Chapter 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

Note: *The work of the present chapter has been published in two separate papers (Corcho et al. 1999 and Corcho et al., 2000)*

2.1 Summary

Ras farnesyltransferase catalyzes the carboxyl-terminal farnesylation of Ras as well as other proteins involved in signal transduction processes. Previous studies demonstrated that its inhibition suppresses the activity of Ras transformed phenotypes in cultured cells, causing tumor regression in animal models. This observation led to the consideration of farnesyltransferase as a target for cancer therapy. The goal of the present work is to assess the bioactive conformation of the peptide Cys-Val-Phe-Met, known to be the minimum peptide sequence that inhibits farnesyltransferase through computational methods. In order to attain this objective, the conformational preferences of four analogs of the peptide were assessed by means of thorough search of their respective conformational spaces, using a simulated annealing as sampling technique. Specifically, the two active analogs: Cys-Val-Tic-Met and Cys-Val-Ψ(CH_2NH)Tic-Met and the two inactive analogs: Cys-Val-Tic-Ψ(CH_2NH)Met and Cys-Val-Aic-Met were selected for the present study. Low energy conformations of the four analogs were classified according to their structural motifs. The putative bioactive conformation of the minimum farnesyltransferase recognition motif was assessed by cross-comparison of the different classes of conformations obtained for the two active and the two inactive analogs. The putative bioactive conformation can be described by two structural motifs: i) a C14 pseudo-ring stabilized by a hydrogen bond between the amino group of Cys^1 and the carboxylate group of Met^4 and a C11 pseudo-ring involving the residues Cys^1 and Tic^3. In addition, the thiol group of Cys^1 side chain of the bioactive conformation points to the carboxylate moiety of Met^4.

Characterization of the subset of low energy minima of a peptide is hampered by the multiple minima problem associated to the roughness of its potential energy surface. In order to demonstrate that the iterative simulated annealing procedure is an effective procedure to overcome these difficulties, the exploration of the conformational space of the peptide Cys-Val-Tic-Met was carried out using the simulated annealing procedure and a random search method. Profile differences were analyzed and discussed in terms of the rotational isomeric model and its usefulness in assessing the degree of completeness of the conformational search. Finally, the iterative simulated annealing method is prone to thermal differences. In the present study the effect of the temperature is analyzed.

2.2 Introduction

Ras are guanosine triphosphatases that play a crucial role in the signal transduction processes involved in cell division. These enzymes, as well as other cytosolic proteins undergo a series of post-translational modifications, including farnesylation by farnesyltransferase (FT) (Schafer et al., 1992). These modifications result in cell membrane association, a required condition for Ras activity. Therefore, since the transforming activity of oncogenic Ras proteins is dependent on their post-translational farnesylation, inhibitors of farnesyltransferase (FTIs) have been shown to be novel promising agents in cancer therapeutics (Gibbs et al., 1997; Mangues et al., 1998; Barrington et al., 1998; Sun et al., 1995). Similarly, the identification of a farnesyltransferase in *Trypanosoma brucei* has opened the way for the development of anti-trypanosomatic chemotherapeutics based on the inhibition of FT (Yokohama et al. 1998).

FT is a heterodimer consisting of a 48-kD α-subunit and a 46-kD β-subunit (Reiss et al., 1990; Chen et al., 1991a; Chen et al., 1991b; Seabra et al., 1991). The crystal structure of the mammalian farnesyltransferase solved at 2.25 Å resolution, reveals the α-subunit as a crescent-shaped seven-helical hairpin domain, whereas the β-subunit exhibits a $\alpha-\alpha$ barrel (Park et al., 1997). The active site is formed by two clefts that intersect at a bound zinc ion. In the crystal structure, the cleft of the α-subunit accommodates a nine-residue peptide from the C-terminus segment of another β-subunit of the crystal that mimics the binding of the Ras substrate. The involvement of the α-subunit in binding the substrate is consistent with the fact that this subunit forms also part of the related enzyme geranylgeranyltransferase type I (GGT-I), which catalyzes the post-translational attachment of the 20-carbon isoprenoid geranyl group (Zhang et al., 1994). On the other hand, the cleft on the β-subunit is lined with highly conserved aromatic residues appropriate for binding the farnesyl isoprenoid diphosphate with the required specificity. FT as well as GGT-I use the corresponding isoprenoid diphosphate derivative to modify the substrate through a thioether linkage to a conserved cysteine located at the protein C-terminus. This cysteine is the first residue of the motif CA_1A_2X found in all FT protein substrates, where A_1 and A_2 represents any aliphatic residue and X a serine, alanine, glutamine, cysteine or methyonine.

Several classes of FT inhibitors have been disclosed to date (Qian, et al., 1997; Sebti et al., 1997; Omer et al., 1997). The replacement of the aliphatic residue A2 for phenylalanine was early shown to prevent farnesylation of the substrate, resulting in the first FT inhibitor disclosed (Goldstein et al., 1991). More extensive structure-activity studies showed the analog Cys-Val-Phe-Met to be the most potent inhibitor among 42 CXXX analogs studied, where X stands for all the possible natural amino acids (Reiss et al., 1991). Subsequent chemical reduction of selected

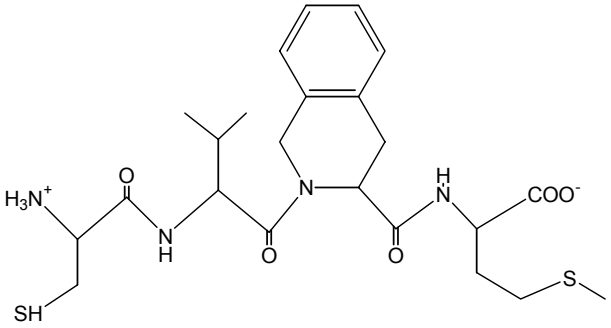
amide bonds in CA_1A_2X-based inhibitors, led to a significant improvement of the inhibitory potency and activity in whole cells (Leftheris et al., 1996). Similarly, analogs containing conformationally constrained amino acid residues were also shown to be potent inhibitors (Byk et al., 1997). In addition, non-peptide inhibitors have also been disclosed (Qian et al., 1997; Sebti et al., 1997).

The goal of the present study is to assess the bioactive conformation of the FT recognition motif CA_1A_2X, using indirect computational methods, i.e. by comparison of the conformational profiles of a small set of selected peptide analogs. The outcome of such a study is twofold. First, to assess the features of a 3D pharmacophore to be directly used for the design and discovery of novel small molecule FT inhibitors. Second, to compare the geometrical and chemical features of the putative bioactive conformation assessed with the structure of the nonapeptide bound to the receptor in the crystal structure, providing evidence to the widely accepted paradigm that the bound conformation of a bioactive peptide is one of its thermodynamically accessible structures and simultaneously validating the usefulness of this type of conformational analysis.

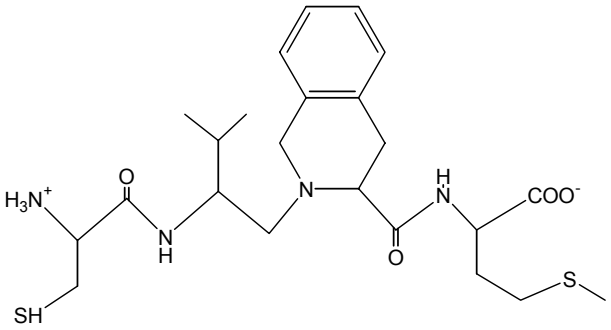
Accordingly, the conformational profiles of a small set of CA_1A_2X-based inhibitors (Leftheris et al., 1996), with A_2 being a conformationally constrained aromatic amino acid residue, were compared in order to assess the conformational requirements of the FT inhibitors. Specifically, two active and two inactive inhibitors were selected for the present study : the active analogs Cys-Val-Tic-Met **(1)** and Cys-Val-$\Psi(CH_2NH)$Tic-Met **(2)** and the inactive analogs Cys-Val-Tic-$\Psi(CH_2NH)$Met **(3)** and Cys-Val-Aic-Met **(4)**, where Tic and Aic are the unnatural amino acid residues (S)-1,2,3,4-tetrahydroisoquinoline-3-carboxylate and 2-aminoindane-2-carboxylic acid), respectively. The *in vitro* FT inhibitory activity of the four peptide analogs selected for the present study are listed in Table 2.1 and depicted in Figure 2.1 (Leftheris et al., 1996).

Table 2.1. In vitro FT inhibitory activity for the peptides selected for the present work.

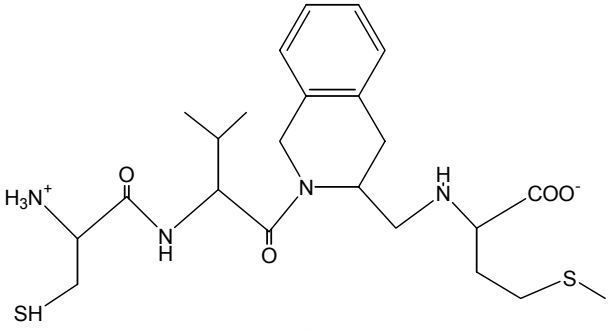
FTIs	IC_{50} (nm)
Cys-Val-Tic-Met	1
Cys-Val-$\Psi(CH_2NH)$Tic-Met	0.4
Cys-Val-Tic-$\Psi(CH_2NH)$Met	5300
Cys-Val-Aic-Met	2900



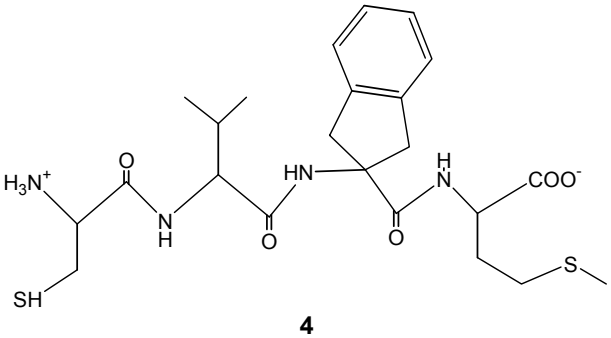
1



2



3



4

Figure 2.1. Molecular structures of the peptides investigated in the present study.

In order to identify the bioactive conformation of these inhibitors, the conformational space of each of these peptide analogs was explored using simulated annealing (SA) in an iterative fashion as sampling technique. The resulting low energy conformations were then classified using a criterion based on the identification of all the possible structural motifs stabilized by hydrogen bonds involving backbone atoms. Finally, cross-comparisons between the different classes of the four peptides studied permitted the identification of only one conformation that is common to the active analogs and that was not found in the inactive ones. This conformation was subsequently considered as the bioactive conformation of this kind of FT inhibitors.

The methodology to search the conformational space of a peptide is critical to accomplish satisfactory results since the potential energy surface of a peptide is characterized by the presence of a large number of valleys (or minima) separated by mountains and ridges (Frauenfelder, et. al. 1991). The minima in the continuously defined potential energy surface constitute a discrete set of microstates, called conformational substates. As a consequence the characterization of the subset of conformations that are thermodynamically or kinetically relevant and contribute to the description of the conformational profile of different peptides is hampered by the difficulties associated to have thorough explorations of the conformational space. This problem is referred in the literature as the multiple minima problem and is still an open question, despite the many techniques that are available to explore the conformational spaces of peptides (Bruccoleri et al., 1990; Scheraga, 1993). In order to demonstrate that simulated annealing procedure is a suitable procedure to characterize the subset of low energy conformations of a peptide we carried out a comparative study of the profile of the peptide Cys-Val-Tic-Met using simulated annealing and a random search procedure.

Some insights of the features of the conformational space can be assessed from the comparison of the density of states profile of such systems. From the few thorough explorations of the conformational space of different flexible molecules published in the past (Auffinger et al., 1990; Schaumann et al., 1990; von Freyberg et al., 1991; Pettit et al., 1991; Perez et al., 1992), it can be deduced that the histogram representing the number of conformations rank ordered by energy of such systems, exhibits a bimodal distribution. More specifically, for conformations with energy close to the global minimum, the distribution exhibits a characteristic exponential growth reaching two maxima at energies relative to the global minimum that depend on the length of the peptide, to continue with an exponential decrease again at higher energies. These features can be described by the random-energy approximation (Derrida et al., 1981) that has been used to describe rough energy landscapes and used to provide a model of the statistical mechanics of protein folding (Bryngelson et al., 1987) where the density of states is assumed to have a Gaussian distribution. On the other hand, the rotational isomeric approximation (Volkenstein, 1969; Perez et al., 1994; Flory, 1969) provides a similar distribution from the assumption that the molecule is composed of N independent rotors (Flory, 1969).

Accordingly, the conformational space of the tetrapeptide Cys-Val-Tic-Met was carried out using an iterative simulated annealing (SA) procedure and a random search. The sampling efficiency of the two searches was assessed by comparing the rank ordered histograms of conformations obtained. These results were used to understand the reliability of the predictions regarding the density of the states of this system that can be done using the rotational isomeric approximation.

The iterative SA procedure is supposed to explore the lowest energy region of the conformational space. This is done by heating fast in order to force the molecule to jump to a different region of the conformational space and then, cooling slowly and minimizing the structure obtained in order to take the molecule to a valley in the potential energy hypersurface. This structure is used as the starting point for the following SA cycle. Thus, when using SA the temperature value at which the structures are minimized could have a critical role in obtaining a low-energy conformation. In this study, different temperatures values have been used in order to assess the variation of the distribution maximum and the best conditions for the method.

2.3 Methods

All the calculations were carried out within the molecular mechanics framework by means of the AMBER program (Pearlman et al., 1991) using the Weiner all-atom force field (Weiner et al., 1986). The four peptides selected for the present work were studied in their zwitterionic form. No explicit solvent was included in the calculations, although an effective dielectric constant of 80 was used to screen electrostatic interactions and no non-bonded interactions cutoff was considered.

The structures of Val-$\Psi(CH_2NH)$Tic and Tic-$\Psi(CH_2NH)$Met where Tic stands for Tic, 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, as well as that of the unnatural amino acid residue Aic, (2-aminoindane-2-carboxylic acid), were constructed using the PREP module of AMBER. ESP charges were computed with the GAUSSIAN94 package by fitting the molecular electrostatic potential (MEP) to point charges located on the nucleii. MEPs were computed at the Hartree-Fock level with a STO-3G basis set and the fitting process was performed by means of the Merz-Kollman procedure (Singh and Kollman, 1984). Initial structures for the four peptides were obtained by energy minimization of the extended conformations using a conjugate gradient algorithm with the convergence criterion set to 0.001 $kcal \cdot mol^{-1} \cdot Å^{-1}$.

The conformational space was explored using a simulated annealing (SA) protocol in an iterative fashion. The method has been described elsewhere (Filizola et al., 1997). The initial minimized structure is quickly heated up to 900 K at a rate of 100 K/ps, in order to force the

molecule to jump to a different region of the conformational space. Subsequently, the 900 K structure is slowly cooled to 200 K at a rate of 7 K/ps and then minimized. The structure obtained at 200 K is stored on a file and used as the starting conformation for a new cycle of SA. In this way an energy rank ordered library of low energy conformations is generated. Low energy conformations are checked for uniqueness by exclusion of those for which at least one of the backbone dihedral angles is different from 60º in respect to the previous conformations already present in the library. The conformational search is continued until the searching procedure reaches a low level of performance in finding new low energy conformations (λ<0.1), being λ an efficiency parameter. This parameter is computed every 100 cycles of SA, and is defined as the number of unique conformations found after N cycles of SA ξ(N) divided by N, and adjusted by a coefficient so that the efficiency parameter is the unity after the first 100 cycles performed, i.e. λ(100) = 1:

$$\lambda(N) = \frac{\xi(N)\cdot 100}{N\cdot\xi(100)} \tag{2.1}$$

The unique conformations within 5 kcal·mol^{-1} above the corresponding lowest energy conformation of each of the peptides were classified according to their structural motifs. The motifs were automatically identified in each structure from the analysis of the hydrogen bonds that involve only atoms of the peptide backbone. The procedure requires to compute for each structure all possible amide hydrogen-carbonyl oxygen distances, as well as the corresponding N-H···O angles. Threshold values used to identify hydrogen bonds were set to 2.5 Å or less for the interatomic distance and to 120° or greater for the N-H···O angle.

The tetrapeptide selected for the second part of the study was Cys-Val-Tic-Met. The peptide was also studied in its zwitterionic form and no explicit solvent was included in the calculations, although an effective dielectric constant of 80 was used to compute the electrostatic interactions. The initial structure of the peptide was generated in an extended conformation and subsequently minimized using the conjugate gradient algorithm with a convergence criterion set to 0.001 kcal·mol^{-1}·Å^{-1}.

The conformational space was explored by two methods: (i) the SA procedure used in an iterative fashion described above and (ii) a random search exploration. The random search was done by generating 12000 inputs for the PREP module of AMBER 4.0 that have the dihedrals angles of the backbone generated randomly. Angles ω_1 and ω_3 were fixed at 180º and ω_2 was modified randomly to exhibit values of 0º and 180º. The structures were then generated separately using the AMBER program and subsequently were minimized using the conjugate gradient algorithm until the 0.001 kcal·mol^{-1}·Å^{-1} convergence criterion was fulfilled.

The set of conformations generated through the random search approach and by SA were checked for uniqueness separately by removing those structures for which at least one of the backbone dihedral angles was different from 60º in respect to any of the previous conformations already stored in the library. The SA conformational search was stopped when the method reached a low efficiency (λ<0.1) as it has been described above.

2.4 Results and Discussion

2.4.1 Conformational analysis of Ras farnesyltransferase inhibitors

As mentioned in the Methods section, the conformational search for each of the four analogs selected for the present study was performed using a simulated annealing protocol in an iterative fashion, as sampling technique. The SA procedure was stopped in all the four cases after 12000 cycles, at which the efficiency of the sampling (λ) was considered lower than the threshold value of 0.1. Figure 2.2 shows the evolution of this parameter along the conformational search.

A summary of the results of the SA process for each peptide is shown in Table 2.2. Following the criterion described in the methods section, 371 conformations were considered unique for the analog Cys-Val-Tic-Met **(1)**. Of these, 291 conformations resulted within a 5 kcal·mol^{-1} range above the lowest energy minimum found, that was obtained at iteration number 889 with an energy of +1.99 kcal·mol^{-1}. The analog Cys-Val-Ψ(CH_2NH)Tic-Met **(2)** showed 531 unique conformations, of which 169 exhibited energies within 5 kcal·mol^{-1} of the lowest energy minimum found. This structure was obtained at iteration number 1653 with an energy value of -0.57 kcal·mol^{-1}. The number of unique conformations obtained for Cys-Val-Tic-Ψ(CH_2NH)Met **(3)** was 537, being 398 of them within the 0-5 kcal·mol^{-1} energy range in regard to the lowest energy conformation found at iteration 460 with an energy value of +4.18 kcal·mol^{-1}. Finally, the conformational search of Cys-Val-Aic-Met **(4)** yielded 532 unique conformations of which 305 belong to the 0-5 kcal·mol^{-1} range. Its lowest energy minimum occurred at iteration 7281 with an energy value of +12.71 kcal·mol^{-1}.

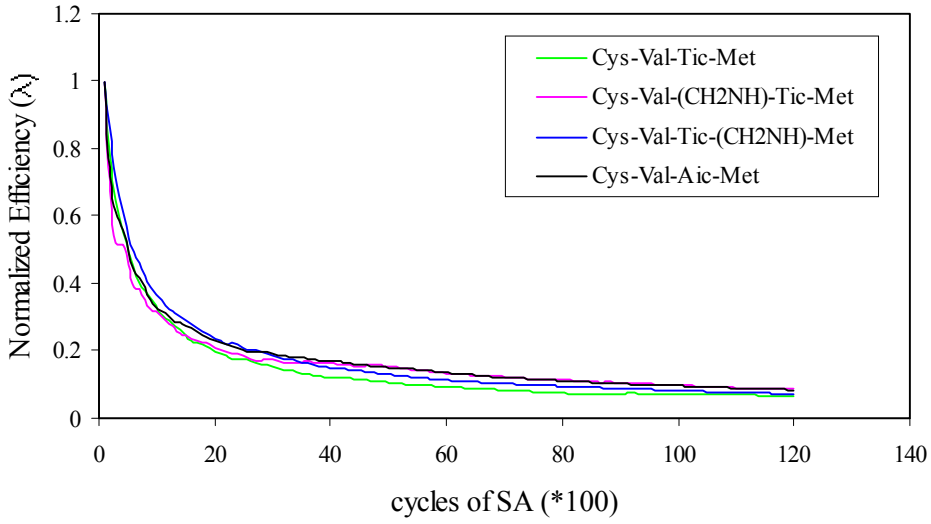


Figure 2.2. Normalized efficiency values for all the peptide searches, computed using the equation reported in the Methods section.

Table 2.2. Computational results of the selected farnesyltransferase inhibitors

FTIs	Unique conformations	Unique conformations in 5 kcal/mol	Number of classes	Global minimum energy value (kcal/mol)
Cys-Val-Tic-Met	371	291	10	+1.99
Cys-Val-Ψ(CH_2NH)Tic-Met	531	169	12	-0.57
Cys-Val-Tic-Ψ(CH_2NH)Met	537	398	11	+4.18
Cys-Val-Aic-Met	532	305	14	+12.71

Low energy conformations characterized within the corresponding 0-5 kcal·mol^{-1} energy range of the four peptides studied were classified into classes according to a criterion based on identification of all the structural motifs they exhibit that are stabilized by hydrogen bonds involving backbone atoms. Since these molecules are four-residue peptides, possible structural motifs are: β-turns involving a hydrogen bond between the amide of residue 4 and the carbonyl oxygen of residue 1 (abbreviated as N4-O1); γ-turns N4-O2 and N3-O1; C8 pseudo-rings N1-O2 and N2-O3; C11 pseudo-rings N1-O3 and N2-O4; and the C14 pseudo-ring N1-O4. Beginning with the lowest energy conformation, all unique conformations for the four analogs studied were assigned to the different classes according to the motif(s) found. The first column of Table 2.3 lists all the different structural motifs that have been found among all the four peptide analogs. The rest of columns

provide information regarding the different peptide analogs and it is organized as follows. Under the name of the analog there are two columns: the first indicates the structural class found to exhibit the corresponding structural motif(s) that identifies the row; whereas the second indicates the percentage of occurrence of the different class among the list of unique conformations. Specifically, for analog **1**, ten different classes were found. Backbone dihedral angles of the representative conformation of each class (i.e. the conformation with the lower conformational energy in the class) characterized for this peptide are listed in Table 2.4.a. The most populated class for this peptide is the so-called bent conformation that does not exhibit any of the standard motifs defined above (class #2). The lowest energy conformation found is the representative conformation of class #1 and exhibits a C14 pseudo-ring with a hydrogen bond between the amino group of Cys^1 and the carboxylate group of Met^4. Classification of the low energy conformations found for analog **2**, yielded 12 different classes. Table 2.4.b lists the backbone dihedral angles of the representative conformations of each of the different classes found for this peptide. Also in this case, the most populated class (class #5) corresponds to bent conformations with no standard motifs involving backbone atoms. The lowest energy conformation exhibits a C11 pseudo-ring motif (class #1) characterized by a hydrogen bond between the hydrogen of the amide group of Val^2 and the oxygen of the carboxylate group of Met^4. In the case of the analog **3**, 11 different classes resulted by identification of structural motifs stabilized by hydrogen bonds involving backbone atoms. Backbone dihedral angles of the representative conformation of each class of this peptide are listed in Table 2.4.c. Analysis of Table 2.3 reveals that the lowest energy conformation found corresponds to a bent conformation with no standard motifs and represents the most populated class (class #1). Finally, classification of the low energy conformations of the analog **4** permits the identification of 14 classes and again the lowest energy minimum can be characterized (class #1) as a bent conformation (Table 2.3). Table 2.4.d shows the backbone dihedral angles of the representative conformation of each class for this peptide.

Table 2.3. Classification of the unique conformations within the 0-5 kcal·mol^{-1} energetic range for the selected FT inhibitors.

	Cys-Val-Tic-Met (1)		Cys-Val-Ψ(CH$_2$NH)Tic-Met (2)	
Structural motifs	**Class #**	**% of unique conformations**	**Class #**	**% of unique conformations**
bent conformation	2	69.4	5	58.6
β-turn(N4-O1)	4	6.2	7	5.3
γ-turn(N4-O2)	9	7.2	—	—
γ-turn(N3-O1)	—	—	—	—
C14(N1-O4)	1	9.6	6	9.5
C11(N1-O3)	3	4.1	2	4.1
C11(N2-O4)	5	0.3	1	8.9
C8(N1-O2)	—	—	—	—
C8(N2-O3)	—	—	9	3.0
β-turn(N4-O1); γ-turn(N3-O1)	—	—	—	—
β-turn(N4-O1); C14(N1-O4)	10	0.3	11	2.4
γ-turn(N3-O1); γ-turn(N4-O2)	—	—	—	—
γ-turn(N4-O2); C14(N1-O4)	8	0.3	—	—
γ-turn(N3-O1); C14(N1-O4)	—	—	—	—
γ-turn(N4-O2); C8(N1-O2)	—	—	—	—
C14(N1-O4); C11(N1-O3)	7	1.4	3	1.2
C14(N1-O4); C8(N1-O2)	—	—	—	—
C14(N1-O4); C11(N2-O4)	6	1.0	4	3.5
C11(N1-O3); C11(N2-O4)	—	—	8	5.3
C11(N1-O3); C8(N2-O3)	—	—	12	1.8
β-turn(N4-O1), γ-turn(N3-O1); C14(N1-O4)	—	—	—	—
γ-turn(N3-O1); γ-turn (N4-O2); C14(N1-O4)	—	—	—	—
C14(N1-O4); C11(N1-O3); C11(N2-O4)	—	—	10	1.2

Table 2.3. Classification of the unique conformations within the 0-5 kcal·mol^{-1} energetic range for the selected FT inhibitors.

	Cys-Val-Tic-Ψ(CH_2NH)Met (3)		**Cys-Val-Aic-Met (4)**	
Structural motifs	**Class #**	**% of unique conformations**	**Class #**	**% of unique conformations**
bent conformation	1	71.6	1	55.7
β-turn(N4-O1)	3	6.5	8	1.6
γ-turn(N4-O2)	4	7.8	6	4.9
γ-turn(N3-O1)	—	—	4	19.0
C14(N1-O4)	2	8.8	3	6.9
C11(N1-O3)	—	—	9	2.0
C11(N2-O4)	8	0.8	—	—
C8(N1-O2)	9	0.8	—	—
C8(N2-O3)	—	—	—	—
β-turn(N4-O1); γ turn(N3-O1)	—	—	13	0.3
β-turn(N4-O1); C14(N1-O4)	5	0.5	7	0.7
γ-turn(N3-O1); γ-turn(N4-O2)	—	—	5	2.0
γ-turn(N4-O2); C14(N1-O4)	6	0.5	2	5.2
γ-turn(N3-O1); C14(N1-O4)	—	—	14	0.7
γ-turn(N4-O2); C8(N1-O2)	11	0.2	—	—
C14(N1-O4); C11(N1-O3)	—	—	—	—
C14(N1-O4); C8(N1-O2)	10	0.5	12	0.3
C14(N1-O4); C11(N2-O4)	7	2.0	—	—
C11(N1-O3); C11(N2-O4)	—	—	—	—
C11(N1-O3); C8(N2-O3)	—	—	—	—
β-turn(N4-O1), γ-turn(N3-O1); C14(N1-O4)	—	—	11	0.3
γ-turn(N3-O1); γ-turn (N4-O2); C14(N1-O4)	—	—	10	0.3
C14(N1-O4); C11(N1-O3); C11(N2-O4)	—	—	—	—

Table 2.4.a. Backbone dihedral angles of the representative conformation of each class of the analog Cys-Val-Tic-Met

Angles	ψ_1	ω_1	φ_2	ψ_2	ω_2	φ_3	ψ_3	ω_3	φ_4
1	-58	179	-113	121	5	-78	-33	180	-122
2	71	177	-94	128	4	-81	-37	-179	-134
3	-118	-177	-139	71	-1	-91	-27	178	-91
4	-55	-179	-58	144	5	-92	23	180	-113
5	-119	-177	70	104	12	-77	-33	-179	-85
6	-53	-179	64	88	11	-76	151	180	66
7	-59	180	-118	87	2	-96	179	178	59
8	176	178	-57	-61	173	-82	76	176	60
9	66	178	-129	86	178	-96	62	-179	-120
10	-63	-179	-60	-47	168	-108	40	-178	54

Table 2.4.b. Backbone dihedral angles of the representative conformation of each class of the analog Cys-Val-$\Psi(CH_2NH)$Tic-Met

Angles	ψ_1	ω_1	φ_2	ψ_2	ω_2	φ_3	ψ_3	ω_3	φ_4
1	82	180	-79	-60	93	-110	-18	179	-72
2	-132	180	-78	-60	96	-103	-28	178	-118
3	-31	-176	-78	-63	92	-110	-14	176	-120
4	49	-175	-84	-64	96	-130	37	176	-130
5	-62	-178	-78	-61	95	-129	32	178	-132
6	-60	177	-137	59	88	-89	-56	180	-163
7	-89	-174	-88	75	97	-77	-56	178	-72
8	-137	178	-141	81	-40	-63	-25	175	-72
9	54	-178	-110	-65	89	-115	179	179	62
10	-51	-172	-139	87	-45	-65	-27	178	-71
11	-60	177	-136	62	84	-134	57	-178	56
12	60	-178	-151	-63	86	-114	-57	-177	-83

Table 2.4.c. Backbone dihedral angles of the representative conformation of each class of the analog Cys-Val-Tic-Ψ (CH_2NH)Met

Angles	ψ_1	ω_1	φ_2	ψ_2	ω_2	φ_3	ψ_3	ω_3	φ_4
1	173	174	-74	147	-2	-85	-59	102	-152
2	-173	-175	-56	-65	179	-75	-52	-177	-99
3	132	-179	-61	148	0	-99	57	161	-120
4	-68	-174	55	76	179	-96	60	-174	-115
5	-55	177	-127	127	2	-94	57	-127	61
6	-65	-177	-67	-50	179	-104	59	-163	85
7	-73	-179	-139	-70	-173	-76	-63	125	-72
8	88	178	-160	101	3	-76	-52	121	-101
9	98	1	-65	144	-176	-79	-61	104	-145
10	102	1	-63	151	0	-89	-47	-173	-109
11	98	1	-64	144	-177	-93	61	-167	-131

Table 2.4.d. Backbone dihedral angles of the representative conformation of each class of the analog Cys-Val-Aic-Met

Angles	ψ_1	ω_1	φ_2	ψ_2	ω_2	φ_3	ψ_3	ω_3	φ_4
1	142	-179	-59	-39	177	57	67	180	-119
2	-60	-178	-66	123	-173	74	-63	173	-133
3	-65	-174	-78	90	180	59	37	-177	-162
4	82	176	-85	69	-174	58	35	180	-79
5	82	177	-83	82	-168	74	-58	172	-124
6	-74	-176	59	77	-177	75	-64	175	-135
7	-166	175	-60	-52	175	-77	11	177	64
8	-70	-178	55	55	180	61	24	177	-137
9	-66	174	-67	-60	169	-73	166	-179	-132
10	-70	177	-92	55	-165	73	-58	170	-77
11	105	-176	68	-76	172	-167	46	179	-162
12	96	4	-61	150	-171	73	-152	-179	-70
13	157	177	-81	75	-176	-179	-44	-178	60
14	-72	179	-85	68	-174	173	-60	175	-152

2.4.2 Characterization of the bioactive conformation

Assessment of the bioactive conformation of the protein-binding motif for FT inhibition requires the characterization of structural classes among the four analogs that are common to the two inhibitors and at the same time not found in the non-binders. For this purpose, cross-comparisons between all the unique conformations of the different structural classes characterized for the four peptide analogs were performed. The analysis yielded class#7 of peptide **1** and class #3 of peptide **2** as the structural classes that fulfill the above mentioned requirements and consequently were considered as the putative bioactive conformation. The structure is characterized by two structural motifs: a C14 pseudo-ring, stabilized by a hydrogen bond between the amide group of Cys^1 and the carboxylate group of Met^4 and a C11 pseudo-ring involving the residues Cys^1 and Tic^3. Figure 2.3 shows the superimposition of the two conformations of class #7 of peptide **1** and class #3 of peptide **2** with the lowest RMS deviation value (0.48 Å when all the backbone atoms are considered). Backbone dihedral angles of the representative conformations of the classes #7 of peptide **1** and #3 of peptide **2** characterizing the putative bioactive conformation can be compared from the results listed in Tables 2.4.a and 2.4.b, respectively.

Both conformations are characterized by a similar disposition of the Cys^1 side chain that could have a crucial role in the inhibition mechanism. Specifically, the hydrogen of the thiol group points to the oxygen of the carboxylate group of Met^4 forming a hydrogen bond. The importance of this disposition of the Cys^1 side chain for the inhibition of FT is confirmed by its opposite orientation in the conformations of the two inactive peptides selected for this study with the lowest values of RMS deviation derived by superimposition of the backbone atoms with those of the bioactive conformation. Figure 2.4 shows the superimposition between the bioactive conformation and those with lowest RMS deviation of the two non-binders (0.53 *Å* for analog **3** and 0.63 Å for analog **4**).

Present results can be compared with previous NMR spectroscopy studies (Stradley et al., 1993) carried out on the heptapeptide Lys-Thr-Lys-Cys-Val-Phe-Met bound to FT. These studies clearly revealed that the sequence Cys-Val-Phe-Met is the segment involved in enzyme binding. Moreover, in the same study it was also suggested that the segment adopts a type I β-turn conformation when bound to the enzyme. Type I β-turn conformations have been characterized for each of the four analogs studied in the present work (see Table 2.3). Specifically, type I and VI β-turns (Möhle et al., 1997) are found among the different classes of conformations of the four analogs, depending on whether the *i+1* dihedral angle was 180° or 0°, respectively. However, this motif does not represent a discriminant feature between binders and non-binders, and cannot be used to characterize the bioactive conformation. Furthermore, not even the different proportion of

β-turn type structures found among the different analogs can be used as discriminant, since they appear to show similar profiles: 6.2% for analog **1**, 5.3% for analog **2** and 6.5% for the non-binder analog 3. Spite of the differences, the differential features between the putative bioactive conformation assessed in this work and a type I β-turn are small and both structures can be well superimposed showing common structural features.

Similar results are obtained from the comparison of the putative bioactive conformation proposed in this work and the structure of the nine-residue peptide Ala^{9}-Val^{8}-Thr^{7}-Ser^{6}-Asp^{5}-Pro^{4}-Ala^{3}-Thr^{2}-Asp^{1}-COOH of the C-terminus segment of a neighboring β-subunit accommodated in the cleft of the α-subunit found in the crystal structure of FT. The crystal structure of the peptide exhibits two consecutive β-turns involving backbone atoms N1-O4 and N2-O5, respectively. The best superimposition between this peptide and the putative bioactive conformation corresponds to the segment Asp^{5}-Pro^{4}-Ala^{3}-Thr^{2}. As in the NMR studies, although the backbone of the putative bioactive conformation is not a β-turn, side chains of the residues thought to play a role in binding exhibit similar 3D features in both structures. Thus, the side chain of Cys^{1} in the putative bioactive conformation exhibits similar conformational accessibility than that of Asp^{5} in the nine-residue peptide crystal and may play a similar role in the interaction with zinc ion. Similarly, the side chain of the putative bioactive conformation of Met^{4}, thought to be important in the inhibition mechanism, points in the same direction of Thr^{2} in the nine-residue peptide crystal structure.

Comparison of the structural features of the putative bioactive conformation assessed in this work with those structures deduced from NMR and X-ray diffraction studies, suggests that the former exhibits structural features similar to those of the experimentally determined. The putative bioactive conformation is characterized by two structural motifs: a C14 pseudo-ring involving residues 1 and 4 and a C11 pseudo-ring involving residues 1 and 3. A structure exhibiting simultaneously both motifs can only be found in the sets of structures of analogs **1** and **2**. However, the two pseudo-rings can be found either isolated or combined with other structural motifs in all the analogs studied, although not with the same propensity among them. Interestingly, the number of unique structures exhibiting a C11 pseudo-ring formed by a hydrogen of the amino group of Cys^{1} and the oxygen atom of the carbonyl group of Tic^{3} shows differences among the four peptides studied. Specifically, a 4.1 % of unique conformations within the 0-5 $kcal \cdot mol^{-1}$ energy range of both inhibitors **1** and **2** shows this motif alone, whereas the two inactive analogs **3** and **4** exhibit percentages of zero and 2.0, respectively. In contrast, no significant trends are observed for a differential propensity to adopt a C14 pseudo-ring, as shown in Table 2.3. After these considerations and taken into account that present calculations were carried out in the absence of the zinc ion, these results suggest that the C14 pseudo-ring may not be an important feature of the bioactive conformation, since the extra interaction would be involved in binding the zinc ion. Accordingly, the propensity to adopt a C11 pseudo-ring between the first and the third residues

can be considered as the most important structural feature that discriminates binders and non-binders and provides the analogs similar structural features to those conformations characterized by experimental studies.

Finally, the fact that the bioactive conformation assessed in the present work is not a β-turn structure may only be a requirement of the analogs studied here, due to the presence of a constrained phenylalanine in all of them. In this regard, it should be considered that in the case of ligand-receptor complexes where none of the backbone atoms of the ligand interact with the receptor, different conformations may lead to the same disposition of the side chains and provide the features necessary for a proper interaction with the receptor. This may happen in the present case where it appears that the ligand-zinc ion interaction with the side chains of the first and fourth residues is one of the most important features for the recognition process with the receptor.

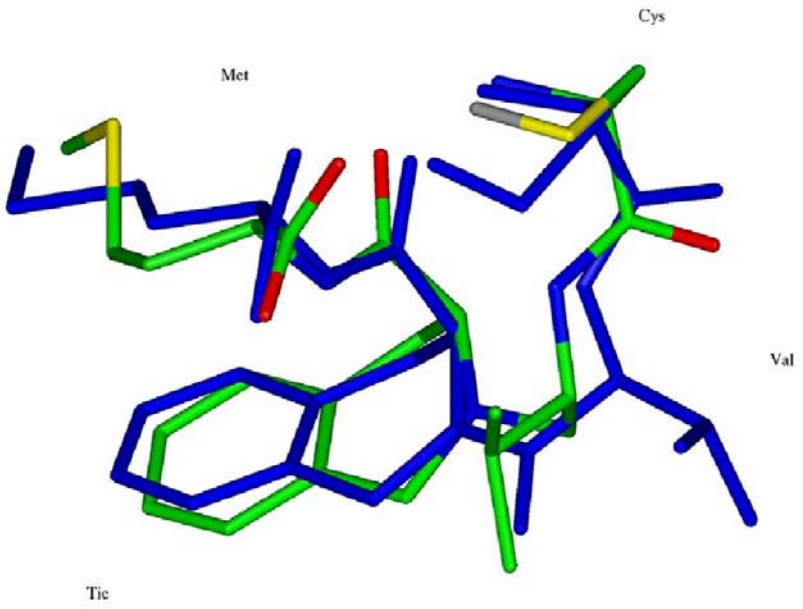


Figure 2.3. Superimposition of the putative bioactive conformation of Cys-Val-ψ(CH_2NH)Tic-Met (colored by atom) with the representative member of the corresponding class of Cys-Val-Tic-Met (colored in blue). RMS distance between the two structures considering all the backbone atoms is 0.48 Å.

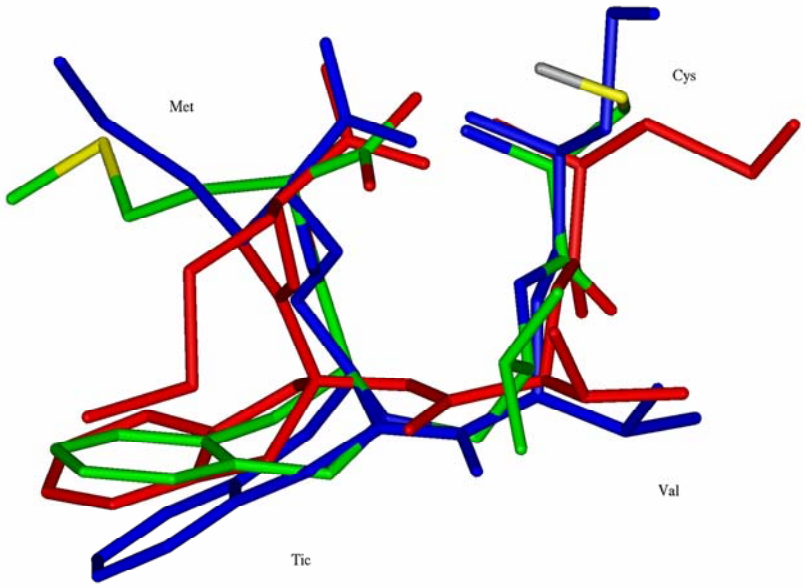


Figure 2.4. Superimposition of the putative bioactive conformation (colored by atom) and the closest of each of the two inactive peptides Cys-Val-Tic-(CH_2NH)Met (blue) and Cys-Val-Aic-Met (red).

2.4.3 Comparison of iterative SA and random conformational searches

Energy distributions of unique conformations after 12000 cycles of iterative SA or the random search are shown in Figure 2.5. Energies are referred to the global minimum that was obtained with the SA procedure. Energy distributions exhibit different shapes: the SA procedure exhibits one maximum at around 4 kcal·mol^{-1} above the global minimum, whereas the random search exhibits a bimodal distribution with one of the maxima at around 7 kcal·mol^{-1} and the other at around 10 kcal·mol^{-1}. These results clearly show that the SA procedure samples more efficiently low energy conformations.

Comparative analysis of the different conformations obtained with the two procedures reveals that it is likely to obtain structures exhibiting the same backbone in both procedures,

however, side chains are always much more relaxed in the structures obtained by the SA procedure. This suggests that although the collection of attainable backbones of the peptide can be sampled using either procedure, the SA protocol provides the most optimized structures within a specific backbone disposition.

From these results and in order to get some insights into the origin of the two peaks of the random search, we proceeded to segregate those conformations exhibiting a hydrogen bond from those that do not exhibit any. Figures 2.8 and 2.9 show energy histograms of all the conformations compared to those with and without hydrogen bonds for the two searches of the conformational space. Figures 2.8 and 2.9 suggest that: conformations that exhibit hydrogen bonds are more abundantly found at energy around the lower energy peak, whereas those that do not have any contribute in a greater extend to the distribution of conformations found at energies around the higher energy peak.

Assessment of the usefulness of the rotational isomeric approximation for predicting the features of the density of states of a specific peptide was performed by computing the energy of the maximum of the distribution U_0. This is computed from the effective number of torsional degrees of freedom *f*; the mean number of rotamers of each rotor *m+1*; and the window energy width, ε_0 (Volkenstein, 1963, Flory, 1969):

$$U_0 = \frac{fm\varepsilon_0}{2} \tag{2.2}$$

In order to get the number of effective rotors, f, it is necessary to carry out a fit of the distribution of states at low energies to the bosonic limit of the distribution in the rotational isomeric model. In the bosonic limit, the density of states has an exponential dependence with the energy, being in the present case: $\ln\Omega = 1.32 + 0.78E_0$.

The second term of this equation is consistent with the energy width of 0.8 kcal·mol^{-1} and the first term is used to compute the value of f = 12. Since the mean value of states for each of the rotors is m=2 and ε_0 = 0.8, results in a value for U_0 of 9.6 kcal·mol^{-1} in reasonable agreement with the position of the second maximum found in the random search of the conformational space.

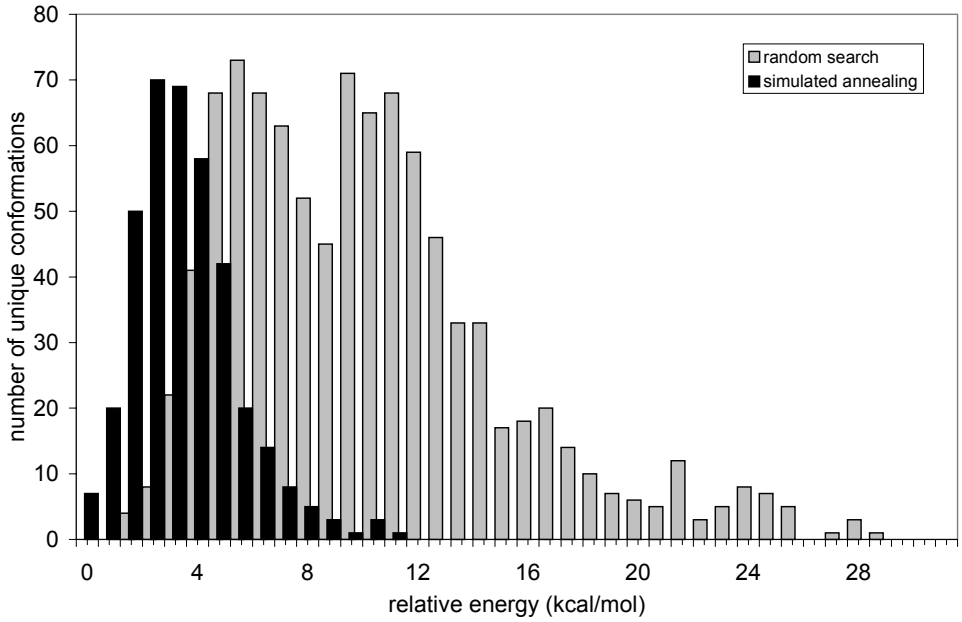


Figure 2.5. Histogram of the unique conformations rank ordered by energy obtained with the iterative simulated annealing procedure (black bars) and random search (light bars) referred to the lowest energy conformation found.

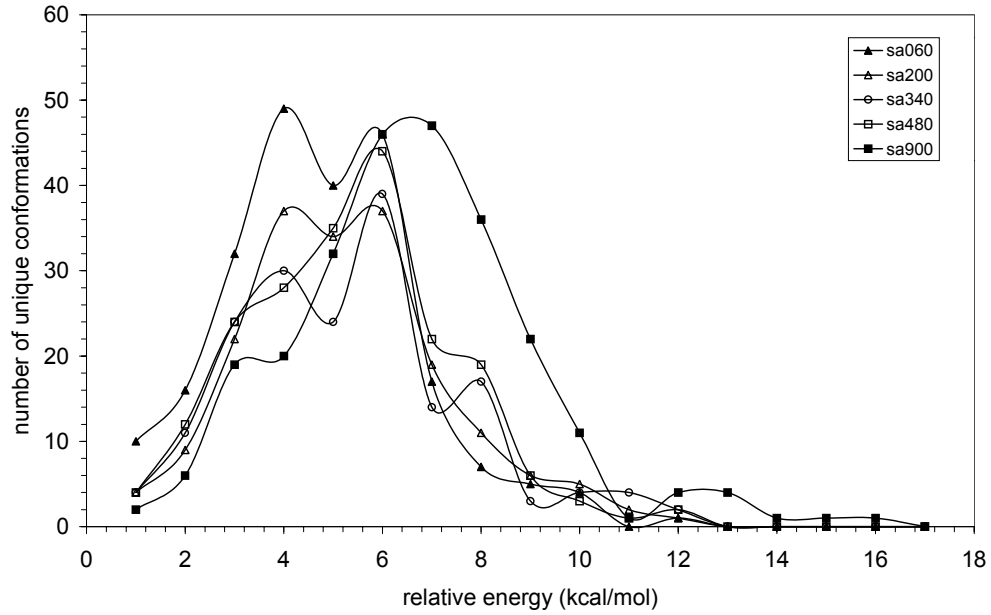


Figure 2.6. Distributions of unique conformations obtained by a simulated annealing protocol at different temperatures of quenching: 900 K (solid squares); 480 K (empty squares); 340 K (empty circles); 200 K (empty triangles); 60 K (solid triangles).

In order to compare the efficiency of both SA and the random strategy, the Eq. 2.1 has been used to assess the evolution of this parameter during the conformational search. The results are shown in Figure 2.7.a for the total unique conformations and Figure 2.7.b for the unique conformations with a potential energy below 5 kcal·mol^{-1}. Figure 2.7.a shows that the initial efficiency of the random strategy is 2.5 larger than that for the SA. Both for the SA and the random strategy the efficiencies decrease in a negative exponential way. On the other hand when the efficiency is calculated using the unique conformations within a 0-5 kcal·mol^{-1} range the efficiency for the SA method is 9.9 larger than the efficiency for the random approach (Figure 2.7.b). Therefore, the SA approach is clearly more efficient when the aim of the conformational search is to obtain a subset of low-energy conformations.

2.4.4 Effect of temperature in the efficiency of SA methodology

In order to understand the effect of the temperature at which the conformations are quenched before minimization when using a SA procedure on the distribution of states, this was varied systematically. Figure 2.6 shows the results of 2000 cycles of SA performed at different quenching temperatures: 900, 480, 340, 200 and 60 K. At 900 K only one peak is observed at around 7 kcal·mol^{-1}. When the temperature is lowered, the position of the maximum shifts towards lower energies and appears splitted with the new peak at around 4 kcal·mol^{-1}.

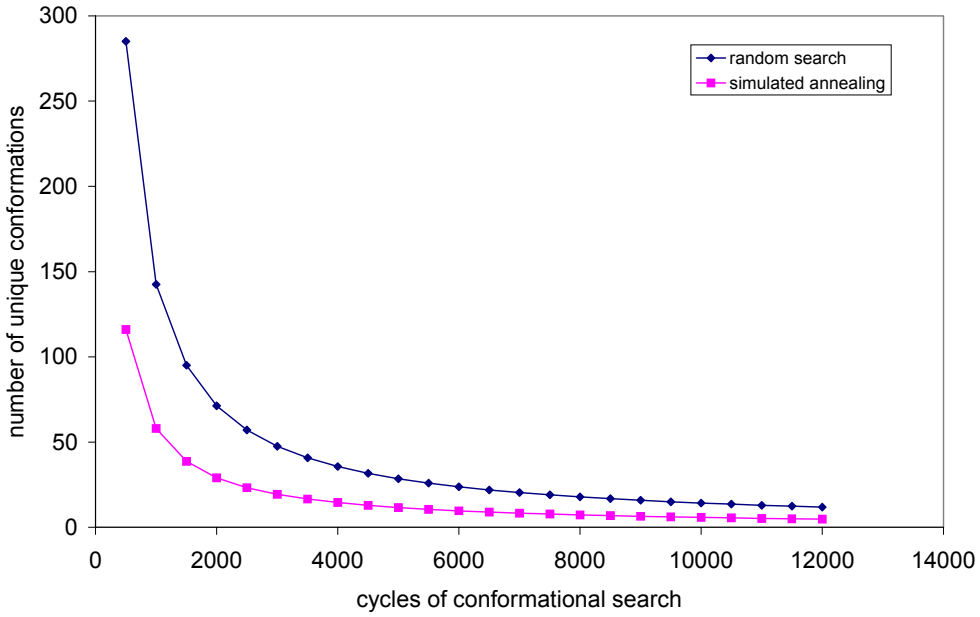


Figure 2.7.a. Absolute efficiency of the random search and the simulated annealing procedure for the Cys-Val-Tic-Met.

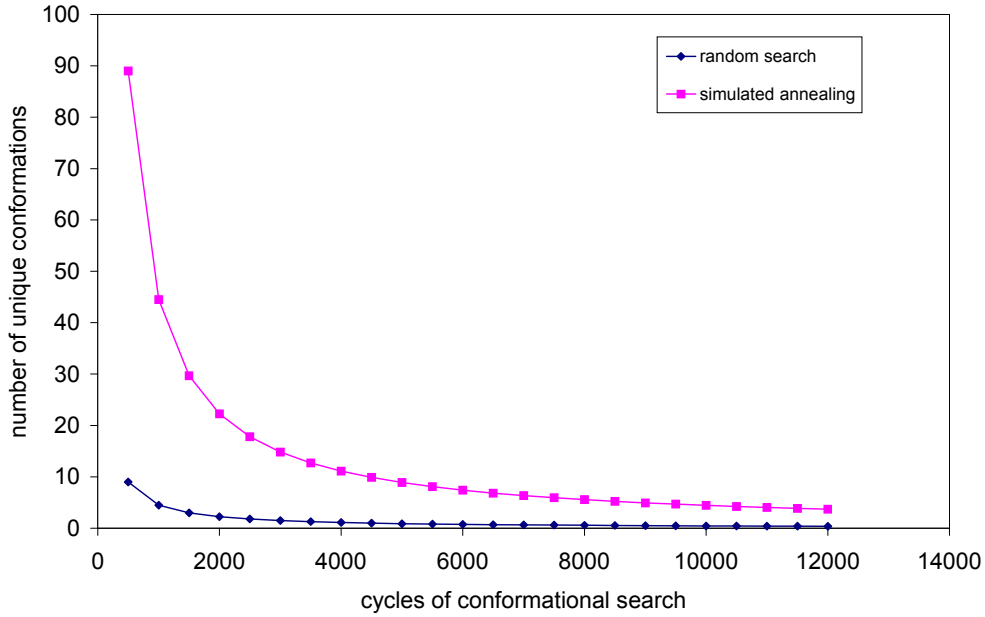


Figure 2.7.b. Absolute efficiency of the random search and the simulated annealing procedure for the unique structures below 5 kcal·mol^{-1} of Cys-Val-Tic-Met.

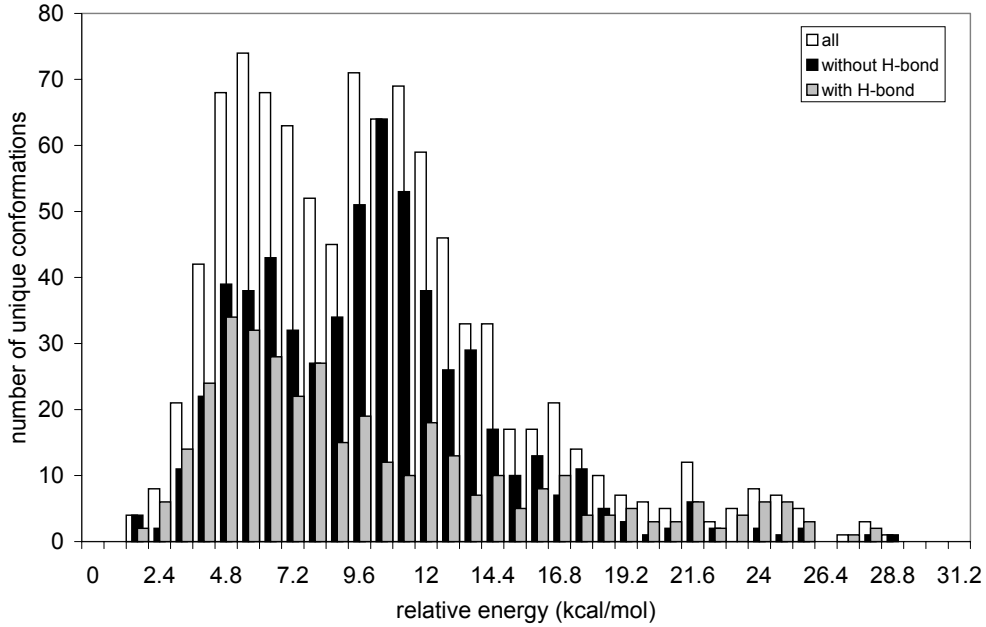


Figure 2.8. Histogram of the unique conformations obtained from the random search (in white). Those conformations that exhibit at least a hydrogen bond are depicted in grey and those conformations not exhibiting a hydrogen bond are depicted in black.

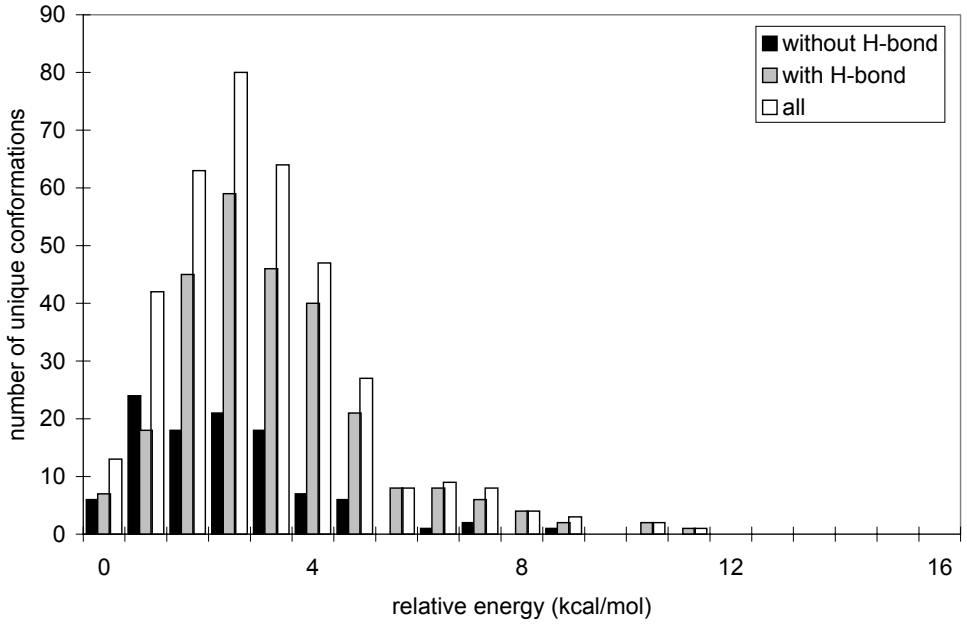


Figure 2.9. Histogram of the unique conformations obtained from the simulated annealing (in white). Those conformations that exhibit at least a hydrogen bond are depicted in grey and those conformations not exhibiting a hydrogen bond are depicted in black.

2.5 Conclusions to chapter 2

Conformational profiles of two active and two inactive FTIs were performed by computational methods. The exploration of the conformational space was carried out by an iterative simulated annealing method. The resulting unique conformations within 5 kcal·mol^{-1} were classified into classes using a structural criterion based on the identification of all the possible motifs stabilized by hydrogen bonds between the backbone atoms. Only one conformation common to the two inhibitors and not present in the inactive peptides was identified by cross-comparisons between the different classes. This conformation, recognized as bioactive one, is characterized by a C14 pseudo-ring stabilized by a hydrogen bond between the amino group of Cys^1 and the carboxylate group of Met^4 and a C11 pseudo-ring involving the residues Cys^1 and Tic^3. In addition, a hydrogen bond between the hydrogen of the Cys^1 thiol group and the carboxylate group of Met^4 was identified. These results can be helpful for the design of new inhibitors of farnesyltransferase.

The iterative simulated annealing used in this study was supposed to explore preferentially the low-energy region of the conformational space of peptides. When comparing this method with

a random strategy the efficacy of iterative simulated annealing in obtaining a subset of low-energy conformation becomes evident.

From the results obtained in the present study it can be deduced that the density of states of the tetrapeptide studied exhibits a bimodal distribution. This is only shown in the random search, since the more efficient SA procedure only provides structures in the lower peak of the distribution. Importantly, the different backbone geometries attained by the peptide could be found in any of the two searches.

The origin of the bimodal distribution can be explained in a very simplistic form, as being the result of the superposition of two different distributions of different sets of conformations depending on the hydrogen bonds that are found in the structure. Furthermore, the maximum of this distribution depends on the way the system has been quenched. Conformations minimized directly from 900 K appear more strained than those quenched at lower temperatures providing different distributions for a limited search.

The fact that the position of maximum predicted by the rotational isomeric approximation lies close to the second maximum found in the random search may be explained on the basis of the nature of the model's approximation. The model assumes that all possible conformations of a system can be generated from all possible combinations of the minima of the different rotors, not contemplating any conformation that could result from a strong intermolecular interaction between different moieties of the chain. Consequently, it is expected that this model describes better the subset of conformations that lie on the second peak of the distribution and this might justify the accuracy of the prediction of the rotational isomeric model. However, it should be noted as an important weakness of the prediction power of the model that in order to make reliable predictions it is required the computation of the effective number of rotors that describe the system. With this information and from pilot calculations performed using a random search it is possible to predict the small range of energy were the global minimum of the peptide is expected to appear.

2.6 References to chapter 2

Auffinger, P. and Wipff, G. *J. Comput. Chem.,* 11, 19, (1990).

Barrington, R.E., Subler, M.A., Rands, E., Omer, C.A., Miller, P.J., Hundley, J.E., Koester, S.K., Troyer, D.A., Bearss, D.J., Conner, M.W., Gibbs, J.B., Hamilton, K., Koblan, K.S., Mosser, S.D., O'Neill, T.J., Schaber, M.D., Senderak, E.T., Windle, J.J., Oliff, A. and Kohl N.E. *Mol. Cell. Biology,* 18, 85-92, (1998).

Bruccoleri, R.E., Haber, E. and Novotry, J. *Protein Folding*, eds. Gierasch, L. M. and King, J. (American Association for the Advancement of Science, Washington, pp. 259-270, (1990).

Bryngelson, J.D. and Wolynes, P.G. *Proc. Natl. Acad. Sci. USA*., 84, 7524, (1987).

Byk, G., Lelievre, Y., Duchesne, M., Clerc, F.F., Scherman, D. and Guitton, J.D. *Bioorg. Med. Chem.,* 5, 115-124, (1997).

Chen, W.J., Andres, D.A., Goldstein, J.L. and Brown, M.S. *Proc. Natl. Acad. Sci. USA,* 88, 11368-11372, (1991).

Chen, W.J., Andres, D.A., Goldstein, J.L., Rusell, D.W. and Brown, M.S. *Cell*, 66, 327-334, (1991).

Corcho, F.J., Filizola, M. and Pérez, J.J. *Chem. Phys. Letters,* 319, 65-70, (2000).

Corcho, F.J., Filizola, M. and Perez, J.J. *J. Biomol. Struct. Dyn.,* 16, 1043-1052, (1999).

Derrida, B. Phys. Rev., B24, 2613, (1981).

Filizola, M., Centeno, N.B. and Perez, J.J. *J. Peptide Sci.,* 3, 85-92, (1997).

Flory, P.J. *Statistical mechanics of chain molecules*; Interscience Publishers: New York, (1969).

Frauenfelder, H., Sligar, S.G., and Wolynes, P.G. *Science,* 254, 1598, (1991).

Gibbs, J.B. and Oliff, A. *Ann. Rev. Pharmacol. Toxicol.,* 37, 143-166, (1997)

Goldstein, J.L., Brown, M.S., Stradley S.J., Reiss, Y. and Gierasch, L.M. *J. Biol. Chem.,* 266 15575-15578, (1991).

Leftheris, K., Kline, T., Vite, G.D., Young, H.C., Rajeev, S.B., Patel, D.V., Patel, M.M., Schmidt, R.J., Weller, H.N., Andahazy, M.L., Carboni, J.M., Gullo-Brown, J.L., Lee, F.Y.F., Ricca, C., Rose, W.C., Yan, N., Barbacid, M., Hunt, J.T., Meyers, C.A., Seizinger, B.R., Zahler, R. and Manne, V. *J. Med. Chem.,* 39, 224-236, (1996).

Mangues, R., Corral, T., Kohl, N.E., Symmans, W.F., Lu, S., Malumbres, M., Gibbs, J.B., Oliff, A. and Pellicer, A. *Cancer Res.,* 58, 1253-1259, (1998).

Möhle, K., Gumann, M. and Hofmann, H.J. *J. Comput. Chem.*, 18, 1415-1430, (1997)

Omer C.A. and Kohl, N.E. *Trends Pharmac. Sci.*, 18, 443-444, (1997)

Park, H.W., Boduluri, S.R., Moomaw, J.F., Casey, P.J. and Beese, L.S. *Science,* 275 1800-1804, (1997).

Pearlman, D.A., Case, D.A., Cadwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A. *AMBER4.0*, University of California, San Francisco, CA, (1991).

Perez, J.J., Villar, H.O. and Arteca, G.A. *J. Phys. Chem*., 98, 2318, (1994).

Perez, J.J.; Villar, H.O.; Loew, G.H. *J. Comp. Aided Molec. Design*., 6, 175, (1992).

Pettit, B.M, Matsunaga, T., Al-Obeidi, F., Gehring, C., Hruby, V.J. and Karplus, M. *Biophys. J.,* 60, 1540, (1991).

Qian, I., Sebti, S.M. and Hamilton, A.D. *Biopolymers,* 43, 25-41, (1997)

Reiss, Y., Goldstein, J.L., Seabra, M.C., Casey, P.J. and Brown, M.S. *Cell,* 62, 81-88, (1990).

Reiss, Y., Stradley, S.J., Gierasch, L.M., Brown, M.S. and Goldstein, J.L. *Proc. Natl. Acad. Sci. USA*., 88, 732-736, (1991).

Schafer, W. R. and Rine, J. *Annu. Rev. Genet.,* 30, 209-237, (1992)

Schaumann, T., Braun, W. and Wüthrich, K. *Biopolymers,* 29, 679, (1990).

Scheraga, H.A. *Computer simulation of biomolecular systems*, Vol. 2, eds. van Gusteren, W.F., Weiner, P.K. and Wilkinson, A.J. (Escom, Leiden, 1993), pp. 231-248.

Seabra, M.C., Reiss, Y., Casey, P.J., Brown, M.S. and Goldstein, J.L., *Cell,* 65, 429-434, (1991).

Sebti S.M. and Hamilton, A.D. *Pharmacol. Ther.,* 74, 103-114, (1997).

Singh, U.C. and Kollman, P.A. *J. Comput. Chem.,* 5, 129-145, (1984)

Stradley, S.J., Rizo, J. and Gierasch, M. *Biochemistry,* 32, 12586-12590, (1993).

Sun, J., Qian, Y., Hamilton, A.D. and Sebti, S.M. *Cancer Res.,* 55, 4243-4247, (1995).

Volkenstein, M.V. *Configurational statistics of polymeric chains;* Interscience Publishers: New York, (1969).

von Freyberg, B.; Braun, W. *J. Comput. Chem.,* 12, 1065, (1991).

Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. *J. Comput. Chem.,* 2, 230-252, (1986)

Yokohama, K., Trobridge, P., Buckner, F.S., Van Voorhis, W.C., Stuart, K.D. and Gelb, M.H. *J. Biol. Chem.,* 273, 26497-26505, (1998).

Zhang, F.L., Diehl, R.E., Kohl, N.E., Gibbs, J.B., Giros, B., Casey, P.J. and Omer, C.A. *J. Biol. Chem.,* 269, 3175-3180, (1994).