The Role of Conditional Hydration on the Thermodynamics of Protein Folding

Raffaele Ragone* and Giovanni Colonna

Department of Biochemistry and Biophysics
II University of Naples, via Costantinopoli 16
80138 Naples, Italy

Received November 19, 1993

The mess of ideas[1-11] about the role played by hydrophobic hydration in protein folding calls for a reassessment. Trying to shed light on this matter, Herzfeld[12] pointed out that protein unfolding should be modeled by a two-step pathway, i.e., fusion of the protein core followed by hydration of interior surface. This is not at all new. Evidences for a crystal-like packing of protein core come from packing density[13-16] as well as from compressibility studies.[17] Such a dissection of the unfolding process implies that the hydration of protein interior has to be modeled by the transfer from a pure liquid phase into water, according to the classical view.[11] However, following Ben-Naim's arguments,[18] purely hydrophobic molecules should not be the model of choice to mimic the nonpolar hydration in proteins. This approach ignores that amino acid side chains are linked to the peptide backbone. Ben-Naim[18] claims that an "ideal" hydrophobicity scale of amino acid residues must reflect their "conditional" solvation in proteins. Far from getting involved in the definition of a new hydrophobicity scale, we try here to throw light upon the implications of the "conditional" hydration of nonpolar surface in protein unfolding.

We propose that the unfolding free energy ($\Delta G_{unf}°$) can be dissected as

$$\Delta G_{unf}° = \Delta G_{fus}° + \Delta G_{pol}° + \Delta G_{np}° \quad (1)$$

Here $\Delta G_{pol}°$ and $\Delta G_{np}°$ represent the polar and nonpolar parts of the hydration free energy. $\Delta G_{fus}°$ is the fusion free energy. The dissection of $\Delta G_{unf}°$ accomplished by Murphy and co-workers[1] identifies enthalpic and entropic contributions not directly imputable to water interaction with the buried protein surface. These contributions are represented by the convergence unfolding enthalpy (ΔH^*) and entropy (ΔS^*), which were observed for some globular proteins.[1] They constitute a residual free energy ($\Delta G^* = \Delta H^* - T\Delta S^*$), which does not depend on the hydration of the protein core. We suggest that $\Delta G_{fus}°$ should be identified with ΔG^*. This is not hard to believe, since ΔS^* actually represents a fusion entropy.[1,19] Moreover, ΔH^* was attributed to hydrogen bonding as well as to dispersion forces within the protein matrix,[1,19] which are typically involved in melting processes. It was also shown[20] that the enthalpy–entropy dissection[1,21] can be described by just one equation, suggesting that ΔH^* and ΔS^* are closely related. As a consequence, from eq 1 we also have

$$\Delta G_{pol}° + \Delta G_{np}° = \Delta C_p°[T - T_h^* - T \ln (T/T_s^*)] \quad (2)$$

$\Delta C_p°$ is the unfolding heat capacity change, which relates the water exposure of buried surface to the hydrophobic hydration.[2,3] T_h^* and T_s^* represent the temperatures at which the specific unfolding enthalpy and entropy of several globular proteins converge, respectively, assuming values indicated by ΔH^* and ΔS^*.[1]

T_h^* and T_s^* were interpreted as the temperatures at which hydrophobic enthalpy (T_h^*) and entropy (T_s^*) vanish, respectively.[1] The original assumption was that they were coincident.[1] It is now clear that T_h^* is lower than T_s^* of about 8 K, with $T_h^* \cong 377$ K and $T_s^* \cong 385$ K.[6,7,20,22,23] The meaning originally attributed to T_h^* does not reconcile with the fact that the hydration enthalpy of small nonpolar molecules vanishes at a much lower temperature ($T_{np} \cong 295$ K)[11] than that invoked for protein unfolding ($T_h^* \cong 377$ K). On the contrary, there is agreement about T_s^*. This is evidenced by Baldwin's liquid hydrocarbon model.[11] A molecular mechanism justifying the existence of convergence temperatures in protein unfolding was proposed by Lee,[5] although Lee's argument has been shown[24] to be mathematically equivalent to the original view.[1,21] For explaining the large difference between T_{np} and T_h^* it was also suggested that the protein hydration resembles more closely that of "compact" gases[23,25-28] or solid diketopiperazines[22,24,29,30] than that of liquids. Accordingly, it has to be explained why proteins or solids are more similar to gases than to liquids.[8]

As can be seen, eq 2 is reduced to the hydration free energy of small nonpolar compounds when $\Delta G_{pol}° = 0$ and $T_h^* \cong 295$ K $= T_{np}$.[11] Thus, it seems likely that the high T_h^* value of proteins is mostly due to $\Delta G_{pol}°$. In particular, since there is no doubt about T_s^* (the temperature at which the hydrophobic entropy vanishes),[1,11] it appears that the polar part of the hydration enthalpy ($\Delta H_{pol}°$) plays a role in determining T_h^*. As a first approximation, we assume that $\Delta H_{pol}°$ largely depends on the hydration of the peptide backbone following protein unfolding. $\Delta H_{pol}°$ is a pure hydration term, devoid of hydrogen bond or dispersion forces effects. Thus, it cannot be represented by ΔH^*, that was indicated as the unique contributor to the polar enthalpy in the model proposed by Lee.[5] The novelty of our proposal relies on the separation of the melting enthalpy from polar contributions. It seems that this has not been fairly accounted for. We suggest that the sum between a polar (negative) and a nonpolar (positive) hydration enthalpy shifts the temperature at which the total hydration enthalpy vanishes to a value near the boiling point of water.

To model the "hydration effect" in proteins, liquid amides should be more suitable than hydrocarbons for reflecting the "conditional" water solvation of amino acid residues. Amides have already been used to evaluate the role played by polar interactions in protein folding.[31] However, this was performed assuming that $\Delta C_p°$ of protein unfolding can be dissected into polar and nonpolar

* Author can be contacted at the following: phone (81)-5665869-5665858 and fax (81)-5665869-5665863.

(1) Murphy, K. P.; Privalov, P. L.; Gill, S. J. *Science* **1990**, *247*, 559–561.
(2) Dill, K. A. *Science* **1990**, *250*, 297.
(3) Dill, K. A. *Biochemistry* **1990**, *29*, 7133–7155.
(4) Muller, N. *Acc. Chem. Res.* **1990**, *23*, 23–28.
(5) Lee, B. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5154–5158.
(6) Baldwin, R. L.; Muller, N. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7110–7113.
(7) Doig, A. J.; Williams, D. H. *Biochemistry* **1992**, *31*, 9371–9375.
(8) Yang, A.-S.; Sharp, K. A.; Honig, B. *J. Mol. Biol.* **1992**, *227*, 889–900.
(9) Muller, N. *Trends Biochem. Sci.* **1992**, *17*, 459–463.
(10) Privalov, P. L.; Gill, S. J.; Murphy, K. P. *Science* **1990**, *250*, 297–298.
(11) Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8069–8072.
(12) Herzfeld, J. *Science* **1991**, *253*, 88.
(13) Klapper, M. H. *Biochim. Biophys. Acta* **1970**, *229*, 557–566.
(14) Chothia, C. *Nature* **1975**, *254*, 304–308.
(15) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151–176.
(16) Bello, J. *Int. J. Pept. Protein Res.* **1978**, *12*, 38–41.
(17) Kharakoz, D. P.; Sarvazyan, A. P. *Biopolymers* **1993**, *33*, 11–26.
(18) Ben-Naim, A. *Biopolymers* **1990**, *29*, 567–596.
(19) Privalov, P. L.; Gill, S. J. *Pure Appl. Chem.* **1989**, *61*, 1097–1104.
(20) Ragone, R.; Colonna, G. *J. Biol. Chem.*, in press.
(21) Murphy, K. P.; Gill, S. J. *Thermochim. Acta* **1990**, *172*, 11–20.
(22) Murphy, K. P.; Gill, S. J. *J. Mol. Biol.* **1991**, *222*, 699–709.
(23) Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1993**, *232*, 660–679.
(24) Murphy, K. P.; Bhakuni, V.; Xie, D.; Freire, E. *J. Mol. Biol.* **1992**, *227*, 293–306.
(25) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1990**, *213*, 375–384.
(26) Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1990**, *213*, 385–391.
(27) Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1992**, *224*, 715–723.
(28) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639–659.
(29) Murphy, K. P.; Freire, E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9335–9338.
(30) Murphy, K. P.; Freire, E. *Adv. Protein Chem.* **1992**, *43*, 313–361.

0002-7863/94/1516-2677$04.50/0

Table 1. Thermodynamic Properties of Amides

amide	A_{np}[a] (Å^2)	A_{pol}[a] (Å^2)	$\Delta C_p°$ [b] (J·mol^{-1}·K^{-1})	$\Delta H_{tr}°(298\ K)$[b] (J·mol^{-1})	T_h[c] (K)
HCONHMe			40 ± 2[d]	−7083 ± 5[d]	475 ± 9
MeCONHMe	179.3	63.8	107 ± 3[d]	−13090 ± 20[d]	420 ± 4
EtCONHMe	218.2	51.6	155 ± 3	−14870 ± 20	394 ± 2
PrCONHMe	245.0	50.1	227 ± 4	−16020 ± 20	369 ± 1
i-Pr-CONHMe	256.9	45.8	222 ± 5	−15790 ± 20	369 ± 2
BuCONHMe	278.4	52.3	286 ± 4	−15030 ± 20	351 ± 1
MeCONHEt	217.8	52.3	163 ± 4	−15480 ± 20	393 ± 3
MeCONHPr	247.1	48.3	230 ± 4	−15760 ± 20	367 ± 1
MeCONH-i-Pr	253.4	44.8	230 ± 4	−17240 ± 20	373 ± 1
MeCONHBu	270.6	47.5	280 ± 1[d]	−14720 ± 30	351 ± 3

[a] Nonpolar and polar surface area evaluated on minimum conformational energy transconformations by the program PCModel by Serena Software. [b] Data taken from ref 34, unless otherwise stated. [c] Calculated by the relationship $T_h = 298\ K - \Delta H°(298)/\Delta C_p°$. The error was evalulated assuming $\Delta(T_h) = [\delta T_h/\delta \Delta H°]\Delta(\Delta H°) + [\delta T_h/\delta \Delta C_p°]\Delta(\Delta C_p°)$. [d] Data taken from ref 35.

contributions.[31] This is still debated.[32] In Table 1 we report both enthalpy and heat capacity changes characterizing the transfer of several liquid amides. These quantities allow the calculation of the temperature (T_h) at which the transfer enthalpy, reflecting both polar and nonpolar hydration, goes to 0 for individual amides (last column of Table 1). As can be seen, T_h is far from T_{np} (295 K), approaching T_h* (377 K) typical of proteins. The strong correspondence between the amide T_h and T_h* of proteins suggests that the unfolding enthalpy at T_h* (ΔH*) actually represents a pure melting enthalpy, given that there the total hydration enthalpy is 0.

We have verified that the amide T_h approaches T_{np} when the fraction of nonpolar surface area becomes larger and larger. In Figure 1 T_h is plotted against the reciprocal number of $-CH_2-$ (n_C). The y-axis intercept (T_h when $n_C = \infty$) is within the experimental uncertainty of T_{np} of liquid hydrocarbons.[11] This value (~309 K) is also very close to that predicted by the hydrogen bonding model proposed by Muller.[4,32] Now, an amino acid residue with a molecular mass typical of globular proteins (~110 Da) contains ~4.8 $-CH_2-$. About 58% of these groups (~2.8 $-CH_2-$ per residue) is buried, since the fraction of buried nonpolar surface in proteins is 0.58, on the average.[33] On this basis, using the coefficients of the linear regression (see figure legend) and $n_C = 2.8$, the amide model predicts that T_h* of protein unfolding should be (392.7 ± 13.4) K. This value is, within the experimental error, close to that typical of proteins (T_h* ≅ 377 K), suggesting that amides are adequate to mimic the hydration of amino acid residues in protein unfolding. The difference between the predicted and the experimental value could also be attributed to the larger fraction of nonpolar surface of amides as compared to that involved in protein unfolding.[5,33] This can be appreciated by considering polar and nonpolar accessible surface areas, which are reported in Table 1.

In conclusion, the energetics of amide transfer into water shows that the enthalpy convergence temperature typical of protein unfolding (T_h*) actually reflects the total enthalpic effects arising from nonpolar and peptide backbone hydration. It is not necessary to introduce a new definition of hydrophobic hydration, a term that we use according to the operational definition of Dill.[2,3] Neither is it required to model protein hydration by gaseous or solid model compounds. The residual unfolding enthalpy (ΔH*) represents a melting contribution arising from hydrogen bonding and dispersion forces within the protein matrix.[1,19] As a consequence, any discrepancy between the different views about the role played by the hydrophobic free energy in protein folding vanishes. The next step should be concerned with the evaluation of the energetics of the protein core melting.

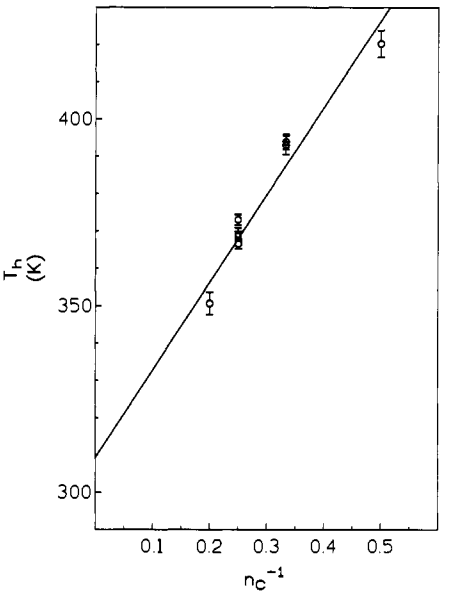

Figure 1. Effect of chain length on T_h of amides. The line represents the linear least-squares regression analysis of data reported in Table 1 [slope = (234.4 ± 20.3) n_C·K, intercept = (309.3 ± 6.1) K, r = 0.975]. Error bars are drawn according to Table 1. N-methylformamide was excluded from the analysis because its carbonyl is linked to a hydrogen, which does not occur in proteins.

Acknowledgment. This work was supported by grants from the "Target Project on Biotechnology and Bioinstrumentation" of the National Researches Council as well as from the Italian Ministry of University and Scientific Research (40% financial aid).

(31) Spolar, R. S.; Livingstone, J. R.; Record, M. T., Jr. *Biochemistry* **1992**, *31*, 3947–3955.
(32) Muller, N. *Biopolymers* **1993**, *33*, 1185–1193.
(33) Miller, S.; Janin, J.; Lesk, A. M.; Chothia, C. *J. Mol. Biol.* **1987**, *196*, 641–656.
(34) Konicek, J.; Wadsö, I. *Acta Chem. Scand.* **1971**, *25*, 1541–1551.
(35) Sköld, R.; Suurkuusk, J.; Wadsö, I. *J. Chem. Thermodyn.* **1976**, *8*, 1075–1080.