



Published in final edited form as:

Biochemistry. 2009 March 31; 48(12): 2777–2787. doi:10.1021/bi801293v.

## Extension of the Tryptophan $\chi^{2,1}$ Dihedral Angle - W3 Band Frequency Relationship to a Full Rotation: Correlations and Caveats<sup>‡</sup>

Laura J. Juszczak<sup>\*,§,+</sup> and Ruel Z. B. Desamero<sup>\*,∂,#</sup>

<sup>+</sup> Chemistry Department, Brooklyn College, The City University of New York, 2900 Bedford Ave., Brooklyn, NY 11210

<sup>#</sup> Chemistry Department, York College, The City University of New York, 94-20 Guy R. Brewer Blvd., Jamaica, NY 1145

<sup>∂</sup> Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461

<sup>§</sup> Department of Physiology and Biophysics, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461

### Abstract

The correlation of the UVRR vW3 mode with the tryptophan  $\chi^{2,1}$  dihedral angle (1-3) has been extended to a full, 360° rotation. The three-fold periodicity of the relationship ( $\cos 3\chi^{2,1}$ ) over 360° results in up to six dihedral angles for a given vW3. Consideration of a Newman plot of dihedral angles for proteinaceous tryptophans taken from the Protein Data Bank shows that sterically hindered ranges of dihedral angle reduce the possible  $\chi^{2,1}$  to one or two. However, not all proteinaceous tryptophans follow the vW3- $\chi^{2,1}$  relationship. Hydrogen bonding at the indole amine, weaker, electrostatic cation- $\pi$  and anion quadrupole interactions and environmental hydrophobicity are examined as possible contributing factors to noncompliance with the relationship. This evaluation suggests that cumulative weak electrostatic and nonpolar interactions characterize the environment of tryptophans that obey the vW3- $\chi^{2,1}$  relationship, matching that of the crystalline tryptophan derivatives used to derive the relationship. In the absence of methods to quantify these weak interactions, measurement of the full width half maximum bandwidth (FWHM) of the W3 band is suggested as a primary screen for evaluating the applicability of the vW3 -  $\chi^{2,1}$  relationship.

### Keywords

tryptophan; protein structure; triose phosphate isomerase; DFT calculations; resonance Raman spectroscopy; dihedral angle; UVRR

---

The utility of spectroscopy with respect to protein structure lies in its predictive value in the absence of crystallographic structure. In UV resonance Raman spectroscopy, tryptophan residues are structural bellwethers because they can be selectively probed at 229 nm. Tryptophan vibrational bands are good markers for hydrophobic interaction (3-6), hydrogen

---

<sup>‡</sup>This work was supported by National Institutes of Health grants GM08153 (RZBD), 5P01GM068036, EB01958, R01 EB-00296 and the W. M. Keck Foundation.

\*To whom correspondence should be addressed. ljuzak@brooklyn.cuny.edu. Telephone: 718 951 5000 × 1426. Fax: 718 951 4607. desamero@york.cuny.edu. Telephone: 718-262-2657. Fax: 718-262-2652.

bonding (5,6) and conformation (1,2,6-8). The intensity ratio,  $I_{1360}/I_{1340}$ , of the tryptophan Fermi 'doublet' (also known as the W7 mode) serves as a good marker for hydrophobic interactions (2,4,9). A rise in the ratio indicates an increasingly hydrophobic environment for the indole ring. Several tryptophan modes, W2, W4, W6, and W17 are sensitive to hydrogen bonding. The stronger the hydrogen bonding the higher is  $\nu W2$ ,  $\nu W4$  and  $\nu W6$  (2) and the lower is  $\nu W17$  (8). One mode, the W3 mode, is of great interest because it tracks the conformation of tryptophan. The W3 mode is a hybrid mode consisting of indole  $C_2=C_3$  stretching (50%),  $N_1-C_2$  (23%) stretching and  $C_2-H$  (19%) bending vibrations (2,6), yet its frequency depends on the  $\chi^{2,1}$  dihedral or torsion angle about the  $C_\alpha-C_\beta-C_3=C_2$  linkage (2). A systematic study of the W3 mode frequency ( $\nu W3$ ) for several crystalline tryptophan analogs revealed that  $\nu W3$  varies with the  $\chi^{2,1}$  torsion angle according to the empirically derived equation (1):

$$\nu W3 = 1542 + 6.7(\cos 3|\chi^{2,1}| + 1)^{1,2} \quad (1)$$

Deuteration of the tryptophan ring  $d_5$  carbon affects the relationship of  $\nu W3$  with the  $\chi^{2,1}$  torsion angle (8), and a similar equation can be written (1,7):

$$\nu Wd3 = 1514 + 7.5(\cos 3|\chi^{2,1}| + 1)^{1,2} \quad (2)$$

These equations are derived using the absolute value of the torsion angle,  $|\chi^{2,1}|$ , within a range of 60-120° (3). It is apparent that the angular dependence for  $\nu Wd3$  is greater than for  $\nu W3$  as the angular coefficient for  $\nu Wd3$  is larger.

In this work, we extend and modify the relationship,  $\nu W3(\chi^{2,1})$ , to cover a full 360° rotation since all dihedral angles are plotted in the same rotational direction. We show that the relationship is not single-valued---and therefore not a function---and has a three-fold periodicity over 360° due to the three-fold steric overlap of the indole residue with moieties bound to  $C_\alpha$ . Thus, a single  $\nu W3$  corresponds to as many as six dihedral angles. This follows intuitively from a consideration of a Newman diagram for dihedral steric clashes as discussed below.  $\nu W3-\chi^{2,1}$  data points for single tryptophan wild-type and mutant proteins, multiple tryptophan proteins where the W3 band has been resolved into components, fibrillar peptides (fd and Pf3 virion coat peptides) and two model tryptophan-containing peptides, the bee venom peptide, melittin, and the caged tryptophan peptide, exendin-4 TC5b, have been added to the original  $\nu W3$  plot for crystalline tryptophan analogs (1,7). This plotting suggests that three ranges of dihedral angle are poorly populated, attributable to steric interference with the protein backbone.

At the same time, the  $\nu W3-\chi^{2,1}$  data points for the proteins and model peptides show that the cosine-based relationship can only provide a rough guide to the tryptophan dihedral angle, even for tryptophans in a hydrophobic, constrained environment similar to that of the crystalline tryptophan derivatives used to derive the  $\nu W3-\chi^{2,1}$  relationship. For some of the peptides and proteins, the UV resonance-Raman determined  $\nu W3$ , coupled with a  $\chi^{2,1}$  dihedral angle taken from an x-ray crystal structure, constitutes a data point that lies off the cosine-dependent curve. In particular, we examine the influence of hydrogen bonding at the indole amine on  $\nu W3$  compliance with the cosine-dependent relationship through density function theory (DFT) calculations. For these calculations, we focus on the *S. cerevisiae* mutant enzyme, TIM Trp90Tyr Trp157Phe, bound to the ligand, 2-phosphoglycolate (PGA), because all other structural information gleaned from the UVRR data correlates with previously published

results for TIM. DFT calculations based on model systems are utilized to understand the deviation of TIM  $\nu W3$  from Eq. (1).

We also examine the weak, noncovalent cation- $\pi$  and anion-quadrupole interactions of proteinaceous tryptophans as the source of noncompliance with the  $\nu W3$ - $\chi^{2,1}$  relationship. Tryptophan dihedral angles for these proteins are taken from x-ray crystal or NMR structures available in the Protein Data Bank while  $\nu W3$  values are taken from the literature or have been measured. We compare the indole  $C_2=C_3$ ,  $C_\beta-C_3$ , and  $C_2-N_1$  bond lengths for the  $\nu W3$ - $\chi^{2,1}$  relationship-compliant versus noncompliant tryptophans. This information is also gleaned from structures in the Protein Data Bank, and is relevant because the  $\nu W3$  depends upon  $C_2=C_3$  and  $C_2-N_1$  stretching, and the  $C_\beta-C_3$  bond length has been found to vary inversely with  $|\chi^{2,1}|$  (1,2). Hydrophobicity of the tryptophan environment is indicated by measure of the W7  $1360\text{ cm}^{-1}$ : $1340\text{ cm}^{-1}$  band ratio (4,5). As the crystalline tryptophan derivatives used to formulate the  $\nu W3$ - $\chi^{2,1}$  relationship are expected to reside in a hydrophobic environment, the W7 ratio was examined for model peptides and relationship-compliant and noncompliant proteins.

The purpose of this analysis is to expose the specific tryptophan interactions and indole structural factors that contribute to deviations of  $\nu W3$ - $\chi^{2,1}$  data points from the relationship given in Eq. (1). We find that numerous weak electrostatic and nonpolar interactions accumulate to create an environment about the indole of tryptophan that is both constrained and hydrophobic, like the environment of the crystalline tryptophan derivatives from which the  $\nu W3$  -  $\chi^{2,1}$  relationship was derived. In the absence of methods to quantify these weak interactions, measurement of the FWHM bandwidth of the W3 band is suggested as a primary screen for evaluating the applicability of the  $\nu W3$  -  $\chi^{2,1}$  relationship.

## Materials and Methods

All reagents used were purchased from either Sigma-Aldrich Co. (St. Louis, MO) or Fluka (Milwaukee, WI) with the exception of glycerol-3-phosphate dehydrogenase, which was purchased from Boehringer Mannheim Ltd. (Indianapolis, IN). The model peptide, exendin-4 TC5b, was synthesized according to the protocol given in (10). *S. cerevisiae* TIM (Trp90Tyr Trp157Phe; hereafter referred to as the TIM mutant or simply TIM) was expressed and purified as described before (11). The kinetic parameters of the mutant are not distinguishable from those of the wild type TIM (12). Enzymatic activity was determined by the conversion of GAP to DHAP in the presence of TIM and glycerol 3-phosphate dehydrogenase as described by Putman and Knowles (13).

TIM enzyme samples employed in UV resonance Raman spectroscopy were at a concentration of 0.3 mM TIM in 50 mM Tris-HCl, 50 mM NaCl buffer, pH 7.8. Four samples were measured: unbound TIM and TIM bound to PGA (5 mM) in the aqueous buffer given above, and deuterium solutions of TIM bound to Pi (50 mM) and G3P (10 mM) with buffer conditions as above. Exendin-4 TC5b (1 mM) was spun cast in a silica gel matrix. Following gelation, the gel was bathed in 15 mM phosphate buffer, pH7.0. The method of silica gel preparation is given in (14). The melittin peptide was dissolved in 20 mM Tris buffer, pH 7.4 to a concentration of 2 mM.

## UV Resonance Raman Spectroscopy

All UVRR spectra were acquired with CW 229 nm excitation at an incident power of 1.8 mW. The UVRR instrument has been described elsewhere (15). For each sample, three acquisitions of three minutes, collected over an  $805$ - $1680\text{ cm}^{-1}$  frequency window, were summed. The sample-containing quartz tube was rastered vertically, spun and chilled to  $16^\circ\text{C}$  to minimize protein exposure to potentially denaturing UV radiation. UVRR peak positions were calibrated

against the known peaks of indene. The resulting calibration-limited accuracy is  $\pm 1 \text{ cm}^{-1}$ . Curvefitting analysis was carried out via the GRAMS/AI software package. All curvefits were carried out with the minimal number of Gaussian-Lorentzian peaks (optimally 50 -70 % Lorentzian) necessary to yield a  $\chi^2 < 3$ .

### Crystallographic Measurements

Dihedral angles, bond lengths, and intermolecular distances were measured from the indicated Protein Data Bank structures using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) (16).

### Raman Data Simulations

A series of calculations were performed using DFT-based simulations with Gaussian03W (17). A DFT approximation implementing the Becke's three-parameter exchange functional (18), in combination with the Lee, Yang, and Parr correlation (19) function (or B3LYP) was used. As a compromise between accuracy and applicability to large molecules, the 6-31G(d) basis set was used. Smaller basis sets are enough in DFT based calculations because the basis functions do not have to describe correlating orbitals (20,21). For all systems the geometry was first optimized before the frequency calculations. All calculated frequencies were scaled by a factor of 0.963 (20,22). GaussView (23) was used to visualize the atomic displacements correlating to each of the simulated Raman peaks. Using normal coordinate analysis, the displacement vectors associated with each different vibrational mode of skatole (6), a tryptophan analog, and skatole- $d^5$  (1) were previously assigned. To identify the W3 or Wd3 mode in the simulated Raman spectra, we compared the visualized atomic displacements to the displacement vector-W3 mode relationship for skatole and skatole- $d^5$ . A series of calculations were made to prove that the assignment was correct. First, a full geometry optimization and frequency calculation at fixed  $|\chi^{2,1}|$  torsion angles of 60.6 °, 78.0 °, 83.5 °, 88.0 °, 98.7 °, 105.1 ° and 116.5 ° were carried out. These  $\chi^{2,1}$  torsion angles were chosen because they correspond to the values used to generate Eq. (1). Second, we made a similar set of calculations using Trp- $d^5$  to generate a relationship similar to Eq. (2). Finally, we performed unconstrained simulations on a Trp molecule, a Trp- $d^5$ , and a Trp analog where the indole N had been replaced by an isotope,  $^{15}\text{N}$ . As defined, the W3 mode should have very minimal contribution from the indole N (2,6).

## Results and Discussion

Four 229 nm.-excited UVRR spectra for the *S. cerevisiae* mutant enzyme, TIM (Trp90Tyr Trp157Phe), in the 800 – 1550  $\text{cm}^{-1}$  range are shown in Figure 1. These include results for the ligand-free enzyme (Fig. 1a, trace 1), and the enzyme bound to the tight-binding transition state analog, PGA (Fig. 1a, trace 2). UVRR results for deuterium solution are also given for TIM ligated to G3P (Fig. 1b, trace 1) and phosphate (Pi)-ligated TIM (Fig. 1b, trace 2). The many peaks represent resonantly-enhanced vibrational modes for the several tyrosines in the TIM sequence, and most importantly, the single tryptophan residue, Trp-168, which is critically situated at the substrate binding loop hinge. The Fig. 1 inset shows a close-up view of the W3 bands for these spectra.

The UVRR data (Figure 1a) present a picture of the immediate environment of Trp-168 in TIM. The intensity ratio,  $I_{1360}/I_{1340}$ , of the W7 mode peaks has been linked to the hydrophobicity of the tryptophan environment, where a value less than one for the ratio corresponds to a hydrophilic environment, and conversely, a value greater than one indicates a hydrophobic one (4). The value of the  $\nu\text{W17}$  ( $876 \text{ cm}^{-1}$ ) for TIM in both the bound and unbound conformations indicates weak hydrogen bonding at the indole amine while the W6

band position at  $1420\text{ cm}^{-1}$  corresponds to a non-hydrogen bonded state (2). Like W3, the W1 mode, located at  $1620\text{ cm}^{-1}$ , depends upon indole ring stretching, but  $\nu\text{W1}$  is buried beneath the intense tyrosine Y8a/b bands at  $\sim 1614\text{ cm}^{-1}$  (not shown) and is therefore undecipherable. Concomitantly, the significance of tyrosine band changes accompanying binding loop motion---such as those for the Y9a mode at  $1173\text{ cm}^{-1}$ ---are difficult to interpret due to the presence of nine tyrosine residues in TIM. However, the importance of the Y208 residue to enzyme activity has been realized (12, 24).

### The W3 Mode ( $1542 - 60\text{ cm}^{-1}$ ) and the Dihedral Angle, $\chi^{2,1}$ for Trp-168 of TIM

While x-ray crystal studies of the bound and unbound binding loop conformers of TIM show that a backbone hydrogen bond network enforces a rigid loop movement between the two binding states (25), subsequent solid-state TIM NMR studies (11,26) and T-jump studies of solution phase TIM (27) show that Trp -168 has some local mobility. The Trp-168 W3 mode ( $\sim 1542 - 1560\text{ cm}^{-1}$ ) is expected to report on the local dihedral angle changes of Trp-168 that accompanying loop motion according to the  $\nu\text{W3}-\chi^{2,1}$  relationship discussed above. The x-ray crystal structure, 1YPI, of apo TIM with Trp-168 in the loop unbound position reveals the  $\chi^{2,1}$  torsion angle to be  $-79.65^\circ$  (28). When TIM is bound to the PGA transition state analog, x-ray crystallography reveals (PDB file: 7TIM), the  $\chi^{2,1}$  torsion angle of Trp-168 to be  $-99.61^\circ$  (25).

### The W3 Mode - Dihedral Angle, $\chi^{2,1}$ Relationship Extends over a Full Rotation

A problem arises when we try to reconcile the observed W3 frequencies and corresponding  $\chi^{2,1}$  torsion angles as derived from Eq. (1) (2) with that obtained from the x-ray crystal structures. Binding of the transition state analogue, PGA, to TIM downshifts the W3 mode from  $1547\text{ cm}^{-1}$  to  $1544\text{ cm}^{-1}$  (Fig. 1a, trace 4). According to Eq. (1), this  $\nu\text{W3}$  corresponds to a change in  $|\chi^{2,1}|$  from  $83^\circ$  to  $78^\circ$ , a  $-5^\circ$  dihedral angle change in going from the apo to holo enzyme conformation. The x-ray diffraction derived dihedral angle for apo TIM is  $-79.6^\circ$  (PDB: 1YPI; (28)), which is within the error bounds defined by Eq. (1). However, for PGA-bound TIM, the x-ray diffraction-obtained value for  $\chi^{2,1}$  is  $-99.6^\circ$  (PDB:7TIM; (25)), far from the value of  $\pm 77.6^\circ$  predicted by Eq. (1). The resulting apo-to-holo dihedral angle change predicted from the crystal structures is  $-20^\circ$ . While the magnitude of the dihedral angle change predicted from crystal structures is four-fold higher, the sign of the change is the same as for the UVRR data.

These discrepancies between  $\chi^{2,1}$  values predicted by the relationship given in Eq. (1) and those available from crystallographic structures suggest that plotting  $\nu\text{W3}$  against x-ray crystallographic dihedral angles for proteins would reveal the general validity of Eq. (1). This is of interest because the relationships reproduced in Eqs. (1) and (2) are often used to predict tryptophan dihedral angles from  $\nu\text{W3}$  results for proteins. Table 1 lists several tryptophan-bearing peptides and proteins with their corresponding  $\nu\text{W3}$  and crystal structure-derived  $\chi^{2,1}$  dihedral angles. According to the convention used in calculating these angles, values are reported within the range of  $\pm 0 - 180^\circ$ . These dihedral angle values have been recalculated so they can be plotted in the same rotational direction. Thus a negative value of  $\chi^{2,1}$  is recalculated as  $360^\circ + \chi^{2,1}$ . Furthermore, given the three-fold redundancy--- $3 \cos \chi^{2,1}$ ---of the relationship given in Eq. 1 over a full rotation, the angular values of  $1-120^\circ$ ,  $121-240^\circ$  and  $241-360^\circ$  are degenerate, and all dihedral angles have been reduced to a value between  $0-120^\circ$  (Table 1).

The reduced  $\chi^{2,1}$  dihedral angles given in Table 1 are plotted against the corresponding UVRR-determined  $\nu\text{W3}$  in Figure 2a (solid triangles). Also included are data points (open squares) for the crystalline tryptophan analogs from which Eq. 1 was derived (1), and data points (open diamonds) for model tryptophan-containing peptides, melittin conformers A and B (labeled 5 and , respectively) and the truncated exendin-4 TC5b W/P cage peptide (labeled 4). Data points

for Trp-bearing proteins that lie far from the fitted curve are also numbered as follows: 1. fd virion coat protein, 2. TIM mutant Trp90Tyr Trp157Phe ligated to PGA, 3. Pf3 virion coat protein. The relationship given by Eq. (1) has been modified (see discussion below) to fit the experimental data points given in Table 1, and is also plotted in Fig. 2a as a solid line. The overlay of the three degenerate  $\chi^{2,1}$  dihedral angle ranges means that data points for proteins (Table 1) define the relationship over one full rotation (Fig. 2a). This has not been previously realized. Due to the  $\chi^{2,1}$  angular degeneracy and the fact that the fit relationship is not single valued, one  $\nu W3$  can map to as many as six  $\chi^{2,1}$  dihedral angles, as can be seen from Fig. 2a. Therefore,  $\nu W3$  would seem to be a poor predictor of the tryptophan  $\chi^{2,1}$  dihedral angle. However, as discussed below, large ranges of dihedral angle values for tryptophan are generally avoided in protein structures, greatly reducing the possible number of dihedral angles for any one  $\nu W3$ .

The  $\nu W3$ - $\chi^{2,1}$  relationship was previously extended to deuterated tryptophan in UV Raman spectroscopic studies of deuterated, solid state tryptophan derivatives (1). The  $\nu Wd3$  - dihedral angle data pairs originally used to define this relationship are reproduced in Fig. 2b (solid circles). Additionally, a single data pair for deuterated, phosphate and G3P ligated TIM is shown (star) as the same dihedral angle as used for aqueous, ligated TIM (Table 1) is used here. A plot of a modified  $\nu Wd3$  - dihedral angle relationship, fit to the data pairs for deuterated tryptophans, is also shown in Fig. 2b (solid line). As for G3P-ligated TIM in aqueous buffer (Table 1 and Fig. 2a), the experimentally determined  $\nu Wd3$  does not place the corresponding crystallographic dihedral angle value squarely on the modified curve fit discussed below. Possible reasons for deviations of protein  $\nu W3$  and  $\nu Wd3$  - dihedral angle data points (Fig. 2 a, b) from the modified Eqs. (1) and (2) are discussed below.

### Modification of the $\nu Wd3$ - $\chi^{2,1}$ Relationship

In the first paper to discuss the  $\nu Wd3$ - $\chi^{2,1}$  relationship, the data was fit to an unspecified third order spline function (2). Subsequently, Eqs. (1) and (2) were introduced as empirical fits to the  $\nu W3$  and  $\nu Wd3$  - dihedral angle data points (1). Expansion of the relationship given in Eq. (1) to cover a full rotation, and using the same rotational sense for all dihedral angles,  $\chi^{2,1}$ , yields

$$\nu W3 = a + b(\cos 3\chi_{360}^{2,1} + 1)^{1,2} \quad (3)$$

Where for all negative values of dihedral angle,  $\chi^{2,1}$

$$\chi_{360}^{2,1} = 360^\circ + \chi^{2,1} \quad (4)$$

And  $a$  and  $b$  are empirically-determined parameters that are determined from fitting to the  $\nu W3$ - $\chi^{2,1}$  data pairs. The physical significance of the  $\cos 3\chi_{360}^{2,1}$  dependence is now discussed.

### Steric Hindrance result in Excluded Values for the Tryptophan $\chi^{2,1}$ Dihedral Angle

A circular plot of the crystallographically determined tryptophan dihedral angles for the proteins given in Table 1 reveals that certain ranges of dihedral angle are favored in protein structures. This plot is given in Figure 3.  $\chi_{360}^{2,1}$  is plotted clockwise and all dihedral angles are plotted as positive angles in the 0 - 360° range. Dihedral angle values for tryptophans found in proteins in water solution (triangles; see Table 1) and those for crystalline tryptophan

derivatives (diamonds) (2) are plotted. Dihedral angle values for two model peptides, melittin and exendin-4 TC5b (Table 1), are indicated by crosses. The three allowed tryptophan dihedral angles for an  $\alpha$  helix, as determined from a dataset of 61 globular proteins in the Protein Data Bank, are also plotted (stars) (29). Concentric with this plot is a Newman diagram looking down the tryptophan  $C_3$ --- $C_\beta$  bond axis. The  $C_3$ == $C_2$  double bond is part of the indole pentagonal ring as is the fourth bond of  $C_3$ . The  $\chi_{360}^{2,1}$  dihedral angle plot clearly shows that some angular ranges are preferred.  $C_\alpha$ , and the protein backbone extending away from it (not shown), are a steric barrier to the free rotation of the bulky tryptophanyl indole, whose plane is roughly perpendicular to that formed by the protein backbone. Dihedral angles that line up the indole plane with the  $C_\beta$ — $C_\alpha$  bond will likely lead to steric clashes as the bulky indole encounters atoms attached to the backbone and the backbone itself. These sterically hindered dihedral angle ranges appear to extend over the ranges of approximately  $0^\circ$  -  $75^\circ$ ,  $150^\circ$  -  $240^\circ$ , and  $300^\circ$  -  $360^\circ$  (Fig. 3). Complementary to these are two sterically 'favored' dihedral angle ranges:  $75^\circ$  -  $150^\circ$  and  $240^\circ$  -  $300^\circ$ , resulting in the elimination of possible dihedral angles for any given  $\nu W3$ . Likely dihedral angles for any one  $\nu W3$  are now reduced to one or two choices, as can be ascertained from the plot of the  $\nu W3$ - $\chi^{2,1}$  relationship in Fig. 2a, suggesting the continued utility of the  $\nu W3$  --  $\chi^{2,1}$  relationship as given by Eqs. (3) and (4).

Consideration of the Newman diagram makes plain the three fold cosine-dependency of the relationship given in Eq. (3). The highest  $\nu W3$  is encountered whenever the indole plane coincides with one of three  $C_\beta$  bonds or when  $\chi^{2,1}$  assumes values of  $0^\circ$ ,  $120^\circ$  or  $240^\circ$ . It is precisely at these angles that steric hindrance for the indole is greatest. Since the cosine depends on  $3\chi^{2,1}$ , the latter three angles become equivalent to  $360^\circ = 0^\circ$ , where the cosine value of these is unity, and the value of the relationship assumes its maximum value. The addition of one to the cosine  $3\chi^{2,1}$  value is necessary to prevent subtractions from the minimal  $\nu W3$  value given by the parameter,  $a$  (See Eq. (3)). The parameter,  $b$ , and the exponent value of 1.2 are empirically determined by fitting to the set of available experimental  $\nu W3$  -  $\chi^{2,1}$  data pairs.

### Exceptions to the Rule: Not all Protein Tryptophans Manifest the $\nu W3$ -- $\chi^{2,1}$ Relationship

Scrutiny of the  $\nu W3$  -  $\chi^{2,1}$  dihedral angle data pairs given in Table 1 and plotted in Fig. 2a shows that not all  $\nu W3$  for proteinaceous tryptophans conform to the  $\nu W3$ - $\chi^{2,1}$  relationship of Eq. (3). In particular, tryptophans found in the fibrillar viral coat proteins of fd (point 1, Fig. 2a) and Pf3 (point 3, Fig. 2a), are far from the curve defined by Eq. (3). The data point for the loop hinge Trp-168 in PGA-ligated TIM (point 2, Fig. 2a) also falls into this set. Even with an estimated uncertainty of  $\pm 10^\circ$  (1), these data pairs would not lie on the curve defined by Eq. (3). Indeed, the  $\nu W3$  for the fd virion coat protein ( $1560\text{ cm}^{-1}$ ) has a value that is outside the range of the  $\nu W3$ - $\chi^{2,1}$  relationship ( $1542$ - $1558\text{ cm}^{-1}$ ). Clearly, factors besides the  $\chi^{2,1}$  dihedral angle influence  $\nu W3$  in both fibrillar and globular proteins. We now consider several parameters as a suitable criterion for reliable application of the  $\nu W3$ - $\chi^{2,1}$  relationship.

### Indole Bond Lengths with Respect to the $\nu W3$ - $\chi^{2,1}$ Relationship

Maruyama and Takeuchi have shown that the bond length of  $C_2$ - $C_3$  for the indole pentagonal ring increases with increases in  $|\chi^{2,1}|$  for crystalline tryptophan derivatives while the bond length of  $C_\beta$ - $C_3$  decreases under these conditions (1). Taking this cue,  $C_2$ - $C_3$ ,  $C_2$ - $N_1$  and  $C_\beta$ - $C_3$  bond lengths of tryptophanyl indoles that greatly deviate from the  $3(\text{cosine})$  function were measured, and compared to bond lengths for tryptophans whose data pairs lie on the curve at the same dihedral angle (Fig. 2a) to look for trends in bond length with respect to data points that fell above or below the  $\nu W3$ - $\chi^{2,1}$  curve. Specifically, the bond lengths for the viral coat protein, fd (Fig. 2a, point 1), are compared to those for melittin W19B and the crystalline tryptophan derivative, *N*-acetyl-L-tryptophan methyl ester (Ac-L-TrpME), whose data points are on the  $\nu W3$ - $\chi^{2,1}$  curve at roughly the same dihedral angle (Fig. 2a). The bond lengths for the tryptophans in PGA ligated TIM (Fig. 2a, point 2) and the viral coat protein, Pf3 (Fig. 2a,

point 3) are compared to those for the tryptophan in gyrase B with clorobiocin, whose data point in Fig. 2a is just above the  $\nu W3-\chi^{2,1}$  curve at  $22^\circ$ ,  $1555\text{ cm}^{-1}$ . These bond lengths are given in Table 2. For this limited data set, the  $C_2-C_3$  bond lengths show no apparent trend for tryptophans whose data points deviate from the  $\nu W3-\chi^{2,1}$  relationship. The  $C_3-C_\beta$  bond length of the fd virion coat protein tryptophan is not significantly different from those of melittin or Ac-L-Trp ME while the  $C_3-C_\beta$  bond lengths of PGA-ligated TIM and Pf3 are shorter than those of gyrase B. The  $C_2-N_1$  bond lengths of fd and PGA-ligated TIM are not significantly different from those of curve compliant proteins. Only the  $C_2-N_1$  bond length of Pf3 shows any appreciable difference from that of curve compliant gyrase B. Indole bond lengths do not appear to be diagnostic of rule-breaking tryptophans. We next consider a specific case of disagreement between the crystallographic-determined tryptophan dihedral angle and that predicted by W3 band measurements, namely that for the PGA-ligated TIM mutant. DFT-based simulations are applied to explain the discrepancy with the expectation that these explanations may be generalized to other proteinaceous tryptophans that do not follow the  $\nu W3-\chi^{2,1}$  relationship.

### DFT Calculations Applied to the Discrepancy between the $\chi^{2,1}$ Torsion Angle Predicted by $\nu W3$ versus X-ray Crystallography for Ligated TIM

To explain the low  $\nu W3$  value obtained for PGA-ligated TIM (Fig. 2A, point 2), we first calculate the vibrational modes of skatole, a tryptophan analog, in order to verify the  $\nu W3-\chi^{2,1}$  relationship, and then proceed with molecular simulations of Trp-168 noncovalent interactions in the liganded and nonliganded forms of TIM. The results of the vibrational calculations indicate (Figure 4 and Table 3) that indeed we have assigned the correct simulated peak for the W3 mode. The simulated Raman data yield the following equation for  $\nu W3$  and  $\nu Wd3$ , respectively:

$$\nu W3 = (1540.2 \pm 1.1) + (6.245 \pm 0.853) (\cos 3 |\chi^{2,1}| + 1)^{1,2} \quad (5)$$

$$\nu Wd3 = (1510.1 \pm 1.1) + (6.041 \pm 0.806) (\cos 3 |\chi^{2,1}| + 1)^{1,2} \quad (6)$$

Where the errors associated with the intercept and coefficient values are the fitting errors. Not only have we reproduced Eqs. (1) and (2), but we have also shown that the isotopic substitution on the indole nitrogen had very little effect on the  $\nu W3$ . Values of  $1539.9\text{ cm}^{-1}$  for the natural Trp and  $1539.4\text{ cm}^{-1}$  for  $^{15}\text{N}$ -Trp were obtained (Table 4). On the other hand, the unconstrained calculations for Trp- $d^5$  yielded a  $\nu W3$  downshifted from that of the natural Trp by  $30\text{ cm}^{-1}$  (Table 4). This peak shift is consistent with what was previously reported (1). The simulated peak positions corresponding to the W6 and W7 tryptophan modes (Table 4) were also identified using the unconstrained calculations involving Trp,  $^{15}\text{N}$ -Trp, and Trp- $d^5$ .

### Noncovalent Interactions at the Trp-168 Indole - Hydrogen Bonding

With the assignment of the W3 mode ascertained, we attempt to reconcile the  $\nu W3$  and x-ray crystallographic dihedral angle obtained for Trp-168 in PGA-ligated TIM. We examine specifically the hydrogen bonding and around Trp-168 in TIM, followed by more generalized considerations of weaker electrostatic interactions, steric hindrance and hydrophobic interactions.

To simulate hydrogen bonding in PGA-ligated TIM, we carried out calculations involving tryptophan and three different hydrogen bonding partners. These are: 1. a phenol, to mimic the



interaction between Trp-168 and Tyr-164, 2. a carboxyl, to mimic the interaction of Trp-168 with Glu-129, and 3. a water molecule. The distances between the indole N–H and the hydrogen bonding partner were varied and fixed for each set of calculations. Varying the distance simulates the change in the extent of hydrogen bonding: the closer the distance, the greater is the hydrogen bonding. The angle of the hydrogen bond was set to 180°. No other constraints were used.

The set of Trp-H<sub>2</sub>O simulations indicates that as the distance between the water molecule and the indole N–H narrows from 4 to 2 Å, the values of  $\nu W3$  and  $\chi^{2,1}$  hardly change (Table 5), and are near the values obtained for the unconstrained Trp simulation (Table 3). Values of  $\nu W3$  are around 1536.7 cm<sup>-1</sup> while the  $\chi^{2,1}$  are about 42.4°. The upshift in  $\nu W6$  as the water is placed increasingly closer to the indole N-H group indicates that the extent of hydrogen bonding increases with closer approach (8). The insensitivity of  $\nu W3$  to hydrogen bond length indicates that the extent of hydrogen bonding cannot explain the observed shifts in  $\nu W3$ . Similar results were observed when either phenol or a carboxylic group was used as the hydrogen bonding partner. This set of observations, and the frequency positions of the hydrogen bond markers, W6 and W17 (876 cm<sup>-1</sup>, Fig. 1A) (2,8), rule out strong hydrogen bonding as an explanation for the misfit of the PGA-bound TIM  $\nu W3$  -  $\chi^{2,1}$  dihedral angle pair to Eq. (3). This supports the earlier finding of Maruyama and Takeuchi (1) that there is no direct correlation between  $\nu NH$  and  $\nu W3$  for the crystalline tryptophan derivatives.

### Noncovalent Interactions at the Trp-168 Indole: Anion-Quadrupole and Cation- $\pi$ Interactions

Here, we explore the possibility that weaker, attractive electrostatic interactions influence the W3 mode. The anion-quadrupole interaction involves the electrostatic attraction between an anion and the partial positive charge surrounding an aromatic ring, here the indole ring (30, 31) while the cation- $\pi$  interaction involves the attractive force between a cation and the  $\pi$  electron cloud on either face of the indole ring (32,33). Typically, the anion in a proteinaceous anion-quadrupole interaction is a glutamate or aspartate while the cation in proteinaceous cation- $\pi$  interactions is either arginine or lysine. As this weaker interaction is difficult to quantify in the presence of stronger hydrogen bond interactions, we look for the influence of these aromatic interactions on tryptophan discrepancies with the  $\nu W3$ - $\chi^{2,1}$  relationship by examining the x-ray crystal structures. The results for the proteinaceous tryptophans discussed above are given in Table 6. We see that for proteins with tryptophans that follow the  $\nu W3$ - $\chi^{2,1}$  relationship such as bacteriorhodopsin, gyrase B with cholobiocin and melittin, electrostatic interactions are common and multiple. It is noteworthy that for the enzymes, TIM and PTPase, the ligand state which best conforms to the  $\nu W3$ - $\chi^{2,1}$  relationship has more than one such electrostatic interaction. While weak, the cumulative effect of these interactions can be significant (31). On the other hand, the relationship-compliant tryptophan in the model peptide, exendin-4 TC5 with a caged tryptophan, has no such electrostatic interactions, while the noncompliant tryptophan in the Pf3 virion coat peptide does. As the evidence accumulates, it becomes clearer that no one factor is the bellwether for compliance with the  $\nu W3$ - $\chi^{2,1}$  relationship. Rather, compliance appears to be based on a complex web of weak interactions that are difficult to quantify individually.

### Noncovalent Interactions at the Trp-168 Indole: Hydrophobicity

The homogeneous composition of the crystalline tryptophan derivatives used to formulate the  $\nu W3$ - $\chi^{2,1}$  relationship would seem to indicate a hydrophobic environment for the constituent indole rings. Hydrophobicity can be spectroscopically estimated from the intensity ratio, 1360 cm<sup>-1</sup>:1340 cm<sup>-1</sup>, which are  $\nu W7$  bands, since a hydrophobic environment yields a band ratio greater than or approximately equal to one (4,5). Estimate of this simple intensity ratio for the crystalline tryptophan derivatives is not straightforward because the W7 bands are not split into doublets but into triplets, quartets or even more component bands (2). Estimate or

measurement of this W7 band ratio for the set of peptides and proteins under discussion here is somewhat more straightforward because W7 doublets are present or readily discernible from the UVRR results. W7 1360  $\text{cm}^{-1}$ :1340  $\text{cm}^{-1}$  band ratios are given in Table 7. The ratios range in value from 0.8 to 1.7, with the fd virion peptide taking the maximum value. As the W7 band ratio for the ligated forms of both enzymes, TIM and PTPase, and horse heart cytochrome C, is lowest in value, the least hydrophobic environment for their respective tryptophans is indicated. W7 band ratios for skatole in solvents of varying polarity show that a band ratio of 0.17 is obtained in dimethyl formamide and 1.0 in benzene (4). Even at a 1360  $\text{cm}^{-1}$ :1340  $\text{cm}^{-1}$  band ratio of 0.8, enzymatic tryptophans would seem to be in a fairly hydrophobic environment. The highest values in W7 band ratio are found for the virion coat peptides, fd and Pf3 (1.7 and 1.5, respectively), but the ratio for PGA-ligated TIM is 0.8, the minimum value found. As the  $\nu\text{W3} - \chi^{2,1}$  data points for these peptides and protein all do not fall on the  $\nu\text{W3} - \chi^{2,1}$  curve, hydrophobicity as measured by the W7 1360  $\text{cm}^{-1}$ :1340  $\text{cm}^{-1}$  band ratio is not a good indicator of compliance with the derived relationship. Consideration of the W7 band ratio for the relationship compliant model peptides and proteins leads to the same conclusion.

### Noncovalent Interactions at the Trp-168 Indole: Steric Hindrance

No one molecular interaction appears to account for the noncompliance of a given tryptophan with the  $\nu\text{W3} - \chi^{2,1}$  relationship. The sum total of all weak interactions will determine whether the  $\nu\text{W3} - \chi^{2,1}$  data point for any single tryptophan will follow the relationship. It should come as no surprise, therefore, that tryptophans with an environment similar to those of the crystalline tryptophan derivatives used to define the  $\nu\text{W3} - \chi^{2,1}$  relationship will comply with it; that is, tryptophans in a constrained environment.

### Tryptophans Compliant with the $\nu\text{W3} - \chi^{2,1}$ Relationship

NMR study of the tryptophan in the model truncated peptide, exendin-4 TC5b (point 4, Fig. 2a), shows that Trp-26 is encased by an extensive hydrophobic network of methyl groups and prolines on both faces of the indole and about its edge (34). The conformation of Trp-26 is also stabilized by hydrogen bonds both at the indole amine and backbone carbonyl and amine (34). This environment is very much like the environment of the crystalline tryptophan derivatives used to define the  $\nu\text{W3} - \chi^{2,1}$  relationship: hydrophobic, constrained and in most cases, with hydrogen bonds to either the indole amine or the backbone carbonyl or amine. The environment of Trps-19A and B of the model peptide, melittin (points 6 and 6, Fig. 2a), is also hydrophobic and constrained. The tryptophans are aligned on the apolar face of the bent rod of 26 amino acids (35). Four peptide rods pack together in a bilayer sandwich, two rods per layer, with the hydrophobic faces packed together. This arrangement persists in the venom sack of the insect and in solution where the positive face of each rod accounts for aqueous solubility and prevents protein aggregation (35). All subunit contacts are hydrophobic, with tight interhelical packing of valine, leucine, isoleucine and tryptophan residues. Lys-23 in the same and an adjacent helix stabilize Trp-19 with cation- $\pi$  interactions. Considering Trp-182 in retinal-bound bacteriorhodopsin, which also follows the  $\nu\text{W3} - \chi^{2,1}$  relationship (data point 36°, 1550  $\text{cm}^{-1}$ , Fig. 2a), the retinal cofactor both creates an hydrophobic environment for the indole ring and sterically hinders it via cation- $\pi$  interactions (Table 6) (36).

### Tryptophans Noncompliant with the $\nu\text{W3} - \chi^{2,1}$ Relationship

The fibrillar structure of the virion coat peptides, fd (point 1, Fig. 2a) and Pf3 (point 3, Fig. 2a), suggests a constrained, crystal-like, environment for their tryptophans. However, the interpretation of fiber x-ray structures of both of these coat proteins is clouded by the lack of resolution in the data (37). Welsh *et al.* (38) construct several possible structural models for Pf3 and Marvin and coworkers (37) refer to the Trp-26 dihedral angle values provided by Raman studies of fd in constructing their structure (39,40). In models of Pf3, Trp-38 is directed

to the central core of the fiber, which is occupied by viral DNA. There, multiple copies of Trp-38 are thought to help neutralize the negative charge of the DNA phosphates via anion-quadrupole electrostatic interactions, or by cation- $\pi$  interactions bridged by a cation (38). Several of the models for Trp-38 place it in a polar environment, which would help explain the noncompliance of Trp-38 with the  $\nu W3 - \chi^{2,1}$  relationship. An earlier x-ray structure for the fd virion coat protein provided a dihedral angle of  $290.1^\circ$  for Trp-26 (41). With  $\nu W3 = 1560 \text{ cm}^{-1}$ , the data point for Trp-26 would lie very far from the  $\nu W3 - \chi^{2,1}$  curve. Subsequent crystallographic study of fd (37) provides a dihedral angle of  $87^\circ$ , bringing the data point for Trp-26 closer to the  $\nu W3 - \chi^{2,1}$  curve (point 1, Fig. 2a). As this angle was determined by reference to the dihedral angle values provided by Raman studies of fd (39,40), there is uncertainty in the x-ray structural position of Trp-26. More to the point, however, is the value of  $\nu W3$  for Trp-26,  $1560 \text{ cm}^{-1}$ , which is outside the possible value range of the  $\nu W3 - \chi^{2,1}$  relationship given in Eq. (3), namely  $1542\text{--}1558 \text{ cm}^{-1}$ . This narrow value range---only  $16 \text{ cm}^{-1}$ ---suggests that  $\nu W3$  bandwidth could be a useful guide in empirically determining steric hindrance for tryptophan, and therefore the utility of the  $\nu W3 - \chi^{2,1}$  relationship for predicting the  $\chi^{2,1}$  dihedral angle.

### W3 Bandwidth as a Predictor of Compliance with the $\nu W3 - \chi^{2,1}$ Relationship

$\nu W3$  bandwidths for several of the peptide and protein data points shown in Fig. 2a are given in Table 7.  $\nu W3$  bandwidths have been equated with the FWHM value for each peak. As published spectra were only available for FWHM measurements in some cases (virion peptides and horse heart cytochrome C), the values given are only a first order approximation. The FWHM value was also estimated from published spectra for each of the crystalline tryptophan derivatives used to construct the  $\nu W3 - \chi^{2,1}$  relationship (2), and these are summarily reported as a range of values,  $9\text{--}15 \text{ cm}^{-1}$ . For these tryptophan derivatives, the widest  $\nu W3$  bandwidth would seem to lie just inside the maximum value of the  $\nu W3 - \chi^{2,1}$  relationship. Estimated FWHM values for relationship-compliant proteins, such as horse heart cytochrome C (42), fall just within the  $\nu W3 - \chi^{2,1}$  relationship bandwidth boundaries while those for noncompliant fd and Pf3 virion coat proteins (43) do not. More valuable are FWHM measurements that can be made on the original data, as for exendin-4 TC5b peptide and melittin. Here, the FWHM bandwidths are  $16 \text{ cm}^{-1}$  and  $15 \text{ cm}^{-1}$ , respectively (Table 7). Perhaps the most instructive measurements are those made on two different states of the same protein where one state follows the  $\nu W3 - \chi^{2,1}$  relationship and the other does not. This is true for the liganded and nonliganded states of the enzymes, TIM and PTPase (44). The FWHM data for these enzymes provide a nice contrast because in one case, for TIM, the nonliganded state is compliant with the  $\nu W3 - \chi^{2,1}$  relationship while for PTPase, the liganded state is relationship compliant. The FWHM measurements in Table 7 are in agreement with the  $\nu W3 - \chi^{2,1}$  plot results. Data points for enzymatic states with the narrower FWHM bandwidths (nonligated TIM, FWHM =  $13 \text{ cm}^{-1}$  and ligated PTPase, FWHM =  $11 \text{ cm}^{-1}$ ) follow the relationship while states with wider FWHM bandwidths (ligated TIM, FWHM =  $16 \text{ cm}^{-1}$  and nonligated PTPase, FWHM =  $17 \text{ cm}^{-1}$ ) do not. Thus, FWHM measurements of  $\nu W3$  provide an estimate of how constrained the tryptophan is with respect to the constrained environment of the crystalline tryptophan derivatives used to construct the  $\nu W3 - \chi^{2,1}$  relationship, and thus can provide a rough guide to the utility of the relationship in predicting the  $\chi^{2,1}$  dihedral angle for any one tryptophan. It should come as no surprise that constraint on the motion of the indole ring of tryptophan, as measured by the FWHM of  $\nu W3$ , is relevant here. As discussed above and illustrated by the Newman projection in Fig. 3, the  $3\cos$  dependence of  $\nu W3$  stems from the three-fold steric hindrance encountered by the indole from moieties bonded to  $C_\alpha$ . Clearly, this relationship would find great utility in Raman crystallographic studies where residue motion is minimized. Even with this spectroscopic guideline, the experimental value of  $\nu W3$  for the fd viral coat protein still lies outside the boundaries of the  $\nu W3 - \chi^{2,1}$  relationship. Science is an ongoing

process, and undoubtedly refinements to this relationship will be made by others, advancing the utility of Raman spectroscopy for making structural predictions.

## Conclusion

The  $\nu W3 - \chi^{2,1}$  relationship (1-3) has been extended to a full,  $360^\circ$  rotation by plotting all dihedral angles in the same rotational direction, i.e., clockwise. As the period of the relationship given in Eq. (3) is three-fold ( $3\cos \chi^{2,1}$ ) the  $\chi^{2,1}$  dihedral angle ranges  $0^\circ - 120^\circ$ ,  $120^\circ - 240^\circ$ ,  $240^\circ - 360^\circ$ , are superimposable. It is clear from this plotting that as many as six dihedral angles correlate to a single  $\nu W3$ . A Newman plot of tryptophan dihedral angles gleaned from protein and tryptophan analog crystallographic data shows, however, that the  $\chi^{2,1}$  dihedral angle ranges,  $0^\circ - 75^\circ$ ,  $150^\circ - 240^\circ$ , and  $300^\circ - 360^\circ$ , are generally avoided because of steric clashes between the indole rings and the protein backbone. This elimination of dihedral angle ranges reduces the number of possible dihedral angles for any  $\nu W3$  to two or even a single angle, suggesting a general utility for the  $\nu W3 - \chi^{2,1}$  relationship. At the same time,  $\nu W3 - \chi^{2,1}$  data points for some proteinaceous tryptophans do not conform to the relationship given in Eq. (3). For one of these proteins, the mutant enzyme TIM (Trp90Tyr Trp157Phe), DFT-based calculations and simulations are used to explore the possibility that hydrogen bonding, a strong, noncovalent interaction, at the indole amine is responsible for noncompliance of the ligated TIM  $\nu W3 - \chi^{2,1}$  data point with the  $\nu W3 - \chi^{2,1}$  relationship. In agreement with Maruyama and Takeuchi (1), our simulation results show no dependence of  $\nu W3$  on hydrogen bonding at the indole amine. The  $\chi^{2,1}$  dependence on  $C_2-C_3$ ,  $C_2-N_1$  and  $C_\beta-C_3$  bond lengths of tryptophanyl indoles, suggested earlier (1), is not found here for proteinaceous tryptophans that obey the  $\nu W3 - \chi^{2,1}$  relationship. The effect of weaker anion-quadrupole and cation- $\pi$  interactions on  $\nu W3$  is evaluated through examination of x-ray crystal structures for several proteins for which  $\nu W3$  is available. This evaluation suggests that multiple, weak electrostatic anion-quadrupole, cation- $\pi$ , and aromatic face-edge interactions coupled with van der Waals interactions act in a cumulative fashion to stabilize the indole side chain of tryptophan. In the absence of methods to quantify or model the cumulative effect of these interactions, the environment about the crystalline tryptophan derivatives used to formulate the  $\nu W3 - \chi^{2,1}$  relationship is considered. That environment is both constrained and hydrophobic. The hydrophobicity of the tryptophan environment can be evaluated from the W7 band ratio,  $1360 \text{ cm}^{-1}:1340 \text{ cm}^{-1}$ , where a value greater than one indicates a hydrophobic environment (5), but band ratio measurements for the set of peptides and proteins considered here showed no correlation with the  $\nu W3 - \chi^{2,1}$  relationship. Estimation of environmental constraint for tryptophan as measured by FWHM of the W3 band showed greater success in predicting compliance with the  $\nu W3 - \chi^{2,1}$  relationship. For model peptides where the indole is stabilized through numerous noncovalent interactions and other relationship-compliant proteins, the  $\nu W3 - \chi^{2,1}$  data point conforms to the  $\nu W3 - \chi^{2,1}$  relationship. As the  $\nu W3$  bandwidth of the  $\nu W3 - \chi^{2,1}$  relationship is only  $16 \text{ cm}^{-1}$ , the FWHM of  $\nu W3$  for the tryptophan under study should be no greater than this. Given this constraint, application of  $\nu W3 - \chi^{2,1}$  relationship for predicting the tryptophan  $\chi^{2,1}$  angle has some chance of success. Clearly, where a spectral bandwidth nearly matches the full range of the relationship used to predict a structural parameters, the error bars are large.

## References

1. Maruyama T, Takeuchi H. Effects of hydrogen bonding and side-chain conformation on the Raman bands of tryptophan-2,4,5,6-d5. *J Raman Spectrosc* 1995;26:319–324.
2. Miura T, Takeuchi H, Harada I. Tryptophan Raman bands sensitive to hydrogen bonding and side-chain conformation. *J Raman Spectrosc* 1989;20:667–671.
3. Takeuchi H. Raman structural markers of tryptophan and histidine side chains in proteins. *Biopolymers* 2003;72:305–317. [PubMed: 12949821]

4. Harada I, Miura T, Takeuchi H. Origin of the doublet at 1360 and 1340 cm<sup>-1</sup> in the Raman spectra of tryptophan and related compounds. *Spectrochim Acta, Part A* 1986;42:307–312.
5. Miura T, Takeuchi H, Harada I. Characterization of individual tryptophan side chains in proteins using Raman spectroscopy and hydrogen-deuterium exchange kinetics. *Biochemistry* 1988;27:88–94. [PubMed: 3349046]
6. Takeuchi H, Harada I. Normal coordinate analysis of the indole ring. *Spectrochim Acta, Part A* 1986;42:1069–1078.
7. Maruyama T, Takeuchi H. Water accessibility to the tryptophan indole N-H sites of gramicidin a transmembrane channel: detection of positional shifts of tryptophans 11 and 13 along the channel axis upon cation binding. *Biochemistry* 1997;36:10993–11001. [PubMed: 9283091]
8. Miura T, Takeuchi H, Harada I. Raman spectroscopic characterization of tryptophan side chains in lysozyme bound to inhibitors: role of the hydrophobic box in the enzymatic function. *Biochemistry* 1991;30:6074–6080. [PubMed: 1646007]
9. Spiro TG, Grygon CA. Applications of ultraviolet resonance Raman spectroscopy to proteins. *J Mol Struct* 1988;173:79–90.
10. Juszczak L. Comparative vibrational spectroscopy of intracellular tau and extracellular collagen I reveals parallels of gelation and fibrillar structure. *J Biol Chem* 2004;279:7395–7404. [PubMed: 14660656]
11. Rozovsky S, McDermott AE. The time scale of the catalytic loop motion in triosephosphate isomerase. *J Mol Biol* 2001;310:259–270. [PubMed: 11419951]
12. Sampson NS, Knowles JR. Segmental movement: definition of the structural requirements for loop closure in catalysis by triosephosphate isomerase. *Biochemistry* 1992;31:8482–8487. [PubMed: 1390632]
13. Putman SJ, Coulson AF, Farley IR, Riddleston B, Knowles JR. Specificity and kinetics of triose phosphate isomerase from chicken muscle. *Biochem J* 1972;129:301–310. [PubMed: 4643318]
14. Ellerby LM, Nishida CR, Nishida F, Yamanaka SA, Dunn B, Valentine JS, Zink JI. Encapsulation of proteins in transparent porous silicate glasses prepared by the sol-gel method. *Science (New York, N Y)* 1992;255:1113–1115.
15. Juszczak LJ, Hirsch RH, Nagel RL, Friedman JM. Conformational differences in CO derivatives of HbA, HbC(βE6K) and HbS(βE6V) in the presence and absence of inositol hexaphosphate (IHP) detected using ultraviolet resonance Raman spectroscopy. *J Raman Spectrosc* 1998;29:963–968.
16. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–1612. [PubMed: 15264254]
17. Frisch, MJ.; T, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Montgomery, JJA.; Vreven, T.; Kudin, KN.; Burant, JC.; Millam, JM.; Iyengar, SS.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, GA.; 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, JE.; Hratchian, HP.; Cross, JB.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, RE.; Yazyev, O.; Austin, AJ.; Cammi, R.; Pomelli, C.; Ochterski, JW.; Ayala, PY.; Morokuma, K.; Voth, GA.; Salvador, P.; Dannenberg, JJ.; Zakrzewski, VG.; Dapprich, S.; Daniels, AD.; Strain, MC.; Farkas, O.; Malick, DK.; Rabuck, AD.; Raghavachari, K.; Foresman, JB.; Ortiz, JV.; Cui, Q.; Baboul, AG.; Clifford, S.; Cioslowski, J.; Stefanov, BB.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, RL.; Fox, DJ.; Keith, T.; Al-Laham, MA.; Peng, CY.; Nanayakkara, A.; Challacombe, M.; Gill, PMW.; Johnson, B.; Chen, W.; Wong, MW.; Gonzalez, C.; Pople, JA. Gaussian 03. Gaussian Inc.; Wallingford, CT: 2004.
18. Becke AD. Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 1993;98:5648–5652.
19. Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a function of the electron density. *Phys Rev B* 1988;41:785–789.
20. Rauhut G, Pulay P. Transferable scaling factors for density functional derived vibrational force fields. *J Phys Chem* 1995;99:3093–3100.

21. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. Ab initio calculations of vibrational absorption and circular dichroism spectra using density functional force fields. *J Phys Chem* 1994;98:11623–11627.
22. Pulay P, Fogarasi G, Pongor G, Boggs JE, Vargha A. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled quantum mechanical (QM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene. *J Am Chem Soc* 1983;105:7037–7047.
23. Dennington, RKT., II; Millam, J.; Eppinnett, K.; Hovell, WL.; Gilliland, R. GaussView. Shawnee Mission, KS: 2003.
24. Sampson NS, Knowles JR. Segmental motion in catalysis: investigation of a hydrogen bond critical for loop closure in the reaction of triosephosphate isomerase. *Biochemistry* 1992;31:8488–8494. [PubMed: 1390633]
25. Lolis E, Petsko GA. Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-Å resolution: implications for catalysis. *Biochemistry* 1990;29:6619–6625. [PubMed: 2204418]
26. Rozovsky S, Jogl G, Tong L, McDermott AE. Solution-state NMR investigations of triosephosphate isomerase active site loop motion: ligand release in relation to active site loop dynamics. *J Mol Biol* 2001;310:271–280. [PubMed: 11419952]
27. Desamero R, Rozovsky S, Zhadin N, McDermott A, Callender R. Active site loop motion in triosephosphate isomerase: T-jump relaxation spectroscopy of thermal activation. *Biochemistry* 2003;42:2941–2951. [PubMed: 12627960]
28. Lolis E, Alber T, Davenport RC, Rose D, Hartman FC, Petsko GA. Structure of yeast triosephosphate isomerase at 1.9-Å resolution. *Biochemistry* 1990;29:6609–6618. [PubMed: 2204417]
29. McGregor MJ, Islam SA, Sternberg MJ. Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. *J Mol Biol* 1987;198:295–310. [PubMed: 3430610]
30. Burely SK, Petsko GA. Weak polar interactions in proteins. *Adv Protein Chem* 1988;39:125–189. [PubMed: 3072867]
31. Jackson MR, Beahm R, Duvvuru S, Narasimhan C, Wu J, Wang HN, Philip VM, Hinde RJ, Howell EE. A preference for edgewise interactions between aromatic rings and carboxylate anions: the biological relevance of anion-quadrupole interactions. *J Phys Chem B* 2007;111:8242–8249. [PubMed: 17580852]
32. Gallivan JP, Dougerty DA. Cation- $\pi$  interactions in structural biology. *Proc Nat Acad Sci U S A* 1999;96:9459–9464.
33. Ma JC, Dougherty DA. The cation- $\pi$  interaction. *Chem Rev* 1997;97:1303–1324. [PubMed: 11851453]
34. Neidigh JW, Fesinmeyer RM, Prickett KS, Andersen NH. Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in the solution and micelle-associated states. *Biochemistry* 2001;40:13188–13200. [PubMed: 11683627]
35. Terwilliger TC, Eisenberg D. The structure of melittin. II. Interpretation of the structure. *J Biol Chem* 1982;257:6016–6022. [PubMed: 7076662]
36. Hashimoto S, Obata K, Takeuchi H, Needleman R, Lanyi JK. Ultraviolet resonance Raman spectra of Trp-182 and Trp-189 in bacteriorhodopsin: novel information on the structure of Trp-182 and its steric interaction with retinal. *Biochemistry* 1997;36:11583–11590. [PubMed: 9305948]
37. Marvin DA, Welsh LC, Symmons MF, Scott WR, Straus SK. Molecular structure of fd (f1, M13) filamentous bacteriophage refined with respect to X-ray fibre diffraction and solid-state NMR data supports specific models of phage assembly at the bacterial membrane. *J Mol Biol* 2006;355:294–309. [PubMed: 16300790]
38. Welsh LC, Symmons MF, Sturtevant JM, Marvin DA, Perham RN. Structure of the capsid of Pf3 filamentous phage determined from x-ray fibre diffraction data at 3.1 Å resolution. *J Mol Biol* 1998;283:155–177. [PubMed: 9761681]
39. Blanch EW, Hecht L, Day LA, Pederson DM, Barron LD. Tryptophan absolute stereochemistry in viral coat proteins from Raman optical activity. *J Am Chem Soc* 2001;123:4863–4864. [PubMed: 11457308]

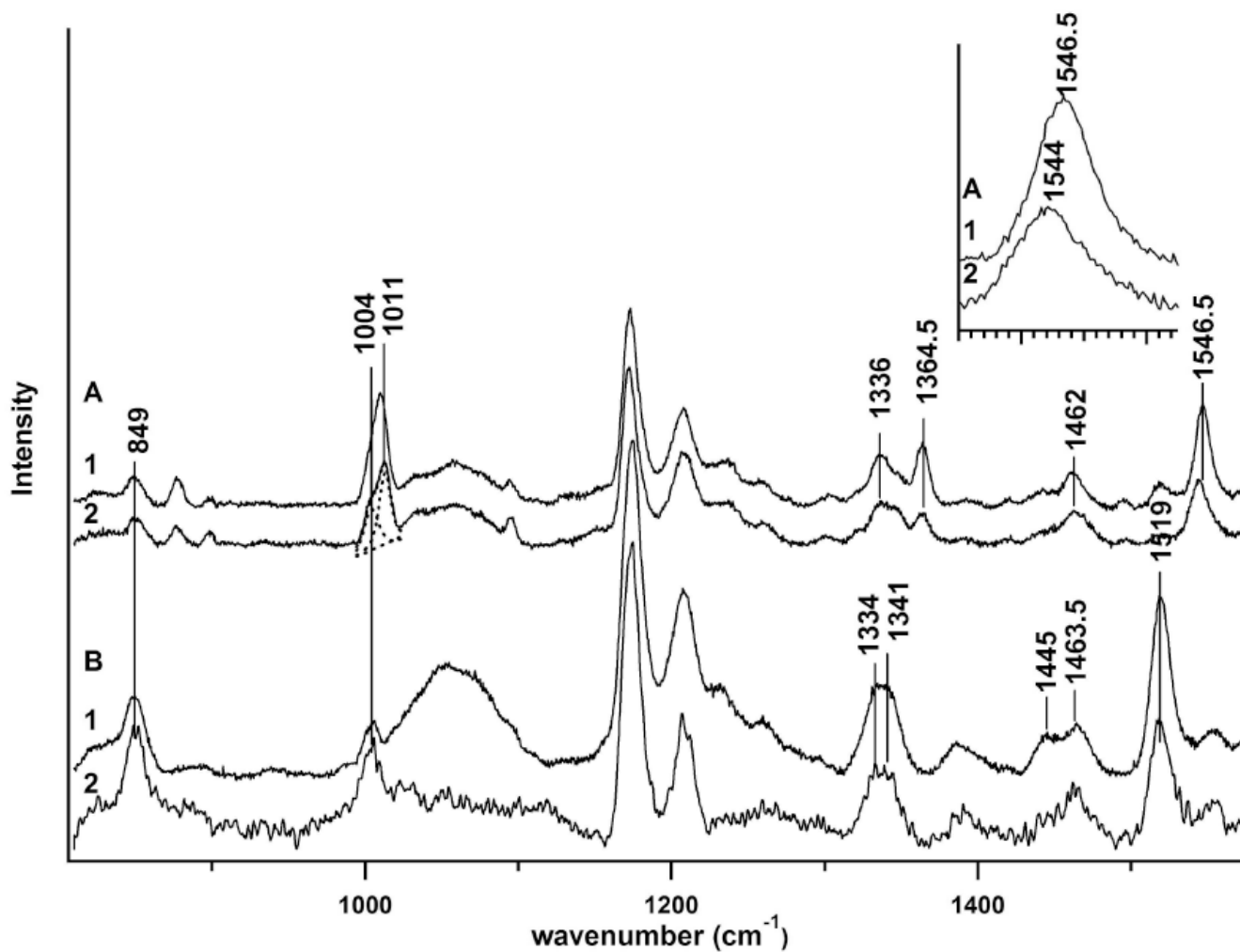
40. Tsuboi M, Overman SA, Thomas GJ Jr. Orientation of tryptophan-26 in coat protein subunits of the filamentous virus Ff by polarized Raman microspectroscopy. *Biochemistry* 1996;35:10403–10410. [PubMed: 8756696]
41. Marvin DA, Hale RD, Nave C, Citterich MH. Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages: Ff (fd, f1, M13), If1 and IKe. *J Mol Biol* 1994;235:260–286. [PubMed: 8289247]
42. Jordan T, Eads JC, Spiro TG. Secondary and tertiary structure of the A-state of cytochrome c from resonance Raman spectroscopy. *Protein Sci* 1995;4:716–728. [PubMed: 7613469]
43. Wen ZQ, Thomas GJ Jr. Ultraviolet-resonance Raman spectroscopy of the filamentous virus pf3: interactions of Trp 38 specific to the assembled virion subunit. *Biochemistry* 2000;39:146–152. [PubMed: 10625489]
44. Juszczak LJ, Zhang ZY, Wu L, Gottfried DS, Eads DD. Rapid loop dynamics of Yersinia protein tyrosine phosphatases. *Biochemistry* 1997;36:2227–2236. [PubMed: 9047324]
45. Austin JC, Wharton CW, Hester RE. An ultraviolet resonance Raman study of dehydrogenase enzymes and their interactions with coenzymes and substrates. *Biochemistry* 1989;28:1533–1538. [PubMed: 2655694]
46. Couling VW, Fischer P, Klenerman D, Huber W. Ultraviolet resonance Raman study of drug binding in dihydrofolate reductase, gyrase, and catechol O-methyltransferase. *Biophys J* 1998;75:1097–1106. [PubMed: 9675211]
47. Rodgers K, Su C, Subramanian S, Spiro T. Hemoglobin R --->T structural dynamics from simultaneous monitoring of tyrosine and tryptophan time-resolved UV resonance Raman signals. *J Am Chem Soc* 1992;114:3697–3709.
48. Skarzynski T, Moody PC, Wonacott AJ. Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å resolution. *J Mol Biol* 1987;193:171–187. [PubMed: 3586018]
49. Lamour V, Hoermann L, Jeltsch JM, Oudet P, Moras D. Crystallization of the 43 kDa ATPase domain of *Thermus thermophilus* gyrase B in complex with novobiocin. *Acta Crystallogr, Sect D: Biol Crystallogr* 2002;58:1376–1378. [PubMed: 12136161]
50. Lafitte D, Lamour V, Tsvetkov PO, Makarov AA, Klich M, Deprez P, Moras D, Briand C, Gilli R. DNA gyrase interaction with coumarin-based inhibitors: the role of the hydroxybenzoate isopentenyl moiety and the 5'-methyl group of the noviose. *Biochemistry* 2002;41:7217–7223. [PubMed: 12044152]
51. Grigorieff N, Ceska TA, Downing KH, Baldwin JM, Henderson R. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J Mol Biol* 1996;259:393–421. [PubMed: 8676377]
52. Stuckey JA, Schubert HL, Fauman EB, Zhang ZY, Dixon JE, Saper MA. Crystal structure of Yersinia protein tyrosine phosphatase at 2.5 Å and the complex with tungstate. *Nature* 1994;370:571–575. [PubMed: 8052312]
53. Fauman EB, Yuvaniyama C, Schubert HL, Stuckey JA, Saper MA. The X-ray crystal structures of Yersinia tyrosine phosphatase with bound tungstate and nitrate. Mechanistic implications. *J Biol Chem* 1996;271:18780–18788. [PubMed: 8702535]
54. Safo MK, Burnett JC, Musayev FN, Nokuri S, Abraham DJ. Structure of human carbonmonoxyhemoglobin at 2.16 Å: a snapshot of the allosteric transition. *Acta crystallographica. Section D, Biological crystallography* 2002;58:2031–2037.
55. Fermi G, Perutz MF, Shaanan B, Fourme R. The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. *J Mol Biol* 1984;175:159–174. [PubMed: 6726807]
56. Bushnell GW, Louie GV, Brayer GD. High-resolution three-dimensional structure of horse heart cytochrome c. *J Mol Biol* 1990;214:585–595. [PubMed: 2166170]
57. Denu JM, Lohse DL, Vijayalakshmi J, Saper MA, Dixon JE. Visualization of intermediate and transition-state structures in protein-tyrosine phosphatase catalysis. *Proc Natl Acad Sci U S A* 1996;93:2493–2498. [PubMed: 8637902]

## Abbreviations

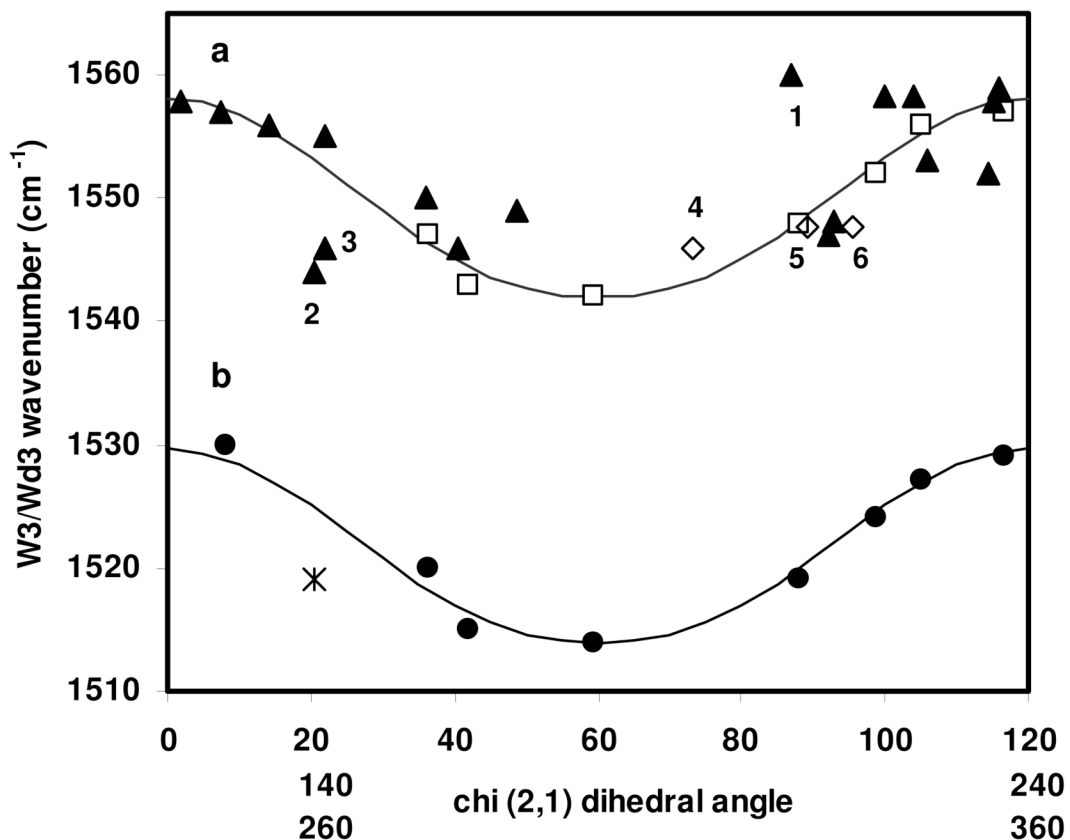
**DFT** density functional theory

<b>DHAP</b>	dihydroxy acetone phosphate
<b>FWHM</b>	full width half maximum
<b>GAP</b>	glyceraldehyde phosphate
<b>G3P</b>	glycerol 3-phosphate
<b>PTPase</b>	protein tyrosine phosphatase
<b>PGA</b>	2-phosphoglycolate
<b>Pi</b>	phosphate
<b>PDB</b>	Protein Data Bank
<b>TIM</b>	triose phosphate isomerase
<b>UVRR</b>	ultraviolet resonance Raman spectroscopy



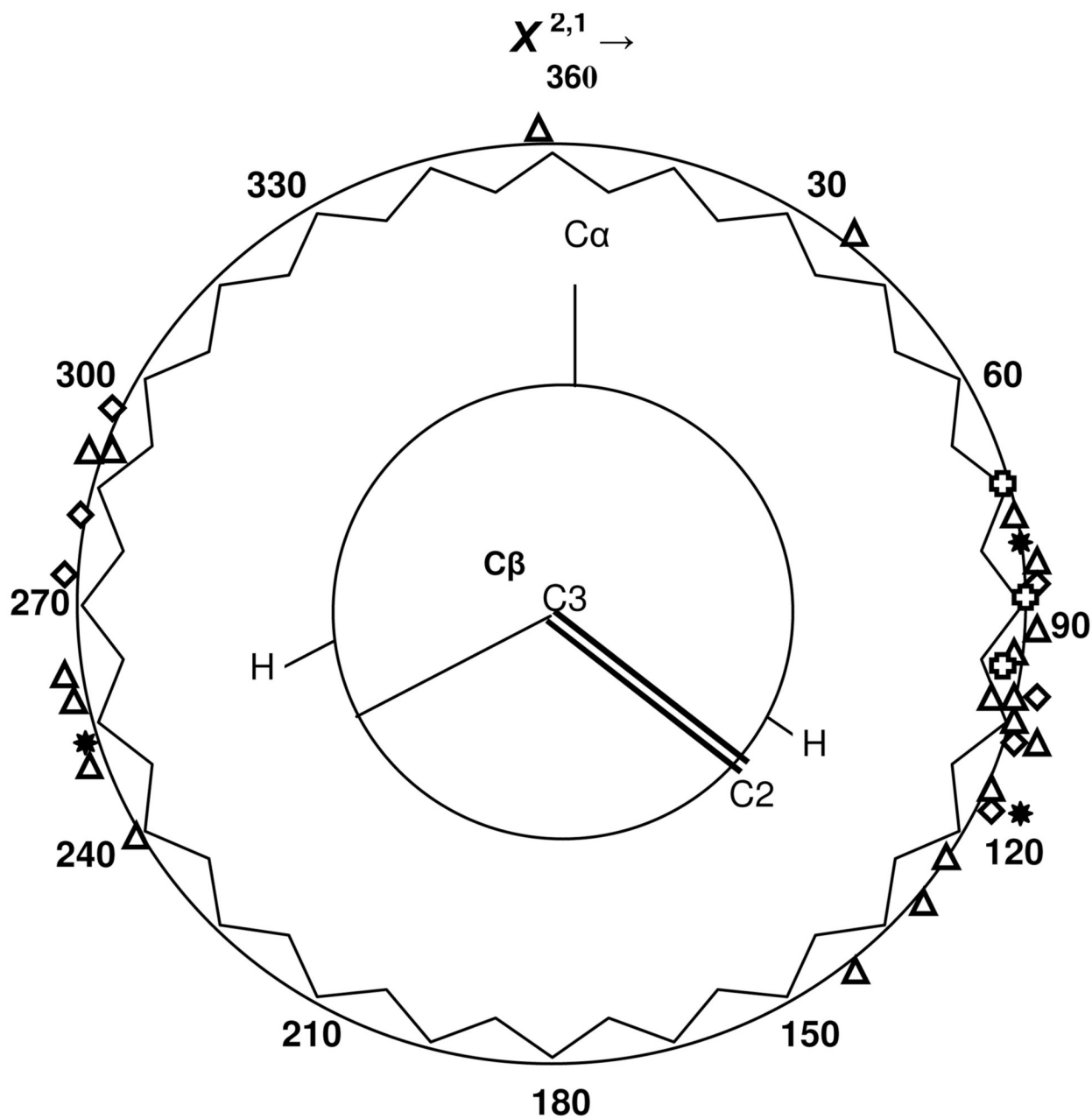


**Figure 1.** UVRR spectra for TIM in aqueous and deuterium solution with and without ligands. **a.** Aqueous solution. 1. ApoTIM 2. TIM ligated to PGA. Inset displays the W3 mode for apoTIM and PGA bound TIM. **b.** Deuterium solution 1. TIM ligated to G3P 2. TIM ligated to phosphate.



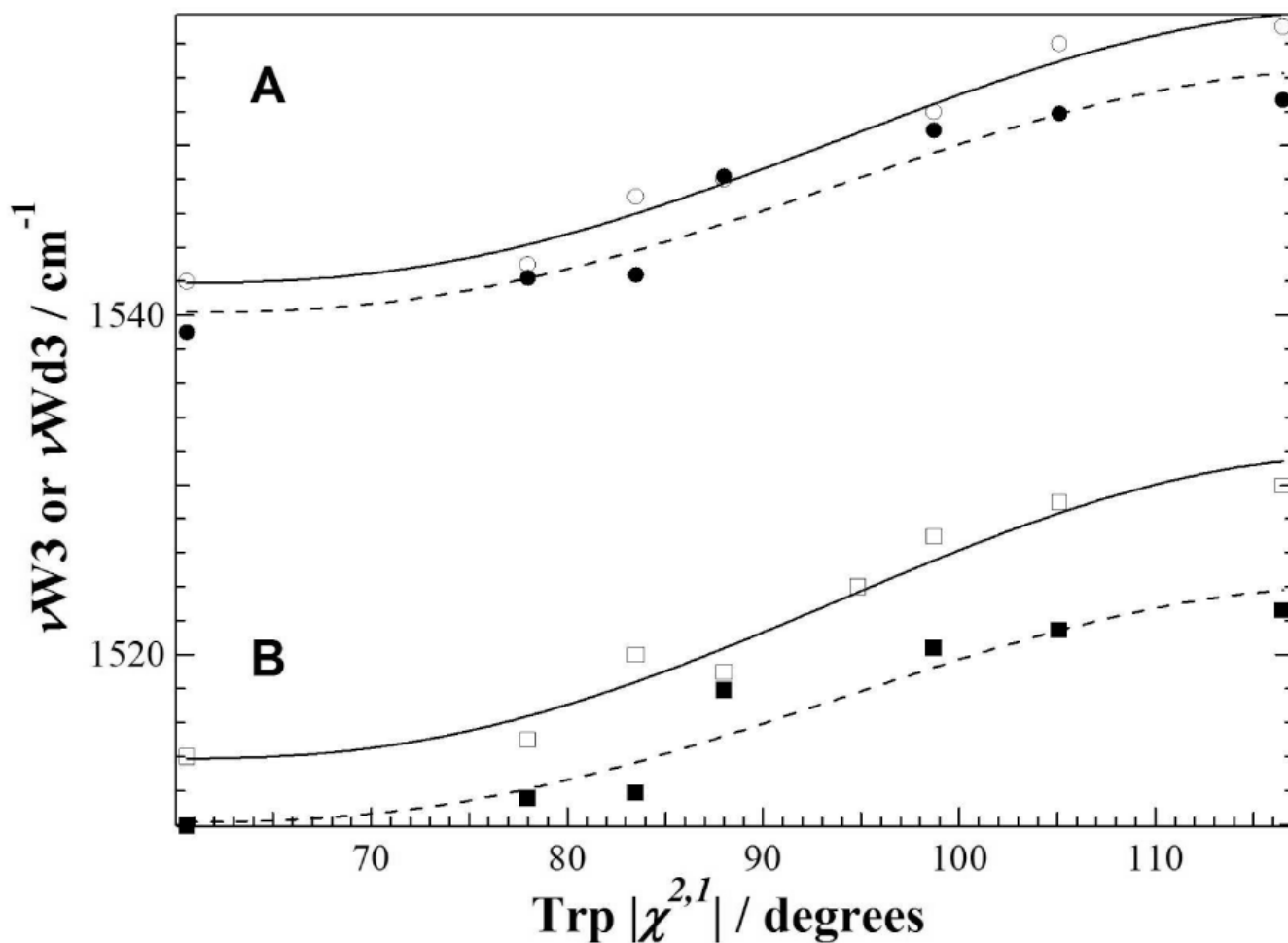
**Figure 2.**

Relationship between  $\nu W3$  and the dihedral angle over a full rotation. **a.** Data points are for crystalline tryptophan derivatives (open squares) (2), proteinaceous tryptophan residues (solid triangles, see Table 1), and model peptides (open diamonds, see Table 1) for which  $\nu W3$  is available (See Table 1).  $\chi^2_{360^\circ}$  is given by Eq. (3) where  $a = 1542 \text{ cm}^{-1}$  and  $b = 7.0$ . As in reference (1), these parameters are empirically adjusted. Numbered points correspond to: 1. fd virion coat protein, 2. TIM mutant Trp90Tyr Trp157Phe ligated to PGA, 3. Pf3 virion coat protein, 4. the truncated exendin-4 TC5b peptide, 5, 6. melittin, two data points corresponding to dihedral angles W19A (point 6) and W19B (point 5) from crystal structure, 2MLT. **b.** Data points are for deuterated tryptophan derivatives (solid circles) (1), and the Pi/G3P ligated, deuterated TIM proteins (cross), which share the ligated, aqueous TIM dihedral angle given in Table 1.  $\nu Wd3$  taken from Fig. 1. The curve fit to this data is given by Eq. (3) where  $a = 1514 \text{ cm}^{-1}$  and  $b = 6.8$ . As in reference (1), these parameters are empirically adjusted.



**Figure 3.**

Newman plot of tryptophan  $\chi^{2,1}$  dihedral angles for proteinaceous tryptophans (triangles) and model peptides (crosses), as given in Table 1, and for crystalline tryptophan derivatives (diamonds) from (2). The three allowed (stars) tryptophan dihedral angles for an  $\alpha$  helix, as determined from a dataset of 61 globular proteins in the Protein Data Bank, are also plotted (29). All negative dihedral angles are plotted as  $\chi^{2,1}_{360^\circ} = 360^\circ + \chi^{2,1}$ . Concentric with the plot is a Newman diagram showing the disposition of the tryptophan indole  $C_2=C_3$  bond with respect to the  $C_\beta-C_\alpha$  bond along the  $C_\alpha-C_\beta-C_3=C_2$  linkage.



**Figure 4.**

Simulated  $\nu W3 = (1540.2 \pm 1.1) + (6.245 \pm 0.853)(\cos 3|\chi^{2,1}| + 1)^{1,2}$  (Table 2,  $\nu W3_{\text{calc scaled}}$ ; broken line trace and solid circles) and  $\nu Wd3 = (1510.1 \pm 1.1) + (6.041 \pm 0.806)(\cos 3|\chi^{2,1}| + 1)^{1,2}$  (Table 2,  $\nu Wd3_{\text{calc scaled}}$ ; broken line trace and solid squares) as a function of the  $\chi^{2,1}$  torsional angles and where the errors associated with the intercept and coefficient values are fitting errors. Results of the simulations are very similar to the published data on the  $\nu W3$  (solid line trace and open circles, from (2)) and the  $\nu Wd3$  (solid line trace and open squares, from (1)). All frequencies reported are calculated using DFT/B3LYP and 6-31g(d) with a scaling factor = 0.963 (20,22).

Table 1

W3 mode frequencies and  $\chi^{2,l}$  torsional angle for Tryptophan in Peptides and Proteins

Protein System/Trp No.	Trp residue	$\nu W3$ (cm <sup>-1</sup> )	$\chi^{2,l}$ , degree	Reduced $\chi^{2,l}$ (0 - 120°)	PDB #
glyceraldehyde 3-phosphate dehydrogenase	84	1557 <sup>d</sup>	127.61	7.6	1GDJ <sup>j</sup>
Gyrase B with Novacin	128	1556 <sup>b</sup>	-106.13	14	1KIJ <sup>k</sup>
Gyrase B with chlorobiocin	170	1555 <sup>b</sup>	-98.1	22	1KZN <sup>l</sup>
Bacteriorhodopsin	182 189	1550 <sup>c</sup> 1553 <sup>c</sup>	36 106	36 106	2BRD <sup>m</sup>
PF3 virion coat protein	38	1546 <sup>d</sup>	142	22	1IFP <sup>n</sup>
fd virion coat protein	26	1560 <sup>e</sup>	87	87	2C0X <sup>o</sup>
apo TIM mutant	168	1546 <sup>f</sup>	-79.6	40.4	1YPI <sup>p</sup>
PGA-TIM mutant	168	1544 <sup>f</sup>	-99.6	20.4	7TIM <sup>q</sup>
apoPTPase	354	1549 <sup>g</sup>	-71.2	48.8	1YPT <sup>r</sup>
Tungstate PTPase		1559 <sup>g</sup>	-4	116	1YTW <sup>s</sup>
COHb	$\beta$ 37 $\beta$ 15 $\alpha$ 14	1547 <sup>h</sup> 1558 <sup>h</sup> 1558 <sup>h</sup>	92 <sup>i</sup> 115 <sup>i</sup> 122 <sup>i</sup>	92 115 2	1LJW <sup>t</sup>
deoxyHb	$\beta$ 37 $\beta$ 15 $\alpha$ 14	1548 <sup>h</sup> 1558 <sup>h</sup> 1558 <sup>h</sup>	93 <sup>i</sup> 104 <sup>i</sup> 100 <sup>i</sup>	93 104 100	4HHB <sup>u</sup>
horse heart cytochrome C	59	1552 <sup>j</sup>	114.5	114.5	1HRC <sup>v</sup>
melittin	19A 19B	1547 <sup>f</sup> 1547 <sup>f</sup>	95.3 89.2		2MLT <sup>w</sup>
exendin-4 TC5b W/P cage	25	1546 <sup>f</sup>	73.2		1JRJ <sup>x</sup>

<sup>a</sup> Austin *et al.*, 1989 (45)

<sup>b</sup> Couling *et al.*, 1998 (46)

<sup>c</sup> Hashimoto *et al.*, 1997 (36)

<sup>d</sup> Wen and Thomas Jr., 2000 (43)

<sup>e</sup> Wen and Thomas Jr., 2000 (43)

<sup>f</sup> This paper

<sup>g</sup> Juszczak *et al.*, 1997 (44)

- <sup>h</sup>Rodgers *et al.*, 1992 (47)
- <sup>i</sup>Jordan *et al.*, 1995 (42)
- <sup>j</sup>Skarzynski *et al.*, 1987 (48)
- <sup>k</sup>Lamour *et al.*, 2002 (49)
- <sup>l</sup>Lafitte *et al.*, 2002 (50)
- <sup>m</sup>Grigorieff *et al.*, 1996 (51)
- <sup>n</sup>Welsh *et al.*, 1998 (38)
- <sup>o</sup>Marvin *et al.*, 2006 (37)
- <sup>p</sup>Lolis *et al.*, 1990 (28)
- <sup>q</sup>Lolis and Petsko, 1990 (25)
- <sup>r</sup>Stuckey *et al.*, 1994 (52)
- <sup>s</sup>Fauman *et al.*, 1996 (53)
- <sup>t</sup>Safo *et al.*, 2002 (54)
- <sup>u</sup>Fermi, G. *et al.*, 1984 (55)
- <sup>v</sup>Bushnell *et al.*, 1990 (56)
- <sup>w</sup>Terwilliger *et al.*, 1982 (35)
- <sup>x</sup>Neidigh *et al.*, 2001 (34)

**Table 2**

Indole bond lengths for Tryptophan-containing Proteins and Derivatives

Tryptophan-containing Species	$r(C_3-C_\beta)/\text{pm}^*$	$r(C_3-C_2)/\text{pm}^*$	$r(C_2-N_1)/\text{pm}^*$
Ac-L-TrpME	150.7	135.4	136.2
Melittin W19B	150.0	136.7	138.5
fd virion coat	149.5	135.8	137.2
Gyrase B/chlorobiocin	152.0	136.3	137.6
TIM/PGA	149.3	136.1	137.8
PF3 virion coat	146.3	136.4	138.9

\* All bond lengths measured from corresponding Protein Data Bank file; see Table 1 for corresponding file; pm=picometer.

Table 3

Observed and calculated W3 mode frequency for Trp ( $\nu$ W3) and Trp- $d^5$  ( $\nu$ Wd3) as a function of the  $|\chi^2-1|$  torsional angle.

$ \chi^2-1 $ , deg	$\nu$ W3, $\text{cm}^{-1}$		$\nu$ Wd3, $\text{cm}^{-1}$	
	obs <sup>a</sup>	calc <sup>b</sup>	obs <sup>d</sup>	calc <sup>b</sup>
60.6	1542	1598.1	1514	1539.0
78.0	1543	1601.5	1515	1542.2
83.5	1547	1601.6	1520	1542.4
88.0	1548	1607.7	1519	1548.2
98.7	1552	1610.5	1527	1550.9
105.1	1556	1611.6	1529	1551.9
116.5	1557	1612.3	1530	1552.7
				calcscaled <sup>c</sup>
				1509.9
				1511.5
				1511.8
				1517.9
				1520.4
				1521.4
				1522.6

<sup>a</sup>From (2)

<sup>b</sup>Based on DFT/B3LYP and 6-31g(d) calculations.

<sup>c</sup>Scaling factor = 0.963 (20,22)

<sup>d</sup>From (1)



Table 4

Isotopic dependence of the calculated W3 mode frequency for Trp ( $\nu_{W3}$ ) and  $\chi^2$  torsional angle.<sup>†</sup>

Samples	$ \chi^2 $ , deg	$\nu_{W3}$ , $\text{cm}^{-1}$	$\nu_{W6}$ , $\text{cm}^{-1}$	high $\nu_{W7}$ , $\text{cm}^{-1}$	low $\nu_{W7}$ , $\text{cm}^{-1}$	$I_{\text{high } \nu_{W7}}/I_{\text{low } \nu_{W7}}$	$I_{\text{new } \nu_{W7}}/I_{\text{low } \nu_{W7}}$	$\nu_{W17}$ , $\text{cm}^{-1}$
natural abundance Trp	39.9	1539.9	1467.2	1359.8	1334.7	0.386	0.735	856.8
Trp with indole N replaced by $^{15}\text{N}$ -trp)	39.8	1539.4	1457.8	1358.3	1332.8	0.318	0.691	849.2
ring deuterated Trp (Trp- $d^5$ )	39.9	1510.0	1376.9	----	----	----	----	781.1

<sup>†</sup> All values presented are based DFT/B3LYP and 6-31g(d) and scaling factor = 0.963 (20,22).

Table 5

Calculated frequencies of some tryptophan modes and  $\chi^{2,1}$  torsional angle for the Trp-water complex. The effect of hydrogen bonding and steric strain. <sup>†</sup>

Distance of water from the indole N-H group, Å	$ \chi^{2,1} $ , deg	$\nu_{W3}$ , $\text{cm}^{-1}$	$\nu_{W6}$ , $\text{cm}^{-1}$	high $\nu_{W7}$ , $\text{cm}^{-1}$	low $\nu_{W7}$ , $\text{cm}^{-1}$	$I_{\text{high}} \nu_{W7} / I_{\text{low}} \nu_{W7}$	$\nu_{W17}$ , $\text{cm}^{-1}$
1.5	40.4	1532.1	1442.1	1360.2	1341.5	1.088	871.6
2.0	42.3	1536.3	1431.2	1359.8	1338.6	0.707	857.5
2.5	44.2	1538.1	1420.9	1359.3	1336.6	0.497	856.1
3.0	42.6	1535.9	1413.2	1358.3	1335.7	0.424	858.2
4.0	42.3	1538.1	1413.7	1359.3	1335.2	0.406	857.8

<sup>†</sup> All values presented are based DFT/B3LYP and 6-31g(d) and scaling factor = 0.963 (20,22).

**Table 6**

## Noncovalent Interactions for Tryptophan in Several Proteins

Protein	PDB #	Molecular Coordinates	Distance (Å)	Interaction
Gyrase B with chlorobiocin	1KZN	W170A CZ2 - R168 ANH1	3.96	Cation- $\pi$
		W170A CH2 - K129A NZ	3.96	Cation- $\pi$
apoTIM + PGA	1YPI	W168 CE2 - K134 N2	3.97	Cation- $\pi$
	7TIM	W168 CZ2 - L174 N	5.79	Cation- $\pi$
		W168 CE2 - K134 N2	4.26	Cation- $\pi$
Yersinia PTPase/apo;WT “ + vanadate	1YPT	W354 CE2 - R409 NH2	4.95	Cation- $\pi$
	2I42 <sup>1</sup>	W354 CD2 - R409NH1	5.91	Cation- $\pi$
		W354 CD1 - T358 OG1	3.42	Anion-quadrupole
Pf3 virion coat	1HFP 1FPSF	W38 - R37 (interstrand)	NA	Cation- $\pi$
Fd virion coat Melittin	1HFJ	F42 - W26	NA	Aromatic face-edge
	2MLT	W19A CZ2 - K23A NZ	4.62	Cation- $\pi$
		W19B CE2 - K23B NZ	4.39	Cation- $\pi$
Bacteriorhodopsin W182	2BRD	W182 CZ2 - RET 216 het C20	4.14	Cation- $\pi$
		W182 CZ2 -RET 216 het C19	5.39	Cation- $\pi$

<sup>1</sup>  
(57)

**Table 7**

## W3 Bandwidth for Tryptophans in Peptides and Proteins

Trp-bearing species	W7 Band Ratio (1360 cm <sup>-1</sup> /1340 cm <sup>-1</sup> )	W3 Bandwidth (FWHM/cm <sup>-1</sup> )
Range of $\nu_{W3}$ - $\chi^{2.1}$ Relationship	-	16 (1542 - 1558)
Crystalline Trp derivatives	-	9 - 15
Exendin-4 TC5b W/P cage	1.1	16
Melittin	1.5	15
Horse heart cytochrome C	0.8	14
Fd virion peptide	1.7	18
Pf3 virion peptide	1.5	23
PTPase + tungstate	0.8	11
apoPTPase	2.0	17
TIM + PGA	0.8	16
apoTIM	1.3	13