ELSEVIER

Biophysical Chemistry 105 (2003) 183–193

Biophysical Chemistry

www.elsevier.com/locate/bpc

Hydrophobic hydrophilic phenomena in biochemical processes

Arieh Ben-Naim*

Department of Physical Chemistry, the Hebrew University, Jerusalem, 91904, Israel

Received 22 July 2002; received in revised form 20 August 2002; accepted 21 August 2002

Abstract

The evolution of concepts developed in the study of the hydrophobic affect is surveyed, within the more general context of solvent-induced effects. A systematic analysis of the solvent-induced contribution to the driving force for the process of protein folding has led to two important modifications in our understanding of these effects. First, the conventional concepts of hydrophobic *solvation* and hydrophobic *interactions* had to be replaced by their respective *conditional* effects. Second, each of the hydrophobic effects has also a corresponding hydrophilic counterpart. Some of the latter effects could contribute significantly to the total driving force for the process of protein folding, and perhaps even dominate the driving force for biochemical processes.

Keywords: Hydrophobic; Hydrophilic; Protein folding

1. Introduction

I met Dr Kauzmann in one of the first Gordon conferences on Water in the late 1960s. It was during our long walks through the woods of New Hampshire that I first learned about hydrophobic interactions—or rather—hydrophobic bonds, as referred to in those days. It was Dr Kauzmann who suggested and encouraged me to write a book on the theory of liquid water [1]. In this book I dedicated the last chapter 1974 to review what was known at that time on the subject of hydrophobic interactions. After writing the book, I have continued to work on the problem of hydrophobic interactions for almost fifteen years, continually corresponding and discussing the subject with Dr Kauzmann. This period has ended with my writing of a book on hydrophobic interactions in 1980 [2].

In this paper I would like to review the evolution of the concepts associated with the so-called hydrophobic effect.

The term Hydrophobic bond was coined by Kauzmann in his 1959 review article [3] entitled: 'Some factors in the interpretation of protein denaturation'. During those years this concept has evolved and mutated into quite a few related concepts.

In the next section we shall survey four model-systems that were designed for studying specifically the hydrophobic phenomena. In Section 3, we turn to a general definition of the solvent-induced contribution to the free energy of any process taking place in a solvent. In Section 4, we shall discuss the result of an analysis of all possible

*Tel.: +972-2651-3742; fax: +972-265-85733.
E-mail address: arieh@batata.fh.huji.ac.il (A. Ben-Naim).

0301-4622/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0301-4622(03)00088-7

solvent-induced effects for the specific process of protein folding. Among these we shall see that Kauzmann's model does not feature, but a close relative of this model is identified. In addition, we find other solvent-induced effects, which may be classified either as hydrophobic or as hydrophilic effects. As we shall see, some of the latter could contribute significantly and perhaps even dominantly, to the driving force for biochemical processes.

2. Model-systems designed for studying solvent-induced contributions to the driving force of protein folding

In this section, we review some of the most basic models that have been used in the study of the hydrophobic effects.

2.1. Kauzmann's model

Kauzmann was obviously puzzled by the fact that some proteins after being denatured, either by increasing the temperature or by addition of a cosolvent, can refold back (or renature) into their original, active form. How does the protein 'know' to refold, from a highly random configuration, into a precise three-dimensional structure, in a surprisingly short time? Moreover, why does this highly-specific process occur only in aqueous media?

Posing the same question in physical-chemical terms: what is the source of the large negative free energy that drives the process of refolding? More specifically, what is the contribution of the water to this driving force?

The first serious attempt to answer these questions was undertaken by Kauzmann in 1959. Kauzmann reviewed all possible factors that might be contributing to the driving force for the process of refolding.

Following a comment made previously by Kirkwood in 1954 [4,5], on the possible role of the solvent in such processes, Kauzmann brought forward the idea of the hydrophobic bond, which, at that time, was conceived as the tendency of non-polar solutes to adhere to each other in aqueous media.

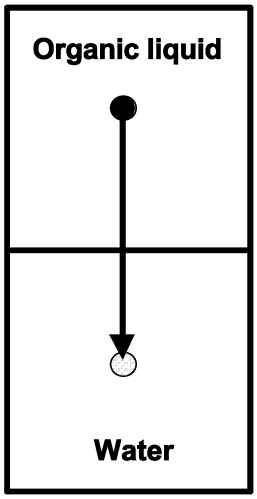

Fig. 1. A non-polar solute is transferred from an organic liquid into water.

The main reason underlying Kauzmann's model is the following:

It was known that the solubility of a non-polar solute is extremely low in water, compared with the solubility of the same solute in an organic liquid. This may be translated into free energies as follows: Consider the transfer of a simple solute, say argon or methane, from an organic liquid, say, ethanol or dioxane, into water. The process is schematically shown in Fig. 1. The very low tendency of the non-polar solute to dissolve in water, relative to an organic liquid, is equivalent to a large positive change in the free energy of the processes depicted in Fig. 1. Kauzmann noted that in the process of protein folding, many non-polar groups, attached to the backbone of the protein that are initially exposed to the solvent become buried in the interior of the protein in the folded form. These non-polar groups that are initially solvated by water, become 'solvated' by the environment of the interior of the protein, in the final folded form.

Hence, Kauzmann argued, the whole process of protein folding can be viewed as a reversal of the process depicted in Fig. 1. Thus, if each non-polar side chain that is transferred from water to the interior of the protein contributes a few kcal/mol to the total driving force, one can expect a very

large negative free energy change originating from all the non-polar side chains being transferred into the interior of the protein, in the folding process. This is essentially the Kauzmann's model for the hydrophobic effect. This model has dominated the biochemical literature for almost thirty years. It obviously provided a simple, though ill understood, explanation for the driving force for many biochemical processes including protein folding, protein–protein association and binding of small ligands to proteins.

In spite of the overwhelming acceptance of the Kauzmann's model, there remained two disturbing questions that could not be resolved with what was known at that time.

First, to what extent is Kauzmann's model appropriate for estimating the contribution of each non-polar side-chain to the total free energy of the process of protein folding? Second, how do we know that this particular effect is the most important contribution to the overall free energy change in the entire process of protein folding? In fact, the first question consists of two parts: one, concerning the choice of the appropriate solute-molecule to represent the non-polar group being transferred into the interior of the protein. The second concerns the choice of the appropriate organic liquid, which is supposed to represent the environment of the interior of the protein.

It should be noted that the solvation free energy is indeed a measure of the extent of affinity between the solute and the solvent. However, non-polar groups hung on the protein are not solvated solely by water. They are, by definition, attached to the backbone of the protein. Hence, they are only partially solvated by the solvent. This, in turn, has a significant effect on the solvation free energy of these groups.

These questions clearly could not be answered before a complete analysis of all possible solvent-induced effects, as well as detailed structures of both folded and unfolded states, became available. We shall return to discuss these questions in Section 4.

2.2. Hydrophobic interaction

The second model introduced and studied in the early 1970s is the pairwise interaction between two simple non-polar solutes in water [6,7]. This has required redefinition of the hydrophobic effect, from solvation of a single solute molecule to interactions between two or more solutes in water. The main idea behind this new model was simple. The transfer of a non-polar solute as depicted in Fig. 1, models the extreme process, where a non-polar group is transferred from an aqueous environment into an organic-liquid-like environment. However, in a real protein folding processes, as well as in many other biochemical processes, polar groups do not always lose *all* their solvation in water and acquire a new solvation environment. Instead, we often observe non-polar groups approaching each other in water, where only *part* of their solvation by water is eliminated. This case is illustrated in Fig. 2.

Fig. 2. Two non-polar molecules approaching each other on water.

The main question posed here is whether one can detect a strong tendency for the two solutes to adhere to each other in water, as compared with other solvent? To answer this question it was suggested [6,7] to study the Gibbs free energy of the system, $G(R)$, of two non-polar groups in water as a function of their relative distance R, as illustrated in Fig. 2. Knowing the function $G(R)$ would tell us all we need to know on this system. At that time there was no way to compute the function $G(R)$, the simulation techniques were in their infancy and the experimental data were very scarce and were not directly relevant to this problem. For instance, Scheraga and his group [8] studied the dimerization free energies of carboxylic acids in water. Kozak et al. [9] studied the second osmotic virial coefficient of some molecules in aqueous solutions, such as alcohol carboxylic acids. Although some qualitative information on hydrophobic interaction could be extracted from these data, these methods were obviously not adequate. The reason is that these types of molecules have both a hydrophobic group and a hydrophilic group. Hence, both the dimerization free

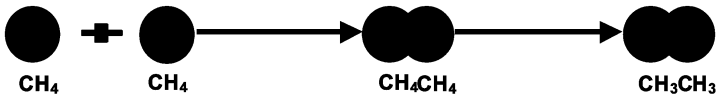

Fig. 3. Two methane molecules are brought from infinite separation to a very short distance, then replaced by one ethane molecule.

energy, and the virial coefficients include contributions from hydrophobic interactions as well as hydrophilic interactions; the separation between the two contributions is not possible (these studies were critically reviewed in Ref. [1]. There was clearly a need for a genuine measure of the extent of association ('dimerization') of two strictly non-polar solutes in water.

Information on this aspect of the problem started to accumulate, first from experimental sources [6], then from theoretical [10,11] and simulated calculations [12–14]. It should be noted that the function $G(R)$, i.e. the free energy of the system as a function of the distance between the two non-polar groups, could not be obtained experimentally. Since the solubility of non-polar solutes such as methane or argon is extremely low in water, one cannot expect to detect any measurable amount of 'dimers' in liquid water. However, simulation of aqueous solutions of simple non-polar molecules

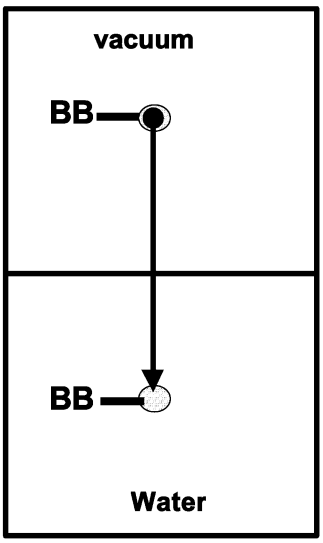

Fig. 4. A non-polar solute, attached to a backbone (BB), is transferred from vacuum into water.

became available. Once we have the function $G(R)$, one can simply define a 'dimer' as the two molecules (or atoms) at a distance R. Thus, the function $G(R)$ may be interpreted as the free energy of 'dimerization', in the sense of the free energy required to form a 'dimer' (i.e. the two monomers at a distance R), from the two monomers at infinite separation.

In spite of these difficulties it was possible to obtain a great deal of information on hydrophobic interactions between pairs, triplets, etc. of non-polar groups in water by replacing a pair of say methane molecules at very short separation, by one ethane molecule [6,7]. The trick is to 'fool', so to speak, the solvent into 'thinking' that the pair of methane molecules at very short distance is indistinguishable from a single ethane molecule. Thus, instead of the required, but unavailable free energy changes:

G(at R)-G(at infinity), one could use the experimental data on the solvation of methane and ethane, namely $\Delta G^*_{\text{ethane}} - 2\Delta G^*_{\text{methane}}$. This process is described schematically in Fig. 3.

2.3. Conditional solvation of a non-polar solute in water; a modified Kauzmann's model

As we mentioned above, Kauzmann's model for studying hydrophobic phenomena dealt with an extreme process: i.e. when the solute loses its solvation upon transfer from water into an organic liquid. However, when detailed analysis is made of all solvent effects on protein folding, one finds that this model is not adequate for two reasons. First, the non-polar groups in the unfolded form of the protein are not fully solvated by water. They are only partially exposed to the solvent. Second, the groups being transferred into the interior of the protein is not *solvated* by the protein. It does *interact* with other parts of the protein, but this interaction is already taken care of by the energy

change in the process, and is not part of the solvent-induced effect. What remains is the *conditional* solvation of a non-polar group. Fig. 4 shows a modified Kauzmann's model. The modified process is now used as a model to the study of the contribution of a single non-polar group being transferred from the unfolded form into the interior of the protein.

The two main differences between the processes of Fig. 1 and Fig. 4 are the following: First, the organic liquid solvating the group is replaced by vacuum. Secondly, the non-polar group is transferred not to pure liquid water, but to a point adjacent to the backbone of a protein. It should be noted that these two modifications to Kauzmann's model result from the general analysis of all the possible solvent-induced effects, details of which may be found elsewhere [7]. These new models have also changed significantly our estimate of the magnitude of the hydrophobic effect as predicted by Kauzmann's model.

The method of obtaining information on the conditional solvation free energies is quite different from the one used to obtain the transfer free energies. In the latter, the free energy change associated with the process of Fig. 1 is obtained from the following difference in solvation free energies: $\Delta G^*_{\text{solute}}$ (in water) $-\Delta G^*_{\text{solute}}$ (in organic liquid). On the other hand, the conditional solvation free-energy of a non-polar group attached to a backbone (Fig. 4), is obtained from the difference: $\Delta G^*_{\text{backbone with the group}} - \Delta G_{\text{backbone without the group}}$. Here, the solvation frees energies of the backbone with and without the group is both evaluated in water. The 'organic liquid' does not feature in this quantity. It has been found that these two quantities as defined above could differ by one or even two orders of magnitude [7].

2.4. Conditional pairwise hydrophobic interactions; intramolecular hydrophobic interactions

We have seen in the last paragraph that Kauzmann's model for solvation has been modified into a conditional free energy of solvation. Similarly, we need to modify the models for studying pairwise (or higher order) hydrophobic interactions. The reason for the modification is the same as

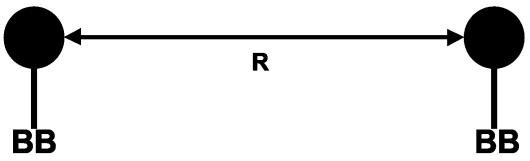

Fig. 5. Two non-polar molecules attached to a backbone (BB) approaching each other in water.

before. Two non-polar groups approaching each other in the folded form of the protein are not 'free' solute, as implied by the model studied in Fig. 2. Instead, we recognize the fact that the two groups are attached to a backbone in both the initial and the final state of the protein-folding process. Thus, instead of the model used in Fig. 2, we need to study the modified model depicted in Fig. 5. Here we follow the free energy of the system of two groups attached to a backbone as a function of their separation.

A great deal of information has been accumulated on the function $G(R)$ pertaining to Fig. 2, but none is known at present on the function $G(R)$ pertaining to Fig. 5.

Instead, limited information on the conditional hydrophobic interaction (previously referred to as intramolecular hydrophobic interaction) has been obtained for a specific distance and a specific backbone [6,7]. Two examples of such models [15,16] are shown in Fig. 6. The two models are designed to provide the same information. In the first process (I), we transfer a methyl group from the para (4) position to the ortho (2) position with respect to a methyl group already situated at position (1). Note that this process involves also changes in the internal energy of the molecule. However, when we do the same process once in vacuum and once in the liquid we can eliminate the changes in the internal energy, and we are left with solvent effect only on this process.

The quantity $\Delta G_{1,2} - \Delta G_{1,4}$ is thus a measure of the indirect or solvent-induced contribution to the change in the free energy of the process I in Fig. 6. If we further assume that the two methyl groups at the 1, 4 positions are far away ('infinity') from each other, i.e. if we can assume that the correlation between the two methyl groups at this distance

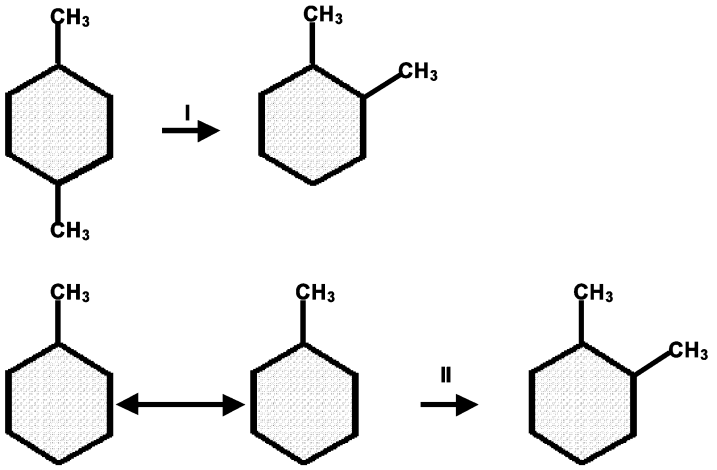

Fig. 6. Two models for studying conditional pairwise hydrophobic interactions (or intramolecular hydrophobic interactions).

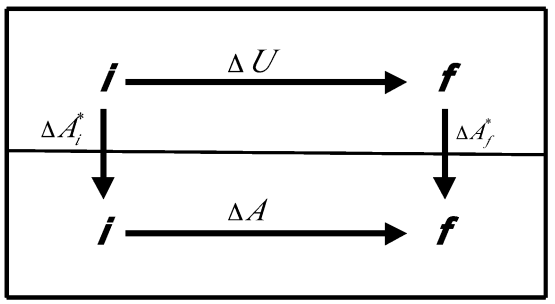

Fig. 7. Thermodynamic cycle shows the relations between the quantities in Eq. (3.7).

is negligible, then we can interpret the measurable quantity as the intramolecular hydrophobic interaction at the ortho position. This approximation is confirmed by the second model process depicted in Fig. 6 (process II), where now we transfer a methyl group from toluene (i.e. a singly methylated benzene ring), to another toluene molecule at position 2. The indirect part of the free energy change for this process is $\Delta G_{1,2}-2\Delta G_1$, Where ΔG_1 is the solvation free energy of the toluene molecule. The approximate values obtained for the two processes indicate that the correlation between the two methyl groups at the para (1,4) positions is indeed negligible.

2.5. Conditional higher-order hydrophobic interactions

So far, we have seen four different model systems that were used to study hydrophobic phenomena. The first two were ad-hoc models suggested long before a general analysis of all possible solvent-induced effects has been undertaken. Once this analysis has been carried out [7,17], it became clear that these ad-hoc model-systems should be modified to take into account the presence of the backbone. This has a significant effect on the solvation of the attached side chains. Clearly one cannot stop at pairwise hydrophobic interactions. Indeed, several studies of hydrophobic interactions among three or more solutes have been carried out [2]. The extreme process of these is the formation of a droplet consisting of many nonpolar solute particles. So far no study of the corresponding higher-order *conditional* hydrophobic interaction has been undertaken.

With this comment we have exhausted all possible model-processes that are pertinent to the study of hydrophobic phenomena. We stress again that the study of the *conditional* processes was initiated only after an inventory of all possible solvent-induced effects has been established [17]. It was quite surprising and unexpected to find that the conditional quantities are quite different from

the unconditional quantities, sometimes the difference amount to one or two order of magnitude. What was even more unexpected is the finding that *hydrophilic* effects, both solvation and interactions seem to be far larger than the corresponding values of the hydrophobic effects. This of course has a significant impact on our efforts to understand the driving forces underlying the highly-specific biochemical processes in aqueous solutions. We shall quote some numerical figures in Section 4.

3. General definition of solvent-induced effects

Consider any process taking place in a solvent, which symbolically is written as:

$$R \rightarrow P \tag{3.1}$$

Where R represents all the reactants, and P represents all the products. This process could be a simple isomerization process, a dimerization of two monomers to form a dimer, or any more complex process. We shall refer to the L.H.S. of the equation as the initial state and to the R.H.S. as the final state of the process. Together R and P are presumed to represent all the solutes involved in the process Eq. (3.1).

When the reaction Eq. (3.1) takes place in vacuum, the total potential energy of the system in the initial and the final states are written as $U_n(R_1,\ldots,R_n)$ and $U_n(P_1,\ldots,P_n)$, respectively. Here $R_1,\ldots,R_n$ is the configuration of all solute molecules in the *initial* state, where R_i is the location vector of the i^{th} atom of the solutes involved in the process. Similarly P_i is used to denote the location vector of the i^{th} atom of the solutes in the *final* state. On the other hand, when the same process takes place in a solvent we write the total potential energy of the system as:

$$U(X^N,R^n)=U_N(X_1,\ldots,X_N)+U_n(R_1,\ldots,R_n)+U(X_1,\ldots,X_N|R_1,\ldots,R_n) \tag{3.2}$$

$$U(X^N,P^n)=U_N(X_1,\ldots,X_N)+U_n(P_1,\ldots,P_n)+U(X_1,\ldots,X_N|P_1,\ldots,P_n) \tag{3.3}$$

For the initial state, Eq. (3.2) and for the final state, Eq. (3.3), respectively. Here $U_N(X_1,\ldots,X_N)$ is the total interaction energy between all N solvent molecules, at a specified configuration $X_1,\ldots,X_N$, X_i being the location and orientation vector of the ith solvent molecule, and $U(X_1,\ldots,X_N|R_1,\ldots,R_n)$ is the total interaction energy between all solute atoms and all solvent molecules. Note that the configuration of all the solute molecules is specified by the *locations* of all solute *atoms*, whereas for the solvent configuration we specify the *location* and *orientation* of all solvent molecules. Normally one also assumes that the total potential energy is pairwise additive. However, for the analysis we describe in the next section, we shall need a weaker assumption. Namely, that only the solute-solvent terms in Eq. (3.2) and Eq. (3.3) are pairwise additive, i.e. we write:

$$U(X_1,\ldots,X_N|R_1,\ldots,R_n)=\sum_{i-1}^{n}\sum_{j-1}^{N}U(X_j,R_i) \tag{3.4}$$

Where $U(X_j,R_i)$ is the interaction energy between the j^{th} solvent molecule and the i^{th} atom of the protein, at the specified configuration. A similar expression applies to the interaction energy between the solvent molecules and the solute atom at the final state.

When applying Eq. (3.1) to the protein folding process, as we shall do in the next section, we always assume that the protein is very dilute in the solvent. Hence, no solute–solute interactions are taken into account. In this case, the potential energy term, $U_n(R_1,\ldots,R_n)$ includes only direct interactions between the atoms of a single protein molecule.

The Helmholtz energy change associated with the process Eq. (3.1) is obtained from the ratio of the corresponding canonical partition functions, i.e.:

$$\exp[-\beta\Delta A]=\exp[-\beta(A_f-A_i)]=\frac{Q_f}{Q_i}=\frac{\int\exp[-\beta U(X^N,P^n)]\mathrm{d}X^N}{\int\exp[-\beta U(X^N,R^n)]\mathrm{d}X^N} \tag{3.5}$$

Here Q_i and Q_f are the partition functions for the initial and final states, respectively, the integrations in both the numerator and the denominator are extended over all the configurations of the solvent molecules only. The coordinates of all the atoms belonging to the solute molecules are presumed to be fixed at the initial and final states.

In Eq. (3.5), $\beta=(k_BT)^{-1}$, where k_B is the Boltzmann constant, and T is the absolute temperature.

Factoring $U_n(P_1,\ldots,P_n)$ from the numerator and $U_n(R_1,\ldots,R_n)$ from the denominator we can rewrite Eq. (3.5) as:

$$\exp[-\beta\Delta A]=\exp[-\beta\Delta U]\exp[-\beta(\Delta A_f^*-\Delta A_i^*)] \tag{3.6}$$

where ΔU is the change in energy associated with the process Eq. (3.1), and ΔA_f^* and ΔA_i^* are the solvation Helmholtz energies of the final and initial states, respectively. We can now define the solvent-induced contribution to the total Helmholtz energy of the process by:

$$\delta A=\Delta A-\Delta U=\Delta A_f^*-\Delta A_i^* \tag{3.7}$$

Thus, for any process the solvent-induced contribution to the Helmholz (or the Gibbs) energy is defined by the difference between the total Helmholtz (or Gibbs) energy and the corresponding energy change for the same process (for more details on the definition of the solvation Helmholtz energy and the relations Eq. (3.5) and Eq. (3.6), the reader is referred to Ref. [7]). This in turn is equal to the difference in the solvation Helmholtz (or Gibbs) energies of the solutes in the final and in the initial states. The relation between all the quantities in Eq. (3.7) is schematically shown in Fig. 7.

It should be noted that each of the two contributions to the total free energy change, as represented in Eq. (3.7), is a very complicated expression, for a process such as protein folding. However, the solvent contribution is vastly more complicated than the direct energy term. The reason is simple. The term ΔU on the R.H.S. of Eq. (3.7) includes all the interactions among all atoms belonging to all solutes that are involved in the process. Note, that this term should not be confused with the thermodynamic change in the internal energy of the system; this normally involves averages over all configurations of the solvent molecules, as well as over all possible conformations of the protein molecule. For a typical protein of say, 150 amino acid residues, this term includes quite a large number of contribution due to all interactions between the atoms of the protein. The term, defined in Eq. (3.7) is far more complicated, since it involves averages over all configurations of the solvent molecules these in turn depend on all the interactions between all the solute–solvent molecules as well as solvent–solvent molecules. Clearly, there is no obvious way of looking 'into the detailed content' of this term, let alone estimating the order of magnitude of these ingredients.

In the next section, we shall outline one possible way of dissecting the solvent-induced contribution to the total free energy change into smaller more manageable ingredients. This has led to two important consequences. First, we have obtained an inventory of all possible ingredients of the solvent-induced effects. Second, the theoretical analysis has indicated a way of studying each of these ingredients by either theoretical or experimental means. This, in turn, has also led to estimating some of the order of magnitudes of the different ingredients, the most unexpected result has been the finding that various *hydrophilic* effects might be playing a more important role in the process of protein folding than the corresponding hydrophobic effects.

4. Dissecting the solvent-induced contribution into its ingredients: the emergence of the hydrophilic phenomena

In this section, we briefly review the procedure for dissecting the quantity δG into its ingredients. Details of this procedure have been described elsewhere [17], and shall not be repeated here. In Section 3, we specified the configuration of the solutes by the locational vectors of all atoms of the solutes. In this section, we treat the case of protein folding process. Here, we choose a more convenient characterization of the configuration of

the protein. Instead of specifying the locations of all atoms, we specify the locations and perhaps also the orientations of all *groups* of atoms such as methyl methylene, hydroxyl, etc. This level of characterization of the protein configuration is more convenient, since it leads us naturally to what has been traditionally referred to as hydrophobic and hydrophilic phenomena.

The general reaction in Eq. (3.1) will now be specified for the protein folding process, namely,

$$\text{Unfolded form} \rightarrow \text{Folded form} \tag{4.1}$$

The total solvent-induced effect is given by:

$$\delta G = \Delta G_F^* - \Delta G_U^* \tag{4.2}$$

where, on the R.H.S. of Eq. (4.2) we have the difference in the solvation Gibbs energy of the protein in the folded and unfolded forms. It should be noted that for our particular analysis, we take one specific conformation of each form, say the fully extended all transforms for the unfolded protein and the exact three-dimensional structure of the protein in the folded form, as determined by X-ray diffraction.

It is clear from Eq. (4.2) that the total solvent effect is determined by the change of the solvation Gibbs free energy of the entire protein in the process of folding Eq. (3.1). This is an immensely complicated expression. Not only because it involves the changes in the coordinates of all atoms of the protein, but also because it includes an average overall configurations of all solvent molecules. Obviously, at this level of description it is impossible to carry out any analysis of the solvent contributions to the driving force. The other extreme and the most detailed description would have been achieved by following the 'fate' of each atom of the protein when the process Eq. (3.1) is carried out (this has been the reason for the description of the protein configurations by their atomic coordinates in the previous section). This may be referred to as the atomic level of analysis. Clearly, such a detailed description will not lead to a feasible analysis. A more traditional and more effective level of description would be the level of groups of atoms, such as methyl, ethyl, hydroxyl, etc. This intermediary level will lead us to identify the traditional hydrophobic and hydrophilic groups, hence, to the corresponding solvent effects.

Our task is now to follow the 'fate' of each group upon the performance of the process Eq. (3.1). We first make a distinction between three broad classes of groups. In each class the groups are characterized according to the extent of the change in their solvation in the process Eq. (4.1). The first class consists of those groups that are initially exposed to the solvent and become buried in the interior of the protein. These are the groups, the solvation of which is completely lost upon folding. The second class contains all the groups, the solvation of which is unchanged in the folding process. Finally, the third class contains all the remaining groups, i.e. all groups, the solvation of which is changed to some intermediate extent between the first and the second class.

What is the contribution of each group to the total solvent induces driving force?

Clearly, the answer depends on both the type of the group and on the class to which it belongs. We present here a brief description of the types of solvent-induced contribution, which together constitutes the inventory of solvent-induced effects.

4.1. Groups belonging to the first class

Each group belonging to the first class will contribute its *conditional* solvation free energy. Since such a group is initially solvated by water, and it loses it solvation upon folding, its contribution will be the loss of the conditional solvation Gibbs energy. It is conditional since this group is attached to the backbone of the protein in both its initial and its final state. It should be stressed, however, that the solvation-in-the-organic-liquid does not feature in this contribution. The reason is that in the final state the groups under discussion are surrounded by other groups of the protein. Hence, the interaction energy of this group with its surroundings belongs to the term ΔU, in Eq. (3.7) and not to the solvent effects δG. It has been found that the typically conditional solvation Gibbs energy for a hydrophilic group is in the order of approximately -6 to -7 kcal/mol.

Whereas, the corresponding quantity for a hydrophobic group is approximately $+0.3$ to $+0.6$ kcal/mol [7]. The latter quantity was found to be strongly dependent on the type of backbone, say, saturated or unsaturated backbone, and could vary significantly, sometimes a few orders of magnitude when the type of the backbone is changed.

4.2. Groups belonging to the second class

Groups in the second class do not experience any change in their solvation. Hence, their contribution to δG will be zero. This can formally be proven, but it is also intuitively quite clear.

4.3. Groups belonging to the third class

This class of groups is the richest in its content. The variety of effects results from the fact that there are different types of groups as well as different extent of changes incurred to the group as a result of the said process. We present here a schematic list of possible solvent effects in this class:

a. Pair correlations between two hydrophobic groups.
b. Pair correlations between two hydrophilic groups.
c. Pair correlations between a hydrophobic and a hydrophilic group.

The above list for *pairs* should be repeated for triplets, quadruplets, etc. then repeated for all possible combinations of the types of groups, and then repeated over all possible configurations of the correlating groups. Clearly, the list becomes very long, indeed there is a plethora of solvent-induced effects. It should be noted that all of these correlations are conditional correlations, i.e. the groups are always attached to the protein backbone, as in the examples shown in figures. At present, we have only very limited information on the magnitudes of some representative cases only. For instance the contribution due to pair correlation between two methyl groups attached to a benzene ring at the positions 1,2, is approximately -0.30 kcal/mol, for two hydroxyl groups at positions 1,3 it has been estimated to be approximately -3.5 kcal/mol [7].

It should be pointed out that in this article the terms hydrophilic and hydrophobic are understood to be part of the solvent-induced effect to the driving force. Direct intramolecular hydrogen-bonding are included in the direct energy change ΔU. Recently, some doubts were raised regarding the dominance of the hydrophobic effect in protein folding, as expressed by a recent review by Dill [18]. For instance, Pace et al. [19] reached the conclusion that 'hydrogen bonding and hydrophobic effects make large comparable contributions to the stability of globular proteins'. Similar conclusions were reached by Makhatadze and Privalov [20] and by Cooper [21]. All these authors refer to the direct intramolecular hydrogen-bonding, which in our nomenclature is a part of the energy change ΔU, and not of δG. Our conclusion, regarding the possible dominance of the hydrophilic effects, such as loss of solvation, or correlations mediated by the solvent, do not include *direct* hydrogen bonding.

In concluding this section we note that the dissection of the total solvent-induced effect to obtain an inventory of all the ingredients was achieved by theoretical arguments. Theory also has provided us with some hints of how to study each of these ingredients by experimental means. It is only when one has acquired a complete inventory of all possible solvent-induced effects, that one can turn to explore their relative magnitudes and hence their relative importance in building up the total driving force for the process of protein folding. Clearly, much work remains to be done on both the experimental and theoretical ground

5. Conclusion

Over forty years ago, when Kauzmann first put forward the hydrophobic effect hypothesis, the driving forces directing the protein to fold into a precise three-dimensional structure, were quite a mystery. In spite of the tremendous effort expended to unravel the sources of this remarkable process, much of the mystery is still with us. Kauzmann's courageous and ingenuous hypothesis did not solve the mystery. Its main impact was to produce a

strong impetus to research on a whole range of topics, from the structure of pure water to aqueous solution of highly complex biomolecules. Kauzmann's ideas did pave the way for an extended study of the role of water in complex process, such as protein folding. As we have seen in the previous section, Kauzmann's original model obviously was not adequate for the study of all the driving forces for protein folding. In fact, the model as it stands cannot account for the contribution of hydrophobic effect in the context of the study of protein folding. As we have seen in Section 3 the hydrophobic effects should be studied through the corresponding conditional solvation. This simply means that we recognize the fact that the hydrophobic groups are attached to the protein-backbone, and are not free to wander in the solution. Furthermore, once we study all possible solvent-induced effects, it becomes clearer that hydrophilic effects are more important than the corresponding hydrophobic effects. For me, personally, Kauzmann's ideas were always a source of excitement and encouragement to study the structure of water, the properties of simple aqueous solutions, and ultimately complex biological fluids and the processed occurring therein.

References

[1] A. Ben-Naim, Water and Aqueous Solutions, Plenum Press, New York, 1974.

[2] A. Ben-Naim, Hydrophobic interactions, Plenum Press, New York, 1980.

[3] W. Kauzmann, Some factors in the interpretation of protein denaturation, in: C.B. Anfinsen, M.L. Anson, K. Bailey, J.T. Edsall (Eds.), Advances In Protein Chemistry, XIV, Academic Press, New York, 1959, pp. 1–63.

[4] W. Kauzmann, Denaturation of proteins and enzymes, in: W.D. McElroy, B. Glass (Eds.), A Symposium On the Mechanism of Enzyme Action, Johns Hopkins University Press, Baltimore MD, 1954, pp. 70–120.

[5] J.G. Kirkwood, The nature of the forces between protein molecules in solution, in: W.D. McElroy, B. Glass (Eds.), A Symposium on the Mechanism of Enzyme Action, Johns Hopkins University Press, Baltimore MD, 1954, pp. 4–23.

[6] A. Ben-Naim, Statistical mechanics study of hydrophobic interactions. I. Interaction between two identical non-polar solute molecules, J. Chem. Phys. 54 (1971) 1387–1404.

[7] A. Ben-Naim, Statistical Thermodynamics for Chemists and Biochemists, Plenum Press, New York, 1992.

[8] E.E Schrier, M. Pottle, H.A. Scheraga, The influence of hydrogen and hydrophobic bonds on the stability of the carboxylic acid dimers in aqueous solution, J. Am. Chem. Soc. 86 (1962) 3444–3449.

[9] J.J. Kozak, W.S. Knight, W. Kauzmann, Solute-solute interactions in aqueous solutions, J. Chem. Phys. 48 (2) (1968) 675–690.

[10] L.R. Pratt, D. Chandler, Theory of the hydrophobic effect, J. Chem. Phys. 67 (8) (1977) 3683–3704.

[11] L.R. Pratt, D. Chandler, Effects of solute-solvent attractive forces on hydrophobic correlations, J. Chem. Phys. 73 (7) (1980) 3434–3441.

[12] A. Geiger, A. Rahman, F.H. Stillinger, Molecular dynamic studies of the hydration of Lennard-Jones solutes, J. Chem. Phys. 70 (1) (1979) 263–276.

[13] M. Pangali, M. Rao, B.J. Berne, Hydrophobic hydration around a pair of apolar species in water, J. Chem. Phys. 71 (7) (1979) 2982–2990.

[14] G. Ravishanker, M. Mezei, D.L. Beveridge, Monte Carlo computer simulation study of the hydrophobic effect: Potential of mean force for $[(CH_4)_2]_{aq}$ at 25 and 50 °C, Faraday Symp. Chem. Soc. 17 (1982) 79–91.

[15] A. Ben-Naim, J. Wilf, A direct measurement of intramolecular hydrophobic interactions, J. Chem. Phys. 70 (1979) 771–777.

[16] J. Wilf, A. Ben-Naim, Intromolecular hydrophobic interaction in light and heavy water, J. Chem. Phys. 70 (1979) 3079–3081.

[17] A. Ben-Naim, Solvent effects on protein association and protein folding, Polymers 29 (1990) 567–596.

[18] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133–7155.

[19] C.N. Pace, B.N. Shirley, M. McNutt, K. Gajiwala, Forces contributing to the conformational stability of proteins, FASEB 10 (1996) 75–83.

[20] G.I. Makhatadze, P.L. Privalov, Hydration effects in protein unfolding, Biophys. Chem. 51 (1994) 291–309.

[21] A. Cooper, Heat capacity of hydrogen-bonding, hydrophobicity, packing and folding thermodynamics, Biophys. Chem. 85 (2000) 25–39.