Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 189 (2007) 90–103
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Periodicity, planarity, and pixel (3P): A program using the intrinsic residual dipolar coupling periodicity-to-peptide plane correlation and phi/psi angles to derive protein backbone structures

Jinbu Wang [a], Joseph D. Walsh [a], John Kuszewski [b], Yun-Xing Wang [a,*]

[a] Protein Nucleic Acid Interaction Section, Structural Biophysics Laboratory, NCI-Frederick, NIH, Frederick, MD 21702, USA
[b] Imaging Sciences Laboratory, Division of Computational Biology, CIT, NIH, Bethesda, MD 20892, USA

Received 19 June 2007; revised 21 August 2007
Available online 6 September 2007

Abstract

We present a detailed description of a theory and a program called 3P. "3P" stands for periodicity, planarity, and pixel. The 3P program is based on the intrinsic periodic correlations between residual dipolar couplings (RDCs) and in-plane internuclear vectors, and between RDCs and the orientation of peptide planes relative to an alignment tensor. The program extracts accurate rhombic, axial components of the alignment tensor without explicit coordinates, and discrete peptide plane orientations, which are utilized in combination with readily available phi/psi angles to determine the three-dimensional backbone structures of proteins. The 3P program uses one alignment tensor. We demonstrate the utility and robustness of the program, using both experimental and synthetic data sets, which were added with different levels of noise or were incomplete. The program is interfaced to Xplor-NIH via a "3P" module and is available to the public. The limitations and differences between our program and existing methods are also discussed.
Published by Elsevier Inc.

Keywords: Residual dipolar coupling; RDC wave; Peptide plane; RDC–PP correlation

---

1. Introduction

Since the end of the last decade, residual dipolar couplings (RDCs) have proved their utility in a number of different areas, including the validation of NMR and X-ray structures [1,2]; direct refinement of structures using distance, torsion angle and RDC restraints [3–7]; direct structure determination with two alignment tensors [8]; and protein folding [8–12]. Since RDCs contain information about the orientation of internuclear bond vectors in relation to the alignment tensor, the RDC constraints are independent from one another.

A caveat in the use of RDCs in the direct refinement protocol is the degeneracy problem [1]. In general, any orientations of a bond vector along a cone about the principle alignment tensor axis and its inversion give rise to the same RDC value [7]. Consequently, the continuum of possible orientations represented by the cones leads to an extremely large number of local minima. These multiple minima make global structure determination by a simulated annealing protocol in which only RDCs are the main source of restraints impractical.

Using peptide planes as structural building units has been reported [13–15]. The solid-state NMR method developed mainly by Opella et al. yields two polar angles, which are not sufficient to specify the orientation of a plane [14]. To overcome this insufficiency, they utilized the Ramachandran energy of phi and psi angles to discern the plane orientations [14]. The approach of Quine and Cross applied an elegant scheme that "connects the dots" of backbone atom positions and used dipolar and ^{15}N chemical shift data to reduce the number of possible orientations of a biplane to four [16]. Mueller et al. used peptide planes as

---

[*] Corresponding author. Fax: +1 301 846 6231.
E-mail address: wangyu@ncifcrf.gov (Y.-X. Wang).

1090-7807/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.jmr.2007.08.018

a unit to interpret RDC data and to refine an existing NOE-based structure [17,18]. Hus et al. employed the concept of a chiral motif (a peptide plane followed by a tetrahedral center) and two alignment tensors to derive the backbone structure of ubiquitin [8].

Opella's group reported observations of a "dipolar wave," in which dipolar couplings measured in both solid- and solution-state NMR show sinusoidal oscillations when plotted against residue numbers in α-helical peptides [19–22]. Dipolar waves have been used to derive the orientation information of alignment tensor coordinates in relationship to the molecular frame using an empirical fitting function [20–22] or an exact analytical expression (RDC–periodicity correlation) [23,24]. The information that can be obtained from dipolar waves is intriguing. The dipolar waves provide a direct correlation between structures and experimentally measured dipolar couplings or RDCs from respective fully/partially aligned samples. In addition to the RDC–periodicity correlation, the RDCs and bond vectors of a peptide plane are also periodically correlated (RDC–planarity correlation) [25]. Both the RDC periodicity and planarity correlations (RDC–PP correlations) provide restraints to define peptide plane orientations with three angles (Fig. 1): the tilt angle δ; the phase angle ρ of the peptide plane normal vector; and the pitch angle γ of an in-plane bond vector in regular secondary-structure regions [25].

The intuitive periodicity correlation between a regular secondary structure and an RDC vanishes in non-regular secondary-structure regions. Nevertheless, the correlation between specific peptide plane orientations and RDCs can be established with the RDC–PP correlations complemented by readily available phi/psi dihedral torsion angles. The combined use of the RDC–PP correlations and phi/psi angle predictions leads to determination of peptide plane orientations. These oriented peptide planes are subjected to constraints of the covalent peptide bond lineage. In essence, peptide plane orientations $O(\delta,\rho,\gamma)$ are used as the protein's structural "pixels," analogous to the computer graphic pixels (contrast, brightness, and hue), and form the basis for determining backbone structures of proteins.

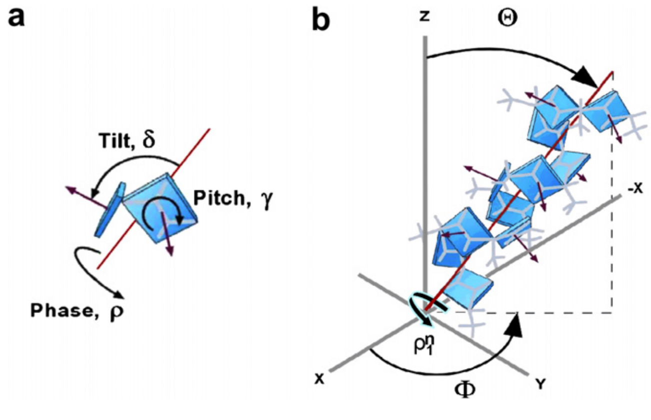

Fig. 1. (a) The definition of the peptide plane orientation tilt (δ^n), phase (ρ^n), and pitch (γ^n). The orientation of the peptide plane normal vector $\hat{n}$ is determined by the tilt and phase in the usual spherical coordinate sense according to $\hat{n} = (\sin\delta^n\cos\rho^n, \sin\delta^n\sin\rho^n, \cos\delta^n)$. The pitch is the clockwise rotation of the bond vector $\hat{r}^{\mathrm{AB}}$ about $\hat{n}$, which determines the pitch of a helix at that plane. (b) An α-helix backbone structure defined by consecutive peptide plane orientations. The plane normal vectors, which define the plane orientations, are indicated with red arrows.

2. Theory

2.1. Periodicity and plane orientation in periodical regular secondary structures

Peptide plane normal vectors can be treated as pseudo-bond vectors, whose orientation varies periodically, similar to the peptide plane orientation in regular secondary structures. A correlation exists between the periodic behavior of these normal vectors and the periodicity in the RDC wave. The RDC–PP correlations encompass the intricate correlations among the RDC periodicity, secondary-structure periodicity, and periodicity of the bond vectors within the peptide plane [25]. These correlations can be exploited once the peptide plane normal vectors are expressed in terms of three angles, namely the phase ρ, tilt δ, and pitch γ angles (Fig. 1a). The peptide plane orientation defined with the RDC–PP correlations represents a unifying principle that allows the unambiguous interpretation of RDC data [13–16,26].

Bond vector orientations directly correlate with their RDCs when the orientation of the alignment frame in relationship to the molecular frame is known [23,24]. The extraction of orientational information from the RDCs is made possible by the explicit analytical equation (Eq. (1)) that expresses the RDC, D_{AB}, in relation to the bond vector in the alignment tensor coordinates [24]. When considering structural elements of known types, such as an α-helix, or a duplex in nucleic acids, D_{AB} can be expressed in terms of the bond vector orientation (δ_i,ρ_i) in relationship to a secondary-structure axis that is oriented at angles (Θ,Φ) with respect to the alignment frame [23,24].

$$\begin{aligned} D_{\mathrm{AB},i} &= C_1(\Theta,\Phi,\delta_i)\cos 2\rho_i + C_2(\Theta,\Phi,\delta_i)\sin 2\rho_i \\ &\quad + C_3(\Theta,\Phi,\delta_i)\cos\rho_i + C_4(\Theta,\Phi,\delta_i)\sin\rho_i \\ &\quad + C_5(\Theta,\Phi,\delta_i) \end{aligned} \qquad (1)$$

Eq. (1) is universally applicable to any periodic structural element. In this equation, $\rho_i = (\rho_1 + 2\pi(i-1)/T)$ is the phase of the bond vector of the ith residue, which is related to the phase of the bond vector of the first residue, ρ_1, and the period $T \approx 3.6$ residues/turn for the α-helix (Fig. 1b) and $T \approx 2$ residues/turn for a β-strand. The slant angle δ_i is the angle the bond vector AB makes with the secondary-structure element axis, and the coefficients $C_k = C_k\ (\Theta,\Phi,\delta_i)$ are functions of both the helical axis orientation and δ_I [24]. We use subscript index i to indicate