Journal of Biomolecular NMR, **17:** 63–77, 2000.
KLUWER/ESCOM

The *cis*proline(*i* − 1)-aromatic(*i*) interaction: Folding of the Ala-*cis*Pro-Tyr peptide characterized by NMR and theoretical approaches

Frederico Nardi*, Johan Kemmink**, Michael Sattler & Rebecca C. Wade***
European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstraße 1, D-69012 Heidelberg, Germany

Received 20 October 1999; Accepted 20 March 2000

Key words: chemical shift, *cis*proline-aromatic interaction, conformational search, molecular dynamics, peptide folding in aqueous solution

Abstract

*Cis*proline(*i*−1)-aromatic(*i*) interactions have been detected in several short peptides in aqueous solution by analysis of anomalous chemical shifts measured by ^{1}H-NMR spectroscopy. This formation of local structure is of importance for protein folding and binding properties. To obtain an atomic-detail characterisation of the *cis*proline(*i* −1)-aromatic(*i*) interaction in terms of structure, energetics and dynamics, we studied the minimal peptide unit, blocked Ala-*cis*Pro-Tyr, using computational and experimental techniques. Structural database analyses and a systematic search revealed two groups of conformations displaying a *cis*proline(*i* −1)-aromatic(*i*) interaction. These conformations were taken as seeds for molecular dynamics simulations in explicit solvent at 278 K. During a total of 33.6 ns of simulation, all the 'folded' conformations and some 'unfolded' states were sampled. ^{1}H- and ^{13}C-chemical shifts and ^{3}J-coupling constants were measured for the Ala-Pro-Tyr peptide. Excellent agreement was found between all the measured and computed NMR properties, showing the good quality of the force field. We find that under the experimental and simulation conditions, the Ala-*cis*Pro-Tyr peptide is folded 90% of the time and displays two types of folded conformation which we denote '*a*' and '*b*'. The type *a* conformations are twice as populated as the type *b* conformations. The former have the tyrosine ring interacting with the alanine α proton and are enthalpically stabilised. The latter have the aromatic ring interacting with the proline side chain and are entropically stabilised. The combined and complementary use of computational and experimental techniques permitted derivation of a detailed scenario of the 'folding' of this peptide.

Abbreviations: Δδ, chemical shift deviation from random coil value; δ_{ring}, ring current effect from an aromatic system; BPTI, bovine pancreatic trypsin inhibitor; CSD, Cambridge Crystal Structure Database; DQF-COSY, double quantum filter correlation spectroscopy; PE-COSY, primitive exclusive correlation spectroscopy; NOESY, nuclear Overhauser and exchange spectroscopy; MD, molecular dynamics; ppm, parts per million; PDB, Brookhaven Protein Data Bank; SCAN3D, WHAT IF structural database; 1D, 2D and 3D, one-, two-, and three-dimensional.

Introduction

The study of local structure formation in peptides is of importance for the understanding of protein–ligand binding and protein folding mechanisms. Formation of native or non-native local structure will influence the rate of protein folding and, by limiting the conformational space accessible to the peptide, the configurational entropy loss of a peptide upon folding or binding to another molecule.

Usually, small peptide fragments are disordered in aqueous solution (Kemmink and Creighton, 1993) and do not have a single energetically favoured conformation: they are in the 'random coil' state (Howart, 1978; Merutka et al., 1995; Wishart et al., 1995; Fiebig

*Present address: AstraZeneca Structural Chemistry Laboratory, S-43183 Mölndal, Sweden.
**Present address: NMR Facility, University of Kent, Canterbury, U.K.
***To whom correspondence should be addressed. E-mail: wade@embl-heidelberg.de

et al., 1996; Smith et al., 1996). Local structure can be detected by deviations from this disordered state, which are indicated by anomalous chemical shifts in NMR spectroscopy (Dyson et al., 1988). One type of local structure interaction that has been detected by this means in short peptides in aqueous solution is the *cis*proline($i-1$)-aromatic(i) interaction. Dyson and colleagues identified this interaction in the peptide Ser-Tyr-Pro-Tyr-Asp-Val, which has a strong propensity to form a type VI turn in aqueous solution (Dyson et al., 1988; Yao et al., 1994; Demchuk et al., 1997; Mohanty et al., 1997). Kemmink and Creighton (1995) showed that in peptides derived from BPTI containing the sequence Xaa-Pro-Aro(i), the aromatic ring can interact strongly with the proline side chain and the aliphatic α proton of residue $i-2$ when the Xaa-Pro peptide bond is *cis*. For the peptide Gly-Ala-Pro-Tyr(i)-Thr-Gly-Ala (which is similar to the sequence of residues 7–13 of BPTI), the *cis*proline($i-1$)-aromatic(i) interaction is characterised by such large anomalous chemical shifts (0.4–0.5 ppm) that these shifts can only be explained by the ring current effect of the aromatic ring. Moreover, in this peptide, the *cis*proline($i-1$)-aromatic(i) interaction competes with an aromatic(i)-amide($i+2$) local interaction (Kemmink and Creighton, 1995; Nardi et al., 1997).

The *cis*proline($i-1$)-aromatic(i) interaction can be expected to be particularly important for folding mechanisms because the isomerisation of Xaa-Pro peptide bonds can be a slow, rate-limiting step in folding. This isomerisation is catalysed by peptidyl prolyl *cis–trans* isomerases, such as cyclophilin A (Fischer et al., 1989). Crystal structures show that cyclophilin A binds to the *cis* forms of Xaa-Pro dipeptides (Zhao and Ke, 1996a) and the succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide substrate (Zhao and Ke, 1996b).

A full characterisation of the *cis*proline($i-1$)-aromatic(i) interaction in short peptides in aqueous solution is difficult to obtain by experimental methods alone. NMR experiments provide chemical shifts and J-coupling constants but no detailed structural picture of this interaction. Crystallographic structures are not available for short linear peptides showing the *cis*proline($i-1$)-aromatic(i) interaction, because of the low occupation of the *cis*proline conformer compared to the *trans*proline conformer. For example, an Asn-Pro-Tyr sequence is present with a *cis*-peptide bond in the crystal structure of ribonuclease A, but the short peptide Ac-Asn-Pro-Tyr-NHMe adopts a *trans* form when crystallised (Montelione et al., 1984). Thus, the goal of the present study is to obtain an atomic-detail characterisation of the structure, dynamics and energetics of the *cis*proline($i-1$)-aromatic(i) interaction by computational analysis of the blocked peptide Ala-*cis*Pro-Tyr in aqueous solution that is consistent with experimental NMR data for this peptide. The Ala-*cis*Pro-Tyr peptide can be considered as a minimal peptide not only in the sense of being the shortest sequence to exhibit a *cis*proline($i-1$)-aromatic(i) interaction, but also in exhibiting features of protein folding in aqueous solution with formation of a fold by burial of hydrophobic atoms and formation of intramolecular hydrogen bonds.

A computational approach to explore the *cis*proline($i-1$)-aromatic(i) interaction has previously been taken by Scheraga and colleagues, who studied the Ac-Asn-Pro-Tyr-NHMe and Ac-Ala-Pro-Tyr-NHMe peptides by conformational free energy calculations in vacuo, and found that type VI β-bend conformations were adopted (Oka et al., 1984). As will be shown, our results provide a much more complete description of the conformational propensities of the Ala-*cis*Pro-Tyr peptide in aqueous solution, which turn out to be more complex than found in this early study. More recently, Demchuk et al. studied peptides containing Aro-Pro-Aro sequences by molecular dynamics simulation (Demchuk et al., 1997a, b) and Ripoll et al. used Monte Carlo simulations to investigate the pH-conformational dependence of the SYPYD peptide (Ripoll et al., 1999). The nature of the *cis*proline($i-1$)-aromatic(i) interaction is sequence dependent, and thus the latter studies are complementary to the work described in the present manuscript.

The procedure we adopted to study the blocked Ala-*cis*Pro-Tyr peptide is as follows. First, we constructed realistic conformations of this peptide by homology to structures in crystallographic databases and by systematic ab initio modelling. These conformations were required to show a *cis*proline($i-1$)-aromatic(i) interaction which was defined by the presence of a computed chemical shift on the H$\alpha(i-2)$ or the H$\gamma 3(i-1)$ protons due to the ring current effect of the aromatic ring. Then, we used molecular dynamics (MD) simulations in explicit solvent to estimate the relative stability of the conformations found and to explore the dynamics of the transitions between them. Finally, we measured the NMR properties of the Ala-Pro-Tyr peptide in aqueous solution and directly compared them with the MD simulations. It should be noted that the MD simulations were performed independently of the NMR experiments and without NMR restraints. Two conformational classes, which

we denote '*a*' and '*b*', were identified in the structure database and systematic searches. Conformations from both groups were stable over long periods in the MD simulations and, by considering populations of both conformational classes, very good agreement was obtained between the NMR experiments and the MD simulations. Further, examination of conformational transitions provides a picture of the folding landscape of this peptide in aqueous solution. Finally, determinants of sequence dependence of the *cis*proline($i-1$)-aromatic(i) interaction were investigated in short comparative simulations of the blocked Gly-*cis*Pro-Tyr peptide.

Materials and methods

The peptide model

The Ala-Pro-Tyr peptide was modelled with the N- and C-termini blocked with acetyl and methylamide groups, respectively.

IUPAC nomenclature (IUPAC Commission, 1970; Markley et al., 1998) is used for atom names and side chain conformations, and main chain conformations are named following Wilmot and Thornton (1990).

The CHARMm22 all-atom force field was employed (Brooks et al., 1983; Molecular Simulations Inc., 1992; Momany and Rone, 1992; MacKerell et al., 1998). The torsional force constant of the proline ψ dihedral angle in the CHARMm22 force field was reduced from 10.46 to 1.046 kJ mol^{-1} to correspond to the value of this parameter for all other residues.

Computation of ^{1}H-NMR properties

The proton chemical shifts ($\Delta\delta$) of a given peptide structure were calculated using the empirically parameterised TOTAL program (Williamson and Asakura, 1993). Due to the lack of random coil shift values for residues preceded or followed by a *cis* peptide bond, we only consider the ring current contribution (δ_{ring}). Indeed, the lack of random coil shift values for the residues adjacent to a *cis* peptide bond and inadequate incorporation of the effect of the *cis-trans* isomerisation into calculation of the chemical shifts led to slightly poorer agreement between computed and experimental values for total compared to ring current chemical shifts. The TOTAL program does not provide the self-δ_{ring} value of the β protons of aromatic residues because these protons are not usually stereospecifically assigned. We introduced this self-contribution by modifying the TOTAL program and re-normalising the aromatic β protons with respect to their random coil chemical shifts. This improved the agreement between experimental and computed δ_{ring} and $\Delta\delta$ values. Thus, it was possible to reproduce the splitting of the β protons of the tyrosine observed in the ^{1}H-NMR spectra for the *cis* form of the Ala-Pro-Tyr peptide. This splitting was mainly due to the tilt of the aromatic ring by ca. $+10°$ from its standard orientation.

Data base analysis

In order to establish a set of realistic aqueous solution conformations for the Ala-Pro-Tyr peptide, Cartesian coordinates from three databases were used:

(1) The Brookhaven Protein Data Bank (PDB) (Bernstein et al., 1977).

(2) The SCAN3D database (Vriend et al., 1994) which contained a representative set of 315 non-homologous high-resolution protein crystal structures selected from the PDB (see Supplementary material, to be obtained from the authors on request).

(3) The Cambridge Crystal Structure Database (CSD) (Allen and Kennard, 1993).

The atomic coordinates of fragments of three amino acid residues that match the sequence motif Xaa-Pro-Aro(i) were extracted from these databases. In the motif, Xaa stands for any amino acid, Aro stands for Tyr or Phe, and i is the index number of the last residue. The sequences of the extracted peptide fragments were mutated to Gly-Pro-Phe, and the dihedral angles, ϕ, ψ, ω, χ_{1-3}, and all proton chemical shifts were computed. Finally, the structures were clustered according to their backbone conformations and their computed chemical shifts (see Supplementary material, http://www-z.embl-heidelberg.de:8080/ExternalInfo/wade/pub/data/JBNMR00_admat).

Systematic search

The conformation of the Ala-Pro-Tyr peptide with standard bond lengths and bond angles is defined by 19 backbone and side-chain dihedral angles. Therefore, to limit the conformational space examined, we only considered conformations in the known sterically accessible regions, such as the α_R, α_L and β regions of the Ramachandran plot (Ramachandran and

Table 1. Conformations sampled in the systematic search

Dihedral angles (°)	Ala	*cis*Pro	Tyr
(ϕ,ψ)	(−60, −40)	(−80, −150)	(−60, −40)
	(−65, 145)	(−90, 0)	(−65, 145)
	(−150, 150)	(−100, 60)	(−150, 150)
	(50, 40)		(50, 40)
ω	180	0	180
χ_1	–	0	−60, 60, 180
χ_2	–	0	90
χ_3	–	0	–

Sasisekharan, 1968), which are listed in Table 1. This makes a total of 144 well-separated peptide conformations that were then energy minimised 10 times with 250 steepest descent steps followed by 250 conjugate gradient steps with an initial step size of 0.5 Å. The relative dielectric constant ϵ was set to infinity and no cut-off was used. Values of 20 and 80 for ϵ were also tested, but did not change the results. Finally, the 144 structures were clustered according to their backbone conformation, their energy, and their computed chemical shifts.

Molecular dynamics protocol

MD trajectories were computed using the ARGOS program, version 6.0 (Straatsma and McCammon, 1990). The CHARMm22 force field together with the TIP3P (Jorgensen et al., 1983) water model was used. Simulations were carried out with periodic boundary conditions at constant pressure and temperature using the Berendsen coupling scheme (Berendsen et al., 1984). The external temperature bath was set to 278 K with a coupling time constant of 0.4 ps; the external pressure bath was set to 1 atm with a compressibility coefficient of 4.53 10^{-10} m^2 N^{-1}. A non-bonded group-based cutoff of 10 Å was used, and the pair list was updated every 10 steps. A timestep of 2 fs was employed and bond lengths were constrained using the SHAKE algorithm (Ryckaert et al., 1977).

All the conformations detected in the conformational searches were taken as starting points for MD simulations. For each simulation, the peptide was energy minimised with 20 steps of steepest descent to remove van der Waals clashes introduced by mutation of the peptide sequence. The peptide was then solvated in a 25 Å side cube of pre-equilibrated TIP3P water molecules at 300 K and 1 atm. The water molecules for which the oxygen atom was closer than 2.6 Å to any non-hydrogen peptide atom were removed. The system was equilibrated during 100 ps of MD simulation at constant pressure and temperature. Coordinates were recorded every 1 ps.

Experiments

The peptide acetyl-Ala-Pro-Tyr-amide was synthesised using 9-fluorenyl-methoxycarbonyl chemistry to 99% purity.

All NMR measurements were performed at 273 K on a 50 mM peptide sample in 90% ^1H_2O/10% ^2H_2O (v/v) at pH 4.6. 2,2-Dimethyl-2-silapentane-5-sulfonic acid sodium salt (DSS) was added as an internal reference.

The ^{1}H-NMR spectra were assigned using standard methods (Wüthrich, 1986; Cavanagh et al., 1996). The scalar coupled spin systems were first identified in TOCSY and DQF-COSY spectra; the sequential connectivity was then determined using NOESY and ROESY experiments. A natural abundance ^{1}H-^{13}C HSQC (Bodenhausen and Ruben, 1980) spectrum was recorded to resolve resonance overlap of the β-γ proline protons. The coupling constants were extracted from 1D and PE-COSY (Müller 1987) experiments. The stereospecific assignments of the methyl protons in the proline and tyrosine residues were obtained from the ROESY spectra and the ^{3}J-coupling constants. In all experiments except the PE-COSY, water suppression was accomplished by WATERGATE (Piotto et al., 1992). In PE-COSY spectra, water was suppressed by presaturation.

^{1}H chemical shifts were compared to the random coil values given by Wishart et al. (1995), which take into account the influence of the conformational restriction imposed by the proline residue (MacArthur and Thornton, 1991). Experimentally based δ_{ring} values for the Ala-*cis*Pro-Tyr(*i*) were derived by comparison of $\Delta\delta$ values with those measured by Kemmink and Creighton (1995) for the Gly-Ala-*cis*Pro-Ala(*i*)-Thr-Gly-Ala peptide and by Wu and Raleigh (1998) for the Gly-Ala-*cis*Pro-Gly(*i*). The two estimates of δ_{ring} for Hγ3($i-1$) agreed within $\pm$0.02 ppm.

Results and discussion

Conformational searches

A summary of the potential plausible conformations of the blocked Ala-Pro-Tyr peptide in aqueous solution that were identified in the database and systematic searches is presented in Table 2.

Proline cis-trans *isomerisation*

All the database analyses confirm the experimentally observed propensity of the aromatic residue to stabilise the *cis*proline isomer in the Xaa-Pro-Aro peptide sequence fragment. The population of the *cis*Pro-Aro motif found in the PDB and in the SCAN3D databases is higher than the average *cis* proline content of 5–7% present in all protein structures (Stewart et al., 1990; MacArthur and Thornton, 1991). Similarly, in the CSD, the population of the *cis*Pro-Aro motif is 58%, whereas the average *cis* proline content is 38% (Miler-White et al., 1992). Experimentally, in aqueous solution about 20% of the Xaa-Pro-Aro peptides occupy the *cis*proline isomer ((Kemmink and Creighton, 1995), and this work). In the absence of the aromatic residue, the population of the *cis*proline isomer in similar peptides under the same conditions ranges from 0 to 8% (Grathwohl and Wüthrich, 1981; Kemmink and Creighton, 1995; Wu and Raleigh, 1998).

Proton ring shifts

In the PDB, the average values of δ_{ring} computed for all the structures matching the Xaa-*cis*Pro-Aro motif are almost null for all protons except H$\alpha(i-2)$, H$\gamma3(i-1)$ and H$\delta3(i-1)$, whose average δ_{ring} value is about -0.17 ppm. Even though the computed absolute average δ_{ring} value of these protons (in the PDB and in the other database searches) is much smaller in magnitude than the one measured in Ala-*cis*Pro-Tyr in aqueous solution (see below), they show the same trends. This confirms the presence of a *cis*proline$(i-1)$-aromatic(i) interaction in the Xaa-*cis*Pro-Aro motif.

All the conformations that do not display a large δ_{ring} value, either on the H$\alpha(i-2)$ or on the H$\gamma3(i-1)$ protons, have δ_{ring} values centred on 0 ppm with a ±0.2 ppm Gaussian-like distribution. Consequently, a conformation with a δ_{ring} value more negative than -0.25 ppm on either the H$\alpha(i-2)$ or the H$\gamma3(i-1)$ proton is considered to display a *cis*proline$(i-1)$-aromatic(i) interaction. This criterion enabled all the structures in the different databases that display a *cis*proline$(i-1)$-aromatic(i) interaction to be found. It is important to emphasise that, in agreement with experiment, none of the conformations detected in any of the databases displayed significant computed δ_{ring} values simultaneously on both the H$\alpha(i-2)$ and the H$\gamma3(i-1)$ protons. Thus, the experimentally determined chemical shifts can only be explained by population of more than one conformation.

Peptide conformations

All the conformations identified are displayed in Figure 1. The peptides that match the motif Xaa-*cis*Pro-Aro in the PDB sample a huge configurational space (Figure 1, bottom left), which is much more restricted when the peptide adopts a *cis*proline$(i-1)$-aromatic(i) interaction (Figure 1, middle and top left). To differentiate between the two types of *cis*proline$(i-1)$-aromatic(i) interaction, we designated the conformations in which $\delta_{ring} < -0.25$ ppm on the H$\alpha(i-2)$ proton as type *a*, and the conformations in which $\delta_{ring} < -0.25$ ppm on the H$\gamma3(i-1)$ proton as type *b*.

A g^+ tyrosine χ_1 rotamer is a common feature of all conformations in the databases with a *cis*proline$(i-1)$-aromatic(i) interaction. This tyrosine rotamer is also the most populated in all protein structures (Ponder and Richards, 1987).

The conformations with a type *a* *cis*proline$(i-1)$-aromatic(i) interaction detected in the database searches can be clustered into three groups: *a*1, *a*2 and *a*3 (see Figure 2 and Table 3). The most populated group, *a*1, forms a $\beta\rightarrow\beta^{cis\mathrm{Pro}}$ type VI turn (Wilmot and Thornton, 1990). In the systematic search, all these type *a* conformations and a further conformer type, *a*4 (see Table 3), were identified at an energy less than 10 kJ mol^{-1} above the lowest energy conformation.

The *cis*proline$(i-1)$-aromatic(i) type *b* conformations identified in the databases and systematic searches cluster into two groups: *b*1 and *b*2, that belong, respectively, to the $\beta^{P}\rightarrow\alpha_R{}^{cis\mathrm{Pro}}$ and $\beta^{E}\rightarrow\alpha_R{}^{cis\mathrm{Pro}}$ type VI turns (Wilmot and Thornton, 1990) which are often called, respectively, type VIa and VIb turns (Lewis et al., 1973).

All the conformers found in the PDB were also found in the SCAN3D database. In contrast, the CSD contains only the *a*1 and *b*1–2 conformations. This is probably due to the conformational restrictions in the large number of cyclic peptides present in the CSD. All the type *a*1 conformers observed in the CSD were in metal-bound antamanide (a cyclic decapeptide). One antamanide structure, which was not metal-bound, displayed a type *b*1 conformer.

Table 2. Summary of the results of the conformational searches

	PDB		SCAN3D		CSD		Systematic search
Number of structures in database							
	4733		315		175 093		
Number of occurrences of							
Xaa-Pro-Aro	4156		230		31		288[b]
Xaa-*cis*Pro-Aro	426	10%[a]	29	13%	18	58%	144[b]
Average δ_{ring} *(ppm) of*[c]							
H$\alpha(i-2)$	−0.17		−0.14		–		−0.26
H$\gamma3(i-1)$	−0.17		−0.30		–		−0.24
Number of occurrences of Xaa-cisPro-Aro for which							
δ_{ring} H$\alpha(i-2) < -0.25$ ppm	83		5		3		9[d]
δ_{ring} H$\gamma3(i-1) < -0.25$ ppm	92		10		15		6[d]

[a]% are of the total population of Xaa-Pro-Aro possible.
[b]The number of conformations generated by the systematic search.
[c]The ring shift is computed with the TOTAL program. The standard deviation of all the δ_{ring} distributions is about 0.35 ppm.
[d]These are from 21 out of the 144 generated conformations, for which the total potential energy is less than 10 kJ mol^{-1} above the lowest energy conformation.

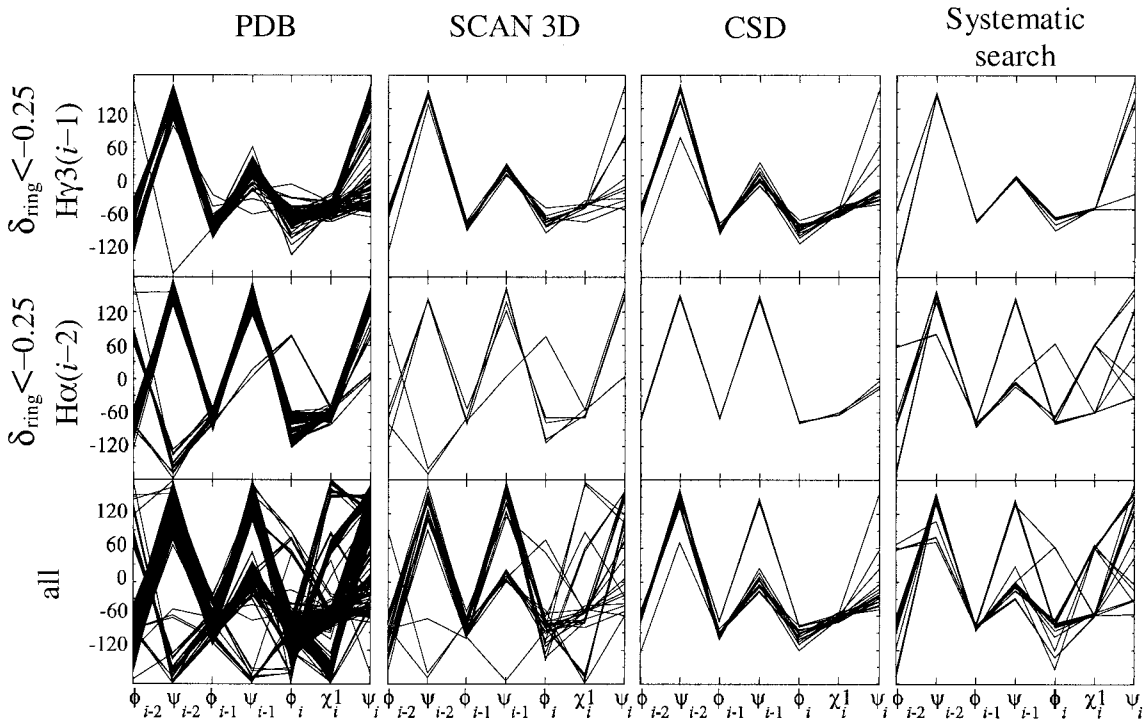

Figure 1. Parallel-coordinate (McClain and Erickson, 1995; Becker, 1997) plot of dihedral angle values (°) versus the dihedral angle sequence of Xaa-Pro-Aro(i) configurations found in the PDB, SCAN 3D and CSD databases and the systematic searches. The top plots show all the configurations with $\delta_{ring} < -0.25$ ppm on the H$\gamma3(i-1)$ protons. The middle plots display all the configurations with $\delta_{ring} < -0.25$ ppm on the H$\alpha1(i-2)$ protons. The bottom plots display all the configurations matching the sequence motif Xaa-Pro-Aro(i). For the systematic search (right-hand plots), only the conformations with energy lower than the lowest energy plus 10 kJ mol^{-1} are shown.

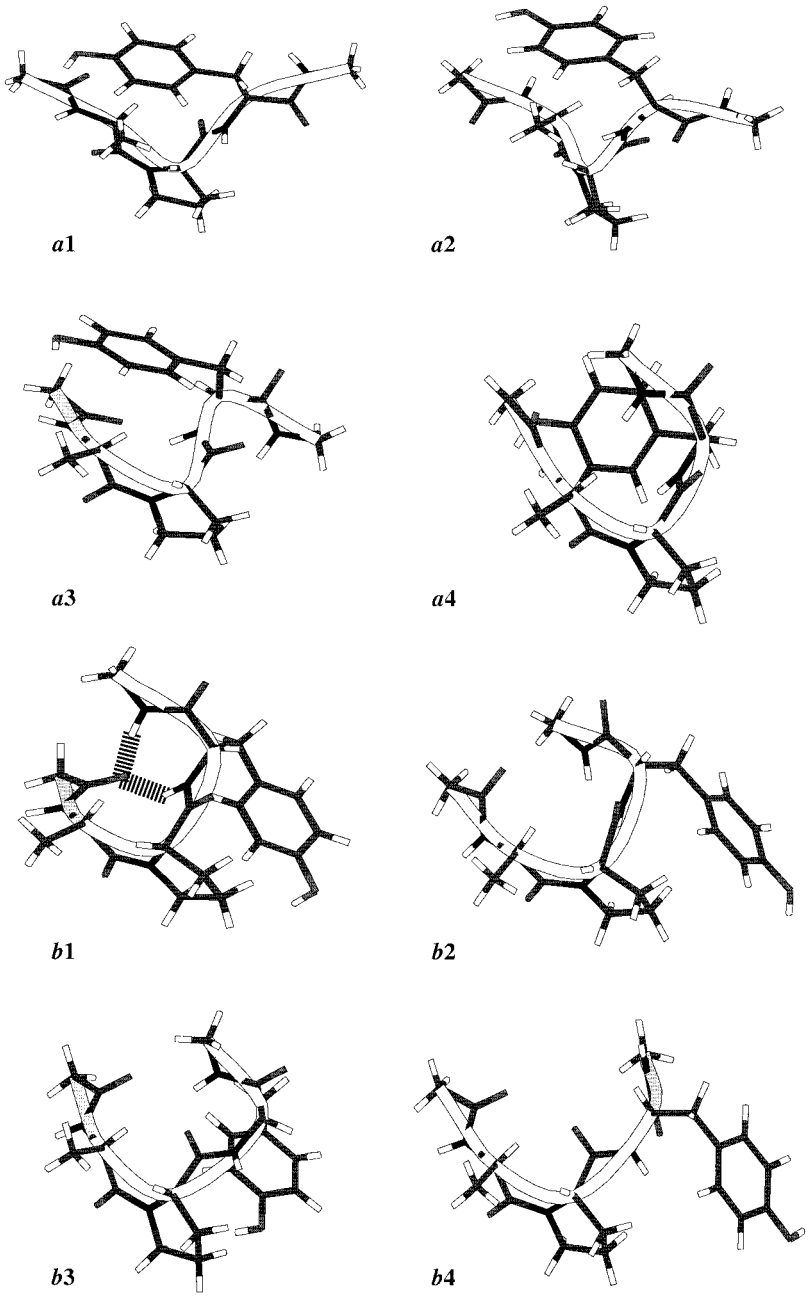

Figure 2. The main conformations of Ala-*cis*Pro-Tyr found in the conformational searches, and in the MD simulations. The *a*1 conformation is dominant; the *b*1 and *b*2 conformations are populated much less but significantly, the other conformations are rare, see Table 2. In the *a*1–4 conformations, the tyrosine ring packs over the main chain of alanine, and displays a $\delta_{ring} < -0.25$ ppm on the H$\alpha(i-2)$ proton. In the *b*1–4 conformations, the tyrosine ring packs over the pyrrolidine ring of the proline, and displays a $\delta_{ring} < -0.25$ ppm on the H$\gamma 3(i-1)$ proton. The *a*1–3 and *b*1 conformations correspond to the residue fragments 54–56 of 1fus, 26–28 of 1gmp, 8–10 of 1tgx and 70–72 of 1lct, respectively, that have been mutated to Ala-Pro-Tyr. The *a*4 conformation corresponds to a possible conformation detected by the systematic search. The *b*2–4 conformations correspond to additional conformations found during the MD simulations. Hydrogen bonding is shown by ||||||||||||.

Table 3. Summary of the Ala-*cis*Pro-Tyr conformations

Conf	Ala ϕ ψ ($i-2$)	*cis*Pro ϕ ψ ($i-1$)	Tyr ϕ ψ (i)	χ_1	PDB	SCAN3D	CSD	Systematic search	MD	Population (%) MD	Population (%) NMR
*a*1	β	β	α_R or β	g^+	✓	✓	✓	✓	✓	69	63
*a*2	ϵ	β	α_R or β	g^+	✓	✓		✓	✓	1	
*a*3	β	α_R	α_L	g^+	✓	✓		✓	✓	1	
*a*4	β	α_R	α_R or β	g^-				✓	✓	1	
*b*1	β^P	α_R	α_R or β	g^+	✓	✓	✓	✓	✓	9	27
*b*2	β^E	α_R	α_R or β	g^+	✓	✓	✓	✓	✓	7	
*b*3	β	β	α_L	g^+					✓	4	
*b*4	β	β	α_R or β	g^-					✓	4	
u		‘unfolded’		*t*	✓	✓	✓	✓	✓	4	10

Stability and sequence specificity of the conformer cisproline($i-1$)-aromatic(i) interaction

Which *cis*proline($i-1$)-aromatic(i) interaction is more stable, type *a* or type *b*? On one hand, the type *b* *cis*proline($i-1$)-aromatic(i) conformations were found more often in all the database searches than the type *a* conformations. On the other hand, MacArthur and Thornton (1991) found that, for the Xaa-*cis*Pro sequence motif, the β→β conformation (39 hits in their set of non-homologous proteins), which mostly results in a type *a* interaction, occurred more often in proteins than the β→α_R conformation (13 hits), which mostly results in a type *b* interaction (see Table 3). Consequently, a simple structural database analysis does not result in a clear prediction of the structural propensity of the Ala-*cis*Pro-Tyr fragment in aqueous solution. Moreover, the statistics in the databases were not sufficient to detect any sequence specificity for type *a* or type *b* interactions.

The NMR studies of Yao et al. (1994) on the Ser-Xaa-Pro-Tyr(i)-Asp-Val peptide series, however, indicate that a Gly($i-2$) stabilises the type *a* interaction, an Aro($i-2$) stabilises the type *b* interaction, and an Ala($i-2$) allows both conformational types to be significantly populated. This indicates that both type *a* and type *b* conformations may be present in the Ala-*cis*Pro-Tyr peptide and this makes it particularly challenging to reproduce the competition between these interactions computationally.

In summary, an ensemble of possible conformations that are likely to be present in Ala-*cis*Pro-Tyr in aqueous solution and which are compatible with the *cis*proline($i-1$)-aromatic(i) interaction were identified from the conformational searches. No sequence specificity could be extracted from the database searches for stabilisation of the type *a* versus the type *b* conformers. The stability of each conformation of Ala-*cis*Pro-Tyr, which can display both type *a* and type *b* conformers, was therefore examined with MD simulations.

Molecular dynamics simulations

Stability of conformer types

The *a*1–4 and *b*1 conformations were taken as starting points for MD simulations. Only the *a*1 and *b*1 conformations remained stable after 100 ps of equilibration. During the equilibration phase, the *a*2 conformation transformed into the *a*1 conformation after 20 ps, and the *a*3–4 conformations unfolded to conformations without any *cis*proline($i-1$)-aromatic(i) interaction after approximately 50 ps. Consequently, two trajectories, A and B, started after the equilibration of the *a*1 and *b*1 conformations, respectively, were recorded. After 1.5 ns of the MD simulation A, the *a*1 conformation transformed, over 200 ps, into the *b*2 conformation crossing an energy barrier of about 20 kJ mol^{-1} caused by the Lennard-Jones interaction between the aromatic ring and the alanine side chain. Observation of such a rare event after so short a simulation period was a chance event. Thus, to observe several such transitions, the trajectories A and B were each extended to 16.8 ns. Clearly, long trajectories could also have been started from extended or random coil conformations without the need for con-

formational searches beforehand, as all subconformers observed in conformational searches were sampled in the simulations. However, the protocol we adopted permitted us to examine the stability of the individual conformers in short simulations without biasing sampling in the longer simulations.

Figure 3 displays the time evolution of the principal parameters along the trajectories A and B, which have been concatenated. The δ_{ring} plots show that the H$\alpha(i-2)$ and H$\gamma 3(i-1)$ protons are alternately up-field shifted. Thus, the simulation confirms the finding in the conformational searches that the H$\alpha(i-2)$ and H$\gamma 3(i-1)$ protons cannot be up-field shifted simultaneously. During the total of 33.6 ns of simulation, the peptide underwent several transitions, but most of the time either the H$\alpha(i-2)$ or H$\gamma 3(i-1)$ proton was up-field shifted. Indeed, considering only the δ_{ring} criterion ($\delta_{ring} < -0.25$ ppm), the type *a* and *b* *cis*proline($i-1$)-aromatic(i) interactions and the unfolded state were present 71%, 25% and 4% of the time, respectively.

Proton ring shifts

In the folded states, the instantaneous value of δ_{ring} on H$\alpha(i-2)$ or H$\gamma 3(i-1)$ fluctuates from 0 to -1.5 ppm, whereas the 100 ps time average values remain constant at around -0.9 ppm on either H$\alpha(i-2)$ or H$\gamma 3(i-1)$. The average values of δ_{ring} over the 33.6 ns MD simulation are -0.55 ppm on H$\alpha(i-2)$ and -0.14 ppm on H$\gamma 3(i-1)$. These values are in very good agreement with the chemical shifts measured by Kemmink and Creighton (1995) on the Gly-Ala-Pro-Tyr(i)-Thr-Gly-Ala peptide, for which we estimated, by comparison of shifts measured in the corresponding peptide in which the aromatic residue has been mutated to alanine, δ_{ring} as -0.65 ppm for H$\alpha(i-2)$ and -0.20 ppm for H$\gamma 3(i-1)$ when the Ala-Pro peptide bond is *cis*. These δ_{ring} values arise mainly because of contributions from the *a*1 and *b*1–2 conformations (see Figure 3).

Peptide conformations

During the 33.6 ns of MD simulation, the peptide sampled all the *a*1–4 and *b*1–2 conformations found in the database analyses and in the systematic search. In addition, two new conformations, *b*3 and *b*4, were sampled during the MD simulations (see Table 3). These two new conformations were not detected in the systematic search because they are of high energy in vacuum. They are stabilised in the MD simulations by the presence of the explicit solvent. Consequently, the MD simulation is consistent with and complementary to the database and systematic searches. Note that all the other MD simulations starting from the *a*2–4 and *b*2 conformations were not extended because they are implicitly included in the A and B MD simulations.

The *a*1 conformation is the most stable conformation overall. In this conformation, the tyrosine ring packs snugly above the alanine backbone aliphatic proton, and the main-chain is partially extended so that all the peptide polar groups are solvated (see Figure 2). This conformer is present 69% of the time and is on average stable over 5 ns time periods. All the other type *a* *cis*proline($i-1$)-aromatic(i) interactions represent only 3% of the total population. The next most populated conformations after the *a*1 conformation are the *b*1–2 conformations in which the aromatic ring packs snugly against the pyrrolidine ring. In the conformation *b*1, which is present 9% of the time and on average stable over a 1 ns time period, the backbone forms a type VIa turn (see Figure 2), and the N-terminal carbonyl oxygen, O($i-3$), forms a transient bifurcated hydrogen bond with the two C-terminal amide hydrogens, H(i) and H($i+1$). This bifurcated hydrogen bond was also observed by Demchuk et al. (1997) in a simulation of a type VI turn peptide. The *b*2 conformer, which is present 7% of the time, differs from the *b*1 conformer in that it lacks intra-residue hydrogen bonds and forms a type VIb turn. The absence of intra-residue hydrogen bonds results in a jump in the intra-peptide Coulombic energy, relative to the *b*1 conformation, of about $+10$ kJ mol^{-1} (see Figure 3). The reduced compactness and the absence of cross-turn hydrogen bonds make the *b*2 conformer slightly less stable than the *b*1 conformer. The *b*3 and *b*4 conformations are each present for only 4% of the simulation time. They have the aromatic ring packed above the $\beta 3$ proton of the pyrrolidine ring instead of the $\gamma 3$ proton as for the *b*1 and *b*2 conformations, and this results in a lower contribution to δ_{ring} on the H$\gamma 3(i-1)$ (see Figure 3).

The proline residue adopted a DOWN pucker 85% of the time. This pucker percentage was also observed for *cis*proline in protein crystal structures (Miler-White et al., 1992). When the aromatic side chain interacts strongly with the pyrrolidine ring in the *b*1–3 conformations, occupation of the DOWN pucker increases to 95%. During the MD simulations, the UP pucker conformation was, however, stable on a 1 ns time scale.

Analysis of the tyrosine χ_1 torsion angle gives populations of the g^+, t and g^- rotamers of 91%,

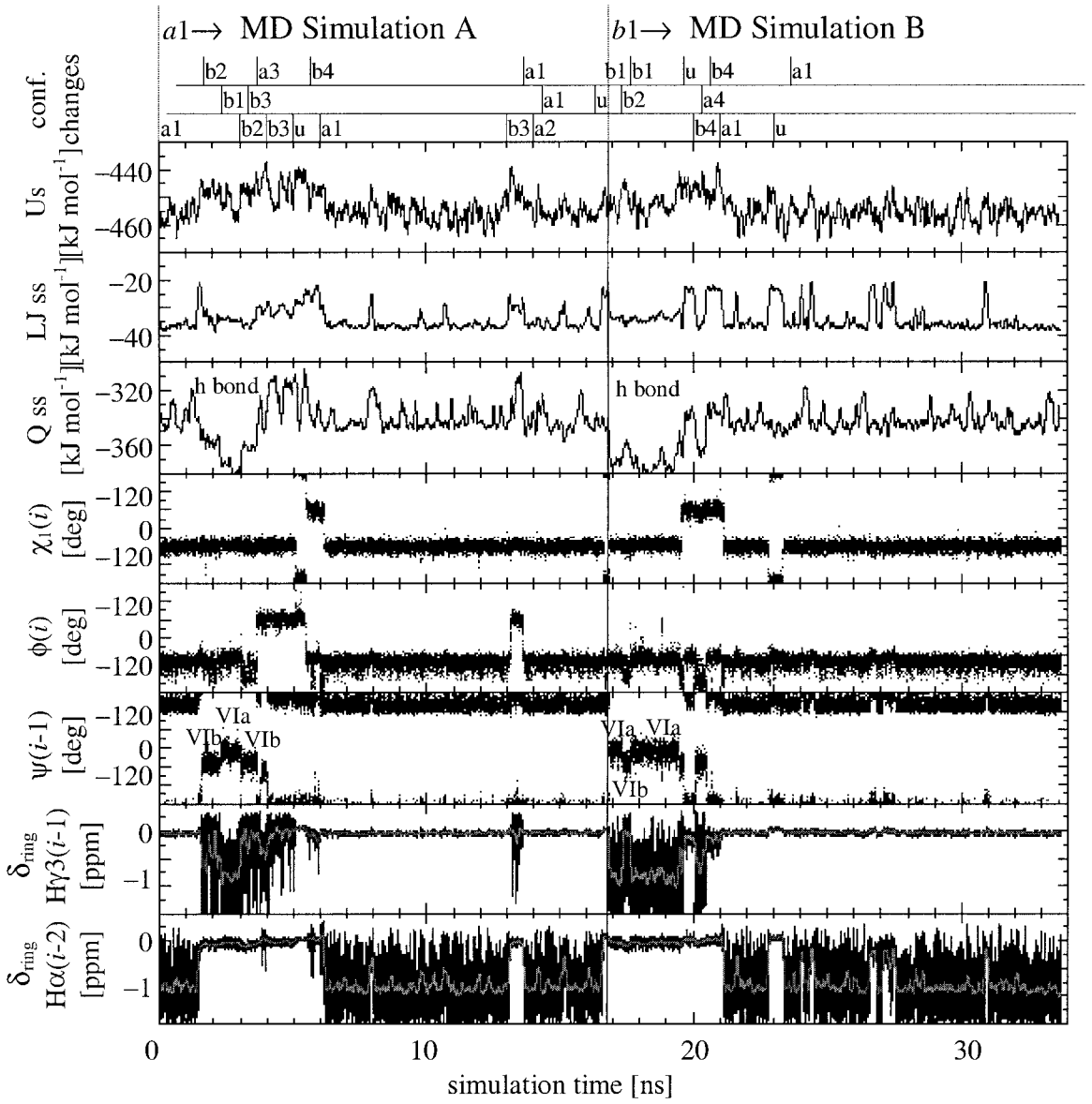

Figure 3. From bottom to top, time evolution plots along the MD simulations A (from 0 ns to 16.8 ns) and B (from 16.9 ns to 33.7 ns) of the instantaneous δ_{ring} values (black) and the 100 ps time average δ_{ring} values (grey) of the α alanine and $\gamma 3$ proline protons; the proline ψ and the tyrosine ϕ and χ_1 torsional angles; and the 100 ps time average intra-solute-Coulombic (Q_{ss}), intra-solute-Lennard-Jones (LJ_{ss}), and total (U_s) peptide potential energies. The conformational changes of the peptide are indicated by '|xx', where '|' marks the time at which the conformation changes and 'xx' indicates the conformation which is adopted.

4% and 5%, respectively. The tyrosine *t* rotamer is the least populated because of the unfavourable side chain solvation energy of a C-terminal tyrosine, and is characteristic of the unfolded state. The folded state was mainly found with a g^+ rotamer. Thus, the *cis*proline($i-1$)-aromatic(i) interaction and the position of the C-terminus dramatically perturb the average distribution of tyrosine χ_1 g^+, *t* and g^- rotamers of 52.3%, 32.7% and 15%, respectively, observed in protein crystal structures (Ponder and Richards, 1987).

*Sequence specificity: Simulations of the Gly-*cis*Pro-Tyr(*i*) peptide*

We performed MD simulations of the Gly-*cis*Pro-Tyr(*i*) peptide to examine the sequence specificity of the *cis*proline($i-1$)-aromatic(i) interaction. In two

3 ns long independent MD simulations of the Gly-*cis*Pro-Tyr(*i*) peptide started from the conformation *b*1, the peptide converted spontaneously to the *a*1 conformation after approximately 1 ns and remained in this conformation for the rest of the MD simulations. In another 3 ns MD simulation started from the *a*1 conformation, the peptide remained in this conformation for the whole simulation. Consequently, the presence of a small residue, such as glycine at the $(i-2)$ position, enhances the stability of the type *a cis*proline$(i-1)$-aromatic(i) interaction, whereas a bigger residue such as tyrosine enhances the type *b* interactions. This is consistent with the experimental observation of a large ring shift ($\delta_{ring} = -1.52$ ppm) on the H$\alpha(i-2)$ proton in the Ala-Gly-*cis*Pro-Tyr(i)-Thr-Gly-Ala peptide (Kemmink and Creighton, 1995). The introduction of the glycine residue allows more space for the tyrosine ring to interact with the aliphatic hydrogens of the $i-2$ residue. Thus, the glycine residue at position $i-2$ entropically favours the type *a cis*proline$(i-1)$-aromatic(i) interaction, whereas a bulky residue such as tyrosine at position $i-2$ reduces the conformational space available for the type *a cis*proline$(i-1)$-aromatic(i) interaction, thus favouring the type *b* interaction. The *b*1–2 conformations were shown to be stable for long times in simulations of the Ala-Tyr-*cis*Pro-Tyr(*i*)-Asp sequence (Demchuk et al., 1997).

Summary of MD simulations

The MD simulations of the Ala-*cis*Pro-Tyr peptide in aqueous solution revealed an ensemble of conformations that were structurally and thermodynamically in good agreement with the experimental observations made by Kemmink and Creighton (1995) on a heptapeptide incorporating this sequence. To validate our computational model, we went on to perform NMR measurements on the Ala-*cis*Pro-Tyr peptide itself and compare them to our computational data.

NMR experiments

1D and 2D spectra were recorded for acetyl-Ala-Pro-Tyr-amide. This peptide, which differs from our model peptide in lacking the methyl group at the C-terminus, was studied due to ease of synthesis and purification. The ^{1}H and ^{13}C chemical shifts are given in Table S1 and ^{1}H-^{1}H ^{3}J-coupling constants and inter-residue ROESY peaks are given in Table S2 of the Supplementary material. The conformational information derived and the NMR measurement from which this was obtained is given in Table 4.

Proline cis-trans *isomerisation*

The peptide adopts *cis* and *trans* forms in aqueous solution. The two forms were clearly identified by the presence of the Hα(Ala)-Hδ(*trans*Pro) and the Hα(Ala)-Hα(*cis*Pro) peaks in the ROESY spectra. The *cis* to *trans* population ratio is 20%:80% and was estimated from the ratio of the integrated peaks of the alanine methyl protons.

Peptide conformations

δ_{ring} was estimated as -0.84 ± 0.01 ppm and -0.32 ± 0.02 ppm for H$\alpha(i-2)$ and H$\gamma3(i-1)$, respectively, where the errors reflect the difference between the two references taken for the δ_{ring} evaluation (see Methods section). As observed in our computations, no single peptide conformation can be adopted that is compatible with such ring current shifts on both protons. Consequently, the measured chemical shift deviations are the result of an average of the chemical shifts of several conformations. Furthermore, the absence of a cross peak between the tyrosine ring protons and either the H$\alpha(i-2)$ or the H$\gamma3(i-1)$ proton in the ROESY spectra is consistent with exchange between several conformations.

An empirical correlation has been observed between the protein backbone conformation and ^{13}Cα and ^{13}Cβ secondary chemical shifts (Spera and Bax, 1991; Wishart and Sykes, 1994). Analysis of the ^{13}Cα chemical shifts shows that in Ala-*cis*Pro-Tyr, Ala and Pro are mainly in extended conformations and Tyr is in equilibrium between helical α_R and extended conformations. The value of the ψ dihedral angle of a proline residue is strongly correlated with the difference between the β and γ ^{13}C chemical shifts ($\Delta\delta_{\beta\gamma}$) (Siemion et al., 1975; Siemion, 1976; Giessner-Prettre et al., 1987). From this, it can be deduced that in the *cis* form $\psi(i-1)$ could be 148° or −28° with an error of ±10°.

MacArthur and Thornton (1991) analysed a non-homologous set of protein structures and showed that *cis*proline has a stronger preference to be in the β region, centred on $B_{cis} = (-76°,159°)$, than to be in the α_R region, centred on $A_{cis} = (-86°,1°)$, (α_R:β = 24%:76%). Consequently, with the following two-state model,

$$\Delta\delta_{\beta\gamma} = [\alpha_R]\Delta\delta_{\beta\gamma}(A) + [\beta]\Delta\delta_{\beta\gamma}(B), \quad (1)$$

where the measured $\Delta\delta_{\beta\gamma}$ is the weighted average of the $\Delta\delta_{\beta\gamma}$ values of the conformations at the centres *A*

Table 4. Summary of the peptide Ala-Pro-Tyr(*i*) conformations at 273 K derived from NMR

Residue	Dihedral angle	Sequence	*Trans*	*Cis*	Technique
Ala	ϕ	$(i-2)$	$-165°$ or $-70°$	$-170°$ or $-70°$	$^3J_{HN-H\beta}$
	ψ		More β	More β	$\Delta\delta^{13}C\alpha$
Pro	ϕ	$(i-1)$	Coil	More β	$\Delta\delta^{13}C\alpha$
	ψ		$-35°$ or/and $150°$	α_R 1°(26%)	
				β 159°(74%)	$\Delta\delta_{\beta\gamma}$
	χ_1		UP and DOWN	20° DOWN	$^3J_{H\alpha-H\beta}$
	χ_2		UP and DOWN	$-35°$ DOWN	$^3J_{H\beta-H\gamma}$
	χ_3		UP and DOWN	25° DOWN	$^3J_{H\gamma-H\delta}$
Tyr	ϕ	(i)	$-160°$ or $-80°$	$-155°$ or $-85°$	$^3J_{H\alpha-H\beta}$
	ψ		Coil	Coil	$\Delta\delta^{13}C\alpha$
			g^+ (48%)	g^+ (84%)	$^3J_{H\alpha-H\beta}$
	χ_1		*t* (7%)	*t* (13%)	ROESY
			g^- (45%)	g^- (3%)	

and *B* of the α_R and β regions, respectively, defined by MacArthur and Thornton (1991), one can estimate the α_R:β population ratio of Ala-Pro-Tyr. Surprisingly, the α_R:β population ratio of the proline residue of the Ala-*cis*Pro-Tyr peptide is almost identical to the ratio extracted by MacArthur and Thornton (1991) from all the *cis*proline residues in their protein dataset. At the centres of the *cis*proline α_R and β regions, $\Delta\delta_{\beta\gamma}$ equals 7.3 ppm and 10.5 ppm, respectively. The measured $\Delta\delta_{\beta\gamma}$ in Ala-*cis*Pro-Tyr is 9.7 ppm. Thus, the α_R:β ratio of the Ala-*cis*Pro-Tyr peptide is estimated as 26%:74%. This finding confirms the β propensity of the *cis*proline previously derived from the $^{13}C\alpha$ chemical shift analysis. Unfortunately, the α_R:β population ratio of Ala-*trans*Pro-Tyr cannot be extracted in the same manner because the $\Delta\delta_{\beta\gamma}$ values of the *trans*proline α_R and β region centres are almost identical.

Comparison between experiment and computation

The agreement between the computational and experimental results is very good, as can be seen in Figure 4. Agreement is at both the structural level (chemical shifts and the ^{3}J-coupling constants) and at the thermodynamic level (population of the conformations).

The respective computed and measured δ_{ring} values are -0.55 ppm and -0.83 ppm for H$\alpha(i-2)$ and -0.14 ppm and -0.30 ppm for Hγ3$(i-1)$. This particular property is usually difficult to compute because it is very sensitive to the sampling of the phase space (Worth et al., 1998) and to the force field parameters (van der Spoel et al., 1996), and because it displays large fluctuations along the MD simulation (see Figure 3, two bottom graphs).

The experimental and computed ^{3}J-coupling constants are well correlated. The standard deviation between the computed and experimental values is 1.1 Hz, which corresponds to 10° standard deviation between the computed and experimental dihedral angles. The computed and measured populations of the tyrosine χ_1 (Figure 4c) and proline ψ (Figure 4d) rotamers are in good agreement. The MD simulation gives 18% and 82% (Table 3) for the proline α_R and β conformation populations, respectively, and the NMR gives 26% and 74% (Table 4), respectively. The ψ dihedral angle of the proline residue and the side-chain rotamer of the tyrosine residue are the main parameters that define formation of the *cis*proline$(i-1)$-aromatic(i) interaction. If the χ_1 dihedral of the tyrosine is *t*, the peptide is unfolded. When the proline is in the extended conformation, the type *a cis*proline$(i-1)$-aromatic(i) interaction is present 90% of the time. When the proline is in the α_R conformation, the type *b* interaction is found 90% of the time.

The population of the type *a* and *b cis*proline$(i-1)$-aromatic(i) interaction conformations and the 'unfolded' state is estimated, from the NMR data, to be 63%±5%, 27%±5%, and 10%±5%, respectively and, from the MD simulations, to be 72%, 24% and 4%, respectively.

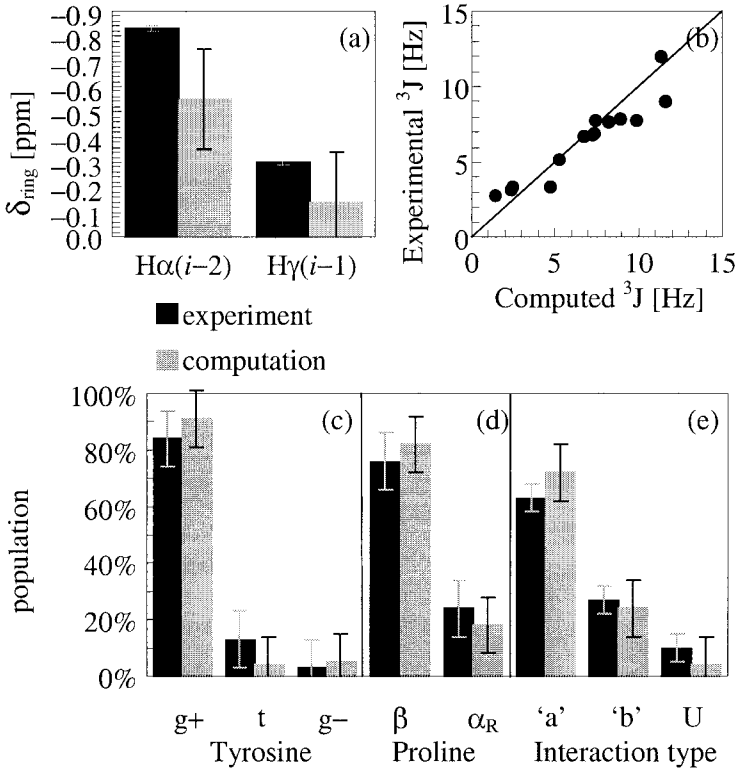

Figure 4. Comparison of the computational and experimental results for Ala-*cis*Pro-Tyr. (a) Hα($i-2$) and Hγ3($i-1$) δ_{ring}. (b) Experimental versus computed ^{3}J-coupling constant. (c) Population of tyrosine χ_1 rotamers, (d) population of the proline α_R and β conformations. (e) Population of the types *a* and *b cis*proline($i-1$)-aromatic(i) interaction conformers and the 'unfolded' state.

Conclusions

The 'folding landscape' of the Ala-*cis*Pro-Tyr peptide in aqueous solution is derived from this study. Two metastable, folded states, type *a* and type *b*, in which the tyrosine ring interacts with the alanine residue and the proline side-chain, respectively, coexist in aqueous solution. Each consists of a distribution of conformers. The two states are present 63% and 27% of the time, respectively, under the conditions of study, and hysteresis is observed upon transition between these two states.

The ratio between the populations of the type *a* and *b* conformations is dependent on sequence. Our simulations of the Gly-*cis*Pro-Tyr(i) peptide show that the type *a* conformation is much more stable relative to the type *b* conformation than for the Ala-*cis*Pro-Tyr(i) peptide. The introduction of a glycine at the ($i-2$) position allows more space for the tyrosine ring to interact with the aliphatic hydrogens of the $i-2$ residue, and thereby appears to entropically favour the type *a* *cis*proline($i-1$)-aromatic(i) interaction. In contrast, a bulky residue, such as tyrosine, at the position $i-2$ reduces the conformational space available for the type *a* *cis*proline($i-1$)-aromatic(i) interaction, and thus the type *b* interaction is favoured.

In this study, the Ala-*cis*Pro-Tyr peptide was chosen as a model peptide for the study of local interactions. It can be considered as a minimal model system exhibiting basic features of protein folding such as formation of hydrophobic contacts and hydrogen bonds. However, its existence in two folded states with hysteretic transitions between them would make Ala-*cis*Pro-Tyr well suited to function in signal transduction or memory mechanisms. Its local interactions could be stabilised by other interactions such as protein binding, long-range inter-residue interactions in a longer peptide chain, or by thermodynamic conditions. In this context, it is interesting to note that related short linear peptides that are potentially involved in tyrosine kinase activation (Li et al., 1997) and in acting as analgesic compounds (Zadina et al., 1997) have been discovered recently.

With regard to protein folding, the non-native property of the *cis*proline($i-1$)-aromatic(i) interaction found in unfolded BPTI may be thought to hinder protein folding. However, the presence of a local interaction, even if it is not native, can still considerably reduce the huge accessible conformation space and thus accelerate folding (Creighton et al., 1996). Furthermore, the BPTI folding intermediate, $[30–51, 5–14]_{Ser}$, with two disulphide bridges, one of which is non-native, may be stabilised by a *cis*proline($i-1$)-aromatic(i) interaction in which the conformation $a1$ dominates according to the chemical shifts measured by van Mierlo et al. (1994). Moreover, the conformational preferences of peptides containing Xaa-Pro sequences should influence their catalysis by peptidyl-prolyl isomerases and thus the rate of prolyl isomerisation upon folding. The crystal structure of the complex of cyclophilin A and the succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide substrate (Zhao and Ke, 1996) shows that the peptide binds in a type *a cis* conformation. Our calculations indicate that this would be the energetically preferred conformation of the *cis* form of this peptide. This supports a mechanism for the enzyme in which it would catalyse prolyl isomerisation by binding to and stabilising the dominant *cis* form of the substrate.

The present study demonstrates that to tackle the difficult problem of peptide folding in aqueous solution, it is useful to employ a wide range of strategies such as database searches, configurational sampling, molecular dynamics simulations and ^{1}H- and ^{13}C-NMR spectroscopy. In this case, the combined use of the complementary database search and simulation methods led to an atomic-detail model that is in remarkably good agreement with experimental measurements. The structural database analysis considerably reduced the size of the conformational space to explore in the peptide in solution, but was not sufficient to estimate the stability of each individual conformation. The MD simulation of the Ala-Pro-Tyr peptide led to a detailed scenario for the folding of the peptide, showing two main folded states.

Acknowledgements

We are very grateful to Tjerk Straatsma for providing the ARGOS program; Richard Jacob for synthesizing the Ala-Pro-Tyr peptide; Mike Williamson and David Case for providing their chemical shift calculation programs; Tom Creighton, Giovanna Musco, Gert Vriend and Graham Worth for helpful advice; and EMBL and the EU (ERBCHBGCT920073) for financial support.

References

Allen, F.H. and Kennard, O. (1993) *Chem. Des. Autom. News*, **8**, 31–37.

Becker, O.M. (1997) *J. Comput. Chem.*, **18**, 1893–1902.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W., Di Nola, A. and Haak, J. (1984) *J. Chem. Phys.*, **81**, 3684–3690.

Bernstein, F., Koetzle, T., Williams, G., Smith, E.J., Brice, M., Rodgers, J., Kennard, O., Shimanouchi, T. and Tasumi, T. (1977) *J. Mol. Biol.*, **112**, 535–542.

Bodenhausen, G. and Ruben, D.F. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.

Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, CA.

Creighton, T.E., Darby, N.J. and Kemmink, J. (1996) *FASEB J.*, **10**, 110–118.

Demchuk, E., Bashford, D. and Case, D.A. (1997a) *Folding Design*, **2**, 35–46.

Demchuk, E., Bashford, D., Gippert, G.P. and Case, D.A. (1997b) *J. Mol. Biol.*, **270**, 305–317.

Dyson, H.J., Rance, M., Houghten, R.A., Lerner, R.A. and Wright, P.E. (1988) *J. Mol. Biol.*, **201**, 161–200.

Fiebig, K.M., Schwalbe, H., Buck, M., Smith, L.J. and Dobson, C.M. (1996) *J. Phys. Chem.*, **100**, 2661–2666.

Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. and Schmid, F.X. (1989) *Nature*, **337**, 476–478.

Giessner-Prettre, C., Cung, M. and Marraud, M. (1987) *Eur. J. Biochem.*, **163**, 79–87.

Grathwohl, C. and Wüthrich, K. (1981) *Biopolymers*, **20**, 2623–2633.

Howart, O.W. (1978) *Prog. NMR Spectrosc.*, **12**, 1–40

IUPAC Commission (1970) *Biochemistry*, **9**, 3471–3479.

Jorgensen, W.L., Chandrasekhar, J. and Madura, J.D. (1983) *J. Chem. Phys.*, **79**, 926–935

Kemmink, J. and Creighton, T. (1993) *J. Mol. Biol.*, **234**, 861–878

Kemmink, J. and Creighton, T.E. (1995) *J. Mol. Biol.*, **245**, 251–260.

Lewis, P., Momany, F. and Scheraga, H. (1973) *Biochim. Biophys. Acta*, **303**, 211–229.

Li, S.C., Songyang, Z., Vincent, S.J., Zwahlen, C., Wiley, S., Cantley, L., Kay, L.E., Forman-Kay, J. and Pawson, T. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 7204–7209.

MacArthur, M.W. and Thornton, J.M. (1991) *J. Mol. Biol.*, **218**, 397–412.

MacKerell, A.D.J., Bashford, D., Bellott, M., Dunbrack, J.R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher III, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D. and Karplus, M. (1998) *J. Phys. Chem. B*, **102**, 3586–3616.

Markley, J., Bax, A., Arata, Y., Hilbers, C., Kaptein, R., Sykes, B., Wright, P. and Wüthrich, K. (1998) *J. Biomol. NMR*, **12**, 1–23.

McClain, R. and Erickson, B. (1995) *Int. J. Pept. Protein Res.*, **45**, 272–281.

Merutka, G., Dyson, H.J. and Wright, P.E. (1995) *J. Biomol. NMR*, **5**, 14–24.
Miler-White, E.J., Bell, L.H. and Maccallum, P.H. (1992) *J. Mol. Biol.*, **228**, 725–734.
Mohanty, D., Elber, R., Thirumalai, D., Beglov, D. and Roux, B. (1997) *J. Mol. Biol.*, **272**, 423–442.
Molecular Simulations Inc. (1992) *QUANTA release 40*, Molecular Simulations Inc., San Diego, CA.
Momany, F.A. and Rone, R. (1992) *J. Comput. Chem.*, **13**, 888–900.
Montelione, G.T., Arnold, E., Meinwald, Y.C., Stimson, E.R., Denton, J.B., Huang, S.-G., Clardy, J. and Scheraga, H.A. (1984) *J. Am. Chem. Soc.*, **106**, 7946–7958.
Müller, L. (1987) *J. Magn. Reson.*, **72**, 191–196.
Nardi, F., Worth, G.A. and Wade, R.C. (1997) *Folding Design*, **2**, S62–S68.
Oka, M., Montelione, G.T. and Scheraga, H.A. (1984) *J. Am. Chem. Soc.*, **106**, 7959–7969.
Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
Ponder, J. and Richards, F. (1987) *J. Mol. Biol.*, **193**, 775–791.
Ramachandran, G. and Sasisekharan, V. (1968) *Adv. Protein Chem.*, **23**, 283–438.
Ripoll, D.R., Vila, J.A., Villegas, M.E. and Scheraga, H.A. (1999) *J. Mol. Biol.*, **292**, 431–440.
Ryckaert, J.-P., Cicotti, G. and Berendsen, H.J.C. (1977) *J. Comput. Chem.*, **23**, 327–341.
Siemion, I.Z. (1976) *Org. Magn. Reson.*, **8**, 432–435.
Siemion, I.Z., Wieland, T. and Pook, K.H. (1975) *Angew. Chem. Int. Ed. Engl.*, **14**, 702–703.
Smith, L.J., Fiebig, K.M., Schwalbe, H. and Dobson, C.M. (1996) *Folding Design*, **1**, R95–106.
Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.
Stewart, D.E., Sartar, A. and Wamler, T.E. (1990) *J. Mol. Biol.*, **214**, 254–260.
Straatsma, T.P. and McCammon, J.A. (1990) *J. Comput. Chem.*, **11**, 943–951.
van der Spoel, D., van Buuren, A.R., Tielemen, D.P. and Berendsen, H.J.C. (1996) *J. Biomol. NMR*, **8**, 229–238.
van Mierlo, C.P.M., Kemmink, J., Neuhaus, D., Darby, N.J. and Creighton, T.E. (1994) *J. Mol. Biol.*, **235**, 1044–1061.
Vriend, G., Sander, C. and Stouten, P.F.W. (1994) *Protein Eng.*, **7**, 23–29.
Williamson, M.P. and Asakura, T. (1993) *J. Magn. Reson.*, **B101**, 63–71.
Wilmot, C. and Thornton, J. (1990) *Protein Eng.*, **3**, 479–493.
Wishart, D. and Sykes, B. (1994) *Methods Enzymol.*, **239**, 363–392.
Wishart, D.S., Bigam, C., Holm, A., Hodges, R. and Sykes, B.D. (1995) *J. Biomol. NMR*, **5**, 67–81.
Worth, G.A., Nardi, F. and Wade, R.C. (1998) *J. Phys. Chem.*, **102**, 6260–6272.
Wu, W.J. and Raleigh, D.P. (1998) *Biopolymers*, **45**, 381–394.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.
Yao, J., Dyson, H.J. and Wright, P.E. (1994) *J. Mol. Biol.*, **243**, 754–766.
Yao, J., Feher, V.A., Espejo, B.F., Reymond, M.T., Wright, P.E. and Dyson, H.J. (1994) *J. Mol. Biol.*, **243**, 736–753.
Zadina, J.E., Hackler, L., Ge, L.J. and Kastin, A.J. (1997) *Nature*, **386**, 499–502.
Zhao, Y. and Ke, H. (1996a) *Biochemistry*, **35**, 7356–7361.
Zhao, Y. and Ke, H. (1996b) *Biochemistry*, **35**, 7362–7368.