Relationship between intramolecular hydrogen bonding and solvent accessibility of side-chain donors and acceptors in proteins

Alexander V. Efimov*, Evgenii V. Brazhnikov

Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

Accepted 9 October 2003

First published online 21 October 2003

Edited by Robert B. Russell

Abstract **This study shows that intramolecular hydrogen bonding in proteins depends on the accessibility of donors and acceptors to water molecules. The frequency of occurrence of H-bonded side chains in proteins is inversely proportional to the solvent accessibility of their donors and acceptors. Estimates of the notional free energy of hydrogen bonding suggest that intramolecular hydrogen-bonding interactions of buried and half-buried donors and acceptors can contribute favorably to the stability of a protein, whereas those of solvent-exposed polar atoms become less favorable or unfavorable.**

Key words: Hydrogen bond; Intramolecular interaction; Protein database analysis; Stability

1. Introduction

Hydrogen-bonding and hydrophobic interactions are the most important driving forces in protein folding, however, the extent to which hydrogen bonds stabilize the folded states of proteins is a contentious issue. On the one hand, hydrogen bonding appears to contribute nothing to protein stability, since protein donors and acceptors swap partners between solvent and protein upon folding, maintaining a similar number of hydrogen bonds in the folded and unfolded states. On the other hand, recent experimental studies and theoretical estimates show that the free energy of interaction varies from −0.4 to −2.0 kcal/mol and higher favoring intramolecular hydrogen bonding [1–6]. Mutational studies of salt bridges (in many salt bridges the partners form concomitant hydrogen bonds) have also given contradictory results. Solvent-exposed salt bridges contribute only marginally to protein stability [7,8], half-buried salt bridges are estimated to be favorable [9], whereas buried salt bridges can be stabilizing [10] or destabilizing [11].

The features of hydrogen bonding in proteins observed in this paper can help to explain some of these contradictory results. We have found that the fraction of H-bonded side chains in proteins is inversely proportional to the solvent accessibility of their donors and/or acceptors. This enables us to conclude that intramolecular hydrogen bonding is favorable for buried and half-buried donors and acceptors and becomes less favorable or even unfavorable for solvent-exposed polar atoms.

2. Materials and methods

For this study, a data set of 88 non-homologous protein crystal structures determined at 2.0 Å resolution or better was used (for multimeric proteins, only one chain was taken into account): 1aew, 1aii, 1aq6(A), 1aw7, 1ax1, 1b4f(A), 1bea, 1bis(A), 1bsl(A), 1elp(A), 1enc, 1fkh, 1gpo(L), 1hle(A), 1ihb(A), 1kvw, 1l99, 1lou, 1lts(D), 1mjw(A), 1opc, 1ova(A), 1swo(A), 1tfg, 1ydc, 2trh(A), 2psp(A), 1be0, 1hrd(A), 1ifb, 1f3z, 1a64(A), 1a6q, 1isu(A), 2ctb, 5cha(A), 9pap, 1msi, 1qau, 1bxa, 1dcs, 1fus, 1plc, 1bsm(A), 1rhs, 1ppt, 1thm, 1nxb, 1rdg, 1st3, 1xyz(A), 2phy, 1bgf, 2end, 1g3p, 1xnb, 1a2p(A), 1aap, 1agi, 1ah7, 1brt, 1crn, 1edm(B), 1ezm, 1hka, 1koe, 1mla, 1noa, 1opd, 1pmy, 1une, 1whi, 2end, 2ovo, 2pkc, 1dpt(A), 1orc, 1hta, 1lit, 1a3h, 1akz, 1btk(A), 1cvl, 1scs, 1ris, 1bxy(A), 1bxe, 1rss.

Accessible surface areas for individual donors and acceptors in protein crystal structures were determined using the WHAT_IF server at http://www.cmbi.kun.nl/gv/servers/WIWWWI/. Hydrogen bonds were identified using the WHAT_IF server's 'Optimal Hydrogen Bonding Network'. Criteria for H-bonding and geometrical parameters of H-bonds used in this program are described in detail in [12]. It should be noted that WHAT_IF deals with H-bond lengths less than 3.5 Å.

The process of formation of an intramolecular hydrogen bond can be represented as

$$D{-}H{\cdots}W_1 + W_2{\cdots}A{-}X \rightleftarrows D{-}H{\cdots}A{-}X + W_1{\cdots}W_2 \quad (1)$$

where D–H and A–X are protein donor and acceptor groups, respectively, and $W_{1,2}$ are two water molecules. If strong competition of water in hydrogen bonding takes place (e.g. in aqueous surroundings), the reaction can proceed to the left. If the hydrogen bond between D–H and A–X groups is not or partly accessible to water molecules, the reaction can proceed to the right. So the accessible surface area of donors and acceptors can be used as the 'effective concentration' of water in hydrogen bonding.

In the present study, the effect of solvent accessibility of donors and acceptors on hydrogen bonding between them is considered. The analysis focuses on the intramolecular hydrogen bonds formed by donors and acceptors of polar side chains. These are predominantly two types of H-bonds, side chain to side chain and side chain to main chain, apart from the intermolecular hydrogen bonds of any type (e.g. interchain hydrogen bonds, hydrogen bonds with water and cofactors) because these are competitive interactions. Donors and acceptors of side chains are considered to be H-bonded if each of them forms at least one intramolecular hydrogen bond. Those which do not form intramolecular H-bonds are considered 'non-H-bonded', although they can participate in hydrogen bonding with water or cofactors. It should be noted that most proteins in the data set do not contain cofactors and some of them have small cofactors that form only a few hydrogen bonds with side chains.

The formation of hydrogen bonds between CO and NH groups of the main chain is a more complex process because α-helices and β-sheets are cooperative systems. In extended hydrogen-bonded structures, a difference is often found between the ease of formation of the first hydrogen-bonded structural unit and the subsequent ones [13].

*Corresponding author. Fax: (7)-095-924 04 93.
E-mail address: efimov@protres.ru (A.V. Efimov).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)01189-X

The influence of water on the hydrogen bonding of main-chain polar atoms may differ from that of side chains, so the analysis of the behavior of main-chain donors and acceptors will be described elsewhere.

3. Results and discussion

We include here data for five side-chain donors (these are atoms of Gln(NE2), Asn(ND2), Lys(NZ), Arg(NE) and Arg(NH1,2); see Fig. 1) and four side-chain acceptors (these are atoms of Glu(OE1,2), Asp(OD1,2), Gln(OE1) and Asn(OD1); see Fig. 2). Side-chain oxygen atoms of Thr(OG1), Ser(OG) and Tyr(OH) can be both donors and acceptors, so they have been analyzed separately (Fig. 3). A total of 4612 H-bonded and 5682 non-H-bonded donors and acceptors of side chains have been identified and used for the analysis.

The histograms presented in Figs. 1–3 show the number of H-bonded and non-H-bonded donors and acceptors in the database as a function of their accessibility to water molecules. It can be seen that in all cases the profiles of the frequency of occurrence of the H-bonded donors and acceptors decrease from left to right, whereas those of the non-H-bonded donors and acceptors increase. The curves below the histograms show that the fractions of the H-bonded donors and acceptors also decrease from left to right in all cases. One may conclude that the lower the solvent accessibility of side-chain donors and acceptors, the higher the probability of formation of hydrogen bonds, and vice versa.

The observed data suggest that there is a strong competition of water in hydrogen bonding, and the formation of intramolecular hydrogen bonds involving donors and/or acceptors with solvent accessibility higher than 40–50% becomes less favorable or unfavorable. On the other hand, the data suggest that the burial of hydrogen bonds is favorable. Apparently, this is because of both the gain in enthalpy arising from the lower dielectric 'constant' of the protein interior and

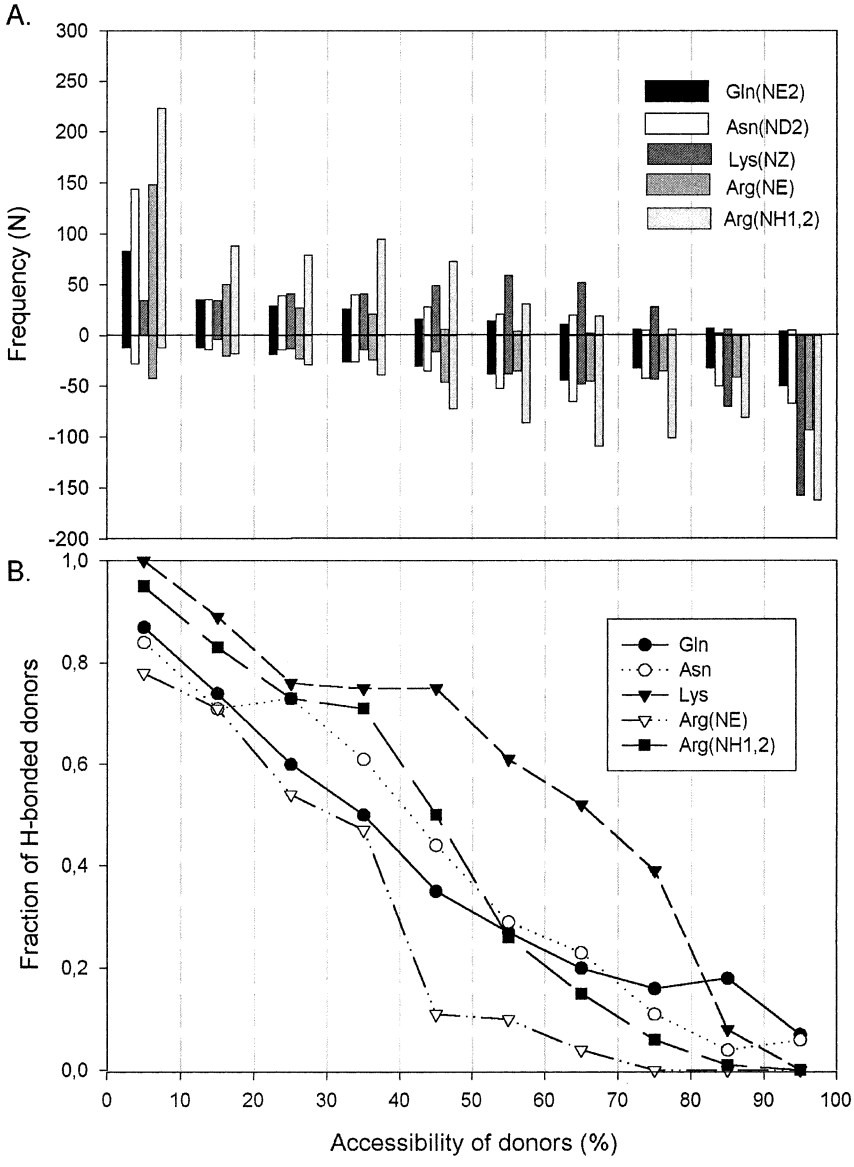

Fig. 1. A: Frequency of occurrence of H-bonded (positive values) and non-H-bonded (negative values) donors (atoms of Gln(NE2), Asn(ND2), Lys(NZ), Arg(NE) and Arg(NH1,2)) in the database versus their accessibility to water molecules. B: Fraction of the H-bonded donors in the database as a function of their solvent accessibility.

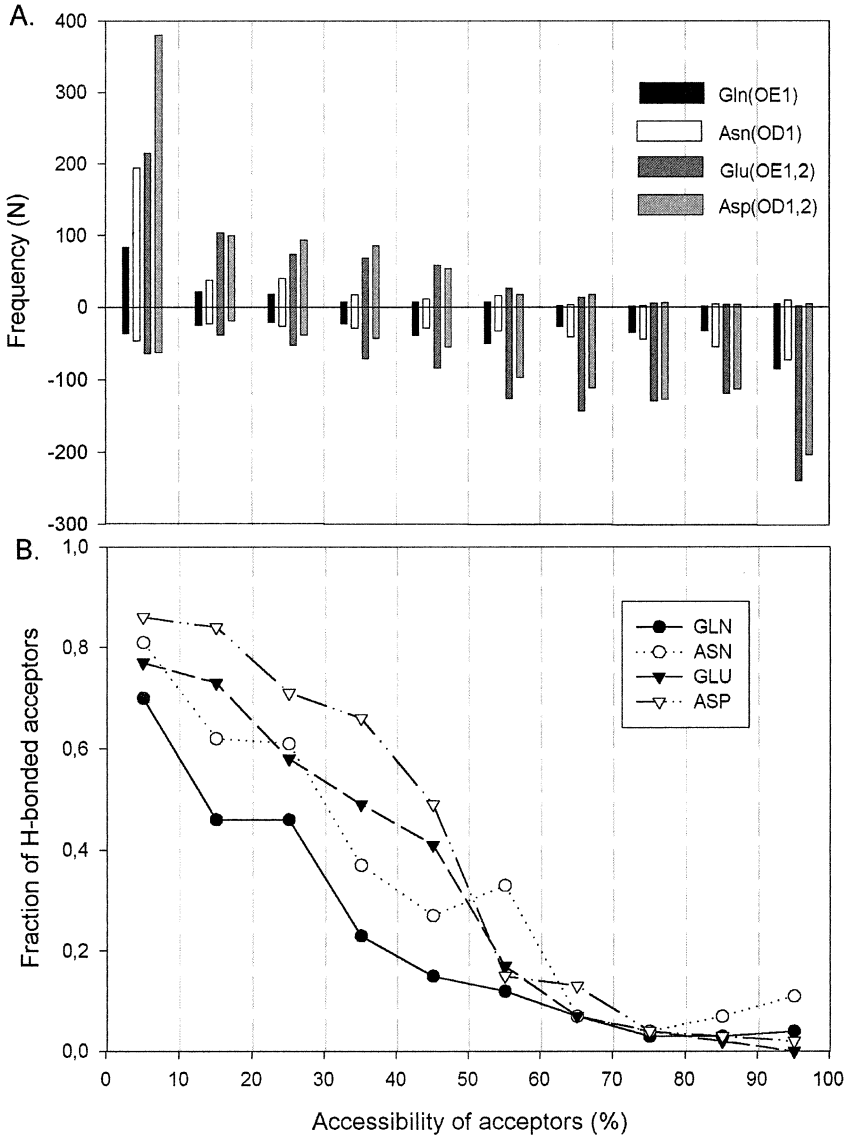

Fig. 2. A: Frequency of occurrence of H-bonded (positive values) and non-H-bonded (negative values) acceptors (atoms of Gln(OE1), Asn(OD1), Glu(OE1,2) and Asp(OD1,2)) in the database versus their accessibility. B: Fraction of the H-bonded acceptors in the database as a function of their solvent accessibility.

the lacking or small competitive hydrogen bonding with water molecules. Qualitatively, our results are in agreement with the estimates of Scheraga and his colleagues [13]. For buried polar groups, the enthalpy of formation of the hydrogen bond was estimated to be −6.0 kcal/mol, but for groups exposed to water the value was estimated to be −1.5 kcal/mol; the corresponding values of the free energy were estimated to be −4.5 to −5.0 kcal/mol and 0 to −0.6 kcal/mol, respectively [13].

In order to estimate the notional free energy of hydrogen bonding, ΔG_{Hb}, we used the approach proposed by Miller et al. [14] for calculation of surface–interior partition coefficients and free energies of transfer of different residues:

$$\Delta G_{\mathrm{Hb}} = -RT\ln f_{\mathrm{ASA}} \qquad (2)$$

where $RT \approx 0.6$ kcal/mol at room temperature and f_{ASA} is the partition coefficient at a fixed accessible surface area (ASA) of a polar atom. In the first approximation, the effective partition coefficient can be calculated as the ratio $N_{\text{H-bonded}}/N_{\text{non-H-bonded}}$ of the number of H-bonded to non-H-bonded donors or acceptors of a given type at a fixed solvent accessibility.

In a simple way, partition coefficients can be calculated from the fraction of H-bonded donors and acceptors. For example, at ASA ≤ 10% the fraction of the H-bonded donors in Fig. 1B ranges from ∼80% (Arg(NE)) to 99% (Lys(NZ)), so f_{0-10} ranges from 80/20 = 4 to 99/1 = 99. In accordance with Eq. 2, the notional free energy ranges from −0.8 kcal/mol to −2.8 kcal/mol. Taking into account that atom NE of Arg usually forms one hydrogen bond and buried atom NZ of Lys can form three H-bonds, the free energy of formation of one intramolecular hydrogen bond by these donors is estimated to vary from −0.7 to −0.8 kcal/mol at a solvent accessibility less than 10%. Calculations for Tyr, Thr and Ser give ΔG_{Hb} values of −0.6, −1.4 and −1.2 kcal/mol, respectively. Taking into account that atom OH of Tyr usually forms one hydrogen bond (see, e.g. [15]) and buried OG atoms of Thr and Ser form on average 1.7 and 1.5 hydrogen bonds (derived

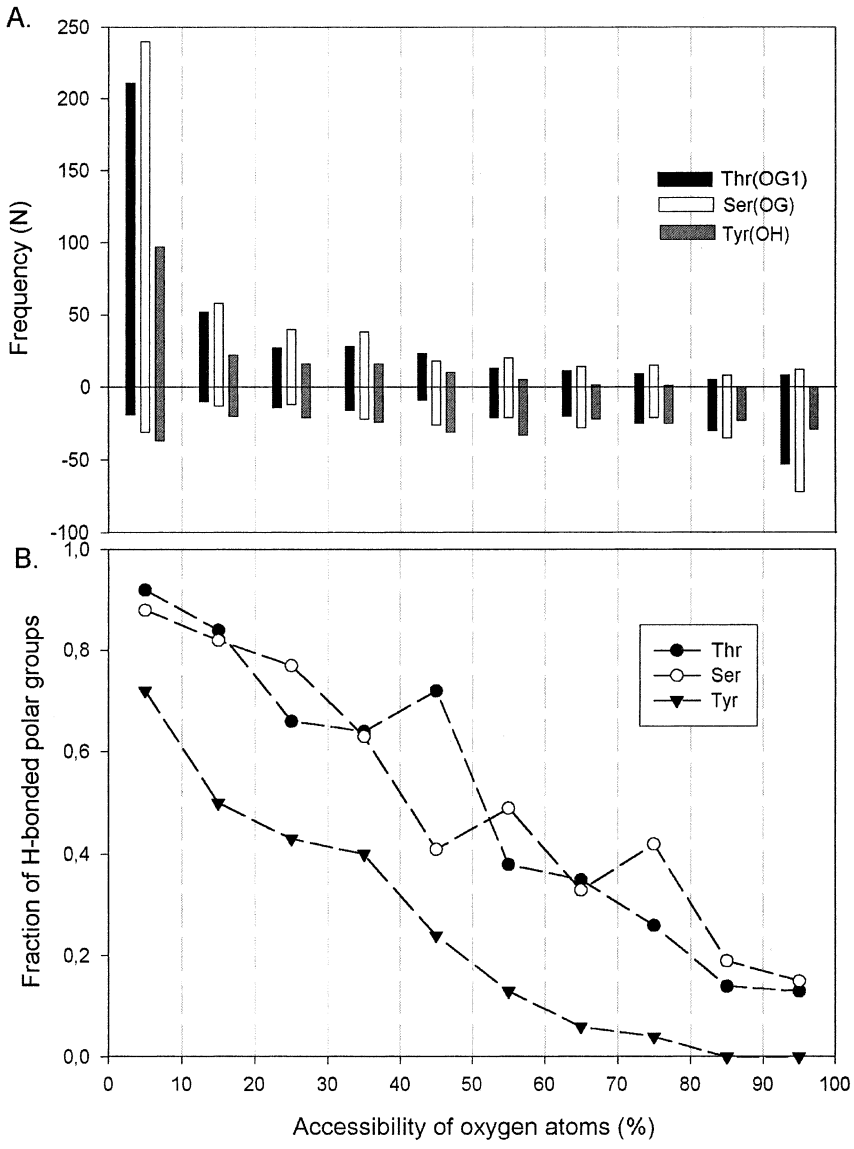

Fig. 3. A: Frequency of occurrence of H-bonded (positive values) and non-H-bonded (negative values) side-chain oxygens of Thr(OG1), Ser(OG) and Tyr(OH) in the database versus their accessibility. B: Fraction of the H-bonded oxygen atoms as a function of their solvent accessibility.

from the data presented in [16]), the free energy of formation of one H-bond is estimated to be −0.6 kcal/mol for Tyr and about −0.8 kcal/mol for Thr and Ser at accessibility of their oxygen atoms less than 10%. Similar estimates for the donors and acceptors of Asn, Asp, Gln and Glu give the values of −0.5 to −0.6 kcal/mol for one hydrogen bond. It should be noted that our estimates of the notional free energy of hydrogen bonding are in good agreement with recent theoretical estimates (the corrected $\Delta(\Delta G) = 2.3 \pm 1.1$ kcal/mol for buried polar groups and 2.0 ± 1.2 kcal/mol for half-buried and exposed polar groups [4]) and recent experimental data (the free energy of interaction varies from −0.4 to −2.0 kcal/mol [1–6,9]). All this demonstrates that intramolecular hydrogen-bonding interactions of buried and half-buried donors and acceptors of polar side chains can contribute favorably to the stability of a protein molecule.

In accordance with Eq. 2, there is a correlation between the fraction of the H-bonded donors and acceptors and their notional free energy of hydrogen bonding. As seen in Figs. 1–3, the fraction of H-bonded donors and acceptors decreases from left to right. When its value becomes equal to 50%, $\Delta G_{Hb} = 0$. In most cases, this is observed at an accessibility of 30–50%. At higher accessibility of the donors and acceptors, the intramolecular hydrogen bonding is unfavorable.

It can be calculated that the accessibility of about 70% of the donors and acceptors participating in intramolecular hydrogen bonding is less than 30%. The observation that buried polar groups in proteins are almost invariably hydrogen-bonded is well known [17–19] (see also [14–16]). One reason for this is that buried polar groups should satisfy their hydrogen-bonding potential. On the other hand, the data presented here suggest that the burial of hydrogen bonds is favorable. It seems very likely that proteins fold so as to form as many buried hydrogen bonds as possible.

Acknowledgements: This work was supported in part by the Russian Foundation for Basic Research (Grant #01-04-48399).

References

[1] Scholtz, J.M., Qian, H., Robbins, V.H. and Baldwin, R.L. (1993) Biochemistry 32, 9668–9676.
[2] Huyghues-Despointes, B.M.P., Klinger, T.M. and Baldwin, R.L. (1995) Biochemistry 34, 13267–13271.
[3] Thorson, J.S., Chapman, E. and Schultz, P.G. (1995) J. Am. Chem. Soc. 117, 9361–9362.
[4] Myers, J.K. and Pace, C.N. (1996) Biophys. J. 71, 2033–2039.
[5] Stapley, B.J. and Doig, A.J. (1997) J. Mol. Biol. 272, 465–473.
[6] Loladze, V.V., Ermolenko, D.N. and Makhatadze, G.I. (2002) J. Mol. Biol. 320, 343–357.
[7] Horovitz, A., Serrano, L., Avron, B., Bycroft, M. and Fersht, A.R. (1990) J. Mol. Biol. 216, 1031–1044.
[8] Dao-pin, S., Sauer, V., Nicholson, H. and Matthews, B.W. (1991) Biochemistry 30, 7142–7153.
[9] Makhatadze, G.I., Loladze, V.V., Ermolenko, D.N., Chen, X.F. and Thomas, S.T. (2003) J. Mol. Biol. 327, 1135–1148.
[10] Tissot, A.C., Vuilleumier, S. and Fersht, A.R. (1996) Biochemistry 35, 6786–6794.
[11] Waldburger, C.D., Schildbach, J.F. and Sauer, R.T. (1995) Nat. Struct. Biol. 2, 122–128.
[12] Hooft, R.W.W., Sander, C. and Vriend, G. (1996) Proteins 26, 363–376.
[13] Nemethy, G., Steinberg, I.Z. and Scheraga, H.A. (1963) Biopolymers 1, 43–69.
[14] Miller, S., Janin, J., Lesk, A.M. and Chothia, C. (1987) J. Mol. Biol. 196, 641–656.
[15] McDonald, I.K. and Thornton, J.M. (1994) J. Mol. Biol. 238, 777–793.
[16] Kondratova, M.S. and Efimov, A.V. (2002) Mol. Biol. (Moscow) 36, 144–151.
[17] Chothia, C. (1975) Nature 254, 304–308.
[18] Chothia, C. (1976) J. Mol. Biol. 105, 1–14.
[19] Richards, F.M. (1977) Annu. Rev. Biophys. Bioeng. 6, 151–176.