Vibrational Analysis of Conformation in Peptides, Polypeptides, and Proteins*

S. KRIMM, *Biophysics Research Division, University of Michigan, Ann Arbor, Michigan 48109*

Synopsis

A vibrational force field for the polypeptide chain has been developed for normal-mode analysis of such molecules. It can reproduce observed frequencies of known structures to within about 5 cm^{-1}. We review the application of this technique to conformational problems in peptides (β-turns and their model compounds), polypeptides [the α_{II}-helix and crystalline poly(glycine II)], and proteins (bacteriorhodopsin and glucagon).

INTRODUCTION

Vibrational spectroscopy (i.e., ir and Raman) has a long history of application to structural studies on polypeptide molecules, but for most of this period the full potential of this technique has not been utilized. This is because interpretations of experimental results have been based mainly on group frequency correlations rather than on more powerful theoretical approaches to the analysis of the spectra. In recent years, as a result of the development of high-speed computers and the refinement of highly predictive vibrational force fields for polypeptides, it has become possible to do meaningful normal-mode analyses on such molecules and thus to make predictions of the vibrational frequencies associated with a proposed structure. As a result of the detailed comparisons that are now possible between observed and calculated frequencies, much stronger structural conclusions can be reached.

In this paper I review some recent work on the refinement of a force field for polypeptides and give examples of its application to the study of conformation in peptides, polypeptides, and proteins.

VIBRATIONAL FORCE FIELD FOR POLYPEPTIDES

Force fields are usually refined by parametrizing a vibrational potential function (in our case, a generalized valence force field) and determining the values of the (force) constants in it that give the best fit to the observed frequencies of relevant molecules of known structure (usually together with their isotopic derivatives). Our studies began with *N*-methylacetamide,[1]

* This is paper number 16 in a series of which Ref. 20 is paper 15.

Biopolymers, Vol. 22, 217–225 (1983)
© 1983 John Wiley & Sons, Inc.
CCC 0006-3525/83/010217-09$01.90

and this force field was then transferred to poly(glycine I)[2] and poly(glycine II).[3] These results formed the basis for a calculation of poly(glycine I) in the revised antiparallel-chain rippled-sheet structure,[4] β-poly(L-alanine) in the antiparallel-chain pleated-sheet structure,[5] and α-helical poly(L-alanine).[6] During this stage of the refinement process, we discovered that transition dipole coupling must be included in the force field if observed amide I and amide II splittings are to be satisfactorily explained.[7,8]

In our most recent work the force fields for poly(glycine I)[9] and β-poly(L-alanine)[10] have been "fine tuned" to give maximum transferability between these molecules and optimum agreement with the spectra of isotopic species. For the parent molecules it is now possible to reproduce observed frequencies to within an average error of about 5 cm^{-1}, an example of which, for β-poly(L-alanine), is shown in Fig. 1. This force field has also shown good transferability to α-poly(L-alanine) (A. M. Dwivedi and S. Krimm, to be published) and has enabled us to show that the ND stretching fundamental in β-poly(L-alanine) is derivable from a three-level Fermi resonance analysis.[11] By using such a force field it should now be possible to calculate normal vibration frequencies with sufficient accuracy to permit in-depth analysis of conformation from spectra. Some applications of this approach to the study of peptides, polypeptides, and proteins are illustrated in what follows.

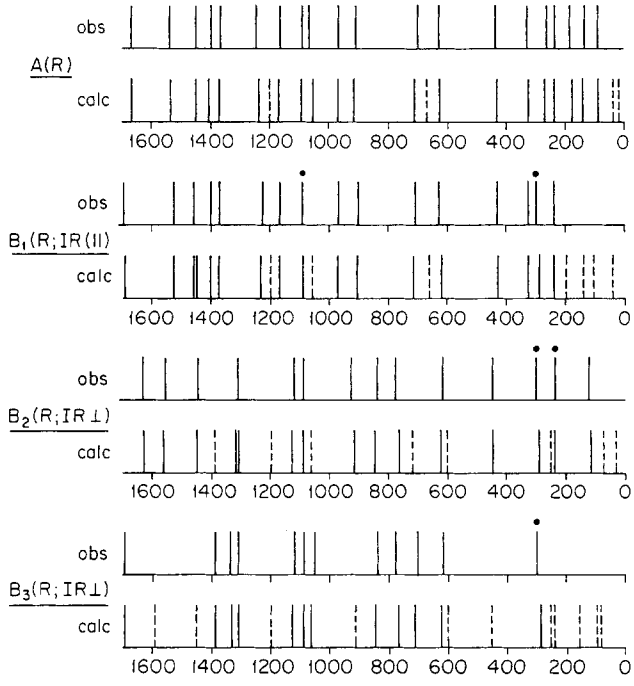

Fig. 1. Observed and calculated frequencies of β-poly(L-alanine).

TABLE I
Calculated Amide Mode Frequencies (in cm^{-1}) for Various Structures

Mode	Structure				
	α[a]	β[b]	β_I	β_{II}	β_{III}
Amide I		[1701][c]			
		1695	1690		1686
	[1663]	1669		1666	
	1659			~1656	
	[1650]	1630	~1640		1646 ± 3
Amide II		[1587]	1575		
	[1546]	1555	1558	1558	1562
	1543			1547	1551
		1534	1536	1540	1539
	1520	1523			
Amide III			1324	1329	1317
			1305	1303	1303
	1287		1299	1297	1291
	1270				
		1247			
		1228			
		1225			
Amide V	[639]	[709]	595	644	671
	628	713	575	607	589
	609	702	574	594	577
	[559]	[699]	570	588	573
				572	573
				571	

[a] From Ref. 6.
[b] From Ref. 5.
[c] Not observed.

PEPTIDES

A common constituent structure in globular proteins is the β-turn, the region in which a polypeptide chain reverses direction by ~180°. It is therefore important to be able to characterize this structure spectroscopically, a task made difficult by the range of conformations involved[12] and the paucity of model compounds available. We have approached this problem by doing normal mode calculations on canonical β-turns and by analyzing the spectra and conformations of peptides known to have β-turn structures.

Our calculated amide mode frequencies of standard β-turns[13,14] (Table I) show two interesting features when compared to frequencies of α-helical and β-sheet structures. First, some of the modes are characteristic of the particular β-turn, thus providing a possible means of identifying these structures. Second, there are β-turn frequencies that occur in the same

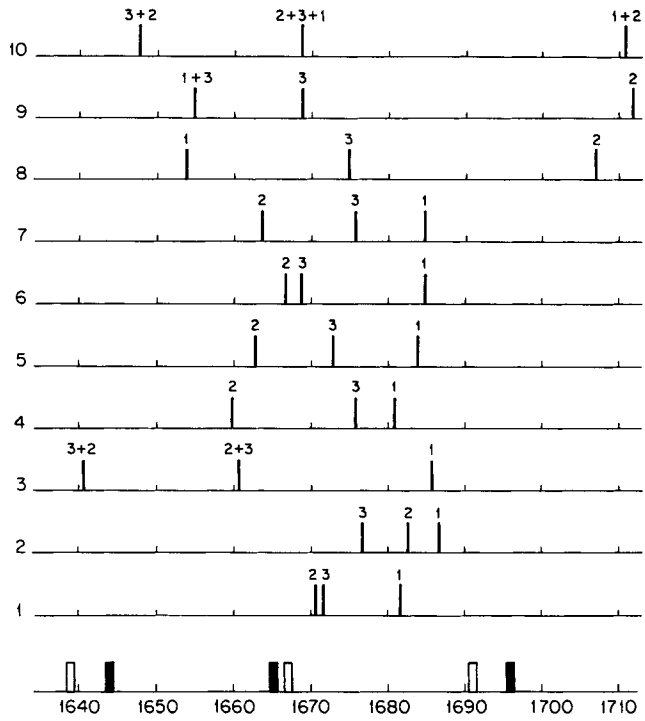

Fig. 2. Calculated frequencies in the amide I region of the 10 lowest energy conformations of *cyclo*(L-alanylglycyl-ϵ-aminocaproyl). Observed ir (solid bars) and Raman (open bars) bands are shown on the bottom line. Numbers above the calculated frequencies represent the peptide groups involved in the vibration: 1 refers to the group between Aca and Ala, 2 to the group between Ala and Gly, and 3 to the group between Gly and Aca.

region as α-helix and/or β-sheet modes, showing that caution may be necessary in assigning bands of globular proteins exclusively to the latter conformations.

These conclusions are given support by the ability of our force field to reproduce observed frequencies of known β-turn structures. The tetrapeptide Cbz-Gly-L-Pro-L-Leu-Gly crystallizes in a type I β-turn, and our calculated amide I and amide II frequencies are in excellent agreement with the observed ir and Raman bands.[15] We have computed the normal modes of the four β-turns of insulin,[16] and we predict frequencies that correspond well with hitherto observed but unassignable Raman bands in the amide I and amide II regions.

We have used this predictive capability to determine, from ir and Raman spectra of a cyclic dipeptide model for a β-turn, which of a number of energetically feasible conformations is present in the crystal. Conformational energy calculations on this molecule, *cyclo*(L-alanylglycyl-ϵ-aminocaproyl), show that it has four different conformations within about 1 kcal/mol of the minimum-energy structure.[17] We have calculated the normal modes of the 10 lowest energy conformations,[18] and we show in Fig. 2 the predicted

TABLE II
Observed and Calculated Frequencies (in cm^{-1}) for Some of the Modes of Antiparallel and Parallel Chain Structures of Crystalline Poly(Glycine II)

Observed		Calculated				
		Antiparallel		Parallel		
Raman	Infrared	A	B	A	E	Mode[a]
3279 w	3279 s	3281	3281	3281	3281	NH s
3257 m		3254	3254	—[b]	—	NH s
2979 s	2977 w	2980	2980	2981	2981	CH_2 as + ss
2940 vs	2935 mw	2936	2936	—	—	CH_2 as
	2850 mw	2853	2853	—	—	CH_2 ss
	2805 w	2803	2803	2803	2803	CH_2 ss + as
	1432 w	1433	1433	1435	1431	CH_2 b
1421 s	1420 m	1422	1423	—	—	CH_2 b
1383 ms	1377 m	1383	1380	1374	1374	CH_2 w
1334 vw	1332 vw	1344	1345	*[c]	*	CH_2 w
952 w		958			*	CH_2 r
864 w	862 vw	872	871	*	*	CH_2 r
673 m		674	678	*	*	CN t, NH ob
566 m		566		*	*	CO ob

[a] s = stretch, as = antisymmetric stretch, ss = symmetric stretch, b = bend, w = wag, r = rock, t = torsion, ob = out-of-plane bend.

[b] No calculated modes.

[c] Modes calculated at very different frequencies.

amide I modes of these structures together with the observed ir and Raman bands. It can be seen that certain conformations (e.g., 1, 2, 4–7) give very poor agreement with the observed frequencies; for others (8–10), the agreement is certainly questionable; and one (3) gives fairly good agreement. Examination of the other amide modes,[18] in particular amide V, shows convincingly that conformation 3 is the structure most likely to be present. Similar analyses permitted us to determine the likely conformations of *cyclo*(L-alanyl-L-alanyl-ϵ-aminocaproyl) and *cyclo*(L-alanyl-D-alanyl-ϵ-aminocaproyl).[19] These studies show that frequencies are often more sensitive to conformational structure than energies, and thus that normal-mode analysis can be a powerful technique in studying peptide conformation.

POLYPEPTIDES

Our force-field refinement was based on normal-mode analyses of poly(glycine I)[4,9] and β-poly(L-alanine)[5,10]; it showed clearly that the former has a rippled-sheet structure, whereas the latter has a pleated-sheet structure. This force field provided a basis for analyzing the spectrum of α-helical poly(L-alanine) (Ref. 6; and Dwivedi and Krimm, to be published),

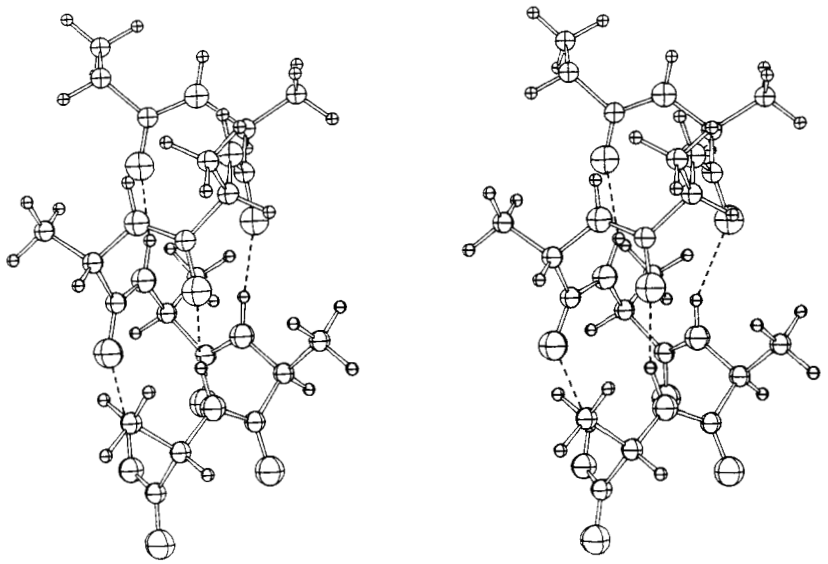

Fig. 3. Poly(L-alanine) in the α_I-helix and α_{II}-helix conformations.

in both the standard α_I as well as modified α_{II} forms (see below). We have recently completed a study of poly(glycine II) that reveals the nature of chain packing in the crystal and confirms the presence in this structure of interchain C—H···O=C hydrogen bonds.[20]

Poly(glycine II) is a 3_1-helix that can presumably pack in parallel[21] or antiparallel[22,23] chain arrangements, in both of which interchain C—H···O=C hydrogen bonds can form.[21,23] Spectroscopic studies[24,25] had suggested that such bonds are present, but stronger evidence is desirable. We have now calculated the normal modes of the full crystal structure,[20] abandoning the point-mass approximation for the CH_2 group[3] and incorporating the possibility of C—H···O=C and related bifurcated N—H···O hydrogen bonds. The results for some of these modes are shown in Table II.

It will be seen that some observed bands are predicted by the antiparallel- but not by the parallel-chain structure, whereas in other cases the frequency agreement for the antiparallel-chain structure is substantially better than for the parallel-chain structure. This provides convincing evidence that the former structure is preferred in crystalline poly(glycine II). The additional bands arise from the fact that strict 3_1-helix symmetry is broken by the proposed presence in the antiparallel-chain structure of a C—H···O=C hydrogen bond in every third residue.[23] The spectroscopic results therefore show that such nonequivalence of residues indeed exists and, from the values of the CH_2 frequencies, that C—H···O=C hydrogen bonds are responsible.

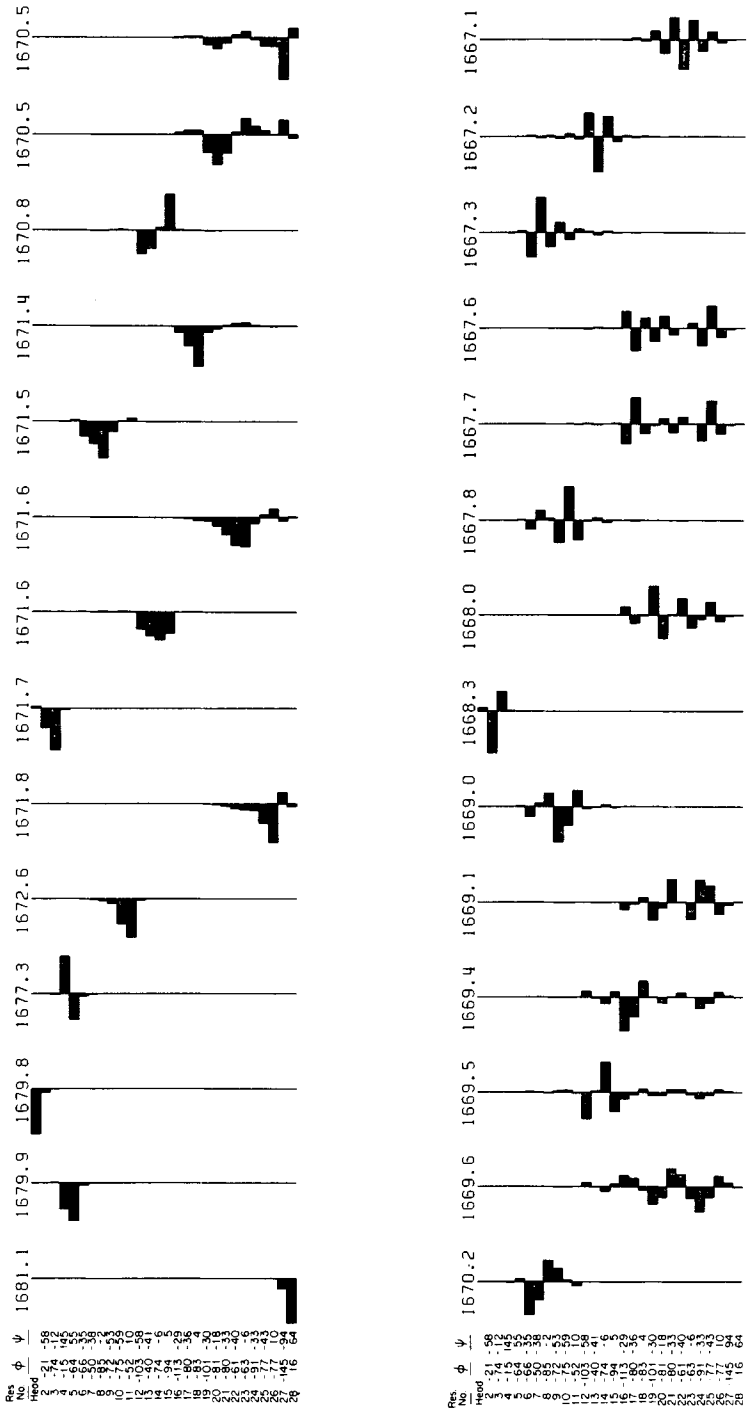

Fig. 4. C=O stretching coordinates (contraction to left, extension to right) in amide I modes of glucagon.

PROTEINS

The ir spectrum of bacteriorhodopsin is generally consistent with the α-helix structures suggested by electron diffraction,[26] but some bands in the spectrum have unusual frequencies.[27] In searching for a possible reason for this, we calculated the frequencies of the α_{II}-helix, a helix with the same n and h as the standard α_I-helix but having slightly different φ,ψ.[28] These helices, with $\varphi = -57.4°$, $\psi = -47.5°$ for α_I and $\varphi = -70.5°$C, $\psi = -35.8°$ for α_{II}, are compared in Fig. 3. Our calculations (Ref 29; and Dwivedi and Krimm, to be published) show that four observed anomalous features of the bacteriorhodopsin spectrum can be accounted for by the presence of α_{II}-helices. In such helices (see Fig. 3), the planes of the peptide groups are tilted such that the N—H bond points inward toward the axis, thus decreasing the distance between NH hydrogens so that they can be in contact with one another. This led us to the speculation[29] that the helical array of NH hydrogen atoms, rather than the previously proposed side-chain pathway, may constitute the framework through which proton transport takes place. Although this is a novel suggestion, it is given a sound structural basis by the detailed spectroscopic analysis.

The ability to calculate the normal modes of globular proteins would open up exciting possibilities for studying the conformations of these molecules. Such a computation has now been done on glucagon.[30] Although an "approximate" force field was used (Dwivedi and Krimm, to be published), in which the side chain is taken as a point mass, the calculation should yield relevant information on amide and skeletal vibrations. The ~600 normal modes have yet to be analyzed, but as an illustration of some of the results, we show in Fig. 4 the distribution of one internal coordinate (C=O stretching) in the amide I modes[30]: in some modes the vibration is localized in one or two peptide groups, whereas in others it is delocalized over (generally α-helical) segments of the polypeptide chain. Similar effects can be seen in other modes. Since glucagon has quite different Raman spectra in crystal, gel, and solution,[31] it should be possible with the help of such calculations to do detailed studies of its conformational states as a function of environment.

This research was supported by National Science Foundation Grants PCM-7921652 and DMR-7800753.

References

1. Jakeš, J. & Krimm, S. (1971) *Spectrochim. Acta* **27A,** 19–34.
2. Abe, Y. & Krimm, S. (1972) *Biopolymers* **11,** 1817–1839.
3. Abe, Y. & Krimm, S. (1972) *Biopolymers* **11,** 1841–1853.
4. Moore, W. H. & Krimm, S. (1976) *Biopolymers* **15,** 2439–2464.
5. Moore, W. H. & Krimm, S. (1976) *Biopolymers* **15,** 2465–2483.
6. Rabolt, J. F., Moore, W. H. & Krimm, S. (1977) *Macromolecules* **10,** 1065–1074.
7. Krimm, S. & Abe, Y. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 2788–2792.
8. Moore, W. H. & Krimm, S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4933–4935.

9. Dwivedi, A. M. & Krimm, S. (1982) *Macromolecules* **15,** 177–185.
10. Dwivedi, A. M. & Krimm, S. (1982) *Macromolecules* **15,** 186–193.
11. Krimm, S. & Dwivedi, A. M. (1982) *J. Raman Spectrosc.* **12,** 133–137.
12. Lewis, P. N., Momany, F. A. & Scheraga, H. A. (1973) *Biochim, Biophys. Acta* **303,** 211–229.
13. Bandekar, J. & Krimm, S. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 774–777.
14. Krimm, S. & Bandekar, J. (1980) *Biopolymers* **19,** 1–29.
15. Bandekar, J. & Krimm, S. (1979) in *Peptides: Structure and Biological Function. Proceedings of the Sixth American Peptide Symposium*, Gross E. & Meienhofer, J., Eds., Pierce Chemical Co., Rockford, Ill., pp. 241–244.
16. Bandekar, J. & Krimm, S. (1980) *Biopolymers* **19,** 31–36.
17. Némethy, G., McQuie, J. R., Pottle, M. S. & Scheraga, H. A. (1981) *Macromolecules* **14,** 975–985.
18. Maxfield, F. R., Bandekar, J., Krimm, S., Evans, D. J., Leach, S. J., Némethy, G. & Scheraga, H. A. (1981) *Macromolecules* **14,** 997–1003.
19. Bandekar, J., Evans, D. J., Krimm, S., Leach, S. J., Lee, S., McQuie, J. R., Minasian, E., Némethy, G., Pottle, M. S., Scheraga, H. A., Stimson, E. R. & Woody, R. W. (1982) *Int. J. Pept. Protein Res.* **19,** 187–205.
20. Dwivedi, A. M. & Krimm, S. (1982) *Biopolymers* **21,** 2377–2397.
21. Ramachandran, G. N., Sasisekharan, V. & Ramakrishnan, C. (1966) *Biochim. Biophys. Acta* **112,** 168–170.
22. Krimm, S. (1966) *Nature* **212,** 1482–1483.
23. Ramachandran, G. N., Ramakrishnan, C. & Venkatachalam, C. M. (1967) in *Conformation of Biopolymers*, Ramachandran, G. N., Ed., Academic Press, New York, pp. 429–438.
24. Krimm, S., Kuroiwa, K. & Rebane, T. (1967) in *Conformation of Biopolymers*, Ramachandran, G. N., Ed., Academic Press, New York, pp. 439–447.
25. Krimm, S. & Kuroiwa, K. (1968) *Biopolymers* **6,** 401–407.
26. Henderson, R. & Unwin, P. N. T. (1975) *Nature [New Biol.]* **257,** 28–32.
27. Rothschild, K. J. & Clark, N. A. (1979) *Science* **204,** 311–312.
28. Miyazawa, T. (1961) *J. Polym. Sci.* **55,** 215–231.
29. Krimm, S. & Dwivedi, A. M. (1982) *Science* **216,** 407–408.
30. Tasumi, M., Takeuchi, H., Ataka, S., Dwivedi, A. M. & Krimm, S. (1982) *Biopolymers* **21,** 711–714.
31. Yu, N.-T. & Liu, C. S. (1972) *J. Am. Chem. Soc.* **94,** 5127–5128.

Received June 20, 1982
Accepted September 2, 1982