© 2004 Nature Publishing Group http://www.nature.com/naturebiotechnology

Keeping dry and crossing membranes

Ariel Fernández

Water sustains life but also imposes constraints on what life should be like. Its high dipole moment, proton donor-acceptor dual nature and rotational freedom make it a powerful former of hydrogen bonds. In soluble proteins, this property places stringent constraints on the way proteins interact[1–4] and fold[5]. I argue here that the manipulation of intramolecularly underdehydrated or underwrapped electrostatic interactions in proteins can be exploited in engineering strategies to create molecules with an enhanced ability to traverse biological membranes, with potential implications for oral delivery of peptide-based drugs.

How water gets dry

An intramolecular electrostatic interaction in a soluble protein prevails only if the competing hydration of the polar or charged groups is precluded by the structural design itself[1–5]. Water tends to attack underprotected intramolecular hydrogen bonds and exposed salt bridges, being particularly 'unforgiving' to interactions exposed to bulk solvent[1–3]. This property dictates that essential parts of protein structure, like the backbone hydrogen bonds, should be 'kept dry' if the structure is to prevail. It also constrains the protein folding[5] and interactions[1–4]. For instance, the failure to completely dehydrate backbone hydrogen bonds intramolecularly makes a soluble protein reliant on binding partnerships[1] for the preservation of its fold.

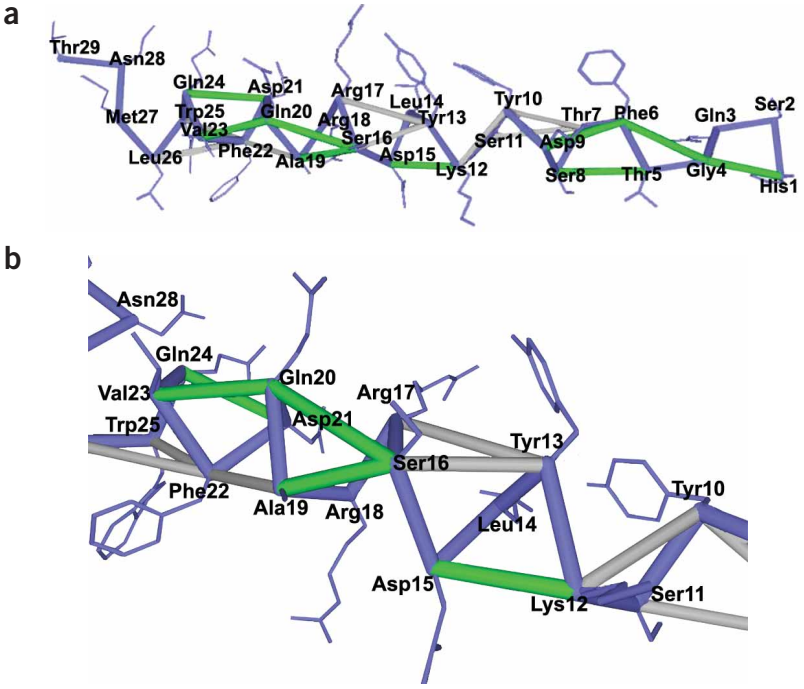

Figure 1 Wrapping of the glucagon molecule in helical conformation (PDB entry 1GCN). (**a**) Full molecule; (**b**) detail. The backbone is indicated by blue virtual bonds joining α-carbons. The well-wrapped backbone hydrogen bonds are indicated as gray segments, and the dehydrons are indicated in green.

When suitable partners are absent, a lack of proper wrapping or protection of the backbone hydrogen bonds may promote aberrant aggregation[3,6] or nonspecific protein-membrane associations[2] as a means to improve the protection of the backbone from water attack in detriment of the information encoded in the primary sequence[6]. This phenomenology may be rationalized after the experimental demonstration that packing defects in proteins are actually sticky, as electrostatic interactions become enhanced and stabilized through the removal of surrounding water[7].

Certain regions of the protein surface are naturally designed to be dry in water. When water is severely confined, as it may be when in proximity to hydrophobic residues forming a protein cavity, a sufficient number of hydrogen bond partnerships are lost, lowering the fluid density at the interface[8]. This unfavorable interface leads to the formation of a gaslike layer (drying) through capillary evaporation[9]. Such constraints make it entropically and enthalpically too costly for water to remain in the confining region, thus inducing a 'drying transition'[9–11].

Because backbone hydrogen bonds are determinants of the basic structural motifs, the α-helix and β-sheet, protein structure can only prevail in water if the majority of such bonds are adequately dehydrated[1–5]. This desolvation requires that nonpolar groups cluster around the amide-carbonyl hydrogen bonds, expelling the surrounding water; thus, stable soluble structure must

Ariel Fernández is in the Indiana University School of Informatics and Center for Computational Biology and Bioinformatics, Indiana University Medical School, 714 N. Senate Avenue, Indianapolis, Indiana 46202, USA, and the Department of Computer Science, The University of Chicago, Chicago, Illinois 60637, USA (permanent addresses); Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871, Japan.
e-mail: arfernan@iupui.edu

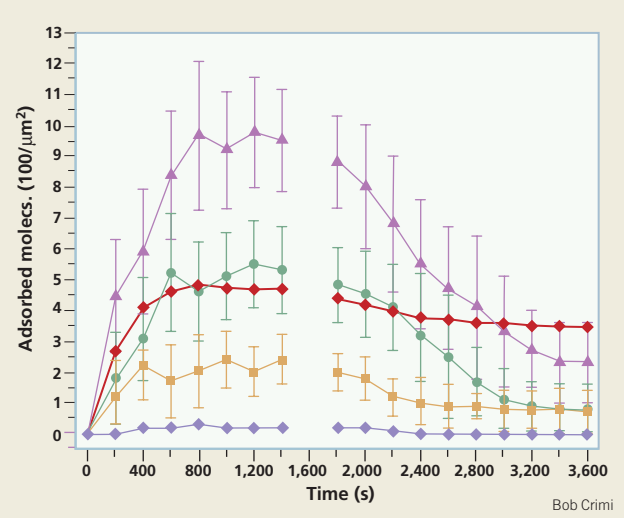

Figure 2 Time dependence of the adsorption uptake and desorption of soluble mutants of the G4V glucagon peptide onto and from a dilauryl phosphatidylcholine bilayer coating a biosensing optical device. The adsorption experiments were conducted at 318 K under controlled hydrodynamic conditions identical to those provided in refs. 2 and 7. Adsorption was assayed by evanescent-field total-reflection spectroscopy to reach equilibrium at 1 μM bulk peptide concentration. This concentration was reduced to 0 μM within the interval 1,400–1,800 s, after equilibration was reached at 1,400 s in all cases. The mutation G4V in all peptides prevents the exposure of the backbone amide, a bottleneck for penetration into the lipid phase. Eight measurements were made for each peptide, and the error bars indicate the dispersion in adsorption uptake values. Comparable dispersions were obtained for the A19V mutation but have been omitted for clarity. Red diamonds, A19V; lavender diamonds, D15N; yellow squares, G4V glucagon; green circles, S16E; purple triangles, S16E/D15E double mutant.

© 2004 Nature Publishing Group http://www.nature.com/naturebiotechnology

thoroughly 'wrap' its hydrogen bonds[1–5] as a means to preserve its integrity. A dehydrated hydrogen bond is more stable (by roughly an order of magnitude in free-energy terms) than one formed in bulk water[12,13].

Dehydrons and wrapping

To assess the significance of this building constraint, Scott and I[13] have systematically investigated the microenvironments for backbone hydrogen bonds in all high-resolution (2 Å or better) Protein Data Bank (PDB) structures. Strikingly, we found that 95% of the stable soluble proteins examined keep at least 92% of their backbone hydrogen bonds 'dry' by completely wrapping them. The extent of wrapping of a hydrogen bond can be assessed by defining two desolvation spheres centered at the α-carbons of the bonded residues and counting the number of side chain nonpolar groups within the spheres[13].

Underwrapped hydrogen bonds in monomeric structure, now termed 'dehydrons'[13], typically signal binding regions[1,7,13]. Nonpolar groups from the binding partner approach preformed dehydrons in the monomeric structure, and in so doing, they strengthen them by markedly lowering the solvent screening of the partial charge on the amide and carbonyl. A similar principle should apply to underwrapped salt bridges, termed here 'desolvons', which engage charged groups. The 'attraction prompted by water removal' is essentially a three-body effect: the nonpolar group is attracted to the hydrogen bond or salt bridge because, in so doing, it expels the water surrounding the interaction, making it both stronger and more stable. This force also has a thermodynamic component, in that the hydration of the amide and carbonyl (or charged groups) is hindered by the approach of a nonpolar group, and thus the state with the broken bond is destabilized[12,13].

In this light, the hydrogen bond and salt bridge could be regarded as hydrophobic entities in the broadest sense[13]. This fact was originally noted by Roseman[12], who proved using a thermodynamic cycle argument that the transfer of an amide and carbonyl group from water to a nonpolar phase is far more favorable when such groups are paired by a hydrogen bond than when transferred in isolation.

It should be possible to 'intermolecularly correct' the structural deficiencies of monomeric protein structure by exogenously dehydrating its hydrogen bonds. This property clearly holds because dehydrons (or desolvons) favor energetically and thermodynamically the removal of surrounding water to an extent comparable to exposed hydrophobic patches[7]. This free-energy decrease is invariably a sufficient compensation for the work required to remove water from the dehydron surroundings, because such water is already partially confined and has lost hydrogen bond partnerships[14] while attached to the dehydron.

Thus, it should not be surprising to find an intimate relationship between protein interactivity and the inherent inability of the monomer to keep itself completely 'dry' where it is most crucial to preserve the structure. This relationship provides a vivid biological realization of the model for hydrophobicity at nanoscales[8].

A clear pattern arises as we examine systematically the surface of soluble proteins and the protein association interfaces. Although binding sites are indeed regions where it is most advantageous to exclude surrounding water, such regions seldom involve patches of overexposed hydrophobes—actually a rare occurrence on the surface of the protein. Binding sites typically involve underwrapped hydrogen bonds[1–3].

Structures that constitute severe 'underwrappers' of their backbone hydrogen bonds almost invariably either correspond to toxins, requiring a profusion of disulfide bridges to preserve their structural integrity, or represent major anomalies, like amyloidogenic proteins[2,3], which undergo major rearrangement and aggregation.

Making peptides go places

Assuming that a protein design principle can be built around the concept of underwrapped electrostatic interactions, there are several implications for protein engineering strategies. One major challenge in the oral delivery of peptide-based drugs is the need to ensure their effective transportation across the cellular membrane and into the intracellular aqueous space[15,16]. Thus, an intramolecularly underwrapped electrostatic interaction might facilitate such transportation, in that a partition coefficient between the lipid (anhydrous) and the aqueous phase may be established as a result of two competing effects: the dehydration of preformed electrostatics and the hydration of the polar groups. The former fosters penetration into the lipid phase, whereas the latter is favored in an aqueous environment. Thus, a conformational switch yielding alternatively the bonded or nonbonded state might effectively generate the partition coefficient required for drug delivery across cellular membranes.

I envision a new generation of peptide-based 'trojan' drugs whose success depends on their autonomous, spontaneous and effective delivery articulated by an unspecific percolation through cellular membranes. Through internal structural change, such peptides must not only become soluble in

© 2004 Nature Publishing Group http://www.nature.com/naturebiotechnology

the lipid but also be capable of modulating their solubility in the nonpolar phase so as to eventually become water soluble, as required for proper delivery. The postulated peptide drug would modulate its water/lipid partition coefficient through conformational change.

Although engineering of a hydrophobic patch would probably provide inadequate partition switching for such a protein because of its high instability in water, engineering of an underwrapped region might be more promising[2]. This is because the hydration of amide and carbonyl, a natural occurrence in the water phase, effectively competes with the dehydration of the hydrogen bond, which engages the amide and carbonyl and takes place in the lipid (anhydrous) phase. Fostering and manipulating this competition might very well become the design principle required to build a peptide-based drug capable of autonomously traversing the cell membrane.

As emphasized earlier, the design potentialities of the dehydron and desolvon concepts need to be assessed with regard to their impact for oral delivery[15,16]. If a peptide chain forms dehydrons or desolvons within a frequently visited conformation, a partition coefficient between the lipid membrane and the intracellular space is defined by two opposing processes: (i) the dehydration of an intramolecular electrostatic interaction and (ii) the hydration of the polar groups associated with the loss of the interaction. Such investigations may shed light on how fundamental interactions between hormone peptides and their membrane-bound receptors take place, because such interactions are known to require specific structural motifs[17] that are unlikely to prevail in bulk water.

Proof of principle

To illustrate these possibilities, I chose glucagon, a 29–amino acid hormone peptide that regulates glucose metabolism[18]. This system is suitable because it represents a good compromise between size and biological relevance. The molecule is assumed to form partially exposed i, i+3 and i, i+4 salt bridges, provided it adopts an α-helical conformation[19]. Furthermore, the i, i+3 and i, i+4 backbone hydrogen bonds involving any of the six poor-wrapping residues (Ala, Ser, Thr, Asp, Asn, Gly)[20] will invariably be dehydrons because of the lack of a large-scale context (that is, lack of tertiary structure) needed to fully dehydrate the helix. Such dehydrons provide sticky regions as the peptide visits the helix conformation in the bulk-water phase[7]. However, such dehydrons are also spots vulnerable to water attack, and thus the helical conformation is expected to be unstable in bulk water. On the other hand, preformed underwrapped electrostatic interactions in water, however ephemeral they may be, are required to ensure membrane penetration through dehydration, and thus mutants have been designed to 'correct' structural deficiencies in glucagon, promoting the occurrence of underwrapped electrostatic interactions. One basic mutation, G4V, appears necessary in any peptide with an adequate water/lipid partition coefficient. The substitution is required to prevent exposure of the backbone amide. These properties of the helical structure of glucagon lead us to suspect that the acquisition of helical structure is required for the penetration into the membrane bilayer, which in turn ensures the structure preservation by bolstering the dehydration of the salt bridges and hydrogen bonds.

Thus, my group has designed mutants intended to partially increase the wrapping of purported salt bridges in the helical state as well as stabilizing the helical conformation by wrapping the backbone hydrogen bonds. The latter condition is required to ensure that the salt bridges have a chance to form in water in the first place.

If G4V glucagon were to adopt in bulk water the same helical structure that it adopts in the crystal (PDB entry 1GCN), the backbone would present a high concentration of packing defects in the region defined by the dehydrons Asp21–Gln24, Gln20–Val23, Ser16–Gln20, Ser16–Ala19 and Lys12–Asp15, as indicated in **Figure 1**. Such a high number of packing deficiencies implies that the helix will not prevail in water unless poor-wrapping residues, such as Ser16 or Ala19, are modified to protect the backbone. On the other hand, the helical conformation can accommodate three desolvons: Lys12–Asp15, Asp15–Arg18 and Arg17–Asp21. Thus, the mutants must be designed to protect the backbone and at the same time to dehydrate or protect the salt bridges, which, being marginally stable in bulk water, would benefit from exogenous dehydration and thus enhance the affinity of the peptide for the lipid phase.

The wild-type helical structure of such a short peptide is unlikely to be stable in water because of the absence of a tertiary structure scaffold required to enhance the dehydration of the backbone hydrogen bonds[5]. However, the helical conformation is undoubtedly sufficiently stable to fulfill the stereochemical demands in glucagon crystallization. Given my design objectives, a high helical stability in water is actually undesirable, because the stickiness of the helix arises precisely as a consequence of its vulnerability to water attack[7].

Mutant glucagon molecules have been used to test the hypothesis that the presence of dehydrons and desolvons facilitate transportation of peptides across lipid phases. Besides the ubiquitous mutation G4V, the assayed mutants were A19V, D15N; S16E and S16E/D15E. No structure for any of the mutant peptides has been reported in the PDB at this time. Given my design objectives, such constraint is probably immaterial: The focus of this study is the creation or deletion of sticky packing defects in the helical conformation—metastable as it may be in bulk water—and the modulation of the vulnerability of the helix to water attack as a means of manipulating the membrane-traversing propensity. Thus, although the mutants and wild type are not likely to fold into a stable helical conformation because of their inability to form tertiary structure, it is precisely the metastability of the helix that confers the propensity to traverse the lipid phase.

We adapted a previously designed setup[2,7] to probe membrane penetration for the different peptides. Thus, protein adsorption onto and desorption from a Langmuir-Blodget dilauroyl phosphatidylcholine bilayer is measured under controlled hydrodynamic conditions with a constant flux fixed at 5×10^{-3} cm^3/s. Adsorption took place at the constant bulk concentration of 1 μM and desorption at 0 μM bulk concentration; variations in the adsorption uptake for different peptides were monitored by evanescent-field total-reflection spectroscopy using an optical biosensing device[2,7], which interrogates the coating of an optical waveguide serving as the floor of the cell.

The adsorption-desorption profiles displayed in **Figure 2** reveal that adsorption equilibration is achieved at t = 1,400 s for all peptides. Obviously, the most desirable peptide for delivery across a lipid phase is one with the highest adsorption equilibrium uptake followed by the most complete desorption, once the bulk concentration is reduced from 1 μM to 0 μM within the 1,400- to 1,800-s interval.

The A19V mutation provides better wrapping to dehydrons 16–19 and 16–20, thus scaffolding the helical structure and thereby fostering the formation of the i, i+3 and i, i+4 desolvons. Furthermore, it contributes to dehydrate and thus stabilize the 15–18 desolvon. Although this mutation triggers an appreciable increase in adsorption uptake in

© 2004 Nature Publishing Group http://www.nature.com/naturebiotechnology

comparison with wild-type levels, it also introduces a hydrophobic surface residue. This increased hydrophobicity precludes proper desorption (**Fig. 2**), and thus, it renders the peptide an unlikely candidate for delivery into the intracellular space.

The D15N mutation obviously removes the 15–18 desolvon, leaving the charged Arg18 exposed to the solvent. Thus, the adsorption uptake is markedly reduced with respect to wild-type G4V levels, making this mutation deleterious for transportation across an anhydrous phase, because that process would markedly increase the self-energy of Arg18.

In contrast, the S16E mutation seems very adequate for across-lipid peptide transport, with its considerable degree of adsorption and desorption. This mutation introduces a new Glu16–Arg17 desolvon and stabilizes the helical structure because it contributes to the wrapping of the 16–19, 16–20 and 12–15 backbone hydrogen bonds.

On the other hand, the S16E/D15E double mutation produces the highest adsorption uptake (~0.15 mM equilibrium concentration in the lipid phase), although the incomplete desorption might signal less efficient delivery into the intracellular space. This mutation contributes to more effectively wrap the backbone (similarly to the S16E mutation) and contributes to further protect or dehydrate the purported salt bridges 12–15, 15–18 and even the new salt bridge Glu16–Arg17, thus increasing the dehydration propensity of the peptide, a trigger for penetration into the lipid phase. On the other hand, the mutation does not introduce hydrophobic residues, which would prevent the reinsertion into the aqueous phase.

Several control mutations have been assayed for their adsorption-desorption behavior. Some mutants were designed to enhance the surface hydrophobicity of the helical conformation (the standard trigger for membrane penetration) as well as to delete desolvons, thus leading to the exposure of unpaired polar side chains. Thus, the adsorption uptake of mutants D15V (or D15A) and D21V (or D21A) was found to be virtually undetectable (of the order of 5 molecules per μm^2, or lower). These mutations remove the desolvons 15–18 and 17–21, respectively, thus markedly increasing the self-energy of Arg18 or Arg17 in the lipid phase. Therefore, although the hydrophobicity is increased with both mutations, the removal of sticky packing defects entails a thermodynamic cost sufficiently high to preclude membrane penetration. On the other hand, the helix-stabilizing mutation G4V provides extra hydrophobicity to the peptide without altering its desolvon pattern and introduces a better wrapping of the 6–4 and 4–1 dehydrons. Despite the enhanced hydrophobicity, the maximum adsorption uptake for this mutant is barely detectable (10 molecules per μm^2), because of the persistent instability of the 12–15 and 15–18 desolvons. These desolvons remain so poorly wrapped that penetration into the lipid phase probably entails the burial of unpaired charges most of the time. On the other hand, no desorption was observed for this mutant, as expected in light of its greater hydrophobicity.

The tests provided argue on purely physicochemical grounds for the importance of partially dehydrated electrostatics as a concept to design peptides susceptible of delivery across anhydrous lipid phases and membranes. The biomedical impact of this design principle remains to be assessed. Thus, I envision a new generation of trojan drugs endowed with conformational switches to camouflage the structural features that make them water-soluble, replacing them with dehydration promoters.

ACKNOWLEDGMENTS

This research was supported in part by an unrestricted grant from Eli Lilly and Company. The work was motivated by conversations with Richard DiMarchi.

1. Fernández, A. & Scheraga, H.A. *Proc. Natl. Acad. Sci. USA* **100**, 113–118 (2003).
2. Fernández, A. & Berry, R.S. *Proc. Natl. Acad. Sci. USA* **100**, 2391–2396 (2003).
3. Fernández, A., Kardos, J., Scott, R.L., Goto, Y. & Berry, R.S. *Proc. Natl. Acad. Sci. USA* **100**, 6446–6451 (2003).
4. Fernández, A., Scott, R.L. & Berry, R.S. *Proc. Natl. Acad. Sci. USA* **101**, 2823–2827 (2004).
5. Fernández, A., Sosnick, T.R. & Colubri, A. J. *Mol. Biol.* **321**, 659–675 (2002).
6. Fändrich, M. & Dobson, C.M. *EMBO J.* **21**, 5682–5690 (2002).
7. Fernández, A. & Scott, R.L. *Phys. Rev. Lett.* **91**, 18102–18106 (2003).
8. Ball, P. *Nature* **423**, 25–26 (2003).
9. Lum, K., Chandler, D. & Weeks, J.D. *J. Phys. Chem. B* **103**, 4570–4577 (1999).
10. Hummer, G., Rasaiah, J.C. & Noworyta, J.P. *Nature* **414**, 188–190 (2002).
11. Fernández, A. *J. Chem. Phys.* **119**, 5315–5319 (2003).
12. Roseman, M.A. *J. Mol. Biol.* **201**, 621–624 (1988).
13. Fernández, A. & Scott, R.L. *Biophys. J.* **85**, 1914–1928 (2003).
14. Carey, C., Cheng, Y.-K. & Rossky, P. *Chem. Phys.* **258**, 415–425 (2000).
15. Goldberg, M. & Gomez-Orellana, I. *Nat. Rev. Drug Discov.* **2**, 289–295 (2003).
16. Leone-Bay, A. *et al. Pharm. Res.* **18**, 964–970 (2001).
17. Unson, C.G. & Merrifield, R.B. *Proc. Natl. Acad. Sci. USA* **91**, 454–458 (1994).
18. Dobbs, R. et al. Science **187**, 544–547 (1975).
19. Sapse, A.M., Rothchild, R., Jain, D.C. & Unson, C.G. *Mol. Med.* **8**, 251–262 (2002).
20. Fernández, A., Scott, R.L. & Scheraga, H.A. *J. Phys. Chem. B* **107**, 9929–9932 (2003).