10 March 2000

Chemical Physics Letters 319 (2000) 65–70

CHEMICAL PHYSICS LETTERS

www.elsevier.nl/locate/cplett

Evaluation of the iterative simulated annealing technique in conformational search of peptides

Francesc J. Corcho [a], Marta Filizola [b], Juan Jesús Pérez [a,*]

[a] Dept. d'Enginyeria Quimica, UPC, ETS d'Enginyers Industrials de Barcelona. Av. Diagonal 647, 08028 Barcelona, Spain
[b] Centro di Ricerca Interdipartimentale di Scienze Computazionali e Biotecnologiche (CRISCEB), Seconda Università degli Studi di Napoli, Via Costantinopoli 16, 80138 Napoli, Italy

Received 21 August 1999

Abstract

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. The iterative simulated annealing procedure was recently proposed as an effective procedure to overcome these difficulties. In the present work results of a thorough exploration of the conformational space of the peptide Ac–Cys–Val–Tic–Met performed by means of the simulated annealing procedure is compared to the results of a random search. Profile differences in the two sets of low energy conformations obtained are analyzed. The results are also discussed in terms of the rotational isomeric model and its usefulness in assessing the degree of completeness of the conformational search. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Potential energy surfaces of peptides are characterized by the presence of a large number of valleys (minima) separated by high mountains and ridges [1]. 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 with a thorough exploration of the conformational space. This problem is referred in the literature as the *multiple minima problem* and it is still an open question, despite the many techniques available to explore the conformational space of peptides [2,3].

Some insights into the features of the conformational space can be assessed from the study 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 [4–9], 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

---

[*] Corresponding author. Fax: +34-93-401-7150

PII: S0009-2614(00)00081-6

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 [10], that has been used to describe rough energy landscapes and used to provide a model of the statistical mechanics of protein folding [11] where the density of states is assumed to have a Gaussian distribution. On the other hand, the rotational isomeric approximation [12–14] provides a similar distribution from the assumption that the molecule is composed of N independent rotors [15].

Knowledge of the density of states profile has implications in the understanding of the efficacy of different procedures used to sample the conformational space as well in the localization of the global minimum and the subset of low energy conformations. Thus, when the conformation of a polypeptide is generated by a random assignment of its dihedral angles, after energy minimization, the most likely situation is to locate a conformation with an energy around the value of the maximum of the density of states profile. Furthermore, a semi-quantitative measure of the extent of the conformational space sampled can be assessed by estimating the number of unique energy minima expected for a given polypeptide as well as the energy corresponding to the maximum of the distribution based on the number and characteristics of the rotors involved [15].

The present work describes the results of two explorations of the conformational space carried out on the tetrapeptide Cys–Val–Tic–Met using two different strategies: a random search and the iterative simulated annealing (SA). 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 states of this system that can be done using the rotational isomeric approximation.

2. Methods

All calculations were carried out within the molecular mechanics framework using the all-atom AMBER 4.0 force field [16]. The tetrapeptide selected for the present work is a conformationally constrained inhibitor of farnesyltransferase with the sequence: Cys–Val–Tic–Met, where Tic stands for the unnatural amino acid (S)-1,2,3,4-tetrahydroisoquinoline-3-carboxylate. The peptide was studied in its zwitterionic form and no explicit solvent was included in the calculations, although an effective dielectric constant of 80 was used to screen electrostatic interactions. Furthermore, no cutoff 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.

The conformational space was explored by two methods: (i) a simulated annealing (SA) procedure used in an iterative fashion and (ii) a random search exploration. The simulated annealing protocol used in the present work has been described elsewhere [17]. The starting extended structure is first minimized and then quickly heated up to 900 K at a rate of 100 K/ps. Subsequently, the structure is slowly cooled up 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 starting conformation for a new cycle of SA. The conformation is heated again at the same rate in order to jump to a different valley and then cooled again slowly. The procedure is subsequently followed in an iterative fashion so to generate an energy rank ordered library of low energy conformations ordered by energy.

The random search was done by generating 12 000 inputs for the PREP module of AMBER for which the backbone dihedral angles had been generated randomly. Angles ω_1 and ω_3 were fixed at 180°, and ω_2 was modified randomly to exhibit values of 0 or 180 degrees. The structures were then generated separately using the AMBER program and subsequently minimized using the conjugate gradient algorithm until the 0.001 kcal/mol convergence criterion was fulfilled.

The set of conformations generated through the random search approach and by SA were checked for uniqueness separately by removing of those structures for which at least one of the backbone dihedral angles was different from 60° in respect to any of the previous conformations already stored in the library. The conformational search was stopped when the method reached a low sampling efficiency

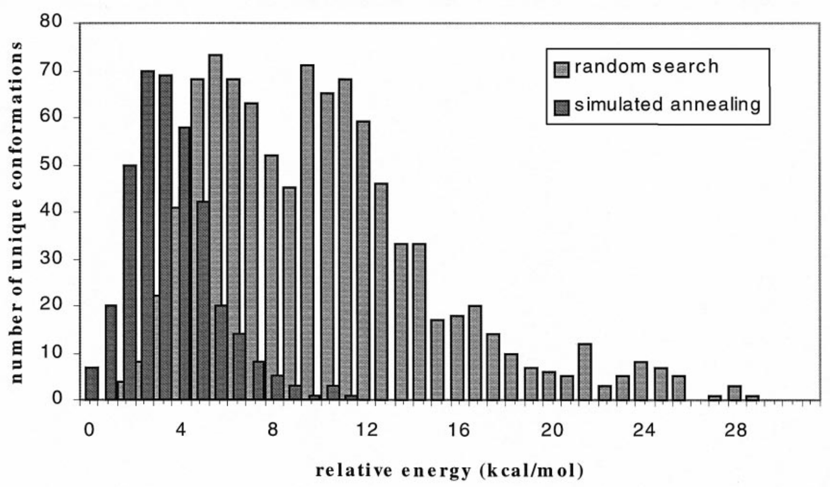

Fig. 1. Histogram of the unique conformations ranked order by energy obtained with the iterative simulated annealing procedure (dark bars) and random search (light bars) referred to the lowest energy conformation found.

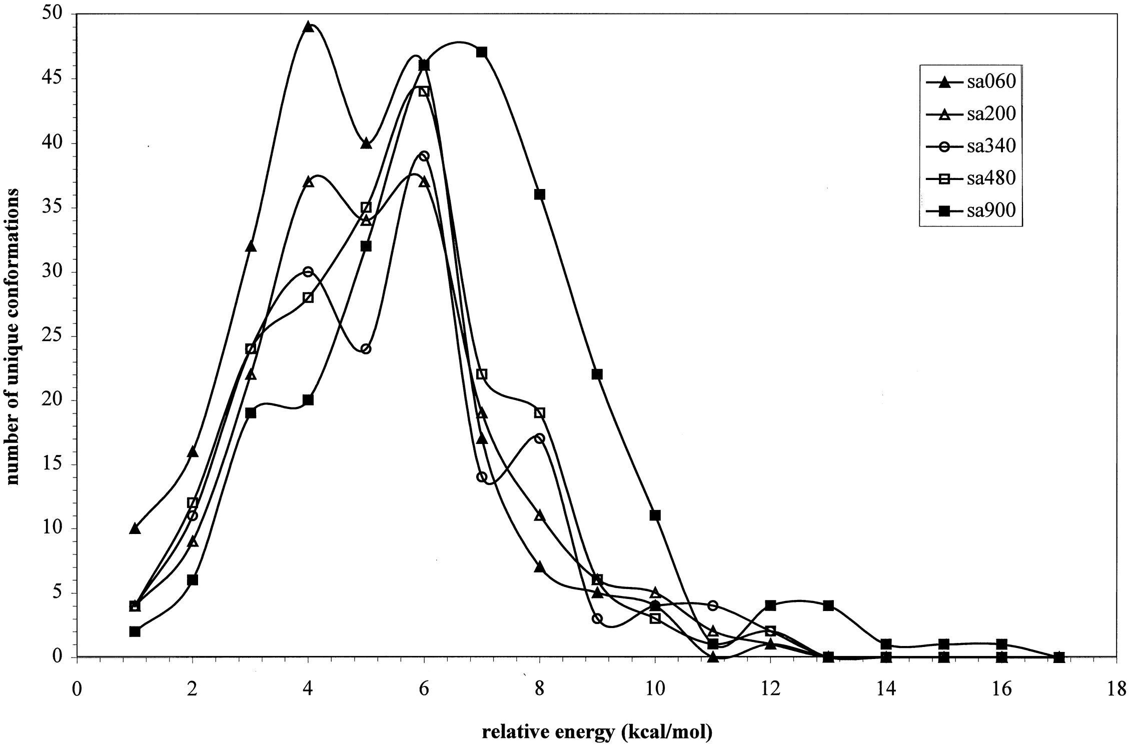

Fig. 2. Distributions of unique conformations obtained by a simulated annealing protocol at different temperatures of quenching: 900 K solid squares; 480 empty squares; 340 empty circles; 200 K empty triangles; 60 K solid triangles.

(< 0.1), defined as a function of the number of unique conformations (ξ) obtained after N cycles of SA.

3. Results and discussion

Energy distributions of unique conformations after 12 000 cycles either using the iterative SA or the random search are shown in Fig. 1. 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 above the global minimum, whereas the random search exhibits a bimodal distribution with one of the maxima at around 7 kcal/mol and the other at around 10 kcal/mol. These results clearly show that the SA procedure samples more efficiently low energy conformations.

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. Fig. 2 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. 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. Therefore when the temperature of quenching in the simulated annealing protocol is decreased, the conformations characterized exhibit lower energies, in any case always at the left hand side peak of the random search.

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

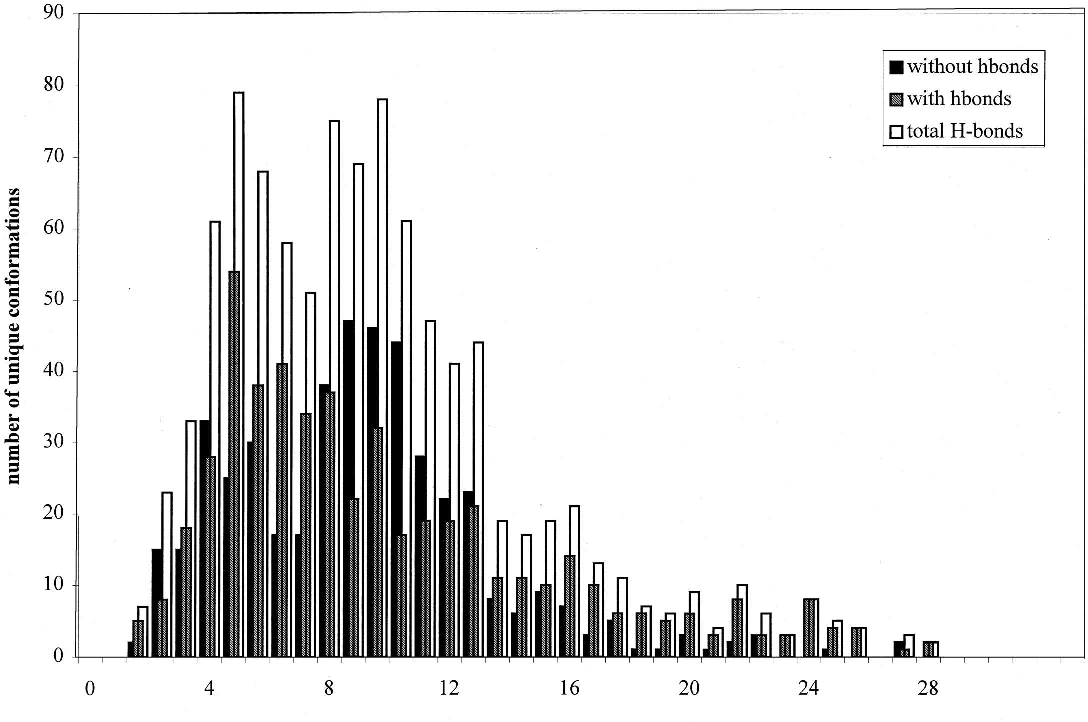

Fig. 3. Histogram of the unique conformations obtained from the random search (in white). Those conformations that exhibit at least a hydrogen bond are depicted in grey and those conformations not exhibiting a hydrogen bond are depicted in black.

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. Figs. 3 and 4 show energy histograms of all the conformations compared to those with and without a hydrogen bonds for the two searches of the conformational space. Figs. 3 and 4 suggest that: conformations that exhibit hydrogen bonds are more abundantly found at energies 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 torsional degrees of freedom, f; the mean number of rotamers of each rotor, $m+1$; and the window energy width, ε_0 [15]:

$$U_0 = fm\varepsilon_0/2$$

In order to get the number of effective rotors, f, it is necessary to carry out a fit of distribution of states at low energies to the bosonic limit of the distribution in the rotational isomeric model. Assuming that the distribution obtained from the SA procedure is more accurate than the other at low energies, this distribution can be used to compute a value for f. In the bosonic limit, the density of states has an exponential dependence with the energy, being in the present case: $\ln \Omega = 1.32 + 0.78 E_0$. The second term of this equation is consistent with the energy width of 0.8

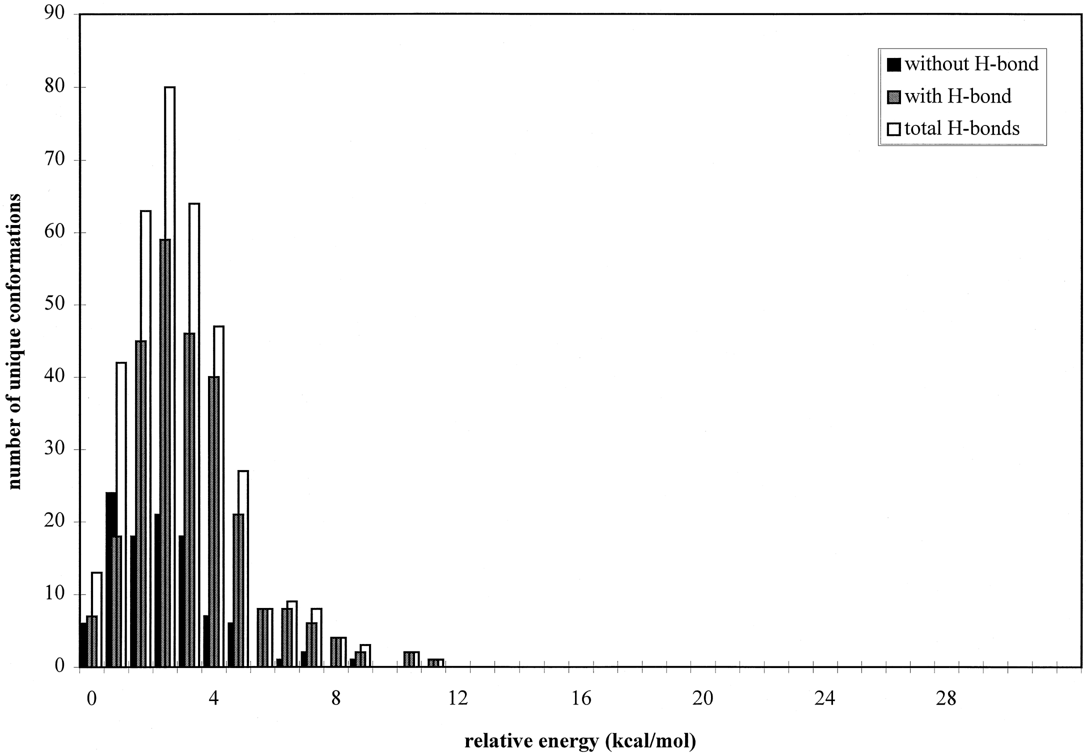

Fig. 4. Histogram of the unique conformations obtained from the simulated annealing protocol (in white). Those conformations that exhibit at least a hydrogen bond are depicted in grey and those conformations not exhibiting a hydrogen bond are depicted in black.

kcal/mol 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 $\varepsilon_0 = 0.8$, results in a value for U_0 of 9.6 kcal/mol in reasonable agreement with the position of the second maximum found in the random search of the conformational space.

4. Discussion

The results discussed previously suggest that the density of states of the molecule 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. This argument is consistent with the fact that conformations found using the SA procedure lie always in this region. 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 on the nature of the approximation of the model. 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 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 for the prediction power of the model that in order to make reliable prediction 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.

References

[1] H. Frauenfelder, S.G. Sligar, P.G. Wolynes, Science 254 (1991) 1598.
[2] R.E. Bruccoleri, E. Haber, J. Novotry, Protein folding, in: L.M. Gierasch, J. King (Eds.), American Association for the Advancement of Science, Washington, 1990, pp. 259–270.
[3] H.A. Scheraga, in: W.F. van Gusteren, P.K. Weiner, A.J. Wilkinson (Eds.), Computer Simulation of Biomolecular Systems, Vol. 2, Escom, Leiden, 1993, pp. 231–248.
[4] P. Auffinger, G. Wipff, J. Comput. Chem. 11 (1990) 19.
[5] T. Schaumann, W. Braun, K. Wüthrich, Biopolymers 29 (1990) 679.
[6] B. von Freyberg, W. Braun, J. Comput. Chem. 12 (1991) 1065.
[7] B.M. Pettit, T. Matsunaga, F. Al-Obeidi, C. Gehring, V.J. Hruby, M. Karplus, Biophys. J. 60 (1991) 1540.
[8] J.J. Perez, H.O. Villar, G.H. Loew, J. Comp. Aided Molec. Design 6 (1992) 175.
[9] A.V. Shah, D.P. Dolata, J. Comput. Aided Molec. Design 7 (1993) 103.
[10] B. Derrida, Phys. Rev. B 24 (1981) 2613.
[11] J.D. Bryngelson, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 84 (1987) 7524.
[12] D.A. Pearlman, D.A. Case, J.C. Cadwell, G.L. Seibel, U.C. Singh, P. Weiner, P.A. Kollman, AMBER4.0, University of California, San Francisco, CA, 1991.
[13] M.V. Volkenstein, Configurational Statistics of Polymeric Chains, Interscience Publishers, New York, 1969.
[14] J.J. Perez, H.O. Villar, G.A. Arteca, J. Phys. Chem. 98 (1994) 2318.
[15] P.J. Flory, Statistical Mechanics of Chain Molecules, Interscience Publishers, New York, 1969.
[16] J.J. Perez, H.O. Villar, G.A. Arteca, J. Phys. Chem. 98 (1994) 2318.
[17] M. Filizola, N.B. Centeno, J.J. Perez, J. Pept. Sci. 3 (1997) 85.