[CONTRIBUTION FROM THE DEPARTMENT OF CHEMISTRY, UNIVERSITY OF MASSACHUSETTS, AMHERST, MASSACHUSETTS]

The Thermodynamics of Protein Denaturation. II. A Model of Reversible Denaturation and Interpretations Regarding the Stability of Chymotrypsinogen

BY JOHN F. BRANDTS

RECEIVED MARCH 30, 1964

A phenomenological model is developed to describe the process of reversible protein denaturation. This model incorporates, in an approximate manner, free-energy contributions arising from hydrogen bonding, hydrophobic bonding, conformational entropy, electrostatic repulsions, and anomalously titrating groups. The unusually large increase in heat capacity when nonpolar side chains are transferred from protein interior to solvent permits one to separate hydrophobic bonding contributions from other contributions and thereby to make estimates of the absolute contribution to protein stability arising from the above factors. The analysis of the chymotrypsinogen data from the preceding paper indicates that this protein is stabilized in nearly equal amounts of hydrogen bonds and by hydrophobic bonds and is unstabilized primarily by a large conformational entropy of unfolding. Numerical estimates are included. The effect of salt and pH on stability suggests that changes in the ionic nature of the solvent must be viewed with regard to *both* local interactions and general long-range electrostatic repulsions in acidic media. The experimental data on ethanol action are shown to be semi-quantitatively consistent with the idea that this agent denatures primarily by weakening hydrophobic bonds and that it may do so in an indirect manner through clathrate disruption and more favorable solvation. A preliminary analysis of urea action shows that there is a poor correlation between the expected effect of urea, as determined from model compound studies, and the experimentally determined effect.

A fairly detailed study of the changes in physical properties[1] of chymotrypsinogen solutions during reversible denaturation has failed to yield any definite information about the nature of the unfolding process. The nominal change in optical rotation and optical rotatory dispersion parameters might be interpreted as indicating that the average spatial arrangement of the peptide groups is left virtually unchanged by the processes which occur during denaturation and consequently suggests that this denaturation may not involve an extensive unfolding of the polypeptide chain. On the other hand, it has been shown[2] that *all* of the "nonexchangeable" protons in native chymotrypsinogen exchange rapidly once reversible denaturation has occurred. This fact might then be used to support the contrary opinion that denatured chymotrypsinogen is extensively unfolded and devoid of all but the most labile interpeptide hydrogen bonds. The small but significant change in specific viscosity and other physical properties likewise provides no conclusive evidence which would permit a choice between these two interpretations nor do they suggest a more comfortable alternative description. In lieu of this, it seems desirable to appeal to thermodynamics as a supplementary tool to gain some additional information about the unfolding process.

In this paper the experimental thermodynamic data from the previous paper are used to support a fairly detailed picture of the reversible denaturation of chymotrypsinogen. The approach which will be used is not novel and the author is particularly indebted to Tanford[3] for the treatment of hydrophobic bonding which shall be used. The development of the model used to characterize the denaturation process is perfectly general and follows closely the approach used previously by Schellman,[4] Scheraga,[5] Kauzmann,[6] and Tanford,[3] whereby the total thermodynamics of denaturation is considered to result from partial contributions from the different force interactions and entropy factors which collectively determine the balance of free energy between the native and denatured forms of the protein. The justification for this procedure has been given previously and will not be repeated here.

The Model

Most noteworthy of the several approaches to the problem of protein stability is that fostered by Kauzmann,[6,7] who pictures the native protein as a tightly folded structure maintained by a large number of secondary interactions between spatially adjacent amino acids, and the denatured protein as a flexible chain of high entropy nearly devoid of intramolecular interactions and resembling a random chain in the extreme limit. The general unfolding process which we shall describe is in accord with this picture and is shown diagrammatically in Fig. 1. The native protein consists predominantly of well-folded segments of polypeptide chain, shown as solid lines, containing amino acid residues whose peptide groups are involved in intramolecular hydrogen bond formation and whose side-chain groups are similarly participating in hydrogen bond interactions, hydrophobic interactions, and possibly other types of interactions depending on the nature of the side chains. There may also exist large or small segments of the chain, shown as dashed lines, which are unfolded in the native protein, and in these segments the residues will be well solvated and involved in no strong intramolecular interactions.

Most of the residues in the denatured protein will be located in "randomly coiled" segments of the polypeptide chain and consequently will be fully solvated. However, we must allow for the possibility that some of the folded structure of the native protein will be retained in the denatured protein because of unusually favorable residue sequence or because of favorable entropic stabilization introduced by suitable positioning of disulfide links, as illustrated in Fig. 1.

This unquestionably represents an oversimplification of the actual denaturation process. We have assumed that the state of any particular residue may be classified as "folded" or "unfolded" on the basis of

(1) J. F. Brandts and R. Lumry, *J. Phys. Chem.*, **67**, 1484 (1963).
(2) P. E. Wilcox, *Biochim. Biophys. Acta*, **34**, 602 (1959).
(3) C. Tanford, *J. Am. Chem. Soc.*, **84**, 4240 (1962).
(4) J. A. Schellman, *Compt. rend. trav. lab. Carlsberg, Ser. chim.*, **29**, 223 (1955).
(5) H. A. Scheraga, *J. Phys. Chem.*, **64**, 1917 (1960).
(6) W. Kauzmann, *Advan. Protein Chem.*, **14**, 1 (1959).
(7) W. Kauzmann in "Mechanism of Enzyme Action," W. D. McElroy and B. Glass, Ed., Johns Hopkins University Press, Baltimore, Md., p. 70.

whether or not it is involved in intramolecular interactions, whereas an exact description of the system must include all gradations of interaction and would, therefore, have to treat each residue in the native and denatured state individually. However, we need not be too concerned with this approximation as long as it is remembered that we are talking about the average state of all residues rather than the particular state of each residue.

It is not necessary to make any assumptions regarding the precise three-dimensional arrangement of the polypeptide chain in folded regions of the protein. It will be assumed that all folded residues will be so located that their peptide hydrogen bonding valences are satisfied intramolecularly and so that their side-chain groups are favorably located with respect to their free-energy requirements, but whether this is best accomplished by the formation of an α-helix or by an amorphous folding of the backbone need not be specified.

Titration of Anomalous Groups.—The well-known effects of pH and ionic strength on the stability of native proteins may be properly attributed to two principal factors:

(1) The presence of "anomalous" groups which have an unusually high or an unusually low dissociation constant in the native protein but a normal dissociation constant in the denatured protein. This situation undoubtedly results from strong local interactions of these groups in the native protein.

(2) Changes in long-range electrostatic interactions as the charge densities on the native and denatured forms are changed by altering the pH or as the extent of charge-screening is changed by variations in the ionic strength.

We will refer to the second situation as the "electrostatic" effect, in keeping with the usual nomenclature, even though it is recognized that anomalous groups may result from predominantly electrostatic interactions also. The primary distinction is then between the relatively weak long-range interactions which occur in the second case and the strong interactions occurring over short distances of the order of van der Waals radii in the first case.

If we regard N_0 as being that form of the native protein which has all of the anomalous groups, which we shall number 1, 2, 3, ..., in the unprotonated form then we may define dissociation constants for the one-step protonation reactions in the following way

$$N_0 + H^+ = N_1 \quad (N_1) = \frac{a_H}{K_N{}^1}(N_0)$$

$$N_0 + H^+ = N_2 \quad (N_2) = \frac{a_H}{K_N{}^2}(N_0)$$

$$N_0 + H^+ = N_3 \quad (N_3) = \frac{a_H}{K_N{}^3}(N_0)$$

where a_H is the hydrogen ion activity and where (N_1) is the concentration of native protein which has only anomalous group 1 protonated, (N_2) the concentration of native protein which has only group 2 protonated, etc. The concentrations of those species having more than one anomalous group protonated may be calculated readily from the one-step dissociation constants, *i.e.*

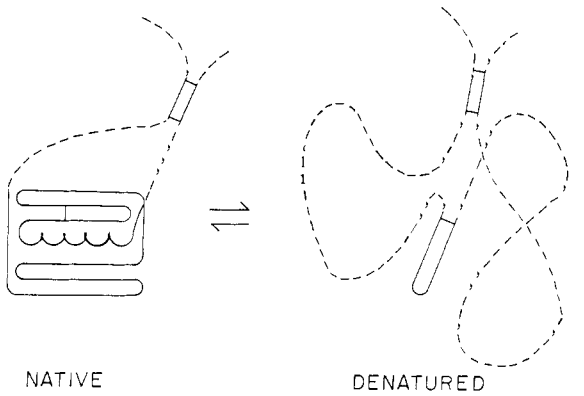

Fig. 1.—Simplified denaturation reaction; the solid lines refer to polypeptide chain containing "folded" residues while the dashed lines refer to fully solvated polypeptide chain containing "unfolded" residues.

$$(N_{1,2}) = \frac{a_H{}^2}{K_N{}^1 K_N{}^2}(N_0)$$

$$(N_{1,3}) = \frac{a_H{}^2}{K_N{}^1 K_N{}^3}(N_0)$$

$$(N_{2,3}) = \frac{a_H{}^2}{K_N{}^2 K_N{}^3}(N_0)$$

$$(N_{1,2,3}) = \frac{a_H{}^3}{K_N{}^1 K_N{}^2 K_N{}^3}(N_0)$$

if it is assumed that the dissociations occur independently.

The total concentration of native protein will be equal to the sum of the concentrations of all ionic species, or

$$(N) = (N_0)\left(1 + \frac{a_H}{K_N{}^1} + \frac{a_H}{K_N{}^2} + \frac{a_H}{K_N{}^3} + \frac{a_H{}^2}{K_N{}^1 K_N{}^2} + \frac{a_H{}^2}{K_N{}^1 K_N{}^3} + \frac{a_H{}^2}{K_N{}^2 K_N{}^3} + \frac{a_H{}^3}{K_N{}^1 K_N{}^2 K_N{}^3} + \cdots\right)$$

This expression may be written more concisely as

$$(N) = (N_0)\prod^{i}\left(1 + \frac{a_H}{K_N{}^i}\right)$$

where the product is taken over all anomalous groups in the protein. An analogous expression may be written for the denatured protein

$$(D) = (D_0)\prod^{i}\left(1 + \frac{a_H}{K_D{}^i}\right)$$

where the K_D values refer to the dissociation constants of the same groups when the protein exists in the denatured state and D_0 is again that form of the denatured protein which has all the anomalous groups in the unprotonated state.

The standard free energy of denaturation, $\Delta F^\circ = -RT \ln (D)/(N)$, will then be properly written as the sum of two terms

$$\Delta F^\circ = \Delta F^\circ{}_0 + \Delta F^\circ{}_{titr} \qquad (1)$$

where

$$\Delta F^\circ{}_0 = -RT \ln (D_0)/(N_0) \qquad (2)$$

$$\Delta F^\circ_{titr} = -RT \ln \prod^{i} \left(\frac{1 + a_H/K_D^i}{1 + a_H/K_N^i} \right) \quad (3)$$

Contributions to ΔF°_0.—The first term on the right of eq. 1, ΔF°_0, is simply the standard free-energy change when 1 mole of native protein containing all anomalous groups in the unprotonated state is transformed into denatured protein containing the anomalous groups in a similar state. It will then include all contributions to ΔF° except that associated with titrating the anomalous groups from the unprotonated state to that state of ionization existing under the solution conditions where ΔF° is calculated. The latter contribution is given by the ΔF°_{titr} term. In the evaluation of the ΔF°_0 term, the precise state of ionization of all of the "normally" ionizing groups is unspecified and will, of course, vary with solution conditions.

The ΔF°_0 term is best evaluated by estimating the contribution from each amino acid residue involved in the unfolding transition and then summing these contributions over all residues. Since the change in electrostatic free energy depends upon long-range interactions among all charged groups, this is best treated as a single contribution, however, so that we have

$$\Delta F^\circ_0 = \sum_i \Delta f^i_{res} + \Delta F_{elec} \quad (4)$$

where ΔF_{elec} is the total change in electrostatic free energy and where Δf^i_{res} is the individual free-energy contribution from residue i, excluding long-range electrostatic effects. The summation is taken over all residues which participate in the transition.

In this approximate treatment, we shall attempt to treat only the more important residue contributions. More specifically, we shall assume that the residue contributions arising from unfolding are associated with the breaking of peptide–peptide hydrogen bonds, changes in the solvation of side-chain groups, and from alterations in conformational entropy which might occur. Equation 4 then becomes

$$\Delta F^\circ_0 = \sum_i (\Delta h^i_H - T\Delta s^i_c + \Delta f^i_{sc}) + \Delta F_{elec} \quad (5)$$

where Δf^i_{sc} is the unitary free-energy change when the side chain of residue i is transferred from its local environment in the native protein to its environment in the denatured protein, Δh^i_H is the change in the peptide hydrogen bond enthalpy of residue i which occurs during the denaturation process and Δs^i_c is the contribution of residue i to the total conformational entropy change during unfolding. It should be noted that Δf^i_{sc} includes only the free energy of transfer of the side chain and, therefore, will not include conformational entropy changes resulting from changes in motional freedom. All conformational entropy changes, including those associated with both the backbone and side-chain portions of the residues, will be included in Δs_c.

In the case of a transition from a *completely folded* to a *completely unfolded* protein, the terms in the summation on the right of eq. 5 will be evaluated for each residue when it transformed from its particular "folded" state to its "unfolded" state, which is here assumed to be the completely solvated state. Equation 5 would then become

$$\Delta F^\circ_0 = \left(N\overline{\Delta h_H} - NT\overline{\Delta s_c} + \sum_{i=1}^{N} \Delta f^i_{tr} \right) + \Delta F_{elec} \quad (6)$$

where $\overline{\Delta h_H}$ is the average enthalpy change for breaking a single peptide–peptide hydrogen bond, $\overline{\Delta s_c}$ is the average increase in conformational entropy per unfolded residue, and N is the total number of amino acid residues in the protein.

In the general case, however, the native protein may contain some residues which are unfolded while the denatured protein may retain some of the folded structure of the native protein so that the above equation would tend to overestimate the residue contributions by assuming complete unfolding. This may be taken into account by modifying eq. 6 so that

$$\Delta F^\circ_0 = p\left(N\overline{\Delta h_H} - NT\overline{\Delta s_c} + \sum_{i=1}^{N} \Delta f^i_{sc} \right) + \Delta F_{elec} \quad (7)$$

where the separate terms in the parentheses are still the contributions assuming complete unfolding but the p factor, which may be regarded as that fraction of the N residues which is actually unfolded during the cooperative transition, will make the expression generally applicable to situations in which only partial unfolding occurs.

The Estimation of $\sum_{i=1}^{N} \Delta f^i_{sc}$.—The evaluation of the change in unitary free energy when a particular side chain is transferred from its local environment in the folded protein to its predominantly solvent environment in the unfolded protein involves a prior decision as to the nature of the environment in the folded protein. In order to simplify matters, we will assume that the side chain of a folded residue must be located either in the interior of the protein, where it will be surrounded by hydrophobic side chains and by portions of the polypeptide chain, or at the protein–solvent interface where it will be in a completely solvated state. Those side chains which are located at the interface will undergo no change in unitary free energy during unfolding since they will also be completely solvated in the unfolded state, whereas those which are transferred from the low dielectric interior of the native protein to aqueous environment in the denatured protein may undergo quite large changes in free energy.

The decision as to the location of a particular side chain in the folded protein must be made with regard to the free-energy requirements of that side chain. Tanford[3] has shown that the unitary free energy of transfer of an amino acid i from some reference solvent to water is composed of two additive free-energy terms—one of which measures the transfer free energy for the peptide portion of the amino acid and the other of which measures the transfer free energy of the side-chain portion of the amino acid. Since the transfer free energy of glycine should adequately approximate the contribution from the peptide portion of any amino acid, Tanford concludes that the unitary free energy of transfer of just the side-chain portion of amino acid i, Δf^i_{tr}, may then be obtained from the difference in transfer free energy for the amino acid i and for glycine, *i.e.*

side chain i (reference solvent) $\rightleftharpoons$ side chain i (H_2O)

$$\Delta f^{i}_{tr} = -RT \ln \left(\frac{X^{i}_{H_2O} X^{gly}_{ref}}{X^{i}_{ref} X^{gly}_{H_2O}} \right) \qquad (8)$$

where the X values refer to mole fraction solubilities of amino acid i and glycine in water and reference solvent as indicated.

If the reference solvent is chosen so that it is an adequate representation of the environment in the interior of the protein, then Δf^{i}_{tr} should adequately approximate the change in unitary free energy when a side chain of type i is transferred from protein interior to solvent (assumed here to be water). Fortunately, the choice of reference solvent does not appear to be too critical. Tanford has shown that several relatively nonpolar reference solvents such as methanol, ethanol, acetone, and butanol lead to values of Δf^{i}_{tr} which differ by only about 300 cal. for large nonpolar side chains. In the present instance, we will use 95% ethanol (by volume) for the reference solvent intended to approximate the protein interior since solubility data over a wide temperature range are available in this solvent.

In this approximate treatment, we shall not attempt to treat small differences in the hydrophobic properties of the various amino acid side chains. In many cases this is not possible since sufficiently accurate data are not available. Tanford's previous treatment shows that a number of amino acids have very similar values of Δf^{i}_{tr} for transfer from absolute ethanol to water at 25°. This is also the case if 95% ethanol is used for the reference solvent except that the values of Δf^{i}_{tr} are less positive by 100–300 cal./mole for a typical side chain using this reference solvent.

All of the commonly occurring amino acids may be divided rather conveniently into six groups according to their relative values of Δf^{i}_{tr}. The approximate values of Δf^{i}_{tr} for each group of side chains at 25° have been estimated from solubility data[8–10] and are given in Table I. This division of side chains into groups does not introduce any additional restrictions into the treatment of transfer free energies since the values of Δf^{i}_{tr} which shall be used in the treatment of hydrophobic interactions must be regarded as being uncertain to the extent of 200–400 cal. for a typical side chain because of the uncertainty in choosing the correct reference solvent, and, in general, the values of Δf^{i}_{tr} for different side chains placed in the same group in Table I lie well within this range of uncertainty. We will therefore treat them as being identical.

It is apparent that those side chains in groups I–IV are all appreciably hydrophobic, as indicated by their large positive transfer free energy from reference solvent to water. We may reasonably assume then that these side chains, when they exist in well-folded regions of the protein, will be preferentially located in the interior of the protein when at all possible. Consequently, it will be assumed that when these residues are changed from a folded to an unfolded state during the conformational transition, their side-chain contribution to $\Delta F°$, *i.e.*, Δf^{i}_{sc}, may be satisfactorily approximated by Δf^{i}_{tr}.

(8) E. J. Cohn and J. T. Edsall, "Proteins, Amino Acids and Peptides as Dipolar Ions," Reinhold Publishing Co., New York, N. Y., 1943.

(9) "Handbook of Chemistry and Physics," Chemical Rubber Publishing Co., Cleveland, Ohio, 1963.

(10) Unpublished data fron this laboratory.

TABLE I

GROUPING OF AMINO ACID SIDE CHAINS ACCORDING TO TRANSFER FREE ENERGIES FROM 95% ETHANOL TO WATER AT 25°[a]

Group	Amino acid side chain	Δf^{i}_{tr}, cal.
I	Tyr	2800
II	Norleu, leu, ileu, phe, pro, tyr	2500
III	Val, meth, lys (nonpolar portion)	1500
IV	Ala, arg (nonpolar portion)	600
V	Gly, asp (NH_2), glu (NH_2), ser, thr, cys	0
VI	Asp, glu, his, and charged portion of arg and lys	<0

[a] The values of Δf^{i}_{tr} are approximate and have been estimated from solubility data.[8–10]

The side chains in group V have transfer free energies which are positive or negative to the extent of only a few hundred calories and are thus not appreciably hydrophobic or hydrophillic. They will show no strong preference for protein interior or solvent interface and should be perfectly happy in either location. Irregardless, the contribution of these side chains resulting from solvation changes will be small, no matter what their location in the folded protein, and it will be assumed that Δf^{i}_{sc} for these side chains will be zero. These side chains will contribute to the total free-energy change *when unfolded*, however, because of altered hydrogen bonding and conformational entropy, as seen in eq. 7.

All of those side chains in group VI will be charged or partially charged in neutral and acidic pH regions and therefore will be extremely hydrophilic. Although solubility data for the charged form of these amino acids in suitable reference solvents is lacking, the cost in free energy for ionizing these groups in low dielectric environments will be prohibitively large, and only in very unusual circumstances would we expect to find these groups located in the interior of a folded protein. If these side chains are then located at the protein surface in the folded state, any changes in solvation during unfolding will be small and we may assume Δf^{i}_{sc} for these groups will be zero also.

Tanford has suggested that the charged portion of long side chains, such as those of lysyl and arginyl, may be positioned at the interface while the hydrophobic portion is still predominantly in the protein interior. He assumed that the hydrophobic portion of the lysyl and arginyl side chains should resemble the valyl and alanyl side chains, respectively, in terms of their hydrophobic properties, and we have accordingly placed these basic side chains in the appropriate hydrophobic group as well as in group VI in order to retain this feature of Tanford's original treatment.

With these assumptions, the side-chain contributions will be

$$\sum_{i}^{N} \Delta f^{i}_{sc} = \sum_{I-IV} \Delta f^{i}_{tr} \qquad (9)$$

where the sum of transfer free energies on the right need only be taken over all of the hydrophobic residues in groups I–IV.

Obviously, this simplification neglects all side-chain interactions other than hydrophobic interactions. It has been postulated on numerous occasions that side-chain hydrogen bonding may contribute significantly to the stability of the folded form of globular proteins. It seems likely that this form of hydrogen bonding will

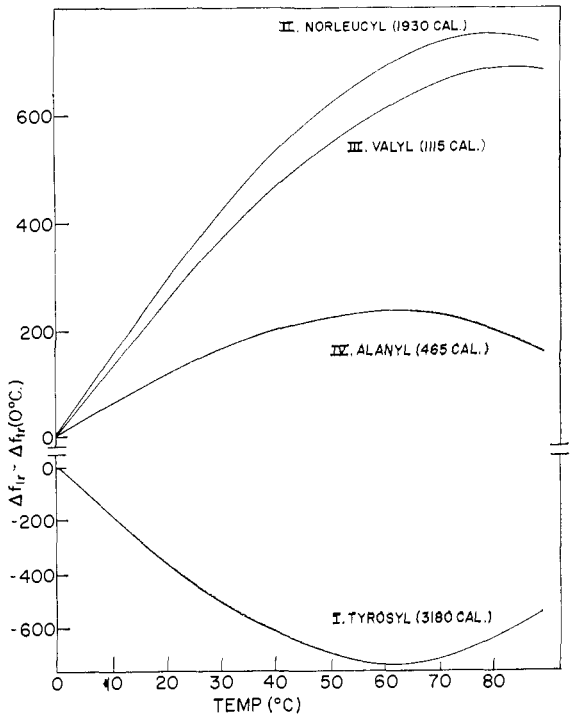

Fig. 2.—The temperature dependence of Δf^{i}_{tr} for side chains in groups I–IV; the values of Δf^{i}_{tr} at 0° are shown in parentheses.

be much less important than peptide hydrogen bonding largely because there are far fewer side-chain groups capable of forming strong hydrogen bonds, but also because the regularity of peptide hydrogen bonding valences provides a great advantage for participation in ordered structures such as the α-helix. Nevertheless, some of the polar side chains will undoubtedly be involved in intramolecular hydrogen bonding and this must be taken into account. It will soon become apparent, in the treatment which follows, that contributions arising from side-chain hydrogen bonding as well as contributions from other side-chain interactions such as weak dispersion forces will appear in the $\overline{\Delta h_H}$ term of eq. 7. Consequently, this term cannot be regarded strictly as a measure of contributions from alterations in peptide hydrogen bonding although it seems likely that this will be the most important single factor controlling the magnitude of $\overline{\Delta h_H}$.

Of particular interest in the present study is the temperature dependence of the Δf^{i}_{tr} terms. Because of the factors of water structure upon which hydrophobic bonding depends, we might anticipate a very strong temperature dependence, and this indeed turns out to be the case. For a great number of amino acids over a wide temperature range, it has been shown[8] that the logarithm of the mole fraction solubility in water may be accurately expressed as a power series involving a constant term and two additional terms dependent on the first and second powers of the absolute temperature. The same expression holds well for expressing the solubilities in 95% ethanol. Therefore, the side-chain transfer terms may be expressed as a power series involving the first three powers of the absolute temperature, *i.e.*

$$\Delta f^{i}_{tr} = A^{i}T + B^{i}T^{2} + C^{i}T^{3} \qquad (10)$$

The coefficients A^{i}, B^{i}, and C^{i} may be obtained by fitting eq. 10 to the experimental values of Δf^{i}_{tr} obtained from solubility data and eq. 8. The constants in eq. 10 have been determined for a representative member of each of groups I–IV from solubility data over the temperature range 0–80°. These are shown in Table II.

TABLE II

VALUES OF THE TEMPERATURE-INDEPENDENT CONSTANTS[a] FOR SIDE-CHAIN TRANSFER OF THOSE SIDE CHAINS IN GROUPS I–IV

Group	A	$B \times 10^{2}$	$C \times 10^{5}$
I (tyr)	100.46	−53.715	77.576
II (norleu)	−24.906	20.265	−31.330
III (val)	−25.746	18.394	−27.354
IV (ala)	−19.000	13.258	−20.790

[a] Calculated from data in ref. 8 and 9.

There are adequate temperature data available which indicates that the norleucyl, leucyl, isoleucyl, and phenylalanyl side chains in group II not only have similar values of Δf^{i}_{tr} at 25° but have very similar temperature dependence of Δf^{i}_{tr} as well. Consequently, in this approximate treatment we shall assume that Δf^{i}_{tr} for all members in group II is identical over the temperature range 0–80°. The norleucyl side chain will be used as the representative member of group II in spite of the fact that it is not one of the naturally occuring forms since it appears to be slightly more representative of all members in the group. Likewise, Δf^{i}_{tr} for valyl and alanyl side chains will be used for those members in groups III and IV, respectively, as indicated in Table II.

The smooth curves in Fig. 2 show the change in the value of Δf^{i}_{tr} for each of the first four groups as the temperature is increased above 0°. The value of Δf^{i}_{tr} at 0° is shown in parentheses. It is seen that the transfer free energy for the norleucyl, valyl, and alanyl side chains becomes considerably more positive as the temperature is increased above 0°, indicating a marked increase in the hydrophobicity of these groups. At high temperature, 60° for alanyl and about 75° for norleucyl and valyl, Δf^{i}_{tr} reaches a maximum value and decreases with further temperature increase.

The tyrosyl side chain has been included in a separate group since it appears to be atypical among the hydrophobic side chains in terms of its temperature dependence of transfer. For tyrosyl, Δf^{i}_{tr} decreases with increasing temperature and approaches a minimum value at about 60°. Thus, the temperature dependence in this case is inverted with regard to the other hydrophobic groups. Undoubtedly this results from the presence of the –OH function on the aromatic ring since the phenylalanyl side chain shows the more characteristic temperature dependence. Although the solubility data for tyrosine in 95% ethanol does not appear as accurate as for the other amino acids, the errors are not of the proper order of magnitude to account for the inversion of the temperature effect. Thus, it seems safe to conclude that the tyrosyl side chain is unusual in its temperature dependence although the reasons for this are far from clear.

The curves of Fig. 2 indicate that the enthalpic and entropic contributions to Δf^{i}_{tr} are extremely temperature dependent since a straight line relationship would result in the absence of a temperature dependence. For instance, the enthalpy for transfer of the norleucyl side chain from 95% ethanol to water is

-2400 cal. at 0° and becomes more positive as the temperature is increased, changing from negative to positive at 50° and increasing above this temperature to a value of $+3300$ cal. at 80°. The entropy of transfer is similarly negative at low temperature, -17 e.u. for norleucyl at 0°, and becomes more positive as the temperature is increased, changing sign at about 75°.

This temperature dependence results from a large difference in the heat capacity of these side chains in water as opposed to the reference solvent. The heat capacity difference may be obtained from eq. 10 and well-known thermodynamic relations and is

$$\Delta C^{i}_{p} = -2B^{i}T - 6C^{i}T^{2}$$

Thus, ΔC_p for the norleucyl side chain is 30 cal./mole deg. at 0° and increases to a value of 90 cal./mole deg. at 80°. This large difference in heat capacity as well as the temperature dependence of the difference is unquestionably a reflection of the state of solvation of the side chain in the aqueous phase since it was noted many years ago[8] that amino acids having nonpolar side chains exhibited unusually high calorimetric values of C_p in aqueous solutions. It seems highly likely that these effects result from the clathrate structures which are known to form about nonpolar groups in water.[6] The large negative entropy of transfer at low temperature indicates a pronounced ordering of solvent molecules in the aqueous phase and this, along with the negative enthalpy of transfer, may be a reflection of the formation of strongly hydrogen-bonded clathrate structures. The high heat capacity suggests that clathrate melting occurs as the temperature is increased and this therefore leads to less negative, and eventually positive, values for the entropy and enthalpy of transfer as the temperature is increased.

This extreme dependence of thermodynamics on temperature is highly reminiscent of the thermodynamics for the denaturation of chymotrypsinogen presented in the previous paper and suggests that the unusually large temperature dependence of the thermodynamics of denaturation may arise from those factors of water structure upon which hydrophobic bonding depends. This possibility will be explored subsequently.

The general expression for the standard free energy of denaturation may then be obtained from the combination of eq. 1, 7, and 9, and this is

$$\Delta F^{\circ} = p\left(N\overline{\Delta h_{H}} - NT\overline{\Delta s_{c}} + \sum_{1\text{-}1V}\Delta f^{i}_{tr}\right) + \Delta F_{elec} + \Delta F^{\circ}_{titr} \quad (11)$$

It would be rather meaningless to attempt to use the above equation to calculate a quantatitive value of ΔF° for chymotrypsinogen or any other protein in view of the number of approximations necessary to derive such an equation. This sort of calculation would necessitate prior evaluation of Δh_H and Δs_c. Independent estimates of these terms from model compound studies will be discussed later, but it should be pointed out here that an error of 100 cal. in the hydrogen bonding term or an error of 0.3 e.u. in the conformational entropy term would lead to an error of 10,000 cal. in the standard free energy of denaturation for a typical transition involving 100 or more residues. Since ΔF° probably does not greatly exceed 10,000 cal. for most denaturation reactions under any set of solution conditions, it seems questionable whether estimates from model compound studies will ever be accurate enough to permit a reliable theoretical calculation of ΔF° even if the precise three-dimensional structure of the protein were known.

Our ambitions are much more modest than this. If the calculated expression for ΔF°, given generally by eq. 11, is equated to experimentally determined values of ΔF°, then certain of the unassigned parameters, such as $\overline{\Delta h_H}$ and $\overline{\Delta s_c}$, may be evaluated with satisfactory accuracy if experimental data are available over rather wide ranges of solution conditions. This will, at the least, serve as a test for the internal consistency of the denaturation model and, at the most, perhaps provide order of magnitude estimates of the various contributions to thermodynamic stability.

Comparison of the Model with the Denaturation of Chymotrypsinogen

pH and Salt Dependence.—There seems little doubt that the effects of pH on protein folding must be closely related to the effects of ionic strength and that both find their origin at the ionizable side chains of the protein. It should be mentioned, however, that von Hippel and Wong[11] have recently suggested that the effects of neutral salts on the collagen → gelatin phase transition arise because of alterations in solvent structure produced by the ions and the corresponding effect which this has on the stability of hydrophobic bonds. Because of the same relative effects of different cations and anions on ribonuclease unfolding, they conclude that the same mechanism might be applicable to globular proteins as well. There is no experimental evidence which necessitates this conclusion, however, and the author feels that there is strong reason to reject it at least in the case of chymotrypsinogen. The experimental evidence indicates that the standard free energy of denaturation varies linearly with the logarithm of the chloride ion concentration down to concentrations as low as 0.003 *M*. At this concentration, only one solvent molecule in about 20,000 is a chloride ion and a significant perturbation of water structure seems very unlikely. In addition, a logarithmic dependence of free energy on concentration, observed also by von Hippel and Wong, is completely unexpected if water structure is the important variable, whereas this is the predicted concentration dependence for strong binding and approximately so for charge-screening at the protein surface.

There is no doubt that at high salt concentration or at high concentration of acid or base, there will be appreciable alterations in water structure which may alter the stability of hydrophobic bonds and the stability of hydrogen bonds as well. In the present study, the mole fraction of water in the solvent is always 0.995 or greater, and the conclusion that the predominant effect of variations in ionic strength, and also pH, will be concerned with direct interactions between the ions and the charged protein side chains seems justified.

If this is the case, then the only terms in eq. 11 which will be appreciably dependent upon the ionic nature of the solvent will be the final two so that the variation in

(11) P. H. von Hippel and K. Wong, *Biochemistry*, **2**, 1387 (1963).

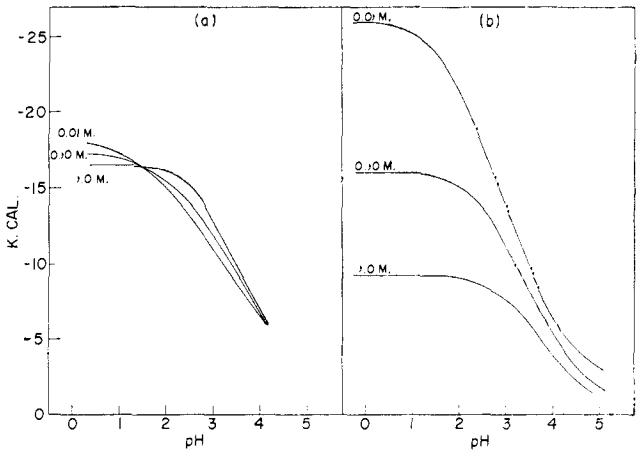

Fig. 3.—Variation in ΔF° with pH changes at constant temperature (25°) and ionic strength: (a) experimental, (b) Debye–Hückel estimate of ΔF_{elec}.

ΔF° as the pH is altered at constant temperature and ionic strength will arise from changes in the electrostatic free energy and from the titration of anomalous groups. There is, of course, no way in which the contribution from anomalous groups can be experimentally separated from the general long-range electrostatic contribution, and this is the principal reason why there is confusion in the literature regarding the relative importance of these two effects. Although no profound conclusions are forthcoming, it is nevertheless instructive to examine the experimental values of ΔF° to see if these are consistent with the notion that either of these two contributions is small and may be neglected with respect to the other.

Since we are concerned here with pH data below pH 3.5, we may reasonably conclude that any anomalous groups will be of the carboxyl variety so that we would expect, to a first approximation, that both K_N and K_D will be independent of temperature for all anomalous groups. If this is the case then ΔF°_{titr}, given by eq. 3, contains no enthalpy contribution and we may conclude that

$$\Delta S^\circ_{titr} = R \ln \prod^{i} \left(\frac{1 + a_H/K_D}{1 + a_H/K_N} \right)^i \qquad (12)$$

Reference to eq. 7 of the preceding paper shows that the total experimental variation in ΔS° as the pH is changed at 0.01 M Cl^- becomes precisely equal to the above expression if it is assumed that chymotrypsinogen contains three (actually, 3.18) anomalous carboxyl groups with a pK of 1.3 ($K_N = 5 \times 10^{-2}$) in the native protein and a normal pK of 4.5 ($K_D = 3 \times 10^{-5}$) in the denatured protein. Closer examination reveals, however, that this good agreement with experiment is partly fortuitous. At higher chloride ion concentrations of 0.1 and 1.0 M the experimental curves begin to show a poor quantitative resemblance to a titration curve (Fig. 8 of preceding paper). More important, it is not possible to obtain calculated entropy curves at two different ionic strengths which show a cross-over point, as found experimentally, simply by adjusting the values of K_N at the different ionic strengths. The anomalous group idea alone does not therefore appear completely consistent with the fact that salt acts as both a stabilizing and unstabilizing agent for native chymotrypsinogen. It is certainly not surprising that this simple interpretation is inadequate since, at acid pH, there are as many as 34 charged groups on the molecule, and the net charge reaches a maximum positive value of 20 so that the complete neglect of electrostatic interactions cannot be justified.

The other extreme position which neglects local interactions and assumes that variations in free energy arise solely from long-range interactions may be treated approximately. If it is assumed that native chymotrypsinogen is spherical with all charged groups uniformly distributed on the surface, then the Debye–Hückel theory will apply. It will also be assumed that the polypeptide chain of denatured chymotrypsinogen is sufficiently extended so that the electrostatic free energy of this conformation may be neglected relative to that of the compact native protein.

The details of similar calculations have appeared numerous times[12] and we will include only the results here. The curves in Fig. 3b show the expected magnitude of the ΔF_{elec} term at 0.01, 0.10, and 1.0 M ionic strength and 25°. These calculations were made assuming that the spherical radius of native chymotrypsinogen is 20 Å. and that the intrinsic pK of the fourteen carboxyl groups is 4.5. For comparison purposes, the total experimental variation in ΔF° with pH, adjusted to an arbitrary scale, is shown at approximately the same values of ionic strength in Fig. 3a. It is seen that the agreement between the experimental results and the expected Debye–Hückel electrostatic factor is extremely poor. The point which deserves particular attention is that the electrostatic model predicts a very large stabilization of the native protein as the ionic strength is increased, amounting to some 17,000 cal. in free energy in going from 0.01 to 1.0 M ionic strength at very acid pH. There is little evidence of this stabilizing effect in the experimental values. We have observed a very small stabilizing effect at low pH, but this effect is inverted at pH 1.5, in direct contrast to the Debye-Hückel prediction.

It does not seem possible that the poor agreement with experiment results solely from our failure to consider the electrostatic free energy of the denatured protein in an explicit fashion. The electrostatic free energy of the denatured form should show the same type of pH and salt dependence as that of the native form except to a much lesser extent because of its more extended conformation so that the effect of including this in the calculated values would be simply to lower the estimated free-energy differences at all ionic strengths in Fig. 3b. The lack of agreement must then result from the fact that the Debye–Hückel picture of native chymotrypsinogen as a spherical molecule with uniformly distributed charge is a very poor approximation of the actual situation. Moreover, any model based solely on long-range electrostatic interactions, no matter what the assumptions regarding shape and precise charge distribution, would give poor agreement with the experimental values since charge-screening is an important feature of all these models,[12] and this will always act to reduce an unfavorable electrostatic situation and thereby stabilize the native protein as the salt concentration is increased at acid pH.

(12) See, for instance, C. Tanford, "Physical Chemistry of Macromolecules," John Wiley and Sons, Inc., New York, N. Y., 1961.

We do not intend to imply by this that the electrostatic contribution is not important in this denaturation reaction, but it does seem rather certain that this does not tell the whole story. The lack of any significant sign of charge-screening in the experimental values is probably an indication that local interactions are rather important in this protein. The placement of carboxyl groups in the native protein must be such that favorable local interactions with other groups occur and that increased ionic strength acts to decrease the strength of these interactions. This effect is presumably large enough to overcome the stabilizing effect which must result from charge-screening of long-range repulsions in these acidic solutions with the result that salt may act as a denaturing agent at certain pH values.

Temperature Dependence and Estimates of p, $\overline{\Delta h_H}$, and $\overline{\Delta s_c}$.—As previously indicated, the denaturation of chymotrypsinogen is well suited for theoretical analysis, since the measurement of a number of physical properties indicates that the conformation of both the native and denatured forms are fairly insensitive to variations in pH, ionic strength, and temperature even though the equilibrium distribution between the two states is strongly dependent on all of these variables. In view of this, we may conclude that the transition always involves the unfolding of the same cooperative structural unit and we may treat the p factor as a constant independent of solution conditions.

The amino acid composition of chymotrypsinogen[13] shows that this protein has 4 amino acid residues with side chains belonging in group I, 51 in group II, 36 in group III, and 26 in group IV out of a total 240 residues. Using this information, the $\sum_{I-IV} \Delta f^i_{tr}$ sum may be readily evaluated from the constants in Table II and eq. 11 then becomes

$$\Delta F^\circ = p(240\overline{\Delta h_H} - 240T\overline{\Delta s_c} - 2289.2T + 18.255T^2 - 0.028128T^3) + \Delta F_{elec} + \Delta F^\circ_{titr} \quad (13)$$

If the above calculated expression for ΔF° is compared with the equivalent experimental expression, given by eq. 8 in the preceding paper, it is seen that the two will become equal at all temperatures if the values of p, $\overline{\Delta h_H}$, and $\overline{\Delta s_c}$ are chosen as follows: p = 0.634, $\overline{\Delta h_H}$ = 800 cal., and $\overline{\Delta s_c}$ = 5.10 e.u. This, of course, also assumes that the ΔF_{elec} and ΔF°_{titr} terms in eq. 13 are so chosen that their sum becomes equal to the pH and salt dependent terms in the experimental expression at all solution conditions. The errors in the assignment of the p, $\overline{\Delta h_H}$, and $\overline{\Delta s_c}$ values from the experimental data are less than 0.01, 10 cal., and 0.1 e.u., respectively.

Thus the model which we have chosen is completely adequate to account for the experimental results over the entire temperature range of study as shown by the good agreement between the calculated curve and the experimental results shown in Fig. 7 of the preceding paper. In view of the rather unusual temperature dependence, this provides some confidence that the model does incorporate those features which actually give rise to this temperature dependence. A certain amount of additional confidence is provided by the indicated values of the p, $\overline{\Delta h_H}$, and $\overline{\Delta s_c}$ terms which are in good agreement with independent evidence in so far as direct comparison can be made.

Schellman[4] has previously concluded that the conformational entropy change for unfolding a single residue in a polypeptide chain should not exceed 7–8 e.u., and Kauzmann[7] has estimated the correct value to be close to 4 e.u. These estimates include only restrictions in rotational freedom in the backbone and contributions from side chains probably would raise these estimates somewhat although perhaps not a great deal. These estimates are naturally rather uncertain but the value of 5.1 e.u. obtained in the present study certainly lies within acceptable bounds.

Schellman also estimated, from the activities of urea in aqueous solution, that the enthalpy change for breaking a peptide–peptide hydrogen bond in water will be about 1500 cal. The more desirable experimental technique of Klotz and Franzen[14] indicates that the correct value should be much closer to zero and that it will probably vary significantly with the location of the residue in the native protein—those residues located in the low dielectric interior of the protein forming stronger hydrogen bonds than those located at the interface. Wada's[15] estimate from studies on synthetic polypeptides places the value at 400 cal., but this is the result of a rather indirect method.

Again, the relatively small value of 800 cal. indicated in the present analysis is entirely reasonable. Even this estimate must probably be regarded as an upper limit of the strength of a peptide–peptide hydrogen bond since, if there are a significant number of side-chain hydrogen bonds in the native protein, then their contribution should also be included in $\overline{\Delta h_H}$ and this would tend to decrease the estimate for peptide hydrogen bond enthalpy. Our failure to consider explicitly the contributions from less important forces, such as ion pairing, should not significantly alter the picture. The indicated contribution to the stability of chymotrypsinogen provided by hydrogen bonding, *i.e.*, $pN\overline{\Delta h_H}$, exceeds 120,000 cal. Thus even if neglected interactions contributing to enthalpy stabilization are of the order of 12,000 cal., this will only introduce an error of 10% in the estimated contribution from hydrogen bonding.

The magnitude of the total contributions from hydrogen bonding, hydrophobic bonding, and conformational entropy are shown in Fig. 4. As stated, hydrogen bonding will contribute about 120,000 cal. of stabilizing free energy to native chymotrypsinogen according to this analysis. The contribution from hydrophobic bonding will be about 105,000 cal. at 0°, and this will increase rapidly at low temperature and level off at about 145,000 cal. at 75° and decrease at higher temperatures. These two important contributions to the free energy are approximately balanced at all temperatures by the unstabilizing free energy arising from the increase in conformational entropy. Thus, although there may be well over 250,000 cal. of stabilizing free energy, the net free energy of denaturation will seldom exceed 10,000 cal. because of the deli-

(13) P. E. Wilcox, E. Cohen, and W. Tan, *J. Biol. Chem.*, **228**, 999 (1957).

(14) I. M. Klotz and J. Franzen, *J. Am. Chem. Soc.*, **84**, 3461 (1962).

(15) A. Wada, *Mol. Phys.*, **3**, 409 (1960).

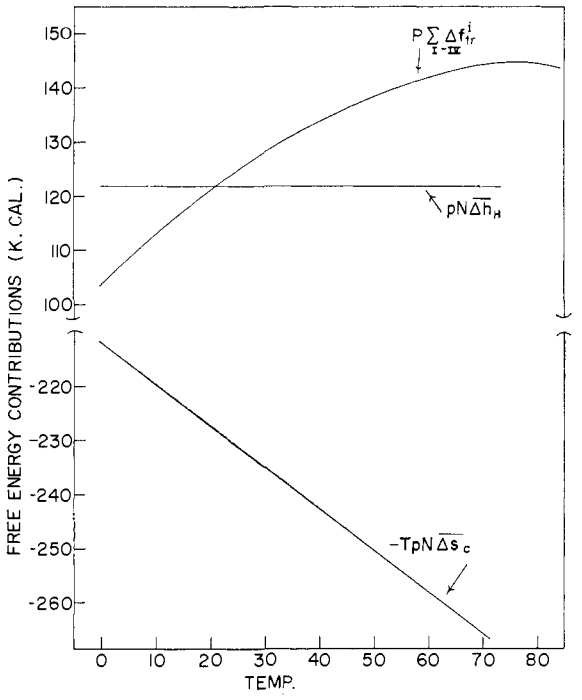

Fig. 4.—Estimates of the total contributions to the stability of native chymotrypsinogen resulting from hydrogen bonding, hydrophobic bonding, and conformational entropy.

cate balance between stabilizing and unstabilizing factors. It is apparent that even though the contributions from factors such as anomalous groups or from electrostatic free energy are small in comparison with the total stabilizing and unstabilizing free energy, they will be large enough to seriously alter the equilibrium as the pH is changed under the unusual conditions of marginal stability.

The occurrence of a temperature of maximum stability, which appears at about 10° for chymotrypsinogen, is another manifestation of the fine balance between stabilizing and unstabilizing factors. The unstabilizing effect of conformational entropy increases linearly with increasing temperature, as shown in Fig. 4. The stabilizing effect of hydrophobic bonding also increases with increasing temperature but to a much lesser extent at high temperature than at low temperature. At high temperature the increase in the unstabilizing effect from conformational entropy is greater than the increase in the stabilizing effect from hydrophobic bonding as the temperature is increased and the usual "thermal denaturation" is observed. At low temperature, the increased stabilization of hydrophobic bonds outweighs the conformational entropy effect so that the native protein becomes more stable as the temperature is increased leading to an "inverted transition."

The p factor indicates that about 63% of the 240 residues in chymotrypsinogen are involved directly in the unfolding process. Although an independent check of this estimate is not available, there is some evidence which indicates that somewhat less than 100% unfolding should be expected. The changes in optical rotation[16] during the activation of chymotrypsinogen to chymotrypsin suggest that there is a randomly coiled portion of the polypeptide chain in the zymogen which folds only after the cleavage of restrictive peptide bonds during the activation process. A preliminary treatment of the reversible denaturation of chymotrypsin in terms of the present model shows that there are a number of additional residues which unfold when the active enzyme denatures, in agreement with the activation studies, but even in this case somewhat less than 100% unfolding occurs. One possible reason for this is that the denatured form is not the completely unfolded form in both cases. If reversibly denatured chymotrypsinogen is transferred to a solution of 8 M urea, there is a marked increase in levorotation[1] which may arise from the further unfolding of particularly stable, ordered regions which cannot be unfolded in water alone. These arguments suggest that native chymotrypsinogen may be partially unfolded and that denatured chymotrypsinogen may be partially folded so that we would not expect 100% participation of all residues in the reversible transition. Beyond this, however, the present estimate of 63% participation cannot be substantiated by independent evidence.

(16) H. Neurath and G. H. Dixon, *Federation Proc.*, **16**, 791 (1957).

Effect of Ethanol and Urea as Denaturing Agents.—Tanford[3] has previously suggested that the ability of ethanol to denature proteins might arise from the effect of this agent on the transfer free energy of hydrophobic side chains from protein interior to solvent. The experimental data and the theoretical analysis which has preceeded affords an opportunity of checking this interesting hypothesis.

If it is assumed that all terms in eq. 11 except the side-chain transfer terms are independent of the concentration of ethanol in the solvent, then the lowering of $\Delta F°$ which results upon the addition of ethanol to the solvent may be expressed as

$$\Delta(\Delta F°)_{EtOH} = -p\sum_{I-IV}(\Delta f^i_{tr} - \Delta f^i_{tr}(c)) \quad (14)$$

where Δf^i_{tr} is again the transfer free energy of side chain i from reference solvent to water, and where $\Delta f^i_{tr}(c)$ is the transfer free energy from reference solvent to the aqueous ethanol solvent containing c% ethanol. This may be written more conveniently as

$$\Delta(\Delta F°)_{EtOH} = -p\sum_{I-IV} r\Delta f^i_{tr} \quad (15)$$

where r is the relative lowering of the free energy, $(\Delta f^i_{tr} - \Delta f^i_{tr}(c))/\Delta f^i_{tr}$.

The existing solubility data show that, to a fair approximation, r for any side chain may be expressed in terms of an equation involving the first two powers of the ethanol concentration so that r/c will be a linear function of the ethanol concentration, *i.e.*

$$\frac{r}{c} = a + bc \quad (16)$$

Typical values of r/c are plotted in Fig. 5 for side chains in group II (norleucyl and leucyl) and for side chains in group III (valyl). It can be seen that, at any particular temperature, the values of r/c are relatively independent of the nature of the side chain being considered. The only marked discrepancy is the valyl point at 0° and 25% ethanol which lies somewhat above the points for leucyl and norleucyl. We shall then obtain an approximate estimate of $\Delta(\Delta F°)_{EtOH}$ by assuming that the r values for all hydrophobic side chains are the same at any particular temperature. Although

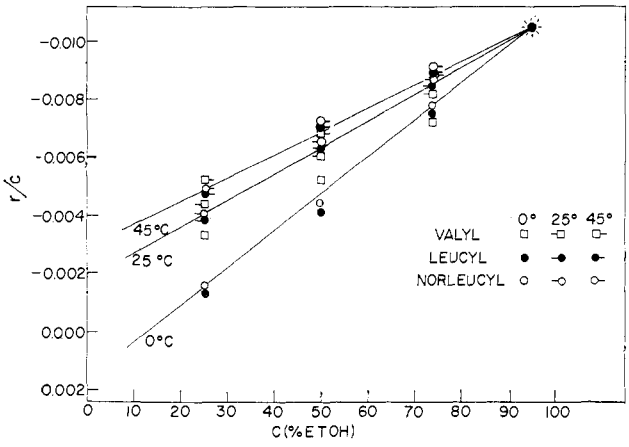

Fig. 5.—Relative lowering of the transfer free energies of hydrophobic side chains by ethanol addition; see text for details.

accurate data on group I and group IV side chains are not available, it should not introduce large errors if we assume that they have similar r values since these two groups combined account for only 15% of the hydrophobic character of chymotrypsinogen.

The r values may then be obtained from eq. 16 and the slopes and intercepts of the plots in Fig. 5. The values of a and b are 1.7×10^{-3} and -13×10^{-5}, respectively, at 0°; -1.4×10^{-3} and -9.5×10^{-5} at 25°; -2.7×10^{-3} and -8.4×10^{-5} at 45°.

The marked temperature dependence of the r factor arises from a large increase in the entropy and enthalpy of transfer of side chains as the ethanol concentration is increased. For the norleucyl side chain at 0°, for instance, the entropy of transfer to water is −17 e.u. and this increases to −10 e.u. for transfer to 25% ethanol and to −3 e.u. for transfer to 50% ethanol. This is undoubtedly an indication that the solvation sphere of the side chain is becoming considerably less ordered as the ethanol concentration is increased, and the significance of this will be considered shortly.

The calculated values of $\Delta(\Delta F^\circ)_{EtOH}$ for chymotrypsinogen have been obtained from eq. 15 and are shown in Fig. 6. This indicates more clearly the expected effect of temperature on ethanol action. At 45°, the addition of 20% ethanol to an aqueous solution of chymotrypsinogen (at constant pH and ionic strength) should lower the value of ΔF° by −12,000 cal., while the addition of the same amount of ethanol at 0° produces a decrease in ΔF° of only −2000 cal. This large temperature coefficient should be readily apparent in the experimental values if the predominant effect of ethanol arises from the hydrophobic transfer term as we have supposed.

The dashed line at 39° shows the experimental curve obtained from a number of measurements discussed in the preceding paper. The agreement with the calculated values is seen to be fairly good, particularly at ethanol concentrations below 10% where the experimental curve at 39° lies between the calculated curves at 45 and 25°. At higher concentration, the experimental curve falls somewhat below the expected values but the over-all agreement is still reasonably good.

The lower dashed curve at 0° indicates the limiting experimental curve from a single experiment which enabled us to conclude that the value of $\Delta(\Delta F^\circ)_{EtOH}$

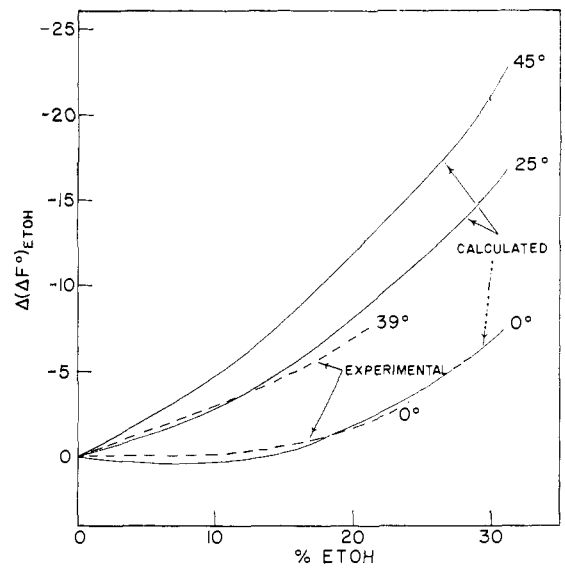

Fig. 6.—The calculated and experimentally determined effect of ethanol on chymotrypsinogen stability.

at 20% ethanol is more positive than −2000 cal. This curve then indicates the line of maximum possible slope, and the actual curve could lie considerably closer to the zero line. Although this 0° curve cannot be considered quantitative, it is adequate to illustrate that the ability of ethanol to denature chymotrypsinogen is considerably suppressed at low temperature, and this provides very strong support for the contention that the effect of ethanol on chymotrypsinogen arises from changes in the transfer free energies of the hydrophobic side chains as Tanford suggested. More complete studies on the effect of temperature on ethanol denaturation of chymotrypsinogen are now underway in this laboratory, and they tend to support this conclusion in somewhat more detail.

These results have some pertinence as to the precise mechanism by which ethanol may weaken hydrophobic bonds. As mentioned earlier, the transfer of hydrophobic side chains from protein interior to water results in a large negative entropy change which suggests that clathrating occurs about these side chains in the solvated state (*i.e.*, in the denatured state). This clathrating is particularly important at low temperature. As ethanol is added to the solvent, the solvation entropy becomes considerably less negative suggesting a disruption of clathrates. The action of ethanol may be correlated with its ability to disrupt these clathrate structures. At low temperature where the clathrate structures are most stable, ethanol is only a very weak denaturing agent. At high temperature, where thermal instability of the clathrates becomes more important, ethanol becomes better able to disorganize the clathrates and provide more favorable accommodation of the hydrophobic side chains in the solvent phase. This leads to an exceedingly large cooperative effect between ethanol and temperature in facilitating the cooperative unfolding of the chymotrypsinogen molecule.

The mechanism of action of urea appears to be more complicated. Simpson and Kauzmann[17] reported earlier that the activation free energy for urea denatura-

(17) R. B. Simpson and W. Kauzmann, *J. Am. Chem. Soc.*, **75**, 5139 (1953)

tion of ovalbumin varied linearly with the logarithm of the urea concentration at urea concentrations above 4 M. This fact suggests that strong binding of urea molecules is taking place. Their results have been interpreted in terms of a direct binding of urea molecules to a number of identical sites on the protein. This would be the denaturation mechanism expected if urea acts exclusively as a peptide hydrogen bond breaking agent.

A simple mechanism such as this will not account for the present thermodynamic results on chymotrypsinogen at low urea concentrations from 0 to 3.2 M. If it is assumed that reversibly denatured chymotrypsinogen has n identical sites for urea binding which are unavailable for binding in the native protein, then it can be shown quite easily that the change in the standard free energy of denaturation when urea is added to an aqueous solution of the protein should obey the expression

$$\Delta(\Delta F^\circ)_{\text{urea}} = -RT \ln [1 + K_u(\text{urea})]^n$$

where K_u is the binding constant of urea to a single site on the protein and where (urea) is the concentration of urea. In the present instance, no single values of K_u and n exist which will adequately represent the experimental results shown in Fig. 12 of paper I. This simple binding mechanism is, therefore, inadequate to describe the effect of urea on chymotrypsinogen and a more complex mechanism must be sought.

Kauzmann[6] also suggested that urea may denature proteins in part by weakening hydrophobic interactions just as we have previously found to be the case for ethanol. Since that time Wetlaufer, *et al.*,[18] and Nozaki and Tanford[19] have attempted to treat this problem quantitatively by studying the effect of urea on the free energy of transfer of hydrophobic side chains. If one were to attempt to interpret the effect of urea on chymotrypsinogen as resulting completely from changes in the side-chain transfer term, a rather obvious inconsistency arises regarding the concentration dependence. The data of Nozaki and Tanford show clearly that the effect of urea on the solubilization of nonpolar side-chains decreases as the urea concentration is increased. Thus, the transfer free energy of all hydrophobic side chains (except alanine) from water to 4 M urea is considerably less than twice as large as for transfer from water to 2 M urea. Our experimental results show, however, that the change in ΔF° produced by a given increment in urea concentration is smaller at low urea concentration than at high urea concentration. Consequently, a quantitative comparison between experiment and theory would show rather poor agreement in this case. Furthermore, Nozaki and Tanford have estimated also the effect of urea on the transfer of the peptide portion of amino acid residues and they conclude that this varies approximately linearly with urea concentration. Thus, even if both the side-chain and backbone contributions were included in the calculated estimate, the agreement with experiment would be rather poor as the urea concentration is varied.

(18) D. B. Wetlaufer, S. K. Malik, L. Stoller, and R. L. Coffin, *J. Am. Chem. Soc.*, **86**, 508 (1964).

(19) Y. Nozaki and C. Tanford, *J. Biol. Chem.*, **238**, 4074 (1963).

The effect of temperature on the ability of urea to denature chymotrypsinogen does provide some suggestive evidence that hydrophobic bonding is involved. It was found that the value of $\Delta(\Delta F^\circ)_{\text{urea}}$ is about twice as large at 30° as at 0° for the same urea concentration (2.3 M). The studies of Wetlaufer, *et al.*, show that the ability of urea to solubilize small hydrocarbons increases quite markedly with increasing temperature. This indicates, as was the case for ethanol, that the solvation entropy of these nonpolar groups is much less negative in aqueous urea solutions than in water. Thus, the strong cooperative effect of temperature and urea found experimentally is similar to to what would be expected if the action of urea arises, at least in part, from clathrate disruption and a more favorable accommodation of hydrophobic side chains in the solvent phase.

Discussion

The preceding analysis shows that the stability of chymotrypsinogen is, in most cases, quite clearly consistent with preconceived notions regarding the strength of force interactions which have arisen rather indirectly from model compound studies. This represents the first instance in which detailed consistency has been demonstrated by direct studies on proteins. Although no strong arguments can be made for the quantitative correctness of the model, the qualitative picture is consistent, and reasonable and further studies should help to establish reliability.

The method used for the separation of the contributions from hydrophobic bonding from other contributions such as hydrogen bonding, etc., is rather arbitrary and is precisely valid only if all other contributions arise from temperature independent enthalpic and entropic factors. This cannot be absolutely correct since the conformational entropy term as well as the hydrogen bond enthalpy term will in reality show at least a small temperature dependence. However, the heat capacity effects associated with clathrating are so very large, as indicated by model compound solubilities, that small temperature dependencies of other terms will not seriously obscure the results. For instance, the total enthalpy change *per residue*, $\Delta H^\circ/p\text{N}$, will change from about -200 cal. at 0° to $+2400$ cal. at 100° for chymotrypsinogen denaturation. The total entropy change per residue, $\Delta S^\circ/p\text{N}$, changes from -1 to $+7$ e.u. over the same temperature range. It seems highly unlikely that a significant amount of these changes results from a temperature dependence of hydrogen bond enthalpy or conformational entropy since the total temperature variation is itself much greater than the absolute contribution which we could reasonably expect from these terms at any temperature.

The author feels that the large heat capacity effects, which have gone unnoticed in many previous studies of denaturation reactions, offer perhaps the most direct evidence to date implicating hydrophobic bonding as an important determinant of protein structure in solution. Other models of nonpolar group participation do not appear as consistent with the experimental results as does the hydrophobic bonding picture. In particular, Klotz[20] has presented evidence which suggests that nonpolar side chains in *native* proteins may

(20) I. M. Klotz, *Science*, **128**, 815 (1958).

be surrounded by a sheath of "frozen" ordered water and that one important consequence of denaturation is that this frozen water is disrupted. If this were the case and if any melting of the frozen water occurs as the temperature is increased, as would be anticipated, we would also expect to see a denaturation reaction characterized by a large heat capacity effect, but in this case the heat capacity change during denaturation should be negative, in contrast to the large positive value found experimentally. In addition, it has been observed that denaturation occurs with a negative ΔS° at low temperature. This means that ΔS° must contain a large negative contribution which, at low temperature, is sufficient to cancel the large positive conformational entropy contribution. However, if the solvent is ordered to a great extent in the native state, as in the Klotz picture, then this should produce an additional large positive entropy contribution which will add to the positive conformational entropy contribution, and it becomes difficult to decide just how denaturation might ever occur with a negative ΔS°.

A positive ΔC_p for any reaction will necessarily lead to a maximum of ΔF° at some particular temperature. For chymotrypsinogen, this stability maximum occurs at a low temperature, about 10°. Evidence has been presented showing that there may be a number of other proteins exhibiting stability maxima at similar temperatures. These will occur at the temperature where ΔS° is zero. In the present picture ΔS° is the sum of a large positive contribution from conformational entropy and a large negative and temperature-dependent entropy contribution arising from clathrating about exposed hydrophobic groups. Consequently, those proteins having a large number of hydrophobic interactions which are solvated during the unfolding process would be expected to have maximum stability at a higher temperature than those proteins which exhibit less hydrophobic character in their transition. Chymotrypsinogen has an average or slightly greater than average content or large hydrophobic side chains so it would be expected that many proteins will show maximum stability below 10° and possibly even below 0° where it will not be observable.

Recent studies of protein stability have tended to place major emphasis on hydrophobic bonding, in contrast to the classical emphasis on hydrogen bonding. In particular, Tanford has suggested[3] that hydrophobic bonds by themselves may be able to provide sufficient free energy to overcome the conformational entropy associated with folding. If hydrophobic bonding were the only important source of stabilization of globular proteins, however, we should expect that denaturation would occur with negative values of ΔH° at all temperatures below 50° because of the predominance of the heat effects associated with clathrating. This is in contradiction to data on any number of denaturation reactions. Our studies show that in order to adequately account for the free energy, enthalpy, entropy, and heat capacity changes during unfolding, one must incorporate some large source of enthalpy stabilization into the model, and in the present picture this arises from a large number of weak intramolecular hydrogen bonds. Alternative interpretations of the enthalpy stabilization are possible, however. Tanford suggested that large enthalpy contributions may arise from a "conformational enthalpy" term associated with potential energy barriers to free rotation in the unfolded polypeptide chain. Such contributions will unquestionably exist, but their magnitude cannot be properly estimated. We prefer to consider this contribution as being inseparable from the direct contribution from hydrogen bond breaking since both contributions are enthalpic and both are a necessary consequence of hydrogen bond rupture in the polypeptide backbone. In this way, the Δh_H term from protein studies should be directly comparable with the enthalpy values obtained from model peptide group interactions since these must also contain contributions from both factors.

The description of native chymotrypsinogen which emerges from this study is that of a well-folded protein containing considerably more than 50% of its residues in the folded state. The native conformation is stabilized in about equal amounts by hydrogen bonds and by hydrophobic bonds and is unstabilized principally by a large conformational entropy factor which nearly balances the total stabilizing free energy under any set of solution conditions. This thermodynamic picture favors the idea that the reversible conformational transition constitutes a gross rearrangement of the polypeptide chain which involves the majority of all residues in the protein. If this is the case, we are still faced with the problem of rationalizing the fact that the transition occurs with only nominal changes in optical rotation. Since the preliminary X-ray model of native chymotrypsinogen[21] shows that this protein contains very little helix, we may tentatively assume that this denaturation involves a transition from a conformation in which the polypeptide chain is arranged in a rigid, amorphous fashion to a conformation in which the chain is flexibly random. Such a situation should lead to a great deal of cancellation of positive and negative rotatory contributions with the result that the net change in optical rotation and optical rotatory dispersion parameters does not provide an adequate measure of the extent of unfolding of the polypeptide chain.

In an earlier paper,[1] two alternative descriptions of the chymotrypsinogen denaturation were presented. One description assumes rather complete solvation of the unfolded protein while the other emphasizes only a limited penetration of solvent into the protein domain. We have concentrated on the first description in this paper. It might be argued that the small change in viscosity precludes any dramatic unfolding reaction such as we have proposed here. The 75% increase in reduced viscosity of chymotrypsinogen during thermal denaturation in water is rather small compared to the fivefold increases commonly observed for urea denaturations. When thermally denatured chymotrypsinogen is transferred to 7 M urea, for instance, a further increase of 100% in reduced viscosity occurs. This may be an indication that the denatured form in water is not closely akin to a random chain. It should be noted that if the domain of the denatured protein is sufficiently small, there is the possibility that the solvation of the hydrophobic side chains will be incomplete so that the model compound analogy with transfer reactions will not be valid. The arguments

(21) J. Kraut, *Proc. Natl. Acad. Sci. U. S.*, **48**, 1417 (1962).

pertaining to the relationship between changes in physical properties and conformational changes are rather involved, however, and the reader is referred to the original paper. They are, at any rate, inconclusive and cannot be used to eliminate or substantiate the present picture. They do serve to emphasize the tentative nature of the conclusions presented here and to point out the need for more extensive studies.

Acknowledgment.—The author wishes to acknowledge financial assistance from the National Institutes of Health (Grant GM 11071) during the course of this investigation.

[CONTRIBUTION NO. 3110 FROM THE GATES, CRELLIN, AND CHURCH LABORATORIES OF CHEMISTRY, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA]

Studies of the Chemistry of Mercury in Aqueous Solution. I. Mercury(I) and Mercury(II) Complexes of Aniline[1]

BY THOMAS H. WIRTH AND NORMAN DAVIDSON

RECEIVED MAY 27, 1964

The formation constant of the 1:1 complex of Hg(II) with aniline in aqueous $NaClO_4$–$HClO_4$ solution has been determined spectrophotometrically. Potentiometric study of this system revealed that in addition to 1:1 and 2:1 complexes of aniline with Hg(II), a 1:1 aniline–Hg(I) complex exists. The ultraviolet absorption spectra of aniline–Hg(I) solutions confirmed the existence of the latter. This is the first confirmed case of a soluble mercurous complex with a nitrogen ligand. The equilibrium constants at 27°, μ = 1.0 M ($NaClO_4$), are: $C_6H_5NH_3^+ + Hg^{+2} = C_6H_5NH_2Hg^{+2} + H^+$, $K_1 = 0.68$; $C_6H_5NH_3^+ + Hg_2^{+2} = C_6H_5NH_2Hg_2^{+2} + H^+$, $K_1^1 = 0.09$; and $2C_6H_5NH_3^+ + Hg^{+2} = (C_6H_5NH_2)_2Hg^{+2} + 2H^+$, $K_2 = 0.45$.

Introduction

Mercury(II) shows an extremely strong affinity for organic nitrogen ligands; in aqueous solution the Hg(II) ion forms stronger complexes with nitrogen bases than do the transition metal ions or even the other ions of the d^{10} electronic configuration. Studies of the binding of Hg(II) by a variety of heterocyclic and aromatic nitrogen ligands have been carried out in this laboratory.[2-4] Interest in the topic was stimulated in part by the role of Hg(II) in the study of biological macromolecules which contain basic nitrogen atoms such as the nucleic acids. That work showed that the interaction of Hg(II) with such ligands is neither straightforward nor easily characterized.

In the present work, the interaction of Hg(II) with the simplest aromatic amine, aniline, was studied to determine whether Hg(II) exhibits an unusually strong affinity for such compounds. Although aqueous, acidic Hg(II)–aniline solutions are not stable with time because of aromatic mercuration, it was possible to measure the stability constants of the labile $C_6H_5NH_2Hg^{+2}$ and $(C_6H_5NH_2)_2Hg^{+2}$ complexes by working rapidly and using appropriate extrapolations. The stability of the 2:1 complex was determined from the potential of a Hg(0) electrode in acidic solutions containing aniline in excess. However, unless the Hg^{+2} ion is strongly complexed, a large fraction of the total Hg(II) reacts with the electrode to form Hg_2^{+2}.[5] Therefore the 1:1 Hg(II)–aniline complex was studied by spectrophotometry in the presence of excess Hg^{+2}. In the potentiometric study, evidence for a Hg(I)–aniline complex, $C_6H_5NH_2Hg_2^{+2}$, was obtained. Independent spectrophotometric evidence confirmed the existence of this soluble mercurous–nitrogen complex.

(1) This work is described in more detail in the Ph.D. Thesis of T. H. W., California Institute of Technology, 1964.

(2) P. Brooks and N. Davidson, *J. Am. Chem. Soc.*, **82**, 2118 (1960).

(3) R. Ferreira, E. Ben-Zvi, T. Yamane, J. Vasilevskis, and N. Davidson, "Advances in the Chemistry of Coordination Compounds," S. Kirschner, Ed., Macmillan, New York, N. Y., 1961, pp. 457–462.

(4) J. Vasilevskis, Ph.D. Thesis, California Institute of Technology, 1963.

(5) S. Hietanen and L. G. Sillén, *Arkiv Kemi*, **10**, 103 (1956).

Experimental

Materials.—All solutions were prepared from redistilled water. Perchloric acid, nitric acid, and sodium nitrate solutions were prepared from reagent grade chemicals. Carbonate-free sodium hydroxide was prepared from a 50% NaOH solution. Acids and bases were standardized by conventional methods. Mercuric perchlorate was prepared from weighed quantities of HgO and an excess of standardized 60% perchloric acid; filtration through sintered glass removed the slight residue. The Hg(II) concentration was determined by Volhard titration with KSCN using $FeNH_4(SO_4)_2$ indicator. The concentration of excess acid was determined by titration with standard base in the presence of a large excess of NaCl, using a glass electrode as indicator.

Sodium perchlorate was prepared by adding 50% NaOH to aliquots of standardized 60% $HClO_4$ until neutral when tested with pH paper, then diluting to volume. When necessary, $NaClO_4$ solutions were carefully titrated for excess acid or base using a glass electrode indicator and a $NaNO_3$ salt bridge.

Aniline perchlorate and nitrate stock solutions for the potentiometric experiments were prepared by distilling aniline directly into an aqueous solution containing a known amount of the appropriate acid. The receiver was stirred continuously, and a nitrogen atmosphere was maintained; ultraviolet radiation was excluded. The amount of aniline added was computed from the weight of the receiver before and after the distillation. If stored in the dark, these acidic solutions are stable over long periods. The aniline perchlorate stock solution for spectrophotometric experiments was prepared from the solid compound. Reagent grade aniline (J. T. Baker) was added to excess 30% $HClO_4$, cooled, and filtered. The crude solid was precipitated from acetone solution with chloroform four times in succession and dried. A weighed sample of the pure, white product was dissolved in dilute perchloric acid and made up to known volume.

Mercurous perchlorate solutions were prepared by treating $Hg(ClO_4)_2$ solutions with metallic mercury. A mercurous nitrate solution was prepared from the reagent grade salt (Baker and Adamson). The acidic solutions were analyzed for total Hg by titration with KSCN after oxidation with $KMnO_4$ and removal of brown MnO_2 with aqueous $FeSO_4$. Analysis for the traces of Hg(II) in the Hg(I) solutions was accomplished potentiometrically with a Pt electrode.[6]

The Potentiometric Method.—A 200-ml. round-bottom flask with an inlet for nitrogen gas contained the experimental solutions. It was immersed in a constant-temperature bath, which was equipped with an underwater magnetic stirrer. The Hg^{+2} concentration was determined from the potential of a J-type mercury electrode *vs.* a saturated calomel reference elec-

(6) A. Jonsson, I. Qvarfort, and L. G. Sillén, *Acta Chem. Scand.*, **1**, 461 (1947).