

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Aliasing Compromises Staggered-Rotamer Analysis of Polypeptide Sidechain Torsions

Jürgen M. Schmidt

## Abstract

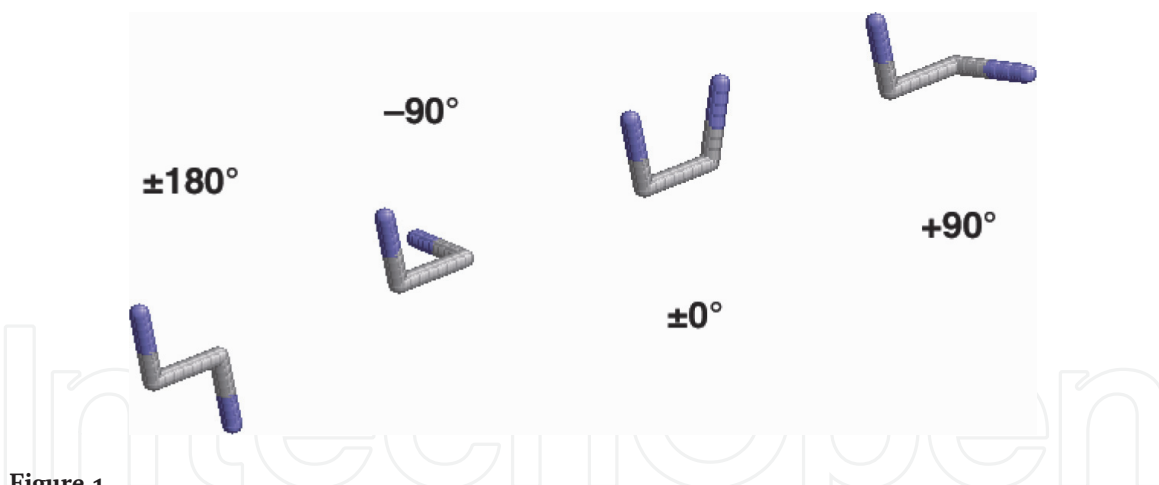
Circular undersampling and the ensuing aliasing effect are demonstrated to compromise nuclear magnetic resonance (NMR)-based molecular torsion-angle analysis referring to experimental  $^3J$ -coupling constants when employing the staggered-rotamer model, also known as Pachler model. This popular model is flawed insofar as it systematically produces counterintuitive probabilities for the two minor constituents out of the total three rotamers, to the effect that the apparent circular mean direction of the molecular bond conformation is inflected about its main rotamer angle, a situation that apparently went unnoticed for more than 50 years. The principal reason for systematic errors lay in the model's ill-conceived attempt to resolve the bimodal  $^3J$ -coupling-angle dependency by a mere three discrete points on the circle, thereby conflicting with the Nyquist-Shannon sampling theorem. An anti-aliasing approach is being offered that helps improve the results.

**Keywords:** circular distribution, directional data, probability density, torsion angle conformation, staggered-rotamer equilibria, discrete Fourier transform, sampling theorem, undersampling, anti-aliasing, vicinal coupling constants,  $^3J$ , aminoacid sidechain, protein structure

## 1. Introduction

Atom-atom bonds in a molecule often give rise to rotational degrees of freedom, also known as torsion angles. A torsion angle, also known as dihedral angle, is formed by three consecutive bonds in a molecule and defined by the angle between the two outer bonds projected onto a plane perpendicular to the central bond (Figure 1).

Finding out how two parts on either side of a rotatable bond relate to each other, that is, assigning a value to the torsion angle, presents one of the challenges in molecular structure determination. A molecule adopting different geometric arrangements—without breaking or making bonds—is said to exhibit distinguishable conformers. Nuclear magnetic resonance (NMR) spectroscopy [2] is uniquely positioned to help characterize not only static molecular structure, but also dynamical processes that involve interconversion between conformers on a short, typically nanosecond timescale. Studying torsion-angle geometry and dynamics by NMR benefits from the measurement of  $^3J$  coupling constants [3]. Typically on the order of a few Hertz, these are magnetic interaction parameters between atoms X and Y in



**Figure 1.**

Basic torsion-angle fragment depicted for various angles of rotation. IUPAC-IUB recommendations [1] assign positive angles to clockwise twists about the central bond of a torsion. Identical torsion-angle sign and value result regardless of viewing down the central bond ‘front-to-rear’ or ‘rear-to-front’.

the three-bond, four-atom fragment X-A-B-Y constituting dihedral angle  $\theta_{XY}$  about bond AB. In essence,  $^3J_{XY}$  coupling interaction is strongest when bonds XA and BY are orientated parallel ( $\theta_{XY} = \pm 180^\circ$ ) and weakest if perpendicular ( $\theta_{XY} \sim \pm 90^\circ$ ). A secondary smaller maximum exists for  $\theta_{XY} = \pm 0^\circ$ .

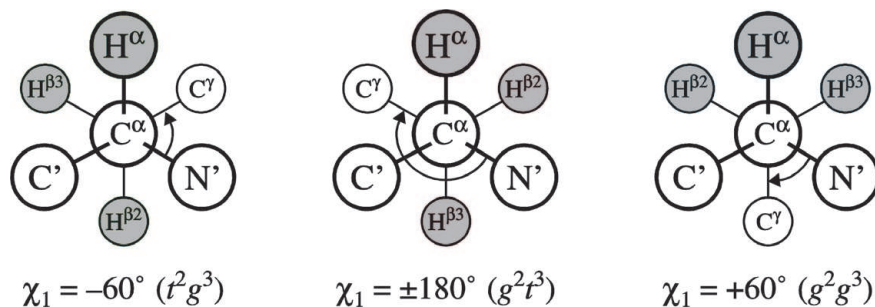
### 1.1 The Karplus curve

Karplus [4] formulated a universal empirical relation as to how a  $^3J$  coupling constant depends on the intervening torsion angle:

$$^3J_{XY}(\theta) = C_0 + C_1 \cos(\theta + \Delta\theta) + C_2 \cos(2\theta + 2\Delta\theta) \quad (1)$$

Karplus coefficient  $C_0$  signifies the *average* coupling strength obtained for a complete torsion-angle revolution. The  $C_2$  term lends the Karplus curve its *bimodal* shape and determines the undulation depth. The  $C_1$  term adds a *unimodal* component that affects the difference between primary and secondary maxima. Typically,  $C_1$  is negative in order for the primary coupling-constant maximum to appear in the so-called *trans* conformation, when the X-A-B-Y angle is at  $\pm 180^\circ$  and the atom-atom interaction is strongest.

We here focus on biomolecular structure analysis of the so-called  $\chi_1$  torsion in aminoacid residues, the constituent units of polypeptide and protein chains [1], where the value of  $\chi_1$  refers to the dihedral angle  $\theta_{N'C_\gamma}$  about the bond between carbons  $C_\alpha$  and  $C_\beta$  (**Figure 2**). It is imperative to understand, that for each of the up to nine possible atom-pair combinations formed from X = N', C', or  $H^\alpha$  (“front”) and Y = C',  $H^{\beta 2}$ , or  $H^{\beta 3}$  (“rear”), the phase parameter  $\Delta\theta$  in the Karplus curve must



**Figure 2.**

Staggered-rotamer states for aminoacid sidechain torsion  $\chi_1$  viewed down its central  $C^\alpha - C^\beta$  bond. Most aminoacids exhibit two  $H^\beta$  protons oriented trans or gauche with respect to  $H^\alpha$  as indicated by a shorthand [5].

be either  $0^\circ$  or  $+120^\circ$  or  $-120^\circ$ . This crucially sets up and determines the invariable three-point sampling of the Karplus curve in the situation of  $\chi_1$ .

## 1.2 The staggered-rotamer model

Pachler [6] employed  $^3J_{\text{H}\alpha\text{H}\beta 2}/^3J_{\text{H}\alpha\text{H}\beta 3}$  coupling data pairs in order to quantify populations of conformational states in ethane-like compounds, such as those encountered in aminoacid sidechains. According to Pachler's model, the torsion  $\chi_1$  in the aminoacid sidechain is considered to adopt one of three, or to hop between two or between all three so-called *staggered rotamers*, characterized by nominal  $\chi_1$  angle values of  $-60^\circ$ ,  $\pm 180^\circ$ , or  $+60^\circ$  (**Figure 2**). An alternative set of angles comprising  $-120^\circ$ ,  $\pm 0^\circ$ , and  $+120^\circ$  would give rise to *eclipsed* rotamers, which, however, are unfavorable in view of molecular energy and, thus, are disregarded for reasons of limited or vanishing lifetime. Both *Pachler model* and *staggered-rotamer model* are synonymous.

The Pachler model allows one to deduce probabilities for each of the three staggered states once a pair of experimental  $^3J$  coupling constants is available. Historically, hydrogen-hydrogen coupling constants  $J_{\text{H}\alpha\text{H}\beta 2}$  and  $J_{\text{H}\alpha\text{H}\beta 3}$  were the first ones accessible to measurement [7], heteronuclear coupling constants were alternatively used later [8]. Accordingly:

$$\begin{aligned} P_{-60} &= (J_{\text{H}\alpha\text{H}\beta 2} - J_{\pm 60^\circ}) / (J_{180} - J_{\pm 60^\circ}) \\ P_{180} &= (J_{\text{H}\alpha\text{H}\beta 3} - J_{\pm 60^\circ}) / (J_{180} - J_{\pm 60^\circ}) \\ P_{+60} &= 1 - P_{-60} - P_{180} \end{aligned} \quad (2)$$

Eq. (2) derives staggered-rotamer probabilities from the measured  $J_{\text{XY}}$  coupling constants by interpolating between specific coupling values  $J_{180^\circ}$  and  $J_{\pm 60^\circ}$ , also designated as  $J_{\text{trans}}$  and  $J_{\text{gauche}}$ , respectively, that correspond to the specified fixed geometries for the  $\theta_{\text{XY}}$  angle.

The Pachler model sets up the experimental observable as the corresponding probability-weighted *state-averaged value of J* (indicated by brackets):

$$\langle J \rangle = P_{180} \cdot J_{180} + P_{+60} \cdot J_{+60} + P_{-60} \cdot J_{-60} \quad (3)$$

Both models, the dependency of  $^3J$  couplings on the torsion angle according to Karplus (continuous model) and the dependency of the torsion-angle distribution on relative  $^3J$  coupling-constant pairs according to Pachler (discrete model), form an indispensable basis of many a biomolecular NMR structure investigation.

## 2. Theory

### 2.1 Continuous-discrete model interconversion

Sets of probabilities, associated with the three staggered rotamers in the traditional Pachler model of aminoacid sidechain torsion-angle variability, transform into equivalent continuous angle distributions and, vice versa, discrete staggered-rotamer probabilities can be computed from any given angular direction in continuous circular space.

Mean direction  $\theta$  and concentration  $R$  of a continuous angle distribution convert into corresponding normalized probabilities for three discrete-state samples evenly distributed on the circle, according to [9],

$$\begin{aligned}
P_{-60} &= (1/3) \{1 + R \cos \theta - \sqrt{3}R \sin \theta\} \\
P_{180} &= (1/3) \{1 - 2R \cos \theta\} \\
P_{+60} &= (1/3) \{1 + R \cos \theta + \sqrt{3}R \sin \theta\}
\end{aligned} \tag{4}$$

Conversely, discrete staggered-rotamer probabilities aggregate into the *circular mean direction* of a continuous angle distribution,

$$\tan \theta = \sqrt{3}(P_{+60} - P_{-60}) / \{3(P_{+60} + P_{-60}) - 2\} \tag{5}$$

and its *circular order parameter* (concentration) runs between 1 (static fixed angle) and 0 (rotational average),

$$R = \left\{ 1 + 3 \left( P_x^2 + P_y^2 - P_x - P_y + P_x P_y \right) \right\}^{1/2} \tag{6}$$

where  $P_x$  and  $P_y$  signify any pair out of the three probabilities  $P_{-60}$ ,  $P_{180}$ , and  $P_{+60}$ .

## 2.2 Three-point discrete Fourier transforms

Three-point discrete Fourier transforms (DFT) using kernel,  $\omega = \exp(-i2\pi/N)$ , such that a phase of  $-i/N$  turns equates to  $120^\circ$  phase rotation if  $N = 3$ , are expressed in matrix form as [10]

$$\begin{bmatrix} M_0 \\ M_1 \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^1 & \omega^2 \\ 1 & \omega^2 & \omega^4 \end{bmatrix} \times \begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix} \tag{7}$$

Using complex operator  $i = (-1)^{1/2}$ , transforms rest on conjugate operators  $\omega^1 = \omega^+ = -(1/2) (1 + i\sqrt{3})$  and  $\omega^2 = \omega^- = -(1/2) (1 - i\sqrt{3})$ , where  $\omega^+ \omega^- = 1$ . Critically,  $\omega^4 = \omega^1$  also. Factors  $(1/2)$  and  $(i\sqrt{3}/2)$  identify with  $\cos(\pi/3)$  and  $i\sin(\pi/3)$ , respectively, such that Eq. (7) recasts as

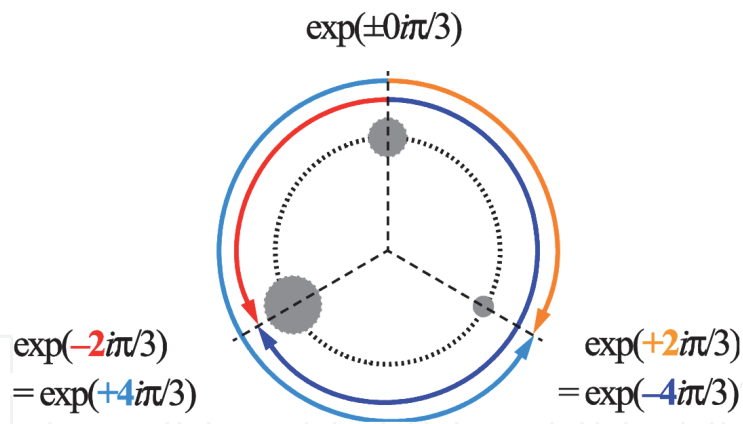
$$\begin{aligned}
M_0 &= m_0 + m_1 + m_2 \\
M_1 &= m_0 - (m_1 + m_2) \cos(\pi/3) - i(m_1 - m_2) \sin(\pi/3) \\
M_2 &= m_0 - (m_1 + m_2) \cos(\pi/3) + i(m_1 - m_2) \sin(\pi/3)
\end{aligned} \tag{8}$$

While constituent  $M_0$  represents pure real data,  $M_1$  and  $M_2$  form a complex-conjugate pair sharing identical real parts. Sums,  $\cos\theta = (m_1 + m_2)$ , and differences,  $\sin\theta = i(m_1 - m_2)$ , respectively, form the real and imaginary components of the *circular direction* in the continuous-discrete model transformation of Eq. (5).

Swapping  $\omega$  and its conjugate yields the inverse transform, including normalization, as

$$\begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega^1 \\ 1 & \omega^1 & \omega^2 \end{bmatrix} \times \begin{bmatrix} M_0 \\ M_1 \\ M_2 \end{bmatrix} \tag{9}$$

Now consider probabilities at the  $N = 3$  equispaced samples on the circle, that is,  $P_{180} = M_0/3$ ,  $P_{-60} = M_1/3$ , and  $P_{+60} = M_2/3$  inserted in Eq. (9), informing the first three *modes*  $p_m = 0,1,2$  of a circular probability density, generically given by



**Figure 3.**  
 Directional ambiguity in circular 3-point systems due to degeneracy in circular modes 1 and 2.

$$p_m = N^{-1} \sum_k (\cos m\theta_k - i \sin m\theta_k) \quad (10)$$

Continuous-discrete model interconversion is transparent regarding Fourier transform. Combining circular modes  $p$ , the transforms of the  $P$  domain, yields the circular mean direction as the *first circular moment*,<sup>1</sup>

$$\tan \theta = (p_1 - p_2) / (p_1 + p_2) \quad (11)$$

and the (squared) circular order parameter, related to auto-correlation [9], as the *second circular moment*,

$$R^2 = p_1 \cdot p_2 \quad (12)$$

The zero-th mode quotes the trivial chance of finding the direction at all,

$$p_0 = 1 = P_{180} + P_{-60} + P_{+60} \quad (13)$$

First and second modes, respectively, then supposedly represent amplitudes of features occurring *once-per-cycle* ( $m = 1$ ) and *twice-per-cycle* ( $m = 2$ ),

$$\begin{aligned} p_1 &= P_{180} - (1/2) \{ P_{-60} + P_{+60} - i\sqrt{3}(P_{-60} - P_{+60}) \} \\ p_2 &= P_{180} - (1/2) \{ P_{-60} + P_{+60} + i\sqrt{3}(P_{-60} - P_{+60}) \} \end{aligned} \quad (14)$$

However, these are conjugates of the *same circular frequency*! Thus, mode  $p_2$  does *not* reflect twice the rate of  $p_1$ . This peculiarity is unique to the 3-point DFT and ambiguates the meaning of *single* and *double* phase advances through the circle (**Figure 3**), because if, and only if  $N = 3$ , the following identities hold:

$$\exp(\mp i2\pi/3) = \exp(\pm i4\pi/3); \cos(\pi/3) = -\cos(2\pi/3); \sin(\pi/3) = \sin(2\pi/3) \quad (15)$$

$$\text{implying } \omega^{+2} = \omega^{-1}, \omega^{-2} = \omega^{+1}.$$

<sup>1</sup> In analogy to linear statistics, circular modes are series of multiple-angle arguments, whereas circular moments are power series in the trigonometric operators. Moments are simple combinations of modes, the inverse does not hold.



Apparently, periodic wrapping equates the double angular speed,  $\omega^2$ , to the conjugate of the single angular speed,  $\omega^-$ , thereby folding—under sign inversion—samples  $M_1$  and  $M_2$  onto each other, as visualized in **Figure 4**.

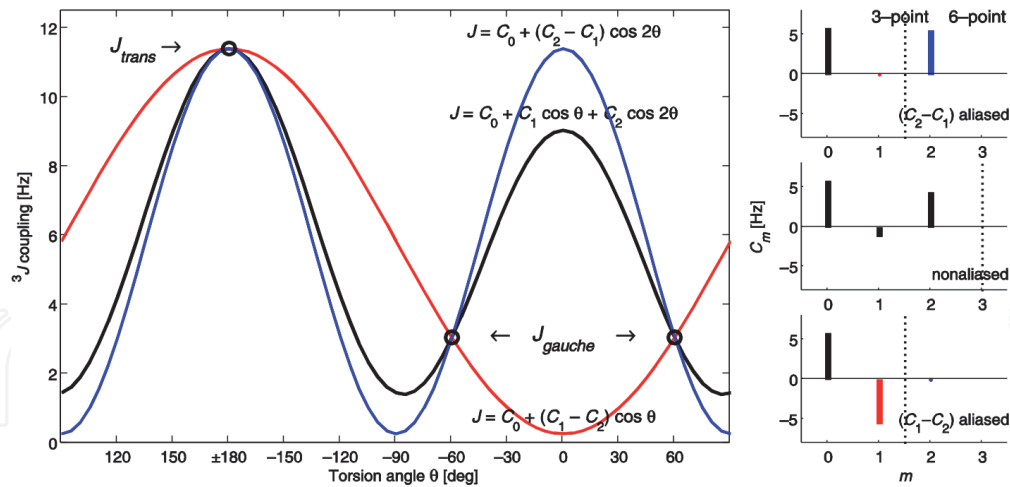
All issues regarding sampling-rate doubling, sign change in the real portion of the complex numbers, and data-pair ambiguity hold for both forward and inverse three-point transforms. Both setups produce identical results, making reconstructing clockwise and counterclockwise arrangements of the circular samples, or  $P_{+60}$  and  $P_{-60}$ , for that matter, somewhat impossible. Entirely independent of input and output data, the effects arise solely from the transformation operator and are, therefore, *model-inherent*.

Regarding the unique case of three staggered rotamers, let us also consider Karplus coefficients  $C_0 = m_0$ ,  $C_1 = m_1$ , and  $C_2 = m_2$  subjected to three-point DFT according to Eq. (7) and obtain precisely those specific  $J$ -coupling values needed to compute the state-averaged  $J$  value of Eq. (3):

$$\begin{aligned} J_{trans} &= J_{180} = C_0 - C_1 + C_2 \\ J_{gauche} &= J_{\pm 60} = C_0 + C_1/2 - C_2/2 \end{aligned} \quad (16)$$

Inverse-transforming the values of  $J_{trans}$  and  $J_{gauche}$ , inserted as  $J_{180} = M_0/3$ ,  $J_{+60} = M_1/3$ , and  $J_{-60} = M_2/3$  into Eq. (9), recovers the three coefficients (i.e., modes) of the Karplus Eq. (1).

Mathematically, this looks clean and entirely reversible. In analogy to the modes of the circular probability density in Eq. (14), the DFT obtained for both  $J_{+60}$  and  $J_{-60}$  inserted in Eq. (3) would necessarily result in complex-conjugate numbers. And yet, deliberately ignored was the fact that theoretical as well as experimentally



**Figure 4.**

Main panel: effective Karplus-curve manifold generated by aliasing  $\cos\theta$  and  $\cos 2\theta$  modes. Typical coefficients ( $C_0$ ,  $C_1$ ,  $C_2$ ) and a phase offset of  $\Delta\theta = 0^\circ$  apply to  $^3J_{\text{H}\alpha\text{H}\beta_3}$  in aminoacid sidechains. The original (non-aliased) Karplus curve is shown in black. Aliased unimodal (red) and bimodal (blue) curves, respectively, result from coefficient  $C_2$  folding onto  $C_1$  and, vice versa,  $C_1$  folding onto  $C_2$ . Associated with staggered rotamers are those invariant focal points  $J_{trans}$  and  $J_{gauche}$  through which all curves pass. Small panels: sampling properties of rotamer models in contexts of 3-point and 6-point discrete Fourier transforms of Karplus curves. Each coefficient  $C_m$  as the  $m$ th circular mode (bars) connects with an effective circular sampling rate in units of  $(2\pi \text{ rad})^{-1}$ ,  $m$  being the periodicity of that component. Bottom: sampling the bimodal coupling-angle dependency at three  $120^\circ$ -equispaced (e.g., staggered-rotamer) angles corresponds to rate  $1.5 \pi^{-1} \text{ rad}^{-1}$  (dotted line), too low to resolve the higher  $C_2$  mode. Consequently,  $C_2$  inverts and folds into the lower frequency band and stacks onto or aliases the negative  $C_1$  mode (red bar). The original  $C_2$  amplitude appears blanked. Reconstructed from such distorted coefficients, the mis-sampled Karplus curve appears purely unimodal. Top: alternatively,  $C_1$  folds under sign inversion onto the  $C_2$  mode (blue bar), generating a purely bimodal curve. Middle: aliasing of the high mode is avoided with a 6-state staggered-eclipsed  $60^\circ$ -equispaced rotamer model that samples the circle at twice the rate of  $3 \pi^{-1} \text{ rad}^{-1}$  (dash-dotted line).

observed  $J$ -coupling constants are generally considered real data and, therefore,  $J$  values for *gauche*<sup>+</sup> and *gauche*<sup>−</sup> states are ordinarily identical. Clockwise and counterclockwise orientations become indistinguishable once again. At any rate, disregarding imaginary components amounts to loss of information, introducing a subtle irreversibility into the staggered-rotamer analysis that cannot be corrected.

### 3. Results

#### 3.1 Worked example: selected sidechain rotamers in a protein

In the course of extensive protein-structure studies by NMR,  $^3J$  coupling constants related to aminoacid sidechain torsions were measured in a variety of proteins. Ongoing investigations target sidechain  $\chi_1$  torsion structure in ribonuclease T<sub>1</sub> (RNase T<sub>1</sub>), an enzyme of 104 aminoacids size, experimental  $^3J$  coupling data of which are deposited with the BMRB [11]. Applying up to nine available coupling constants as experimental constraints, two models, a continuous single-state torsion angle and the discrete staggered-rotamer populations, were least-squares fitted for each aminoacid in the enzyme. Exemplifying the present issues, RNase T<sub>1</sub> contains three histidines, each of which adopts a different predominant staggered state with only little dispersion about their respective circular mean direction. Accordingly, the torsions lock  $\chi_1$  conformations near  $\pm 180^\circ$  in His27 (as evident from a large *trans*- $^3J_{N'C_\gamma}$  coupling constant, **Figure 2**), near  $+60^\circ$  in His40 (large *trans*- $^3J_{H\alpha C_\gamma}$ ), and near  $-60^\circ$  in His92 (large *trans*- $^3J_{C'C_\gamma}$ ).

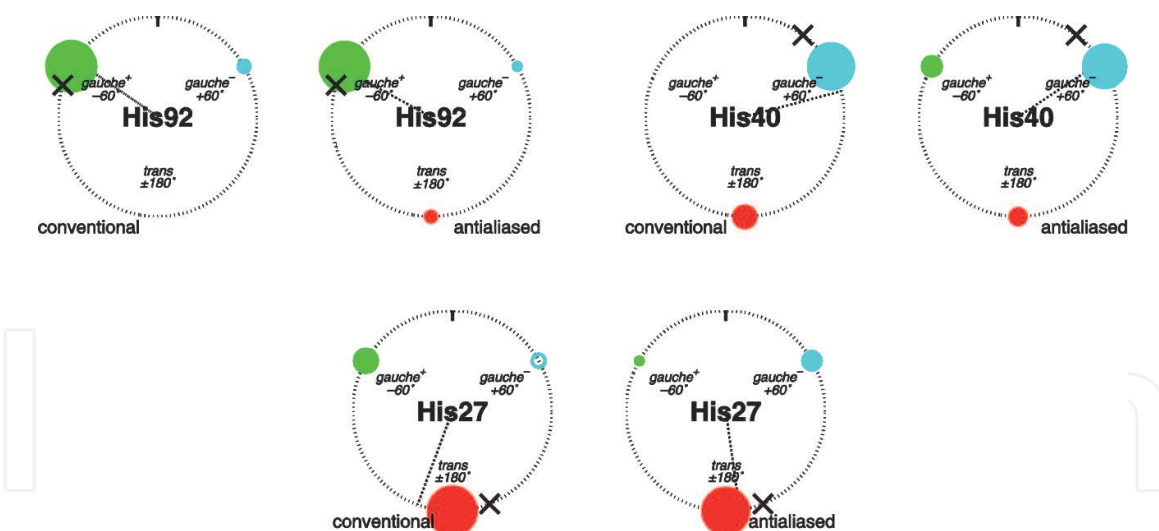
The predominant  $\chi_1$  rotamer in His92 is populated at  $P_{-60} = 92\%$ , with only negligible or minor contributions from  $P_{180} = 0\%$  and  $P_{+60} = 8\%$ , which together make for a circular mean direction of  $-56^\circ$  (**Figure 5**). The single-state model converges at  $-72^\circ$ , somewhat “mirrored” or inflected about the ideal  $-60^\circ$  staggered state.

The His27 torsion noticeably tilts away from the ideal  $\pm 180^\circ$  staggered angle, circular mean direction from the discrete model being  $-160^\circ$ . However, a  $\chi_1$  value of  $+158^\circ$  results from fitting a continuous single-state angle and would infer a primary  $\pm 180^\circ$  and a secondary  $+60^\circ$  rotamer, broadly at 3:1 proportion [9]. Yet, populations emerging from staggered-rotamer analysis ( $P_{-60} = 23\%$ ,  $P_{180} = 89\%$ ,  $P_{+60} = -12\%$ ) suggest  $-60^\circ$  as the more significant secondary constituent. This, together with a decidedly negative probability for the  $+60^\circ$  rotamer, results in an “opposite” circular mean direction, again inflected about the constituent  $\pm 180^\circ$  state.

Similarly, discrete staggered-state populations in His40 emerge as  $P_{-60} = 0\%$ ,  $P_{180} = 22\%$ , and  $P_{+60} = 78\%$ , making for a circular mean direction of approximately  $+76^\circ$ , while  $\chi_1$  converges at  $+36^\circ$  in the continuous model. In the staggered-rotamer model, a mean direction deviating from  $+60^\circ$  toward smaller angle values would command a significant contribution from the  $-60^\circ$  state to the overall average. Once again, the discrete model suggests the less plausible  $\pm 180^\circ$  conformer.

Even though both fitted models agree as regards the main conformer in each case, state populations of the minor conformers in the staggered-rotamer model seem always somewhat counterintuitive. Deviating from a single ideal staggered state, the inflection phenomenon in the discrete model consistently manifests as a false apparent mean direction in the rotationally opposite sense, off in terms of degrees by—approximately—as much as the *correct* mean direction derived from the continuous model deviates from that nearest ideal staggered state. How do we know it is the staggered-rotamer model delivering wrong results and not the continuous model? Because were the angle tilted to the “opposite/other side,” the





**Figure 5.**

Conventionally optimized discrete staggered-rotamer probabilities for histidines in RNase T<sub>1</sub>, and the effect of anti-aliasing applied to these. Probabilities  $P_{\pm 180}$  (red),  $P_{-60}$  (green), and  $P_{+60}$  (blue) shown as area-proportional circles, and dials indicate the effective circular mean direction obtained from these. Crosses mark torsion angles that best fit the J-coupling data using a continuous variable-angle single-state model. Anti-aliasing improves the probabilities insofar as apparent mean directions for the modified sets are closer to a staggered state than for the original and also agree better with torsion angles from the single-state fit result. Anti-aliasing helps inflect the apparent mean direction about the nearest staggered angle toward the opposite side. Most noticeably, His27 is an example of a predominant trans rotamer with clockwise deflection of its apparent circular mean direction, chiefly due to the conventionally fitted negative probability of the  $+60^\circ$  state (open circle). Anti-aliasing reverts the apparent mean into a counterclockwise direction and also inverts the implausible value of  $P_{+60}$ .

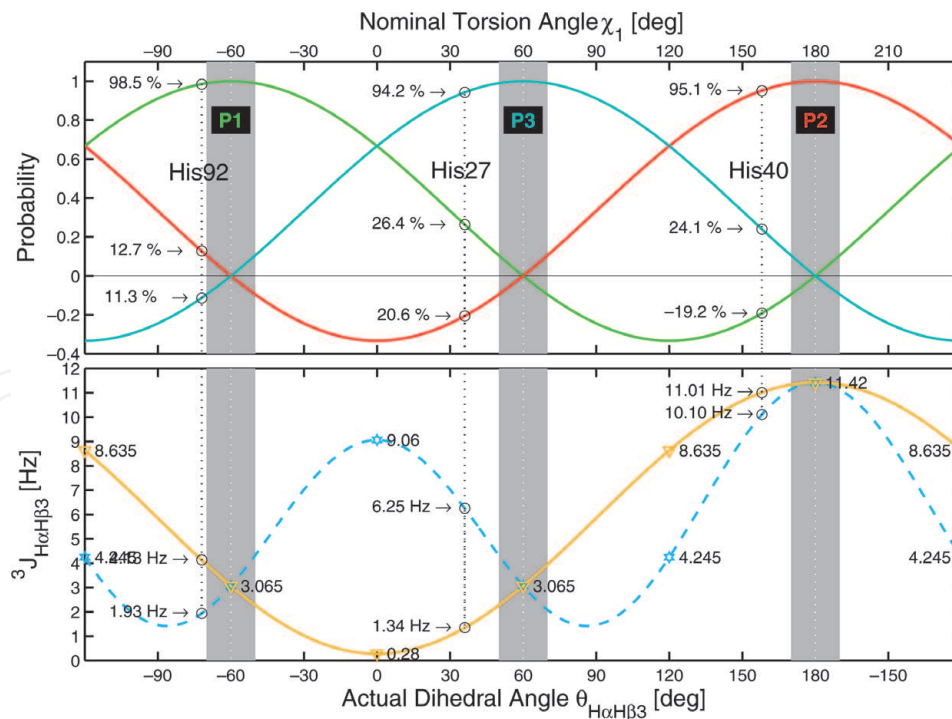
corresponding set of probabilities (top panel of **Figure 6**) would generate a value of  $J$  typically in disagreement with the experimental value, for example, for  $^3J_{\text{H}\alpha\text{H}\beta 3}$ , as read off the Karplus curve for the given angle (bottom panel of **Figure 6**).

Such discrepancy is not limited to specific aminoacid residue types; rather, it is seen—without exception—in all rotamer studies. Revisiting  $^3J$  data collected for a different protein, *Desulfovibrio vulgaris* flavodoxin [12], confirmed suspicion that a systematic issue would be at work. Invariably, counterintuitive admixtures of rotamers are calculated for all non-ideally staggered sidechain conformations.

### 3.2 A compromise fix: anti-aliasing

Anti-aliasing signifies any procedure that attempts to ameliorate adverse effects from aliasing due to coarse sampling. Anti-aliasing applied in, for example, image processing helps smoothen jagged lines and edges that result from digitizing continuum data into discrete samples [13]. Generally, such approaches form weighted averages over a number of adjacent data points and, therefore, are irreversible data manipulation.

Anti-aliasing applied to the discrete rotamer model helps restore some bimodal curve feature in the distorted coupling angle relationship and would also minimize unrealistic negative excursions in the probability parameters. Considering that Eq. (2) introduces into the traditional interpretation of staggered rotamers a strict one-to-one correspondence between each coupling and precisely one, and only one, staggered-state population, this permits the remaining two population parameters to be manipulated at will without interfering with that respective correspondence, offering an opportunity to diminish differential probability and improve the population parameters generally. The approach taken here represents the improved symmetrical, balanced variant of the range-bound probability fitting devised in [9] by manipulating all probability parameters equally and simultaneously:



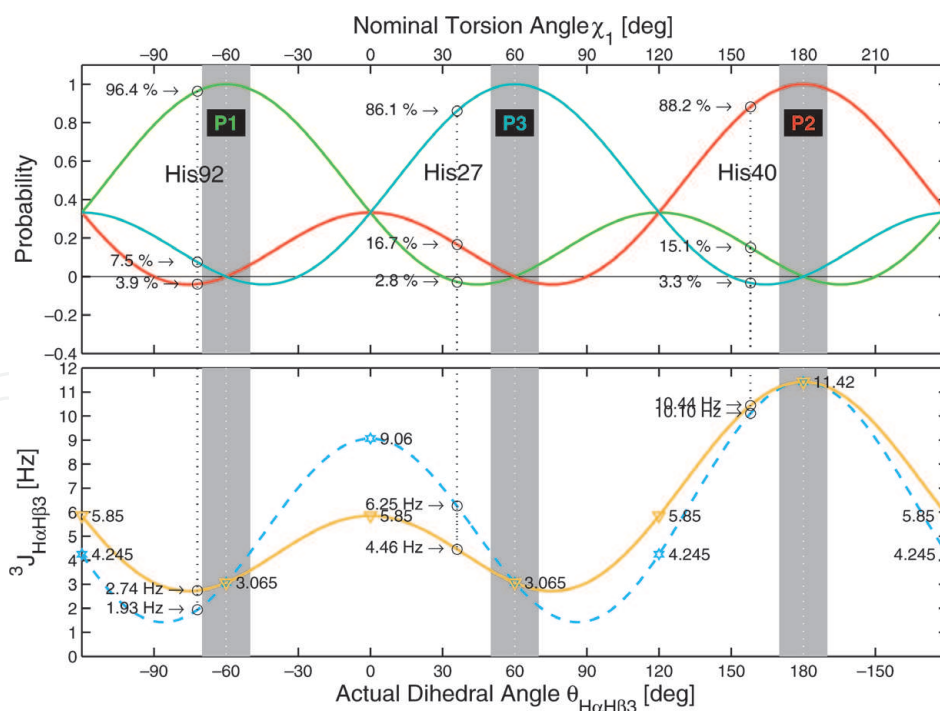
**Figure 6.** Top: concerted change in staggered-rotamer probabilities  $P$  with varying circular mean direction for torsion  $\chi_1$  according to Eq. (2); bottom: state-averaged aminoacid  $^3J_{H\alpha H\beta_3}$  coupling derived from these probabilities when aggregated according to Eq. (3). The decidedly unimodal constructed curve (solid line) tracks exclusively the trans rotamer probability  $P_2$ , thus, failing to reproduce the bimodal angle dependency of  $J$  in direct response to an angle sweep according to the original Karplus curve (dashed line) of Eq. (1). Annotated RNase-histidine examples highlight large differential probability between the minor rotamer pairs (top) as well as counterintuitive coupling values (bottom) as seen for all deflections from ideal staggered states (shaded zones).

$$\begin{aligned}
 {}^{aa}P_{-60} &= P_{-60} + z(-P_{-60}P_{180}/2 - P_{-60}P_{+60}/2 + P_{180}P_{+60}) \\
 {}^{aa}P_{180} &= P_{-60} + z(-P_{-60}P_{180}/2 + P_{-60}P_{+60} - P_{180}P_{+60}/2) \\
 {}^{aa}P_{+60} &= P_{+60} + z(+P_{-60}P_{180} - P_{-60}P_{+60}/2 - P_{180}P_{+60}/2)
 \end{aligned} \tag{17}$$

Underlying the anti-aliasing principle in Eq. (17) is the observation that the smallest probability derived from the Pachler model is always too small, if not negative, while that of the main rotamer is always too large. Therefore the small  $P$  value is being raised (one positive increment) at the expense of both other parameters, from which half the amount each is being taken off (two negative half increments). Applying the correction in turn to each set gives rise to the anti-aliasing matrix in Eq. (17). Normalization is critically ensured as each product term added (or subtracted) is being subtracted (or added) elsewhere.

Values between 0.5 and 1.5 for the anti-aliasing parameter  $z$  range form a reasonable compromise between such diverse effects as differential probability, negative probability, and unimodality of the coupling-angle dependence [9]. By choosing  $z = 1$ , all eclipsed-state populations amount to one-third, making these situations indistinguishable from the complete rotational-average limit. Contrasting the conventional staggered-rotamer model, the anti-aliased model typically inverts and reduces differential probability between the two minor rotamers (Figure 7).

Aliasing affects primarily the probabilities, yet, consequently also any back-calculated—as opposed to experimental—coupling values. Only the mean conformation-averaged calculated  $J$  couplings remain invariant when inserting anti-aliased probabilities  ${}^{aa}P$  into Eq. (3), equaling those obtained with the unmodified  $P$  values.



**Figure 7.**

Top: staggered-rotamer probabilities for an angle sweep anti-aliased using  $z = 1$  in Eq. (17); bottom: dependency of the state-averaged  $^3J_{\text{H}\alpha\text{H}\beta 3}$  coupling according to Eq. (3) when using anti-aliased staggered-rotamer probabilities, re-introducing into the coupling-angle dependency a certain bimodal feature (solid line), approaching that in the original Karplus curve (dashed line). Other graph features as described in Figure 6 legend.

In an attempt to improve the coupling estimates, anti-aliasing can be implemented in two possible ways:

- Applied as post-processing procedure, the fitted probabilities are being modified (anti-aliased) according to Eq. (17), and coupling estimates in Eq. (3) amended in accordance with the updated values of  $P_{x,y,z}$ . As least-squares parameter optimization already minimizes the discrepancy between observed and calculated  $J$  coupling values, any subsequent parameter modification will invariably increase the overall fit error in  $J$ . After all, we are interested less in the fit error than we are in plausible and acceptable molecular descriptions.
- Applied during data fitting, anti-aliasing and optimization of the probability parameters will interfere and cancel somewhat. Again, optimization strives to minimize the difference between observed and calculated  $J$  values; only this time, the instantaneous fit error feeds back on the continual probability-parameter adjustment, counteracting the anti-aliasing due to the model-inherent divergence between calculated and true coupling-angle dependence.

In practice, parameters optimized following option (ii) converged nearer to those obtained in the original, aliased fit, while more significant corrective effects are seen when adopting option (i).

As circular folding tends to equalize the meaning of both first and second mode coefficients of the Karplus curve, one might equally (or alternatively) employ jointly the  $C_2$  coefficient with an additional folded (mirrored and inverted)  $C_1$  coefficient. This would connect with twice the angular speed, which happens to

coincide with the single speed of opposite sign (Eq. (14)). This condition is identical to anti-aliasing using  $z = 2$ .

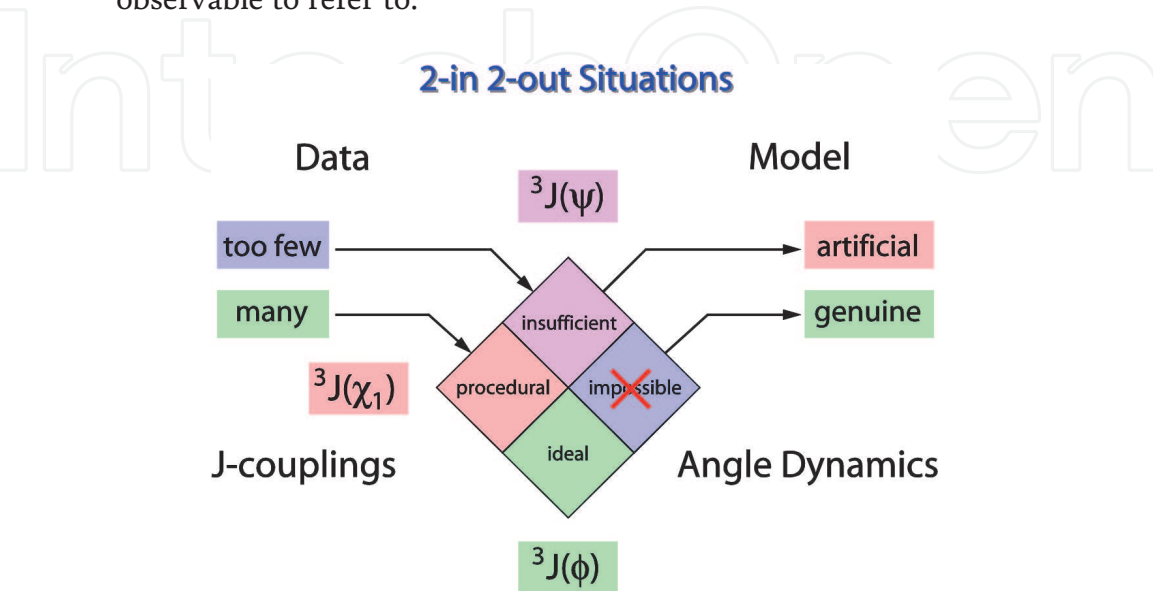
Anti-aliasing is irreversible, as the sequence order in two recursions,  $z = +1/-1$  vs.  $-1/+1$ , does make a difference. The latter results in two equal  $J_{trans}$  maxima, much like applying  $z = +2$  once, and four  $J_{gauche}$  values for all other  $60^\circ$  intervals. The probabilities are more peaked at the top and flatter at the bottom. The former sequence produces flat tops in both  $P$  and  $J$  profiles and peaked bottoms, even though  $J_{trans}$  and  $J_{gauche}$  are correctly reproduced, and all eclipsed states show  $J_{mean}$ .

As multiple couplings are considered in real applications, the final outcome is a balanced compromise between all contributing parts.

#### 4. Discussion and conclusion

In principle, conversion of data into a model encounters one of four possible situations (Figure 8), depending on the ratio between the number of independent experimental observations and the number of model parameters to determine. Focusing here on aminoacid geometry:

- i. Normally, if many data are available and accurate, the analysis outcome is most likely reliable, as is the case with the polypeptide mainchain  $\phi$  torsion angle, where six  $J$ -coupling constants were collected, associated with phases densely spaced at  $60^\circ$  intervals on the circle, allowing to determine the torsion-angle parameter with high accuracy [14].
- ii. A number of model variables that exceeds the number of experimental observables renders analysis generally impossible, owing to too few data.
- iii. At most, interpretation of insufficient data would produce some artificial result. Indications are that this might be the case with polypeptide mainchain  $\psi$  torsion analysis, where a mere three  $J$ -coupling constants are accessible for sampling rotamer probabilities associated with three phases at  $120^\circ$  intervals, similar to the present work, yet one more experimental observable to refer to.



**Figure 8.**  
Possible pathways for conversion of experimental data into abstract or theoretical model parameters.



- iv. Finally, the present case of polypeptide sidechain  $\chi_1$  torsion analysis would not normally suffer any data shortage in determining a single angle parameter, given a principal theoretical parameter over-determination from up to nine experimental observables. Rather, the analysis procedure itself produces the notoriously distorted artificial outcome. This work demonstrated the mathematical reasoning as to why implausible probabilities are being obtained regardless of an overwhelming data supply.

Attempting to extract meaningful parameters of mean direction and dispersion by digitizing the bimodal  $J$ -coupling angle dependency through a “bottleneck” of only three staggered-rotamer angles must appear error-prone. In many ways, this seems tantamount to the proverbial “square peg in a round hole.” It is hoped that the issues and pitfalls connected with circular three-point sampling and transforms, demonstrated here in a biochemical context, would inspire a fresh look at applications in other disciplines in mathematics, physics, or engineering.

## Acknowledgements


Frank Löhr (Goethe-Universität Frankfurt am Main) is thanked for stimulating discussion and his continual interest in  $J$ -coupling studies.

## Author details

Jürgen M. Schmidt  
School of Biosciences, University of Kent, Canterbury, United Kingdom

\*Address all correspondence to: [jmschmidt@btconnect.com](mailto:jmschmidt@btconnect.com)

## IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 



## References

- [1] IUPAC-IUB commission on biochemical nomenclature. *Journal of Molecular Biology*. 1970;**52**:1-17
- [2] Evans JNS. *Biomolecular NMR Spectroscopy*. 1st ed. Oxford: Oxford University Press; 1995
- [3] Schmidt JM, Löhr F. Refinement of protein tertiary structure by using spin-spin coupling constants from nuclear magnetic resonance measurements. In: Faraggi E, editor. *Protein Structure*. Rijeka, Croatia: Intech; 2012. pp. 95-120. ISBN: 978-953-51-0555-8. Available from: <http://www.intechopen.com/books/protein-structure>
- [4] Karplus M. Vicinal proton coupling in nuclear magnetic resonance. *Journal of the American Chemical Society*. 1963; **85**:2870-2871
- [5] Wagner G, Braun W, Havel TF, Schaumann T, Go N, Wüthrich K. Protein structures in solution by nuclear magnetic resonance and distance geometry—The polypeptide fold of the BPTI determined using two different algorithms, DISGEO and DISMAN. *Journal of Molecular Biology*. 1987;**196**: 611-639
- [6] Pachler KGR. Nuclear magnetic resonance study of some  $\alpha$ -amino acids—II. Rotational isomerism. *Spectrochimica Acta*. 1964;**20**:581-587
- [7] Nagayama K, Wüthrich K. Structural interpretation of vicinal proton-proton coupling constants  $^3J_{H\alpha H\beta}$  in the basic pancreatic trypsin inhibitor measured by two-dimensional  $J$ -resolved NMR spectroscopy. *European Journal of Biochemistry*. 1981;**115**:653-657
- [8] Kessler K, Griesinger C, Wagner K. Peptide conformations 42. Conformation of side chains in peptides using heteronuclear coupling constants obtained by two-dimensional NMR spectroscopy. *Journal of the American Chemical Society*. 1987;**109**:6927-6933
- [9] Schmidt JM. Transforming between discrete and continuous angle distribution models: Application to protein  $\chi_1$  torsions. *Journal of Biomolecular NMR*. 2012;**54**:97-114
- [10] Bracewell RN. *The Fourier Transform and its Applications*. 2nd International ed. New York: McGraw-Hill; 1986
- [11] Schmidt JM, Löhr F.  $^3J$  coupling constants related to  $\chi_1$ -torsions in Ribonuclease T<sub>1</sub> from *Aspergillus oryzae*. Biological Magnetic Resonance Data Bank. BMRB Entry 26067 pending release on publication. Available from: [http://www.bmrwisc.edu/data\\_library/summary/index.php?bmrwId=26067](http://www.bmrwisc.edu/data_library/summary/index.php?bmrwId=26067)
- [12] Pérez C, Löhr F, Rüterjans H, Schmidt JM. Self-consistent Karplus parametrization of  $^3J$  couplings depending on the polypeptide side-chain torsion  $\chi_1$ . *Journal of the American Chemical Society*. 2001;**123**:7081-7093
- [13] Foley JD, van Dam A, Feiner SK, Hughes JF, Phillips RL. *Introduction to Computer Graphics*. 13th Reprint. Boston: Addison-Wesley Publishing Company Inc.; 1997
- [14] Schmidt JM, Blümel M, Löhr F, Rüterjans H. Self-consistent  $^3J$  coupling analysis for the joint calibration of Karplus coefficients and  $\phi$ -torsion angles. *Journal of Biomolecular NMR*. 1999;**14**:1-12