Enthalpies of Transfer for Proteins from Aqueous to Water-Alcohol Solutions: A Test for Models of Residues Exposed to Solvent

M. Y. SCHRIER, G. DEFOSTER, E. E. SCHRIER and M. STARZAK, *Chemistry Department, State University of New York at Binghamton, Binghamton, New York 13901*

Synopsis

The enthalpies of transfer of proteins from aqueous solution to alcohol–water solutions are used as probes of solvent-accessible surface for these proteins. Enthalpies of transfer to 10 wt% ethanol solutions are determined calorimetrically for the native proteins ribonuclease A, lysozyme, and ovalbumin. Ribonuclease A and lysozyme are reduced and carboxamidomethylated to produce configurations in which interior residues of the native protein are exposed to the solvent; enthalpies of transfer are determined for these species. These data are then compared with enthalpies of transfer for the constituent amino acids of the proteins. The enthalpies of transfer for the residues are used to generate a maximal enthalpy of transfer that can be compared with the enthalpies of transfer for the reduced, carboxamidomethylated proteins. The residue amino acid enthalpies are coupled with probabilities that each residue is an exterior residue to predict an enthalpy of transfer for the native protein that can be compared with the observed enthalpy. The probabilities developed by Wertz and Scheraga and Lee and Richards, and Chothia are then compared on their ability to predict the native enthalpies of transfer for the protein. The Wertz–Scheraga model gives the better fit of this data in all cases.

INTRODUCTION

The interactions between globular proteins and solvent play a central role in establishing the native structure of the protein.[1] For a given primary structure, the protein is expected to seek a conformation in which polar amino acids or atoms appear at the surface of the globular protein moiety, while the nonpolar amino acids or atoms seek the protein interior. Once the total structure is established in this manner, many of the solution properties of the protein will be dictated by amino acids or atoms on the outer surface.

Even though x-ray crystallographic structures are known for many proteins, these structures do not provide an exact picture of the protein surface in solution. However, these structures can be used to generate such surfaces which are called "accessible surfaces."

Lee and Richards[2] generated an estimate of accessible surface area from the crystal structure of a number of proteins. The surface area

Biopolymers, Vol. 25, 119–134 (1986)

CCC 0006-3525/86/010119-16$04.00

was determined by the accessibility of each region of the protein to a sphere with radius equal to the solvent molecule radius. The sphere was rolled over the x-ray surface to establish regions of the surface that could be reached by this sphere. Shrake and Rupley[3] modified the Lee–Richards calculations to include van der Waals radii for side chains. Their results were consistent with those of Lee and Richards. Since the models focused on atoms located on the accessible surface, both authors noted the presence of a large fraction of nonpolar atoms at the surface. Some of these atoms were associated with polar residues.[3]

Chothia[4] used the Lee–Richards[2] approach to establish probabilities that each amino acid in the protein would be found at the accessible surface. These probabilities were calculated by comparing the observed surface area in the protein with the surface area for the amino acid when it was present as a tripeptide of the form Gly-AA-Gly, where AA is the amino acid residue of interest. For each amino acid residue, the number that were 100 and 95% surface inaccessible could be determined. Averages were computed from such information for 12 different proteins. Thus, for example, a total of 176 leucine residues in the 12 proteins yielded 28 residues that were completely surface inaccessible and 80 residues that were 95% inaccessible. The probability of finding leucine on the surface was

$$p_o = 1 - p_i = 1 - 80/176 = 0.545$$

Using this information, surface-dependent properties of an arbitrary protein could be determined using these probabilities and the numbers of residues in the protein.

Wertz and Scheraga[5] approached the problem of accessible protein surface area in an alternate way. They defined the outer surface of the protein using three sets of orthogonal probe lines. As each probe line intersected the protein surface, the residue of contact was recorded. The number of times a given residue was the first or last residue encountered by a probe line determined its classification as a surface or interior residue. From this information, they determined probabilities that a given residue was on the outside. A total of 20 different proteins was used to establish these probabilities.

The Lee–Richards–Chothia probabilities (LRC) and the Wertz–Scheraga (WS) probabilities differ. To test the reliability of either set, some protein property that depends on the surface residues must be examined in the framework of these probabilities. Enthalpies of transfer for a protein can be determined by observing integral enthalpies of solution for the protein in each of the solvent systems. The enthalpy of transfer is then determined as the difference in these two enthalpies of solution. For example, Almog et al.[6] determined enthalpies of transfer for ribonuclease A from dilute buffer to aqueous solutions con-

taining guanidinium chloride, urea, calcium chloride, lithium chloride, and sodium chloride over a concentration range. These data, in conjunction with data for the transfer of model compounds, are used to estimate the number of residues exposed in a denaturation process.

In this work, enthalpies of transfer for proteins from buffered aqueous solutions to solutions containing alcohols are used to probe the accessible surface of the protein. The presence of alcohols as solvent molecules is expected to change the solvent structure about the protein while introducing molecules that can have a different accessibility to the surface. Such effects will appear as a change in the observed enthalpy of solution. The differential change is the enthalpy of transfer. Since this enthalpy is dictated by interactions between the surface residues of the protein and the solvent, the experimental enthalpy of transfer can be compared with enthalpies of transfer for the separate residues and the probabilities that these residues appear on the surface. Thus, protein enthalpies of transfer, in conjunction with residue enthalpies of transfer, provide an experimental test of the two accessible surface models.

The effectiveness of such measurements requires that the structure of the protein in solution remain the same for each of the solvent systems, since the residue probabilities are based on the native structure of the proteins. The alcohols methanol, ethanol, and *n*-propanol induce a change in the protein transition temperature.[7] This change is linear in the concentration in the 0–20 wt % range for ethanol and methanol, while propanol exhibits a slight curvature at larger concentrations. Brandts and Hunt[8] studied the denaturation of ribonuclease in ethanol solutions. They noted maxima or minima in the thermodynamic parameters for denaturation in the 10–15 wt % range using free energies determined from spectroscopic measurements. Both sets of data suggest that there are no major changes in the native structure, i.e., denaturation, to at least 10 wt % alcohol.

The effectiveness of the experiments also depends on determinations of enthalpies of transfer for model amino acid residues. This requires the assumption that the enthalpies of transfer for all residues can be added to find the total enthalpy of transfer. In addition, it requires that enthalpies of transfer for different portions of the residue can be summed to give the total enthalpy of transfer for this residue. Thus, the enthalpies of transfer for N-acetyl amino acids are the summation of the enthalpy for the amino acid and the N-acetyl group. The enthalpy of transfer for the N-acetyl group can then be determined and subtracted from the enthalpy of transfer for the model compound. The assumption implicit in this approach have been discussed by Stimson and Schrier.[9]

This paper will consider the relationship between enthalpies of transfer for protein and enthalpies of transfer for their amino acid residues in two ways. Some proteins will be reduced to permit full

exposure of the amino acid residues of the protein. The enthalpy of transfer observed for the reduced protein can then be compared with the summation of enthalpies of transfer for the constituent amino acids. The summation is expected to describe the limit of full exposure of the residues. This summation can then be compared with the actual enthalpy of transfer for the reduced protein.

The enthalpy of the native protein can be predicted from enthalpies of transfer for the constituent amino acids if the probabilities that each residue appears on the surface are known, i.e., the enthalpy of transfer of each residue weighted by the probability that it is a surface residue gives the enthalpy of transfer for the native protein. In this case, two distinct types of model for the determination of the accessible surface area give two sets of probabilities. By using each of the probability sets in conjunction with the residue amino acid enthalpies of transfer, the two models can be compared. Enthalpies of transfer for ribonuclease A and its constituent amino acids are presented. The reduced carboxamidomethylated ribonuclease A enthalpy of transfer is compared with the result calculated by summing enthalpies of transfer for the residues. The ΔH_{tr} for the native protein is compared with the ΔH_{tr} calculated using residue ΔH_{tr} data and model probabilities. Finally, the proteins lysozyme and ovalbumin are analyzed and compared with the residue amino acid data and the model probabilities.

EXPERIMENTAL

Ribonuclease A and ovalbumin were purchased from Sigma Chemical Company, while salt-free lysozyme was purchased from the Worthington Biochemical Corporation. All were used without additional purification.

The reduction and carboxamidomethylation procedure[10] was adapted from the method of White.[11] The procedure was repeated to insure maximal reduction.

The purity of the reduced material was checked with discontinuous electrophoresis on polyacrymalide gels. Reduced and carboxamidomethylated (RCAM) lysozyme, for example, yielded a single band for four aliquots with R_f values between 0.26 and 0.34. This compares with R_f values between 0.52 and 0.61 for the unreduced lysozyme.

Enthalpies of solution were determined for solutions that contained $0.15M$ KCl to minimize salt effects. The solutions for ribonuclease A and lysozyme were maintained at pH 2. Although lysozyme has a stable native structure at this pH, the RCAM modification dissolved with difficulty. Ovalbumin was dissolved in solutions of pH 7.80 with $0.15M$ KCl and $0.025M$ Tris. Since residue amino acid enthalpies used solutions of pH 2.00 and KCl concentrations of $0.15M$, an exact correlation between native ovalbumin data and residue data may contain errors associated with enthalpy of transfer differences at the two different pH values.

Water for all experiments was glass-distilled twice. Ethanol (95%) was obtained from U.S. Industrial Chemicals Co. The methanol and *n*-propanol from Fisher Chemical were dried over molecular sieves, stored under nitrogen, and filtered through nucleopore filters. KCl was reagent grade (J.T. Baker Chemical Company) and tris-(hydroxymethyl) amino methane (Tris) was reagent-grade Sigma trizma base. The model amino acid residues were obtained from Sigma Chemical Company, Vega Biochemical Company, Mann Research Laboratories, and Cyclo Chemical Company.

The LKB 8700 calorimeter was calibrated and samples added to the calorimeter ampules were corrected for water accumulation during transfer using a linear extrapolation technique.[10]

RESULTS

Ribonuclease A

Although the alcohols produce a regular variation in the transition temperature in the 0–20 wt % range,[7] enthalpies of transfer were determined in the 0–20 wt % range to verify that this experimental variable also varied regularly in this range. Enthalpies of transfer to solutions containing ethanol of different weight percentages from aqueous solutions were measured for this verification. The variation of ΔH_{tr} with weight percent ethanol is shown in Fig. 1. The enthalpy of transfer increases monotonically over the entire range.

The appearance of a positive enthalpy of transfer that increased with increasing ethanol weight percent is consistent with the notion that the native protein tertiary structure is optimized for an aqueous environment. The ethanol may alter the structure of water around the protein or it may simply limit the water in contact with the surface

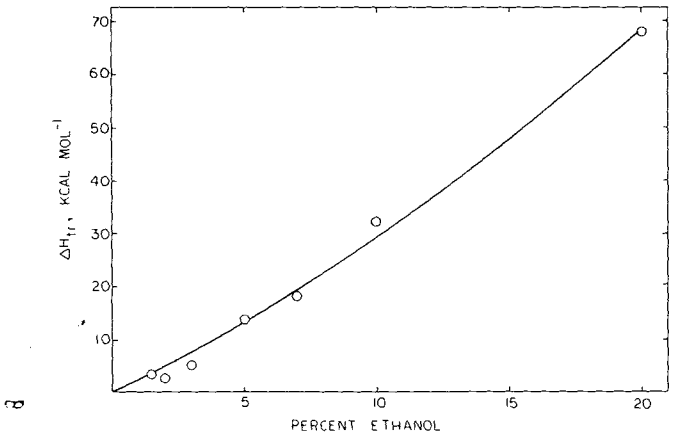

Fig. 1. Dependence of ΔH_{tr} for ribonuclease A on ethanol concentration.

at a given time. In either case, the system is less energetically stable in the ethanol solutions.

There are no major changes in the plot of ΔH_{tr} versus ethanol weight percent. This suggests that the protein has undergone no major structural change that might alter its surface area. Such changes have been suggested by spectroscopic measurements in the regime above 10 wt % alcohol.[8] However, such changes are not observed with these calorimetric measurements.

With these data, it is impossible to determine whether the change in ΔH_{tr} is due entirely to the effects of the solvent on the accessible protein surface. The structure of the protein might also change linearly with the ethanol weight percent so that the surface area increased in some regular way. Since proteins generally tend to exhibit sharp transitions between distinct stable configurations, the possibility of contributions from graded changes in the accessible surface area of the protein in these experiments is assumed to be insignificant.

Model Amino Acid Residues

Enthalpies of transfer for each of the amino acids that comprise the ribonuclease A are determined to provide some measure of the enthalpy of transfer for the protein if it existed in a configuration in which each residue was accessible to the solvent. Model residues, the N-acetyl amino acids or the amides of these acids, were used to determine these enthalpies of transfer from buffered aqueous solution to buffered 10 wt % solutions. The results are tabulated in Table I. The ΔH_{tr} are corrected for ΔH_{tr} for the N-acetyl and amide groups.

The model amino acid data provide some information on the potential location (interior or exterior) of each amino acid in a native protein. For example, ΔH_{tr} increases linearly with side-chain length as shown in Fig. 2. The longer side chains are hydrophobic and are expected to be less stable in the more polar aqueous solutions. The increase in ΔH_{tr} indicates increased hydrophobic bonding in the alcohol–water solutions relative to the aqueous solutions. This is consistent with the observations of Schrier et al.,[7] who observed a decrease in transition temperature with increasing alcohol weight percent. They attributed this change to increased hydrophobic binding between the side chains. In the case of the model residues, the hydrophobic binding is intermolecular rather than intramolecular.

Although the residues studied can be classified in terms of other intrinsic characteristics in addition to chain length, there is no apparent correlation between the observed ΔH_{tr} and the different classes of amino acids. In Fig. 3, amino acids are classified as polar, basic, acidic, acidic and amide, sulfur-containing, and ring-containing. ΔH_{tr} varies significantly for residues within each of these categories.

The experimental observations of ΔH_{tr} for the amino acid residues indicate that side-chain length is the major factor in directing the residue to the interior or exterior of the native protein structure.

TABLE I
Enthalpies of Transfer for Model Compounds and Amino Acid Residues from H_2O to 10% Ethanol at 25°C

Amino Acid Residue	Model Compound	ΔH_{tr}(model) cal mol^{-1}	ΔH_{tr}(corr) cal mol^{-1}	ΔH_{tr}(residue) cal mol^{-1}
Ala	Acetyl· · ·—NH_2	918±35	505±6	413±41
Arg	Acetyl· · ·—OH	764±9	834±7	−70±16
Asn	Acetyl· · ·—OH	1127±14	834±7	293±21
Asp	Acetyl· · ·—OH	1292±11	834±7	458±18
½ Cys	Acetyl· · ·—OH	1290±20	834±7	456±27
Gln	Acetyl· · ·—OH	1230±2	834±7	396±9
Glu	Acetyl· · ·—OH	1430±8	834±7	596±15
Gly	Acetyl· · ·—OH	958±17	834±7	124±24
His	Acetyl· · ·—OH	1285±4	834±7	451±11
Ile	Acetyl· · ·—NH_2	1093±17	505±6	588±23
Leu	Acetyl· · ·—NH_2	1424±7	505±6	919±13
Lys	Acetyl· · ·—OH	1148±3	834±7	314±10
Met	Acetyl· · ·—NH_2	1451±18	505±6	946±24
Phe	Acetyl· · ·—NH_2	1750±11	505±6	1245±27
Pro	Acetyl· · ·—NH_2	1098±5	505±6	593±11
Ser	Acetyl· · ·—NH_2	825±7	505±6	320±13
Thr	L-Amino acid	543±9	87±39[a]	456 ±48
Tyr	Acetyl· · ·—NH_2	1473±42	505±6	968±48
Val	Acetyl· · ·—NH_2	1348±11	505±6	843±17

[a] This correction factor is derived from ΔH_{tr} values for glycine, *N*-acetylglycine, and *N*-acetyldiglycine.

Model Predictions for Ribonuclease

The data for ribonuclease A, RCAM ribonuclease A, and the model amino acid residues can now be combined to develop several aspects of surface accessibility. First, the enthalpy of transfer for the RCAM ribonuclease A can be compared with an enthalpy of transfer calcu-

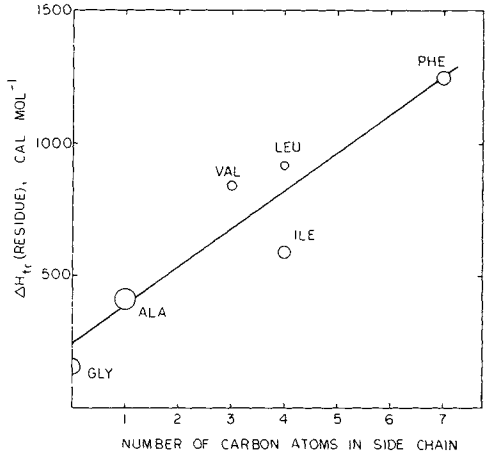

Fig. 2. Variation of ΔH_{tr} (residue) with carbon number for amino acid residues containing hydrophobic side chains.

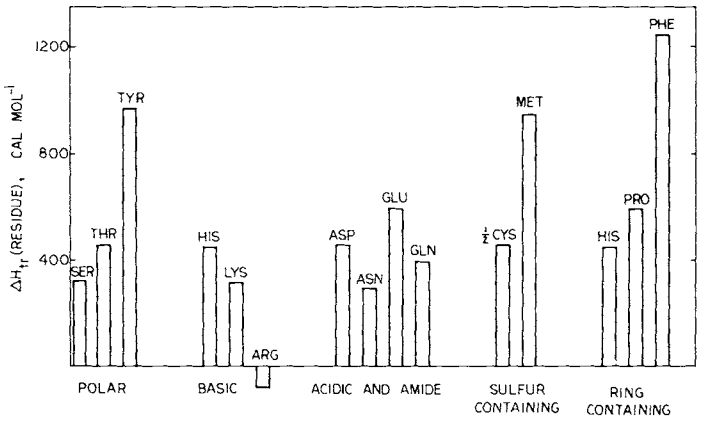

Fig. 3. Comparison of ΔH_{tr} (residue) for various classes of amino acid side chains.

lated assuming that all amino acid residues are accessible to the surface. The amino acid residues for ribonuclease A are tabulated in column 1 of Table II; the Wertz–Scheraga[5] and Lee–Richards–Chothia[2,4] probabilities are tabulated in Table III. Assuming each of these residues is completely exposed to the solvent during the transfer process gives a total enthalpy of transfer of 60,950 cal/mol for the totally exposed ribonuclease (Table IV). This is the predicted maximal enthalpy of transfer.

TABLE II
Amino Acid Residues in Ribonuclease A, Chicken Egg Lysozyme, and Ovalbumin

Residue	Ribonuclease A	Lysozyme	Ovalbumin
Ala	12	12	35
Arg	4	11	15
Asn	10	13	0
Asp	5	8	32
½Cys	8	8	6
Gln	7	3	0
Glu	5	2	52
Gly	3	12	19
His	4	1	7
Ile	3	6	24
Leu	2	8	31
Lys	10	6	20
Met	4	2	15
Phe	3	3	19
Pro	4	2	15
Ser	15	10	38
Thr	10	7	15
Trp	0	6	3
Tyr	6	3	10
Val	9	6	32

TABLE III
Wertz–Scheraga (WS) and Lee–Richards–Chothia (LRC) Probabilities That Amino Acid Residues Appear as Surface Residues

Residue	WS	LRC
Ala	0.48	0.618
Arg	0.51	0.987
Asn	0.58	0.877
Asp	0.63	0.855
Cys	0.17	0.500
Gln	0.65	0.932
Glu	0.62	0.816
Gly	0.59	0.637
His	0.30	0.827
Ile	0.21	0.396
Leu	0.23	0.545
Lys	0.69	0.968
Met	0.24	0.600
Phe	0.13	0.500
Pro	0.65	0.822
Ser	0.51	0.773
Thr	0.62	0.774
Trp	0.14	0.733
Tyr	0.36	0.850
Val	0.28	0.458

TABLE IV
Summation of ΔH_{tr} (Residue) for Amino Acid Residues of Ribonuclease A[a]

Amino Acid Residue	Number × ΔH_{tr} (AA) (cal/mol)	ΔH_{tr} Contribution (cal/mol)	
		WS Model	LRC Model
Ala	4956	2379	3063
Arg	−280	−143	−276
Asn	2930	1699	2570
Asp	2290	1443	1958
½Cys	3648	620	1824
Gln	2772	1802	2584
Glu	2980	1848	2432
Gly	462	273	294
His	1804	541	1492
Ile	1764	370	699
Leu	1838	423	1002
Lys	3140	2167	3040
Met	3784	908	2270
Phe	3735	486	1868
Pro	2372	1542	975
Ser	4800	2448	3710
Thr	4560	2827	3529
Tyr	5808	2091	4936
Val	7587	2124	3475
Totals	60,950	25,848	41,445

[a] Maximal ΔH_{tr} and ΔH_{tr} calculated using the probabilities of Wertz-Scheraga (WS) and Chothia (LRC).

Since the RCAM ribonuclease is expected to assume a structure where the residues are exposed to solvent, experimental enthalpies of transfer for this species can be compared directly with the predicted maximum. ΔH_{tr} for RCAM ribonuclease is 50,441 cal/mol. The two observations show that approximately 83% of the maximal surface is actually exposed in the RCAM ribonuclease.

The enthalpy of transfer for the native compound, 29,255 cal/mol, indicates that 0.48 ± 0.06 of the maximal accessible surface is actually exposed to solvent. The remainder of this maximal surface is buried in the interior of the native protein. The accessible surface area of the native protein is only 0.58 ± 0.016 of the surface area observed for the RCAM ribonuclease.

The enthalpy of transfer for the native protein can also be predicted using the probabilities for appearance at the surface developed by Wertz and Scheraga[5] and Chothia.[4] Since the probabilities developed by each technique differ, the observed data for the native protein and the amino acid residues can be used to compare the two probability sets as models for accessible surface.

The probabilities and enthalpies of transfer for each of the models are tabulated in Table IV. The WS model predicts a native enthalpy of transfer of 25,848 cal/mol, which can be compared with the experimental value of 29,255 cal/mol. The result is equivalent to exposure of 0.42 of the maximal accessible surface area. The Chothia probabilities predict an enthalpy of transfer of 41,445 cal/mol, which corresponds to an exposure of 0.68 of the maximal accessible surface area.

The WS solvent-accessible fraction of 0.42 lies just within the error limits determined from the native protein and model amino acid enthalpy of transfer measurements and is within 12% of this value. The Chothia model predicts a solvent-accessible fraction (0.70), which is outside the experimental limits and differs from the experimental fraction by 46%. The data clearly favor the probabilities developed by Wertz and Scheraga.

Lysozyme Experiments

Enthalpies of transfer for lysozyme and RCAM lysozyme are listed in Table V. The studies include the variation in enthalpies of transfer from aqueous buffer to solutions with different weight percent ethanol. In addition, they include enthalpies of transfer for the lysozyme from aqueous solution to solutions containing 10 wt % methanol and 10 wt % *n*-propanol. Both sets of measurements are repeated for the RCAM lysozyme. These data permit an analysis of enthalpy changes as ethanol weight percent is varied and a comparison of such enthalpy changes with different cosolvent alcohols.

The enthalpy of transfer increases as the weight percent of ethanol increases for both the native lysozyme and the RCAM lysozyme. There

TABLE V
Summation of ΔH_{tr} (Residue) for Amino Acid Residues of Lysozyme[a]

Amino Acid Residue	Number × ΔH_{tr} (AA) (cal/mol)	ΔH_{tr} Contribution (cal/mol)	
		WS Model	LRC Model
Ala	4956	2379	3063
Arg	−770	−393	−760
Asn	3809	2209	3340
Asp	3664	2308	3133
½Cys	3648	620	1824
Gln	1188	772	1107
Glu	1192	739	973
Gly	1488	878	948
His	451	135	373
Ile	3528	741	1397
Leu	7352	1691	4007
Lys	1884	1300	1824
Met	1892	454	1135
Phe	3735	486	1868
Pro	1186	771	975
Ser	3200	1632	2474
Thr	3192	1979	2471
Trp	7092	993	5198
Tyr	2904	1045	2468
Val	5058	1416	2317
Totals	60,649	22,155	40,135

[a] Maximal ΔH_{tr} and ΔH_{tr} calculated using the probabilities of Wertz–Scheraga (WS) and Chothia (LRC).

are no major changes that might signal a significant change in the accessible surface area of the protein. Thus, for both lysozyme moieties, the enthalpies of transfer are expected to provide a reasonable measure for the determination of surface area.

The enthalpies of transfer are significantly larger at each weight percent for the RCAM lysozyme. Since the RCAM lysozyme is expected to have a greater number of nonpolar residues exposed to solvent, these results are consistent with those observed for amino acids with such exposed side chains.

Enthalpies of transfer to solutions containing 10 wt % methanol, ethanol, and propanol increase in opposite directions for the native and RCAM lysozyme. For native lysozyme, methanol solutions generate the largest enthalpy of transfer. The RCAM lysozyme, however, generates the largest enthalpy of transfer when it is transferred to the *n*-propanol solution. However, the actual fraction of alcohol molecules that compete with the water molecules for accessibility to the surface is described by the mole fraction of these alcohols. The solvent size rather than the solvent polarity dictates the relative magnitudes of the enthalpies of transfer. In the RCAM lysozyme, enthalpy of transfer increases with decreasing polarity of the alcohol. This trend

is consistent with that observed for ribonuclease A and lysozyme for increasing weight percent of alcohol. As the proportion of less polar solvent increases, an increase in the enthalpy of transfer to the alcohol–water system is observed. Enthalpies of transfer per mole of protein per mole fraction of each alcohol are:

$$\Delta H_{tr} \text{ (propanol)} = 609.4$$

$$\Delta H_{tr} \text{ (ethanol)} = 516.7$$

$$\Delta H_{tr} \text{ (methanol)} = 410.2$$

The corresponding enthalpies for the RCAM lysozyme are:

$$\Delta H_{tr} \text{ (propanol)} = 2350$$

$$\Delta H_{tr} \text{ (ethanol)} = 1736$$

$$\Delta H_{tr} \text{ (methanol)} = 1034$$

The results are similar to those observed with increasing weight percent of ethanol. The less polar alcohols produce the larger enthalpy of transfer. The alcohols increase hydrophobic bonding to produce the observed increase.

Since both the lysozyme and RCAM lysozyme increase regularly in the weight percent regime studied, the system is again assumed to be suitable for a comparison of the experimentally observed protein enthalpies of transfer and the total enthalpies determined from the enthalpies of transfer for the residues and the WS and LRC probabilities. The predicted enthalpies of transfer from each set of probabilities can then be compared with the maximal enthalpy of transfer deduced as a summation of enthalpies of transfer for all the residues. This maximal value of 60,649 cal/mol cannot be compared with the observed value for the RCAM lysozyme, since transfers of this species involved 10 wt % solutions without 0.15 KCl to promote dissolution. Thus, the observed RCAM lysozyme enthalpy of transfer (72.9 kcal/mol) cannot be used for a quantitative comparison, although it does illustrate the large increase in enthalpy of transfer for the RCAM lysozyme.

The enthalpies of transfer for the chicken egg lysozyme calculated for the WS and LRC probabilities are 22,155 and 40,135 cal/mol, respectively. These enthalpies indicate fractions of accessible surface of 0.37 and 0.65, respectively. The fraction of accessible surface calculated from the observed enthalpy of transfer for the native lysozme and the maximal predicted enthalpy is 0.36. The WS model differs from the experimental prediction of the lysozyme by only 2.7%, while the LRC model differs from this prediction by 81%. Again, the WS model

provides a more accurate prediction of the experimental enthalpy of transfer.

Ovalbumin

Ovalbumin presented several difficulties which restrict its usefulness for accurate quantitative results. The protein is easily denatured and difficult to dissolve during the calorimetric experiments. To minimize such problems, the experiments were run in 0.15M KCl at a pH of 7.8. Since the residue enthalpies of transfer were determined only at pH $=2$, the comparison between them and the observed enthalpy of transfer for the native protein can contain systematic errors.

The maximal enthalpy of transfer deduced from the residue enthalpies of transfer is 222,125 cal/mol (Table VI). The WS[5] model predicts an enthalpy of transfer of 85,741 cal/mol for the native protein while the LRC[2,4] model predicts an enthalpy of transfer of 144,103 cal/mol. These values are compared with an experimental enthalpy of transfer of 88,850 cal/mol. Despite the difficulties inherent in the ovalbumin studies, the WS model appears in the proper range while the LRC prediction is again significantly larger.

TABLE VI
Summation of ΔH_{tr} (Residue) for Amino Acid Residues of Ovalbumin[a]

Amino Acid Residue	Number $\times$ ΔH_{tr} (AA) (cal/mol)	ΔH_{tr} Contribution (cal/mol)	
		WS Model	LRC Model
Ala	14455	6938	8933
Arg	−1050	−536	−1036
Asn	0	0	0
Asp	14656	9233	12531
½Cys	2736	564	1550
Gln	0	0	0
Glu	30992	19215	25289
Gly	2356	1390	1501
His	3157	947	2611
Ile	14112	2964	5588
Leu	28489	6552	15527
Lys	6280	4333	6079
Met	14190	3406	8514
Phe	23655	3075	11828
Pro	8895	5782	7312
Ser	12160	6102	9400
Thr	6840	4241	5294
Trp	3546	496	2579
Tyr	9680	3485	8228
Val	26976	7553	12355
Totals	222,125	85,741	144,103

[a] Maximal ΔH_{tr} and ΔH_{tr} calculated using the probabilities of Wertz–Scheraga (WS) and Chothia (LRC).

DISCUSSION

The enthalpies of transfer of proteins from buffered aqueous solutions to alcohol–water solutions are used as experimental observables for the determination of the proportion of maximal solvent surface available in native proteins. These data are used in two ways. First, enthalpies of transfer are determined for the constituent amino acids of the proteins. These enthalpies then provide a measure of the total enthalpy of transfer for a given protein under conditions in which all its residues are exposed to the solvent for each of the solutions. This predicted maximal enthalpy of transfer is then correlated with 100% exposure of the surface residues. Model amino acids are used to simulate the accessible surface area of the amino acid as it would appear in an extended primary protein structure.

To ascertain whether maximal accessible surface calculated in this way was reasonable, ribonuclease A and lysozyme were reduced and carboxamidomethylated to produce a new protein whose accessible surface area would approach the area predicted by the residue summation, i.e., the enthalpy of transfer observed for the RCAM protein would approach that calculated from the total protein residues. For RCAM ribonuclease A, the observed enthalpy of transfer was 83% of the total residue enthalpy of transfer. RCAM lysozyme produced an enthalpy of transfer that exceeded the maximal value calculated from the residue data by 20%. However, since the enthalpy of transfer took place between aqueous solutions and salt-free ethanol–water solutions, a direct comparison is not possible. The ovalbumin molecule has only a single cystine linkage. The reduction of this cystine linkage was not expected to produce a significant increase in the accessible surface area, and for this reason, RCAM ovalbumin was not produced for these experiments.

The enthalpies of transfer for the native proteins and the amino acid residues also permitted a comparison of two distinct techniques that were developed to determine solvent-accessible surface area. For all three proteins studied, the Wertz–Scheraga model gave enthalpies of transfer that were significantly lower than those determined for the LRC model. More significantly, the WS probabilities gave a more accurate fit of the native protein enthalpy of transfer for all three proteins.

An examination of the probabilities for WS and LRC in Table III shows that all the LRC probabilities are larger than the corresponding WS probabilities. This might suggest that the LRC probabilities could provide a more accurate picture of the experimental data if the criterion used by Chothia for a surface residue were made more stringent. It is interesting that the three enthalpies of transfer calculated using the LRC model can be reduced to those calculated with the WS model by factors between 0.56 and 0.60. However, there are fundamental

differences between the probabilities calculated by each method. Although a reduction of 0.60 in the LRC enthalpy of transfer gives a more accurate enthalpy of transfer, this correlation cannot be extended to the individual probabilities that produce this result. For example, the tryptophan probabilities are 0.14 for the WS model and 0.733 for the LRC model. For glycine, the respective probabilities are 0.59 and 0.637. Consistent scaling between the two models is impossible.

The LRC model determines the accessible surface area by rolling a solvent molecule over all atoms found at the surface. Since the side chains can protrude into solution, the relative area associated with such side chains can be large. This was noted by Shrake and Rupley,[3] who observed an increase in the probabilities when the effects of side chains were analyzed in more detail. In the WS model, the grid lines must intersect the group a minimum number of times if the group is to qualify as a surface group. A side chain perpendicular to the surface would be intersected by a minimal number of grid lines parallel to the chain, and this would produce a lower surface probability. The LRC model would be forced to roll over the entire group, which would yield the larger probability.

The difference between enthalpies of transfer calculated by the two models is amplified by the fact that the amino acids with large side chains that are overestimated by the LRC theory are also those with the larger enthalpies of transfer. These nonpolar surface groups contribute significantly to the total enthalpy of transfer in each of the models. This indicates that the structure of solvent about these groups in the alcohol–water mixture is less effective than the simple aqueous solution in stabilizing the protein.

The data presented here show that enthalpies of transfer from aqueous solutions to alcohol–water solutions can serve as a measure of accessible surface area when these enthalpies of transfer are correlated with enthalpies of transfer for the constituent amino acids of the protein. These data probe the accessible surface of the native protein as it exists in solution, and this makes the data particularly useful.

We acknowledge the support of the National Institute of Health (GM 1176217).

References

1. Némethy, G., Peer, W. & Scheraga, H. (1981) *Annu. Rev. Biophys. Bioeng.* **10**, 459–498.
2. Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55**, 379–400.
3. Shrake, A. & Rupley, J. A. (1973) *J. Mol. Biol.* **79**, 351–371.
4. Chothia, C. (1976) *J. Mol. Biol.* **105**, 1–12.
5. Wertz, D. & Scheraga, H. (1978) *Macromolecules* **11**, 9–15.
6. Almog, R., Schrier, M. & Schrier, E. (1978) *J. Phys. Chem.* **82**, 1703–1707.
7. Schrier, E., Ingwall, R. T. & Scheraga, H. (1965) *J. Phys. Chem.* **69**, 298–303.

8. Brandts, J. F. & Hunt, L. (1967) *J. Am. Chem. Soc.* **89**, 4826–4838.
9. Stimson, E. & Schrier, E. (1975) *Biopolymers* **14**, 509–520.
10. Defoster, G. (1982) Thesis, State University of New York at Binghamton.
11. White, F. H. Jr. (1972) *Methods Enzymol.* **25B**, 342–387.

Received October 9, 1984
Accepted July 8, 1985