PROTEINS: Structure, Function, and Genetics 29:203–211 (1997)

Structure of the LAV6 Peptide: A Nucleation Site for the Correct Receptor-Induced Refolding of the CD4-Binding Domain of HIV1 gp 120

Almut Lindemann,[1]* Volker Kinzel,[1] Paul Rösch,[2] and Jennifer Reed[1]
[1]*Department of Pathochemistry, German Cancer Research Center, D-69120 Heidelberg, Germany*
[2]*Lehrstuhl für Biopolymere, Universität Bayreuth, D-95447 Bayreuth, Germany*

ABSTRACT **LAV44 and LAV15 (lymphadenopathy-associated virus) peptides of the CD4-binding region of gp 120 per se bind to the CD4 receptor (Reed and Kinzel, Biochemistry 30: 4521–4528, 1991; Lasky et al., Cell 50:975–985, 1987). Depending on the environment, the LAV peptides exhibit the ability to switch cooperatively between β-sheet and helical conformation when solvent polarity is changed past a critical point. This property, which is dependent on the amino acid sequence LPCR, is crucial for receptor binding (Reed and Kinzel, Proc. Natl. Acad. Sci. U.S.A. 90:6761–6765, 1993). Structure determination with 2D-NMR-spectroscopy reveals that LAV6 peptide (sequence: TLPCRI) has a well-defined structure, partially exhibiting inverse γ-turn conformation in aqueous solution. Quantitative evaluation of the NMR data discloses 90% trans-conformation for the peptide bond between leucine and proline. The Ψ- and Φ-angles fall into the typical range for amino acids located in turns. On the other hand, the amino acid sequence C-terminal to the LPCR tetrad has been shown to fold atypically in the absence of these residues. All these results show that the short sequence of LAV6 peptide, with the central amino acids LPCR, displays a matrix-independent structure and may, therefore, act as a conformational template for forming secondary structure in the intact CD4-binding domain of gp 120. Proteins 29:203–211, 1997.**

Key words: peptide folding; NMR; gp 120; CD4 binding domain; inverse γ-turn; nucleation site

INTRODUCTION

A 15 residue[1] sequence within the principal CD4-binding domain of the *env* glycoprotein gp 120 from HIV1[2] has the property of undergoing a highly cooperative conformational change from β-sheet to helix when medium polarity is lowered past a critical point.[1] This behavior is interesting for at least two reasons: first, because the cooperative all-or-nothing character of the transition correlates with the ability of synthetic peptides containing this sequence to bind to CD4-expressing cells, and second, because this behavior is conserved thoughout the HIV1 strains tested despite 50% variability in the sequence, providing an ideal model system for the interactions involved in governing a precise case of protein folding/refolding in solution. Previous work has shown that the helical form of the peptide is a 3_{10}- rather than an α-helix,[4] and that there is no helix formation at all in the absence of the N-terminal tetrad LPCR, the response to a reduction in medium polarity then being simply an increase in the content of β-sheet.[3]

The idea that the nascent chain of an unfolded protein might begin to refold at specifically designed points of collapse, or "nucleation sites"[5,6] was introduced over 20 years ago and has been repeated in various forms and guises since.[7–11] Evidence exists, on the one hand, that some short peptides are capable of folding into a surprisingly well-defined conformation in isolation,[12–19] while on the other hand, recent pulse labeling experiments have shown that discrete secondary structure elements are formed at the earliest observable point in the refolding processes, within 5 ms.[20,21] This strongly suggests that the folding of these elements is catalyzed by preformed micro-structures, i.e., nucleation sites.

As the 3_{10}-helix consists essentially of a series of stacked, overlapping β-bends, it has been proposed that a strongly favored β-bend sequence may act as a conformational template, favoring the formation of a 3_{10}-helix which might then rearrange as an α-helix.[3,22] The LPCR tetrad is well placed to act as a nucleation site as it is: a) possessed of a high theoretical β-bend forming propensity[23]; b) associated with a C-terminal 3_{10}-helix; and c) an absolute requirement for helical conformation in the peptide.

If this is the case, the LPCR sequence alone should be capable of folding into a well-defined template structure. We have used ^{1}H-NMR to structurally examine the LPCR sequence plus the native N- and

*Correspondence to: Almut Lindemann, Department of Pathochemistry, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany.
Received 13 December 1996; Accepted 24 April 1997

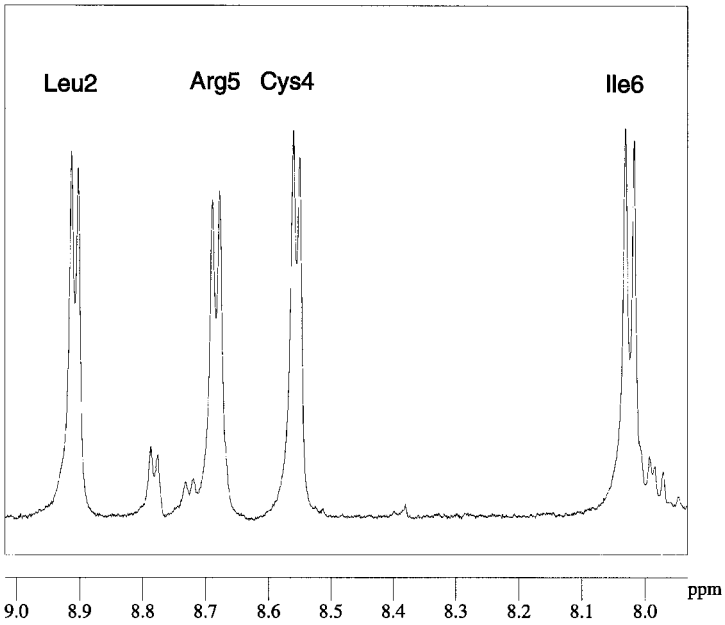

Fig. 1. NH region of ^{1}H-NMR spectrum at 278 K of 10 mM LAV6 peptide in 10 mM sodium phosphate buffer, pH 4.5.

TABLE I. ^{1}H Sequential Assignments for LAV6 Peptide†

Res.	HN	HA	HB	HG	Others	Shift	Conf.
Thr 1		3.82	4.05	1.29		0.54	I
Leu 2	8.91	4.64	1.69, 1.62	1.56	HD* = 0.93	−0.3	90%
Pro 3		4.47	2.28, 1.89	2.02	HD1 = 3.84, HD2 = 3.64	−0.06	
Cys 4	8.55	4.62	3.17, 2.95			−0.07	
Arg 5	8.68	4.41	1.81, 1.72	1.60	HE = 7.23, HH1* = 6.92, HH2* = 6.50, HD* = 3.16	−0.07	
Ile 6	8.02	4.04	1.79	1.40, 1.12	HD* = 0.86	0.13	
Thr 1		3.79	4.06	1.32			II
Leu 2	8.79	4.36	1.61	1.45	HD* = 0.90		10%
Pro 3		4.53	2.35, 2.14	1.94, 1.77	HD1 = 3.57, HD2 = 3.46		
Cys 4	8.78	4.68	3.23, 3.07				
Arg 5	8.73	4.41	1.82, 1.73	1.60	HE = 7.25, HD* = 3.16		
Ile6	8.00	4.04	1.79	1.39, 1.11	HD* = 0.86		

†Dissolved in 10 mM sodium phosphate buffer, pH 4.5 at 278 K. DSS is used as an external standard. The chemical shifts for the trans-(I) and cis-conformation (II) are shown. Asterisks mark protons with degenerate chemical shifts. Additionally, the $C_\alpha H$ upfield shift relative to the random coil-values[51] are shown in column 7 (shift).

C-terminal residues, and have observed it to have a strong preference for a particular bend type conformation in solution. Thus, a potential nucleation site has been identified, shown to have the necessary characteristics of independent folding, and demonstrated to be essential for the subsequent correct formation of a secondary structure domain.

MATERIALS AND METHODS

Peptide Synthesis

The 6-residue peptide with the sequence TLPCRI is part of the switch domain from the CD4-binding region of gp 120, running from Thr 399 to Ile 404. It was synthesized on an automated peptide synthesizer (ABI 433) using standard F_{moc} protocols.[24,25] Peptide chain assembly was performed using in situ activation of amino acid building blocks by 2-(1H-benzotriazole-1-yl)-1.1.3.3.-tetramethyluronium hexafluorophosphate (HBTU; Nova Biochem GmbH, Bad Soden/Ts, Germany). The purity was checked with reverse phase-HPLC (Komasil C18, 5 µm) and MALDI (Finnigan MAT VISION2000). Both methods showed a single signal, characteristic of a homogeneous sample. The lyophilized product was stored at −20°C.

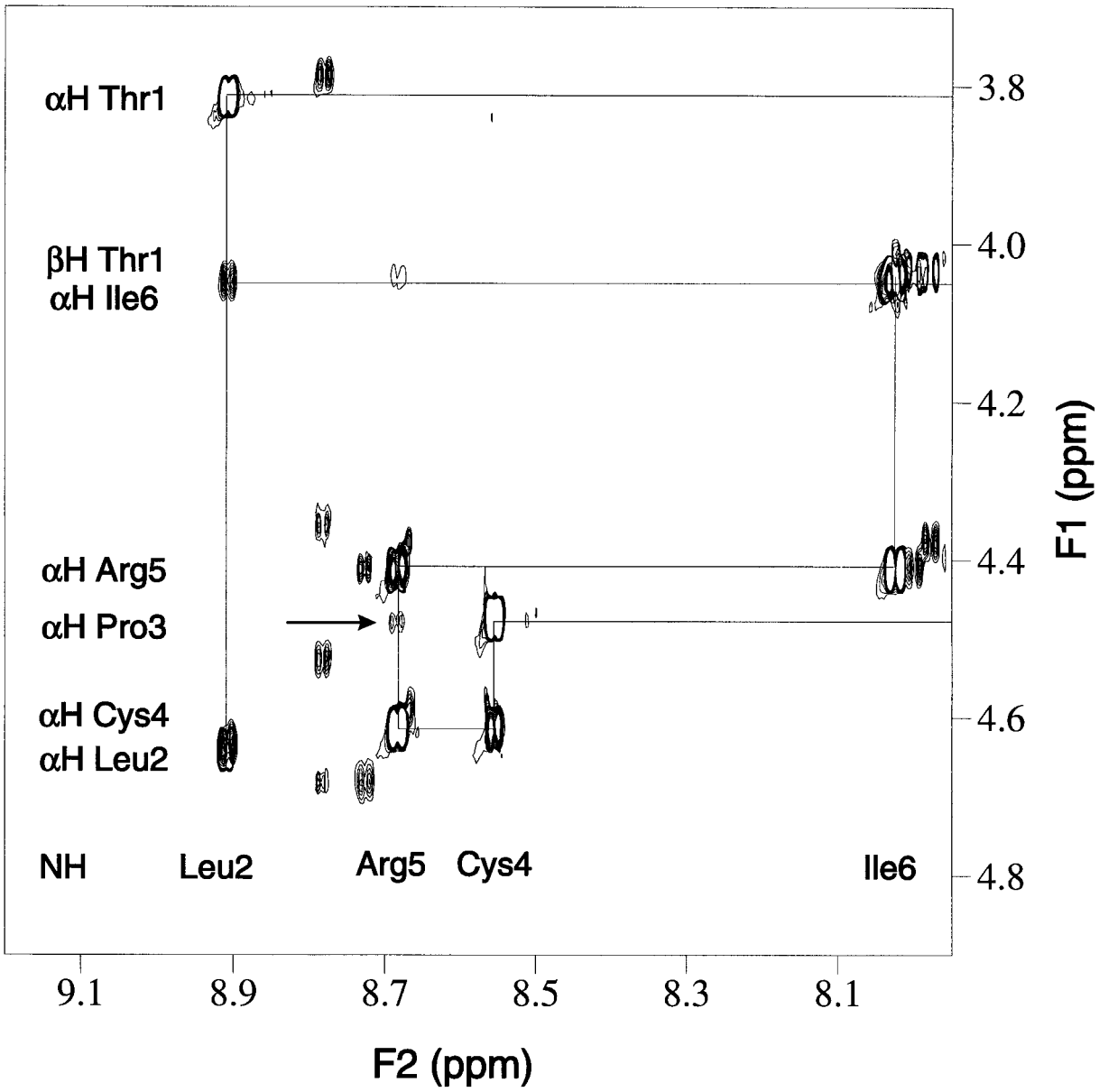

Fig. 2. 200 ms NOESY spectrum at 278 K of 10 mM LAV6 in 10 mM sodium phosphate buffer, pH 4.5. Sequential $C_\alpha H$-NH connectivities (chaintracing) in the fingerprint region for all residues of the trans form are shown. Additionally the NOE between NH of Arg5 and $C_\alpha H$ of Pro3 is marked by an arrow.

NMR Spectroscopy

Peptide samples were dissolved in 0.5 ml 10 mM sodium phosphate buffer, pH 4.5 $H_2O/^2H_2O$ (9:1) to a final concentration of 10 mM. All NMR experiments were performed on standard Bruker AMX600 and AMX400 spectrometers at 278 K. 2,2-dimethyl-2-silapentane sulfonate (DSS; Sigma-Aldrich Chemie GmbH, Diesenhofen, Germany) was used as an external standard, accuracy ±0.01 ppm. Temperature control and compensation in sample heating between TOCSY and other spectra were ensured by the VT1000 temperature unit calibrated for the different experiments. The water resonance was presaturated by continuous coherent irradiation at the resonance frequency prior to the reading pulse. Standard pulse sequences were used for double-quantum-filtered correlated spectroscopy (DQF-COSY),[26] nuclear Overhauser enhancement spectroscopy (NOESY),[27] and total correlation spectroscopy (TOCSY).[28] Mixing time in NOESY experiments was 200 ms and 80 ms in the TOCSY experiments. Quadrature detection was used in both dimensions

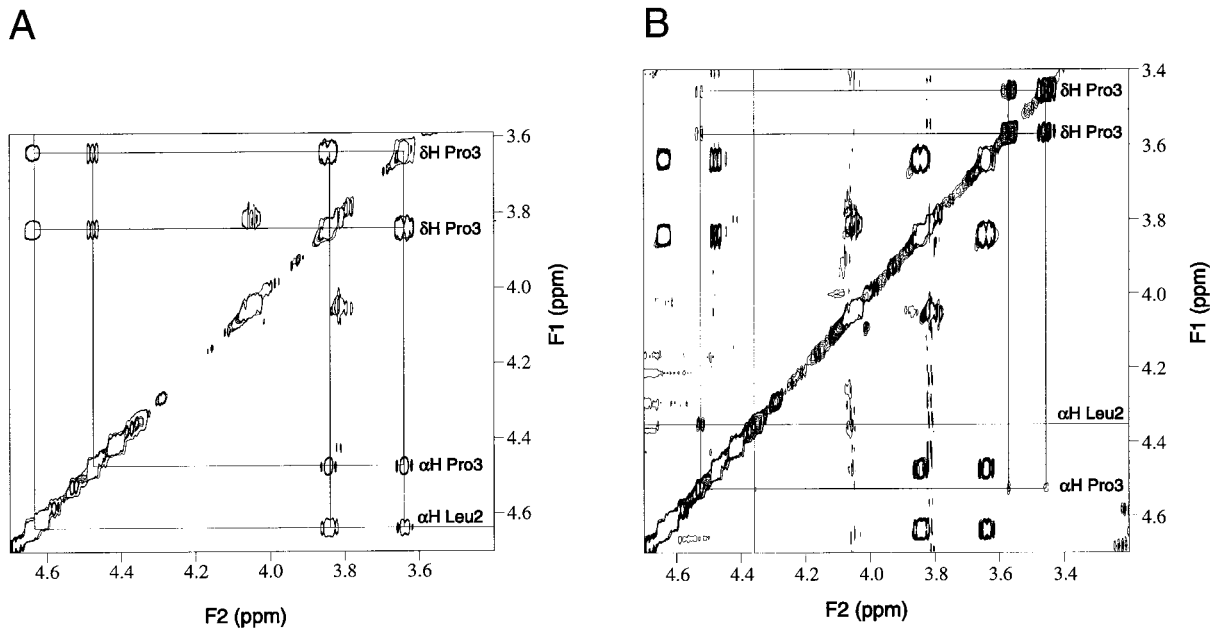

Fig. 3. Part of the aliphatic region of the NOESY spectrum (200 ms, 278 K) of 10 mM LAV6 in 10 mM sodium phosphate buffer, pH 4.5. **A:** The typical NOEs between $C_\alpha H$ of Leu2 and $C_\delta H$ of Pro3 in trans-conformation are marked. **B:** For the cis-conformation, the NOE between the α-protons of Leu2 and Pro3 is characteristic and can be found in a lower intensity (10% cis-conformation).

employing the time proportional phase incrementation (TPPI) technique.[29] The sweep width in ω_1 and ω_2 was 6039 Hz. In one-dimensional experiments, 32 K data points were collected, and 4 K data points were collected in ω_2 and 512 data points in ω_1, in two-dimensional experiments. Zero filling to 1 K data points in ω_1 was applied. The FID (free induction decay) was multiplied with a squared sinebell function phase shifted by $\pi/4$ for TOCSY and NOESY spectra, and a sinebell function phase shifted by $\pi/32$ for COSY spectra. Baseline correction to the 7th order and phase correction to the 2nd order was used. All data processing was done with the NDee program package.[30]

Structure Calculation

Distance information was obtained from NOESY spectra. Ring-protons from Pro3 were used to calibrate the intensities of the NOESY cross peaks. NOE cross-peak intensities were divided into three categories according to their intensities: strong, 0.18–0.27 nm; medium, 0.18–0.33 nm; weak, 0.18–0.5 nm.[31] Pseudoatom correction was used to adjust distances that involve non-stereospecific atom groups like methyl protons; in this case, 0.05 nm were added to the upper distance limits.[32] Dihedral angle constraints were obtained from 1D-NMR spectrum taking into account scalar $^3J_{HN\alpha}$ coupling constants less than 6.5 Hz using the Karplus equation.[33] X-PLOR 3.1 package[34] was used to obtain the initial three-dimensional structure by restrained molecular dynamics calculation. For ab initio simulated annealing and refinement, standard protocols were used.[35] Starting from an extended template structure, 50 ps molecular dynamics at 1,000 K were used, followed by cooling to 100 K in 30 ps and 200 steps of energy minimization. In the final round of the refinement, electrostatic and van der Waals' interactions were treated explicitly using the CHARMM potential,[36] as described previously.[37] In order to be certain that no artefacts could arise from the potential function used, and that the fold is an intrinsic feature of the sequence, two independent calculations, with and without experimental restraints, were carried out.

Structural Analysis

The final structures were analyzed with respect to stable idealized elements of regular secondary structure using the Define Secondary Structure of Proteins (DSSP) program package[38] and PROMOTIV.[39] For visualization of structure data, the SYBYL program package, version 6.0 (TRIPOS Association) was used.

RESULTS AND DISCUSSION

LAV6 is a highly soluble peptide, and one- and two-dimensional NMR spectra of a 10 mM peptide solution showed very good resolution. Two sets of NH-signals occur between 8.95 and 7.9 ppm (Fig. 1). After integration, it was found that the ratio of the main and the minor resonances is 9:1. This was interpreted as originating from a major and a minor structured species of the peptide. From the TOCSY

and DQF-COSY spectra, it was possible to identify all spinsystems (Table I) and make the sequence-specific assignment from NOESY spectrum via the NH-C_αH chaintracing for the main (Fig. 2) and the secondary conformation. The backbone connectivities through NH resonances broke at Pro3, but in the aliphatic region of the 2D spectra, the typical NOEs for the cis- and trans-isomers at Pro3 were visible.[40] The main signals (populated by 90%) show NOEs between the α-proton of Leu2 and the δ-protons of Pro3, which are typical for the trans-conformation of proline (Fig. 3A), whereas in the second set of signals (populated by 10%), the NOE between the α-protons of Leu2 and Pro3 is observed, as expected for the cis-conformation (Fig. 3B). Therefore, the two sets of NMR signals are clearly identified as resulting from the cis- and trans-conformation of Pro3.

The energy barrier for the cis-trans isomerization of proline compared to other peptide bonds is lowered from 20 kcal/mol to 13 kcal/mol.[41] One reason for this is the longer peptide bond between proline and the preceding residue, which results from the redistribution of charge and the lack of resonance stabilization on loss of the NH proton.[42] Therefore, the type of the amino acids preceding and following proline influence the cis-trans ratio of peptide bonds involving prolines.[43,44] For short oligopeptides, the cis-trans equilibrium for X-Pro can vary between 3:7 and 1:9.[43,45] The equilibrium is affected by the bulkiness of the side chain of X,[46] and by the protonation-deprotonation reaction of the carboxylic acid terminus (in dipeptides), and therefore, of the pH of the solvent.[47,48] The LAV6 peptide shows a 1:9 cis:trans ratio which is a rather high trans portion, but might still be considered as within the extreme end of the normal limits.

In proteins, only trans imides (X-Pro) are found at the beginning or the end of helices, whereas cis imides appear primarily in bends and turns (73%).[49,50] This suggests a specific structural role for the type of X-Pro bonding. Also, there are some strong sequence preferences for cis and trans isomers in proline containing proteins.[51] In the case of cis proline, the N-terminal neighboring position is strongly enriched with the amino acids Tyr, Phe, Thr, and Leu (50% occurrence for cis-Pro, compared with 16% for trans-Pro and 14% for all X-Pro).[51,52] With this in mind, the cis:trans ratio of 1:9 for LAV6 (TLPCRI) where a Leu preceeds the Pro is, in fact, unusually high. As shown below, LAV6 has a well-defined set of structural characteristics in solution. The unexpected high 90% trans-isomer is a further indication of restricted conformational freedom.

Tendencies for forming certain secondary structure elements were estimated from the NMR data employing the C_αH-chemical shift index[53] and the $^3J_{HN\alpha}$ coupling constants.[33] Helical secondary structure and turns cause an upfield shift of the C_αH resonances compared to the random-coil values.[54]

TABLE II. NOE and Molecular-Dynamics Statistic*

Parameter	Value
No. of NOE	
Total	102
Intraresidual $\|i–j\| = 0$	74
Sequential $\|i–j\| = 1$	24
Medium range $\|i–j\| = 2, 3, 4, 5$	4
Dihedral angle restraints	2
Average energy	
E_{tot}	37 ± 33 kJ mol^{-1}
E_{bond}	40 ± 2 kJ mol^{-1}
E_{angle}	112 ± 14 kJ mol^{-1}
$E_{inproper}$	7 ± 2 kJ mol^{-1}
E_{VDW}	-120 ± 3 kJ mol^{-1}
E_{NOE}	145 ± 16 kJ mol^{-1}
E_{elec}	-147 ± 25 kJ mol^{-1}
RMSD of ideal geometry	
Bond	0.000 ± 0.000 nm
Angle	0.95 ± 0.06 deg
Improper	0.47 ± 0.07 deg
NOE	0.008 ± 0.000 nm
RMSD	
Backbone	0.048 nm
Heavy atoms	0.123 nm

*Energy contributions of the calculated structures of LAV6 and deviations from the standard geometry, E_{tot}, total energy; E_{VDW}, van der Waals' energy; E_{NOE}, effective NOE energy term resulting from a soft square-well potential function. The negative electrostatic energy (E_{elec}) is evidence for the validity of the structure determination procedure.[59] All calculations were carried out using the standard X-PLOR force-field and energy terms.[34] The values are mean values of 80 refined structures. The RMSD (root mean square deviation) for all calculated structures is given for the peptide backbone and all atoms excluding hydrogen (heavy atoms).

The four central C_αH resonances in LAV6 (Leu2 to Arg5) were shifted slightly upfield (Table I). Additionally the $^3J_{HN\alpha}$ coupling constants of Leu2 and Cys4 (6.3 Hz, 6.1 Hz), which are correlated with the torsion angle Φ in polypeptides,[33] show turn- or 3_{10}-helix-typical Φ-angles in the LAV6 peptide.

Secondary structure prediction algorithms for the residues Cys402 to Val414 of gp120 of HIV1 found a high helix-forming potential.[1] The N-terminus of this area overlaps with the residues Cys4, Arg5, and Ile6 of LAV6. This result fits very well with the experimental findings in the LAV6 peptide (C_αH-chemical shift index, $^3J_{HN\alpha}$ coupling constants) and indicates that the small peptide itself has a potential to form a helix-compatible structure and possibly initiate the folding of the N-terminal helix in LAV44 or LAV15.

One hundred and two NOEs are observed in the NOESY spectrum (Table II), of which 28 are interresidual. Four NOEs are found between the protons of Pro3 and Arg5 (Fig. 2: NH Arg5—C_αH Pro3), evidence for the spatially close position of these amino acids. These four medium-range NOEs, together with the NH-NH NOE between Arg5 and Ile6, and the two torsion angles for Leu2 and Cys4, determine

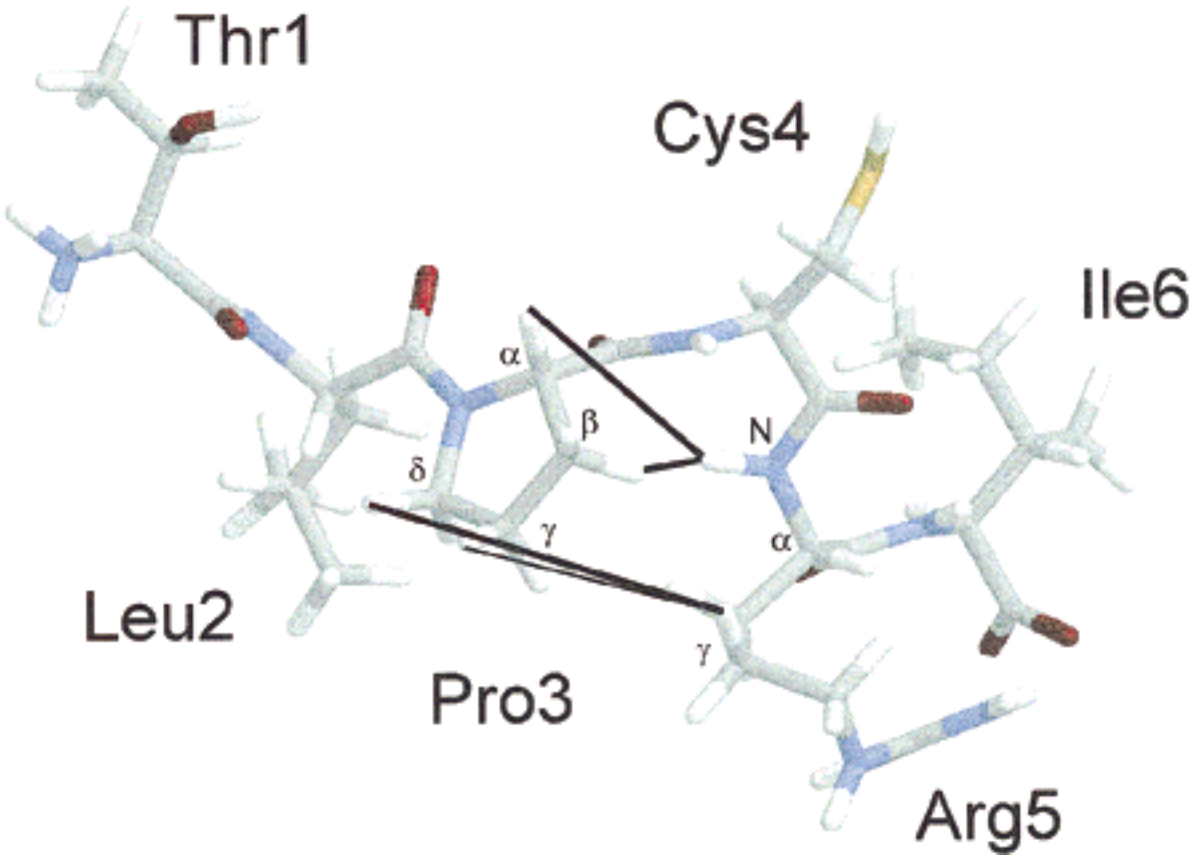

Fig. 4. One representative structure of LAV6 with all visible NOEs between Leu2 and Pro3 used for structure calculations (thick lines). Because of the t_1-noise in the NOESY spectrum, the NOEs between the γ-protons of the two residues cannot be resolved and used for structure calculations (thin line).

the secondary structure of LAV6. An NOE between the γ-protons of Pro3 and Arg5 was expected from structure calculations, but was not clearly observable in the NOESY spectrum because of t_1-noise in the relevant region (thin line in Fig. 4).

A restrained simulated annealing (SA) and SA refinement produced a well-defined three-dimensional structure of LAV6 (Figs. 5, 6B; Table II). In all structures none of the restraints was violated (volation <0.05 nm). The RMSD (root mean squre deviation) for 80 calculated structures was 0.048 nm for the backbone of the peptide and 0.123 nm for all non-proton atoms, evidence for a homogeneous and well-defined family of structures.

The structural data were analyzed with the DSSP (define secondary structure of proteins) program package.[38] In almost all resulting LAV6 structures (82%), a characteristic chirality from residue Leu2 to Cys4 was found. The dihedral angle α(i) [defined between C_α(i–1), C_α(i), C_α(i + 1), C_α(i + 2)] for Leu2 and Pro3 was negative [$-180° < \alpha(i) < 0°$] and positive [$0° < \alpha(i) < 180°$] for Cys4, respectively. A bent structure around Cys4, i.e., a region with a curvature of more than 70° was found in 28% of all structures. An idiosyncracy of the DSSP program is that chirality is only given for amino acids with one N-terminal and two C-terminal flanking residues; for this reason, a value for Arg5 and Ile6 would not be given, even where chirality is clearly present. But the C-terminal residues are involved in forming secondary structure, as shown by the participation of these residues in turn patterns: in 38% of all LAV6 structures, the program found an inverse γ-turn for residue Pro3 to Cys4, and in 68%, the beginning of the inverse γ-turn starts one position later (Cys4-Ile6).

This is, in fact, a lower limit for the turn structure present, as the DSSP program defines a γ-turn as three consecutive residues (i, i + 1, i + 2) which form a hydrogen bond between CO of residue i and NH of i + 2, where the Φ, Ψ angles are only defined for residue i + 1.[38] This is an overly restrictive definition, as the distance between the hydrogen bonding partners is often reported to be longer than ideal, and this hydrogen bond is not of major importance to the stability of the turns.[55,56] In the case of LAV6, the difference between the resulting structures with and without the inverse γ-turns at the two positions is very small. In the structures where no inverse γ-turn was classified, the distance for the protons forming

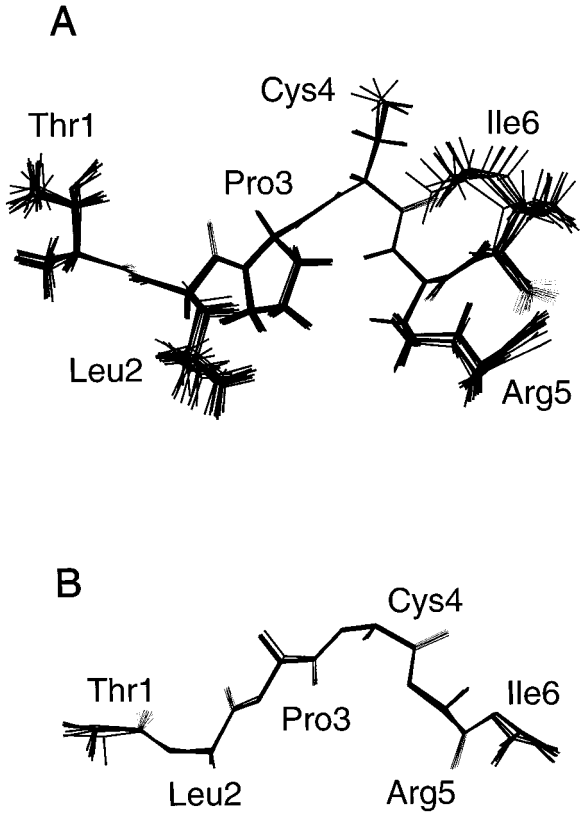

Fig. 5. Ten superposed structures of backbone atoms were obtained from restrained molecular dynamics calculation (**A:** with sidechains; **B:** backbone). These structures showed the lowest energy and a bend structure for the central residues LPC, with a negative chirality for Leu2 and Pro3, and a positive chirality for Cys4, after analyzing the structures with the DSSP (define secondary structure of proteins) program package.[38]

the hydrogen bond is only a little longer than 0.22 nm, the distance up to which DSSP defines an inverse γ-turn.[38] The global geometry, however, for all structures is equal. The RMSD (root mean square deviation, Table II) and the Ramachandran plot for the family of structures (Fig. 6) with well-defined Φ, Ψ angles for all residues is evidence for this.

Pro3, one of the central amino acids, takes part in all observed medium-range NOEs, and is thus involved in building up the LAV6 secondary structure. It is important to note that Pro3 itself is not able to form this structure, because short-range interactions dominate the conformational properties of proline containing peptides.[57] In Pro-X dipeptides, proline stabilizes an extended structure rather than a bent one, if X has a small sidechain like X = Gly, Ser, or Ala. Bent structures are favored for bulky sidechains, for example X = Leu, Val, Phe, Asp, Asn.[58–60] The reason for this is a minimization of the backbone-backbone interaction in the case of small sidechains, and the possibility of separation of the sidechains in the bend conformation for bulky sidechains.[58] In the case of LAV6, Pro3 is followed by Cys4. The size of the sidechain of Cys is comparable with that of Ser, and therefore an extended conformation would be expected. In contrast, the experimental structure of LAV6 shows a change of the chirality from Leu2 to Cys4 visible in a bent arrangement (seen in Fig. 5) with less conformational freedom. Additionally the Φ- and Ψ-angles for Pro3 are within allowed narrow ranges at one side of the β-region (Fig. 6B).[61] This demonstrates that the structure is not merely determined by the presence of Pro3 and the following residue alone, but by the combination

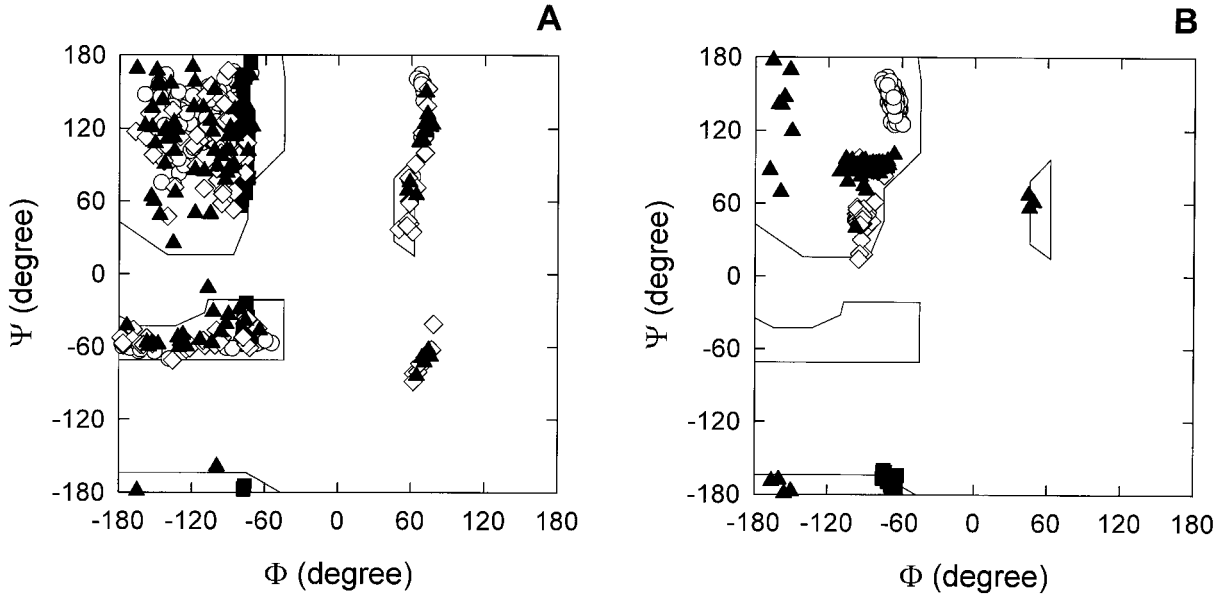

Fig. 6. Ramachandran plot of LAV6 for 80 calculated structures with distance and angle restraints from NMR spectra (**B**) and without structural restrains (**A**). ○ Pro3, ■ Cys4, ◇ Arg5, ▲ Ile6. The extreme difference in the areas occupied in the Ramachandran plot, after the NOEs are included, indicates the defined nature of the peptide structure in solution.

of all amino acids. Therefore, LAV6 has a complex structure with important properties of medium-range interactions in solution which are able to play a specific role in the context of the CD4 binding domain of gp120. The high conservation of the tetrad LPCR for nearly all known mutants of the HIV1 gp120 protein, as taken from the SwissProt data bank, is evidence for the importance of this sequence, and thus the described structure is of general importance for the interaction of HIV1 with CD4.

CONCLUSION

The peptide TLPCRI has been found to adopt a discrete bend structure in solution. The core of this sequence, the tetrad LPCR, is found at the N-terminus of the 15-residue switch domain of gp120, which retains the ability to bind to CD4-expressing cells.[1] In the native protein, the LPCR tetrad is followed C-terminally by residues which alter their secondary structure from β-sheet to a 3_{10}-helix in a cooperative manner when medium polarity is reduced; this conformational switch appears necessary for CD4 binding.[3] When the N-terminal LPCR residues are not present, the switch domain is incapable of folding into a 3_{10}-helix, even when the medium is totally apolar.[3] The LPCR tetrad thus has all the characteristics of a nucleation site,[5–11] inducing correct protein folding: a) the tetrad with flanking residues can adopt a defined, stable structure in isolation in solution; b) the Φ-angles adopted spontaneously are compatible with those of a 3_{10}-helix; c) native folding of subsequent residues in the domain does not occur without this sequence.

It is of special interest that with this sequence, we not only have a "smoking gun," so to speak, in that it has been shown that the micro-structure is required for proper folding, but also evidence that nucleation sites may, in some cases, play an active role even after initial chain folding is completed, acting as latent triggers for functional conformational change.

ACKNOWLEDGMENT

We thank Dr. Rüdiger Pipkorn for the LAV6 peptide preparation, and Dr. Gabi Seidel for helpful discussions. This work was supported by a grant from the Deutsche Forschungsanstalt für Luft- und Raumfahrt, Projektträger Gesundheitsforschung des BMBF.

REFERENCES

1. Reed, J., Kinzel, V. A conformational switch is associated with receptor affinity in peptides derived from the CD4-binding domain of gp120 from HIV I. Biochemistry 30:4521–4528, 1991.
2. Lasky, L.A., Nakamura, G., Smith, D.H., Fennie, C., Shimasaki, C., Patzer, D., Berman, P., Gregory, T., Capon, D.J. Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor. Cell 50:975–985, 1987.
3. Reed, J., Kinzel, V. Primary structure elements responsible for the conformational switch in the envelope glycoprotein gp120 from human immunodeficiency virus type 1: LPCR is a motif governing folding. Proc. Natl. Acad. Sci. U.S.A. 90:6761–6765, 1993.
4. Stosch, A.V., Jiménez, M.A., Kinzel, V., Reed, J. Solvent polarity-dependent structural refolding: A CD and NMR study of a 15 residue peptide. Proteins 23:196–203, 1995.
5. Wetlaufer, D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. Proc. Natl. Acad. Sci. U.S.A. 70:697–701, 1973.
6. Wetlaufer, D.B. Nucleation in protein folding—confusion of structure and process. TIBS 15:414–415, 1990.
7. Tsong, T.Y., Baldwin, R.L., McPhie, P. A sequential model of nucleation-dependent protein folding: kinetic studies of ribonuclease Am. J. Mol. Biol. 63:453–475, 1972.
8. Sachs, D., Schechter, A.N., Eastlake, A., Anfinsen, C.B. An immunologic approach to the conformational equilibria of polypeptides. Proc. Natl. Acad. Sci. U.S.A. 69:3790–3794, 1972.
9. Rose, G.D., Winter, R.H., Wetlaufer, D.B. A testable model for protein folding. FEBS Lett. 63:10–16, 1976.
10. Kim, P.S., Baldwin, R.L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. Annu. Rev. Biochem. 51:459–489, 1982.
11. Montelione, G.T., Scheraga, H.A. Formation of local structures in protein folding. Acc. Chem. Res. 22:70–76, 1989.
12. Brown, J.E., Klee, W.A. Helix-coil transition of the isolated amino terminus of ribonuclease. Biochemistry 10:470–476, 1971.
13. Bundi, A., Andreatta, R.H., Wüthrich, K. Characterisation of a local structure in the synthetic parathyroid hormone fragment 1-34 by ^{1}H nuclear-magnetic-resonance techniques. Eur. J. Biochem. 91:201–208, 1978.
14. Boesch, C., Bundi, A., Oppliger, M., Wüthrich, K. 1H nuclear-magnetic resonance studies of the molecular conformation of monomeric glucagon in aqueous solution. Eur. J. Biochem. 91:209–214, 1978.
15. Dyson, H.J., Cross, K.J., Houghton, R.A., Wilson, I.A., Wright, P.E., Lerner, R.A. The immunodominant site of a synthetic immunogen has a conformational preference in water of a type II reserve turn. Nature 318:480–483, 1985.
16. Haas, E., Montelione, G.T., McWherter, C.A., Scheraga, H.A. Local structure in a tryptic fragment of performic acid oxidized ribonuclease A corresponding to a polypeptide chain-folding initiation site detected by tyrosine fluorescence lifetime and proton magnetic resonance measurements. Biochemistry 26:1672–1683, 1987.
17. Reed, J., Kinzel, V., Cheng, H.-C., Walsh, D.A. Circular dichroic investigations of secondary structure in synthetic peptide inhibitors of cAMP-dependent protein kinase: A model for inhibitory potential. Biochemistry 26:7641–7647, 1987.
18. Reed, J., Hull, W.E., von der Lieth, C.-W., Kübler, D., Suhai, S., Kinzel, V. Secondary structure of the Arg-Gly-Asp recognition site in proteins involved in cell surface adhesion. Evidence for the occurrence of nested beta-bends in the model hexapeptide GRGDSP. Eur. J. Biochem. 178:141–154, 1988.
19. Reed, J., deRopp, J.S., Trewhella, J., Glass, D.B., Liddle, W.K., Bradbury, E.M., Kinzel, V., Walsh, D.A. Conformational analysis of PKI(5-22)amide, the active inhibitory fragment of the inhibitor protein of the cyclic AMP-dependent protein kinase. Biochem. J. 264:371–380, 1989.
20. Miranker, A., Robinson, C.V., Radford, S.I., Aplin, R.T., Dobson, C.M. Detection of transient protein folding populations by mass spectrometry. Science 262:896–900, 1993.
21. Jennings, P.A., Wright, P.E. Formation of a molten globule intermediate early in the kinetic folding pathway of apomoglobin. Science 262:892–895, 1993.
22. Perczel, A., Foxman, N., Fasman, G. How reverse turns may mediate the formation of helical segments in proteins: An x-ray model. Proc. Natl. Acad. Sci. U.S.A. 89:8210–8214, 1992.
23. Chou, P.Y., Fasman, G.D. Empirical predictions of protein conformation. Annu. Rev. Biochem. 47:251–276, 1978.
24. Merrifield, R.B. Solid phase peptide synthesis I. The

synthesis of a tetrapeptide. J. Am. Chem. Soc. 85:2149–2154, 1963.
25. Carpino, L.A., Han, G.Y. The 9-fluorenylmethoxycarbonyl aminoprotecting group. J. Org. Chem. 37:3404–3409, 1972.
26. Aue, W.P., Bartholdy, E., Ernst, R.R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. J. Chem. Phys. 64:2229–2249, 1976.
27. Kumar, A., Ernst, R.R., Wüthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem. Biophys. Res. Commun. 95:1–6, 1980.
28. Bax, A., Davies, D. Practial aspects of two-dimensional transerve NOE spectroscopy. J. Magn. Reson. 65:355–360, 1985.
29. Marion, D., Wüthrich, K. Application of phase-sensitive two-dimensional correlated spectroscopy (COSY) for measurement of proton-proton spin-spin coupling constants. Biochem. Biophys. Res. Commun. 113:967–974, 1983.
30. Herrmann, F. NDEE—ein Programmsystem zur Auswertung mehrdimensionaler NMR-Spektren. In: "GWDG Forschung und wissenschaftliches Rechnen: Beiträge zum Heinz-Billing Preis 1994." Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen mbH, Vol. XL. Plesser, T., Wittenburg, P. (eds.). Göttingen, Germany: 1995:147–158.
31. Clore, G.M., Gronenborn, A.M., Nilges, M., Ryan, C.A. Three-dimensional structure of potato carboxypeptidase inhibitor in solution. A study using nuclear magnetic resonance, distance geometry, and restrained molecular dynamics. Biochemistry 26:8012–8023, 1987.
32. Wüthrich, K., Billeter, M., Braun, W. Pseudo-structures for the 20 common amino acids for the use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. J. Mol. Biol. 169:949–961, 1983.
33. Pardi, A., Billeter, M., Wüthrich, K. Calibration of the angular dependence of the amide proton-Cα proton coupling constants, $^3J_{HN\alpha}$ in a globular protein. J. Mol. Biol. 180:741–751, 1984.
34. Brünger, A. "X-PLOR, 3.1 Manual." New Haven, CT: Yale University Press, 1993.
35. Nilges, M., Clore, G.M., Gronenborn, A.M. Determination of three dimensional structure from interproton distance data by dynamical simulated annealing from a random array of atoms. FEBS Lett. 239:129–139, 1988.
36. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimalization, and dynamics calculations. J. Comput. Chem. 4:187–217, 1983.
37. Bayer, P., Kraft, M., Ejchart, A., Westendorp, M., Frank, R., Rösch, P. Structural studies of HIV-1 Tat protein. J. Mol. Biol. 247:529–535, 1995.
38. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
39. Hutchinson, E.G., Thornton, J.M. PROMOTIF—A program to identify and analyze structural motifs in proteins. Protein Sci 5:212–220, 1996.
40. Wüthrich, K. Sequential ^1H-^1H distances. In: "NMR of Proteins and Nucleic Acids." New York: Wiley, 1986:122–125.
41. Schulz, G.E., Schirmer, R.H. Structural implications of the peptide bond. In: "Principles of Protein Structure." New York: Springer-Verlag, 1979:17–26.
42. Kartha, G., Ashida, T., Kakudo, M. The crystal structure of t-amyloxycarbonyl-L-prolyl-L-prolyl-L -prolin. Acta Crystallogr. B 30:1861–1866, 1974.
43. Brandts, J.F., Halvorson, H.R., Brennan, M. Consideration of the stability that the slow step in protein denaturation is due to cis-trans isomerism of proline residues. Biochemistry 14:4953–4963, 1975.
44. Dyson, H.J., Rance, M., Houghton, R.A., Lerner, R.A., Wreight, P.E. Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements of the formation of a reverse turn. J. Mol. Biol. 201:161–200, 1988.
45. Schmidt, F.X. Propyl isomerase: Enzymatic catalysis of slow protein folding reactions. Annu. Rev. Biophys. Biomol. Struct. 22:123–143, 1993.
46. Grathwohl, C., Wüthrich, K. The X-Pro peptide bond as an NMR probe for conformational studies of flexible linear peptides. Biopolymers 15:2025–2041, 1976.
47. Grathwohl, C., Wüthrich, K. NMR studies of the molecular conformations in the linear oligopeptides H-(L-Ala)$_n$-L-Pro-OH. Biopolymers 15:2043–2057, 1976.
48. Grathwohl, C., Wüthrich, K. NMR studies of the rates of proline cis-trans isomerization in oligopeptides. Biopolymers 20:2623–2633, 1981.
49. Richardson, J.S., Richardson, D.C. Amino acid preferences for specific locations at the end of α helices. Science 240:1648–1652, 1988.
50. Stewart, D.E., Sarkar, A., Wampler, J.E. Occurrence and role of cis peptide bonds in protein structures. J. Mol. Biol. 214:253–260, 1990.
51. Frömmel, C., Preissner, R. Prediction of prolyl residues in cis-conformation in protein structures on the basis of the amino acid sequence. FEBS 277:159–163, 1990.
52. Richardson, J.S., Richardson, D.C. Principles and patterns of protein conformation. In: "Prediction of Protein Structure and the Principles of Protein Conformation." Fasman, G.D. (ed.). New York: Plenum Press, 1990:1–98.
53. Wishart, D.S., Sykes, B.D., Richards, F.M. The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31:1647–1651, 1992.
54. Merutka, G., Dyson, H.J., Wright, P.E. "Random coil" chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. J. Biomol. NMR 5:14–24, 1995.
55. Némethy, G., Scheraga, H.A. Stereochemical requirements for the existence of the hydrogen bonds in β-bends. Biochem. Biophys. Res. Commun. 95:320–327, 1980.
56. Rose, G.D., Gierasch, L.M., Smith, J.A. Turns in peptide and proteins. Adv. Protein Chem. 37:1–109, 1985.
57. Scheraga, H.A. On the dominance of short-range interactions in polypeptides and proteins. Pure Appl. Chem. 36:1–8, 1973.
58. Zimmermann, S.S., Scheraga, H.A. Influence of local interaction on protein structure. I. Conformational energy studies of N-acetyl-N′-methylamides of Pro-X and X-Pro dipeptides. Biopolymers 16:811–843, 1977.
59. Venkatachalapathi, Y.V., Venkataram Prasad, B.V., Balaram, P. Conformational analysis of small disulfide loops. Spectroscopic and theoretical studies on a synthetic cyclic tetrapeptide containing cysteine. Biochemistry 21:5502–5509, 1982.
60. Rao, B.N.N., Kumar, A., Balaram, H., Ravi, A., Balaram, P. Nuclear Overhauser effect and circular dichroism as probes of β-turn conformations in acyclic and cyclic peptides with Pro-X sequence. J. Am. Chem. Soc. 105:7423–7428, 1983.
61. Mac Arthur, M.W., Thornton, J.M. Influence of proline residues on protein conformation. J. Mol. Biol. 218:397–412, 1991.