Relation between calculated amide frequencies and solution structure in Ala-X peptides

ARTICLE *in* BIOPOLYMERS · JANUARY 1990

Impact Factor: 2.39 · DOI: 10.1002/bip.360300512 · Source: PubMed

3 AUTHORS, INCLUDING:

Robert Wayne Williams
Uniformed Services University of the H…

83 PUBLICATIONS **2,689** CITATIONS

Relation between Calculated Amide Frequencies and Solution Structure in Ala-X Peptides*

ROBERT W. WILLIAMS,[†1] JAMES L. WEAVER,[1] and ALFRED H. LOWREY[2]

[1]Department of Biochemistry, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, Maryland 20814-4799, and [2]Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375-5000

SYNOPSIS

Computational techniques have been used to aid interpretation of observed systematic shifts in the amide III frequencies of Ala-X peptides. Optimized structures and frequencies have been calculated for Ala-X peptides using GAUSSIAN86/88 with the 4-31G basis, MOPAC, and normal mode methods based on empirical force fields. We observe the following: (1) Frequencies calculated using scaled GAUSSIAN86 force constants correlate well with the experimental results. (2) Structures of the Ala-X peptides optimized by GAUSSIAN show a clear trend toward lower values of the dihedral angle ϕ as the X side chain becomes larger, while structures optimized here using semiempirical and empirical force fields do not show trends. (3) Computational changes in peptide conformations from β-sheet to α-helix produce large changes in both amide I and amide III frequencies that are inconsistent with the experimental results. (4) Computational changes in the dihedral angle ϕ of Ala-Ala produce a change in the amide III frequency consistent with the experimental results. (5) The experimental frequency shifts cannot be attributed directly to the effects of changing mass.

INTRODUCTION

In the preceding paper,[1] we show that the amide III frequencies of the neutral Ala-X peptides shift to a lower frequency as the side chain of amino acid X becomes larger. Related observations about a difference in the amide III frequency of homopolymers with alanine and glutamate side chains have been made by Krimm and co-workers.[2,3] The cause of this shift is not clear. In this paper we identify computational techniques that may lead to an explanation of this change in frequency.

Computational techniques have proven to be useful in interpreting spectroscopic observations.[3,4] There are three generally accepted computational methodologies: molecular mechanics using empirically determined force fields, self-consistent field methods with semiempirical parameterization of interactions, and ab initio calculations using calibrated basis sets. Much of the focus of this research since the earliest developments of computational molecular orbital theory has been the study of structural characteristics of polypeptides and proteins. Most of this work has been formulated in terms of isolated molecules in vacuum. This formulation has been successful in reproducing spectral features for crystalline peptides.[3,5-7] Calculations on peptides in environments other than a vacuum are more difficult and require more computational resources.

In this paper we compare our results on Ala-X peptides using the GAUSSIAN86 ab initio system, the MOPAC semiempirical package, and normal-mode calculations based on the empirical force field developed by Krimm and his co-workers.[3] We observe that ab initio results provide a good correlation with the experimental observations of peptides in aqueous solutions. The results also provide prelim-

CCC 0006-3525/90/5-60599-10 $04.00
Biopolymers, Vol. 30, 599–608 (1990)

* The opinions or assertions contained herein are the private ones of the authors and are not to be construed as official or reflecting the views of the U.S. Department of Defense or The Uniformed Services University of the Health Sciences.

† To whom correspondence should be addressed.

inary evidence that the experimentally observed shifts in amide III frequencies[1] are due to small systematic changes in the prefered value of the dihedral angle ϕ that depend on the X amino acid sidechain.

Diem and co-workers[8-10] have explained the basis for the conformational sensitivity of the amide III vibrations. Using Raman and vibrational CD measurements of L-alanine-L-alanine, D-alanine-L-alanine, and a number of deuterated isotopermers, they show that the in-plane N-H bend is strongly coupled to bending modes of the two methine protons, and suggest that changes in conformation change the relationships of these vibrations and hence the way in which they mix. These changes give rise to shifts in the frequencies commonly assigned to the amide III vibrations. Although these coupled modes may not all strictly speaking be amide III vibrations, there is ambiguity in the definition, and we refer to these modes as amide III vibrations when their potential energy distributions include a greater than 4% contribution from the in plane N-H bend.

The Ala-X peptide studied here contains two amino acids and one peptide group. Recent ab initio results[6,7] obtained from glycine and alanine dipeptides (that contain one amino acid and two peptide groups) have shown that hydrogen bonding and changes in conformation produce changes in the force constants and dipole moment derivatives. Our observations about the effects of side-chain composition and conformation in the Ala-X peptides on force constants, and a consideration of the solution environment, will appear elsewhere.

METHODS

Revision C of GAUSSIAN86[11] (G86) was executed on the CRAY X-MP at the National Cancer Institute Supercomputer Center (NCISC). Optimizations were performed using restricted Hartree–Fock (RHF) calculations with the standard 4-31G basis. The first two optimized structures, for Ala-Ala with β-sheet and α-helix conformations, were obtained starting with the neutral (nonzwitterion) structures of Ala-Ala optimized by MOPAC. Subsequent Ala-X peptides were optimized starting with this structure of Ala-Ala optimized by G86. Symmetry constraints were not used on the Gly-Gly or other peptides. Frequency calculations were performed with the 4-31G basis and analytically computed second derivatives of the energy. The 4-31G basis was chosen because it is computationally efficient and yields relatively accurate information about trends in force constants and frequencies.[4]

Quadratic force constants were scaled using a single scale factor, chosen to give reasonable agreement between calculated and observed frequencies for carboxylic acid, amide I, and amide III groups. The G86 cartesian force matrix in units of Hartrees · Bohr^{-1} was converted to mydnes · Å^{-1} and scaled by 81.06% by multiplying the matrix by a factor of 15.56919 × 0.8106. Force constants were also scaled using the procedure of Fogarasi and Balázs,[12] with five differential scale factors as modified by Cheam and Krimm[6,7]: 0.80 for hydrogen stretches, C^{β} bends, skeletal, and CO bends, and NC$^{\alpha}$C deformation; 0.88 for all other stretches; 0.86 for NH bends; 0.76 for CH bends; and 0.77 for out-of-plane bends and torsions.

Version 4.0 of MOPAC was obtained from the Quantum Chemistry Program Exchange (QCPE no. 455; Department of Chemistry, Indiana University, Bloomington, IN) and was installed on the VAX 8350 computer at the USUHS. MOPAC was executed using the AM1 Hamiltonian[13] with the options PRECISE, GNORM = 0.06 (or lower), SCFCRT = 0.0000001 (or lower), and FORCE.[14,15] The first optimized structure for Ala-Ala in the β-sheet conformation was obtained starting with ϕ and ψ angles of 180°. The first optimized structure for Ala-Ala in the α-helix conformation was obtained starting with $\phi = -57°$ and $\psi = -47°$. Subsequent Ala-X peptides were optimized starting with these structures of Ala-Ala. MOPAC calculations were performed on neutral non-zwitterionic peptides because N-terminal hydrogens migrated to the C-terminal carboxylate during optimizations of helical zwitterions.

Version 20 of CHARMm[16] was executed on the VAX 8650 at the NCISC using the TOPH19.INP topology file and the PARAM19.INP parameter file, describing amino acid geometry and point charges, provided at the NCISC. The peptides optimized were in the zwitterion form. Optimizations were performed with 50 cycles of steepest descents (SD) followed with 2000 steps of adopted basis Newton–Raphson (ABNR) energy minimization algorithms. The optimized structures were obtained with an energy tolerance of 0.0000000.

ORTEP-II was obtained from Oak Ridge National Laboratory,[17] and was executed on a Cromemco System 420 computer to produce stereo drawings of peptide structures and vibrational displacements.

Normal Mode Analysis

Normal mode analysis was performed using programs designed to implement the scaled quantum

mechanical force field (SQMFF) methodology[4] kindly provided by Lt. Floyd Cordell, USN, and Professor James Boggs at the University of Texas, Austin, and a general vibrational analysis system distributed by the Quantum Chemistry Program Exchange (QCMP no. 012). They were modified to run on a Cromemco System 420. The important features are the force matrix transformation and scaling, which allows a general use of force fields, including cartesian force matrices generated by G86 and the internal coordinate force field from Krimm and Bandekar.[3] The standard potential energy distribution (PED) allows assignment of calculated fundamental frequencies for the molecules.

The force field, taken from Table VI of Krimm and Bandekar,[3,18] is formulated in terms of vibrational internal coordinates. We use the force field values for β-poly(L-alanine). This force field (which is incomplete for some side chains and end groups) was augmented with force constants obtained from other sources. Values for C- and N-terminal groups are from Susi and Byler,[19] and from the averages of our ab initio calculations on Gly-Gly and Ala-Val. Diagonal force constants taken from our ab initio calculations are OH str 7.149, COH b 0.6669, amino NH str 6.7, carboxylic acid C=O str 11.96, carboxylic acid CO str 5.62, NH_2 wag 0.11; interaction terms are: carboxylic acid C=O str CO str 1.18, carboxylic acid C=O str CC_α str 0.335. Values for the valine side chain are from Bandekar and Krimm.[20] Values for the serine side chain are from Susi et al.[21] Values for the phenylalanine side chain were taken from those for toluene reported by Snyder and Painter.[22] The values used in these calculations are available from the authors upon request.

The MMADS system of programs kindly provided to us by Dr. George Famini at SMCCR-RSP-C CRDEC, Aberdeen Proving Grounds, MD 21010, was used to construct and edit the initial coordinates for the Ala-Ala peptide.

RESULTS

Optimizations to Stationary Points

The conformation of the Ala-X peptide optimized by each computational method is shown in Figure 1. Completely optimized geometries were obtained using G86 for Gly-Gly, Ala-Gly, Ala-Ala, and Ala-Val. A completely optimized geometry was also obtained for Ala-Phe. This was started with a partially optimized geometry for Ala-Leu using G86 with 15 cycles on the IBM 3090 at NIH, and completed over

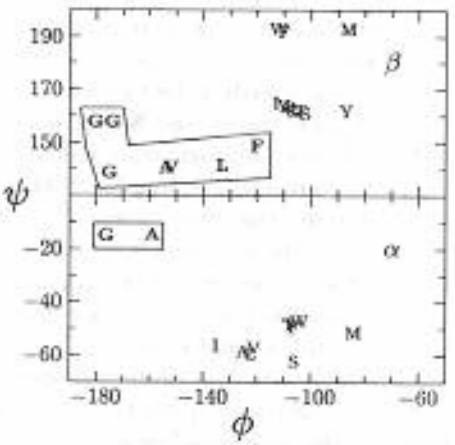

Figure 1. ϕ-ψ map of neutral L-Ala-L-X peptides. Coordinates optimized by GAUSSIAN86/88 are indicated by bold one-letter codes for X enclosed in boxes. The mark for G86 optimized Gly-Gly, GG, is included. Coordinates optimized by MOPAC are indicated by the uppercase one-letter code for X.

a period of about three CPU months using GAUSSIAN 88 with the SCF = DIRECT option on the VAX 8810 at the USUHS. As a practical note, the optimization and frequency calculations for Ala-Val required 67 CRAY X-MP CPU hours. Optimizations using the MOPAC package for initial geometries in the α-helix and β-sheet conformation were performed on Ala-X where X is Gly, Ala, Ser, Thr, Val, Ile, Leu, Met, Phe, Tyr, and Trp.

A partially optimized structure was obtained for Ala-Leu starting with the Ala-Val geometry. Some relatively large changes were observed in the Ala-Leu conformation during the first 15 cycles, including a 20° shift in ϕ shown in Figure 1. An additional 15 cycles produced only minor changes in conformation, with no further changes occurring in ϕ. A partially optimized structure was also obtained for Ala-Phe.

Geometries optimized by GAUSSIAN shown in Figure 1 indicate that there may be systematic trend, as a function of the X side chain, in the conformational preference for ϕ in this series of peptides. Starting with the same initial conformations, G86 and MOPAC produce different optimized conformations. There does not appear to be a systematic trend in conformations optimized by MOPAC. Conformations obtained from optimizations using CHARMm (not shown) were all virtually the same with respect to ψ and ϕ; for the β-structure these were $162 \pm 4°$ and $-159 \pm 3°$, respectively.

Using G86, the RHF energies favor (by about 6 kcal/mol) the α-helix conformation for the series

of Ala-X neutral structures. Our preliminary G86 calculations on this series of peptides with a −1 charge (with ionized carboxylate groups), and with a +1 charge (with a protonated N-terminal amine) show that the β-sheet conformation has a lower energy (unpublished results). Using MOPAC with the AM1 Hamiltonian, the final heats of formation slightly favor (by about 1 kcal/mole) the β-sheet conformation for the series of Ala-X neutral structures. Two exceptions, Ala-Thr and Ala-Tyr, exhibit a preference for the helical structure by about the same energy. However, the preferred structure of these peptides in solution cannot be directly derived from these results, since studies of solvated model peptide systems have shown that solvation makes an important contribution to the configuration entropy that cannot be neglected in calculations of the free energy differences between conformers.[23]

Frequencies from G86 Conformations

Frequencies based on these optimizations can be calculated using different force fields. It is important to realize that G86 generates a force matrix that can be treated in the same way as an empirical force matrix such as that provided by Krimm and Bandekar (K&B). These force matrices are used as input to programs using the Wilson[24] GF vibrational analysis. Figure 2 shows frequencies calculated from the G86-optimized structures using the force fields produced by G86 and the force field from Krimm and Bandekar.[3] When the G86 force constants are scaled by a single factor to fit one peptide (see discussion below) the calculated amide I frequencies for four peptides differ from the observed frequencies by less than 1 cm^{-1} (Figure 2). The correlation coefficient for these four points is 0.98. The amide I frequencies calculated using the K&B force constants (Figure 2) show a correlation with experiment of 0.81, and the agreement is relatively good compared with other methods (see below), considering that this force field was developed from solid state homopolymers.

Scaling of G86 Force Constants

The force matrices from G86 were each scaled by a single factor of 81.06% to give agreement between the calculated β-sheet Ala-Ala amide I frequency and the experimentally observed frequency at 1680 cm^{-1}. While this factor is generally lower than differential scale factors chosen by others,[6,7] particularly for C=O stretches (0.88) and N-H bends (0.86), the lower value may reflect the need to find

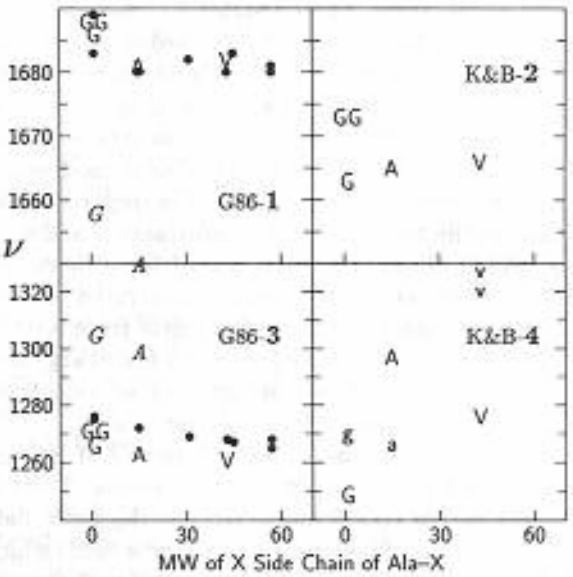

Figure 2. Amide I and III frequencies calculated for the series of peptides Ala-X with geometries optimized in β-sheet conformation. Several sets of force constants were used to calculate frequencies, those calculated by G86 and the set from K&B.[3] Calculated frequencies are marked by uppercase one-letter codes for X. Experimentally observed frequencies are shown by the solid circles. Frequencies calculated from G86 helix geometries are marked by italic one-letter codes. The set of cartesian coordinate force constants calculated for each peptide was scaled by 81.06%. This factor was choosen so that the calculated amide I frequency for β-sheet Ala-Ala was the same as the observed frequency for Ala-Ala. (1) Amide I, G86 force constants. The frequency calculated for Gly-Gly is shown (GG) overlapping the observed frequency for Gly-Gly. The mark for helical Ala-Ala (*A*) is at 1650 cm^{-1}. (2) Amide I, K&B force constants. (3) Amide III, G86 force constants. (4) Amide III, K&B force constants. Uppercase letters represent frequencies with a greater than 10% contribution from the PED of N-H in plane bend. Lowercase letters represent a 3–10% PED.

scale factors that yield accurate predictions of frequencies for peptides in aqueous solutions.

Our ab initio calculations show that hydrogen bonds between water and these polar groups decreases the C=O stretch force constant considerably while increasing the N-H bend force constant (R. Williams and A. Lowrey, unpublished). The single scale factor 81.06% gives a reasonable prediction of experimentally observed frequencies that are at the focus of this paper. Agreement between calculated and observed frequencies for the amide III and carboxylic acid C=O stretch is relatively good. The calculated amide III frequencies are between 5 and 10 cm^{-1} lower than the observed frequencies (Figure

2). The calculated carboxylic acid C=O frequency for helical Ala-Ala is at 1735 cm^{-1} while the experimentally observed frequency, for Ala-Ala in 1 *M* HCl, is at 1727 cm^{-1} with a bandwidth at half height of 35 cm^{-1}. Another choice of scale factor, to make the calculated α-helix Ala-Ala amide I frequency agree with the observation, would make the calculated β-sheet amide I and carboxylic acid frequencies high and outside of the range in which they are normally observed. This does not indicate that Ala-Ala is in the β-sheet conformation in solution, however. The amide I and carboxylic acid frequencies depend on other variables in the peptide conformation that must be studied before such a conclusion can be made.

When we used the differential scale factors, the calculated amide III frequency was the same as that from the single scale factor. However, the amide I and carboxylic acid frequencies were 46 and 60 cm^{-1} higher than those observed, respectively.

The amide I C=O stretch observed[1] at 1680 cm^{-1} was chosen as the reference for the scale factor calculation because it has a strong force constant and its frequency is virtually unchanged by the small geometry changes observed in the alanine peptide upon going from vacuum to solution.[25,26] We included water molecules with hydrogen bonds, constructed with a length of 2.03 Å and an angle of 175°, to the amide carbonyl oxygen of our peptide in our normal mode calculations using the K&B force field, and observed that the vibrational coupling with the water molecule did not shift the calculated amide I frequency by more than 3 cm^{-1}. Cheam and Krimm[6,7] observed that hydrogen bonding with the amide C=O group lowers the diagonal C=O stretch force constant (as well as producing a change in the magnitude and sign of interaction terms). Our experimental results[1] show that strong acid and base do not significantly shift the amide I frequency. In 1 *M* KOH and at neutral pH the amide I for Ala-Ala is at 1680 cm^{-1}, while in 1 *M* HCl, where the carboxylate group is neutralized, the amide I is at 1685 cm^{-1}. We suggest that the average effects of solvation on the amide I frequency may be approximated when scale factors are chosen to yield a calculated frequency that agrees with the frequency observed from the peptide in water.

MOPAC Calculated Frequencies

Figure 3 shows frequencies calculated by the MOPAC package for structures optimized from initial β-sheet and α-helix conformations. With respect to each conformation, the variation of the amide I and amide III frequencies show no evidence of a systematic trend. For the geometries optimized for each peptide Ala-X the shifts in the frequencies and the contributions from the PED of each amide mode do not exhibit systematic variation as a function of the molecular weight of the side chain of amino acid X. With respect to the experimentally observed frequencies, the amide related frequencies calculated by the MOPAC package are displaced approximately 200 cm^{-1} higher. With respect to conformation, the MOPAC frequencies exhibit a systematic shift between the optimized β-sheet conformation and the α-helical conformation for each peptide.

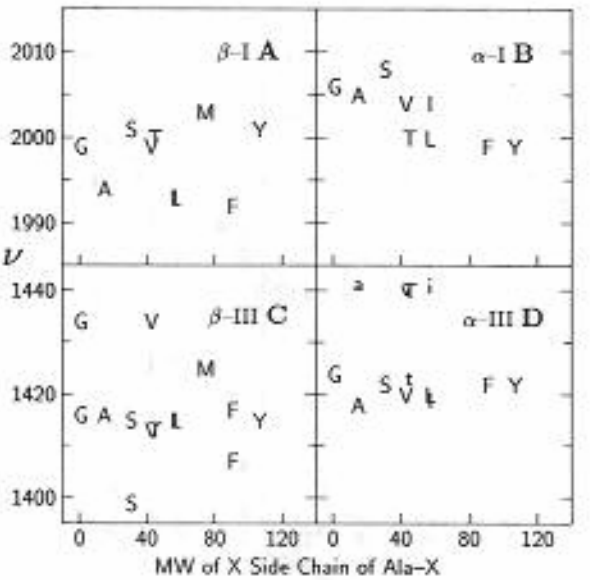

Figure 3. Amide I and III frequencies calculated by MOPAC for the series of neutral peptides Ala-X. (A) β-sheet conformation, amide I. (B) Helix conformation, amide I. (C) β-sheet conformation amide III. (D) Helix conformation, amide III. Amide III frequencies with a greater than 10% contribution from the PED of in plane N-H bend are marked with uppercase letters. Those with between 5 and 10% of this PED are marked with lowercase letters.

Using the optimized geometries from MOPAC it is possible to calculate frequencies and PEDs using the K&B force field. Figure 4 shows these results. Again, the calculated amide III frequencies show a lack of systematic variation with respect to molecular weight of the side chain of the X amino acid. However, the frequencies calculated with this force field are much closer in value to the observations. We also observe that these results do not reproduce the conformation (β-sheet to α-helix) dependent shifts observed in Figure 3.

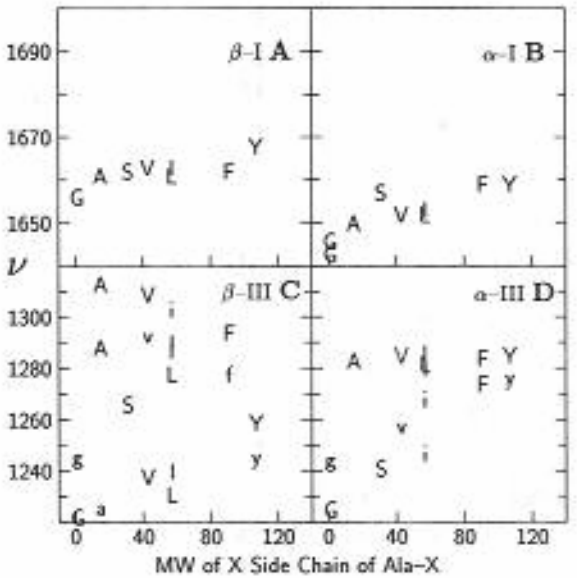

Figure 4. Amide I and III frequencies calculated for the series of neutral peptides Ala-X with MOPAC optimized geometries and force constants from K&B.[3] (A) Amide I frequencies, β-sheet conformation. (B) Amide I frequencies, α-helix conformation. (C) Amide III frequencies, β-sheet conformation. Uppercase letters represent frequencies with a greater than 10% contribution from the PED of N-H in plane bend. Lowercase letters represent a 3–10% PED. One frequency from Ala-Gly (g) at 1198 cm^{-1}, and two for Ala-Ser at 1328 cm^{-1} (s) and 1159 cm^{-1} (S) are omitted. (D) Amide III frequencies, α-helix conformation.

CHARMm Calculated Frequencies

There was no systematic variation in CHARMm calculated amide III frequencies as a function of the X side-chain molecular weight. The amide III frequencies appear to be nearly constant at about 1410 cm^{-1} for most of the X residues in the β-sheet conformation.

Amide III and Amide I Frequency Shifts

Early vibrational analysis of polypeptides[27] has shown that the amide III band is very sensitive to small changes in ϕ and ψ. Diem and his co-workers[8-10] have shown that this sensitivity is due to variations in the coupling of in-plane N-H bending with bending vibrations of methine protons on the two adjacent α-carbons. To investigate this conformational dependence, and its relationship to contributions from these three vibrations, we calculated amide I and amide III frequencies, and the potential energy distributions for these modes, as a function of the dihedral angle ϕ starting with the G86 optimized structure of Ala-Ala. As can be seen in Figure 5, several frequencies in the amide III region can have a significant contribution from in plane N-H bending. Our operational definition of the amide III mode here simply requires that N-H in plane bending contributes 5% or more to the PED. One of these calculated amide III frequencies shifts to lower values with a magnitude similar to or greater than the experimentally observed change, while the other amide III frequencies and the amide I frequency remain relatively constant. Contributions from in plane N-H bending and methine proton bending to the PED (not shown) for the shifting frequencies also change, while contributions for the constant frequencies remain relatively constant.

The vibrational displacements of these frequencies are shown in Figure 6. Here it can be seen that the helix conformation amide III frequency observed

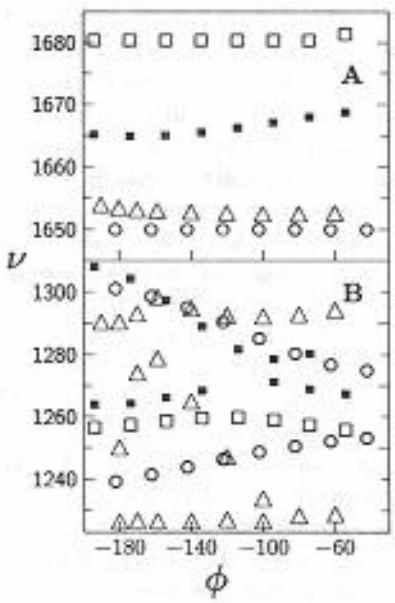

Figure 5. Amide I (A) and III (B) frequencies calculated as a function of ϕ using extended and helical conformations with different sets of force constants. Several of the calculated frequencies in the amide III region have significant contributions from the in-plane N-H bend. Some of these frequencies decrease as ϕ is changed from −170 to −80, while the amide I frequency remains relatively constant. Frequencies between 1200 and 1320 cm^{-1} with a greater than 5% contribution from the PED of in plane N-H bending are included in B. Force constants were taken from our GAUSSIAN86 results for β Ala-Ala (G86β), for helical Ala-Ala (G86α), and from K&B.[3] Open boxes: G86β geometry, G86β force constants; closed boxes: G86β geometry, K&B force constants; open circles: G86α geometry, G86β force constants. Open triangles: G86α geometry, G86α force constants.

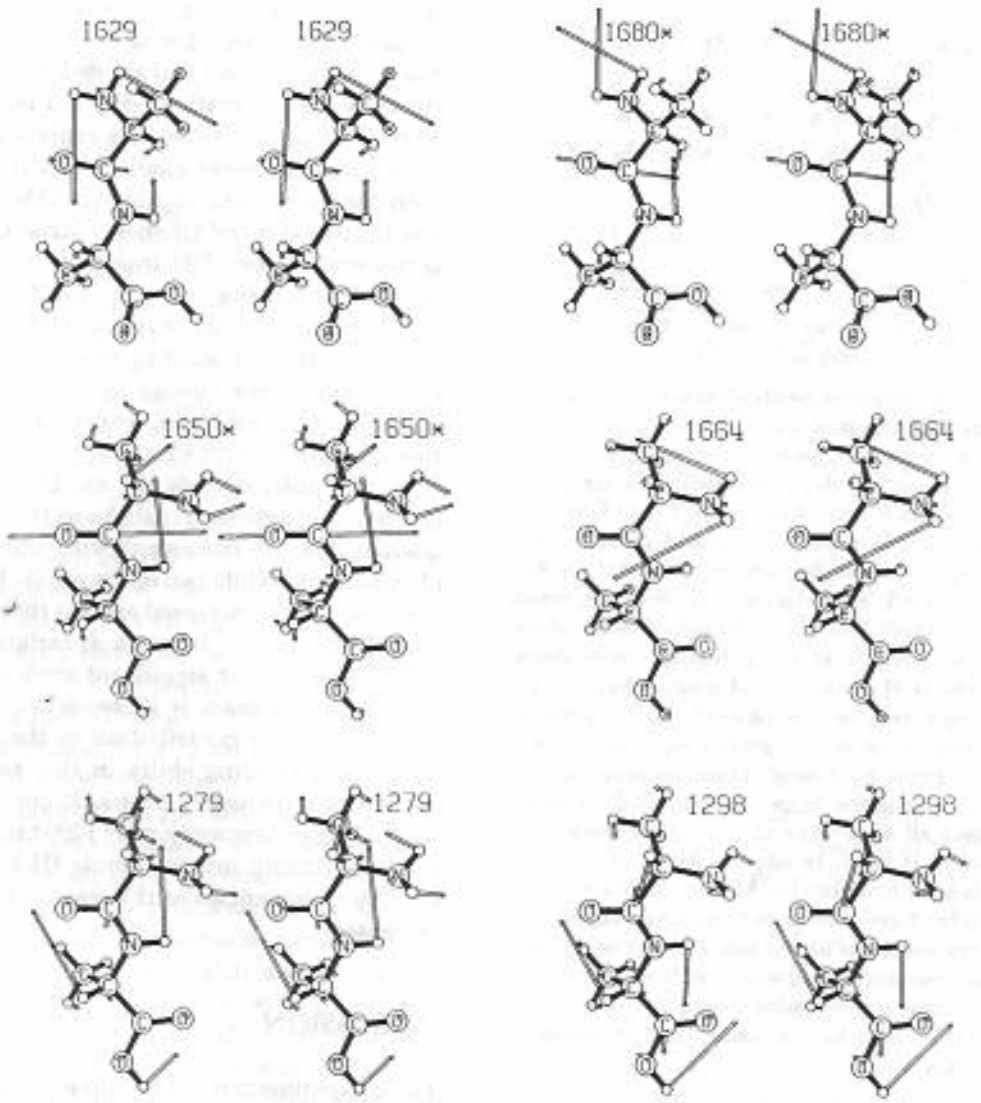

Figure 6. Ortep drawings of L-Ala-L-Ala, showing vibrational displacements of normal modes for amide I and amide III frequencies calculated by GAUSSIAN86. Magnitudes of the displacements shown here have been multiplied by a factor of five. Asterisks following the frequency indicate that the phase of all displacements were rotated 180° for clarity. Top row: Frequencies numbered 49 and 50 calculated from the β-sheet conformation, showing a splitting of the amide I mode at 1629 and 1680 cm^{-1}. Middle row: Frequencies numbered 49 and 50 calculated from the α-helix conformation, showing that the PED of the amide I mode has shifted completely to frequency 49 at 1650 cm^{-1}. Bottom row: Frequencies numbered 36 and 37 calculated from the α-helix conformation, showing splitting in the amide III frequency. Frequency 36 at 1279 cm^{-1} shifts to lower values as ϕ is rotated towards zero.

to shift, at 1279 cm^{-1}, is coupled to both the $C_{\alpha,1}$-H and $C_{\alpha,2}$-H bends in an in phase vibration where all of these displacements are in the same general direction. The other amide III frequency, at 1298 cm^{-1}, has a $C_{\alpha,2}$-H displacement that is out of phase with respect to the $C_{\alpha,1}$-H and N-H displacements. Our ab initio calculations on Ala-Gly using the 4-31G** basis (unpublished) indicate that the Raman intensity of this all-in-phase amide III vibration is significantly higher than the out of phase modes. This in phase mode also appears to always have the highest N-H bend contribution to the PED. A similar distribution of the amide III modes was observed using the K&B force field on the β-sheet structure.

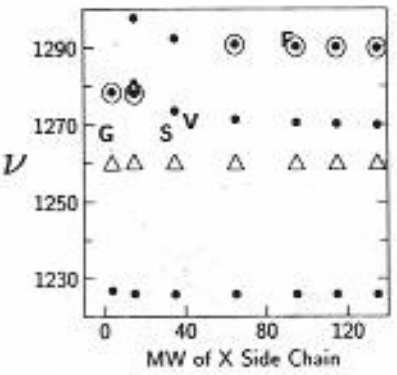

Figure 7. Mass dependence of calculated amide III frequencies. The side chain of the X amino acid in Ala-X was changed while the remainder of the structure was held constant. In one set of calculations using the structures and force constants for Ala-Ala obtained with G86, only the mass of the C-terminal amino acid β-carbon was changed, giving a series of identical conformations of Ala-Ala with a superficially altered mass. Solid circles represent three frequencies calculated from the helix Ala-Ala structure with a contribution in the potential energy distribution from the N-H in plane bend greater than 7%; the open circles (superimposed on the solid circles) represent these same frequencies with a greater than 18% contribution from in plane N-H bend. Triangles represent the one frequency calculated from the extended (β-sheet) structure; these all have more than a 21% contribution from in plane N-H bend. In another set of calculations using the extended structure for Ala-Ala optimized using MOPAC, and the force constants from K&B[3] augmented with other force constants for the side chains, the X amino acid side chain was exchanged with a different side chain while the remaining conformation was held constant. The bold single-letter codes for the amino acid X represent these frequencies.

Changes in ψ, going from the β-sheet to α-helix conformations, appear to produce changes in the amide I frequencies that are inconsistent with the experimental observations.

We investigated two conformations of the C-terminal carboxylic acid that lead to stationary points, one represented by the structure shown at the top of Figure 6 where the hydroxyl C-O bond is *cis* to the backbone (COH-*cis*), and one represented by the structures shown in the middle and bottom of Figure 6 where the hydroxyl C-O bond is *trans* to the backbone (COH-*trans*). Ab initio optimizations and frequency calculations were performed on both of these conformations in the extended forms of Ala-Gly and Ala-Ala. The COH-*trans* conformations had slightly lower energies, and values of ϕ that were about 5° more negative, than those with COH-*cis*. COH-*cis* conformations of Ala-Gly and Ala-Ala had $\phi = -175.4°$ and $-154.0°$, respectively, while COH-*trans* conformations had $\phi = -179.8°$ and $-160.0°$, respectively. Amide I and amide III frequencies from these two conformations differed by only 3 cm^{-1}. The ab initio conformations represented in Figure 1 are all of the lower energy COH-*trans* structures with the exceptions of Ala-Gly, Ala-Ala, and Ala-Val in the extended (β-sheet) structures; This was done because the COH-*trans* structure for Ala-Val is not yet available.

To investigate the hypothesis that the experimentally observed shift in amide III frequencies is due simply to the change in mass associated with each different side chain, we constructed several series of peptides with identical backbone conformations and different side chains. The results, shown in Figure 7, indicate no shifts in the amide III frequency that are consistent with the experimental observations. While two of the amide III frequencies calculated using increased mass at the β_2 carbon shift down about 8 cm^{-1} between β_2 carbon massis of 12 and 40, no further significant shift is observed as the side-chain mass is increased to 130. Furthermore, the major contribution to the PED from in plane N-H bending shifts in this range from the mode with a frequency near 1270 cm^{-1}, to the mode with a higher frequency near 1290 cm^{-1}, suggesting that the Raman intense amide III band may shift to higher frequencies as the mass of the β-carbon is increased.

DISCUSSION

These experimental[1] and computational results have the following general implications: Peptides with only two amino acids exhibit a conformational preference for extended structures in aqueous solutions. Small variations in the side-chain dependent conformational preferences of these peptides can be measured using Raman spectroscopy. Empirical and semiempirical force fields currently neglect forces in the backbone dihedral angles that are side-chain dependent. Results from this and related studies may provide information that can be used to refine empirical force fields.

There are four hypothetical explanations for the experimentally observed[1] trend in amide III frequencies: (1) The shift may be due simply to differences in side-chain mass, independent of conformational changes. We show here that the experimental shift cannot be reproduced with computational changes in the side-chain mass, but we have grown increasingly aware that the computa-

tional shifts in both the frequencies of particular amide III related normal modes and the contributions from N-H and methine C-H bending to the PEDs of these modes depend on the conformations and force constants used in these calculations. We cannot yet reject the possibility that a properly scaled and refined force field, or that a consideration of water, will give support to this hypothesis. This particular hypothesis must be rejected by a greater preponderance of evidence than is currently available. (2) The frequency shift may be due to a shift in the conformation equilibrium. Specifically, we suspect that the experimentally observed shift in frequency is due to a shift in the conformational preference for the dihedral angle ϕ. The explanation for the conformational sensitivity of the amide III vibrations provided by Diem and co-workers,[8-10] the trend in ϕ produced by G86 and G88, and the computational dependence of one of the amide III frequencies on ϕ shown here all support this possibility. (3) The frequency shift may be due to a side-chain dependent shift in force constants for vibrations related to the amide III modes. Our current set of quantum-mechanical force constants, for Ala-Gly, Ala-Ala, and Ala-Val (not shown here) do not show trends that support this hypothesis, but this data is not adequate to provide strong evidence either for or against it. We are close to obtaining force constants for other peptides in this series that will help resolve this question. This series of force constants will also provide information that may be useful in the refinement of empirical force fields. (4) The frequency shift may be due to a shift in the vibrational coupling between the X side chain and the N-H and methine C-H bends that is independent of conformation. There is currently no evidence to support this hypothesis. As can be seen in Figure 6, and in the related figures for amide III vibrations of Ala-Ala and Ala-Val in the extended conformation (not shown), displacements related to amide III vibrations do not appear to be strongly coupled to vibrational modes of the side chains. However, this is a possibility that requires investigation when the SQMFFs for Ala-Phe and Ala-Leu become available.

Early Raman studies of alanine peptides showed that spectra of solid state dialanine are similar to the spectra of dialanine in aqueous solution.[28] The amide III frequencies for solid state and solution Ala-Ala were observed at 1265 and 1276 cm^{-1}, respectively.[28] Our solid state and solution spectra of Ala-Ala have amide III frequencies at 1261 and 1276 cm^{-1}, respectively, and amide I frequencies at 1682 and 1680 cm^{-1}, respectively.[1] It is significant that the amide I frequency for solid and solution Ala-Ala is essentially the same. This suggests that the conformation of these solid and solution states of Ala-Ala are similar in the sense that they may be represented in the same region on the ϕ-ψ map. This can be contrasted to observations made by Gupta and Gupta[5] that both experimental and calculated amide I frequencies for solid Ala-Ala are significantly different from those for solid Ala-Ala · HCl, which has a different conformation.

Frequencies calculated from structures optimized by MOPAC and CHARMm do not show a systematic shift in the amide III frequency that reproduces the experimental[1] results. With respect to the results from MOPAC, this may reflect that the parameterization of interactions does not adequately include the effect of the side chain on the backbone conformation. These and related results may be of some value in refining the parameters used in empirical and semiempirical methods. With respect to the results from CHARMm, it appears to be well known that the potential functions used in molecular mechanics approaches are crude at the level of predicting vibrational frequencies and are not considered meaningful at this time.

Our results are consistent with observations made by others that the SQMFF methodology is useful for obtaining force constants that are transferable to other molecules. When the force matrix obtained from Ala-Ala in the β-sheet conformation was used to calculate frequencies for Ala-Ala in the α-helix conformation, the amide frequencies obtained were similar to those calculated using the force matrix obtained from the α-helix conformation.

The amide I frequency calculated at 1650 cm^{-1} for Ala-Gly and Ala-Ala in the α-helix conformation agrees with the amide I frequencies generally observed in mostly helical proteins.[29] This may be a fortuitous result of our choice of scale factor, but it suggests that the scale factor may be close to being correct. Computational changes in peptide conformation from β-sheet to α-helix produce large changes in both amide I and amide III frequencies. This suggests that the conformational equilibrium for the Ala-X peptides does not shift between β-sheet and α-helix in solution as a function of the X side chain. In accordance with this possibility, the experimentally observed shift in the amide III frequency may reflect a much smaller degree of conformational change, such as the change in ϕ investigated here.

This work was supported in part by the Strategic Defense Initiative Organization Medical Free Electron Laser Program (at USUHS, Medical Applications of Advanced La-

ser Technology Project grant number GM7160), in part by the Office of Naval Research, and by a generous grant of supercomputer time on the CRAY X-MP at the National Cancer Institute Supercomputer Facility in Frederick, MD.

We thank Lt. Floyd Cordell, USN, and Professor James Boggs at the University of Texas, Austin, for kindly providing programs to implement the SQMFF methodology; Dr. George Famini at Aberdeen Proving Grounds, MD, for kindly providing the MMADS programs; and Professor Samuel Krimm for reading the manuscript and making some helpful suggestions.

REFERENCES

1. Weaver, J. & Williams, R. W. (1990) *Biopolymers*, preceding paper.
2. Sengupta, P. K. & Krimm, S. (1985) *Biopolymers* **24,** 1479–1491.
3. Krimm, S. & Bandekar, J. (1986) *Adv. Protein Chem.* **38,** 183–364.
4. Fogarsi, G. & Pulay, P. (1985) in *Vibrational Spectra and Structure: A Series of Advances*, Vol. 14, Durig, J. R., Ed., Elsevier, New York, pp. 125–219.
5. Gupta, M. K. & Gupta, V. D. (1978) *Ind. J. Biochem. Biophys.* **15,** 413–416.
6. Cheam, T. C. & Krimm, S. (1989a) *J. Mol. Struct.* **193,** 1–34.
7. Cheam, T. C. & Krimm, S. (1989b) *J. Mol. Struct.* (Theochem.) **188,** 15–43.
8. Oboodi, M. R., Alva, C. & Diem, M. (1984) *J. Phys. Chem.* **88,** 501–505.
9. Diem, M., Oboodi, M. R. & Alva, C. (1984) *Biopolymers* **23,** 1917–1930.
10. Roberts, G. M., Lee, O., Calienni, J. & Diem, M. (1988) *J. Am. Chem. Soc.* **110,** 1749–1752.
11. Frisch, M. J., Binkley, J. S., Schlegel, H. B., Raghavachari, K., Melius, C. F., Martin, R. L., Stewart, J. J. P., Bobrowicz, F. W., Rohlfing, C. M., Kahn, L. R., Defrees, D. J., Seeger, R., Whiteside, R. A., Fox, D. J., Fleuder, E. M. & Pople, J. A. (1984) *Gaussian 86*, Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh, PA.
12. Fogarasi, G. & Balázs, A. (1985) *J. Mol. Struct.* (Theochem.) **133,** 105–123.
13. Dewar, M. J. S., Zoebisch, E. G., Healy, E. F. & Stewart, J. J. P. (1985) *J. Am. Chem. Soc.* **107,** 3902–3909.
14. Dewar, M. J. S. & Ford, G. P. (1977) *J. Am. Chem. Soc.* **99,** 1685–1691.
15. Dewar, M. J. S., Ford, G. P., McKee, M., Rzepa, H. S., Thiel, W. & Yamaguchi, Y. (1978) *J. Mol. Struct.* **43,** 135–138.
16. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4,** 187–217.
17. Johnson, C. K. (1976) *ORTEP-II: A FORTRAN Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations*, Oak Ridge National Laboratory ORNL-5138.
18. Dwivedi, A. M. & Krimm, S. (1984) *J. Phys. Chem.* **88,** 620–627.
19. Susi, H. & Byler, D. M. (1980) *J. Mol. Struct.* **63,** 1–11.
20. Bandekar, J. & Krimm, S. (1988) *Biopolymers* **27,** 885–908.
21. Susi, H., Byler, D. M. & Gerasimowicz, W. V. (1983) *J. Mol. Struct.* **102,** 63–79.
22. Snyder, R. W. & Painter, P. C. (1981) *Polymer* **22,** 1629–1632.
23. Brady, J. & Karplus, M. (1985) *J. Am. Chem. Soc.* **107,** 6103–6105.
24. Wilson, E. B., Decius, J. C. & Cross, P. C. (1955) *Molecular Vibrations*, McGraw-Hill, New York.
25. Rossky, P. J. & Karplus, M. (1979) *J. Am. Chem. Soc.* **101,** 1913–1937.
26. Pettitt, B. M. & Karplus, M. (1985) *Chem. Phys. Lett.* **121,** 194–201.
27. Hsu, S. L., Moore, W. H. & Krimm, S. (1976) *Biopolymers* **15,** 1513–1528.
28. Sutton, P. & Koenig, J. L. (1970) *Biopolymers* **9,** 615–634.
29. Williams, R. W. & Dunker, A. K. (1981) *J. Mol. Biol.* **152,** 783–813.

Received January 22, 1990
Accepted April 16, 1990