Conformation Dependence of the $C^{\alpha}D^{\alpha}$ Stretch Mode in Peptides. II. Explicitly Hydrated Alanine Peptide Structures

Noemi G. Mirkin, Samuel Krimm
LSA Biophysics, University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109-1055

Received 2 March 2009; revised 29 April 2009; accepted 29 April 2009
Published online 7 May 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.21250

ABSTRACT:

Our previous studies of the potential utility of the $C^{\alpha}D^{\alpha}$ stretch frequency, ν(CD), as a tool for determining conformation in peptide systems (Mirkin and Krimm, J Phys Chem A 2004, 108, 10923–10924; 2007, 111, 5300–5303) dealt with the spectroscopic characteristics of isolated alanine peptides with α_R, β, and polyproline II structures. We have now extended these ab initio calculations to include various explicit-water environments interacting with such conformers. We find that the structure-discriminating feature of this technique is in fact enhanced as a result of the conformation-specific interactions of the bonding waters, in part due to our finding (Mirkin and Krimm, J Phys Chem B 2008, 112, 15268) that $C^{\alpha}—D^{\alpha}\cdots O$(water) hydrogen bonds can be present in addition to those expected between water and the CO and NH of the peptide groups. In fact, ν(CD) is hardly affected by the latter bonding but can be shifted by up to 70 cm^{-1} by the former hydrogen bonds. We also discuss the factors that will have to be considered in developing the molecular dynamics (MD) treatment needed to satisfactorily take account of the influence of outer water layers on the structure of the first-layer water molecules that hydrogen bond to the peptide backbone.

Correspondence to: Samuel Krimm; e-mail: skrimm@umich.edu
Contract grant sponsor: National Science Foundation
Contract grant number: CHE-0517905

Keywords: hydrated peptides; vibrational spectroscopy; conformation; CD stretch

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

INTRODUCTION

Vibrational spectroscopy has been an important tool for studying the conformations of peptide chains.[1] Until now, the application of this technique has concentrated on the main normal mode frequencies of the peptide group, such as the amide I mode (mainly C=O stretch) (1600–1700 cm^{-1}), which is somewhat dependent on the φ(CNC$^{\alpha}$C) and ψ(NC$^{\alpha}$CN) dihedral angles about the C$^{\alpha}$ atom but is particularly sensitive to the φ,ψ-dependent frequency splitting resulting from a transition dipole coupling interaction between peptide groups.[2–5] Although this general peptide-group region of the spectrum (1200–1700 cm^{-1}) has been very useful, it is important to continue uncovering the structural potential of other normal modes that could provide conformational information.

Toward this goal, we have pointed out[6] that the frequency of the C$^{\alpha}$D$^{\alpha}$ stretch mode, ν(CD), could be a tool for the determination of peptide conformation. It is a highly localized vibration in a clear region of the fundamental spectrum, ~2300 cm^{-1}, both of the peptide and of water, which makes it particularly adaptable to studies of peptide systems in aqueous environments. Our initial note[6] dealt mainly with ab initio calculations of the isolated alanine dipeptide (ADP), $[CH_3CONH]_1C^{\alpha}D^{\alpha}(CH_3)[CONHCH_3]_2$, designated Ala-1, in the (most likely) α_R, β, and polyproline II (P)[7,8] conformations. A subsequent work[9] reported similar studies

of ν(CD) in all the intervening isolated peptides through Ala-7 in these three uniform (and some mixed) φ,ψ conformational states. The protocol for an experimental structural analysis was envisioned as follows. For a particular Ala-*n*, *n* samples would be synthesized, the first with D^{α} substituted only on residue 1, the IR (or Raman) spectrum providing $\nu(CD_1)$ (or its overtone, if possible, to enhance frequency separations). The second sample would have D^{α} substituted only on residue 2, providing $\nu(CD_2)$; and so on. This set of observed bands, together with the calculated ν(CD)s for each conformation, would be the basis for determining the structural state of the peptide. The results of our analysis[9] indeed indicated that such discrimination was possible for uniform structures on the basis of at least three spectral properties, the value of $\nu(CD_1)$, the values of the successive $\Delta\nu_{ij}$, and the value of $\nu_n - \nu_1$. Although these studies provided a "proof of principle" of the methodology, the restriction to the isolated molecules limited the analyses to nonaqueous environments.

Because a special goal of this technique is to enable aqueous solution studies of peptide conformation, we have extended preliminary studies[6] and undertaken comparable ab initio analyses of the above three ordered peptide conformations in the presence of interacting water molecules. In particular, the aim was to determine the effects of hydrogen bonding by water molecules on the values of ν(CD) and how this interaction supplements the local conformation dependence in the isolated molecule that determines the C^{α}—D^{α} bond length, r(CD), and thus the value of ν(CD).[9] Because of the extensive nature of such calculations, we limited these studies to the Ala-1 through Ala-4 peptides, but we believe that the basic effects are evident from this set. We also consider how to extend such analyses to longer chains, both in terms of the peptide length and the treatment of the interacting water molecules. In practical terms, of course, studies will be needed to evaluate many experimental issues, such as solution concentration requirements, relative background intensity, and combination band contributions.[10]

Calculations

As in our previous studies,[6,9] the ab initio calculations of frequencies were done with the B3LYP functional and the (equilibrium planar-peptide-predicting[11]) 6-31 + G* basis set, using Gaussian 03.[12] (Test MP2 calculations, as well as trials with the B3PW91 functional, give qualitatively the same results, thus justifying our use of the B3LYP protocol to reveal the underlying physical bases of the interactions.) The geometries were fully optimized within the φ,ψ dihedral angle constraints of the three local conformations being studied. (Such constraints, which could affect lower frequencies, are not expected to influence high-frequency relatively localized stretching vibrations.[13]) Two sets of angles were used: set A comprised typical φ,ψ values, used in our prior studies of the isolated peptides,[6,9] viz., $\alpha_R(-60^\circ, -40^\circ)$, $\beta(-134^\circ, 145^\circ)$, and $P(-75^\circ, 145^\circ)$; set B, used in just a few cases to test the effect of small angle variations, was chosen from the results of a joint molecular dynamics (MD) and NMR study of trialanine,[14] viz., $\alpha_R(-80^\circ, -50^\circ)$, $\beta(-120^\circ, 130^\circ)$, and $P(-60^\circ, 140^\circ)$. (We note that the latter values may be no more compelling to our study, since all covalent bond lengths were constrained in the MD calculations.[14]) The quoted values of r(CD) are those of r(CH), but the values of ν(CD) are (unscaled) normal-mode frequencies from Wilson GF method[15] calculations in which H^{α} is replaced by D^{α}. As noted earlier,[9] although the frequencies are obtained on the fully C^{α}-deuterated molecules, these results are applicable to the individually deuterated species, because we have shown that there is no coupling between the localized $C^{\alpha}D^{\alpha}$ stretch modes on adjacent C^{α} sites in the isolated molecules and the bonded waters do not provide any additional pathways for this to occur.

To establish a reliable protocol for the ab initio calculations of the hydrated peptide structures, we examined a wide range of possible water-peptide models. It was quickly evident that by representing the hydration with only a continuum solvent, whether by the Onsager reaction field (not found to be satisfactory) or the polarizable continuum model (PCM), it was (as is well known[16,17]) impossible to reproduce the special effects of hydrogen bonding by explicit water molecules. With various specific water structures (to be described in detail below), we found that the addition of a PCM water continuum surrounding the hydrated system had some effect on ν(CD), and we therefore examined this added feature in extensive calculations of representative hydration numbers for the conformers. This started with application of the default PCM method, which replaces hydrogen-bearing heavy atoms by united atoms. However, it became evident that it is necessary to treat the H^{α} atom individually, PCM(H^{α}), in order to reliably determine changes in r(CD) (whereas adding individual N hydrogens has only a minor impact and the addition of explicit H atoms to the bonded waters was found to make no difference). Although some of these results are presented, we concentrate here on the ν(CD) of the minimally bound systems (to be explained below). This was not found to compromise the main conclusion of this investigation, namely that the methodology does in fact permit distinguishing between hydrated structures of the three conformations.

Table I $C^{\alpha}D^{\alpha}$ Stretch Frequencies[a] of Isolated and Various Hydrated β-Conformation Alanine Dipeptide Structures

	Water Structures[b]							
	CO_1a	CO_1b	$[NH]_1$	$[CO]_2$	$[NH]_2$	r(CD)[c]	ν(CD)[d]	r(D···O)[e]
Isolated						1.0948	2276	
1	1[f]					1.0944	2281	
2			1			1.0945	2279	
3	1		1			1.0939	2284	
4				1		1.0947	2278	
5			1	1		1.0943	2281	
6		1			(1)[g]	1.0876	2348	2.22
7		1		1	(1)[g]	1.0873	2350	2.20
8	1	1				1.0872	2352	2.23
9		1	1[h]	1	1[i]	1.0891	2330	2.77, 2.69
						1.0892[j]	2330[j]	2.76, 2.77
10	1	1	1	2	1	1.0889	2331	2.75, 2.77
						1.0890[j]	2331[j]	2.72, 2.80
11	1[k]	1[k]	1[k]	1[k]	2[k]	1.0906	2313	3.17, 2.92

[a] B3LYP/6-31 + G* calculation; set A $\varphi, \psi = -134^{\circ}, 145^{\circ}$.
[b] With reference to $[CH_3CONH]_1C^{\alpha}D^{\alpha}(CH_3)[CONHCH_3]_2$. CO_1a refers to water *cis* to $[CH_3]_1$ across $[CO]_1$; CO_1b refers to water *cis* to $[NH]_1$ across $[CO]_1$.
[c] Length in Å.
[d] Frequency in cm^{-1}.
[e] Distances < 3.40 Å. First in pair is to CO_1b water, second to $[NH]_2\cdots O$ water.
[f] Water at CO_1a position.
[g] Water was initially placed hydrogen bonded to $[NH]_2$ but it then optimized to CO_1b position.
[h] Water on $[NH]_1$ also participates in hydrogen bond to $[CO]_2$.
[i] Water also hydrogen bonds to water on CO_1b.
[j] B3LYP/6-31 + G* PCM(H^{α}) calculation (see text).
[k] Three additional water molecules hydrogen bond to these six peptide-interacting waters.

RESULTS AND DISCUSSION

In discussing the results of the ab initio calculations, we first concentrate in detail on the behavior of the ADP–water systems, because these results provide a framework in which the properties of the longer peptides can be more easily assessed. There have been a number of quantum-mechanical (QM) studies of the structure[16–21] and dynamics[22] of ADP/explicit-water systems, but none has included an investigation of the possible specific C^{α}-H^{α}/water interactions that we have analyzed.

Alanine Dipeptide

Our calculations show that ν(CD) is affected to a minor extent by hydrogen bonding of water to the CO and NH of the peptide group but much more so by the specific water structure at the C^{α}—D^{α} bond, with its resulting development of C^{α}—$D^{\alpha}\cdots O$ hydrogen-bonding interactions.[23] We illustrate the nature of these separate effects by first discussing systematic studies of different water arrangements in the β-conformation (set A φ,ψ) of the ADP. In all these arrangements, the peptide and water systems were otherwise fully optimized.

In Table I, we present r(CD), ν(CD), and r(D···O) for 11 ADP–water structures (plus data on the isolated ADP). A number of conclusions emerge from these results.

1. Hydrogen bonding of a water molecule to the CO or NH of a peptide group in such a way that there is no possibility of close interaction with D^{α} produces no major change in ν(CD) from that of the isolated molecule [for which r(CD) = 1.0948 Å and ν(CD) = 2276 cm^{-1}]. This is seen in structure **1** (see Figure 1), in which the water occupies a position on the side of $[CO]_1$ *cis* to $[CH_3]_1$, referred to as CO_1a and involving one of the lone-pair orbitals on this O, with r(CD) = 1.0944 Å and ν(CD) = 2281 cm^{-1}. With waters hydrogen bonded to the CO_1a and $[NH]_1$ positions, as in structure **3**, there is a small decrease in r(CD) and therefore a small upward shift in ν(CD), to 2284 cm^{-1}, indicating a very small neighbor effect on the C^{α}—D^{α} bond resulting from the altered electronic properties of the adjacent fully hydrogen-bonded peptide group.

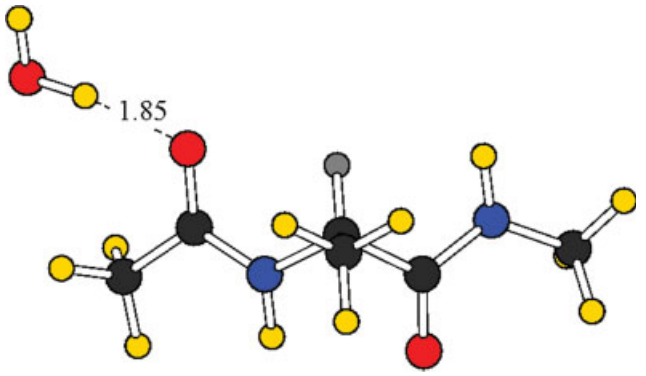

FIGURE 1 β-Conformation alanine dipeptide with water hydrogen bonded in the CO_1a position (structure **1** of Table I).

Including structures with waters bound to C^{α}-adjacent $[NH]_1$ and $[CO]_2$ groups, that is, structures **1–5**, such ν(CD) are found at 2281 ± 2 cm^{-1} [with r(CD) of 1.0944 ± 0.0002 Å].

2. When a close $D^{\alpha}\cdots O$ distance occurs (such as in structures **6**, **7**, and **8**), indicative of $D^{\alpha}\cdots O$ hydrogen-bond formation,[23] r(CD) decreases significantly and ν(CD) increases, in these cases to the range of 2348–2352 cm^{-1}. The nature of this interaction is seen in structure **6** (see Figure 2), where although the single water was initially placed in a hydrogen-bond position at $[NH]_2$ it optimized to a new hydrogen-bonding position at $[CO]_1$, referred to as the CO_1b position and involving the other lone-pair orbital on this O. This results in a close approach of the water O to D^{α}, at a distance of 2.22 Å (a value associated with an "extreme" van der Waals contact[24]) and with one of its lone pairs clearly oriented in the direction of D^{α}, a significant contraction of r(CD) to 1.0876 Å, and an increase in ν(CD) to 2348 cm^{-1}. (As already noted,[23] the interaction of this water O with $[NH]_2$ is very weak, and thus this cannot be considered a traditional N—H···O hydrogen bond: its $[N]_2H\cdots O$ distance is 2.79 Å, significantly longer than the ~1.9 Å of such hydrogen bonds found in the more fully hydrated structures.)

Except for the early infrared studies of polyglycine II,[25–28] it had not been thought that, from a spectroscopic view, the C—H···O hydrogen bond occurred in peptides, but it is now a well-established attractive interaction,[29] being recognized as occurring in (nonzwiterionic) amino acids[30] and indeed found in proteins.[31–33] The unusual bond contraction and stretch frequency increase on hydrogen bonding are now also understood to be initiated by the dominant contracting force on the bond generated by the antiparallel orientation of the C—H dipole derivative and the acceptor electric field[34–36] (in distinction to the lengthening force from the parallel orientation, with bond extension and frequency decrease, in traditional hydrogen bonds such as N—H···O), and more generally to depend on properties of the charge density derivatives of the donor.[37,38] The energy of the C^{α}—$H^{\alpha}\cdots O$ hydrogen bond in the amino acid alanine has been given as 2.10 kcal/mol.[30] Although the peptide value will depend on the $D^{\alpha}\cdots O$ distance, an estimate of its optimum β-ADP value can be obtained from the difference in energies of structures **1** and **6** and is found to be 2.31 kcal/mol.[23] This is probably a maximum value, because there is a small $[N]_2$ H···O attractive interaction,[23] but it is in good agreement with the amino acid value. Such an interaction energy is not trivial, especially if it is multiply present, as in peptides, and should be considered not only in the stabilization of protein structures[31,39] but also in the general interaction of proteins with their aqueous environment.[23]

Additional evidence that the interaction is basically of a hydrogen bond rather than of a steric nature comes from an examination of the changes in charge distribution between relevant structures. In Table II, we show the Mulliken charges (which should be qualitatively if not quantitatively indicative of the changes) of the isolated, CO_1a, and CO_1b structures. Shown in *me*, these figures denote the gain or loss of electrons on atoms or groups upon structural modification. From the isolated to the CO_1a structure, the changes in charge distribution in the $[CO]_1$ bond as well as in the terminal CH_3 group as a result of the water hydrogen bond are understandable, as well as its nonlocalization as reflected in the change on C^{α}. In comparing the CO_1b and CO_1a structures, we see a further loss of electrons

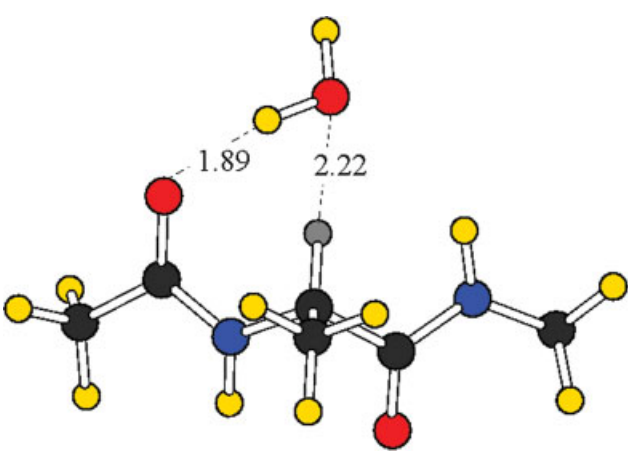

FIGURE 2 β-Conformation alanine dipeptide with water hydrogen bonded in the CO_1b position (structure **6** of Table I).

Table II Shifts in Mulliken Charges (in *me*) of *β*-Conformation Alanine Dipeptides from Isolated to CO_1a and from CO_1a to CO_1b Singly Hydrated States

Atom or Group	$I^a \rightarrow CO_1a^b$	$CO_1a^b \rightarrow CO_1b^c$
$[CH_3]_1$	23	2
C_1	−42	−60
O_1	90	−8
N_1	4	14
H_1	−6	1
C^{α}	−59	15
H^{α}	−1	−10
$(C^{\alpha})CH_3$	−1	106
C_2	7	−18
O_2	−3	0
N_2	−8	−32
H_2	0	−23
$[CH_3]_2$	−7	20

[a] Isolated ADP: see Table I footnote.
[b] Structure **1** of Table I (Figure 1).
[c] Structure **6** of Table I (Figure 2).

from C_1, a modest change on $C^{\alpha}H^{\alpha}$, a huge increase in electronic charge on the C^{α} methyl group, and a small change on the $[NH]_2$ group, consonant with the above-noted weak interaction with the water. Such significant and broadly distributed electronic redistributions in the ADP molecule are much more likely to be a consequence of the additional $H^{\alpha}\cdots O$(water) wave function overlap in a hydrogen bond than of a traditional van der Waals steric interaction.

3. The extent of a ν(CD) upshift is related to the change in r(CD), which is determined (in part) by the value of $r(D\cdots O)$ and therefore the strength of this interaction. Thus, the short $r(D\cdots O) = 2.20$ Å of structure **7,** with r(CD) = 1.0873 Å, results in ν(CD) = 2350 cm^{-1}, whereas the longer pair of $r(D\cdots O) = 2.92$ Å and 3.17 Å of structure **11**, with r(CD) = 1.0906 Å, results in ν(CD) = 2313 cm^{-1}. Because the latter $r(D\cdots O)$ are significantly longer than the $r(D\cdots O) = 2.4$ Å associated with "normal" van der Waals interactions,[24] the changes in ν(CD) cannot be associated primarily with steric effects and must be attributed predominantly to the effects of the electrical interactions with the two water molecules. If we associate hydrogen bonding with significant overlap of the H and O wave functions, the interaction of structure **11** is only borderline in this respect judging from the onset of overlap in the water dimer only at an $H\cdots O$ distance of less than ∼2.94 Å.[40] Structures **9** and **10**, with intermediate $D^{\alpha}\cdots O$ distances and ν(CD)s, would thus be described as having weak C^{α}—$D^{\alpha}\cdots O$ hydrogen bonds. It is of interest to note that the addition in structure **8** (see Figure 3) of a CO_1a water to the CO_1b water of structure **6**, both of which hydrogen bond to $[CO]_1$, produces only a small increase in ν(CD) (from 2348 to 2352 cm^{-1}). This probably results from the enhancement in hydrogen bonding of the two waters to $[CO]_1$, indicated by the increase in the r(CO) of 1.240 Å of structures **1** and **6** to r(CO) = 1.250 Å of structure **8**, similar to the effect of the change in ν(CD) from the isolated ADP, with r(CO) = 1.231 Å, to that of the CO_1a structure, that is, from 2276 to 2281 cm^{-1}.

4. The value of ν(CD) will depend not only on $r(D\cdots O)$ but also on the angle that the water molecule makes with the C^{α}—D^{α} bond, because this determines the component of the electric field along the bond (as well as the interaction of the O lone pair with D). This may account for the small difference between the ν(CD)s of structures **6** (2348 cm^{-1}) and **7** (2350 cm^{-1}), even though both share the same CO_1b bonding (although the influence of the hydrogen bonding of the second water to $[CO]_2$ cannot be discounted).
5. A comparison of ν(CD) of structure **6**, 2348 cm^{-1} for bonding at CO_1b, with that of structure **9**, 2330 cm^{-1} for similar bonding at CO_1b, also demonstrates the sensitivity of this frequency to the specific $D^{\alpha}\cdots O$ geometry: the additional hydrogen bond of the $[NH]_2$ water to the water bonded to $[CO]_1$ (see Figure 4) has resulted in the CO_1b water O moving away from D^{α} (2.77 Å vs 2.22 Å) and the $[NH]_2$ water O also developing a $D\cdots O$ interaction (2.69 Å), both longer distances resulting in a net weaker hydrogen-bonding effect and the lengthening of r(CD) (from 1.0876 to 1.0891 Å), and thus the above-mentioned decrease in frequency.

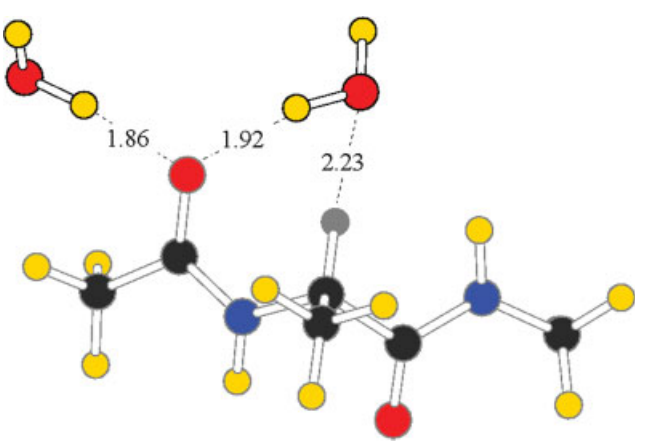

FIGURE 3 *β*-Conformation alanine dipeptide with waters hydrogen bonded in the CO_1a and CO_1b positions (structure **8** of Table I).

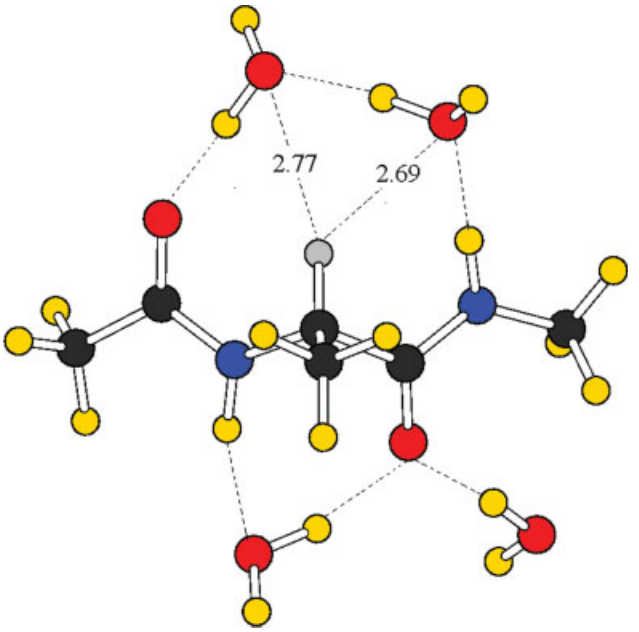

FIGURE 4 β-Conformation alanine dipeptide with four hydrogen-bonded waters (structure **9** of Table I).

6. The increase from four bonded waters in structure **9** to six in structure **10** has made essentially no change in the value of ν(CD) as a result of the similar values of r(CD) resulting from similar configurations of the waters (see Figure 5). However, the addition of three more waters in structure **11** (in perhaps only one of the possible configurations) has, without changing the general arrangement of peptide-bonding waters, increased the six-water $r(D\cdots O)$ distances from 2.75 to 3.17 Å and 2.77 to 2.92 Å with the resultant decrease in ν(CD) from 2331 cm^{-1} to 2313 cm^{-1}. This indicates that water molecules beyond the first-bonded layer can have an important effect on the first-layer water interaction with D^{α}. The best way of taking such variations with water environment into account will be through MD calculations in an enlarged explicit-water environment, to be considered in more detail below.
7. The addition of the PCM(H^{α}) solvent continuum to the calculations of the four-water (9) and six-water (10) structures has made no change in ν(CD), although this is not generally true for other conformations and varying water structures. This variability of the solvent continuum results should be minimized in the above-mentioned MD calculations.

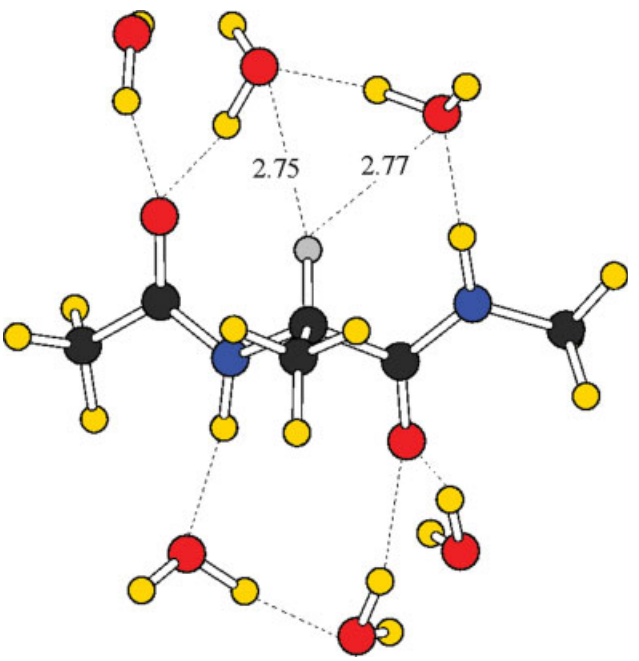

FIGURE 5 β-Conformation alanine dipeptide with six hydrogen-bonded waters (structure **10** of Table I).

The above results clearly demonstrate the importance of the specific structure of water around D^{α} in establishing the distinct C^{α}—$D^{\alpha}\cdots O$ interaction patterns that lead to the modified r(CD) and thus ν(CD). This was also found for the other ADP conformations, whose relative energies also depend on the water structure, as shown in Table III. This is seen particularly for the β and P structures, β being more stable than P by 0.3 kcal/mol for $ADP(H_2O)_4$, whereas the energy of β is significantly higher, by 1.0 kcal/mol, than that of P for $ADP(H_2O)_6$. The very much higher relative energies of the hydrated than isolated α_R structures emphasize the dominating dependence on water structure. This can be clearly seen by comparing the peptide-group frequencies of the various hydrated structures, given in Table IV: those of α_R are generally less shifted from the isolated values than is the

Table III Relative Energies (in kcal/mol) of Isolated and Hydrated Alpha-, Beta-, and Polyproline II-Conformation Alanine Dipeptide Structures

	Conformation		
Hydration State	α_R	β	P
ADP – set A[a]	3.6	0	1.2
$ADP(H_2O)_4$ – set A[a]	15.7	0	0.3
$ADP(H_2O)_4$ – set B[b]	16.4	0	1.8
$ADP(H_2O)_6$ – set A[a]	19.9	1.0	0
$ADP(H_2O)_6$ – set B[b]	20.2	0	0.3

[a] α_R: $\varphi, \psi = -60°, -40°$; β: $\varphi, \psi = -134°, 145°$; P: $\varphi, \psi = -75°, 145°$ (Ref. 9).

[b] α_R: $\varphi, \psi = -80°, -50°$; β: $\varphi, \psi = -120°, 130°$; P: $\varphi, \psi = -60°, 140°$ (Ref. 13).

Table IV Peptide Group Frequencies (in cm^{-1}) of Isolated and $(H_2O)_n$-Hydrated Alanine Dipeptide Conformations[a]

Structure	$[CO]_1$	$[NH]_1$	$[CO]_2$	$[NH]_2$
α_R	1757	3606	1767	3608
$\alpha_R(H_2O)_4$	1718	3571	1746	3528
$\alpha_R(H_2O)_6$	1715	3568	1689	3499
β	1734	3593	1749	3627
$\beta(H_2O)_4$	1700	3441	1720	3406
$\beta(H_2O)_6$	1677	3396	1702	3366
P	1751[b]	3594	1756[b]	3612
$P(H_2O)_4$	1714	3490	1684	3405
$P(H_2O)_6$	1687	3401	1693	3382

[a] $[CH_3CONH]_1C^{\alpha}D^{\alpha}(CH_3)[CONHCH_3]_2$.
[b] Mixed modes.

case for the β and P conformations, indicative of weaker peptide hydrogen bonds to water in the former case resulting from the spatial constraints on such hydrogen-bond geometries that the fixed φ,ψ of these α_R conformations impose on the adjacent water molecules. Thus, the ~12–16 kcal/mol additional increase over the energies of β or P is attributable to the ability of the latter structures to form stronger peptide-group hydrogen bonds to adjacent water molecules.

The ν(CD) of hydrated species is given in Table V, and it is seen that the relative C^{α}—$D^{\alpha}\cdots O$ bond strength as measured by ν(CD) is not necessarily related to the relative energy of the species. Thus, although ν(CD) is higher in β than in P for both $ADP(H_2O)_4$ and $ADP(H_2O)_6$, the energy of β is lower than that of P for the four-water system and the reverse for the six-water structure (although for the set B parameters[14] we find that the frequency order is the same, 2322 cm^{-1} for β and 2265 cm^{-1} for P, for $ADP(H_2O)_6$).As seen for α_R, the four- and six-water structures cannot only have different ν(CD) (2284 cm^{-1} and 2281 cm^{-1}) but another six-water structure of very slightly higher total energy (0.20 kcal/mol) can have significantly stronger $D^{\alpha}\cdots O$ hydrogen bonds, indicated by a very much higher ν(CD) = 2300 cm^{-1}. For P, the four- and six-water structures also have different ν(CD), and in general the PCM(H^{α}) frequencies are different from the noncontinuum values, in some cases higher and in others lower.

Although our specific choice of φ,ψ is somewhat arbitrary at this point, a comparison of the $ADP\text{-}(H_2O)_6PCM(H^{\alpha})$ values of ν(CD), α_R = 2282 cm^{-1}, β = 2331 cm^{-1}, and P = 2293 cm^{-1}, indicates that these conformations (particularly the β and P) should be distinguishable by the suggested experimental protocol (of course, the actual frequency values will be modified by anharmonic corrections for this system,[41] and possible overtone observations would enhance the differences). It is interesting to note that the water interactions have enhanced the structural frequency differences over those of the isolated molecules. This is indicative of the conformation-specific interactions of the bonding water, emphasizing the necessity of considering the combined peptide-water hydrogen-bonded system as the characteristic structural entity, and indeed determining the likely structure in aqueous solution.[42] This makes it more important that a complete analysis of the expected spectra include a specification of the distribution of water structures about a given peptide conformation, which of course would derive from the suggested MD calculations.

Alanine Tripeptide (ATP), Tetrapeptide (ATeP), and Pentapeptide (APP)

To gain a deeper understanding of the issues raised with the ADP and how to handle longer chains, we first present the results of the ATP with different numbers of bound waters with and without a continuum solvent treatment. These are shown in Table VI for the β conformation. To have a direct comparison with the ADP, we focus on structures with at least a starting CO_1b–water hydrogen bond (see Figure 6). (As seen from the table, a starting CO_1a ATP-$(H_2O)_6$ structure is also at a minimum, but its total energy is much higher (by 6.17 kcal/mol), the $\nu(CD_1)$ are different (2324 vs 2338

Table V $C^{\alpha}D^{\alpha}$ Stretch Frequencies[a] of Hydrated Alpha-, Beta-, and Polyproline II-Conformation Alanine Dipeptide Structures

Structure	r(CD)[b]	ν(CD)[c]	$r(D\cdots O)$[d]
α_R	1.0957	2266	
$\alpha_R(H_2O)_4$	1.0939	2284	2.63
$\alpha_R(H_2O)_4$ R[e]	1.0942	2289	2.90
$\alpha_R(H_2O)_6$ (1)[f]	1.0944	2281	2.41, 3.13
$\alpha_R(H_2O)_6$ (1)[f]R[e]	1.0942	2285	3.02, 3.08
$\alpha_R(H_2O)_6$ (2)[f]	1.0924	2300	2.61, 2.84
$\alpha_R(H_2O)_6$ (2)[f]R[e]	1.0946	2282	3.12
β	1.0948	2276	
$\beta(H_2O)_4$	1.0891	2330	2.77, 2.69
$\beta(H_2O)_4$ R[e]	1.0892	2330	2.77, 2.76
$\beta(H_2O)_6$	1.0889	2331	2.75, 2.77
$\beta(H_2O)_6$ R[e]	1.0890	2331	2.72, 2.80
P	1.0939	2284	
$P(H_2O)_4$	1.0925	2296	3.03, 2.84
$P(H_2O)_4$ R[e]	1.0922	2300	2.97, 2.83
$P(H_2O)_6$	1.0937	2284	3.11, 2.96
$P(H_2O)_6$ R[e]	1.0932	2293	3.10, 2.94

[a] B3LYP/6-31 + G* calculation; set A $\varphi,\psi = -134^{\circ}, 145^{\circ}$.
[b] Length in Å.
[c] Frequency in cm^{-1}.
[d] Distances < 3.40 Å. First in pair is to CO_1b water, second to $[NH]_2\cdots O$ water.
[e] B3LYP/6-31 + G* PCM(H^{α}) calculation (see text).
[f] Two $\alpha_R(H_2O)_6$ optimized structures were found, (1) being 0.20 kcal/mol more stable than (2).

Table VI $C^{\alpha}D^{\alpha}$ Stretch Frequencies[a] in Various Hydrated *β*-Conformation Alanine Tripeptide Structures

Structure	$r(CD_1)$[b]	$\nu(CD_1)$[c]	$r(CD_2)$[b]	$\nu(CD_2)$[c]	$r(D\cdots O)$[d]
β	1.0946	2281	1.0946	2280	
$\beta(H_2O)_6$ a[e]	1.0896	2324	1.0898	2330	2.38; 2.87, 2.77
$\beta(H_2O)_6$ a[e] R[f]	1.0908	2319	1.0894	2327	2.47; 2.80, 2.81
$\beta(H_2O)_6$ b[g]	1.0887	2338	1.0899	2321	2.98, 2.40; 2.89, 2.79
$\beta(H_2O)_6$ b[g] R[f]	1.0892	2331	1.0894	2325	3.00, 2.51; 2.81, 2.81
$\beta(H_2O)_9$ b[g]	1.0894	2332	1.0866	2333	2.35, 3.23; 2.69, 2.86
$\beta(H_2O)_9$ b[g] R[f]	1.0885	2338	1.0885	2336	3.30, 2.32; 2.69, 2.87

[a] B3LYP/6-31 + G* calculation; set A $\varphi, \psi = -134°, 145°$.
[b] Length in Å.
[c] Frequency in cm^{-1}.
[d] Distances < 3.40 Å. First set: to CD_1; second set: to CD_2.
[e] Water on $[CO]_1$ initially placed at position a (see Figure 1).
[f] B3LYP/6-31 + G* PCM(H^{α}) calculation (see text).
[g] Water on $[CO]_1$ initially placed at position b (see Figure 2).

cm^{-1}), and the frequency differences, $\Delta\nu_{12}$, are not the same (6 vs -17 cm^{-1}). Clearly, the "initial" water-molecule structure (for this minimum number of waters) has a nontrivial influence on the arrangement of the subsequent water molecules). For ATP-$(H_2O)_6$ (see Figure 6), although the configuration of the two waters at CD_1 is similar to that in the ADP (see structure **9**, Figure 4), a closer $D^{\alpha}\cdots O$ distance in the ATP (2.40 Å vs 2.77 Å) is obviously responsible for the higher $\nu(CD_1)$, viz., 2338 cm^{-1} versus 2330 cm^{-1}. The PCM(H^{α}) calculation results in changes to both $\nu(CD_1)$ and $\nu(CD_2)$, with a large change in $\Delta\nu_{12}$ from -17 cm^{-1} to -6 cm^{-1}. The ATP-$(H_2O)_9$ structure (see Figure 7) again results in

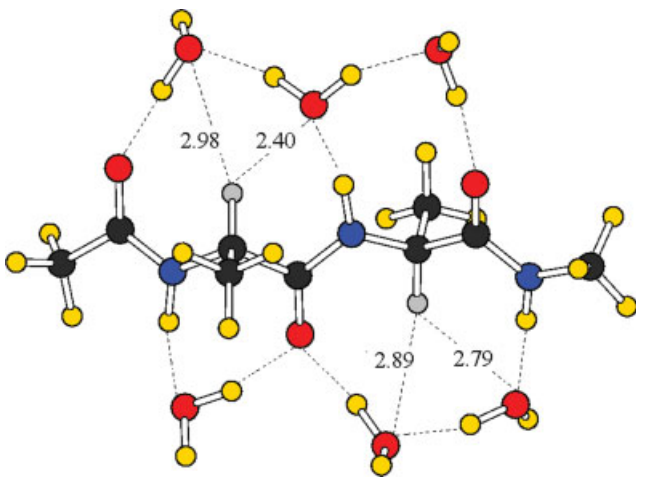

FIGURE 6 *β*-Conformation alanine tripeptide with six hydrogen-bonded waters.

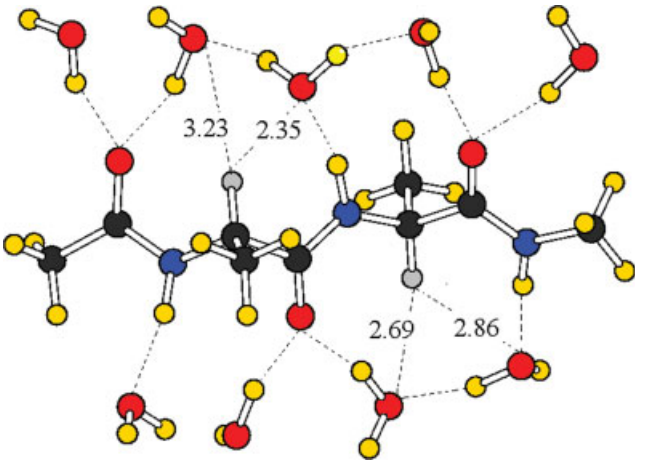

FIGURE 7 *β*-Conformation alanine tripeptide with nine hydrogen-bonded waters.

changes in $\nu(CD_1)$ and in $\Delta\nu_{12}$ as well in the PCM(H^{α}) calculation.

To get a relatively consistent comparison between peptides of increasing length, we have examined (non-PCM) minimally hydrated structures, that is, ones in which at least one water molecule was hydrogen bonded to each peptide group CO and NH. The results are shown in Table VII, and it is evident that the three different conformations can be distin-

Table VII Frequency Increments ($\Delta\nu_{ij}$)[a] in $C^{\alpha}D^{\alpha}$ Stretch of Minimally Hydrated[b] Uniform Ala-*n*[c] Peptides

		n			
	ij	1	2	3	4
α_R	ν_1[d]	2284	2272	2272	2267
	12		9	2	37
	23			−1	−30
	34				−11
	$\Delta\nu_{1n}$		9	1	−4
β	ν_1[d]	2330	2338	2299	2336
	12		−17	41	−2
	23			−5	2
	34				−1
	$\Delta\nu_{1n}$		−17	36	−1
P	ν_1[d]	2296	2290	2313	2303
	12		32	−3	9
	23			−28	−21
	34				−8
	$\Delta\nu_{1n}$		32	−31	−20

[a] $\Delta\nu_{ij} = \nu_j - \nu_i$ in a B3LYP/6-31 + G* calculation; set A φ, ψ.
[b] One water molecule hydrogen bonded to each peptide group CO and NH: ADP-$(H_2O)_4$, ATP-$(H_2O)_6$, ATeP-$(H_2O)_8$, and APP-$(H_2O)_{10}$.
[c] *n* refers to number of C^{α} atoms in $CH_3[CONHC^{\alpha}D^{\alpha}(CH_3)]_n$-$CONHCH_3$.
[d] Frequency of first $\nu(CD)$, at $[CH_3CONH]_1C^{\alpha}D^{\alpha}$ end.

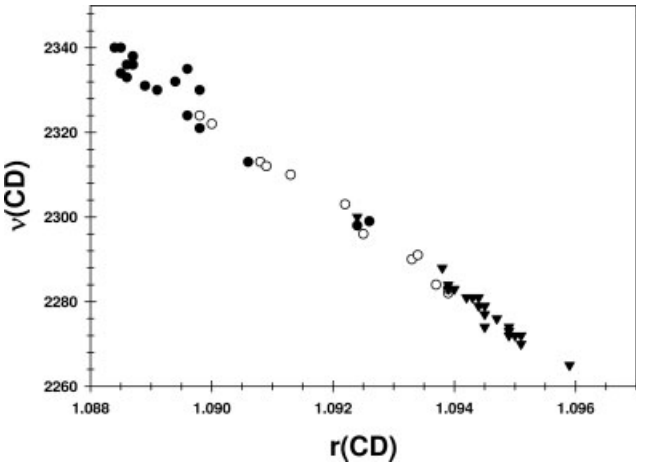

FIGURE 8 ν(CD) vs r(CD) for hydrated alanine peptide α_R (solid triangles), β (solid circles), and polyproline II (open circles) conformations.

guished from each other on the same basis as was true for the nonhydrated peptides,[9] namely, the pattern of values of $\nu(CD_1)$, $\Delta\nu_{ij}$, and $\nu_n - \nu_1$. It is also seen that the hydration significantly enhances the differences in these values from those of the nonhydrated species. We must now turn to the important question of how best to represent the actual hydration state of a peptide in aqueous solution.

Peptide Hydration State in Solution

It is strikingly clear from the above studies that the value of ν(CD) is directly related to the specific nature of the C^{α}—$D^{\alpha}\cdots$O interactions that are formed in aqueous solution, which determine the value of r(CD) and therefore ν(CD). It is also the case that a truly representative picture of the water structure surrounding any given peptide conformation cannot be adequately determined from a limited explicit-water ab initio calculation, even with inclusion of a water continuum surrounding the system, because the hydrogen-bonding structure to the peptide depends very sensitively on many-neighbor-water effects. Determining such structure reliably can best be done through MD calculations, and we turn now to the considerations that will have to be involved in such simulations.

The entry point to such an endeavor must be the determination of the distribution of r(CD), which eliminates (currently common) bond-length-constrained MD analyses. The results of our present ab initio studies of hydrated systems provide an explicit r(CD)-to-ν(CD) conversion, presented in Figure 8, the points from $r(CD) \approx 1.0930$–1.0960 overlapping those of the previous nonhydrated systems.[9] Thus, an analysis of the r(CD) distribution from the ensemble of MD trajectories for a given conformation, representative of the range of water structures and C^{α}—$D^{\alpha}\cdots$O hydrogen-bonding interactions, should provide a reasonable prediction of the ν(CD) band profile for that peptide structure. Fluctuations in conformation about the canonical φ,ψ are also testable.

The accuracy of such a protocol will depend on the ability of the molecular mechanics (MM) force field used in the MD simulations to reproduce the complete physical features of the hydrogen-bonding interaction, both for the peptide and the water. Unfortunately, current standard force fields are deficient in this respect, treating the hydrogen bond in a variety of ways but neglecting the essential electronic and structural changes associated with its formation. We have shown in our spectroscopically determined force field (SDFF) methodology that the needed added ingredient is to account for the changes in charge distribution with change in structure, which we accomplish through charge and dipole flux terms in the force field.[43] This is important not only in the intramolecular peptide force field[43,44] but even more so in the intermolecular interactions of the peptide and C^{α}—D^{α} groups in forming hydrogen bonds with water, in which an energy term representing the overlap of wave functions of the hydrogen-bonding partners needs to be included.[40] Future force field development must strive to incorporate such fundamental physical properties of interacting molecular systems.

CONCLUSIONS

Our previous studies of isolated α_R, β, and polyproline II alanine peptides demonstrated that these conformations could be distinguished through their $C^{\alpha}D^{\alpha}$ stretch frequency, ν(CD), differences.[9] This work extends these studies to explicitly hydrated species and shows that this discrimination is in fact enhanced by subtle differences in the water structure around the peptide related to C^{α}—$D^{\alpha}\cdots$O(water) in addition to peptide-water hydrogen bonding. Specifically: (1) Standard C=O$\cdots$H(water) and N—H$\cdots$O(water) hydrogen bonds have very little effect on ν(CD). (2) Short $D^{\alpha}\cdots$O distances, when possible, lead to such hydrogen bonds, resulting in shorter C^{α}—D^{α} bond lengths and higher ν(CD) (up to $\sim$70 cm^{-1}). (3) Longer $D^{\alpha}\cdots$O distances reduce these changes, to a point where H and O wave function overlap becomes negligible and changes are only attributable to electric field effects from the interacting water molecules. The latter, of course, are significantly dependent on the geometry of the interaction. (4) Results from varying the number of explicitly interacting water molecules show that the structure of the first-layer peptide-hydrogen-bonding waters is signifi-

cantly affected by the nonadjacent waters, thus limiting the predictive abilities of such ab initio calculations.

The last conclusion indicates that the optimum computation of the $\Delta\nu$(CD) associated with a given peptide conformation in an aqueous environment will have to be obtained from an MD calculation. Current standard force fields have a limited ability to achieve this goal, which will require the implementation of our SDFF protocol.[43] As a start, however, we are currently exploring the basic features of such MD simulations with available force fields, using the approach of developing initial MM-derived water configurations as starting points for subsequent QM optimization of selected multilayer water structures.[45]

REFERENCES

1. Krimm, S.; Bandekar, J. Adv Protein Chem 1986, 38, 181–364.
2. Krimm, S.; Abe, Y. Proc Natl Acad Sci USA 1972, 69, 2788–2792.
3. Moore, W. H.; Krimm, S. Proc Natl Acad Sci USA 1975, 72, 4933–4935.
4. Cheam, T. C.; Krimm, S. Chem Phys Lett 1984, 107, 613–616.
5. Lee, S.-H.; Krimm, S. Chem Phys 1998, 230, 277–295.
6. Mirkin, N. G.; Krimm, S. J Phys Chem A 2004, 108, 10923–10924.
7. Tiffany, M. L.; Krimm, S. Biopolymers 1968, 6, 1379–1382.
8. Krimm, S.; Tiffany, M. L. Israel J Chem 1974, 12, 189–200.
9. Mirkin, N. G.; Krimm, S. J Phys Chem A 2007, 111, 5300–5303.
10. Kinnaman, C. S.; Cremeens, M. E.; Romesberg, F. E.; Corcelli, S. A. J Am Chem Soc 2006, 128, 13334–13336.
11. Mirkin, N. G.; Krimm, S. J Mol Struct (Theochem) 1995, 334, 1–6.
12. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennuccci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Lui, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chem, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C. 02, Gaussian Inc., Waallingford, CT, 2004.
13. Mirkin, N. G.; Krimm, S. J Phys Chem A 2004, 108, 5438–5448.
14. Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. J Am Chem Soc 2007, 129, 1179–1189.
15. Wilson, E. B.; Decius, J. C.; Cross, P. C. Molecular Vibrations; McGraw-Hill: New York, 1955.
16. Wang, Z.-X.; Duan, Y. J Comput Chem 2004, 25, 1699–1716.
17. Jalkanen, K. J.; Degtyarenko, I. M.; Niemenen, R. M.; Cao, X.; Nafie, L. A.; Zhu, F.; Barron, L. D. Theor Chem Acc 2008, 119, 191–210.
18. Rossky, P. J.; Karplus, M. J Am Chem Soc 1079, 101, 1913–1937.
19. Cheam, T. C. J Mol Struct 1993, 295, 259–271.
20. Jalkanen, K. J.; Suhai, S. Chem Phys 1996, 208, 81–116.
21. Han, W.-G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. J Phys Chem B 1998, 102, 2587–2602.
22. Kalko, S. G.; Guàrdia, E.; Padró, J. A. J Phys Chem B 1999, 103, 3935–3941.
23. Mirkin, N. G.; Krimm, S. J Phys Chem B 2008, 112, 15267–15268.
24. Fraser, R. D. B.; MacRae, T. P. Conformation in Fibrous Proteins; Academic Press: New York, 1973, p 140.
25. Krimm, S. Nature 1966, 212, 1482–1483.
26. Krimm, S.; Kuroiwa, K.; Rebane, T. Conformation of Biopolymers, Vol. 2; Academic Press: London, 1967, p 439.
27. Krimm, S.; Kuroiwa, K. Biopolymers 1968, 6, 401–407.
28. Dwivedi, A. M.; Krimm, S. Biopolymers 1982, 21, 2377–2397.
29. Barnes, A. J. J Mol Struct 2004, 704, 3–9.
30. Scheiner, S.; Kar, T.; Gu, Y. J Biol Chem 2001, 276, 9832–9837.
31. Derewenda, Z.; Lee, L.; Derewenda, U. J Mol Biol 1995, 252, 248–262.
32. Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. Proc Natl Acad Sci USA 2001, 98, 9056–9061.
33. Manikandan, K.; Ramakumar, S. Proteins 2004, 56, 768–781.
34. Qian, W.; Krimm, S. J Phys Chem A 2002, 106, 6628–6636.
35. Qian, W.; Krimm, S. J Phys Chem A 2002, 106, 11663–11671.
36. Hermansson, K. J Phys Chem A 2002, 106, 4695–4702.
37. Qian, W.; Krimm, S. J Phys Chem A 2005, 109, 5608–5618.
38. Qian, W.; Krimm, S. J Mol Struct (Theochem) 2006, 766, 93–104.
39. Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. J Am Chem Soc 2000, 122, 4750–4755.
40. Mannfors, B.; Palmo, K.; Krimm, S. J Phys Chem A 2008, 112, 12667–12678.
41. Wang, J. J Phys Chem B 2009, 113, 1813–1816.
42. Poon, C.-D.; Samulski, E. T. J Am Chem Soc 2000, 122, 5642–5643.
43. Palmo, K.; Mannfors, B.; Mirkin, N. G.; Krimm, S. Biopolymers 2003, 68, 383–394.
44. Palmo, K.; Mannfors, B.; Mirkin, N. G.; Krimm, S. Chem Phys Lett 2006, 429, 628–632.
45. Bour, P.; Keiderling, T. A. J Chem Phys 2003, 119, 11253–11262.

Reviewing Editor: Laurence Nafie