Peptide bonds revisited

For many years it has been almost dogma in the scientific community that in protein structures the planar peptide bond occurs predominantly in the *trans* conformation[1]. The occasional occurrence of a peptide bond in *cis* conformation was, in most cases, noted as a curiosity of the respective structure. This is remarkable since it became clear almost 20 years ago that the *cis*/*trans*-isomerization of peptide bonds on the N-terminal side of proline plays an important role in the folding process of a protein[2]. Systematic studies of peptide bond conformations have been hampered by the limited amount of structural information available[3], and have so far mainly focused on proline residues[4]. With more three-dimensional structures of proteins at hand today, the notion is slowly emerging that *cis* peptide bonds are by no means a curiosity, and that they may even be important determinants for the function of proteins.

We have analyzed a non-redundant set of 571 proteins from the Brookhaven protein data base[5], the structures of which have been determined crystallographically to a resolution of 3.5 Å or better. Within this set, only one in ~360 peptide bonds is reported to be in the *cis* conformation (Table 1). Most of these instances (>90%) occur where the peptide bond is an imide (Xaa–Pro) rather than an amide bond (Xaa–nonPro).

An obvious discrepancy exists, however, between the fraction of *cis* peptide bonds observed and what can be predicted from free enthalpy values. Free enthalpy differences between the *cis* and *trans* conformations have been reported to lie anywhere between 2.0 kJ mol^{-1} for Xaa-Pro bonds and 10.0 kJ mol^{-1} for Xaa–nonPro bonds[3]. Based on these values and assuming thermodynamic equilibrium at 293 K, about 30% of all Xaa–Pro bonds in acyclic peptides should occur in the *cis* conformation, and so should about 1.5% of all Xaa–nonPro bonds. These numbers are larger — by a factor of 6 for Xaa-Pro and by a factor of ~50 for Xaa–nonPro — than values determined from analyzing protein structures (Table 1).

It is noteworthy that there is a significant correlation between the resolution of the structure solved and the number of *cis* peptides detected. High resolution structures (<2.0 Å) contain almost twice the number of Xaa–Pro bonds than medium and low resolution structures (≥2.5 Å) and almost four times the number of Xaa–nonPro bonds. This striking resolution dependence as well as the above mentioned discrepancy between the fraction of *cis* bonds observed and expected lead us to suspect that many *cis* peptide bonds have not been recognized as such in the determination and refinement of the structures, especially at resolutions ≥2.5 Å. Moreover, most of the refinement programs, which have been widely used in recent years (such as X-PLOR[6]), only allow for the possibility of a *cis* conformation of an Xaa–Pro bond but will, unless specified explicitly, force any other peptide bond into the *trans* conformation. Huber and Steigemann realized this as a potential problem as early as 1974[7].

In proteins in which non-proline *cis* peptide bonds have been unequivocally identified, they often occur at or near functionally important sites and are very likely involved in the function of the molecules. One example is coagulation factor XIII[8], in which an Arg–Tyr *cis* peptide bond has been found near the active site, and a Gln–Phe *cis* peptide bond at the dimerization interface of the molecule. This strongly suggests a functional role for them, as has been proposed for the *cis* peptide bonds in carboxypeptidase A, dihydrofolate reductase and most recently in the intein gyrA[9,10].

In conclusion, we would like to emphasize two points: first, the importance of the *cis* conformation of peptide bonds in protein structures, especially if it is a Xaa–nonPro peptide bond, and second, the possibility that many *cis* peptide bonds may have passed unnoticed due to the limited resolution of the data and to the refinement protocol used.

Table 1 Occurrence and frequency of reported peptide bonds in the Brookhaven Protein Data Base (ref. 5)[1]

	All	<2.0 Å	2.0 Å–2.5 Å	≥2.5 Å
Number of proteins	571	291	184	96
Number of peptide bonds	153,209	72,567	52,194	28,448
Xaa–Pro	7,413	3,407	2,566	1,440
Xaa–nonPro	145,796	69,160	49,628	27,008
Number of *cis* peptide bonds	427 (0.28%)	232 (0.32%)	140 (0.27%)	55 (0.19%)
Xaa–Pro	386 (5.21%)	205 (6.02%)	129 (5.03%)	52 (3.61%)
Xaa–nonPro	41 (0.028%)	27 (0.039%)	11 (0.022%)	3 (0.011%)

[1]The basis of this study is a non-redundant set of 571 protein structures determined by X-ray crystallography. The set has been assembled using the PDBSELECT algorithm[11], with the conditions that the resolution of the structure be at least 3.5 Å and that the pairwise amino acid identity between molecules be less than 25%. In accordance with Stewart *et al.*[3], we excluded the entry 1HDS (sickle cell hemoglobin) from the list, but in contrast to that study, we refrained from including homologous structures even if they contained non-homologous peptide bond conformations. A peptide bond was defined to be in the *cis* conformation when the dihedral angle ω was between -45° and +45°. Shown are the data for all 571 proteins and for subsets that were determined at different resolutions.

Manfred S. Weiss, Andreas Jabs and Rolf Hilgenfeld

Institute of Molecular Biotechnology, Department of Structural Biology and Crystallography, Beutenbergstr. 11, D-07745 Jena, Germany.

Correspondence should be addressed to R.H. email: hilgenfd@imb-jena.de

1. Ramachandran, G.N. & Sasisekharan, V. *Adv. Prot. Chem.* **23**, 283–437 (1968).
2. Schmid, F.X. & Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **75**, 4764–4768 (1978).
3. Stewart, D.E., Sarkar, A. & Wampler, J.E. *J. Mol. Biol.* **214**, 253–260 (1990).
4. MacArthur, M.W. & Thornton, J.M. *J. Mol. Biol.* **218**, 397–412 (1991).
5. Bernstein, F.C. *et al. J. Mol. Biol.* **112**, 535–542 (1977).
6. Brünger, A.T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
7. Huber, R. & Steigemann, W. *FEBS Lett.* **48**, 235–237 (1974).
8. Weiss, M.S., Metzner, H.J. & Hilgenfeld, R. *FEBS Lett.* **423**, 291–296 (1998).
9. Stoddard, B.L. & Pietrokovski, S. *Nature Struct. Biol.* **5**, 3–5 (1998).
10 Klabunde, T., Sharma, S., Telenti, A., Jacobs, W.R. Jr. & Sacchettini, J.C. *Nature Struct. Biol.* **5**, 31–36 (1998).
11 Hobohm, U., Scharf, M., Schneider, R. & Sander, C. *Prot. Sci.* **3**, 409–417 (1992).