ELSEVIER

Biophysical Chemistry 120 (2006) 24 – 31

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Role of amino acid properties to determine backbone τ(N–Cα–C′) stretching angle in peptides and proteins

S.M. Malathy Sony, K. Saraboji, N. Sukumar [1], M.N. Ponnuswamy *

Department of Crystallography and Biophysics, University of Madras, Guindy campus, Chennai — 600025, India

Received 8 June 2005; received in revised form 27 July 2005; accepted 28 July 2005
Available online 2 November 2005

Abstract

The analysis of the basic geometry of amino acid residues of protein structures has demonstrated the invariability of all the bond lengths and bond angles except for τ, the backbone N–Cα–C′ angle. This angle can be widened or contracted significantly from the tetrahedral geometry to accommodate various other strains in the structure. In order to accurately determine the cause for this deviation, a survey is made for the τ angles using the peptide structures and the ultrahigh resolution protein structures. The average deviation of N–Cα–C′ angles from tetrahedral geometry for each amino acid in all the categories were calculated and then correlated with forty-eight physiochemical, energetic and conformational properties of amino acids. Linear and multiple regression analysis were carried out between the amino acid deviation and the 48 properties. This study confirms the deviation of τ angles in both the peptide and protein structures but similar forces do not influence them. The peptide structures are influenced by physical properties whereas as expected the conformational properties influence the protein structures. And it is not any single property that dominates the deviation but the combination of different factors contributes to the τ angle deviation.

Keywords: Bond angle deviation; Amino acid properties; Correlation analysis; Peptides; Ultrahigh resolution protein structures

1. Introduction

During the past few decades, a number of technical and methodological improvements have taken place in macromolecular crystallography. High-intensity synchrotron radiation sources, efficient two-dimensional detectors and cryogenic techniques are employed in collecting high atomic resolution protein structures of increasing size and complexity. The level of accuracy of these three-dimensional protein structures provides relevant information on biological function and catalysis of the particular enzymes under study. From these unbiased studies, evidences for general and fundamental molecular properties can be derived. Recent studies on the geometry of the polypeptide backbone and on its experimental electron density have demonstrated the potential of crystallography at ultrahigh resolution [1]. In fact, the deviations in backbone bond angle τ(N–Cα–C′) [2], the non-planarity of peptide bond [3], the pyramidalization of carbonyl carbon atom [4] and the lengthening of CO bonds when the CN bond shortens [1] were all identified from the ultrahigh resolution protein structures. These analyses were made keeping in mind the increased need for accurate, efficient, and reliable methods to model protein structures for the structural genomics efforts [5].

Studies on the peptide group were carried out owing to its properties, such as the dipole moment, the planar geometry, and the relatively high rotational barrier around C–N bond, which determine the conformation of polypeptides and proteins [4]. It is the Cα atom, the most important locus for evaluating distortion of covalent geometry in protein structures, that joins sidechain with backbone and respond to both and especially to their compatibility [6]. Database analysis carried out by Karplus [2] showed that the average values of the bond lengths, bond angles, and ω-torsion angles match the ideal values based on small-molecule crystal structures but for well-populated main-chain conformations, these parameters vary over ranges of about 0.015 Å, 4° and 7° (up to 8.8° for the τ angle), respectively. The highly variable interpeptide bond angle τ has relatively large effects on structure because the

* Corresponding author. Tel.: +91 44 2235 1367; fax: +91 44 2230 0122.
E-mail address: mnpsy2004@yahoo.com (M.N. Ponnuswamy).
[1] Presently working in NE-CAT, Building 436E, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA.

0301-4622/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2005.07.012

peptide group magnifies the resulting atomic shifts of ~10° to a 0.31 Å shift in the relative positions of 1–3 related Cα-atoms [2]. This shift may affect any crucial non-covalent interaction that is likely to take place and therefore analysis of τ angle is necessitated.

In the high resolution protein structures this τ(N–Cα–C′) bond angle deviation from tetrahedral geometry was correlated with the change in the conformation of the backbone torsion angles ϕ (N–Cα) and ψ (Cα–C′). It was well established as early as in 1965 by Ramakrishnan and Ramachandran [7] that the allowed domain of the conformation map [(ϕ,ψ) map] of the dipeptide unit increases when τ(N–Cα–C′) angle is increased from 110° to 115°. But in the mid-1990s it was concluded that the widening or contraction of τ angle actually depends on the backbone conformation (ϕ,ψ) of the protein structure [8–11,2]. As one moves from one point in (ϕ,ψ) space to another, bond angle τ vary in a characteristic manner which was consistently found to be contracted in extended forms, intermediate in C_7^{eq}, and relatively large in the helical-bridge regions of peptide (ϕ,ψ) space, with variations far exceeding 10° [8,2]. If large variations in an angle of this kind are neglected in peptide modeling along a chain of hundreds of residues, this may have significant implications for the spatial presence of a protein, or the modeling of an active site [8].

Although deviations from standard geometry are seen in small molecule crystal structures (e.g., [12,13]), and some conformational dependence has been inferred [14], these data are limited and the general relevance of such deviations has remained uncertain. Contradiction was also observed with respect to the role of aminoacid residue type influencing the τ angle variation, but some workers [15] observe only weak dependence whereas Ramachandran et al. [16] and Momany et al. [17] do find differences. Most notable difference is observed for *Gly* and *Pro* having high values [18] and *Val* and *Ile* having low values [2]. Aminoacids are subjected to certain constrained conditions imposed by their physical and chemical properties, which might play an important role in protein structure and function [19]. With these points in mind, the following work was initiated.

The τ(N–Cα–C′) bond angle for the twenty naturally occurring aminoacid residue types were retrieved from small molecule peptide structures and its average deviation from tetrahedral geometry was calculated. This was then correlated with various physicochemical, energetic and conformational properties of the aminoacids. A careful study of τ(N–Cα–C′) angles on aminoacids and peptides will give an idea about the maximum range to which the angle can stretch in individual aminoacids. Such a study has been undertaken as they are relevant in modeling of polypeptides and protein chains and also in protein refinement.

2. Materials and methods

To enhance our understanding of peptide energetics, it is important to determine the covalent geometry changes as a function of aminoacid properties. These results can, in turn, stimulate theoretical studies to provide explanations for these observations [2]. Calculated geometries have the advantage of providing information on structure–property relations which is difficult to obtain experimentally. Various procedures for deriving and applying molecular geometries have been described in the literature [11]. These geometries are obtained from small-molecular studies, surveys of protein structures, relationships to other proteins, or from experimental observations made from X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy [5]. Herein, a statistical approach is adopted, in which the τ angle average deviation for each of the twenty aminoacid residues are related with various physicochemical, energetic and conformation properties of aminoacids through single and multiple correlation factors [20–23].

In order to access the deviation of backbone τ(N–Cα–C′) bond angle with respect to various aminoacid properties for small-molecular peptide structures, the dataset (referred henceforth as 'peptide dataset') were chosen as follows: Cambridge Structural Database (CSD: [24]), update 5.24 (November 2002), was searched for peptides, excluding data from cyclic, disordered or D-aminoacids containing peptides, as well as structures with the crystallographic error factor $R > 0.08$ and *Cys* containing peptides there should be no disulphide bridge. The τ(N–Cα–C′) bond angles were retrieved and its average deviation from tetrahedral geometry (d) was calculated using the formula:

$$d = \sum \frac{|\tau - 109.5|}{N} \tag{1}$$

where N is the total number of observables and the value 109.5 is the standard angle for the tetrahedral geometry, in radians.

The single correlation between the τ angle average deviation, d (dependent quantity) and aminoacid properties (independent quantities) was calculated using the familiar expression:

$$r = \frac{N\sum XY - \left(\sum X \sum Y\right)}{\left\{\left[N\sum X^2 - \left(\sum X\right)^2\right]\left[N\sum Y^2 - \left(\sum Y\right)^2\right]\right\}^{1/2}} \tag{2}$$

where r is the correlation coefficient, N, X, and Y are the number of data, average deviation and aminoacid properties, respectively. To evaluate the validity of the correlation coefficient, the so-called null hypothesis, i.e., the hypothesis in which the variables are not correlated, was tested by using Student's t distribution. The statistical test yields a p-value, which represents the probability that random sampling would result in a correlation coefficient. Under this hypothesis there is no correlation between the two variables; p-values <0.05 allow one to reject the null hypothesis at the 95% confidence level [1]. The combined effect of these aminoacid properties towards the τ angle deviation are analysed using multiple regression. Multiple correlation coefficients were determined by standard procedures [25,26].