The amino acid dipeptide: Small but still influential after 50 years

Jan Hermans[1]

Department of Biochemistry, University of North Carolina, Chapel Hill, NC 27599-7260

The conformation of a polypeptide chain can be described with good accuracy in terms of a unique set of values of backbone torsion angles—two for each amino acid residue—called ϕ and ψ (Fig. 1); bond lengths and bond angles are given fixed canonical values, and the peptide bond linking successive residues also is fixed as a planar structure. As was shown by Ramachandran and coworkers nearly 50 y ago (1), it is easy to map distributions of polypeptide conformations expressed in terms of the torsion angles ϕ and ψ in 2D graphs, which have since become known as Ramachandran plots. These authors analyzed the conformations available to a single residue in a polypeptide chain in terms of a simple model that included, besides a single amino acid residue, parts of the neighboring residues as far as the immediately preceding and succeeding α-carbon atoms and, thus, included the two planar peptide groups; to this model they gave the (unsystematic) name of dipeptide. When they assumed standard atomic radii and disallowed conformations with atomic overlap, they found that relatively few combinations of the two variable torsion angles produced favorable structures without atomic overlap (shaded gray in Fig. 1). In PNAS, Avbelj and coworkers (2) present new information on the distribution of peptide conformation in solution.

Ramachandran plots for dipeptides of residues with longer side chains were found to be very much like that for alanine. (Although plots for glycine and proline are quite different.) Conformations of experimentally known helical structures—such as right-handed α-helix, β-structure, and polyproline-II helix—were all found to lie within the favorable range of the plot (Fig. 1, black circles). As high-resolution structures of proteins have become known, the residue conformations have been found to distribute according to the outlines in the Ramachandran diagram, with most residues in favored regions; fewer residues in allowed, but less favored, regions; and very few residues with disallowed conformations (3).

In contrast with the plentiful distributions in folded structures, we have scant information about the conformation of polypeptide chains in aqueous solution. Avbelj and coworkers (2) now show, >50 y after the question was raised, that the conformations of dipeptides in aqueous solution may be represented in terms of three

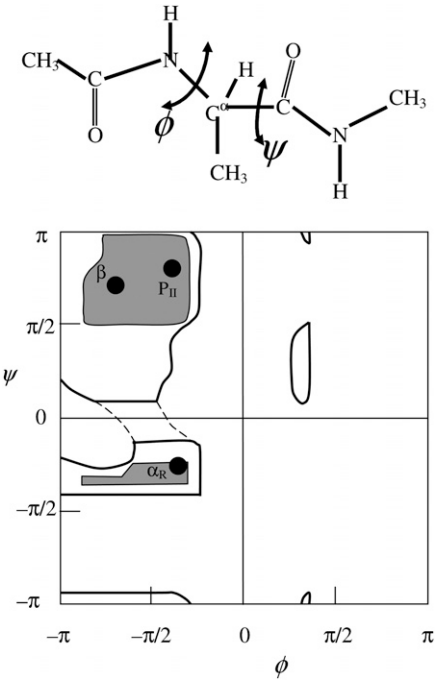


Fig. 1. (*Upper*) Alanine dipeptide (*N*-acetyl-alanyl-*N*′-methylamide). The two torsion angles forming the principal degrees of freedom are indicated. (*Lower*) Ramachandran plot showing so-called allowed and favored conformations of the alanine dipeptide. Three major helical conformations are also indicated.

structures—α_R, β, and P_{II}—that can be separated spectroscopically and present the measured proportions of each structure at equilibrium for 19 amino acids, including glycine. This result is an important new contribution: A characterization of the disordered, unfolded conformation of polypeptide chains is required to describe the properties of "natively unfolded" proteins (4) and for an understanding of the process of formation of the folded protein from the unfolded state.

Avbelj and coworkers (2) find significant differences in the proportion of the three conformations between dipeptides of different amino acids and have been able to measure changes with composition, temperature, side chain ionization state, and solvent composition, that are well known to affect the conformation of unfolded chains (5). The conformation of residues in short chains distributes approximately as that of residues in dipeptides. The interaction between residues in short chains is minimal because the C^{α} atoms of successive residues are separated by three bonds, one of which is the stiff peptide bond (6). As chains become longer, medium-range cooperative interactions favor formation of stretches of α-helical structure. However, the extent of helix formation requires a favorable amino acid composition, in terms of the intrinsic helix propensity of the residues and factors such as the presence of helix-stabilizing interactions between side chains. These conditions have been established by extensive investigations (e.g., ref. 7). As chains become longer still, additional long-range interactions between side chains are possible. In particular, attraction between hydrophobic side chains can lead to formation of collapsed but not highly ordered structures—so-called molten globules. These molten globules may also form as intermediates in the process of forming the biologically active, folded conformation of a protein from the unfolded state (8).

As yet it remains to completely relate these differences in conformational preference to differences in molecular structure and molecular interactions, something that is best approached with the aid of molecular simulations. However, the various force fields that are in wide use do not agree well among each other on the conformational distributions of alanine and glycine dipeptides in aqueous solution (9). Avbelj and coworkers (2) point out that the details of the newly measured distributions should serve as a key reference for producing a refined force field, which may then be used with increased confidence in simulations of unfolded polypeptides in solution. Most importantly, one may expect that this increased accuracy will significantly improve the accuracy of simulation of folding of small proteins with the use of atomic representation and explicit solvation, which has recently been achieved thanks to increases in computer power and improvements in simulation methods (10, 11). Routine structure determination of small proteins by simulated folding, as an alternative to X-ray crystallography and NMR spectroscopy, appears to be just around the corner. The accuracy of force fields used in these simulations will then be of greatest concern.

Author contributions: J.H. wrote the paper.

The author declares no conflict of interest.

See companion article on page 1794 in issue 5 of volume 108.

[1]E-mail: hermans@med.unc.edu.

1. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99.
2. Grdadolnik J, Mohacek-Grosev V, Baldwin RL, Avbelj F (2011) Populations of the three major backbone conformations in 19 amino acid dipeptides. *Proc Natl Acad Sci USA* 108:1794–1798.
3. Lovell SC, Word JM, Richardson JS, Richardson DC (2000) The penultimate rotamer library. *Proteins* 40: 389–408.
4. Radivojac P, et al. (2007) Intrinsic disorder and functional proteomics. *Biophys J* 92:1439–1456.
5. Hermans J Jr, Puett D, Acampora G (1969) On the conformation of denatured proteins. *Biochemistry* 8:22–30.
6. Brant DA, Flory PJ (1965) The configuration of random polypeptide chains. II. Theory. *J Am Chem Soc* 87:2791–2800.
7. Scholtz JM, Baldwin RL (1992) The mechanism of α-helix formation by peptides. *Annu Rev Biophys Biomol Struct* 21:95–118.
8. Ptitsyn OB (1995) Molten globule and protein folding. *Adv Protein Chem* 47:83–229.
9. Hu H, Elstner M, Hermans J (2003) Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. *Proteins* 50: 451–463.
10. Day R, Paschek D, Garcia AE (2010) Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins* 78:1889–1899.
11. Shaw DE, et al. (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346.