Volume 288, number 1,2, 192–196 FEBS 10050 August 1991

ADONIS 001457939100758O

Occurrence of bifurcated three-center hydrogen bonds in proteins

Robert Preißner[1], Ursula Egner[2] and Wolfram Saenger[1]

[1] *Institut für Kristallographie, Freie Universität Berlin, Takustr. 6, D-1000 Berlin 33, Germany* and [2] *Schering AG, Müllerstr. 178, D-1000 Berlin 65, Germany*

Received 11 April 1991; revised version received 19 June 1991

Analysis of 13 high-resolution protein X-ray crystal structures shows that 1204 (24%) of all the 4974 hydrogen bonds are of the bifurcated three-center type with the donor X-H opposing two acceptors A1, A2. They occur systematically in α-helices where 90% of the hydrogen bonds are of this type; the major component is (n+4)N-H···O=C(n) as expected for a 3.6_{13} α-helix, and the minor component is (n+4)N-H···O=C(n+1), as observed in 3_{10} helices; distortions at the C-termini of α-helices are stabilized by three-center bonds. In β-sheets 40% of the hydrogen bonds are three-centered. The frequent occurrence of three-center hydrogen bonds suggests that they should not be neglected in protein structural studies.

Protein structure; Bifurcated hydrogen bond; Three-center-hydrogen bond; Statistical analysis

1. INTRODUCTION

The importance of the two-center hydrogen bonds X-H···A for the stabilization and function of biological macromolecules is well documented [1]. Only little attention has been paid to bifurcated three-center hydrogen bonds, although about 20–25% of all the hydrogen bonds in crystal structures of small biological molecules are of this type [2–4]. In a three-center hydrogen bond, two acceptors A1, A2 oppose the donor X-H; the hydrogen atom lies in the plane defined by the atoms X, A1, A2 so that $\theta 1 + \theta 2 + \theta 3 \sim 360°$, Fig. 1. The symmetrical bond with $r1 \sim r2 \sim 2.2(1)$ Å and $\theta 1 \sim \theta 2 \sim 135(10)°$ is less frequently observed than the asymmetrical bond. In the latter, the major component is 'strong' and approaches a geometry as known for the two-center X-H···A bonds, $r1 \sim 2.03(13)$ Å, $\theta 1 \sim 151(15)°$, and the minor component is 'weaker', $r2 \sim 2.51(18)$ Å with the associated angle more acute, $\theta 2 \sim 117(19)°$.

Although three-center (bifurcated) hydrogen bonds are discussed in the literature [1,5], they were not subjected to a detailed study. In this paper, the analysis of a sample of 13 well-refined X-ray structures of proteins shows that this type of hydrogen bonding occurs systematically in α-helices and β-sheets and requires consideration in model building studies.

2. MATERIALS AND METHODS

This analysis is based on well-refined X-ray protein crystal structures with a resolution of at least 1.6 Å. In the 13 proteins given in Table I, three-center bonds were identified when angles θ1, θ2 were greater than 90° and distances r1, r2 were in the range 1.6–3.0 Å [2–4]. These cutoff-values are more relaxed compared to earlier studies [1–3] in order to ensure that the minor components of three-center bonds are also considered [4]. The angular cutoff $\theta 2 \geq 90°$ comprises about 85% of all experimentally determined hydrogen bonds to carbonyl oxygens [6], and the cutoff $r2 \leq 3.0$ Å accounts for 90% of all the hydrogen bonds with $r2 \leq 3.2$ Å in the present sample of 13 proteins. The relatively long distance $r2 \leq 3.0$ Å is justified because the hydrogen bond is primarily electrostatic and diminishes slowly with 1/r; at 3 Å, it is still 40% of a 'strong' hydrogen bond with an H···A separation of 1.8 Å [7].

3. RESULTS AND DISCUSSION

Of the total of 4974 hydrogen bonds, 1204 (24%) are of the bifurcated three-center type; in Table I, they are grouped into 12 different classes according to donor and acceptor groups belonging to main chain (MC), side chain (SC) or water (W). In 86% of the 1204 three-center hydrogen bonds, the donor groups are main chain -NH, and 55% belong to the class MC-(MC,MC) where both acceptors are main chain C=O. The others are found in much less populated classes, of which five are just above the level of statistical error, 2.9%, (MC-(MC,W), MC-(SC,W), MC-(MC,SC), SC-(SC,W), SC-(W,W)).

In the 13 proteins listed in Table I, 620 peptide N-H groups are engaged in α-helix hydrogen bonding. 571 or 92% of these are of the three-center type, with 460 or 74% of the class MC-(MC,MC). They form a typical zigzag pattern with N-H on one side and O=C on the other, Fig. 2a,b. The major component is the common (n+4)N-H···O=C(n), and the minor component is (n+4)N-H···O=C(n+1), an interaction typical for 3_{10} helices. The three-center hydrogen bond geometry in α-helices varies from very unsymmetrical to symmetrical, depending on distortions in backbone geometry. Two-center bonds within the α-helices are frequently found

Correspondence address: Institut für Kristallographie, Freie Universität Berlin, Takustr. 6, D-1000 Berlin 33, Germany. Fax: (49) (30) 8386702.

Published by Elsevier Science Publishers B.V.

Table 1: Distribution of Three-Center Hydrogen Bonds Occurring in 13 Well-Refined Proteins.

		Proteins[a]														
		CCR	CPA	CSE	EBX	GRS	MBD	OVO	PAZ	PRK	PTP	RNT	SGA	TLN		
% α		57	38	32	-	36	79	21	17	29	12	16	8	42		
% β		-	15	20	53	31	-	29	56	34	30	35	66	23		
residues		112	307	345	62	478	153	56	123	279	223	104	181	316		
solv. mol.[b]		46	315	434	111	523	388	31	93	178	168	121	200	173		
resol. [Å]		1.5	1.54	1.2	1.4	1.54	1.4	1.5	1.55	1.5	1.34	1.5	1.5	1.6		
R-factor[c]		.190	.190	.178	.176	.186	.188	.199	.180	.167	.171	.143	.126	.213		
reference		9	10	11	12	13	14	15	16	17	18	19	20	21		
		Three-Center Hydrogen Bonds[d]														
Don.	Acc.														sum	%
MC	MC MC	21	80	84	9	118	78	11	24	61	29	22	36	94	667	55
MC	MC W	2	9	20	5	22	1	2	4	11	10	6	9	7	108	9
MC	SC W	1	6	19	1	16	5	2	4	9	10	12	7	10	102	8
MC	MC SC	5	7	22	5	11	2	3	5	12	7	3	8	6	96	8
SC	SC W	1	4	9	2	6	7	1	0	2	3	5	2	6	48	4
SC	W W	3	0	10	5	8	12	0	1	2	0	0	3	1	45	4
MC	W W	0	5	9	0	9	4	0	2	3	1	3	2	1	39	3
SC	MC W	2	2	7	1	4	4	0	0	5	2	2	1	0	30	2
MC	SC SC	0	4	3	0	3	2	1	1	3	2	1	2	4	26	2
SC	MC MC	0	4	3	1	2	1	0	2	1	2	1	1	1	19	2
SC	MC SC	1	2	2	1	0	1	0	2	4	1	0	1	1	16	1
SC	SC SC	1	1	0	0	1	0	1	0	0	0	0	0	4	8	1
															1204	100

[a] CCR = cytochrome C, CPA = carboxypeptidase, CSE = subtilisin Carlsberg - eglin, EBX = erabutoxin B, GRS = glutathione reductase, MBD = myoglobin, OVO = ovomocuid, PAZ = pseudoazurin, PRK = proteinase K, PTP = Beta trypsin, RNT = ribonuclease T1, SGA = proteinase A, TLN = thermolysin

[b] Solvent molecules in the crystal asymmetric unit

[c] Crystallographic refinement criterion, $R = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$ with F_{obs} = observed and F_{calc} = calculated structure amplitudes

[d] The analysis was performed using atomic coordinates from the 13 protein X-ray structures [9–21] refined to high resolution with a total of 2739 amino acids. For the analysis of two-center and three-center hydrogen bonds, the positions of N-H hydrogen atoms were calculated with MOLEDT [22] on the bisector determined by C-N-C for peptide bonds and sidechains of HIS and TRP. For GLN and ASN, the amide $-NH_2$ hydrogen atoms were placed in the amide plane at C-N-H angles of 120°, and those of the ARG guanidinium group were positioned accordingly.

Hydrogen atoms of $-OH$ and $-NH_3^+$ groups are not included in this analysis due to rotational freedom. Crystallographically determined water oxygen atoms were positioned in a 3 Å shell around the protein molecules under consideration of space-group symmetries to yield a total of 3196 water molecules for the 13 proteins.

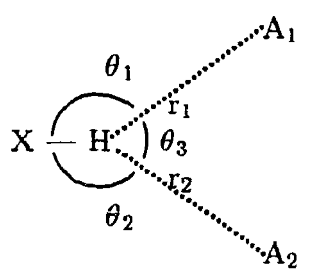

Fig. 1. Geometry of the three-center hydrogen bond.

with amino acid side chains or water molecules participating as third partners, implying that 'true' two-center bonds are rare.

The systematic occurrence of three-center hydrogen bonds in α-helices is associated with distortions observed at their C-termini, called α_{C1} and α_{C2}. In both distortions, the C=O of amino acid (n) accepts three hydrogen bonds from the N-H in positions (n+3), (n+4), and (n+5). In α_{C1} (Fig. 2a), these bonds are of the three-center type with minor components from N-H groups in (n+3) and (n+5), and the major component from (n+4); the amino acid in position (n+4) is predominantly Gly. In α_{C2} (Fig. 2b), a three-center bond is donated by N-H in (n+3), but the hydrogen bonds from (n+4) and (n+5) are of the two-center type. In the β-pleated sheet structures of the 13 proteins listed in Table I, 570 peptide N-H groups are involved in hydrogen bonds, of which 229 or 40% are of the three-center type. The major component is between different strands, and the minor component is intraresidue (intrastrand) (n)N-H···O=C(n), with angle θ2 close to 90°, Fig. 2c. The other most frequent classes of three-center bonds in β-sheets are MC-(MC,SC), MC-(MC,W) and MC-(SC,W). A strong contribution to the MC-(SC,X) classes is due to backbonding of $(n)O_{SC}\cdots H\text{-}N_{MC}(n)$, where amino acids Glu and Asp predominate [1].

121 (10%) of the three-center hydrogen bonds are found associated with β-turns. These are frequently located at the periphery of globular proteins, and consequently a high proportion (49%) of all interactions in β-turns are of the classes MC-(X,W).

The geometry of all the three-center hydrogen bonds is analyzed in Fig. 3. There are only few symmetrical hydrogen bonds where r1 ~ r2, and the unsymmetrical geometry predominates by far. The distribution for the major component r1 rises sharply at 1.8 Å, peaks at r1 = 2.10(23) Å and drops slowly at 2.4 to 2.6 Å; the distribution for the minor component r2 shows a steady increase from 2.3 Å to a maximum H···A distance of 2.8 Å, and falls off sharply at 2.9 Å with a mean of 2.68(22) Å which confirms the selected cutoff-criteria;

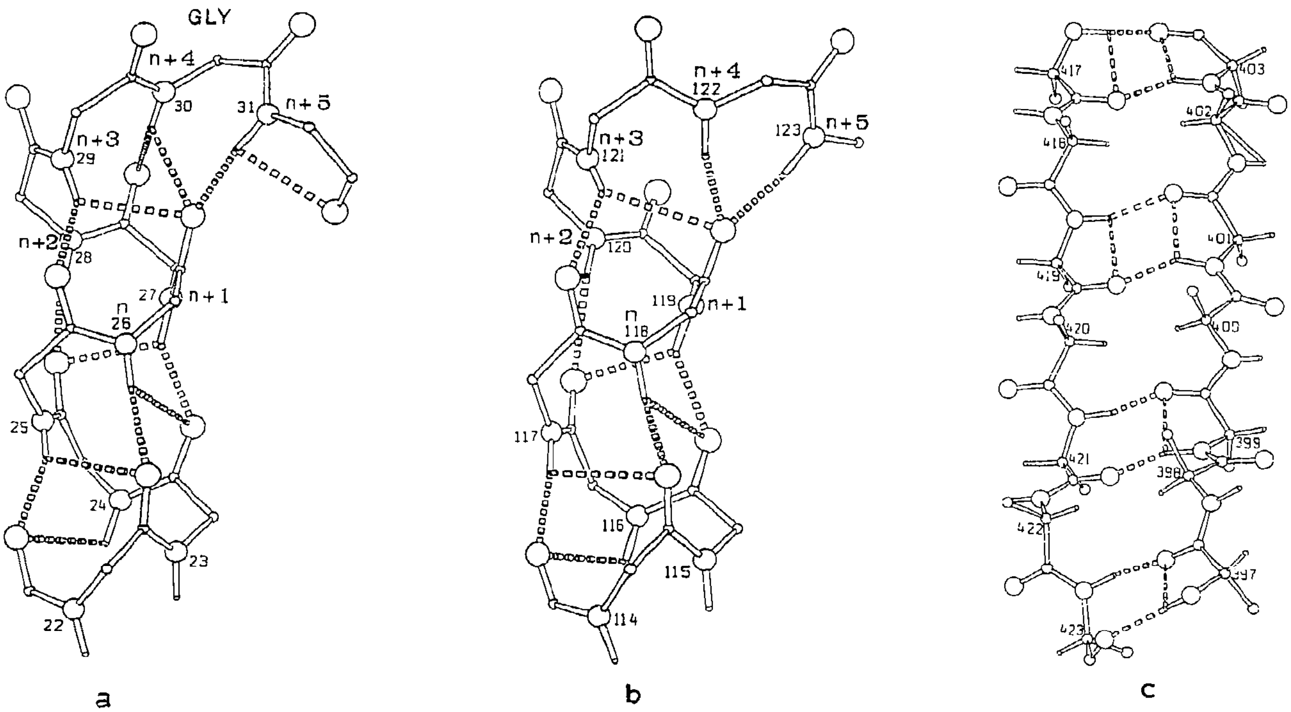

Fig. 2. Plots of α-helix terminus distortions α_{C1} (a; RNT) and α_{C2} (b; CPA), and antiparallel β-pleated sheet (c, GRS), drawn with Schakal [24]. Dashed lines indicate hydrogen bonds, the N-H groups are numbered sequentially and, in a,b, the N-H groups of the residues defining the distortions are numbered (n) to (n+5). In a,b only main chain atoms are drawn, in c Cα atoms are drawn with 4 substituents. In the proteins listed in Table I, there are 15 distortions of type α_{C1}; in the majority (13), residue (n+4) is Gly, one is His and one is Asn; torsion angles φ, ψ are 80°, 30° for (n+4); −80°, −10° for (n+3); −60°, −30° for (n+2). There are 6 distortions of type α_{C2}, with torsion angles similar as for α_{C1} but φ, ψ for (n+4) are −60°, −60°.

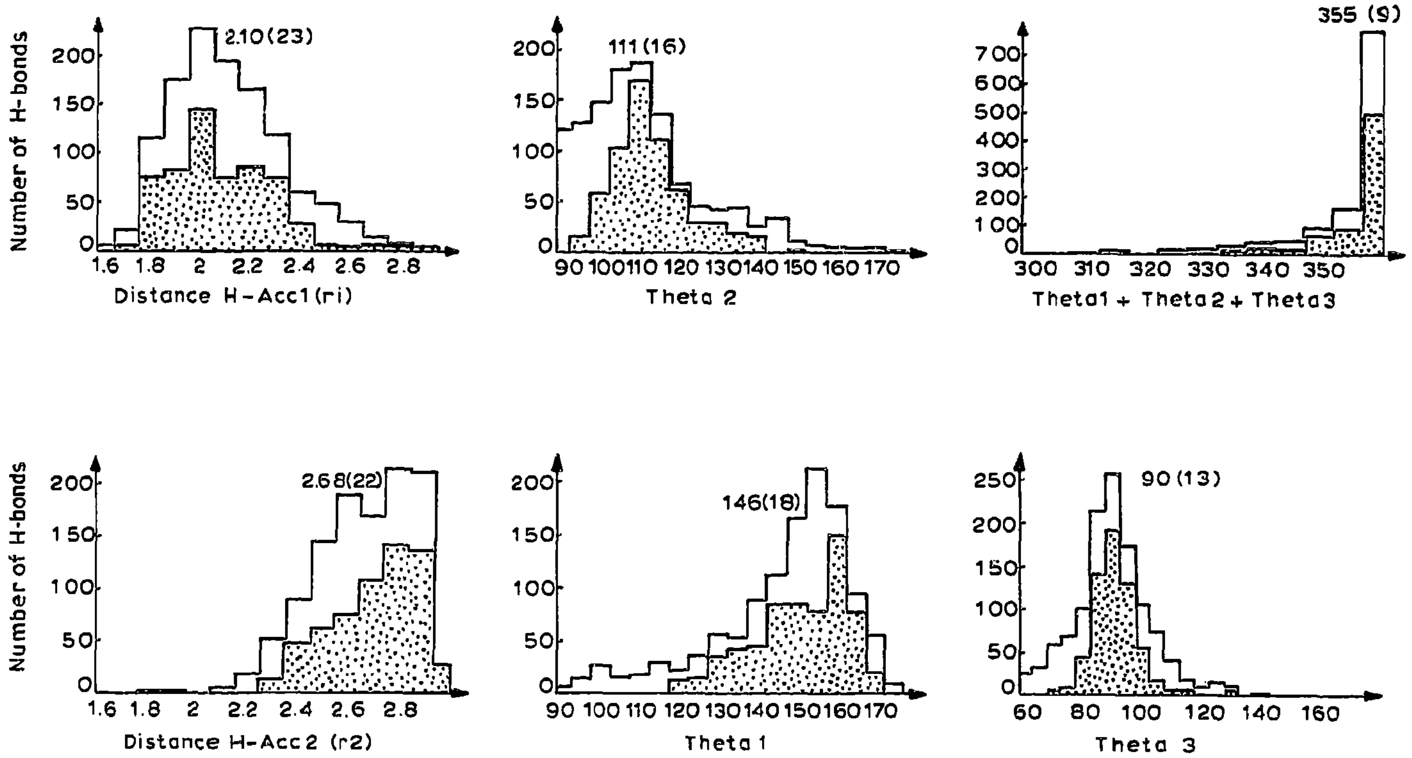

Fig. 3. Histograms describing the frequency distribution of H···A distances r1 and r2 (Å); of X – H···A angles, $\theta 1$ and $\theta 2$; of angle A1···H···A2, $\theta 3$; of the sum of angles $\theta 1 + \theta 2 + \theta 3$. Thick lines indicate the total number of three-center hydrogen bonds, shaded areas give the distribution in α-helices only.

the secondary peak at 2.6 Å is due to the contributions from the β-strands. The distributions in the angles $\theta 1$, $\theta 2$ differ characteristically. For $\theta 1$, the number of data points increases slowly from 100° to a maximum at 155° and then drops steeply at 170°, with a mean of 146(18)°. $\theta 2$ shows a distribution peaking at 110°, mean 111(16)°, and in β-sheets $\theta 2$ tends to be closer to 90° than in α-helices. For $\theta 3$ the distribution is narrow with mean 90(13)°. This is expected because $\theta 3$ is determined by the separation of acceptors A1, A2 which is confined by covalent bond geometry (e.g. in α-helices) and/or by van der Waals separation between these atoms. The sum of angles $\theta 1 + \theta 2 + \theta 3$ peaks sharply at 360°, mean 355(9)°, with 87% greater than 350°, as required when the hydrogen atom is located within the plane defined by atoms X, A1, A2.

Since 24% of all the hydrogen bonds in proteins are of the three-center type, they cannot be neglected in structural studies and should be included in theoretical approaches used in the prediction of protein three-dimensional structures. They will be of importance in the study of structural fluctuations and hence in dynamical simulations because one donor opposes two acceptors simultaneously and is already in an intermediate state between two two-center bonds, X-H···A1 and X-H···A2. Experimental evidence for transient formation of three-center hydrogen bonds is provided by the neutron diffraction study of vitamin B_{12} hydrate [8], and by local distortions of α-helices in several protein crystal structures which served as snapshots to illustrate the occurrence of three-center bonds in the disruption or formation of α-helices [9].

Three-center hydrogen bonds are not treated adequately in the current computer programs that are used to investigate hydrogen bonding interactions. In the DSSP program [5] which was designed to assign secondary structure elements, the energy cutoff at -0.5 kcal/mol for hydrogen bonding would not consider many of the minor components. In glutathione reductase (GRS) for instance, (118)NH donates a three-center bond with major component to O=C(114) (r1 = 2.22 Å, $\theta 1 = 146°$), and with minor component to O=C(115) (r2 = 2.44 Å, $\theta 2 = 119°$); in the DSSP program, the energies of the major and minor components are calculated as -1.7 and -0.2 kcal/mol, respectively, and consequently this hydrogen bonding interaction is considered to be of the two-center type. In addition, DSSP does not limit hydrogen bonds to the 'forward' direction ($\theta \geq 90°$) so that 'backward' geometries are treated as hydrogen bonds although they are impossible stereochemically.

Acknowledgements: This work was supported by the Deutsche Forschungsgemeinschaft (Schwerpunkt 'Protein Design') and by the Fonds der Chemischen Industrie.

REFERENCES

[1] Baker, E.N. and Hubbard, R.E. (1984) Prog. Biophys. Mol. Biol. 44, 97-179.

[2] Taylor, R., Kennard, O. and Versichel, W. (1984) J. Am. Chem. Soc. 106, 244-248.

[3] Ceccarelli, C., Jeffrey, G.A. and Taylor, R. (1981) J. Mol. Struct. 70, 255-271.

[4] Jeffrey, G.A. and Saenger, W. (1991) Hydrogen Bonding in Biological Structures, Springer, Heidelberg, in press.

[5] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577-2637.

[6] Boobbyer, D.N.A., Goodford, P.J., McWhinnie, P.M. and Wade, R.C. (1989) J. Med. Chem. 32, 1083-1094.

[7] Singh, V.C. and Kollman, P.A. (1985) J. Chem. Phys. 83, 4033-4040.

[8] Savage, H. (1986) J. Biophys. Soc. 50, 947-980.

[9] Sundaralingam, M.C. and Sekharudo, Y.C. (1989) Science 244, 1333-1337.

[10] Ochi, H., Hata, Y., Tanaka, N., Kakudo, M., Sakurei, T., Aihara, S. and Morita, Y. (1983) J. Mol. Biol. 166, 407-418.

[11] Rees, D.C., Lewis, M. and Lipscomb, W.N. (1983) J. Mol. Biol. 168, 367-387.

[12] Bode, W., Papamokos, E. and Musil, D. (1987) Eur. J. Biochem. 166, 673-692.

[13] Smith, J.L., Corfield, P.W.R., Hendrickson, W.A. and Low, B.W. (1988) Acta Crystallogr., Sect. A 44, 357-368.

[14] Karplus, P.A. and Schulz, G.E. (1987) J. Mol. Biol. 195, 701-729.

[15] Phillips, S.E.V. and Schoenborn, B.P. (1981) Nature 292, 81-82.

[16] Bode, W., Epp, O., Huber, R., Laskowski, M. Jr. and Ardelt, W. (1985) Eur. J. Biochem. 147, 387-395.

[17] Petratos, K., Banner, D.W., Beppu, T., Wilson, K.S. and Tsernoglu, D. (1987) Febs Lett. 218, 209-214.

[18] Betzel, C., Pal, G.P. and Saenger, W. (1988) Acta Crystallogr., Sect. B 44, 163-172.

[19] Chambers, J.L. and Stroud, R.M. (1979) Acta Crystallogr., Sect. B 35, 1861-1874.

[20] Martinez-Oyandel, J. (1991) PhD Thesis, Freie Universität Berlin and Universidad de Concepcion (Chile).

[21] Moult, J., Sussman, F. and James, M.N.G. (1985) J. Mol. Biol. 182, 555-566.

[22] Holmes, M.A. and Matthews, B.W. (1982) J. Mol. Biol. 160, 623-639.

[23] DISCOVER, User Manual Version 2.6 (1990) BIOSYM Technologies, San Diego, CA, USA.

[24] Keller, E. (1988) Schakal88, Kristallographisches Institut der Albert-Ludwigs-Universität Freiburg, Germany.