PhD Thesis

Computational Studies on the Structure and Dynamics of Bioactive Peptides

by Francesc J. Corcho Sánchez

Barcelona, September 2003

ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA INDUSTRIAL DE BARCELONA
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Computational Studies on the Structure and Dynamics of Bioactive Peptides

PhD Thesis submitted to obtain the degree of
Doctor in Sciences

This work has been carried out at the Chemical Engineering department under the supervision of Prof. Juan Jesús Pérez González and Dr. Josep Cantó Silva

Signed:

Prof. Juan Jesús Pérez González **Dr. Josep Cantó Silva**

Barcelona, September 2003.

Dedicated to the ones I love

"When a theory is highly successful and becomes firmly established, the model tends to become identified with 'reality' itself and the model nature of the theory becomes obscured."

HUGH EVERETT
The Many-Worlds Interpretation of Quantum Mechanics

"Science is not the same thing as the reality it describes. There is always a gap between reality and the description of it."

TOM SIEGFRIED
The Bit and The Pendulum

"Viewing the world as a computer-analyzing natural phenomena in terms of information processing can provide a new framework explaining how the world works, just as the clock and steam engine have in the past.[…] Information is more than a metaphor – it is a new reality."

TOM SIEGFRIED
The Bit and The Pendulum

Agraïments

Aquesta tesi no hagués estat possible si no hagués estat per l'ajuda i el suport de molta gent. Voldria agrair a tots els que m'han ajudat a fer possible l'arribada a bon terme del repte que sempre suposa escriure una tesi. Si em deixo algú li demano per avançat les meves disculpes.

En primer lloc vull agrair als meus pares haver-me proporcionat la formació humana, espiritual i científica que ha sigut imprescindible per a la realització de la tesi. Vull agrair als meus germans haver-me servit com a referència i estímul en el meu esforç; a la meva esposa, Soraya, haver-me donat el seu suport i el seu amor a tothora i especialment en els moments més complicats i al meu fill Jan haver donat una nova embranzida a les meves determinacions i haver donat molts moments d'alegria al final d'aquest llarg procés. També vull agrair als meus sogres el seu tarannà afable i comprensiu i la meva neboda Sofia el ser tan eixerida i simpàtica. Els petits detalls són els que fan que la vida sigui molt més fàcil.

D'una manera molt significativa vull agrair al meu director de tesi, el professor Juan Jesús Pérez, per haver confiat en la meva feina i haver tingut la paciència d'analitzar els resultats amb deteniment. També vull remarcar que ha estat per a mi molt important disposar del seu coneixement i tenir-lo com a referent professional. Vull també agrair especialment al Dr. Josep Cantó, co-director d'aquesta tesi, per ajudar-me a arribar-hi al final. Ell ha exercit moltes vegades, sense saber-ho, del meu *alter ego* científic, escoltant les meves reflexions en veu alta. Les discussions mantingudes amb ell han sigut peça clau del desenvolupament d'aquest treball. De tota manera, qualsevol errada que hi trobeu en aquest treball són el resultat de la meva ignorància i no són en cap cas responsabilitat de qualsevol altra persona. Vull també agrair a l'Oriol Llorens per tots els bon moments que hem compartit durant aquest temps i per ésser un exemple de treball i constància i desitjo que ben aviat pugui incorporar-se de nou al grup. També voldria agrair a la Dra. Marta Filizola per donar-me les primeres lliçons de modelització molecular i de com convertir els resultats en un article. La seva col·laboració va ésser cabdal per a la conclusió de la primera part d'aquesta tesi. El Dr. Xavier Salvatella també mereix una menció especial pel seu ajut en la resolució dels espectres de ressonància magnètica nuclear dels pèptids estudiats.

Finalment voldria agrair a tots els meus companys del grup, Daniel Badia, l'Arnau Cordomí, el David Jímenez, la Marta Pinto i la Lourdes Roset, per haver fet que les hores de feina d'aquests anys resultessin molt més amenes.

Table of Contents

Abbreviations ..... **xi**

Preface ..... **xiii**

1. An Introduction to Molecular Modeling and Computer-aided Drug Design **1**
1.1 Molecular Modeling and Computational Chemistry 3
1.2 Quantum Mechanics and Molecular Mechanics 3
1.3 Force Fields 4
1.4 Energy-Minimizing Procedures 6
1.4.1 Steepest Descent Method 6
1.4.2 Conjugate Gradient Method 6
1.5 Use of charges and solvents 7
1.5.1 Periodic Boundary Conditions 7
1.5.2 Ewald Summation Techniques 9
1.5.3 Particle Mesh Ewald 10
1.6 Conformational Analysis of a Peptide: Multiple Minima Problem 12
1.7 Assessment of the Bioactive Conformation 14
1.8 Folding Studies on Peptides 16
1.9 Computer-aided Drug Design 16
1.10 References to chapter 1 21

2. Assessment of the Bioactive Conformation of the Farnesyltransferase Protein Binding Recognition Motif by Computational Methods and Evaluation of Iterative Simulated Annealing Technique in Conformational Search of Peptides **23**
2.1 Summary 25
2.2 Introduction 26
2.3 Methods 30
2.4 Results and Discussion 32
2.4.1 Conformational analysis of Ras farnesyltransferase inhibitors 32
2.4.2 Characterization of the bioactive conformation 39
2.4.3 Comparison of iterative SA and random conformational searches 42
2.4.4 Effect of temperature in the efficiency of SA methodology 45
2.5 Conclusions to chapter 2 47
2.6 References to chapter 2 49

3. Conformational Study of 1-amino-2-phenylcyclopropanecarboxylic Acid Derivatives and the Search of Novel Leads for Bradykinin Antagonism **51**
3.1 Summary 53
3.2 Introduction 53
3.3 Methods 56
3.4 Results and Discussion 60
3.4.1 Conformational Analysis of Hoe 140 60
3.4.2 B_2 Pharmacophore Revisited 63
3.4.3 Search of the Pharmacophore in Databases 66
3.4.4 Conformational Analysis of c_3Phe Derivatives of BK 68
3.4.5 Pharmacophore Fulfillment by c_3Phe Fragments of BK 80
3.5 Conclusions to chapter 3 82
3.6 References to chapter 3 85

4. Comparative Analysis of the Conformational Profile of Substance P using Simulated Annealing and Molecular Dynamics **89**
4.1 Summary 91
4.2 Introduction 91
4.3 Methods 92

4.3.1 Simulated Annealing Calculations 93
4.3.2 Molecular Dynamics Calculations 93
4.3.3 Conformation Classification 94
4.3.4 Clustering of the structures 97
4.3.5 Transtition Analysis 99
4.4 Results 100
4.4.1 Iterative simulated annealing of the peptide 101
4.4.3. Molecular Dynamics in a water box at 300 K 121
4.4.4. Molecular Dynamics in a water box at 400 K 133
4.5. Discussion 144
4.6. Conclusions to chapter 4 150
4.7 References to chapter 4 152

5.Conformational Study of Substance P through Molecular Dynamics and NMR Spectroscopy in Different Solvents 155
5.1 Summary 157
5.2 Introduction 157
5.3 Methods 159
5.3.1 NMR experiments 159
5.3.1.1. NMR structure determination in water 159
5.3.1.2. NMR structure determination in methanol 159
5.3.2. Molecular dynamics 159
5.3.2.1. Molecular dynamics of SP in water 160
5.3.2.2. Molecular dynamics of SPOH in water 161
5.3.2.3. Molecular dynamics of SPOH in methanol 162
5.3.2.4. Molecular dynamics of SPOH in DMSO 165
5.3.2.5 Conformation classification 167
5.3.2.6. Clustering of the structures 167
5.3.2.7 Transition analysis 167
5.4. Results 168
5.4.1 NMR experiments 168
5.4.1.1. Study of SPOH in water 168
5.4.1.2. Study of SPOH in methanol 170
5.4.2. Molecular dynamics 174
5.4.2.1. Molecular dynamics of SP in water 174
5.4.2.2. Molecular Dynamics of SPOH in water 191
5.4.2.3. Molecular dynamics of SPOH in methanol 204
5.4.2.4. Molecular dynamics of SPOH in DMSO 212
5.6. Discussion 222
5.7. Conclusions to chapter 5 236
5.8 References to chapter 5 238

6. General Conclusions 241

Annex. Results of the Probability Distribution Function of Dihedral Angles for the Simulations on Substance P 247

Abbreviations

3D	Three-dimensional
Å	Angstrom
Ac_3c	1-aminocyclopropanecarboxylic acid
Aic	2-aminoindane-2-carboxylic acid
Ala (A)	Alanine
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
BK	Bradykinin
C	Carbon
c_3Phe	1-amino-2-phenylclyclopropanecarboxylic acid
Cys (C)	Cysteine
D	Dalton
DMSO	Dymethil sulfoxide
Eq.	Equation
FFT	Fast Fourier Transform
fs	Femtosecond (10^{-15} seconds)
FT	Farnesyltransferase
FTI	Farnesyltransferase inhibitor
GGT	Geranylgeranyltransferase
Gln (Q)	Glutamine
Glu (E)	Glutamic acid
Gly (G)	Glycine
H	Hydrogen
His (H)	Histidine
Hyp	Hydroxyproline
IC_{50}	Inhibitory concentration 50%
Ile (I)	Isoleucine
K	Kelvin
kcal	Kilocalories
Leu (L)	Leucine
Lys (K)	Lysine
MD	Molecular dynamics
MEP	Molecular electrostatic potential
Met (M)	Methionine
N	Nitrogen

NMR	Nuclear magnetic ressonance
NOE	Nuclear Overhauser effect
ns	Nanosecond (10^{-9} seconds)
O	Oxygen
Oic	(2S, 3aS, 7aS)-octahydroindole-2-carboxylic acid
PBC	Periodic boundary conditions
Phe (F)	Phenylalanine
PME	Particle Mesh Ewald
Pro (P)	Proline
ps	Picosecond (10^{-12} seconds)
RMS	Root mean square
SA	Simulated annealing
SDS	Sodium dodecyl sulfate
Ser (S)	Serine
SP	Substance P (amidated at C-terminal)
SPOH	Free acid form of substance P (deamidated C-terminal)
TFE	Trifluoethanol
Thi	β-(2-thienyl)-alanine
Thi	β-(2-thienyl)-alanine
Thr (T)	Threonine
Tic	1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid
Trp (W)	Tryptophan
Tyr (Y)	Tyrosine
Val (V)	Valine
Xaa (X)	Unspecified aminoacid
ε	Dielectric constant

Preface

The present work focuses on the exploration of the conformational space of biological active peptides in different conditions with the aim of characterizing their conformational profile. Different techniques have been used within the molecular mechanics framework. A first hurdle is encountered in the exploration as the exhaustiveness of the exploration and the definition of a criterion for stopping the procedure were not well defined at the beginning of the present work. A solution to this problem is presented in chapter 2 for the iterative simulated annealing.

Determining the bioactive conformation is a requirement for the design of peptidomimetics in computer-aided drug design. The bioactive conformation can be simplified by the moieties that are known to interact with the receptor and the relative distances between these moieties, and this schematic entity is termed pharmacophore. The pharmacophore can be used to screen three-dimensional databases of molecules for the search of peptidomimetics. The compounds obtained in the search can be subsequently tested for activity and a new group of lead compounds can be thus identified. In chapter 3 an example of this procedure is described with the aim of obtaining a new group of antagonists of bradykinin, a peptide hormone 9-residue-long.

Peptides have been traditionally considered as flexible molecules especially in polar solvents like water. This flexibility is difficult to measure by experimental techniques and as a consequence peptides have been regarded as molecules lacking of structure under such conditions. On the other hand, biological active peptides are known to interact with their receptors in a preferred conformation, often termed as the bioactive conformation. The most accepted hypothesis for explaining the interaction of peptides and their receptors is the induced fit. Thus, the peptide will exhibit in solution conformational motives that are part of the conformation adopted in the receptor. Subsequent to the binding of the peptide to the receptor, the conformation in solution will be modified in order to optimize the interaction of the peptide to the receptor. Therefore, a contradiction appears to exist between the need for certain degree of structure in the peptide prior to the receptor binding and the inherent lack of structure of peptides in solution. It has been argued in some instances that the binding of the peptide to the biological membrane is a prerequisite for the adoption of the bioactive conformation and the subsequent binding to the receptor. In chapters 4

and 5 this hypothesis will be criticized and an alternative hypothesis will be presented.

Recently, it has been reported several instances were the folding of peptides has been simulated by means of long molecular dynamics trajectories. A problem arises when the wealth of conformations obtained has to be classified in terms of their respective degree of folding. In chapter 4 a novel methodology is described for the classification and grouping of the peptide structures based on the presence of structural motifs and the similarities among them. This methodology can prove very useful as it almost automated and it does not present any limitation regarding the size of the peptide or protein of study.

In order to follow what are the problems arising when the size and the flexibility of peptides are increased, the sequence of chapters of the present study is presented with increasing size. Thus, in chapter 2 the conformational profile of 4-residue long farnesyltransferase inhibitors is studied by means of the iterative simulated annealing procedure. The first part of chapter 3 deals with a bradykinin antagonist, Hoe 140. Although Hoe 140 with 10 residues is larger than the 4-residue farnesyltransferase inhibitors, the conformational diversity of the peptide is only considered for the last 5 residues. This simplification of the peptide is carried out in order to compare with the conformational profile obtained for the group of bradykinin analogs presented in the second part of the chapter: 5-residue long peptides containing 1-amino-2-phenylclyclopropanecarboxylic acid, a conformationally restricted residue. Finally, chapters 4 and 5 are dedicated to a 11-residue neuropeptide: substance P. The increase in size provoked a change in the methodology. Indeed, the conformational profile of the peptide has been studied by means of iterative simulated annealing and extensive molecular dynamics trajectories. This has permitted the comparison between both methodologies and to derive conclusions to the kind of information that can be obtained through these different methodologies for the exploration of the conformational space of peptides.