

Interior turns in globular proteins

George D. Rose & William B. Young

Department of Biological Chemistry, Hershey Medical Center,
Pennsylvania State University, Hershey, Pennsylvania 17033, USA

Lila M. Gierasch

Department of Chemistry, University of Delaware, Newark,
Delaware 19711, USA

Reverse turns are specific sites in proteins at which the polypeptide chain changes its overall direction¹⁻⁶. This category of secondary structure enables the chain to turn a corner, and its frequent occurrence^{7,8} is the geometric basis for the ultimate globular shape of the protein. β -Turns in particular are comprised of four consecutive residues with a stereochemistry⁹ that constrains the turn to be polar. In consequence, turns are almost always situated at the surface of the protein, in contact with solvent water. We have searched proteins of known structure and find that, on occasion, a turn may be buried within the hydrophobic interior of the molecule. In every instance of a buried turn, one or more solvent molecules were also found in a hydrogen-bonded complex with main-chain atoms of the turn residues. These bound water molecules appear to function as an integral part of the protein structure.

β -Turns, as defined by Venkatachalam⁹, consist of four consecutive residues with a hydrogen-bond joining the C=O of the first turn residue (i) to the N-H of the last residue ($i+3$); this steric arrangement requires that any hydrogen bond involving the remaining two residues ($i+2$, $i+3$) be formed with partners outside the turn. An isolated β -turn is already highly polar due to these unpaired N-H and C=O groups. In the context of a protein, polarity is further enhanced because turn formation is promoted at chain sites where the average linear hydrophobicity is locally minimal^{10,11}. Thus, it is hardly surprising that turns are almost always situated at the molecular surface^{1,4,10-12}, in contact with solvent water.

The Brookhaven data base¹³ was screened for interior turns in refined structures that include solvent coordinates. Twenty-two proteins meeting these conditions were examined, and six with interior turns were found, including three homologous serine proteases (see Table 1). Each turn satisfied the following criteria: (1) The dihedral angles for the middle residues ($i+1$, $i+2$) are near their ideal values. Following Venkatachalam⁹, ideal values are taken to be $(-60, -30)$ and $(-90, 0)$ for type I and $(-60, 120)$ and $(80, 0)$ for type II. (2) The distance from the carbonyl oxygen of the first turn residue (i) to the amide nitrogen of the last turn residue ($i+3$) is small enough to permit hydrogen-bonding (<3.4 Å). An interior turn in phospholipase was allowed as an exception because in this case the distance from C=O ($i-1$) to N-H ($i+3$) is 3.15 Å, forming a five-residue, hydrogen-bonded turn. (3) The residues were required to be at a locally minimal radius of curvature⁶, indicating the

presence of a turn site. These are comparatively stringent criteria, as evidenced by the small $C^\alpha(i) - C^\alpha(i+3)$ distances which are well below the usual 7 Å cutoff^{2,8}.

A detailed atomic description of the protein surface was generated according to the method of Lee and Richards¹⁴. In this approach, the molecular surface is defined by those atoms that are accessible to a water-sized probe; remaining atoms comprise the interior. We have used an implementation of this method due to Connolly¹⁵ that is tailored for use with computer graphics.

For each turn, the fraction buried was determined by subtracting the solvent-accessible surface area from total surface area and then dividing by total surface area. Calculated values are listed in Table 1, columns 7-9. By way of illustration, the stippled surface in Fig. 2a depicts the solvent-accessible fraction of the lysozyme turn, residues 54-57, while Fig. 2b shows the total surface that results when these residues are viewed in isolation. Comparison of the two situations indicates that the turn is indeed buried, an interpretation confirmed by Table 1; 84% of the turn's surface is rendered inaccessible by the surrounding presence of the protein.

The proteins used in this study and their Brookhaven file names include phospholipase A₂ (1BP2)¹⁶, γ -chymotrypsin A (2GCH)¹⁷, lysozyme (6LYZ)¹⁸, β -trypsin (3PTP)¹⁹, staphylococcal nuclease (1SNS)²⁰ and trypsinogen (1TGB)²¹. Visualization was achieved with an Evans and Sutherland graphics system using the GRAMPS interpreter written by T.J. O'Donnell. Interatomic distances and angles were calculated directly from X-ray coordinates.

The turns under study are all type I and type II β -turns, and Table 1 lists pertinent information about each.

A striking finding is the presence of one or more tightly complexed water molecules in the vicinity of each turn. The water molecule associated with the lysozyme turn is illustrated in Fig. 2, while in Fig. 3 it is displayed with its hydrogen-bonded neighbours from both the turn and the rest of the protein.

Each of the water molecules bound to an interior turn has at least three hydrogen bonds. Moreover, each polar group in these turns is hydrogen bonded either to a water molecule or to another interior polar group of the protein. Instances of bifurcated hydrogen bonds to these groups were also observed. Table 2 lists water ligands together with all hydrogen-bonded partners closer than 3.4 Å.

To assess whether these complexed water molecules are themselves buried within the protein, their number density was calculated using all protein atoms within 8 Å of each solvent molecule. In Table 2, the pair of numbers in parentheses gives the range of number densities computed for all water molecules in each protein. The table also gives the number density of individual water molecules that are hydrogen bonded to an internal turn. It is apparent that these latter values are at the high end of the range and resemble corresponding values found for the most deeply buried α -carbons in globular proteins¹¹. Thus, water molecules in complex with interior turns are themselves deeply buried within the protein and not part of the bulk solvent.

Fig. 1 Stereoview of a β -turn in lysozyme, residues 54-57, shown in the context of a containing hexapeptide segment, residues 53-58. Broken lines depict hydrogen bonds, which connect C=O(54) to N-H(57) in the turn, as well as C=O(53) to N-H(58) and N-H(53) to C=O(58) in the hexapeptide.

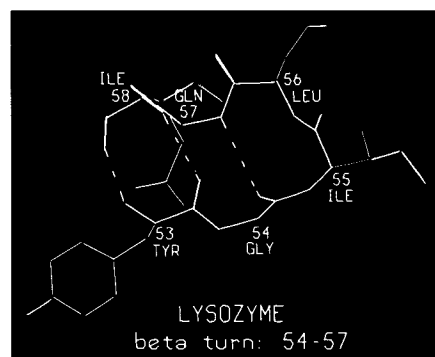
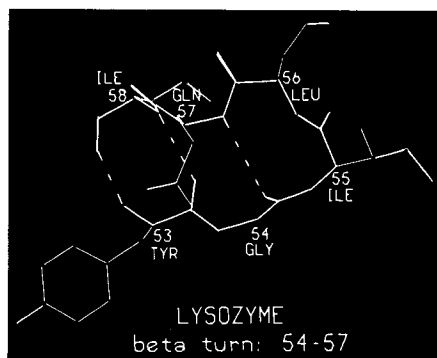
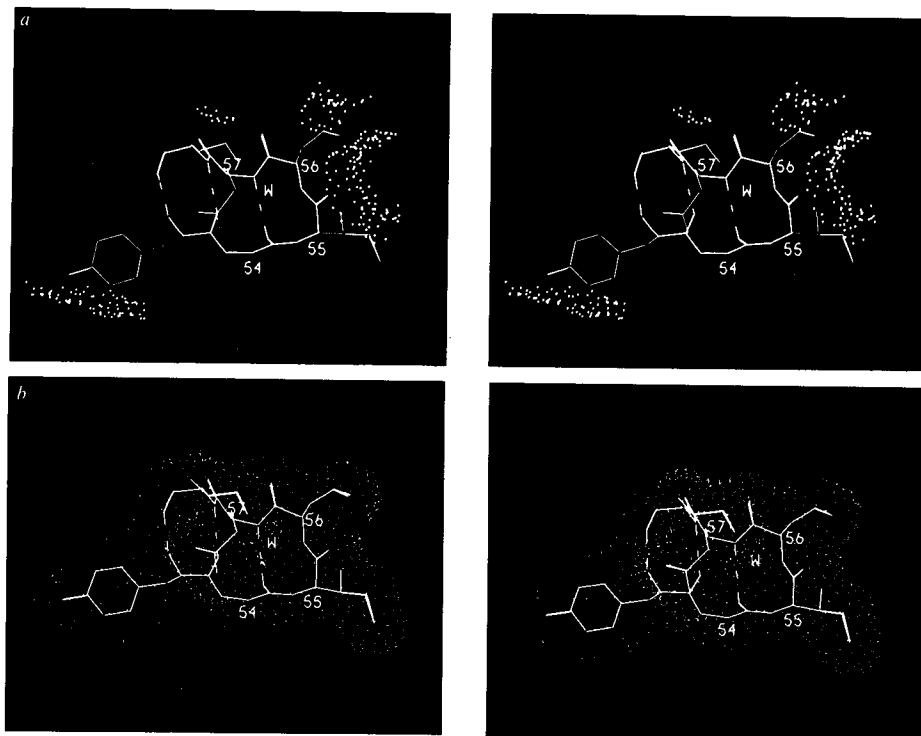


Table 1 Interior turns

Residues in turn	Sequence	Dihedral angles	C=O...N distance	C ^α (i)-C ^α (i+3) distance	Atoms in turn	Accessible surface area	Total surface area	% Buried
1. γ -Chymotrypsin A (2GCH) 27-30	Trp-Pro-Trp-Gln	$i+1$: -59, -25 $i+2$: -81, -12 type I	2.94	5.61	44	19.39	447.33	95.7
2. γ -Chymotrypsin A (2GCH) 194-197	Asp-Ser-Gly-Gly	$i+1$: -53, 144 $i+2$: 83, -20 type II	3.12	5.76	22	12.34	268.92	95.4
3. β -Trypsin, diisopropylphosphoryl inhibited (3PTP) 27-30	Val-Pro-Tyr-Gln	$i+1$: -68, -15 $i+2$: -91, -3 type I	3.14	5.80	35	74.19	368.08	79.8
4. β -Trypsin, diisopropylphosphoryl inhibited (3PTP) 194-197	Asp-Ser-Gly-Gly	$i+1$: -46, 142 $i+2$: 90, -13 type II	3.11	5.85	22	8.72	266.97	96.7
5. Trypsinogen (1TGB) 27-30	Val-Pro-Tyr-Gln	$i+1$: -53, -21 $i+1$: -99, 0 type I	3.22	5.74	35	68.04	350.61	80.6
6. Trypsinogen (1TGB) 194-197	Asp-Ser-Gly-Gly	$i+1$: -47, 145 $i+2$: 97, -15 type II	3.32	5.95	22	53.78	269.55	80.0
7. Bovine pancreas phospholipase A ₂ (1BP2) 26-29	Gly-Cys-Tyr-Cys	$i+1$: -90, -22 $i+2$: -111, -30 type I	4.93	6.07	28	29.96	369.94	91.9
8. Hen egg lysozyme (6LYZ) 54-57	Gly-Ile-Leu-Gln	$i+1$: -62, -31 $i+2$: -108, 7 type I	3.33	4.73	29	53.90	334.90	83.9
9. Staphylococcal nuclease (1SNS) 19-22	Asp-Gly-Asp-Thr	$i+1$: -4, -80 $i+2$: -86, 25 type II'	2.54	4.95	27	74.52	268.64	72.3

Fig. 2 a, Stereoview of the lysozyme β -turn, residues 54-57, showing the solvent-accessible surface. Solvent access to the segment is calculated with a program by Connolly¹⁵ and is depicted by the stippled surface. The turn is seen to be buried within the interior of the protein: only 5 of the 29 atoms in turn proper (residues 54-57) can establish even limited contact with solvent, that is, C ^{γ 1} and C ^{γ 2} in Ile 55 and C ^{γ} , C ^{δ 1} and C ^{δ 2} in Leu 56. The position of a tightly complexed water molecule, visible in the X-ray map, is indicated by a W. **b**, Stereoview of the lysozyme turn, residues 54-57, treated as an isolated segment, with solvent-accessible regions depicted as stippled surface. In the absence of the surrounding protein, the entire turn can be solvated.



The loci of buried waters are known to be conserved in related proteins^{22,23}. The present study includes two internal turns within each of three serine proteases. Both turns are found in regions of conserved sequence in the serine proteases²⁴. The disposition of the water molecules in complex with these turns exhibits a similarly conserved pattern, as shown in Tables 1 and 2, with five buried water molecules hydrogen-bonded to most of the same donors and acceptors in each protein.

Groups from many of these buried turns assume functional roles in their respective proteins. Residues 194–197 in the serine proteases contain active site Ser 195 (ref. 25). In phospholipase, the entire turn is part of the calcium binding loop¹⁶, and in the staphylococcal nuclease turn, calcium is bound to O^{δ1} of Asp 21 (ref. 20). Finally, the lysozyme turn and its bound water impinge on that enzyme's active site²⁶.

A turn situated in a hydrophobic environment such as the interior of a protein must find a way of neutralizing its polar groups without recourse to bulk solvent. In addition, hydrogen-bonding partners for the peptide groups of the middle residues must be found external to the turn itself, for reasons mentioned previously. These constraints have been satisfied for the turns in the present study by incorporating water molecules into the structure as though they were prosthetic groups. With this addition, buried chain segments are apparently able to adopt a favourable turn geometry in what would otherwise be an unfavourable environment.

The present findings represent a specific case of structural hydration, as defined by Kuntz and Kauzmann²⁷. Edsall and McKenzie²³ have also discussed these issues, and proposed four categories of water molecules in globular proteins. One of their

Table 2 Water hydrogen bonds

Residue or water no.				Group	Distance	No. density (range)	Residue or water no.				Group	Distance	No. density (range)		
1. γ -Chymotrypsin A (27–30)						(11–124)	5. Trypsinogen (27–30)						(16–122)		
66W	27	N–H	3.11		717W		27	N–H	2.80					103	
	27	C=O	2.79				27	C=O	2.96						
	Water	121	2.77				25	N–H	3.29						
17W	28	C=O	2.78	111			26	N–H	3.29						
	30	C=O	2.86				Water	604	2.91						
	69	N–H	2.83				Water	709	3.19						
22W	30	C=O	3.10	100	716W		28	C=O	2.86					107	
	67	C=O	2.85				30	C=O	2.94						
	70	N–H	3.06				69	N–H	2.86						
2. γ -Chymotrypsin A (194–197)						(11–124)	708W						110		
9W	194	C=O	2.97	117			66	C=O	3.26						
	139	C=O	2.76				69	N–H	3.19						
	Water	14	2.88				70	N–H	2.79						
26W	196	C=O	2.76	108			Water	709	3.05						
	53	C=O	3.13		6. Trypsinogen (194–197)						(16–122)				
	45	O γ	2.98		401W	194	C=O	3.12					113		
3. β -Trypsin (27–30)						(7–121)	406W						107		
28W	27	N–H	2.79	103			139	C=O	3.23						
	27	C=O	2.96				196	C=O	2.82						
	26	N–H	3.14				45	N–H	3.12						
	25	N–H	3.30				45	O γ	2.51						
	Water	51	2.91				53	C=O	2.97						
	Water	52	3.15		7. Phospholipase A ₂ (26–29)						(10–117)				
42W	28	C=O	2.89	113	22W		26	C=O	3.07						96
	30	C=O	2.73				29	C=O	3.00						
	69	N–H	2.79				120	N ζ	2.64						
41W	30	C=O	2.88	107		Water	73	2.87							
	65	C=O	2.65		26W	26	N–H	3.30					95		
	69	N–H	3.32			26	C=O	2.87							
	70	N–H	2.96			24	C=O	3.19							
52W	27	C=O	3.35	113		117	N–H	3.34							
	30	C=O	3.19			118	N–H	2.72							
	70	C=O	2.96		5W						95				
	Water	28	3.15			28	C=O	3.13							
	Water	46	2.82			30	N–H	2.99							
4. β -Trypsin (194–197)						(7–121)	8. Lysozyme (54–57)						(10–106)		
37W	194	C=O	2.87	117	1W		56	N–H	3.09						106
	139	C=O	2.85				53	C=O	2.94						
	30	O ϵ 1	3.33				91	O γ	3.25						
	Water	16	2.86		9. Staphylococcal nuclease (19–22)						(46–122)				
69W	196	C=O	2.83	108	11W		20	N–H	2.57						108
	197	C=O	3.38				19	O γ 1	3.04						
	45	O γ	2.50				41	O γ 1	2.53						
	53	C=O	3.37				42	C=O	2.56						

The table lists all water ligands for each buried turn together with other possible hydrogen-bonding partners closer than 3.4 Å. Water molecules are of the form *n*W, where *n* is their Brookhaven index number¹³. The number density, shown in the far right column, is the calculated number of protein atoms contained within an 8 Å sphere centred at the given water molecule. A high number density indicates that the water is well buried within the protein¹¹. A single table entry for each protein, shown in parentheses, gives the range of number densities observed for all the water molecules in that protein. Waters associated with buried turns are seen to be at the high end of this range in each case.

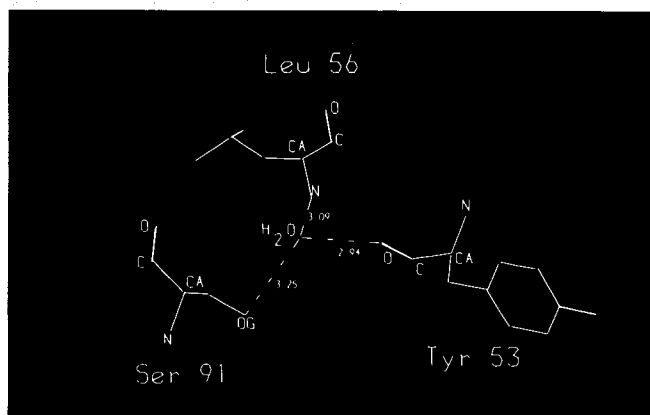


Fig. 3 A specific water molecule, visible in the X-ray map as HOH¹ (ref. 18), is illustrated together with its protein ligands. The water is hydrogen bonded to the C=O in Tyr 53, to the N—H in Leu 56, and to the O γ in Ser 91. These three residues are shown projected onto the O(53)—N(56)—O γ (91) plane; the water oxygen is 0.78 Å below the plane.

categories is the internal water found within folded peptide chains²³.

The finding that interior turns are accompanied by bound water molecules may also have application to membrane proteins. The membrane resembles the protein's own interior to the extent that both are environments of low polarity as a consequence of close-packed alkyl groups and sparing water. Further, the fact that proteins do execute turns within the membrane can be inferred from cases where a protein penetrates the lipid bilayer and re-emerges on the same side without completely traversing the membrane^{28,29}.

Because peptide chain turns are constrained to be polar by their geometry, a turn within an apolar lipid is expected to be at least as thermodynamically unfavourable as the buried turns in the present study. It is possible that individual water molecules can contribute to the stability of a protein within a

membrane by providing analogous site-specific polarity. Such water molecules would then constitute a reversible mechanism to modulate a membrane protein's transition from the polar cytoplasm to the apolar lipid.

We thank Richard Lee, Tony Kossiakoff, Lyndon Hibbard and Robert Wood for useful advice. This work was supported by NIH grants GM 29458 (G.D.R.) and GM27616 (L.M.G.), a Research Career Development Award to G.D.R., a Research Service Award for short-term research training of medical students to W.B.Y., and grants from Merck, Sharpe & Dohme Laboratories and from Smith, Kline & French to L.M.G. This paper is dedicated to Walter Kauzmann on the occasion of his retirement.

Received 10 January; accepted 10 June 1983.

1. Kuntz, I. D. *J. Am. chem. Soc.* **94**, 8568–8572 (1972).
2. Lewis, P. N., Momany, F. A. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2293–2297 (1971).
3. Crawford, J. L., Lipscomb, W. N. & Schellman, C. G. *Proc. natn. Acad. Sci. U.S.A.* **70**, 538–542 (1973).
4. Levitt, M. & Chothia, C. *Nature* **261**, 552–558 (1976).
5. Richardson, J. S. *Adv. Protein Chem.* **34**, 167–339 (1981).
6. Rose, G. D. & Seltzer, J. P. *J. molec. Biol.* **113**, 153–164 (1977).
7. Rose, G. D. & Wetlaufer, D. B. *Nature* **268**, 769–770 (1977).
8. Chou, P. Y. & Fasman, G. D. *J. molec. Biol.* **115**, 135–175 (1977).
9. Venkatchalam, C. M. *Biopolymers* **6**, 1425–1436 (1968).
10. Rose, G. D. *Nature* **272**, 586–590 (1978).
11. Rose, G. D. & Roy, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4643–4647 (1980).
12. Smith, J. A. & Pease, L. G. *CRC crit. Rev. Biochem.* **8**, 315–399 (1980).
13. Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535–542 (1977).
14. Lee, B. K. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
15. Connolly, M. *Quantum Chemistry Program Exchange Bull.*, Program 429, **75** (Indiana University Chemistry Department, Bloomington, 1981).
16. Dijkstra, B. W., Kalk, K. H., Hol, W. G. J. & Drenth, J. *J. molec. Biol.* **147**, 97–123 (1981).
17. Cohen, G. H., Silverton, E. W. & Davies, D. R. *J. molec. Biol.* **148**, 449–479 (1981).
18. Diamond, R. *J. molec. Biol.* **82**, 371–391 (1974).
19. Chambers, J. L. & Stroud, R. M. *Acta crystallogr.* **B35**, 1861–1874 (1979).
20. Colton, F. A., Hazen, E. E. Jr & Legg, M. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2551–2555 (1979).
21. Fehllhammer, H., Bode, W. & Huber, R. *J. molec. Biol.* **111**, 415–438 (1977).
22. Edsall, J. T. *Am. chem. Soc. Symp. Ser.* **127** (1980).
23. Edsall, J. T. & McKenzie, H. A. *Adv. Biophys.* **16**, 53–183 (1983).
24. Greer, J. *J. molec. Biol.* **153**, 1027–1042 (1981).
25. Stroud, R. M., Kossiakoff, A. A. & Chambers, J. L. *Rev. Biophys. Bioengng* **6**, 177–193 (1977).
26. Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. *The Enzymes* **7**, 665 (1972).
27. Kuntz, I. D. & Kauzmann, W. *Adv. Protein Chem.* **28**, 239–345 (1974).
28. Dailey, H. A. & Strittmatter, P. *J. biol. Chem.* **256**, 3951–3955 (1981).
29. Brunner, J. *et al. J. biol. Chem.* **254**, 1821–1828 (1979).