

Potential energy surface and kinetics of the helix–coil transition in a 33-peptide

Giorgia Brancolini · Alessandro Venturini ·
Francesco Zerbetto

Received: 24 October 2006 / Accepted: 22 November 2006 / Published online: 20 January 2007
© Springer-Verlag 2007

Abstract The CHARMM force field is used to calculate 36 minima along the potential energy surface of the helix–coil conversion of Ac-A₁₄KG₃A₁₄K + 2H⁺ and to interconnect them through 35 transition states. The energy barriers are used to give the rate constants of interconversion between the conformers and the relevant kinetic equations are then solved. Fair to good agreement with the data obtained by drift time spectrometry experiments (D. T. Kaleta, M. F. Jarrold, *J. Am. Chem. Soc.* 125:7186, 2003) is obtained under a simple hypothesis of the initial conformer distribution.

Keywords Gas-phase kinetics · Conformational transitions · Molecular modeling · Peptides · Potential energy surfaces

1 Introduction

Over the last decades the fundamental biological role and the complexity of proteins have stimulated a number of important experimental and conceptual developments. On the computational side, the combination

of empirical force fields and molecular dynamics (MD) has produced substantial advances in our comprehension of their properties and functions. The widespread use of MD stems from the necessity of including thermal energy in the appraisal of the properties of these large systems and from the presence of additional complicating features such as water molecules and ions, which are characterized by high mobility. This computational approach differs greatly from that used for smaller molecules, which is dominated by the characterization of potential energy surfaces in terms of minima and transition states. One of the implicit ideas in the use of MD is that general aspects, rather than details, of the landscape govern the dynamics, even, or perhaps especially, when major structural rearrangements, such as folding–unfolding transitions, occur. On the top of it, the characterization of the full potential energy surface of even a small protein is a daunting prospect that requires a remarkable amount of patience, to say the least. And yet, while MD simulations do provide important answers, the amount of data produced is often impressive and wading through them may not convey the straightforward notion that stationary points of the potential energy surface do.

Experimental techniques allow now the study of small, or even medium size, unsolvated proteins and provide a host of information on their global shape and dynamics [1,2]. Indeed, coupled with variations of temperature, drift time spectrometry of molecules volatilized through electrospray is a very powerful technique able to give information on the shape of the molecules and on the kinetics that takes place between the isomers/conformations generated during the “Coulomb explosion” of the droplets [3]. The technique also strips the water molecules, which simplifies the systems both

Contribution to the Fernando Bernardi Memorial Issue.

G. Brancolini · A. Venturini (✉)
ISOF, Consiglio Nazionale delle Ricerche,
via Gobetti 101, 40129 Bologna, Italy
e-mail: A.Venturini@isof.cnr.it

G. Brancolini
e-mail: brancolini@isof.cnr.it

F. Zerbetto (✉)
Dipartimento di Chimica “G. Ciamician”,
Università di Bologna, V. F. Selmi 2, 40126 Bologna, Italy
e-mail: francesco.zerbetto@unibo.it

in the analysis of the experiments and in the computer simulations [4–7]. With this approach, Kaleta and Jarrold studied Ac-A₁₄KG₃A₁₄K + 2H⁺, a medium size polypeptide with 33 residues [1,2]. The two A₁₄ segments, forming half of the molecule, prefer to form helices [8]. The central G₃ link introduces a low propensity towards a helix in the mid of the polypeptide [9,10]. The authors hoped to observe, and actually observed, Helix–Turn–Helix transitions as the temperature in their apparatus increased from 153 to 413 K. In the electrospray experiments, the peptides are generated in conformations of high energy either during the desolvation process or when the ions are collisionally heated as they enter the drift tube. When the temperature of the drift tube is 153 K, three main conformations appear in the spectrum. They become 2 at 233 K, and are reduced to one at 413 K. MD simulations found a coiled-coil structure ~7–8 kcal mol⁻¹ more stable than an extended helix conformation. The analysis of calculated cross sections showed that the coiled-coil gives the dominant, shorter drift-time peak, while the extended helix is responsible for the longer drift-time peak. The MD simulations did not reveal any low-energy conformations that could account for the middle feature in the distribution.

The study of the Potential Energy Surface, PES, of this rather large unsolvated peptide offers the opportunity to bridge the gap between MD simulations and the characterization of potential energy surfaces in terms of minima and transition states (actually saddle points). In particular, the design of the molecule, and the reasonably well-defined coordinate along which the dynamics occurs, allows the reduction of the number of degrees of freedom that must be explored (although in the end, 36 minima and 35 interconnecting transition states are presented out of the plethora that were located). The dynamics is here introduced a posteriori in two steps. First, the energy and the vibrational frequencies of the stationary points provide the input for the activated complex theory that gives the rate constants of interconversion between adjacent minima. Then the kinetic equations are solved to give the change of molar fraction in time. The present approach is related to that carried out in a recent study of two methyl capped dipeptides, *N*-acetyl tryptophan methyl amide, and, *N*-acetyl tryptophan amide [11].

2 Computational background

The characterization of the potential energy surface of the peptide was carried out with the Charmm(21.7) force field [12–14] and Path algorithm as implemented in the

Tinker program [15–17]. Minima and transition states were fully optimized and convergence to the stationary points was obtained when the RMS of the gradient was less than 0.0002 kcal mol⁻¹ Å⁻¹. At each side of the transition states, the minima were identified by relaxing the structures. Activated complex theory was used to evaluate the rate constants, *k*, at several temperatures, *T*,

$$k = \frac{k_B T}{h} \frac{Q_{\text{TS}}}{Q_{\text{react}}} \exp\left(-\frac{\Delta E_{\text{act}}}{k_B T}\right) \quad (1)$$

where *k_B* and *h* are Boltzmann and Planck constants, *Q* is the partition function, calculated in the harmonic approximation at the chosen temperature, of either the minimum (reactant) or the transition state, and Δ*E_{act}* is the activation energy of the reaction.

The helix-coil transition between 36 intermediates generates a set of 36 differential equations that in matrix form read

$$\frac{d\mathbf{X}}{dt} = -\mathbf{k}\mathbf{X} \quad (2)$$

where **X** is the vector of the fractions of the 36 species that vary in time, and **K** is a 36 × 36 square non-symmetric matrix of rate constants. The sum of the matrix elements of each column of **k** is zero. Since **k** does not change in time and is not a function of **X**, the solution of Eq. 2 can be cast as [18]

$$\mathbf{X}(t) = \sum_{i=1}^N \alpha_i \mathbf{C}_i \exp(\lambda_i t) \quad (3)$$

where **C_i** is the *i*th eigenvector of matrix **k** associated to the λ_{*i*} eigenvalue, and α_{*i*}'s are a set of scalar coefficients determined by the initial conditions, α = **C**⁻¹**X**(**t** = **0**). The eigenvalues, λ_{*i*}, are all negative and real for a kinetic problem, apart from one that is zero and determines **X**(*t* = ∞), i.e. the equilibrium composition.

In order to perform some comparison with the drift time experiments, it was necessary to calculate the surface area of the minima. The more compact a structure the fewer the collisions and the more rapid its travel. The correlation between a molecular structure and the gas phase mobility has been thoroughly addressed before [19,20]. Here we use a simple hard-sphere approach where the surface area accessible to the gas atoms is calculated by numerical integration [21,22]. To each conformer was then associated a drift time proportional to its surface accessible area that made the first peak coincide with the experiments. This allows the simulation of a stick spectrum where the lines were further broadened by a Gaussian function with a linewidth of 60 μs. The approach must be intended as semi-quantitative. It

lacks the variation of cross-section due to two phenomena: the first is the interconversion between conformers that still takes place during the flight in the drift tube, the second, and perhaps less important feature, is the variation of the surface area that is caused by the temperature dependence of the large amplitude motions of each species even in the absence of conversion.

3 Results and discussion

The potential energy curve: Figure 1 illustrates the potential energy curve of the helix-coil transition of Ac-A14KG3A14K + 2H⁺. Odd-numbered structures are minima, even-numbered structures are transition states. The global minimum is structure number 43 [1,2].

This structure and those with the same surface area are the first to appear in the drift time experiments and correspond to structure “a” identified experimentally. Analogously, on the left side, structures 1–3 19–27, correspond to structures of type “b” [1,2]; on the right side, minima from 49 to 59, are the structures of type “c” [1,2]. Apart from the three regions that find a match with the experiments (see below), also structures 69 and 71 at the end of Fig. 1 are energetically distinct and may be observable. However, their surface areas are so similar to those of species of type “a” that they only contribute to this experimental peak.

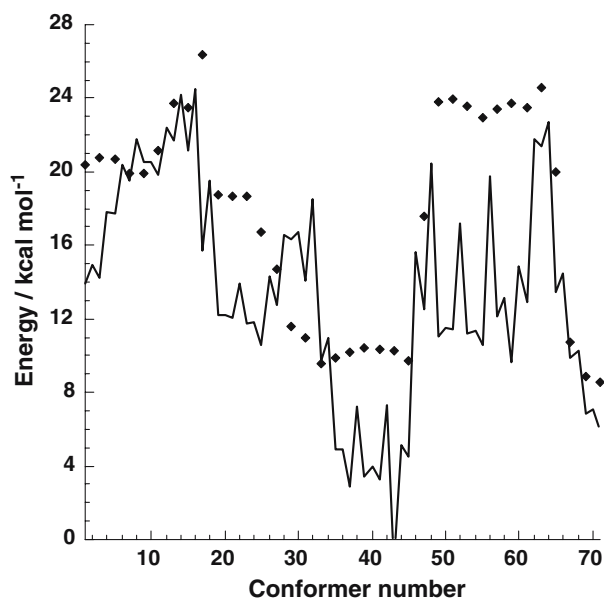


Fig. 1 Energies of minima and transition states of Ac-A14KG3A14K + 2H⁺. The values of the coordinates on the y-axis of the diamonds are proportional to the accessible surface areas of the stable conformers

The minima: Figure 2 shows a pictorial representation of a few selected minima. They are 1, 15, 25, 37, 43, 59, 63, and 71. They were chosen because of their representativeness and show that all the main helix-coil motives (α -helices, 3_{10} helices, π -helices, turns, and coils) are present and exchanged along the folding pathway.

Rearrangements and hinges: The set of Fig. 3 shows the variation of the Φ and Ψ angles along the folding-coil coordinate. Most angles do not change along the reaction path, but a few vary strongly and allow the identification of the residues that work as hinges in the folding-unfolding process.

Along the entire pathway, we identify 11 hinges. They are A1, A8, A12, A14, K15, G16, G17, A21, A27, A32, and K33 (A for Alanine, G for Glycine and K for Lysine). Table 1 reports the hydrogen bonds formed by each of the 13 hinges along the reaction pathway. The alanines of the polypeptide tend to form $i, i+4$ interactions (α -helices) and $i, i+5$ interactions (π -helices), although some $i, i+3$ (3_{10} -helices) are present. The propensity of alanine and lysine to form α -helices is well-known, as is the tendency of glycine to form turns. Table 1 also allows us to examine the dynamics associated to the angle variation of Fig. 1 from the initial conformer to the final.

A1 participates to a helix at high energy, which is then partly forfeited in favor of end-to-end interactions in the most stable region.

A8 is mainly involved in π - and α -helices; in the less stable regions, angle variations occur upon removal of the helix formed by A8 with peptides towards the C-terminus.

A12 is involved in a π -helix with residues towards the N-terminus. Along the pathway, its H-bonding occasionally switches to the C-terminus.

A14 is involved in a turn with residues towards the N-terminus. Angle variations occur in order to form a π -helix towards the C-terminus.

K15, in the initial, higher energy structures, the residue is involved in short range $i, i \pm 2$ interactions, angle variations occur when it becomes part of a helix.

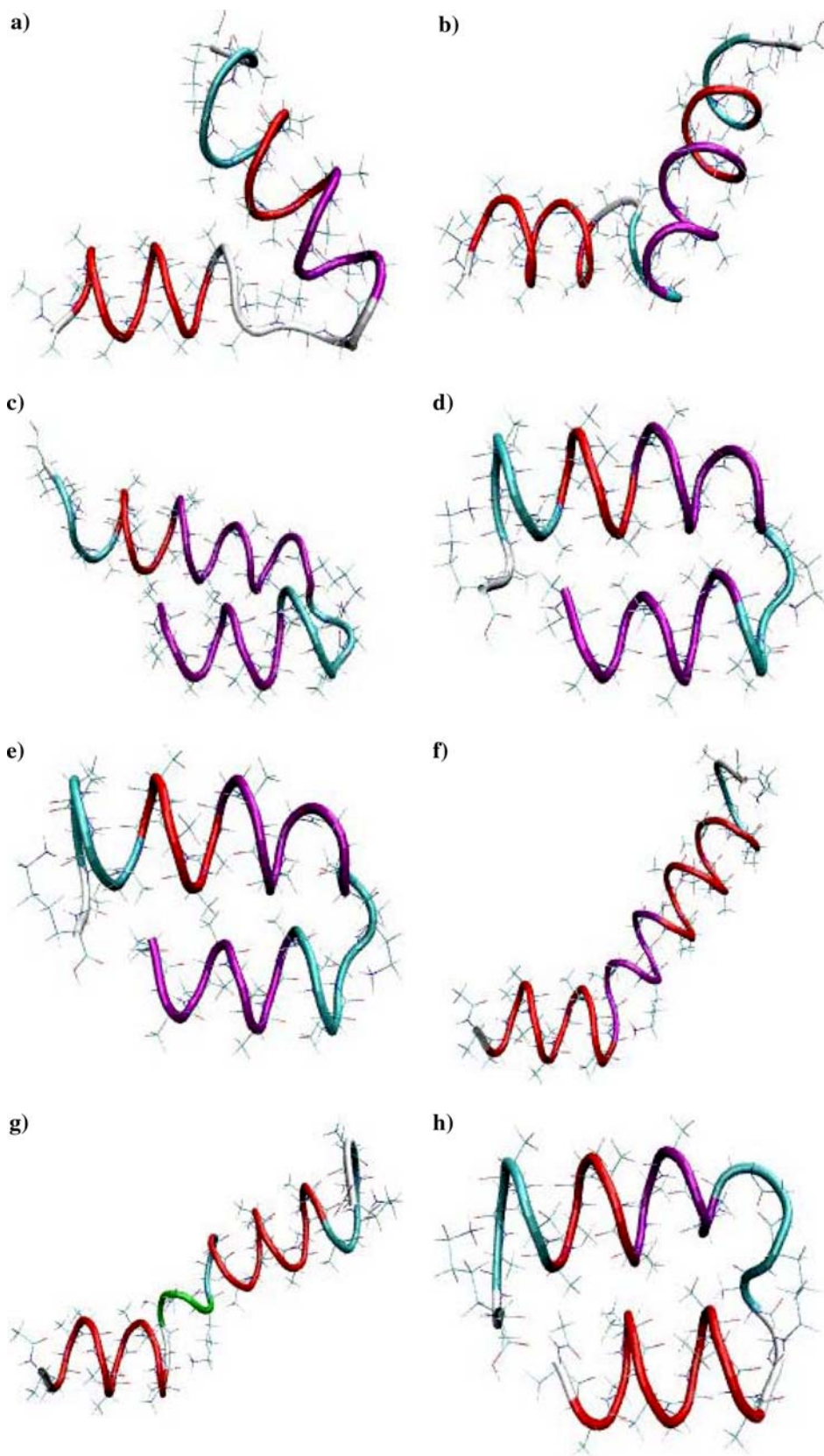
G16 in the low energy region forms helices, while in the high-energy regions forms short range $i, i \pm 2$ interactions.

G17, in the more stable region, forfeits the H-bonds with the N-terminus.

A21 tends to swap between π - and α -helices with residues located towards the N-terminus.

A27, in the more stable region, forfeits a π -helix towards the C-terminus.

Fig. 2 a minimum 1, b minimum 15, c minimum 25, d minimum 37, e minimum 43, f minimum 59, g minimum 63, h minimum 71. Color coding: α -helices are purple, 3_{10} helices are green, π -helices are red, turns are cyan, coils are white



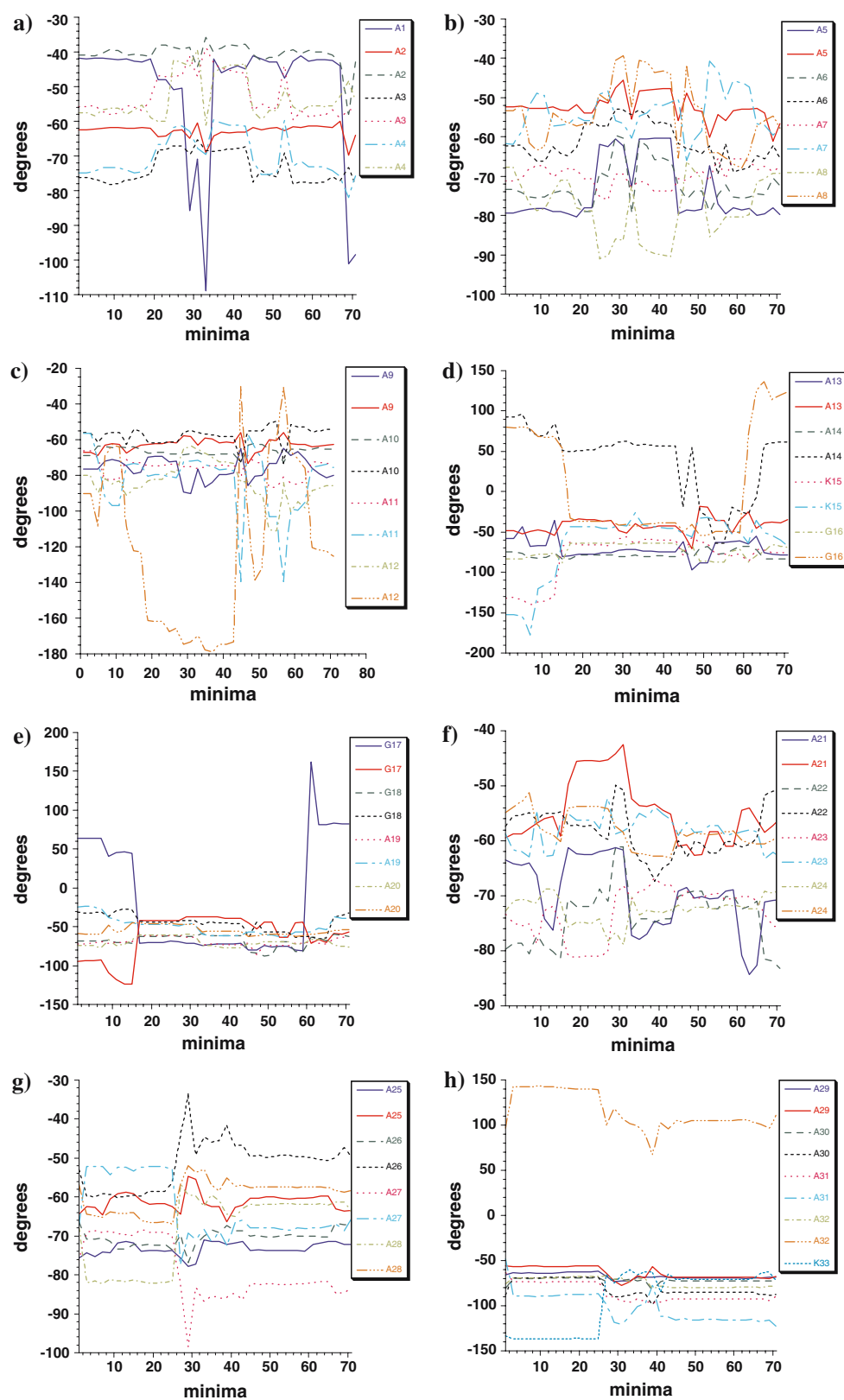


Fig. 3 Variation of the Φ and Ψ angles of the 33 peptides along the reaction coordinate. In the legends, the initial letter identifies the type of residue (A for Alanine, G for Glycine and K for

Lysine), the numbering is progressive, with the Φ angles before the Ψ angles. For the initial residue, N terminus, there is only Ψ , for the final residue, C-terminus, there is only Φ

Table 1 Summary of the interactions of the hinge residues along the folding/unfolding pathway

Hinge residue	Stationary points	H-bonded to residue X		Hinge residue	Stationary points	H-bonded to residue X	
		NH _{hinge} ---OC _X	CO _{hinge} ---HN _X			NH _{hinge} ---OC _X	CO _{hinge} ---HN _X
A1	1–23		A5	G16 (contd)	17–31		A20
			A6		33–43		A20
	25–45		A5				A21
	29	A26			45		A20
	31	A32			47–51		A21
	33–43	K33			53–59	A12	A21
	47–51		A5		61	A12	
			A6		63–65	A13	A18
	53		A5		67–71	A13	
			A6				
A8	1–3	A3	A13	G17	1–13	K15	A21
	5–17	A3	A14		15	A14	A21
	19–23	A3				K15	
	25	A4			17–25	A19	A21
	27	A4			27–29		A21
		A3			31	A14	A21
	29–31	A4			33–35	A14	A21
	33	A3					A22
	35–43	A4			37–43		A22
	45	A3	A13		45		A21
	47	A3	A13		47		A22
			A14		49–51	A13	A22
	49–55	A3	A14		53–55	A12	A22
	57	A3	A14		57–59	A13	A22
	59–71	A3	A14		61–65	K15	A22
					67–71	A13	A21
A12	1–51	A7		A21	1–31	G17	A26
	53–55	A7	G16		33–35	G16	A26
	57		G16			G17	
			K15		37–43	G16	A26
	59–63	A7	G16		45	G17	A26
			K15		47–59	G16	A26
	65–71	A7	K15		61–65		A26
			67–71	G17	A26		
A14	1–3	A9		A22	1–31	G18	A27
	5–13	A8			33–65	G17	A27
	15	A8	G17		67–71	G18	A27
	17	A8	G18				
	19–45	A9					
	47	A8	G18		1–37	A21	A31
	49–55	A8	G19		39	A21	A31
	57	A7	G19				A32
	59–65	A8	G19		41–71	A21	A31
	67–71	A8			29		A1
K15	1–15	A13	G17	A27	1–25	A22	A32
	17–45	A13	G19		27–71	A22	
	47	A13	A20				
	49–51	A9	A20		1–25	A27	
	53–55	A8	A20		27–37		
	57–63	A12	A20		31		A1
	65	A13	A20		39	A26	
	67–71	A13	A19		41–71		
G16	1–7		G18	K33	1–25		
			A19		27–71	A31	
	9–15		A19		33–43		A1
			A20				

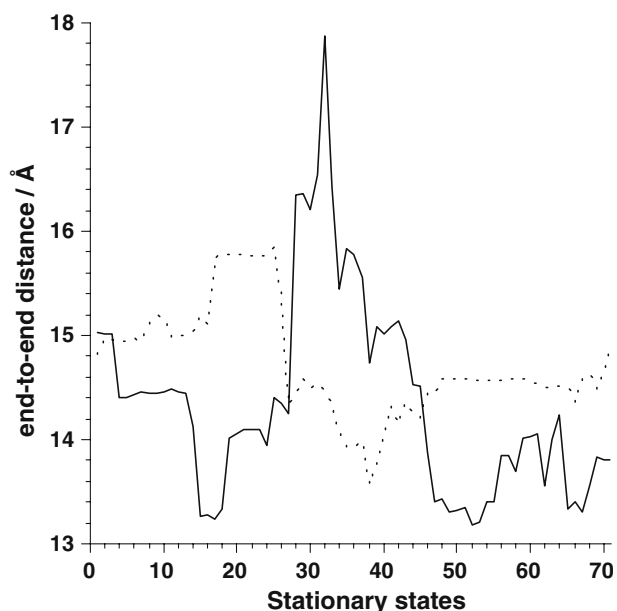


Fig. 4 End-to-end distance of A1–A14 (solid line) and A19–A32 (dotted line)

A32 has one structure, **31**, where there is a H-bond with the N-terminus residue.

K33, in the more stable region, forms end-to-end H-bonds with residue A1.

The global mechanism: Figure 4 provides further simple insight into the helix-coil transition mechanism of this 33-peptides. Two opposing forces drive the conformational transitions: the two alanine chains favor the formation of helices, while the other residues tend to favor coils. Both of the (Ala)₁₄ fragments will therefore remain helices as much as they can. In order to maximize the interchain interactions, however, they are forced to re-arrange their helices. A global index of such re-arrangements is the end-to-end distance of the two chains that shrink and elongate along the entire folding pathway. In the Ala1–Ala14 segment, the distance varies by up to ~ 5 Å and is the longest when the helix is partly forfeited to a turn; in the Ala19–Ala32 segment, the variation is smaller and reaches ~ 2.5 Å. This oligo-peptide is only apparently symmetric: the behaviour of two chains is unbalanced by the central KG₃ moiety and by the terminal lysine residue so that the two alanine helices may be thought to play the accordion but at different, opposing tunes. Indeed, when one is in the most extended conformation, the other is moving towards its most compact.

The distribution of conformers: One of the important results of the experiments performed by Kaleta and Jarrold [1,2], is the presence of three types of

conformers at low temperatures. In the spectra, the species that are most stable appear first. They have the most compact structures. This is confirmed by the comparison of the positions of the diamonds in Fig. 1 (their y-coordinates are proportional to the surface areas of the conformers).

The set of structure, **49–59**, on the right of the most stable conformer have the largest accessible surfaces and will appear last in the drift time experiments. Their energy is ~ 10 kcal mol⁻¹ above that of the most stable structure, a value slightly in excess of that found by Kaleta and Jarrold of 7–8 kcal mol⁻¹ in their MD simulations [1,2].

The set of structures on the left, **1–3 19–27**, have intermediate surface accessible areas and, although their stability is similar to that of the “slowest” structures, they should arrive second at the detector. This set of structures was not actually observed in the MD simulations, perhaps because they are more labile owing to their lower number and the rather low barriers that convert them to more stable structure. The experimental increase of temperature from 153 to 413 K reduces the number of structures and ultimately brings them down to only one.

It is instructive to evaluate the equilibrium composition resulting from the data used in Fig. 1, at temperatures higher than those used in the drift time spectrometry experiments. Such temperatures may correspond, in an effective way, to that experienced by the oligo-peptide when the population of conformers is created with the electrospray. Of course, the process of Coulomb explosion is much more complicated than the generation of equilibrium at high temperature. Figure 5 shows the distribution of conformers at high temperatures. As explained in the computational section, in order to obtain a resemblance with the experiments, to each conformer we associate a drift time proportional to its accessible area, and make the time of the first peak coincide with that of the experiment. The population of each species is then broadened by a Gaussian function. The effective temperatures explored were 1,000, 2,000, and 3,000 K. At 1,000 K, the three peaks system is barely present. The experimental ratios observed at the end of the drift tube are similar to those estimated from the Boltzmann population at 2,000/3,000 K. However, the central peak is broader than in the drift time experiments. This implies that the surface areas of the conformers are more spread than what observed experimentally.

The kinetics: Our intent is not to simulate directly the experiments, but only to obtain qualitative insight into the accuracy of the potential energy surface. It is, however, rewarding to find that the simple

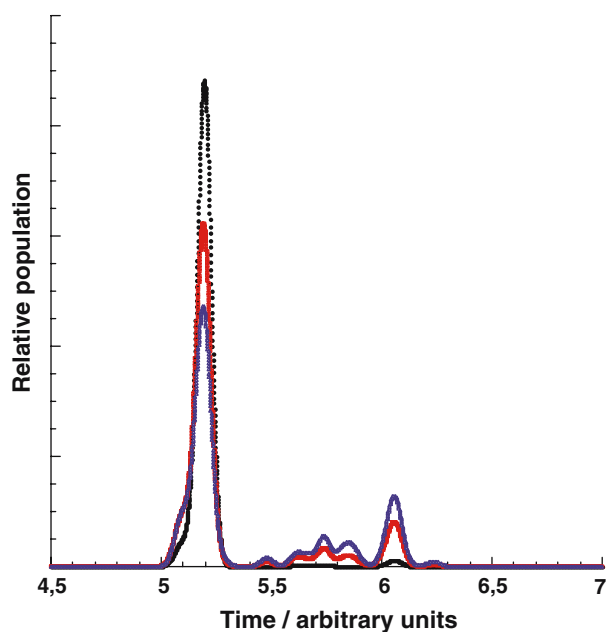


Fig. 5 Boltzmann populations obtained at 1,000 (*black*), 2,000 (*red*), 3,000 (*blue*) K from the potential energies of Fig. 1. The time in the x-axis is proportional to the accessible surface area, also given in Fig. 1

expedient of assuming an effective high-temperature, together with the data of Fig. 1, yields a trend rather similar to experiments where the population created by electrospray is locked by low temperature in the drift tube. If anything, the result vouches for the reliability of the simulations. Moreover, one can take the Boltzmann population at 2,000 K as $\mathbf{X}(t=0)$ and simulate the subsequent dynamics through Eqs. 2 and 3. In Eq. 3, the course of the reaction is mainly governed by the lowest eigenvalue, which is zero, and gives the equilibrium composition, and by the second lowest eigenvalue, since its contribution to the kinetics takes the longest time to taper off. When only one non-zero eigenvalue contributes, the kinetics is expressed as a single exponential and is due to the interconversion between two states. The quality of approximating the reaction dynamics to the second lowest eigenvalue only can be estimated by the ratio of λ_2/λ_3 , which must be $\ll 1$ [23]. Table 2 shows both λ_2 and the λ_2/λ_3 ratio at five temperatures, a ratio of 0.1 is deemed sufficient to give a two-state system. At all temperatures, apart from the highest one, Ac-A14KG3A14K + 2H⁺ behaves as a two-state system. This means that one state goes into another without accessing multiple pathways or straying off.

The timescale of the experiment is milliseconds. At 153, the first peak appears ~ 5 milliseconds and the tail of last one is over at 6.5 milliseconds [1,2]. At this

Table 2 Temperatures and dominant exponent, λ_2 , in the kinetics of Ac-A14KG3A14K + 2H⁺. The λ_2/λ_3 determines the accuracy of treating the kinetics as a single exponential process

T (K)	λ_2 (s ⁻¹)	λ_2/λ_3
153	-0.36E-05	0.0086
233	-0.44E+00	0.0040
313	-0.27E+04	0.0519
353	-0.40E+05	0.0976
413	-0.66E+06	0.3143

temperature, the molecular fractions that enter the drift tube are frozen, since the inverse of λ_2 is of the order of days.

From 313 K onward, the values of λ_2 are so large that equilibration occurs rapidly on the timescale of the experiments. The calculations are therefore not entirely accurate since the experiments still show the presence of a smaller side peak at 313 K.

Instructive is the case of the kinetics at 233 K, which is shown in detail in Fig. 6 (starting from the populations obtained from an effective temperature of 2,000 K). On the right side, the kinetics of the first 18 conformers, **1–35**, shows that they rapidly dwindle to zero. Notice that in this simulation, time is really in s⁻¹ and no fetch factor is used. On the left side, while the most stable structure increases its fraction rapidly, species **71** and **69** are unable to overcome the barrier of transition state **64**. These two structures, however, do not contribute to the secondary peak observed experimentally because their cross section is very close to that of the global minimum. The secondary peak, instead, is given by species **59**, which collects its “neighbors” on the left and is not able to overcome the barrier due to transition state **48**.

The calculated kinetics is faster than what observed experimentally. Although the differences could arise from the initial distribution of conformers that we use (i.e., the equilibrium composition at 2,000 K) we wondered if small changes of the potential energy surface would lead to significant improvements in the agreement with the experiments. We decided to stabilize species **59** by 2, 3 and 4 kcal mol⁻¹, and rerun the kinetics at 313 K. The molar fractions calculated after 5 milliseconds of kinetics were used to synthesize the spectra in Fig. 7 in a manner similar to that of Fig. 5.

An increased stabilization between 3 and 4 kcal mol⁻¹ brings experiments and simulations in good agreement. Notice that this amounts to decreasing the energy gap between the structures to that observed in the MD simulations by K&J [1,2].

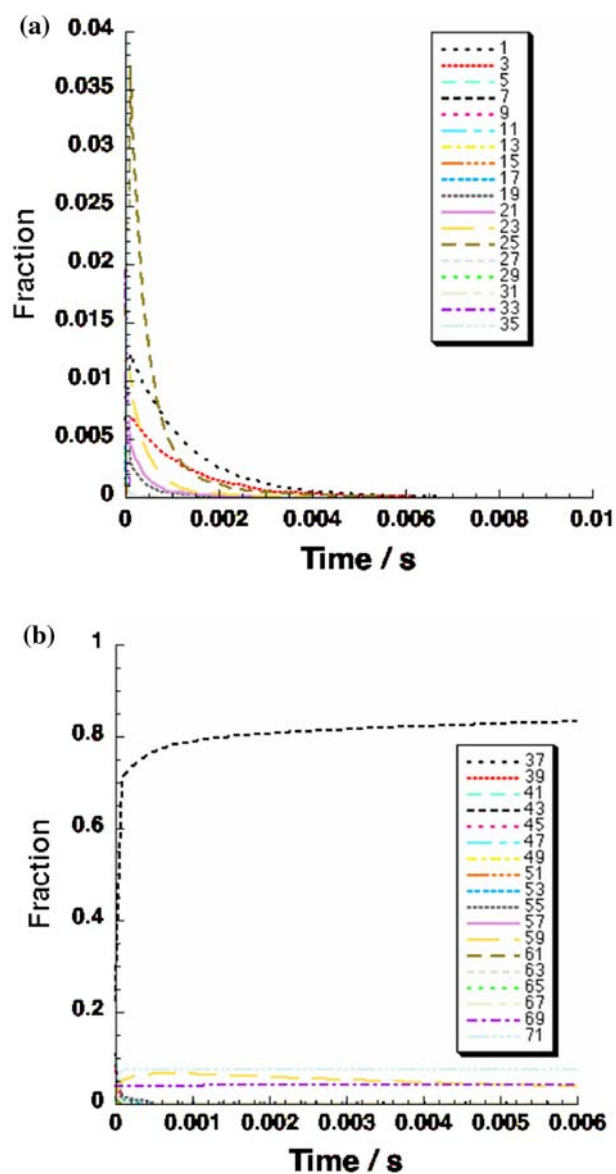


Fig. 6 Kinetics at 233 K (a) of the first 18 conformations, 1–35, (b) of the remaining 17 species, 37–71

4 Conclusion

We have calculated the pathway of the helix-coil conversion of a 33-peptide investigated by drift-time spectrometry. The molecule has an *ad hoc* central Gly₃ moiety with high propensity to give turns and two Ala₁₄ strands that form helices. The pathway entails 71 fully interconnected points (36 minima and 35 transition states) where a continuous exchange between various types of helices and turns appear. Although chemically identical, the location along the chains of the alanines makes them play a variety of structural roles, with some of them behaving as hinges for the helix-coil transition. The potential energy surface was tested in a kinetic

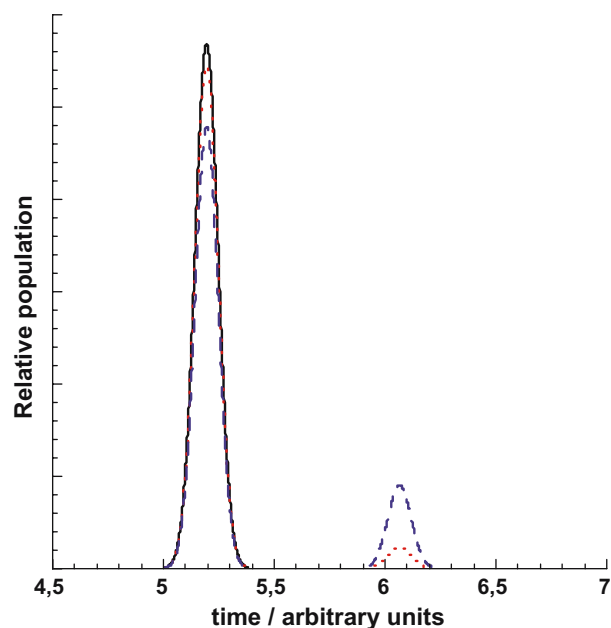


Fig. 7 Populations at 313 K, after 5 milliseconds of kinetics obtained stabilizing structure **59** by 2 kcal mol⁻¹ (black), 3 kcal mol⁻¹ (red), and 4 kcal mol⁻¹ (blue) K

model that gave reasonable agreement with the temperature dependent experiments, especially after one of the most stable points was stabilized by 3 kcal mol⁻¹.

Acknowledgment AV acknowledges partial financial support by ESF-SONS-BIONICS.

References

1. Kaleta DT, Jarrold MF (2003) *J Am Chem Soc* 125:7186–7187
2. Kinnear BS, Kaleta DT, Kohtani M, Hudgins RR, Jarrold MF (2000) *J Am Chem Soc* 122:9243–9256
3. Gaskell SJ (1997) *J Mass Spectry* 32:677–688
4. Kaltashov IA, Fenselau C (1997) *Proteins. Struct Func Genet* 27:165–170
5. Wyttenbach T, Bushnell JE, Bowers MT (1998) *J Am Chem Soc* 120:5098–5103
6. Arteca GA, Velázquez I, Reimann CT, Tapia O (1999) *Phys Rev E* 59:5981–5986
7. Ruotolo BT, Verbeck GF, Thomson LM, Gillig KJ, Russell DH (2002) *J Am Chem Soc* 124:4214–4215
8. Hudgins RR, Ratner MA, Jarrold MF (1998) *J Am Chem Soc* 120:12974–12975
9. Chakrabarty A, Baldwin RL (1995) *Adv Prot Chem* 46: 141–176
10. Hudgins RR, Jarrold MF (2000) *J Phys Chem B* 104:2154–2158
11. Evans DA, Wales DJ, Dian BC, Zwier TS (2004) *J Chem Phys* 120:148–157
12. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M M (1983) *J Comput Chem* 4: 187–217

13. Foloppe N, MacKerell Jr AD (2000) *J Comput Chem* 21: 86–104
14. Banavali N, MacKerell Jr AD (2000) *J Comput Chem* 21: 105–120
15. Ponder JW, Richards FJ (1987) *J Comput Chem* 8:1016–1024
16. Kundrot C, Ponder JW, Richards FJ (1991) *J Comput Chem* 12:402–409
17. Dudek MJ, Ponder JW (1995) *J Comput Chem* 16:791–816
18. Steinfeld JI, Francisco JS, Hase WL (1998) *Chemical kinetics*, 2nd ed. Prentice Hall, New Jersey (1998)
19. von Helden G, Hsu MT, Gotts N, Bowers MT (1993) *J Phys Chem* 97:8182–8192
20. Shvartsburg AA, Jarrold MF (1996) *Chem Phys Letters* 261:86–91
21. Connolly ML (1983) *J Appl Cryst* 16:548–558
22. Connolly ML (1985) *J Am Chem Soc* 107:9186–9187
23. Berezhkovski A, Szabo A (2004) *J Chem Phys* 121:9186–9187