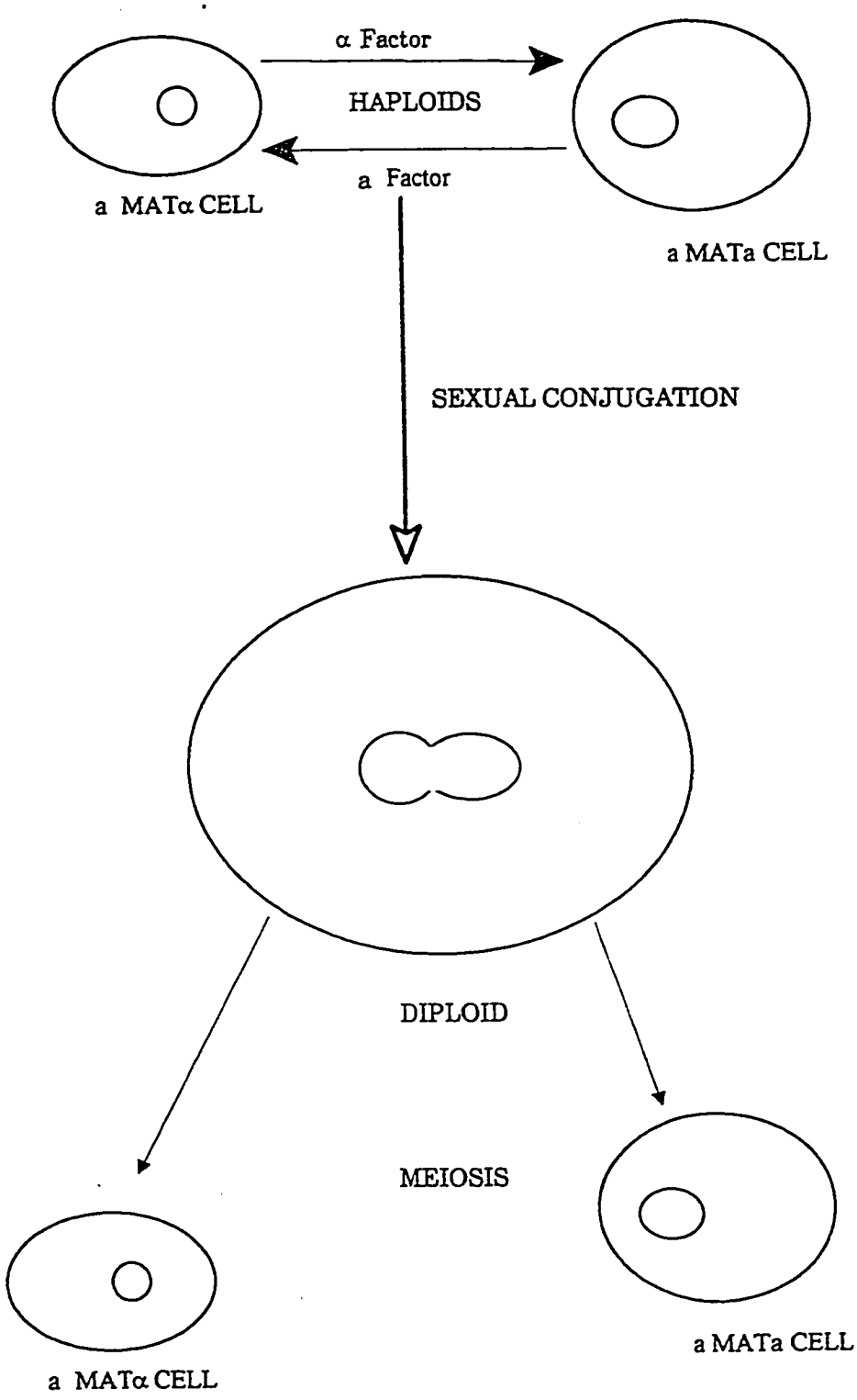


Figure 1. The *S. Cerevisiae* diploid to haploid alternation and sexual conjugation.

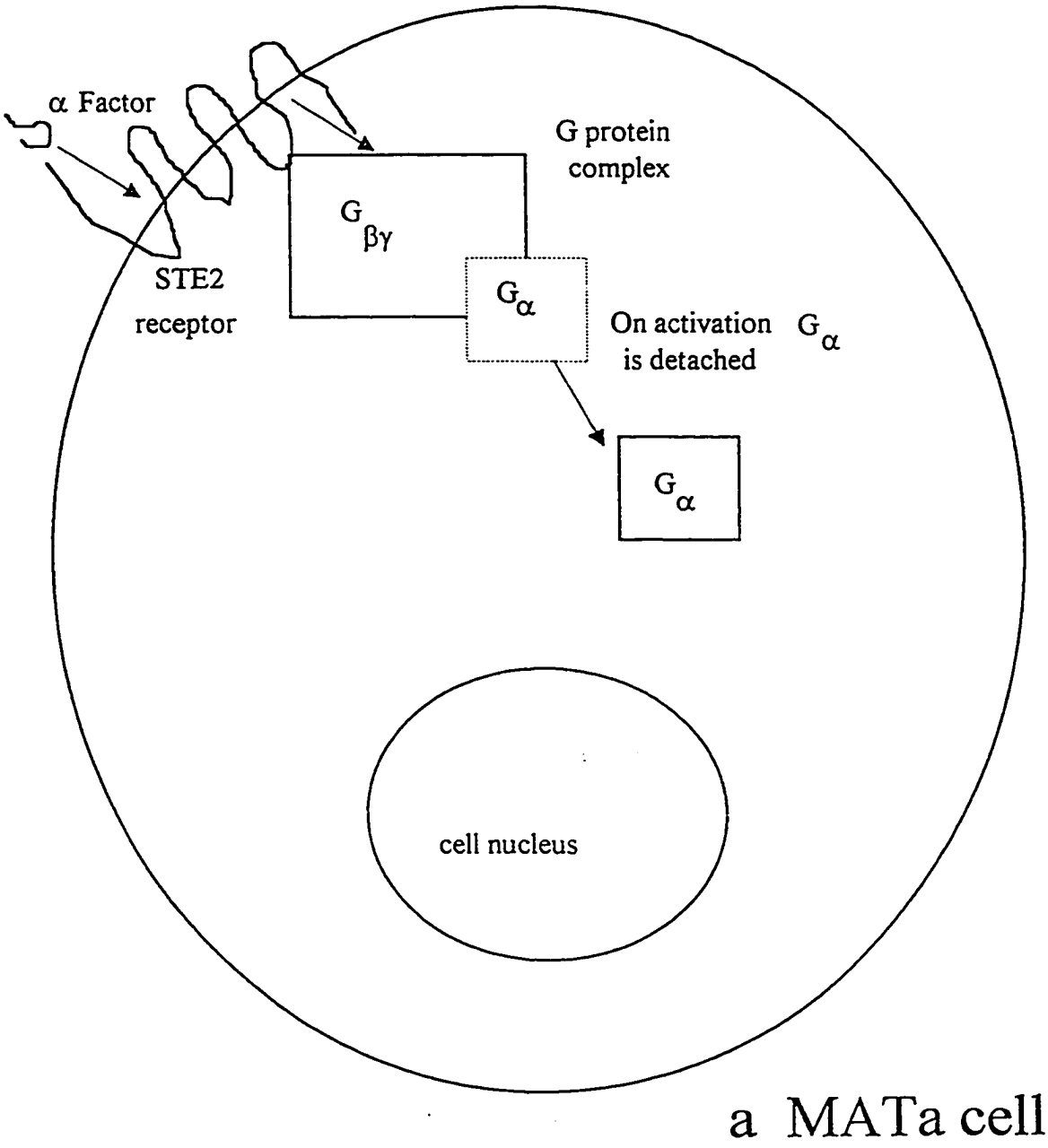
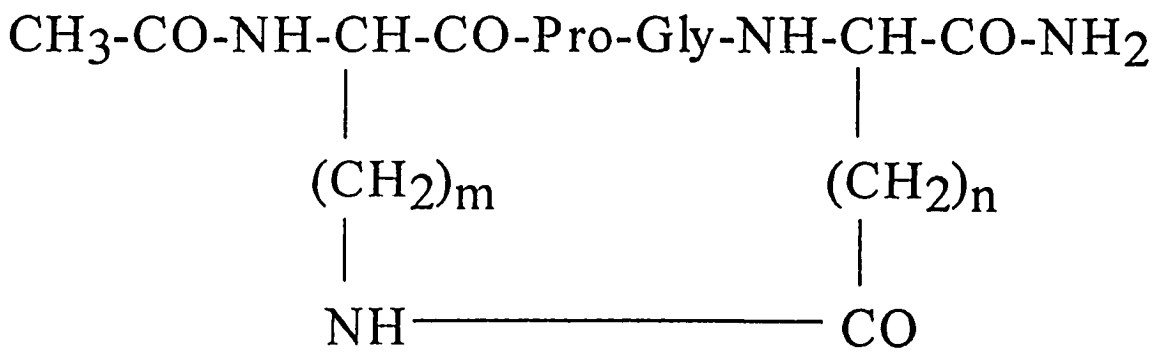


Figure 2

The effect of α Factor on a *MAT***a** cell

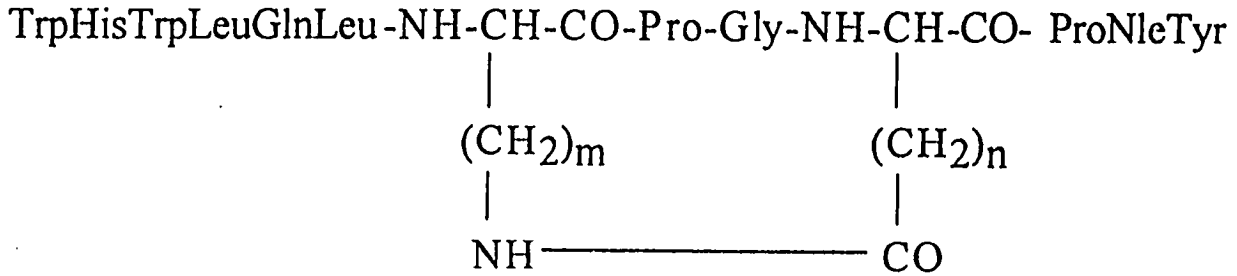
Figure 3a
The formulas of the cyclized tetrapeptides (modeling the turn region).



Peptide	Number of methylenes		Amino acid in Position	
			i	*i+3*
$Tetra_{42}$	m = 4	n = 2	Lys	Glu
$Tetra_{32}$	m = 3	n = 2	Orn	Glu
$Tetra_{22}$	m = 2	n = 2	Dab	Glu
$Tetra_{12}$	m = 1	n = 2	Dpr	Glu
$Tetra_{41}$	m = 4	n = 1	Lys	Asp
$Tetra_{31}$	m = 3	n = 1	Orn	Asp
$Tetra_{21}$	m = 2	n = 1	Dab	Asp
$Tetra_{11}$	m = 1	n = 1	Dpr	Asp

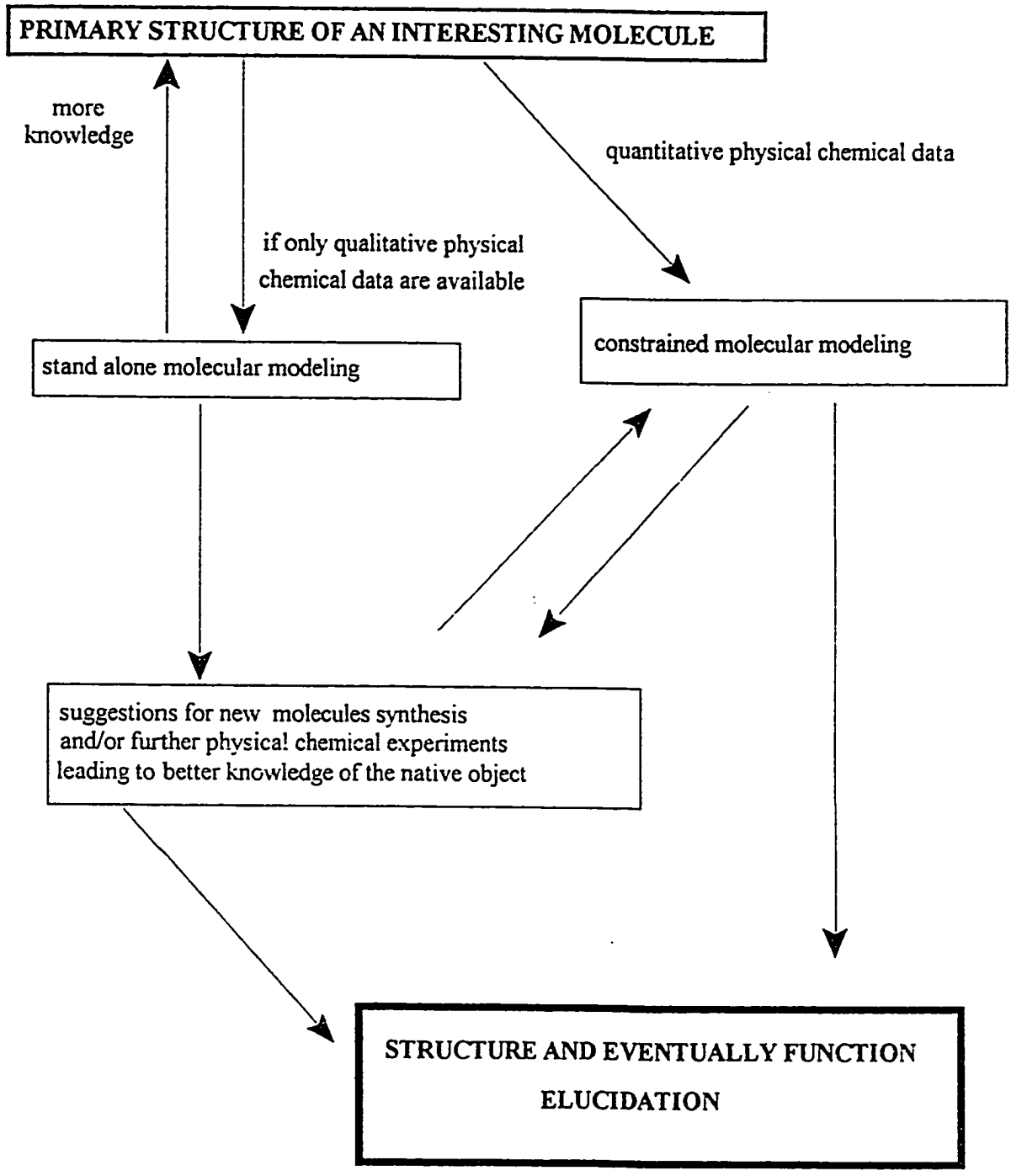
Figure 3b

The formulas of the cyclized tetrapeptides (modeling the turn region) and of the residue 7 to residue 10 side chain lactam cyclized tridecamer pheromone analogs.



Peptide	Number of methylenes		Amino acid in Position	
			i	*i+3*
C42	m = 4	n = 2	Lys	Glu
C32	m = 3	n = 2	Orn	Glu
C22	m = 2	n = 2	Dab	Glu
C12	m = 1	n = 2	Dpr	Glu
C41	m = 4	n = 1	Lys	Asp
C41	m = 3	n = 1	Orn	Asp
C41	m = 2	n = 1	Dab	Asp
C41	m = 1	n = 1	Dpr	Asp

Figure 4
A simplified flow chart illustrating the conceptual process of possible usages of modeling in the understanding of peptide conformations.



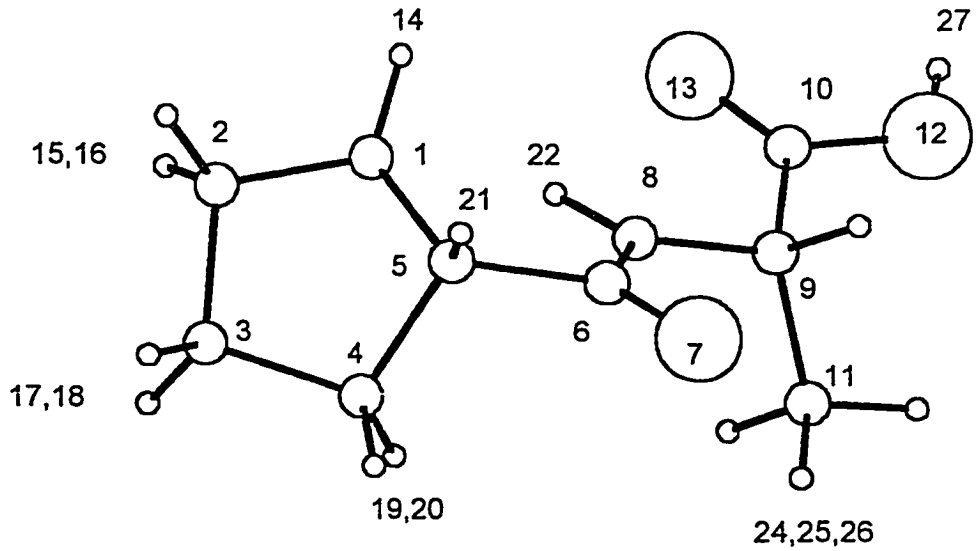


Figure 5a. *L*-Pro-*L*-Ala conformation #1 (Table 1a, legend);
ψ(Pro) = -17.6°, φ(Ala) = -158.4°.

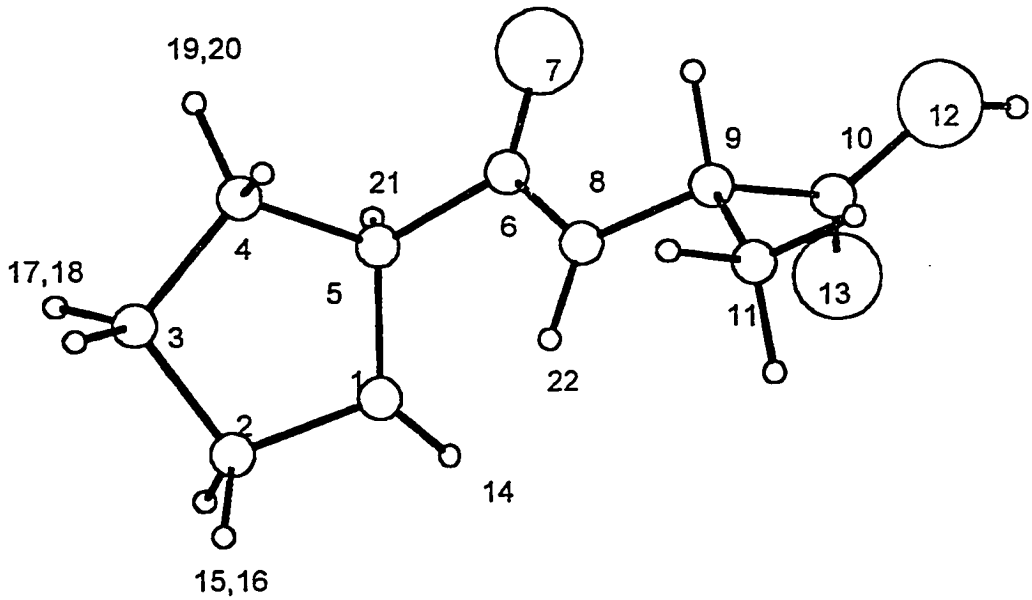


Figure 5b. *L*-Pro-*L*-Ala conformation #2; ψ(Pro) =-18.5°;φ(Ala) = -85.0°.

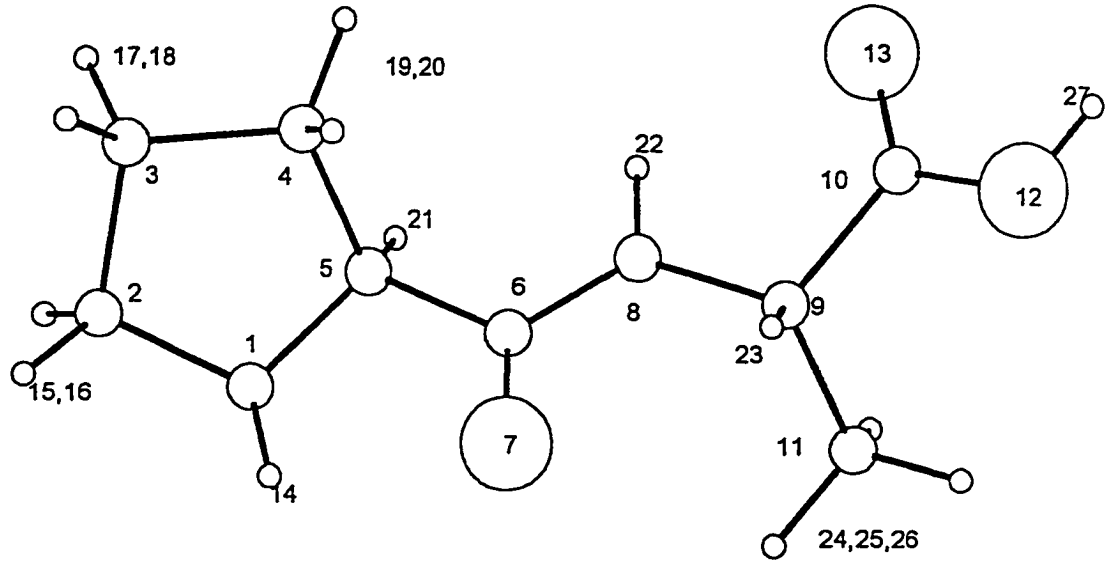


Figure 5c. *L*-Pro-*L*-Ala structure #3; ψ(Pro) = 160.0°;φ(Ala) = -161.0°.

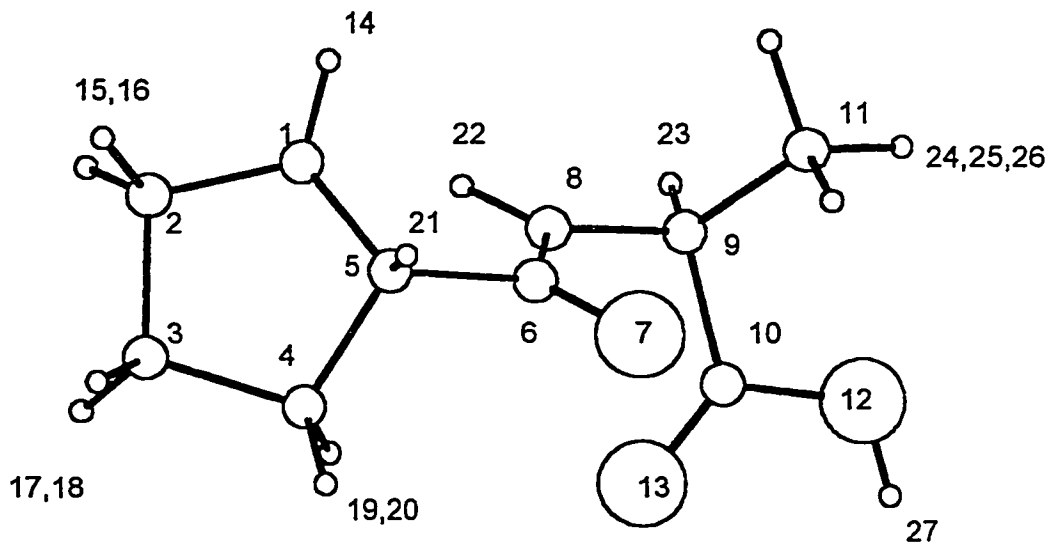


Figure 5d. *L*-Pro-*L*-Ala conformation #4; ψ(Pro) = -21.4°; φ(Ala) = 66.4°.

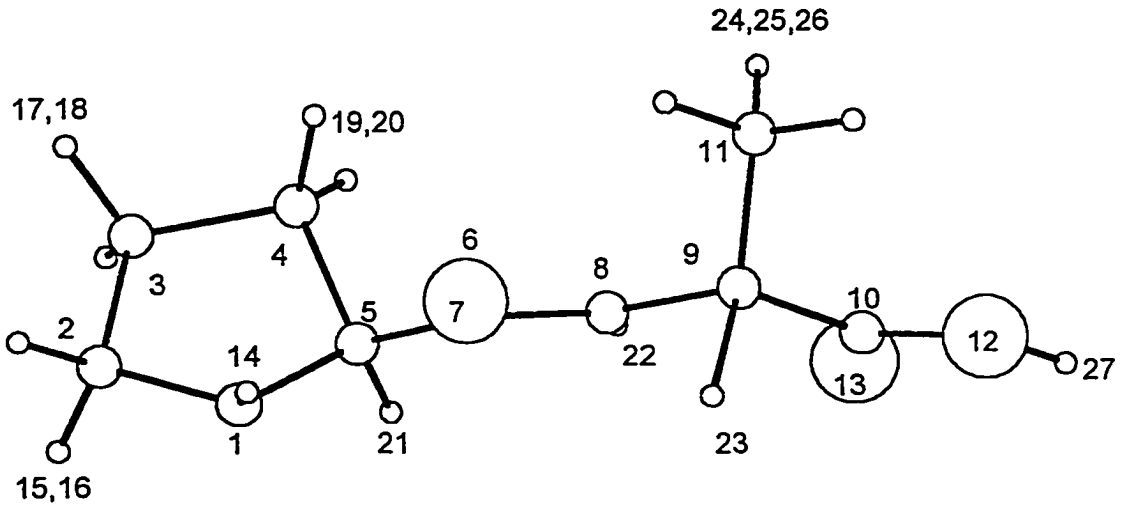


Figure 6a. *L*-Pro-*D*-Ala conformation #1; ψ(Pro) = 155.6°; φ(Ala) = 162.8°.

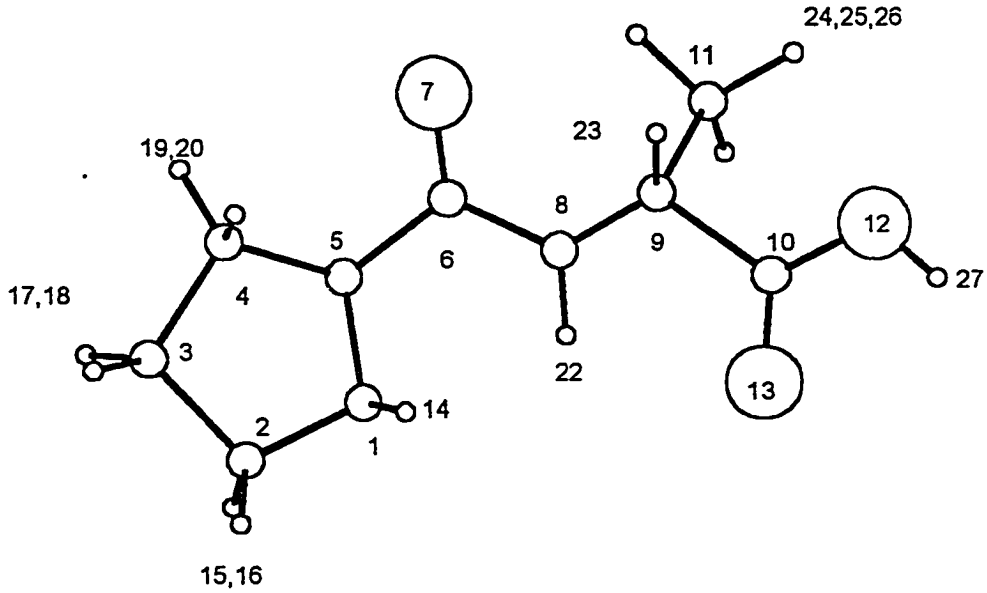


Figure 6b. *L*-Pro-*D*-Ala conformation #2: ψ(Pro) = 33.4°; φ(Ala) = 160.3°.

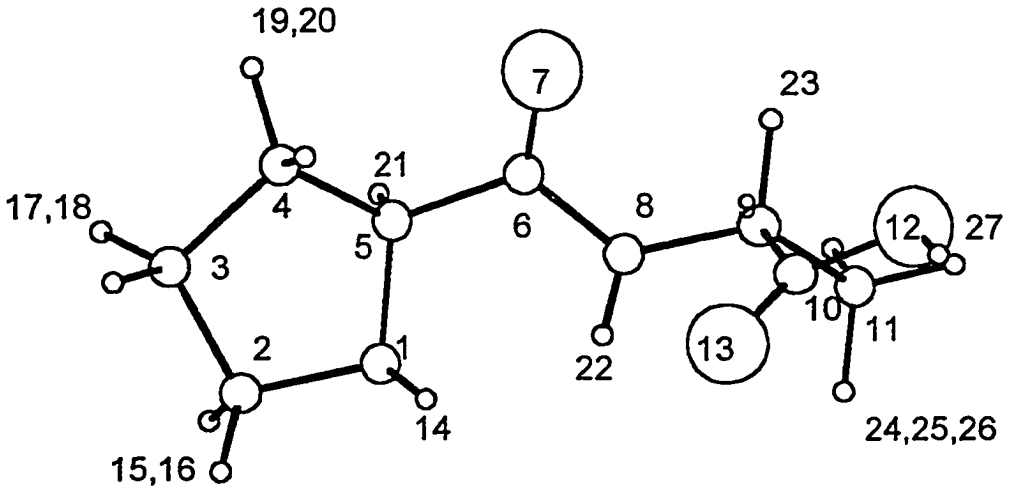


Figure 6c. *L*-Pro-*D*-Ala conformation #3: ψ(Pro) = 37.5°; φ(Ala) = 85.1°.

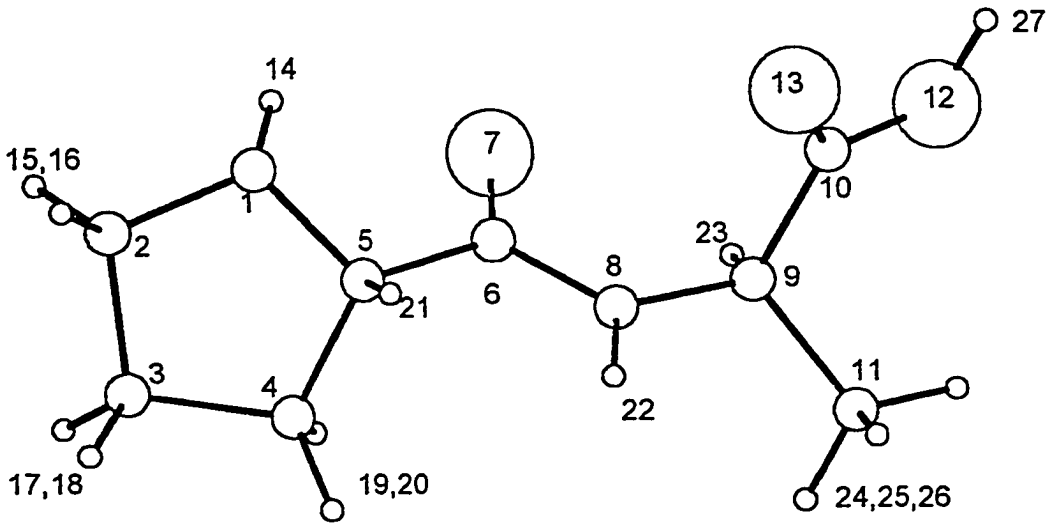


Figure 6d *L*-Pro-*D*-Ala conformation #4: ψ(Pro) =149.3°; φ(Ala) = 85.3°.

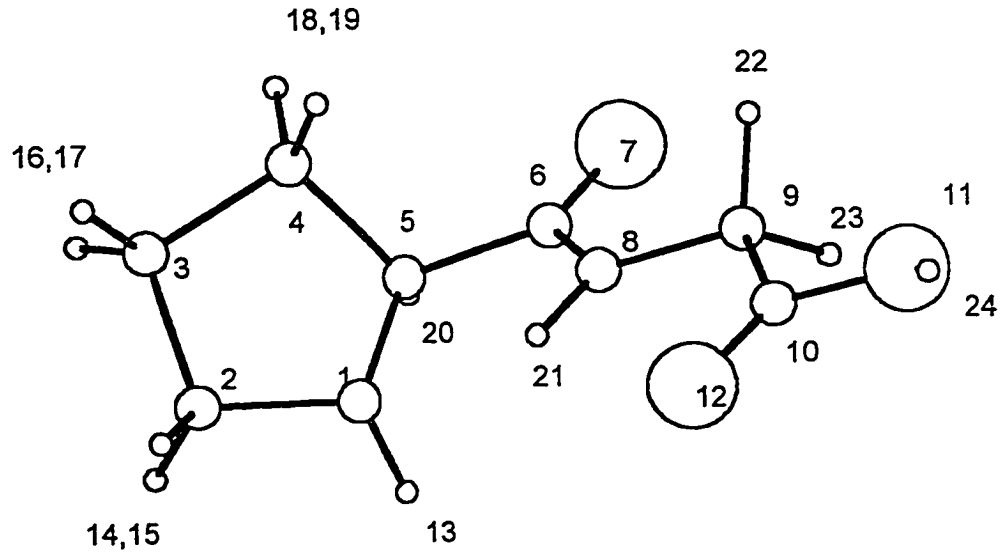


Figure 7. *L*-Pro-Gly conformation #1; ψ(Pro) = -18.5°; φ(Gly) = -175.5°.

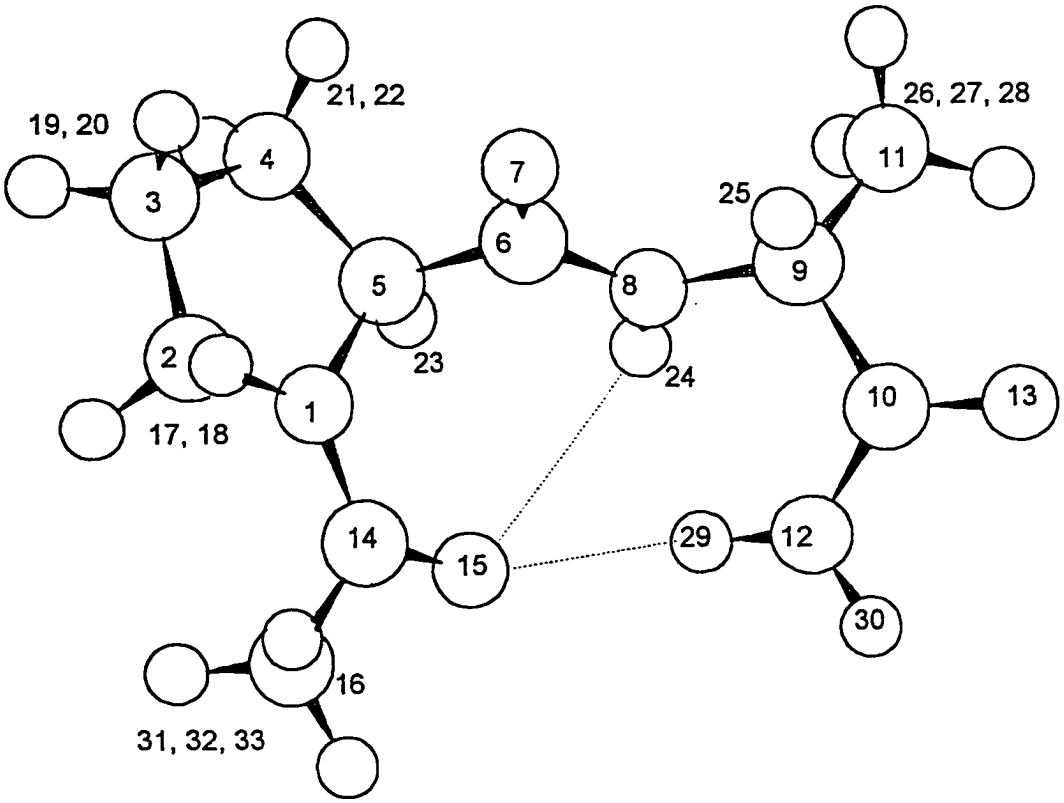


Figure 8a. The lowest energy structure found for the Ac-*L*-Pro-*D*-Ala-NH_2 molecule (structure #1 in Table 6a). The bifurcated hydrogen bond involving hydrogen atoms H24 and H29 and the oxygen atom O15 is displayed. This feature seems to be the reason for the very low energy exhibited by this type II β-turn conformation.

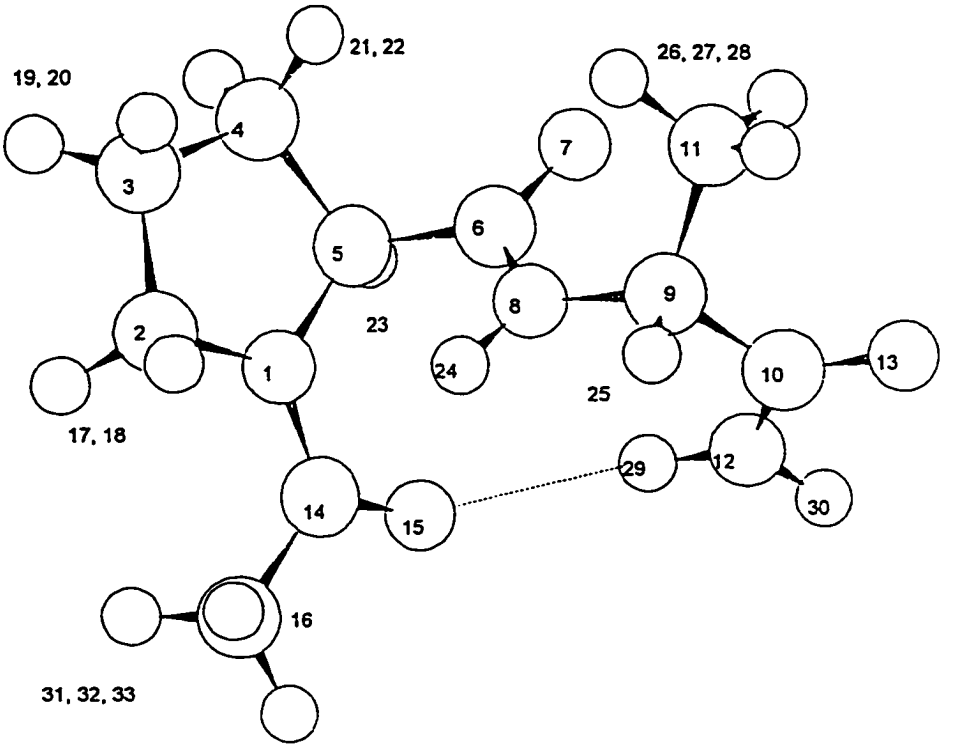


Figure 8b. Type I β-turn like structure for the Ac-*L*-Pro-*D*-Ala-NH_2 (structure #2 in Table 6a). The hydrogen bond involving H29 and O15 displayed seems to provide the reason for this conformation being the second lowest energy structure exhibited by the *D-alanine* dipeptide although it has no nonblocked counterpart.

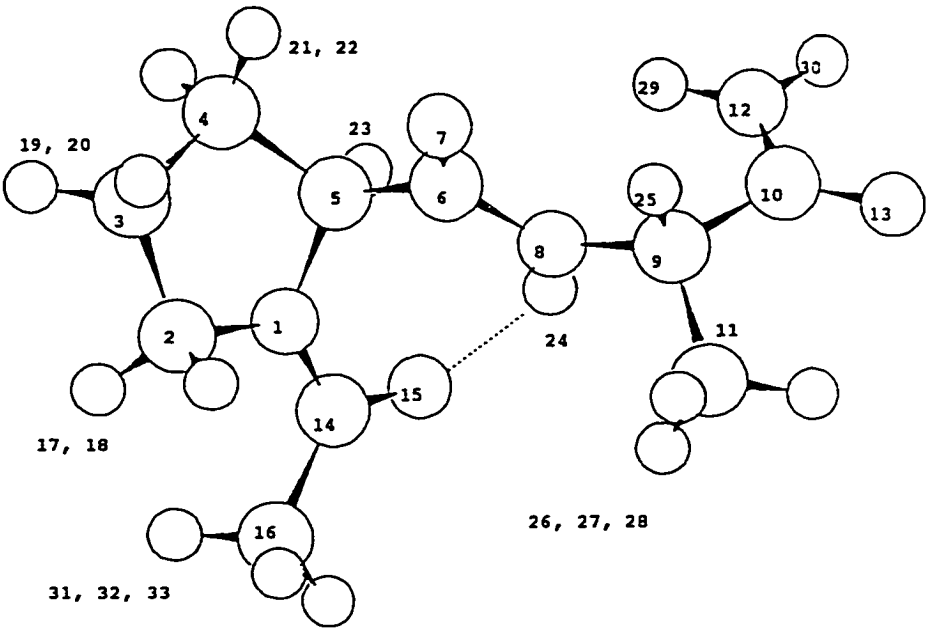


Figure 9a. The lowest energy structure found for the Ac-*L*-Pro-*L*-Ala-NH_2 molecule (structure #1 in Table 6b). The hydrogen bond involving the atoms H24 O15 is displayed. Although the four atoms C14, O15, H24. and N8 are far from being colinear, this H-bond is very strong as judged by the other characteristics described in the text. Its strength is probably the basis of the low energy exhibited by this conformation.

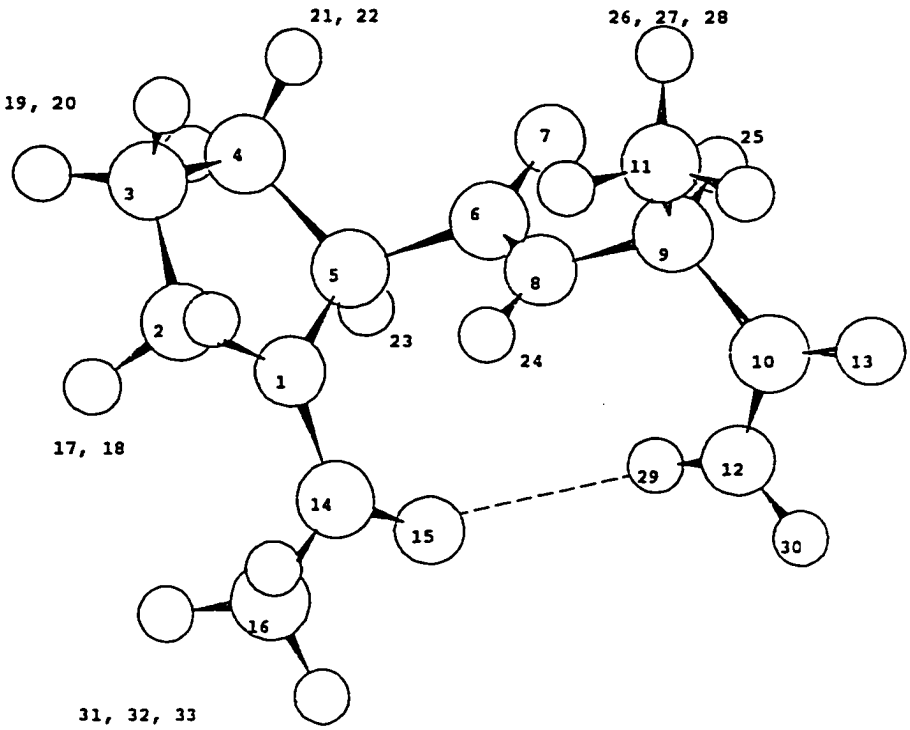


Figure 9b. The second lowest energy structure found for the Ac-*L*-Pro-*L*-Ala-NH_2 molecule (structure #2 in Table 6b) exhibits the attributes of a type I β-turn. The hydrogen bond involving the atoms H29 and O15 that closes the characteristic turn ring is displayed.

Figure 10.The conformations obtained by stand alone molecular modeling for the cyclic tetrapeptides having Glu at position 4 (*i+3*).

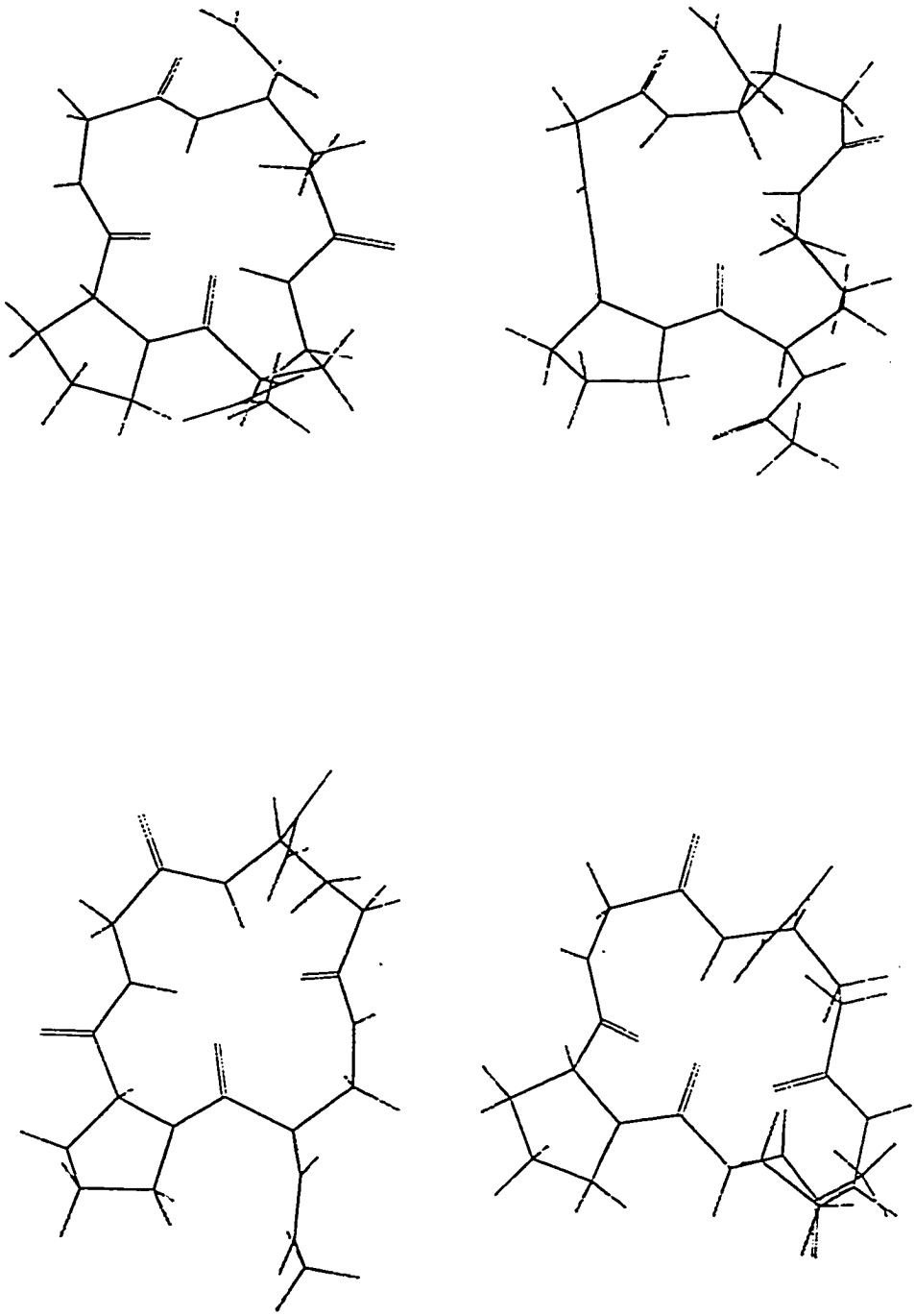
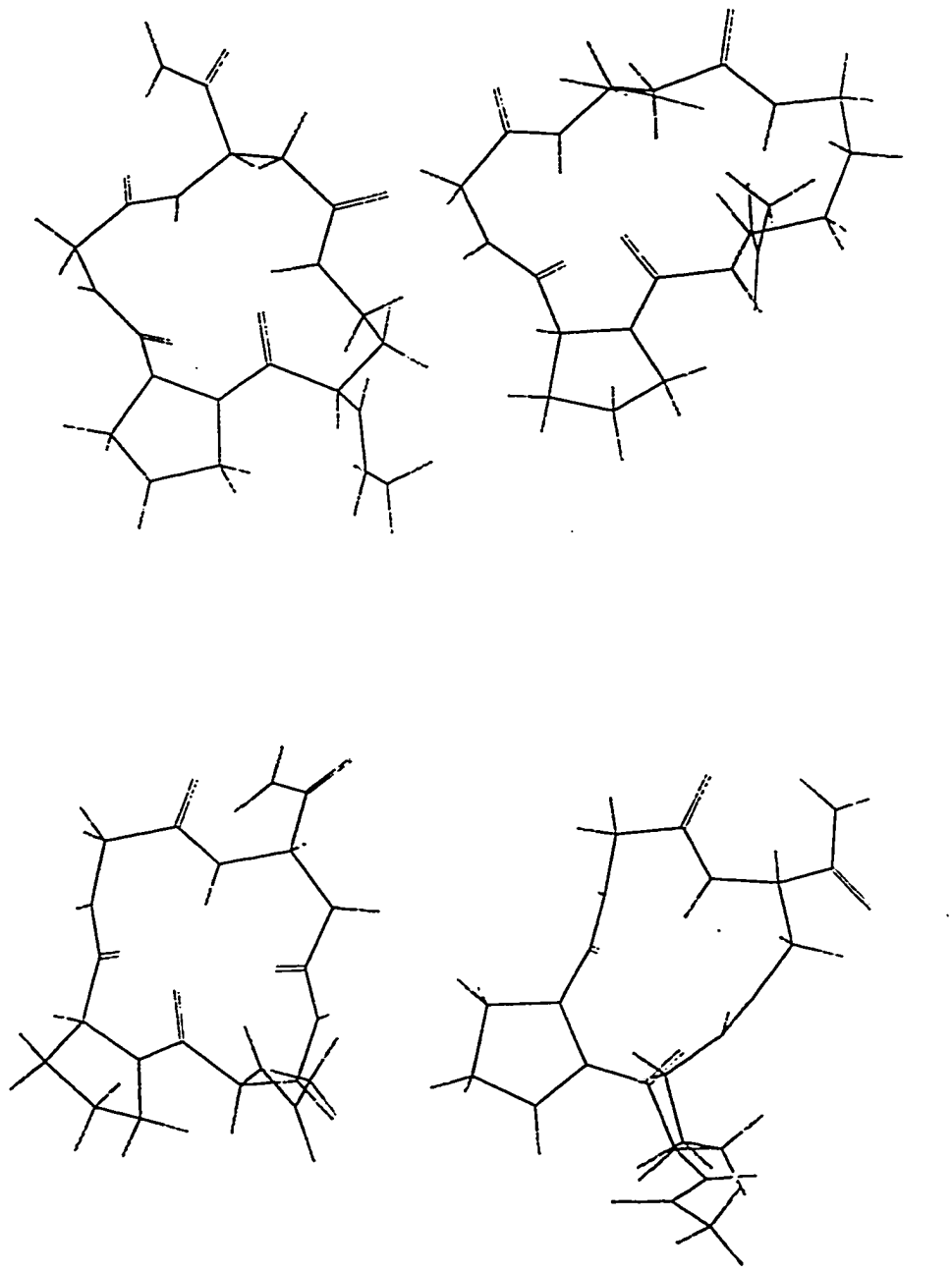


Figure 11.The conformations obtained by stand alone molecular modeling for the cyclic tetrapeptides having Asp at position 4 (*i+3*).



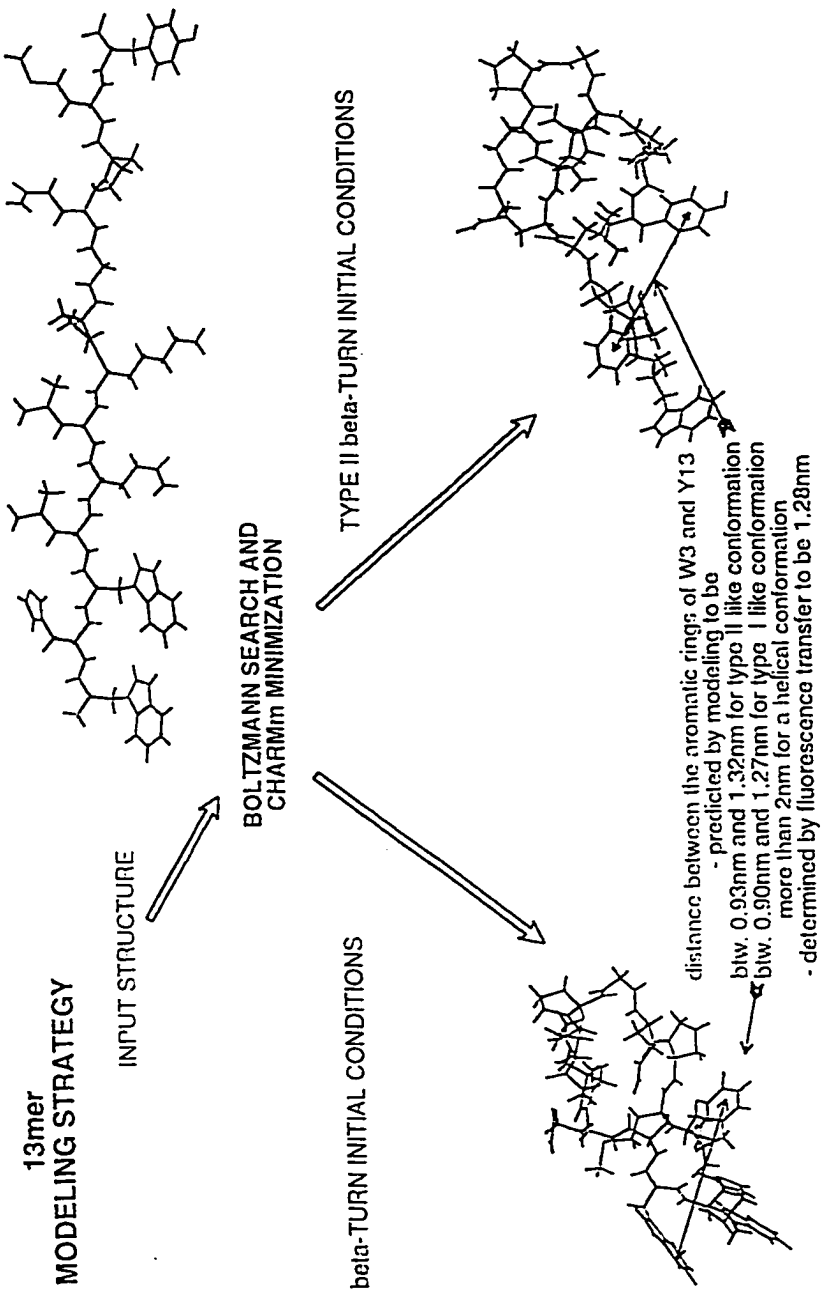
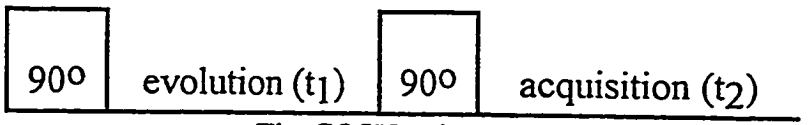


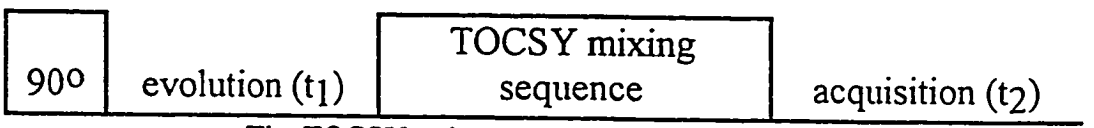
Figure 12

Figure 13. Pulse sequences used for assignment and quantitative conformational measurements of peptides.

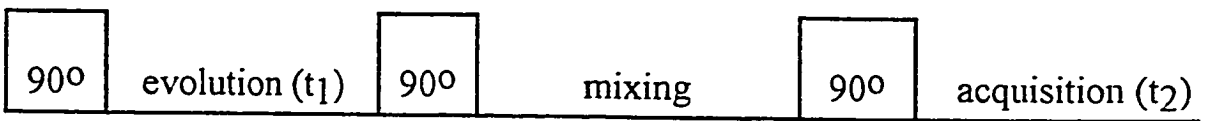


The COSY pulse sequence.

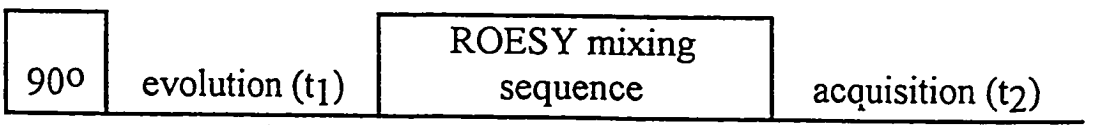
The first pulse creates single quantum coherences. During the t_1 single and double quantum coherences evolve in the constant magnetic field. The second 90° degrees pulse insures that the "readable" (i.e., that can be determined experimentally) coherences contain information that links J-coupled spins. To be more specific, this pulse creates coherences of one spin containing t_1 oscillating factors depending on the spin frequency of a J-coupled partner. During the acquisition period, the signal from the sample is measured and stored as a computer file.



The TOCSY pulse sequence.



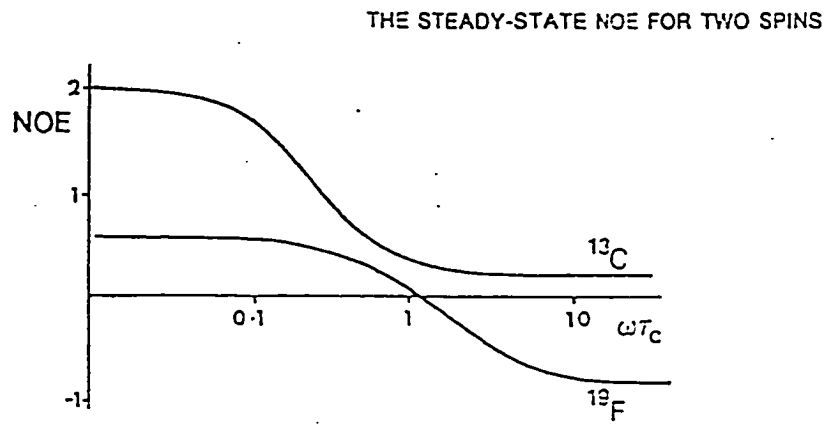
The NOESY pulse sequence.



The ROESY pulse sequence.

The difference between the ROESY and TOCSY pulse sequences is only technical (the pulse widths of the mixing sequence pulses are adjusted differently) but from an experimental point of view the results are dramatically different: the ROESY gives only through space correlations, while the TOCSY gives multiple through bond correlations.

Figure 14. The conformations obtained by stand alone molecular modeling for the cyclic tetrapeptides having Glu at position 4 (*i+3*).



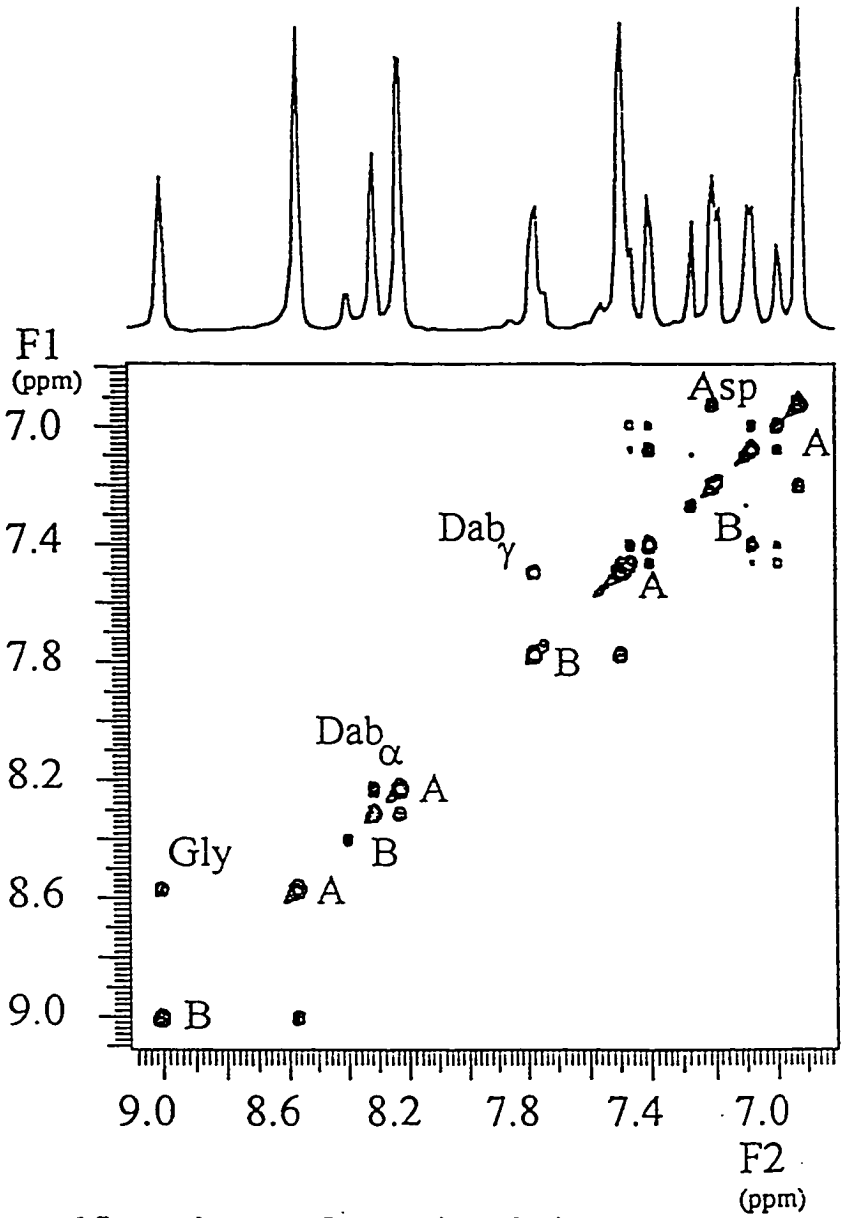


Figure 15. The conformational isomers as seen in the regular ROESY cross peaks. The Dpr$^{\alpha}$CH- Pro$^{\alpha}$CH cross peaks are highlighted.

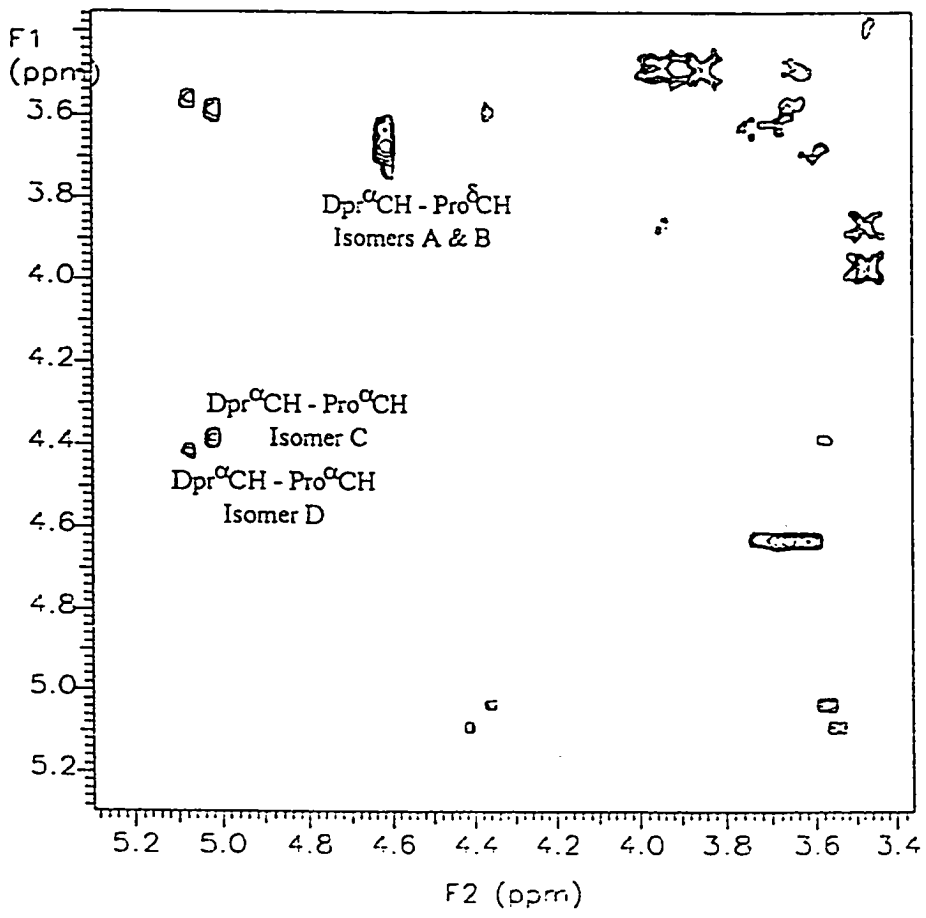


Figure 16. The conformational isomers as seen in the regular ROESY cross peaks. The Dpr$^{\alpha}$CH- Pro$^{\alpha}$CH cross peaks are highlighted.

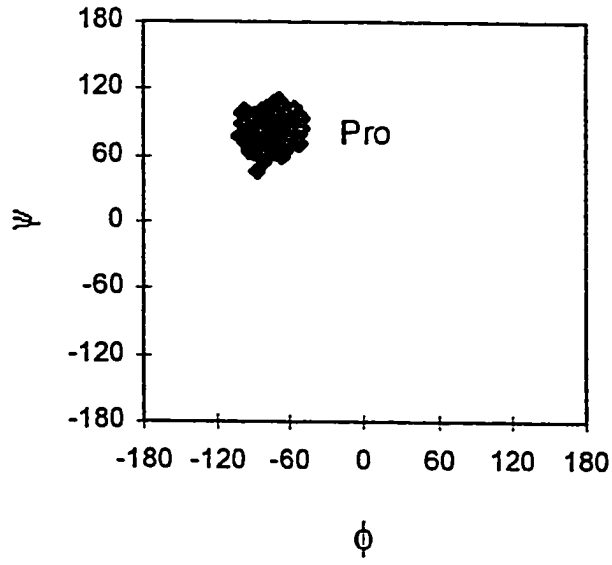


Figure 17a. Ramachandran map of the backbone dihedral angles of the DMSO dynamics structures for $Tetra_{42}$

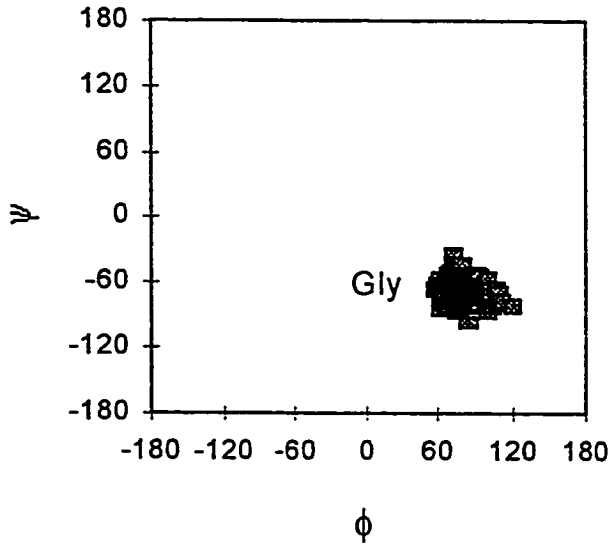


Figure 17b. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structures for $Tetra_{32}$.

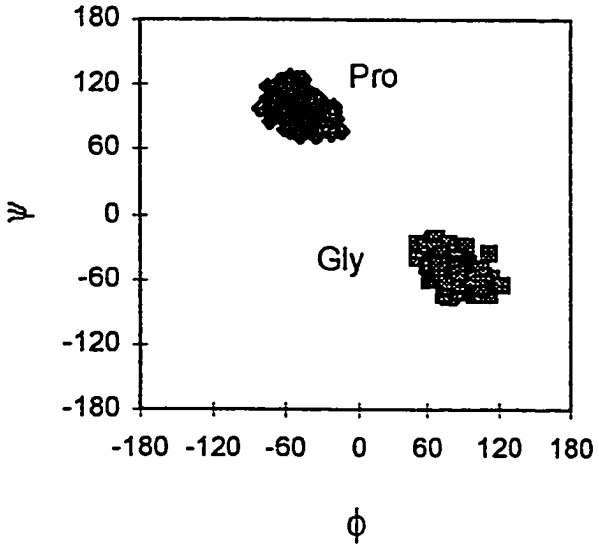


Figure 17c. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structures for $Tetra_{31}$.

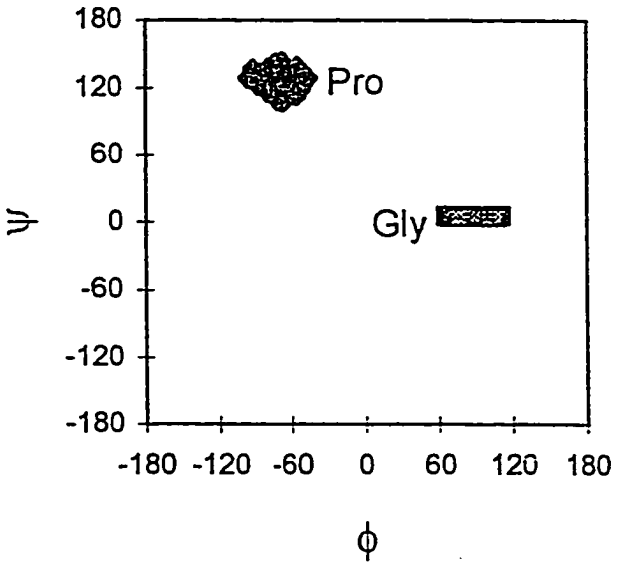
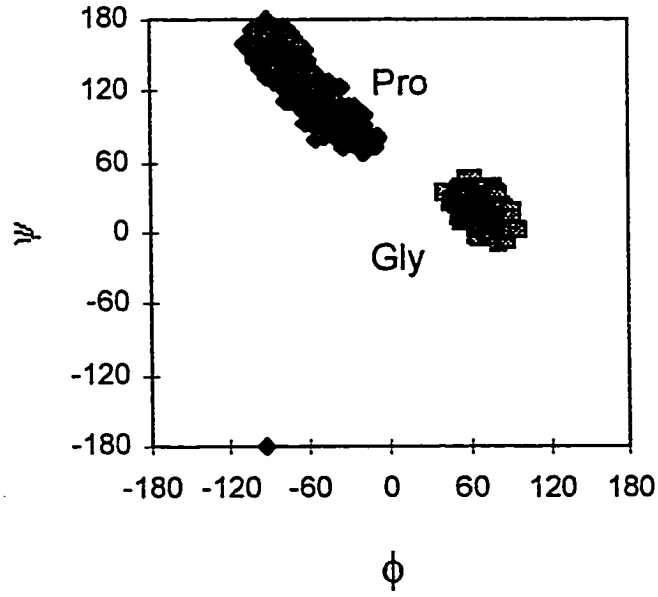


Figure 17d. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structures $Tetra_{22}$.



Figures 17e. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structures for $Tetra_{12}$.

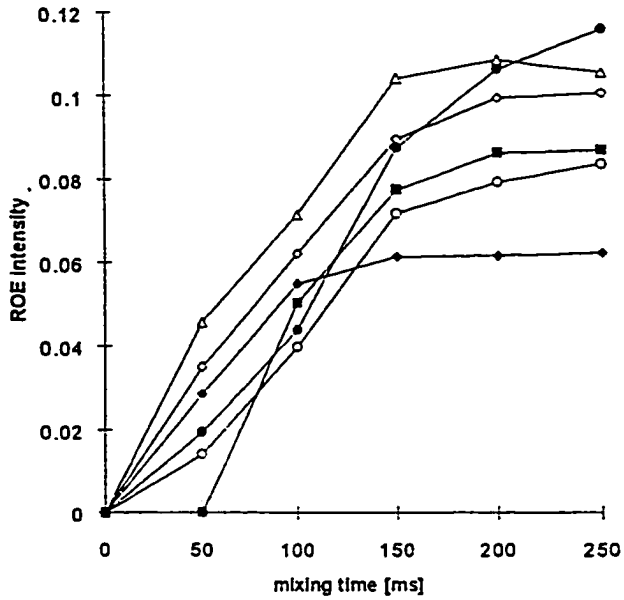


Figure 18. ROE buildup curves.

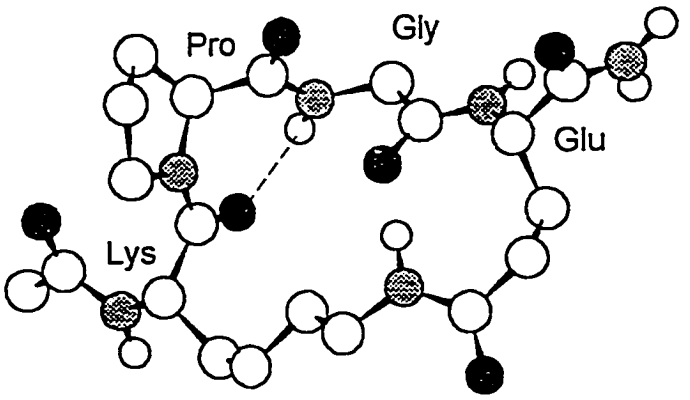


Figure 19a. The models of $Tetra_{42}$ cyclized tetramer representing likely DMSO solution conformations as resulting from the usage of physical chemical data in the molecular modeling studies.

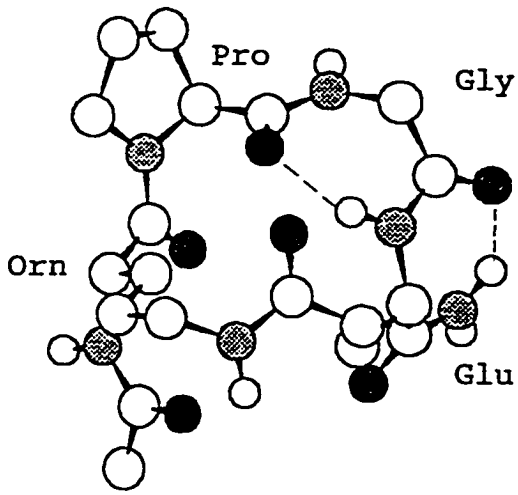


Figure 19b. The models of $Tetra_{32}$ cyclized tetramer representing likely DMSO solution conformations as resulting from the usage of physical chemical data in the molecular modeling studies.

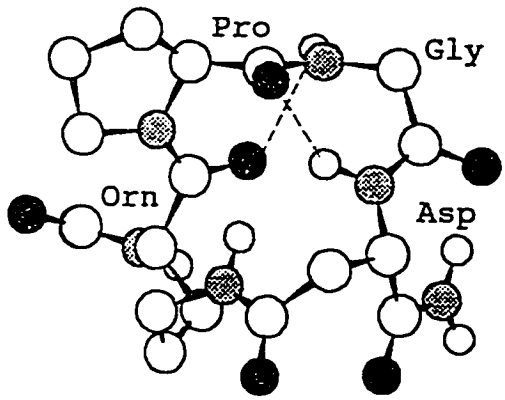


Figure 19c. The models of $Tetra_{31}$ cyclized tetramer representing likely DMSO solution conformations as resulting from the usage of physical chemical data in the molecular modeling studies.

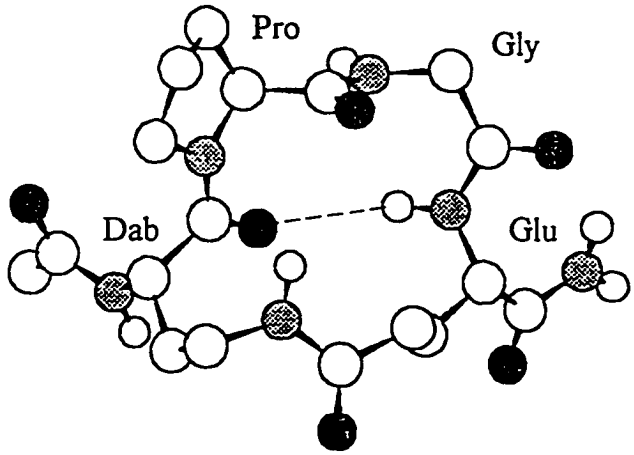


Figure 19d. The models of $Tetra_{22}$ cyclized tetramer representing likely DMSO solution conformations as resulting from the usage of physical chemical data in the molecular modeling studies.

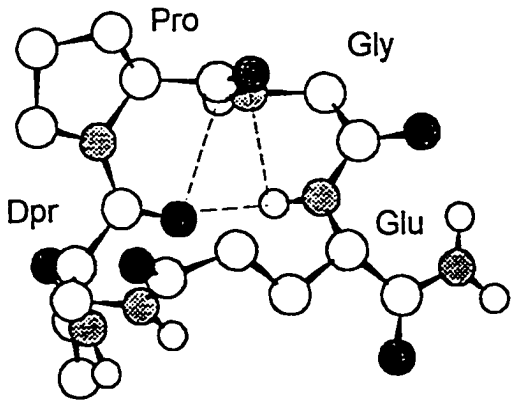


Figure 19e. The models of $Tetra_{12}$ cyclized tetramer representing likely DMSO solution conformations as resulting from the usage of physical chemical data in the molecular modeling studies.

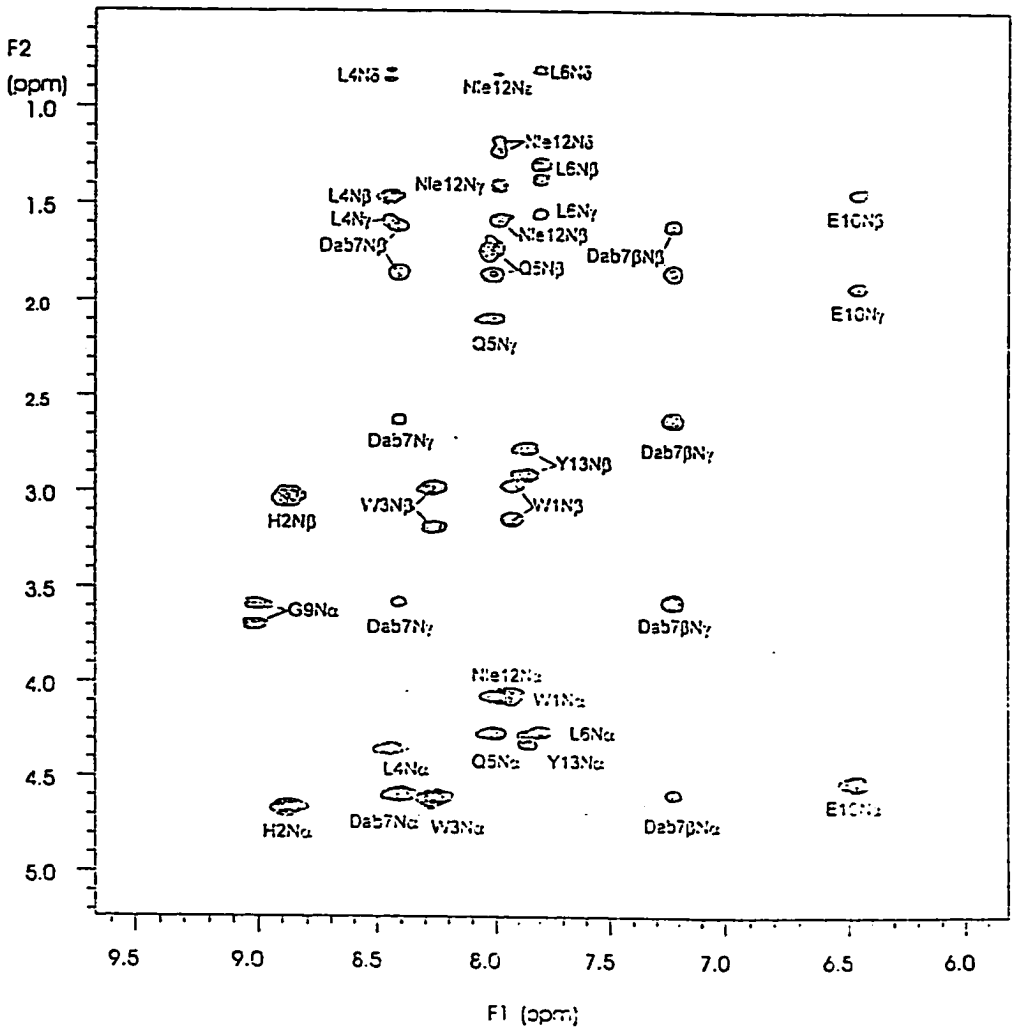


Figure 20. Typical TOCSY connectivities - the case of C22.

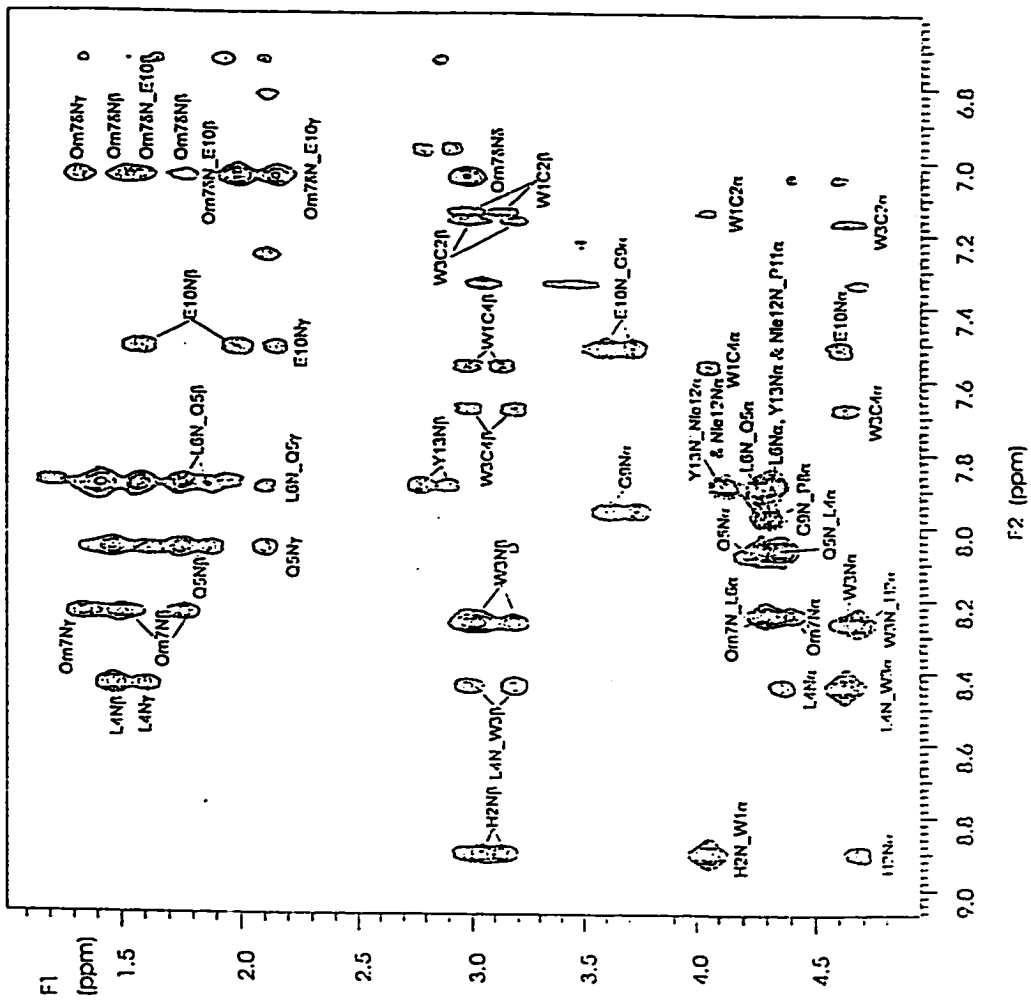


Figure 21. Typical spectrum that illustrates some NOESY cross peaks for C32. Quantitative analysis of these spectra is the basis for NMR assisted molecular modeling

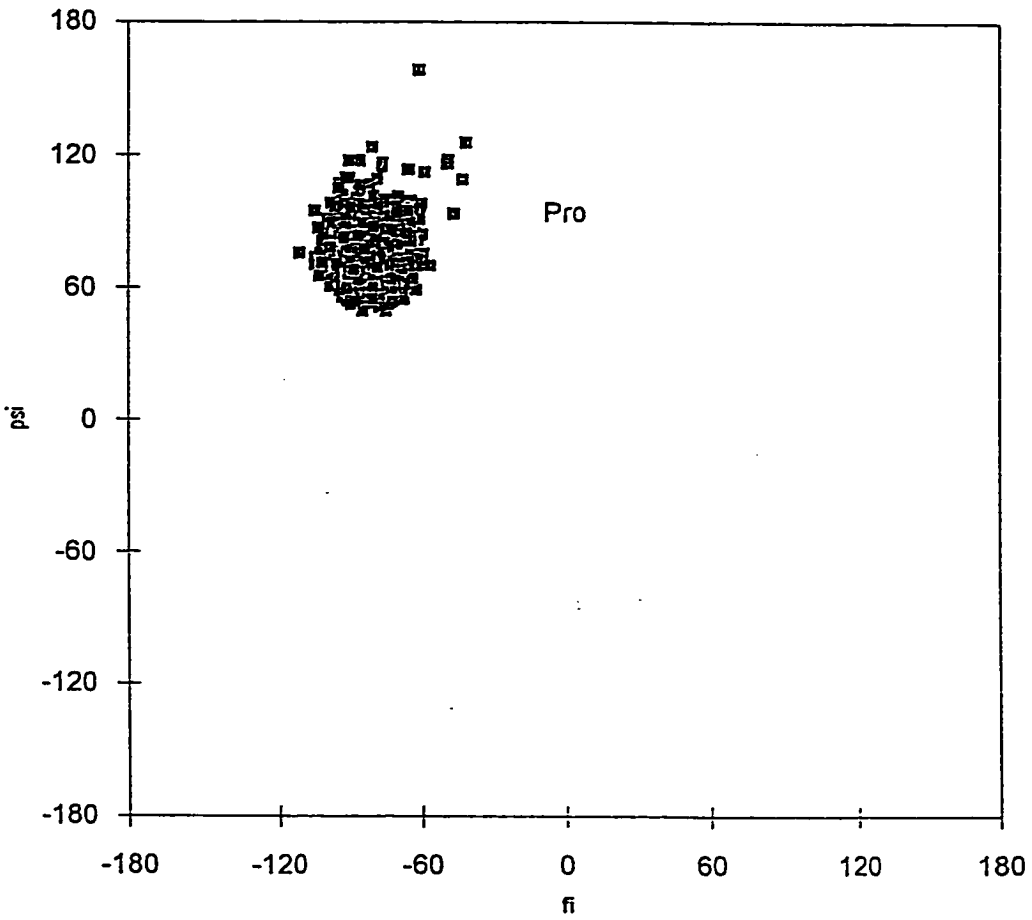


Figure 22a. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structure C42; Dihedral angles (ϕ,ψ) for the Pro and/or Gly residues from 1000 discrete conformers observed during a 60 ps simulation are included in the plots.

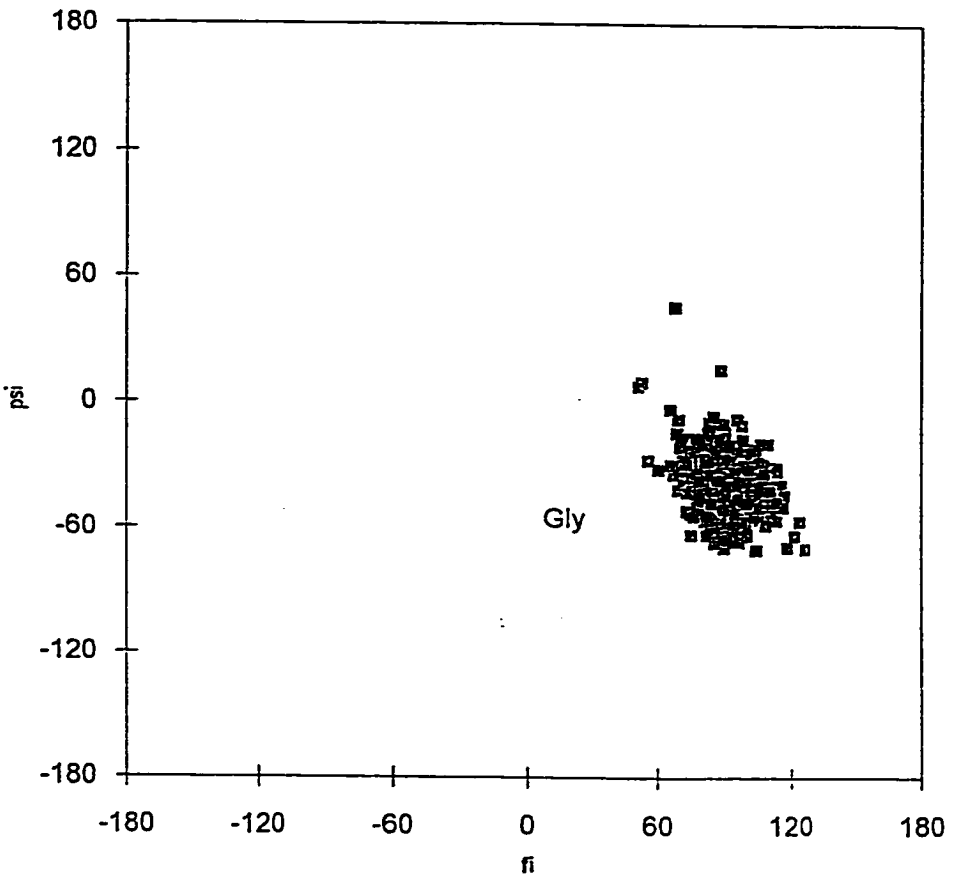


Figure 22b. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structure C32; Dihedral angles (ϕ,ψ) for the Pro and/or Gly residues from 1000 discrete conformers observed during a 60 ps simulation are included in the plots.

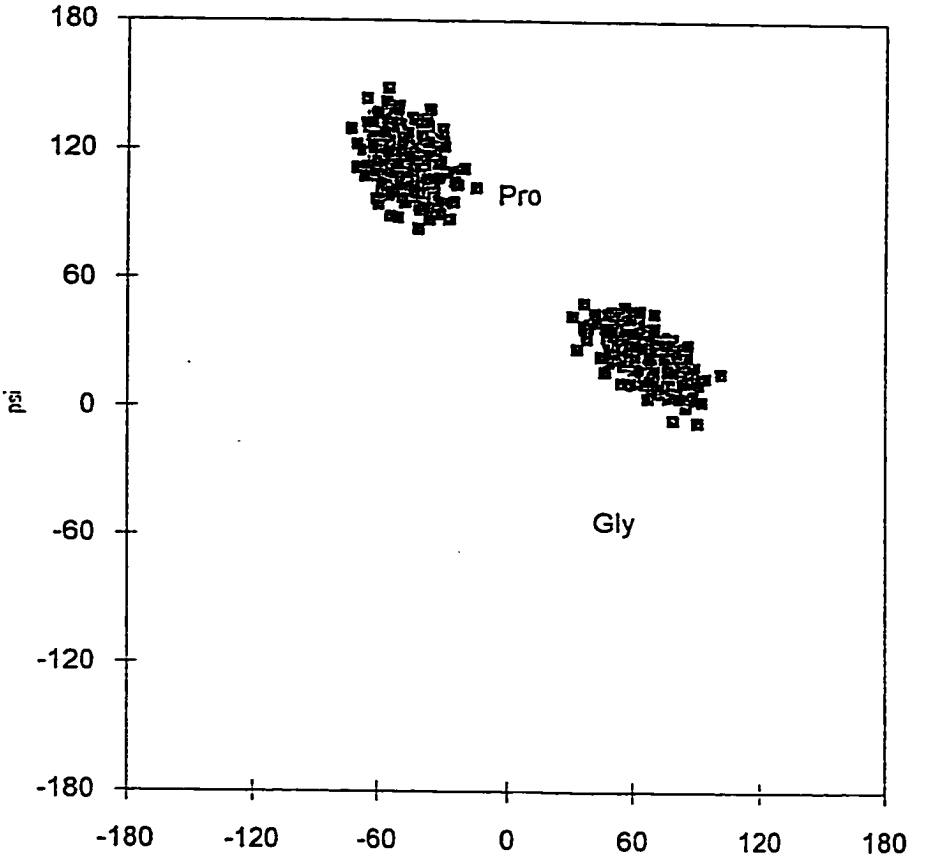


Figure 22c. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structure C22; Dihedral angles (ϕ,ψ) for the Pro and/or Gly residues from 1000 discrete conformers observed during a 60 ps simulation are included in the plots.

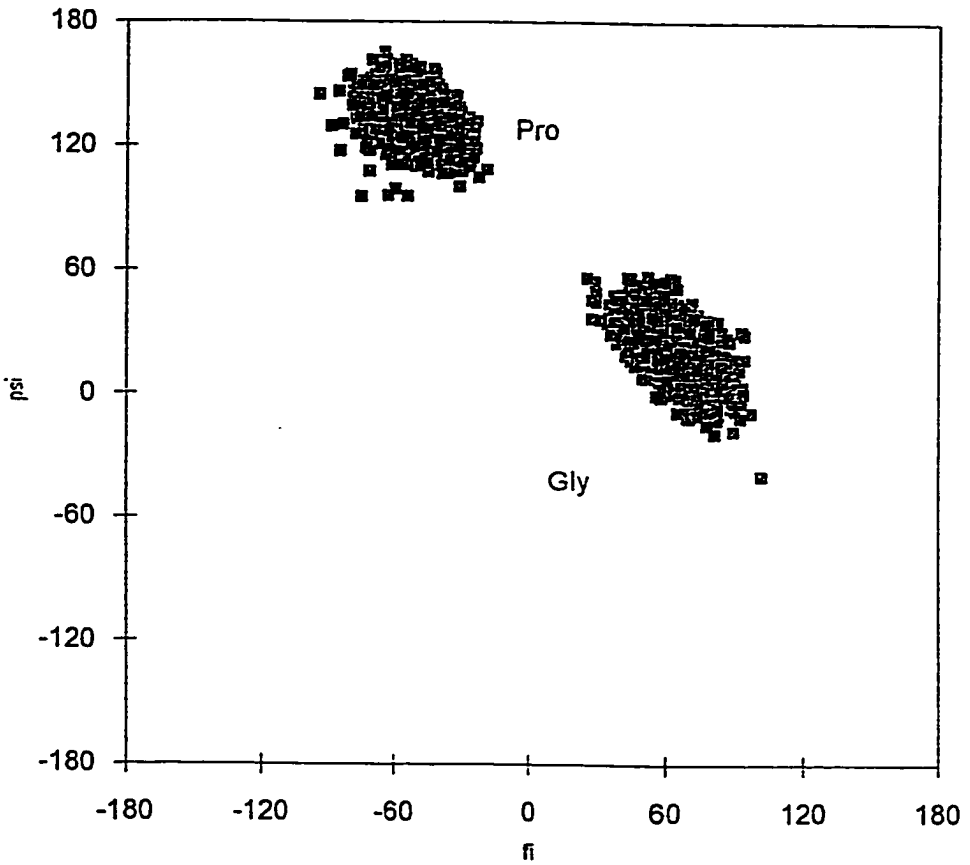
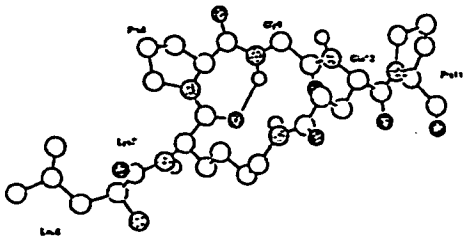
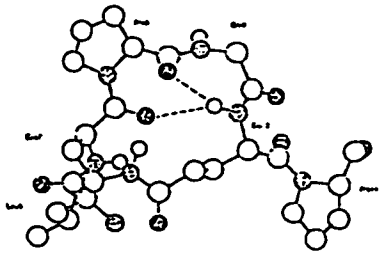


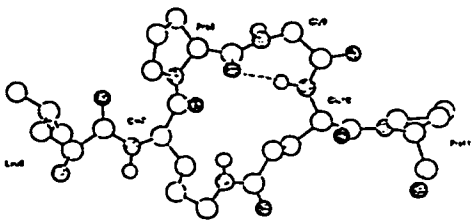
Figure 22d. Ramachandran Maps of the backbone dihedral angles of the DMSO dynamics structure C12; Dihedral angles (ϕ,ψ) for the Pro and/or Gly residues from 1000 discrete conformers observed during a 60 ps simulation are included in the plots.



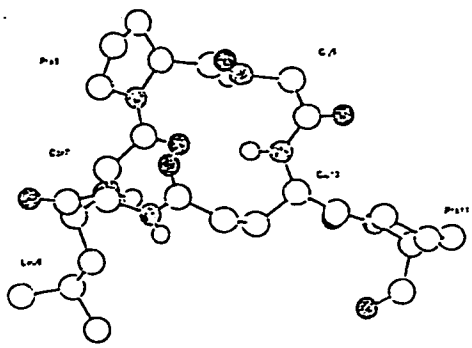
C42



C23



C32



C12

Figures 23a to 23d. Models of the likely DMSO solution conformations of the constrained regions of the cyclized pheromone analogs, as resulting from the usage of physical chemical data in the molecular modeling studies.

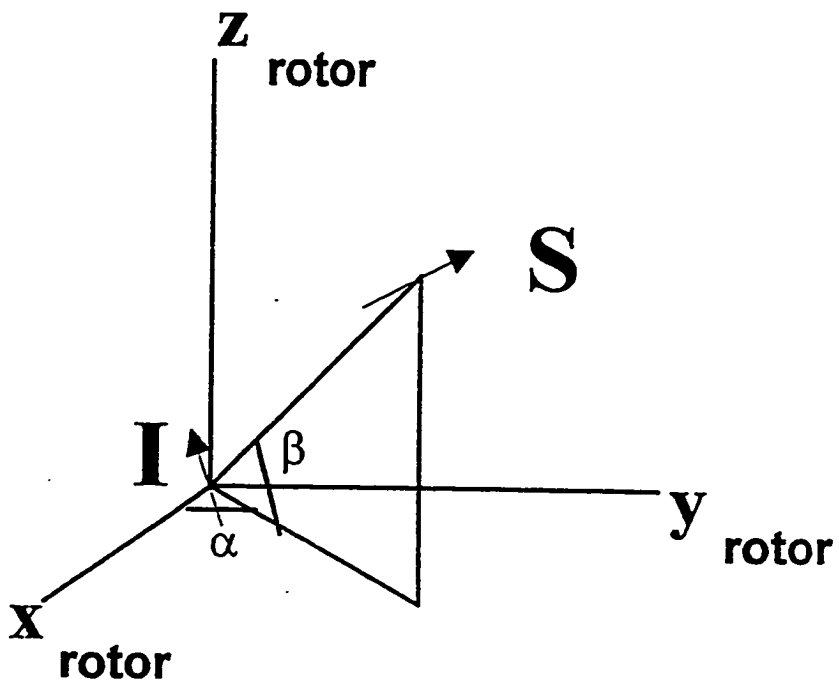


Figure 24. The two spins setup considered in the "classical" REDOR experiment is presented.

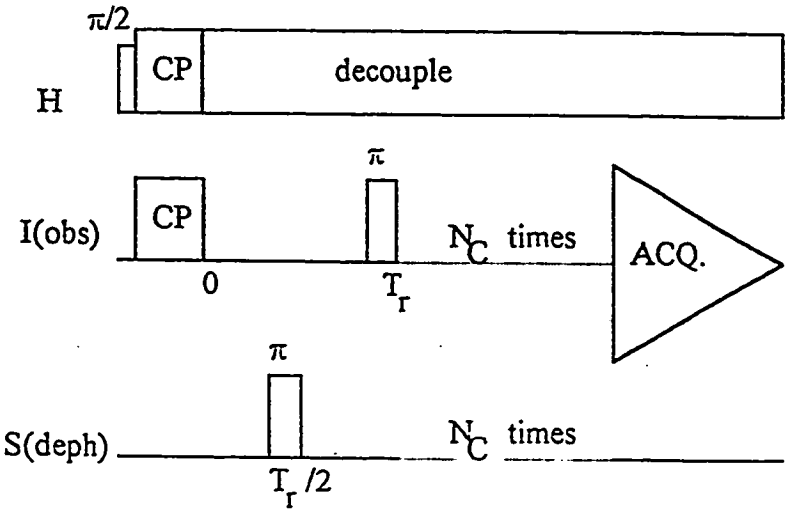


Figure 25. The REDOR pulse sequence

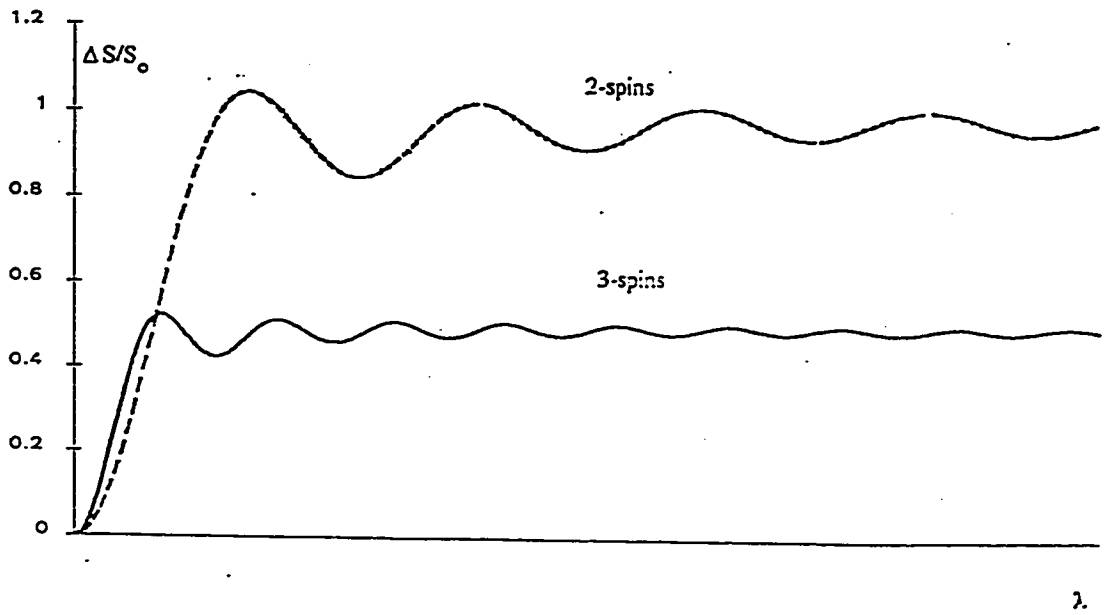


Figure 26. A REDOR effect simulation curve.

Figure 27. Three spin effect as a function of the γ angle, for different values of the REDOR parameter.

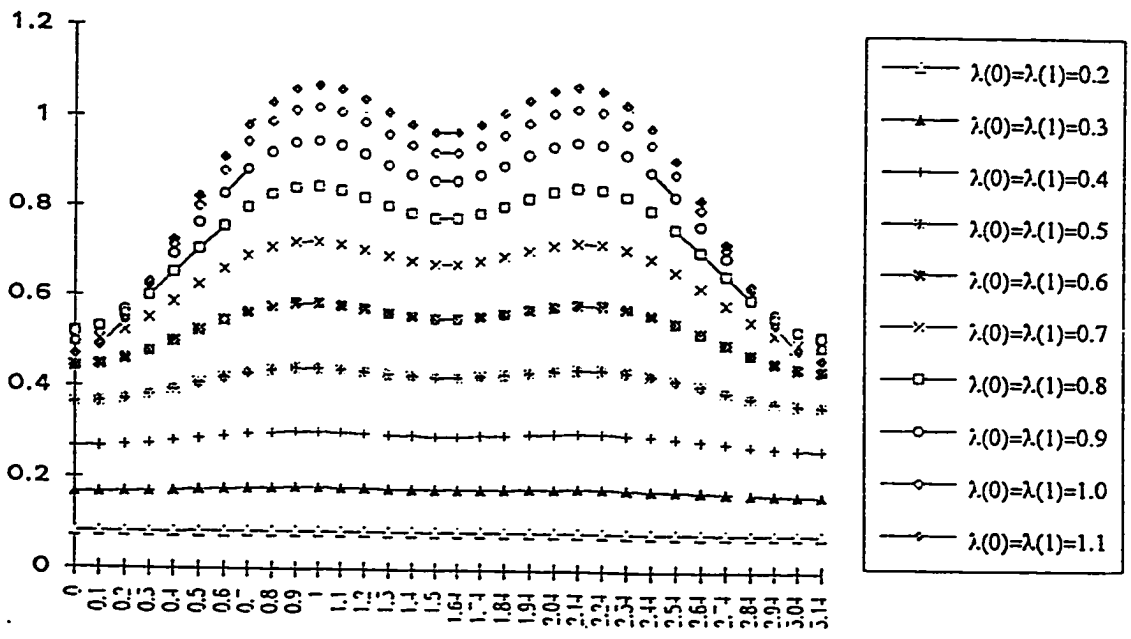
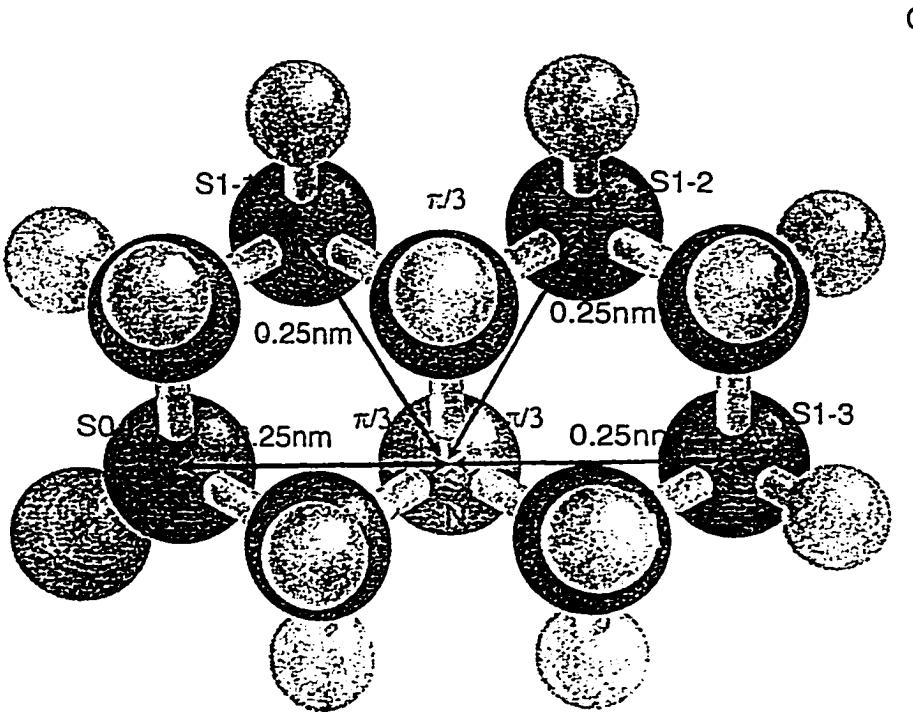
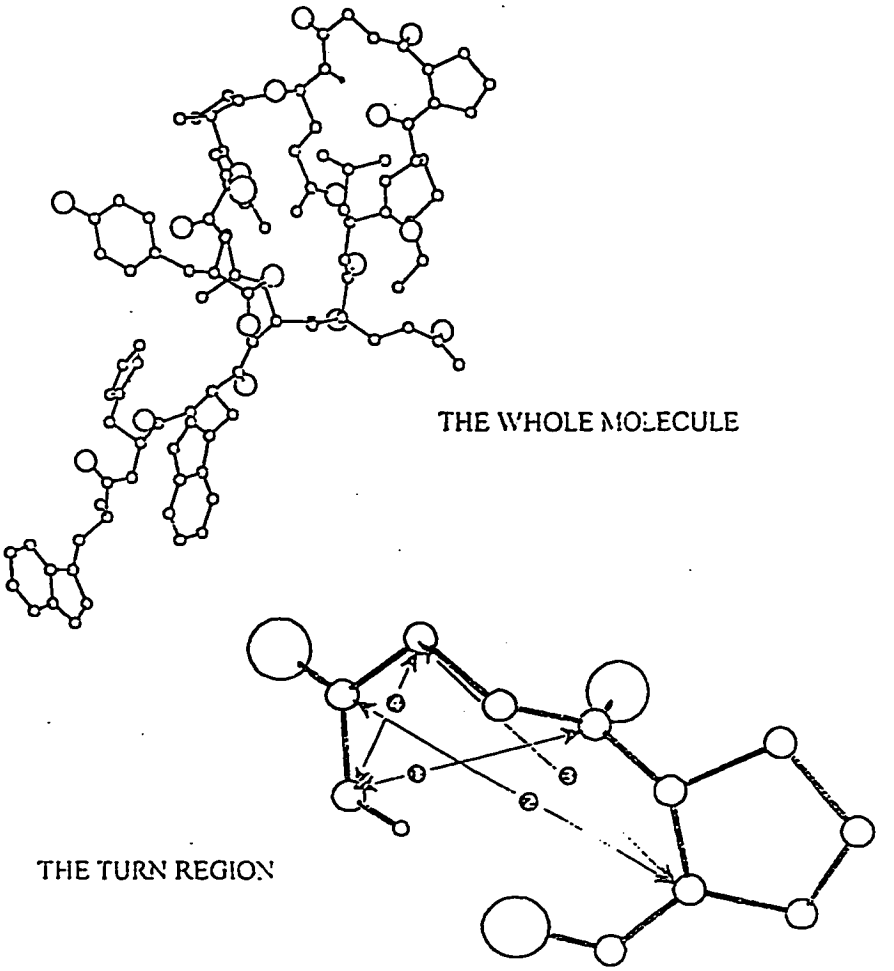


Figure 28. Some labeled compounds that might be used to test the 3-spins REDOR experiment proposed in this thesis could contain the following structure.





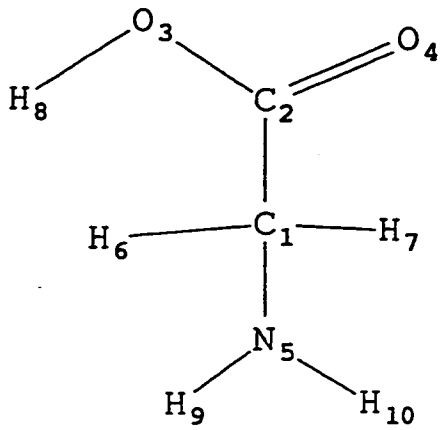
CHARMm22 MINIMIZED MOLECULAR MODEL OF THE α-FACTOR
WITH TYPE II β-TURN AT RESIDUES 8 AND 9 INITIAL CONDITIONS

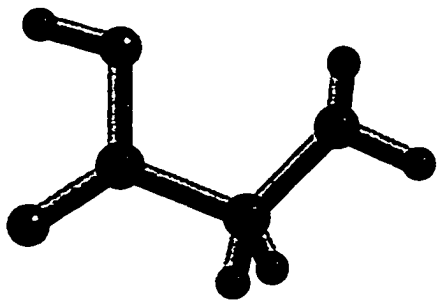
Figure 29

The design of a REDOR experiment. For purposes of calibration, one distance must be known: in our case, distance (4) is 2.54Å irrespective of the peptide conformation. The other distances must be chosen so as to distinguish between possible conformations. The four labeled compounds described in chapter 9 were designed to measure the four distances explicitly exhibited in this figure on a type II β-turn conformation (which is better suited for examination.)

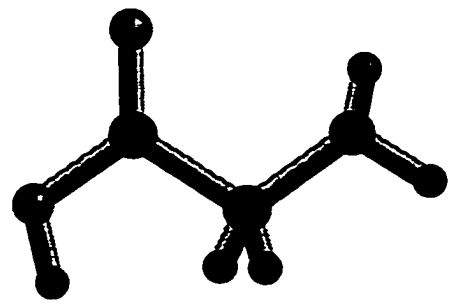
Figure 30

Numbering of the atoms in the *ab initio* calculations on glycine considered in chapter 10.



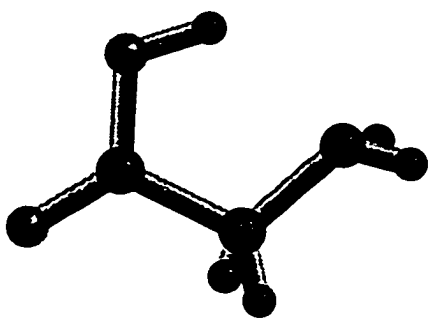


Structure 2



Structure 1

Structure 3



Structure 4

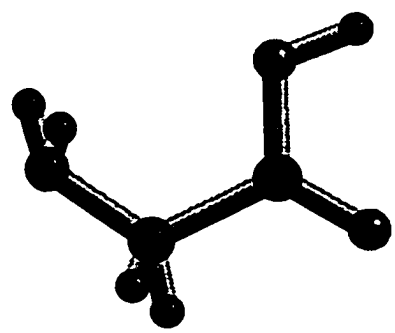


Figure 31
The optimized conformations found for glycine using four different levels of perturbation theory.

REFERENCES

1) Naider, F. and Becker, J.M. (1986) *CRC Critical Reviews in Biochemistry*, **21**, issue 3, 225-248;

2) Sprague, G. F. and Thorner, J. (1993) in *The Molecular Biology of the Yeast Saccharomyces cerevisiae*, (Broach J.R., Pringle J.R. & Jones E.W., Eds.) 2nd Ed., pp 657-744 Cold Spring Harbor Laboratory Press;

3) Gounarides, J. S. (1993) Ph.D. Thesis presented at the City University of New York in August;

4) Naider, F.; Gounarides, J.; Xue, C.-B.; Bargiota, E. and Becker, (1992) J.M. *Biopolymers*, *32*, 335-339;

5) Shenbagamurthi, P.; Kundu, B.; Raths, S.K.; Becker, J.M. and Naider, F. (1985) *Biochemistry*, *24*, 7070;

6) Momany F. A. and Rone R. (1992) J. Comp. Chem, **13**, pp888-900;

7) Yang, W.; McKinney, A.; Becker, J. M.; Naider, F. (1995) *Biochemistry*, **34**, 1308-1315;

8) Marepalli, H. R., Yang, W.; Joshua, H., Becker, J. M.; Naider, F. (1995) *Int. J. Peptide Protein Res.*, **45**, 418-429;

9) Stawarz, B, Genest, M., Genest, D. (1992) *Biopolymers*, **32**, 633-642;

10) Chung, J., Tolman J. R., Howard, K.P. and Prestegard, H.J. (1993) *J. Magn. Res.* **B102**, 137-147;

11) Stradley, S.J., Rizo, J., Bruch, M. amd Gierasch, L.M. (1990) *Biopolymers*, **29**, 263-287;

12) MacKay, V., Manney, T. R. (1974) *Genetics*, **76**, 273-288;

13) Hagen, D. C., McCaffrey, G., Sprague, G.G., Jr. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1418-1422;

14) Blumer K.J. and Thorner (1991) *J. Ann. Rev. Physiol.* **53**, 37;

15) Godchaux, W., Zimmerman, W. F. (1979) *J. Biol. Chem.*, **254**, 7874-7884;

16) Sibley, D.R., Renovic, J.L., Caron, M.G., Lefkowitz, R.J. (1987) *Cell*, **48**, pp913-922;

17) Moscona-Amir, E., Henis, Y. I., Yechiel, E., Barenholz, Y., Sokolovsky, M. (1986) *Biochemistry*, **25**, 8118-8124;

18) Rizo, J. amd Gierasch, L.M. (1992) *Annu. Rev. Biochem.*, pp387-418,;

19) Monahan, M. W., Amoss, M. S., Anderson, H. A. (1973) *Biochemistry* **12**, pp 4616-4620;

20) Friedinger, R.M., Veber, D.F., Perlow, D.S., Brooks, J.R., Saperstein, R. *Science*, **210**, pp656-658

21) Coy, D. H., Jiang, N.-Y., Kim, S. H., Moreau J.-P., Lin J.-T., Frucht H., Qian, J.-M., Wang, L.-W., Jensen, R. T. (1991) *J. Biol. Chem.* **266**, pp16411-16447;

22) Sawyer, T.K., Sanfilippo, P.J., Hruby, V.J., Engel, M.H., Heward, C.B. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5754-5758;

23) Sawyer, T.K., Hruby, V.J., Darman, P.S., Hadley, M.E. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1751-1755;

24) Hill, P.S., Smith, D.D., Slaninova, J., Hruby, V.J., (1990) *J. Am. Chem. Soc.* **112,** 3110-3113; Hruby V.J. and Smith C.W. (1987) Structure-Activity Relationships of Neurohypophyseal Peptides, Chapter 4 in The Peptides, Academic Press Inc. (London) Vol. **8**, pp77-207;

25) Venkatachalam, C. M. (1968) *Biopolymers*, *6*, 1425;

26) Abel, M.G., Zhang, Y.L., Lu, H.-S., Naider, F. and Becker, J.M. submitted to *Biochemistry*.

27) Gullion, T. and Schaefer, J. (1989) *Adv. Mag. Res.*, **13**, 57.

28) Levine, I. N., (1983) *Quantum Mechanics*, Third Edition, Allyn and Bacon, Boston, London, Sydney, Toronto;

29) Sapse A.M.; Mallah-Levy, L.; Daniels, S.B. and Erickson, B.W. (1987) *J. Am. Chem. Soc.* **109**, 3526;

30) Sapse, A.M., Daniels,S.B. and Erickson, B.W. (1988) *Tetrahedron* **44**, No.3, 999;

31) Schäfer, L., Newton, S.Q., Cao, M., Peeters, A., Alsenoy, C. Van, Wolinski, K. and Momany, F.A. (1993) *J. Am. Chem. Soc.* **115**, 272-280;

32) Insight manuals (1995) Biosym Corporation;

33) Sybyl manuals (1995) Tripos Corporation;

34) Neuhaus, D., Williamson, M. (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH: Weinheim

35) Matter, H., Kessler, H. (1995) *J. Am. Chem. Soc.* **117**, 3347-3359;

36) Mierke, D. F. and Kessler H, (1992) Biopolymers, **32**, pp877-882;

37) Bax, A. and Davis, D. G. (1985) *J. Magn. Reson.* **63**, 207-213;

38) Kataoka T., Beusen D.B., Clark J.D., Yodo M., Marshall G.R. (1992) Biopolymers, **32**, pp1519-1533;

39) Momany F. A. and Rone R. (1992) J. Comp. Chem, **12**, pp326-341;

40) Momany F.A., Rone R., Kunz H., Frey R.F., Newton S.Q. and Schäfer, L. (1993) J. Comp. Chem **15**, 3, 435-449;

41) Jelicks, L.A., Naider, F.R., Shenbagamurthi, P., Becker. J.M. and Broido, M.S. (1988) *Biopolymers*, **27**, 431;

42) Gounarides, J., Broido, M.S., Becker. J.M. and Naider, F.R. (1993) *Biochemistry*, **32**, 908-917;

43) Richardson, J. S. and Richardson, D. C. (1990) Principles and Patterns of Proteins Conformation pp 1-98 in Prediction of Protein Structure and the Principles of Protein Conformation, G. D. Fasman ed., Plenum Press, New York and London, Second Edition;

44) Gaussian 90, Revision I, Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S. and Pople, J. A. (1990) Gaussian, Inc., Pittsburgh PA,;

45) Schlegel, H.B. (1982) *J. Comput. Chem.*, **3**, 214;

46) Benedetti, E. (1977) *Fifth American Peptide Symposium* pp 257-273;

47) Antohi, O., Naider, F. and Sapse, A.-M. (1996) *J. Molec. Struct. (Theochem)* **360** 99;

48) Marepalli, H.R., O. Antohi, Becker, J.M. and Naider, F. (1996) *J. Am. Chem. Soc.*, **118**, 6531;

49) Antohi, O., Marepalli, H.R., Yang, W., Becker, J.M. and Naider, F. (1997) *Biopolymers*, **45**, 21;

50) Garbow, J. R., Breslav, M., Antohi, O. and Naider, F. (1994) *Biochemistry*, **33**, 10094-10099.

51) Garbow, J. R. and McWherter C., *J. Am. Chem. Soc.*, **115**, 238, (1993).

52) QUANTA3.2 (1993) *User Manuals*, Molecular Simulations Inc.;

53) Jeener, J. (1971) *Ampère International Summer School*, Basko Polje, Yugoslavia; Croasmun, W. R. and Carlson, R. M. K. (1987) *Two-Dimensional NMR Spectroscopy*, VCH Publishers

54) L. Braunschweiler and R. R. Ernst (1983) *J. Magn. Reson.* **53**, 521; A. S. Edison, F. Abildgaard, W. M. Westler, E. S. Moobery, and J. L. Markley (1994) in *Methods in Enzymology* **239C**, 3

55) J. L. Markley and M. Kainosho (1993) in *NMR of Biological Molecules* G. C. K. Roberts, ed p101, Oxford Press, New York

56) A. S. Edison, F. Abildgaard, W. M. Westler, E. S. Moobery, and J. L. Markley (1994) in *Methods in Enzymology*, **239C**, 3

57) Bothner-By A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* (1984) **106**, 811-813;

58) Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R. R. (1987) *J. Am. Chem. Soc.*, **109**, 607-609;

59) Griesinger, C.; Ernst, R. R. (1987) *J. Magn. Reson.* **75**, 261-271;

60) Kessler, H. *Angew. Chem. Int. Ed. Eng.*, (1982) **21**, 512-523;

61) Hruby, V. J. (1982) *Life Sci., 31*, 189-199;

62) De Grado, W. F. (1988) *Adv. Protein Chem., 39*, 51-123;

63) Bitar, K. G., Somogyvari-Vigh, A., Coy, D. H., *Peptides*, (1994) *15*, 461-466;

64) Farlie, D. P.; Abbenate, G.; March, D. R. (1995) *Current Med. Chem., 2,* 654-686;

65) Sawyer, T. K. (1995) *Adv. Peptide-Based Drug Design,* Taylor, M. D.; Amidon, G. L., Eds.; ACS Prof. Ref. Books, ACS, Washington, D. C. pp 387-422;

66) Gulays, J., Rivier, C., Perrin, M.; Koerber, S. C., Sutton, S., Corrigan, A., Lahrichi, S. L., Craig, A. G., Vale, W., Rivier, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 10575-10579;

67) Arttamangkul, S., Murray, T. F., Delander, G. E., aldrich, J. V. (1995) *J. Med. Chem. 38*, 2410-2417

68) Gilon, C., Halle, D., Chorev, M., Selinger, Z., Byk, G. (1991) *Biopolymers*, **31**, 745-750;

69) Hermkens, P. H. H., Dinther, T. G. V., Joukema, C. W., Wagenaars, G. N., Ottenheijm, H. C. J. (1994) *Tet Lett. 35*, 9271-9274;

70) Bankowski, K., Manning, M., Seto, J., Haldar, J., Sawyer, W. H. (1980) *Int. J. Peptide Protein Res., 16,* 382-391.

71) Sawyer, T. K.; Hruby, V. J.; Darman, P. A.; Hadley, M. E. (1982) *Proc. Natl. Acad. Sci. U.S.A., 79,* 1751-1755;

72) Schiller, P. W., Nguyen, T. M. D., Lemieux, C., Maziak, L. A. (1985) *J. Med. Chem., 28,* 1766-1771;

73) Schiller, P. W., Nguyen, T. M. D., Maziak, L. A., Wilkes, B. C., Lemieux, C. (1987) *J. Med. Chem., 30,* 2094-2099;

74) Mierke, D. F. Schiller, P. W., Goodman, M. (1990) *Biopolymers, 29,* 943-952;

75) Felix, A. M.; Heimer, E. P.; Wang, C. T.; Lambros, T. J.; Fournier, A.; Mowels, T. F.; Maines, S., Campbell, R. M.; Wegrzynski, B. B.; Toome, V.; Fry, D.; Madison, V. S. (1988) *Int. J. Peptide Protein Res., 32,* 441-454;

76) Fry, D. C.; Madison, V. S.; Greeley, D. N.; Felix, A. M.; Heimer, E. P.; Frohman, L.; Campbell, R. M.; Mowels, T. F.; Toome, V.; Wegrzynski, B. B. (1992) *Biopolymers, 32,* 649-666.

77) Kapurniotu, A.; Taylor, J. W. (1995) *J. Med. Chem. 38,* 836-847;

78) Chorev, M., Roubini, E., McKee, R. L., Gibbons, S. W., Goldman, M. E., Caulfield, M. P., Rosenblatt, M. (1991) *Biochemistry, 30,* 5968-5974;

79) Charpenteir, B., Pelaprat, D., Durieux, C., Dor, A., Reibaud, M., Blanchard, J. C. and Roques, B. P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1968-1972;

80) Charpenteir, B., Durieux, C., Pelaprat, D., Dor, A., Reibaud, M., Blanchard, J. C., Roques, B. P. (1988) *Peptides*, **9**, 835-841;

81) Charpenteir, B., Dor, A., Roy, P., England, P., Pham, H., Durieux, C., Roques, B. P. (1989) *J. Med. Chem.*, **32**, 1184-1190;

82) Ning, Q., Ripoll, D. R., Szewczuk, Z., Konishi, Y. and Ni, F. *Biopolymers*, (1994) **34**, 1125-1137;

83) Bienstock, R. J., Koerber, S. C., Rizo, J., Rivier, J. E., Hagler, A. T. and Gierasch, L. M. (1992) *Peptides : Chemistry and Biology, Proceedings of Twelfth American Peptide Symposium;* Smith, J. A. & Rivier, J. E., Ed.; ESCOM, Leiden, pp 262-264;

84) Reddy, D. V.; Jagannadh, B.; Dutta, A. S.; Kunwar, A. C. *Int. J. Peptide Protein Res.* (1995) **46**, 9-17;

85) States, D. J.; Haberkorn, R. A.; Ruben, D. J. (1982) *J. Magn. Reson.* **48**, 286-292

86) Pease, L. G., Deber, C. M. and Blout, E. R. (1973) *J. Am. Chem. Soc.* **95**, 258-260;

87) Rizo, J., Koerber, S. C., Bienstock, R. J., Rivier, J., Hagler, A. T. and Gierasch, L. M. (1992) *J. Am. Chem. Soc., 114*, 2852-2859;

88) Karplus, M. J., *J. Chem. Phys.* (1959) **30**, 11-15;

89) Bystrov, V. F., (1976) *Prog. Nucl. Magn. Reson. Spectrosc., 10*, 41-81;

90) Veber, D. F. (1981) *Peptides: Synthesis, Structure and Function,* Rich, D. H.; Gross, E.; Ed.; Pierce Chem. Co., Rockford, Illinois, pp 685-694;

91) Freidinger, R. M. (1981) *Peptides: Synthesis, Structure and Function*, Rich, D. H., Gross, E. Ed., Pierce Chem. Co., Rockford, Illinois, pp 673-683;

92) Sawyer, T. K., Cody, W. L., Knittle, J. J., Hruby, V. J., Hadley, M. E., Hirsch, M. D., O'Donohue, T. L. (1983) *Peptides: Structure and Function* Hruby, V. J.; Rich, D. H. Ed.; Pierce Chem. Co., Rockford, Illinois, pp 323-331;

93) Schiller, P. W. (1984) *The Peptides* Udenfriend, S.; Meienhofer, J. Ed.; Academic Press, NY, **6**, pp 219-268;

94) Hruby, V. J., Al-Obeidi, F., Kazmierski, W. M. (1990) *Biochem J., 268,* 249-262;

95) Rizo, J.; Gierasch, L. M. (1992) *Ann. Rev. Biochem., 61,* 387-418;

96) Chou, P. Y.; Fasman, G. D. (1978) *Adv Enzymol., 47,* 45-148;

97) Gierasch, L. M., Deber, C. M., Madison, V., Niu, C-H. and Blout, E. R. (1981) *Biochemistry, 20,* 4730-4738;

98) (b) Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* John Wiley & Sons, Inc., New York, pp 117-129;

99) Nemethy, G.; Printz, M. P. (1972) *Macromolecules, 5,* 755-758;

100) Smith, J. A. and Pease, L. G. (1980) *Critical Reviews in Biochemistry,* **15,** 315-399.

101) Felix, A. M., Heimer, E. P., Wang, C. -T., Lambros, T. J., Fournier, A., Mowels, T. F., Maines, S., Campbell, R. M., Wegrzynski, B. B., Toome, V., Fry, D. C. and Madison, V. S. (1988) *Int. J. Peptide Protein Res., 32,* 441-454;

102) Felix, A. M., Wang, C. -T., Campbell, R. M., Toome, V., Fry, D. C. and Madison, V. S. (1992)

Peptides: Chemistry and Biology, Proceedings of Twelfth American Peptide Symposium; Smith, J. A. and Rivier, J. E., Ed.; ESCOM, Leiden, pp 77-79;

103) Osapay, G. and Taylor, J. W. (1990) *J. Am. Chem. Soc.*, **112**, 6046-6051;

104) Osapay, G., Gulyas, J., Profit, A. A., Gulyas, E. S. and Taylor, J. W. (1992) *Peptides : Chemistry and Biology, Proceedings of Twelfth American Peptide Symposium;* Smith, J. A.; Rivier, J. E., Ed.; ESCOM, Leiden, pp 239-240;

105) Houston, Jr M. E.; Gannon, C. L.; Kay, C. M.; Hodges, R. S. (1995) *J. Peptide Sci.* **1**, 274-282;

106) Loumaye E., Thorner, J. & Catt,K.H., (1982) *Science,* **218,** 1323-1325;

107) Wakamatsu, K., Okada, A., Suzuki, M., Higashijima, T., Masui, Y., Sakakibara, S. & Miyazawa, T. (1986) *Eur. J. Biochem.* **154**, 607-615;

108) Wakamatsu, K., Okada, A., Miyazawa, T., Masui, Y., Sakakibara, S., & Higashijima, T., (1987) *Eur. J. Biochem.* **163**, 331-338;

109) Gounarides, J.S., Xue, C.-B., Becker, J.M., & Naider, F. (1994) *Biopolymers* **34**, 709-720;

110) Bax, A. & Davis, D. G. (1985) *J. Magn. Reson.* **65**, 355-360;

111) Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & WÜthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117**, 479-485;

112) Macura, S., & Ernst, R. R. (1980) *Mol. Phys.* **41**, 95-117;

113) T. Gullion and J. Schaefer, (1989) *J. Magn. Reson.*, **81**, 196;

114) McDowell, L. M., and Schaefer, J. (1996) *Current Opinion Struct. Biol.*, **6**, 624;

115) McDowell, L. M., Klug, C. A., Beusen, D. D. and Schaefer, J. (1996) *Biochemistry*, **35**, 5395;

116) A. W. Hing and Schaefer, J. *Biochemistry*, **32**, 7593 (1993);

117) Marshall, G. R., Beusen, D. D., Kociolek, K., Redlinski, A. S., Leplawy, M. T., Pan Y. and Schaefer J., (1990) *J. Am. Chem. Soc.*, **112**, 963;

118) Mueller, K. T., (1995) *J. Magn. Reson.*, **A113**, 81;

119) Mehring M., High Resolution NMR in Solids, Springer-Verlag, Berlin - Heidelberg - New York (1983);

Naito, A., Nishimura, K., Tuzi S. and Saito, H. (1994) *Chem. Phys. Lett.*, **229**, 506;

120) Maricq, M. M. and Waugh, J. S. (1979) *J. Chem. Phys.* **70**, 3300;

121) Brink, D. M., and Satchler, G. R. (1968) *Angular Momentum*, Oxford: Clarendon Press;

122) Gradshtein, I. S. and Ryzhyk, I. M. (1980) *Table of Integrals, Series and Products*, 4th edition, Academic Press, New York London Toronto Sydney San Frncisco;

123) Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T., (1988) *Numerical Recipes in C*, Cambridge University Press, New York - New Rochelle - Melbourne - Sydney;

124) Levitt, M. H., Raleigh, D. P., Creuzet F. and Griffin, R. G. *J. Chem. Phys.*, **92** (11), 6347 (1990).

125) Weintraub, O., Vega, S., Hoelger, Ch. and H. H. Limbach (1994) *J. Magn. Reson.*, **A109**, 14;

126) Beachy, M. D., Chasman, D., Murphy, R.B., Halgren, T. A. and Friesner, R. A. (1996) *Journal of the American Chemical Society*, 119, 5908;

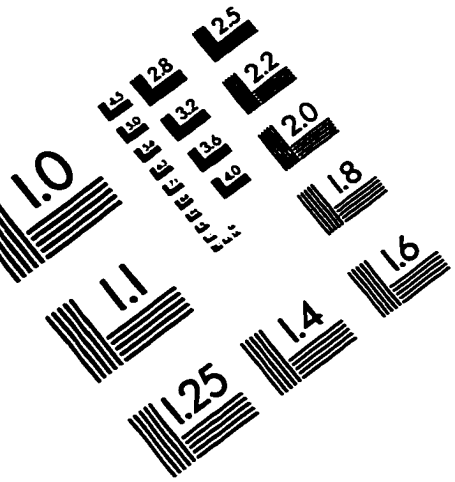
127) Ramek, M. (1990) *International Journal of Quantum Chemistry*, **17**, 45-53;

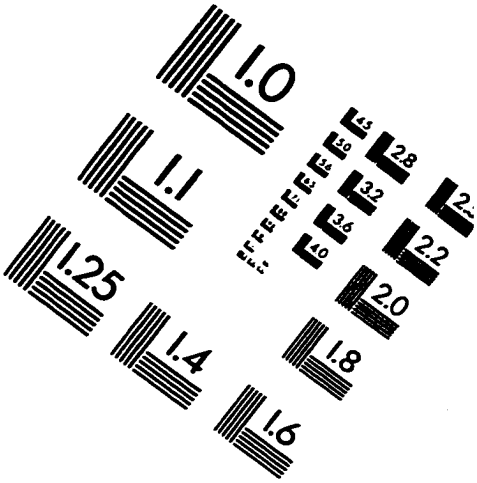
128) Ramek, M., Cheng, V. K. W., Frey, R.F., Newton, S.Q. and Schäfer, L. (1991), J. Molecular Structure, **235**, 1-10;

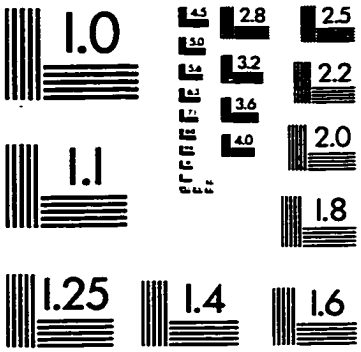
129) Møller, C. and Plesset, M. S. (1934) Phys. Rev. **46**, 618;

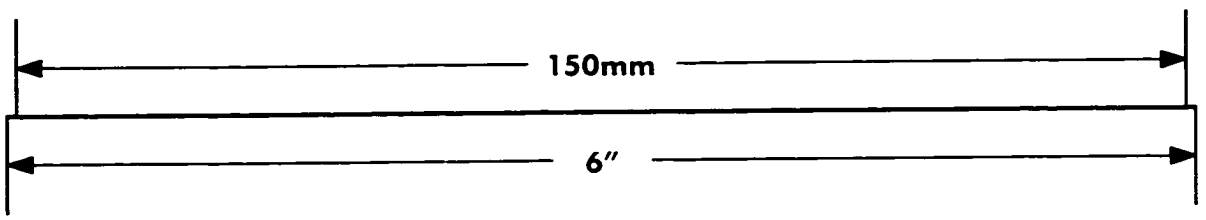
130) Goetz, J. M. and Schaefer, J., (1997) *J. Magn. Reson.*, **127**, 147-154.

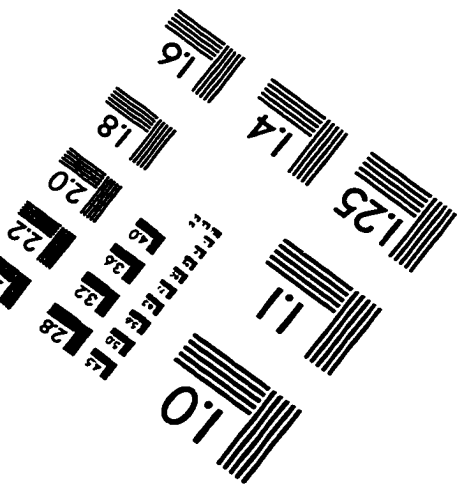
IMAGE EVALUATION TEST TARGET (QA-3)

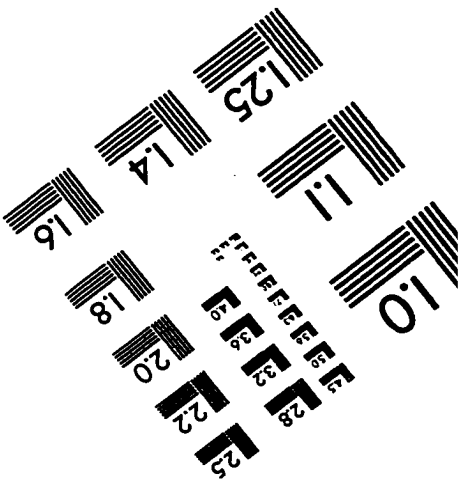












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989