Amino Acids (2007) 33: 689–693
DOI 10.1007/s00726-006-0463-2
Printed in The Netherlands

Amino Acids

Using C′ deviation to study structures of central amino acids in peptide fragments

H.-Y. Tang and **Z.-G. Zhang**

School of Basic Medicine, Peking Union Medical College, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Beijing, China

Received September 15, 2006
Accepted October 15, 2006
Published online November 30, 2006;

Summary. In this investigation, we attempted to study the backbone geometry of amino acids in peptides using C′ deviation. Diameters of distribution were used to describe the various atomic structures, and scatter graphs provided visual evaluation. The length of peptide fragments and the secondary structure of amino acids in the central position of the peptide fragments were also analyzed. The results showed that the atomic distribution of the central amino acids of five-residue peptide fragments was much more restricted than that of their corresponding three-residue peptide fragments. In identical three-residue fragments, atoms of central amino acids with different secondary structures, were distributed in distinct areas.

Keywords: Backbone geometry – C′ deviation – Peptide fragment – Fragment length – Secondary structure

Introduction

It is well established that the conformation of polypeptides and proteins are determined by their amino acid properties (Esposito et al., 2000a). Recent studies on the geometry of amino acids in peptides usually calculated the deviation of backbone torsion angles $\phi(\mathrm{N}-\mathrm{C}\alpha)$ and $\Psi(\mathrm{C}\alpha-\mathrm{C}')$ (Petrescu et al., 2000) or backbone bond angles $\tau(\mathrm{N}-\mathrm{C}\alpha-\mathrm{C}')$ (Esposito et al., 2000b). The $\tau(\mathrm{N}-\mathrm{C}\alpha-\mathrm{C}')$ bond angle deviation from tetrahedral geometry was correlated with the change in conformation of the backbone torsion angles $\phi(\mathrm{N}-\mathrm{C}\alpha)$ and $\Psi(\mathrm{C}\alpha-\mathrm{C}')$ (Malathy Sony et al., 2006). In this study, we attempted to use C′ deviation to study the geometry of amino acids in peptides.

For each amino acid, N, Cα, and C′ are on the backbone of the polypeptide, and they lie on two neighboring peptide planes (Fig. 1). N and Cα lie on the same peptide plane together with $\mathrm{C}\alpha_{-1}$, C'_{-1} and O_{-1} from the previous amino acid in the linear sequence of the peptide. Cα, C′ and O, together with N_{+1} and $\mathrm{C}\alpha_{+1}$ from the next amino acid, are on the same peptide plane. Currently, the dihedral angle between two neighboring peptide planes could be described by backbone torsion angles (ϕ, Ψ) or bond angle (τ). The relative distribution of atoms in the same peptide plane is highly restricted. On each peptide plane, besides backbone torsion angles (ϕ, Ψ) or bond angle (τ), the distribution of C′ relative to that of the previous peptide plane could also reflect the dihedral angles between these two peptide planes directly.

Three-dimensional structures of proteins are determined by their amino acid sequences (Anfinsen, 1973). Much work has been done to study the relationship between peptide fragments and their configurations (Kabsch and Sander, 1984; Chothia and Lesk, 1986; Bystroff et al., 1996; Frishman and Argos, 1996; Crooks et al., 2004). In this investigation, we employed C′ deviation to study the structural diversity of central amino acids in small peptide fragments. The structural diversity of the amino acid in the center of five-residue peptide fragments and their corresponding three-residue fragments was studied by comparing the distribution of C′. In identical three-residue peptide fragments, central amino acids with different secondary structures were also studied. Our results showed that the distribution of C′ of the central amino acids from five-residue peptide fragments was much more restricted than that from their corresponding three-residue peptide fragments which constitute the cores of the five-residue fragments. In identical three-residue fragments, atoms of central amino acids, which have different secondary

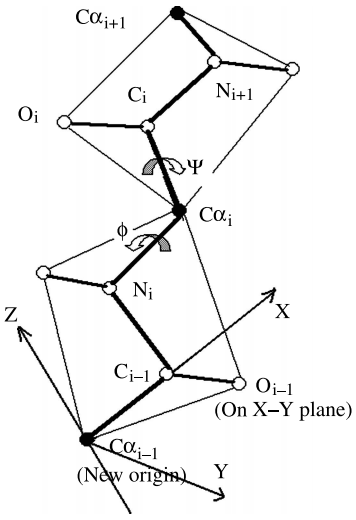

Fig. 1. Peptide plane and the frame of reference. Here C_i means atom C from residue i in linear sequence, and C_{i-1} means atom C from the previous one of residue i. The origin of the frame of reference is $C\alpha_{-1}$, the X axis is from $C\alpha_{-1}$ to C'_{-1}, and the X–Y plane is $C\alpha_{-1}$, C'_{-1}, O_{-1}

structures, were restricted to distinct areas. This study confirms that local sequence similarity could affect structures of amino acids greatly.

Materials and methods

Dataset

The dataset that we used was a non-redundant subset of the Protein Data Bank (PDB) (Berman et al., 1998). The similarity between any two sequences in the dataset was less than 40%. Three-dimensional structures of proteins in PDB were mostly determined by X-ray crystallography (Smyth and Martin, 2000) or nuclear magnetic resonance (NMR) spectroscopy (Kay, 2005). X-ray crystallography is currently the most favored technique for structure determination of proteins and biological macromolecules. Many programs, such as X-PLOR (Badger et al., 1999), SHELXL, NUCLSQ, have been developed to refine the structure data from X-ray crystallography or NMR spectroscopy. X-PLOR seemed to be the best in "comparison of the programs NUCLSQ, PROLSQ, SHELXL93 and X-PLOR" carried out by Schuerman (Schuerman et al., 1996). In order to counteract the errors in different structure determining methods, 1217 structures in the dataset were all determined by X-ray crystallography and refined by X-PLOR.

A fragment library was created to store all the three- and five-residue peptide fragments extracted from the dataset. Information on the secondary structure of central amino acids of these fragments was also deposited in the library.

Transformation of coordinates

Coordinates of C′ need to be transformed in conformity with the relative coordinates of the previous peptide plane which is selected as the reference plane. The reference plane contains two atoms $(N, C\alpha)$ from the same amino acid containing C′, and three atoms $(C\alpha_{-1}, C'_{-1}, O_{-1})$ from the previous amino acid. $C\alpha_{-1}, C'_{-1}$, and O_{-1} are selected as the frame of reference for transformation of coordinates, because not only $C\alpha_{-1}$, C'_{-1}, and O_{-1} are on the same peptide plane with N and Cα, but also because this could reflect the relative distribution of these two neighboring amino acids in the peptide (Fig. 1). The transformation could be divided into three steps:

1. Use $C\alpha_{-1}$ as the new origin.
2. Use $C\alpha_{-1}$ to C'_{-1} as the new X axis.
3. Use $C\alpha_{-1}$, C'_{-1}, O_{-1} as the new X–Y plane.

Since N and Cα are on the same peptide plane with the frame of reference, coordinates of this two atoms were highly restricted after transformation (Fig. 2). The transformed coordinates are based on the coordinates of the previous amino acids, so they also reflect the interrelationship between these two amino acids.

Results

Fragment length

Peptide fragments longer than five amino acids were infrequent in the library. Hence we compared the structure

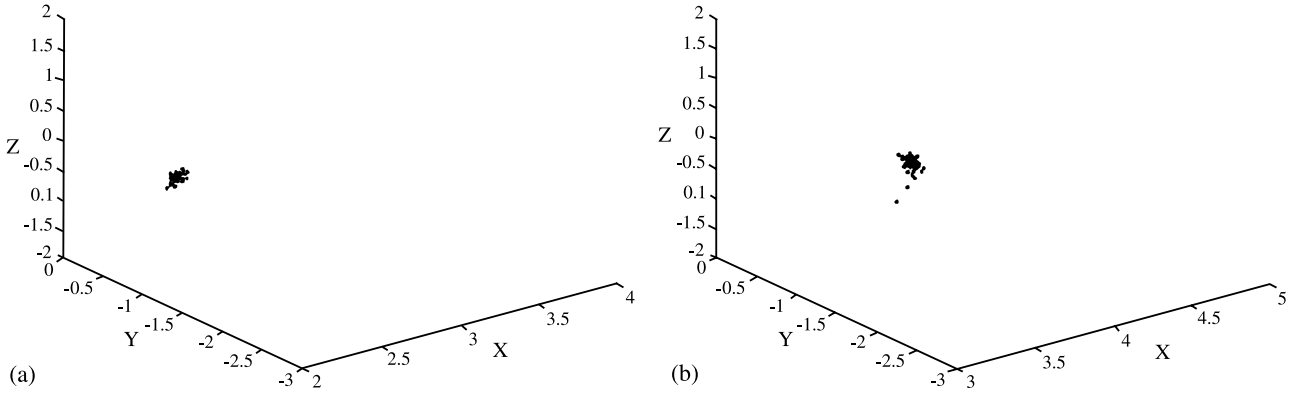

Fig. 2. Distribution of atoms "N" and "Cα" of the central amino acid "K" in the fragment "AKN". It is evident that both "N" and "Cα" show very restricted areas of distribution. **a** Distribution of atom N from residue "K" in fragment "AKN"; **b** Distribution of atom Cα from residue "K" in fragment "AKN". Units: angstrom

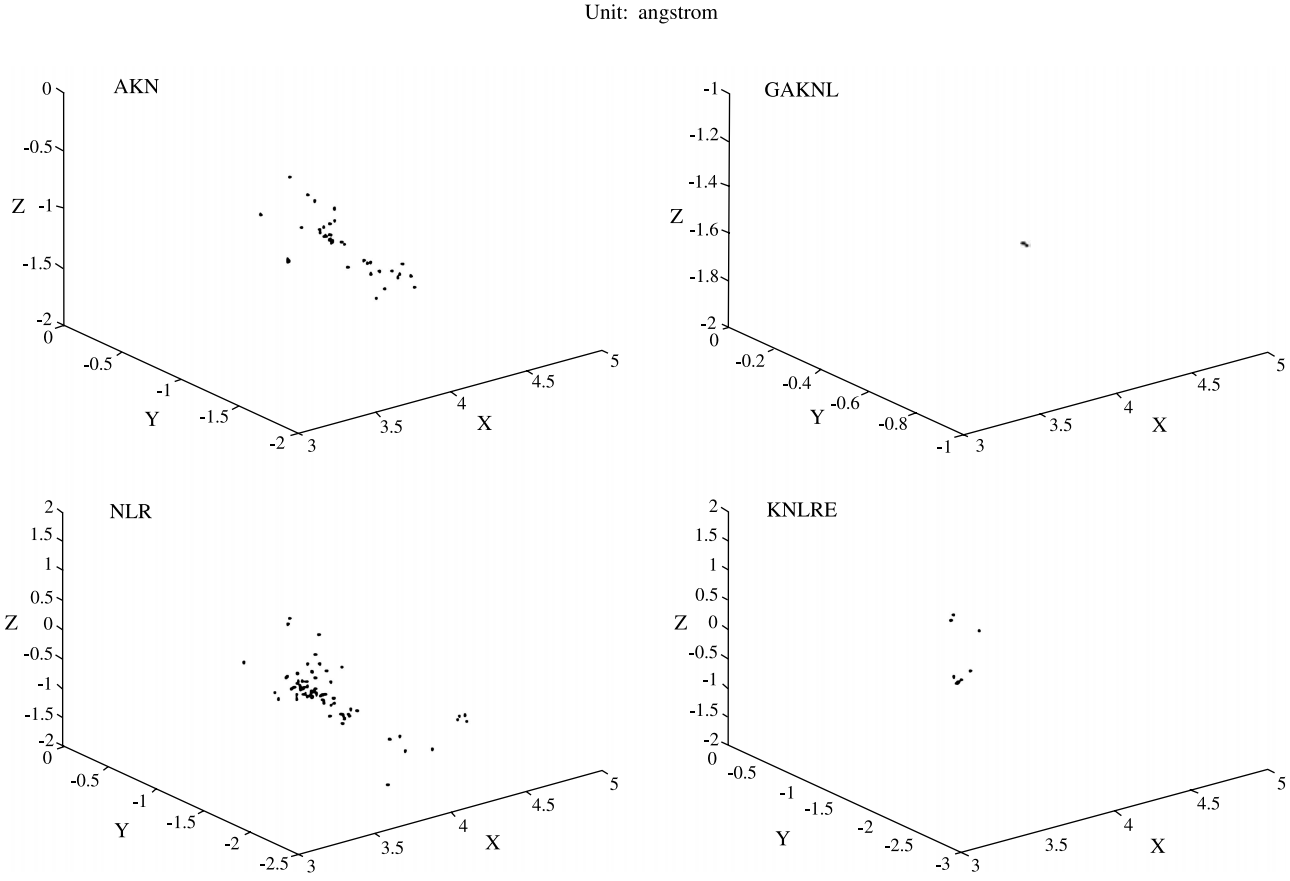

Fig. 3. Distribution of atom C′ from central amino acids. It is obvious that the distribution is much more restricted in five residue peptide fragments than that in their corresponding three residue peptide fragments

of the central amino acid in three- and five-residue peptide fragments to study how fragment length affects the structure of central amino acids. Figure 3 demonstrates that structure diversity in five-residue peptide fragments was remarkably less than that in the corresponding three-residue peptide fragments. Table 1 shows the diameters of C′ distribution of central amino acids in five amino acid fragments and that in their corresponding three amino acid fragments. Obviously, the distribution of C′ was more restricted in five amino acid fragments than in three amino acid fragments.

Table 1. Diameter of C′ distribution of central amino acids in five- and three-residue peptide fragments

Secondary structure	Fragments (Five)	Diameters (Å)	Fragments (Three)	Diameters (Å)
HELIX	AAAEA	0.2112	_AAE_	2.7725
HELIX	AAVEE	0.5656	_AVE_	2.6150
HELIX	AIGYA	0.4176	_IGY_	0.4530
HELIX	AKNLR	0.0296	_KNL_	2.0848
HELIX	EELKA	0.2930	_ELK_	2.2541
HELIX	ELVAI	0.5771	_LVA_	2.1709
SHEET	ELVAI	1.0941	_LVA_	1.6316
HELIX	GAKNL	0.0264	_AKN_	1.1608
HELIX	IEILR	0.2555	_EIL_	1.7253
HELIX	ILRDE	1.1555	_LRD_	1.8089
HELIX	KAAAE	0.2828	_AAA_	2.7725
HELIX	KNLRE	0.8889	_NLR_	2.3901
HELIX	LNKLL	0.3109	_NKL_	1.6926
HELIX	LREAI	0.2063	_REA_	2.6121
SHEET	VKVGD	0.9309	_KVG_	2.2458
HELIX	YADSV	0.0131	_ADS_	2.2324

Secondary structure

The central amino acids in identical fragments with different secondary structures have distinct distribution areas. Since the central amino acids in identical five length fragments mostly have the same secondary structures, we therefore used three-residue peptide fragments to study the relationship between secondary structure and structure diversity. Figure 4 shows the distribution of C′ of central amino acids in several three amino acid fragments. Obviously, the distribution of C′ in HELIX is remarkably different from that in SHEET.

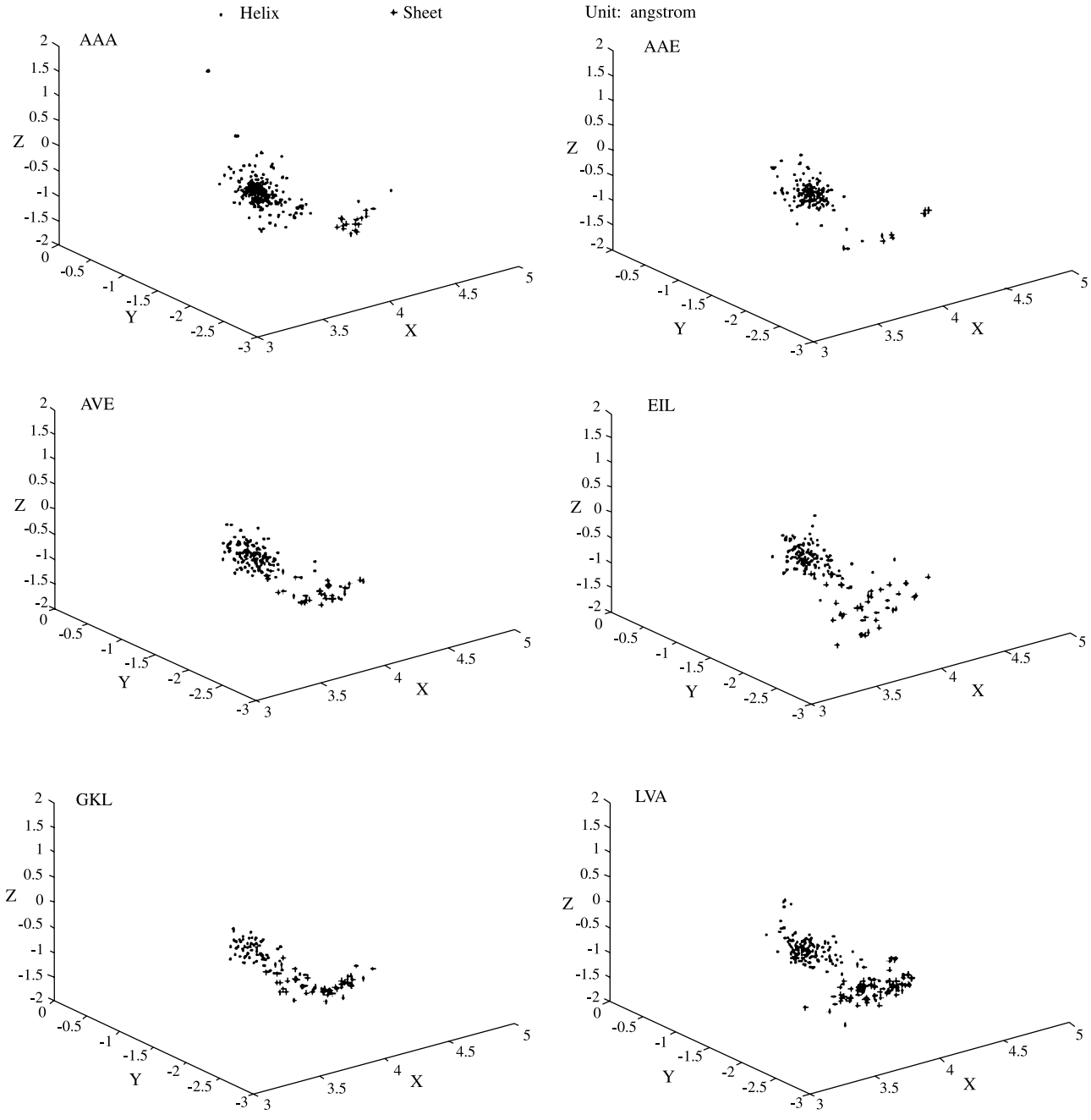

Fig. 4. Distribution of C′ from the central amino acids with different secondary structures. Mostly, the distributions show distinct patterns. Yet sometimes, like the last one, distribution areas might overlap

Discussion

C′ deviation, a new method to study the backbone geometry of amino acids in peptides was illustrated in this paper. The structural diversity of amino acids could be easily analyzed by comparing the distribution of atom C′. The results are more evident visually and more tangible than backbone torsion angles (ϕ and Ψ) or backbone bond angles τ.

Relationships between fragment length as well as secondary structure and structure diversity of single amino acids were studied in this paper. Longer length of peptide fragments could restrict the distribution of atoms to a greater degree, probably because longer length leads to more stable local structural environment and local interaction. The atoms of central amino acids in identical fragments with different secondary structures show distinct

distribution areas though with some exceptions – sometimes their distribution areas could overlap (last one in Fig. 4). This situation might be due to insufficient representation of non-local interaction when secondary structure only is taken into account. Some of the non-local interactions, such as –S–S–, are very complicated, and their effects on the structure of amino acids need to be analyzed further more.

The results indicate that the structure of central amino acids would be greatly restricted if the similarity between two peptide fragments is high enough. This conclusion would be quite useful for protein structure prediction. A new prediction method is under development base on it.

Acknowledgements

We would like to thank Prof Chenyu Jiang and Dr Xiaolin Yang for their help and suggestions. We also thank Prof Wilson Lo for his valuable advice. This project is supported by Chinese Medical Bords of New York, Inc #03-787.

References

Anfinsen C (1973) Principles that govern the folding of protein chains. Science 181: 223–230

Badger J, Kumar RA, Yip P, Szalma S (1999) New features and enhancements in the X-PLOR computer program. Proteins 35: 25–33

Berman HM et al. (1998) The protein data bank. Nucleic Acids Res 28: 235–242

Bystroff C, Simons KT, Han KF, Baker D (1996) Local sequence-structure correlations in proteins. Curr Opin Biotechnol 7: 417–412

Chothia C, Lesk AM (1986) The relation between the divergence of sequence and structure in proteins. EMBO J 5: 823–826

Crooks GE, Wolfe J, Brenner SE (2004) Measurements of protein sequence-structure correlations. Proteins Struct Funct Bioinform 57: 804–810

Esposito L, Vitagliano L, Zagari A, Mazzarella L (2000a) Pyramidalization of backbone carbonyl carbon atoms in proteins. Protein Sci 9: 2038–2042

Esposito L, Vitagliano L, Zagari A, Mazzarella L (2000b) Experimental evidence for the correlation of bond distances in peptide groups detected in ultrahigh-resolution protein structures. Protein Eng 13: 825–828

Frishman D, Argos P (1996) Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. Protein Eng 9: 133–142

Kabsch W, Sander C (1984) On The use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. Proc Natl Acad Sci USA 81: 1075–1078

Kay LE (2005) NMR studies of protein structure and dynamics. J Magn Reson 173: 193–207

Malathy Sony SM, Saraboji K, Sukumar N, Ponnuswamy MN (2006) Role of amino acid properties to determine backbone τ(N−Ca−C′) stretching angle in peptides and proteins. Biophys Chem 120: 24–31

Petrescu AJ, Calmettes P, Durand D, Receveur V, Smith JC (2000) Change in backbone torsion angle distribution on protein folding. Protein Sci 9: 1129–1136

Schuerman GS, Smith CK, Turkenburg JP, Dettmar AN, Van Meervelt L, Moore MH (1996) DNA-drug refinement: a comparison of the programs NUCLSQ, PROLSQ, SHELXL93 and X-PLOR, using the low-temperature d(TGATCA)-nogalamycin structure. Acta Crystallogr D Biol Crystallogr 52: 299–314

Smyth MS, Martin JH (2000) X-ray crystallography. Mol Pathol 53: 8–14

Authors' address: Zheng-Guo Zhang, School of Basic Medicine, Peking Union Medical College, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, China,
Fax: +8610-65296436, E-mail: z.g.zhang@ieee.org