Biochemistry

 Volume 17, Number 15 July 25, 1978

Role of Hydrophobicity in the Binding of Coenzymes†

Joel Janin* and Cyrus Chothia*

Appendix: Translational and Rotational Contribution to the Free Energy of Dissociation

Joel Janin

ABSTRACT: We calculate the loss of surface area accessible to solvent associated with coenzyme binding in *Clostridium* flavodoxin, in dogfish lactate dehydrogenase, and in lobster glyceraldehyde-3-phosphate dehydrogenase. The coenzymes are nearly buried in the complexes and lose on the order of 600 Å^2, while the proteins lose a similar amount of accessible surface area. Some of the loss can be attributed to conformation changes in the protein, at least in the case of lactate dehydrogenase, where we show that the apoenzyme has a larger accessible surface area than the holoenzyme. Using known correlations with the hydrophobic contribution to the free energy, we demonstrate that hydrophobicity is the major source of stabilization free energy in FMN binding to flavodoxin and in NAD binding to the two dehydrogenases: it contributes 25 to 30 kcal/mol to the free energy of dissociation, more than required in order to compensate for the loss of six degrees of translational/rotational freedom by the coenzyme.

Hydrophobic, electrostatic, and van der Waals forces are involved in all the various types of interactions made by the polypeptide chain of a protein: interactions with itself to fold into a globular structure; association with other chains to form multisubunit complexes; and the binding of small ligands. Thus when structural data are available from x-ray crystallography, the geometrical arrangement of polar atoms shows the presence of intra- and intermolecular hydrogen bonds and charge interactions and the volume of the Voronoi polyhedron around each atom describes the atomic packing (Richards, 1974; Chothia & Janin, 1975). The role of hydrophobicity can be assessed using the concept of accessible surface area (Lee & Richards, 1971). For a given protein atom this is the area of the surface over which the center of a water molecule can be placed while it is in van der Waals contact with the atom and not penetrating any other protein atom. Each square angström of protein accessible surface that is removed from contact with the solvent gives a hydrophobic free energy of 25 cal (Chothia, 1974). How do these different forces create the specific strong bonds that are essential for biological systems? From an analysis of the structure of the interfaces that occur between protein monomers, we concluded that hydrophobicity is the major force stabilizing protein–protein association; van der Waals forces and hydrogen bonds (i.e., complementarity) play a selective role in that, while they contribute little to the stability of correct associations, they prevent incorrect associations by large unfavourable enthalpies (Chothia & Janin, 1975; Janin & Chothia, 1976; Chothia et al., 1976).

In this paper we consider the thermodynamics of the binding of small molecules by a protein. The description given above for the thermodynamics of protein–protein association cannot be simply extended to the association of proteins with small molecules. The loss of translational and rotational entropy is the major term unfavorable to association. This loss of entropy is proportional to the logarithm of the molecular weights and so does not vary much with their size: for the associations of two proteins, it is 23–30 kcal/mol and for the binding of a small molecule by a protein 17–22 kcal/mol (see Appendix). In the cases that we have studied, oligomeric proteins compensate for this loss by burying at their interface ~1400 Å^2 of accessible surface. This releases ~35 kcal/mol of hydrophobic free energy (Chothia & Janin, 1975). However, accessible surface area is proportional to the two-thirds power of the molecular weight and so for small molecules it will only be a few hundred square angströms. Therefore it is not clear how, on binding to a protein, small molecules can bury sufficient surface for hydrophobic energy to overcome the loss of their translational and rotational entropy.

Here we examine this problem by analyzing the structure of three protein–coenzyme complexes: that of dogfish lactate dehydrogenase (LDH)[1] and lobster glyceraldehyde-3-phos-

* From the Unité de Biochimie Cellulaire du Département de Biochimie et Génétique Microbienne, Institut Pasteur, 75724 Paris, Cedex 15, France (J.J. and C.C.), and from the Department of Chemistry, University College London, London, WC1, England (C.C.). *Received November 2, 1977.*

[1] Abbreviations used: LDH, lactate dehydrogenase; GPDH, D-glyceraldehyde-3-phosphate dehydrogenase; NAD, nicotinamide adenine dinucleotide; FMN, flavin mononucleotide.

TABLE I: Surface Area Buried in Protein-Nucleotide Complexes.

	coenzyme accessible surface area[a] (Å²)		surface area buried in complex[b] (Å²)		
	free	bound	coenzyme	protein	total
flavodoxin	637	145	492	244	736
holo-LDH	850	75	775	447	1222
holo-GPDH[c]	817	186	631	433	1064

[a] Accessible surface area of FMN or of NAD calculated for free or bound coenzymes in the same conformation. NAD has a different conformation in LDH and in GPDH (Rossmann et al., 1975) so that its "free" accessible surface area is slightly different. [b] Accessible surface area of the coenzyme (resp. protein) alone, minus that of the coenzyme (resp. protein) in the complex, assuming the conformation to be the same. [c] Average of the "green" and "red" subunits, the value being nearly identical. The calculation is made for isolated subunits, not for the tetrameric enzyme. The accessible surface area of bound NAD is actually slightly smaller in the tetramer, due to an interaction made with a neighbouring subunit (Moras et al., 1975).

phate dehydrogenase (GPDH), both of which bind nicotinamide adenine dinucleotide (NAD), and *Clostridium* flavodoxin which binds flavin mononucleotide (FMN). Descriptions of the atomic structure of the holo form of all three proteins and the apo form of LDH have been published (Smiley et al., 1971; Buehner et al., 1974; Burnett et al., 1974; Adams et al., 1973). The holo and apo forms of LDH differ in conformation. Residues 98–120 form a loop between a strand of β-pleated sheet and an α helix. In the apo form this loop sticks out into solution while in the holo form it folds down to cover the NAD in the active site pocket (White et al., 1976).

Methods

The Atomic Coordinates. The atomic structure of dogfish apo-LDH (M4 coenzyme) has been determined by Adams et al. (1973), that of the ternary complex of LDH with NAD and pyruvate (here called holo-LDH) by Smiley et al. (1971), that of lobster holo-GPDH by Buehner et al. (1974), and that of oxidized *Clostridium* flavodoxin by Burnett et al. (1974). The atomic coordinates used here for apo-LDH, holo-LDH, and flavodoxin are from the Cambridge Data Bank. GPDH atomic coordinates are a gift of Dr. Rossmann; they include atomic positions for two subunits of the tetrameric enzyme, labeled "red" and "green" and related by a molecular twofold symmetry (Moras et al., 1975). We submitted the four structures to 64 cycles of energy refinement (Levitt, 1974), using a program of M. Levitt. In each case, the root mean square atomic movement remains within the experimental errors of the atomic coordinates: it is 0.17 Å for flavodoxin, and 0.3 Å for LDH and GPDH, the structures of which have been studied at a lower resolution than for flavodoxin.

Accessible Surface Area Calculations. Accessible surface areas were calculated from the atomic coordinates following the procedure of Lee & Richards (1971) and using a computer program written by Dr. M. Levitt. Accessible surface areas were calculated for the holo proteins, for the holo proteins with the coenzymes removed, for the isolated coenzymes, and for apo-LDH.

Results

The Accessible Surface Area of the Proteins. The accessible surface area of flavodoxin is 6850 Å². We have shown that the accessible surface area A of all *monomeric* globular proteins is a simple function of its molecular weight M (Chothia, 1975).

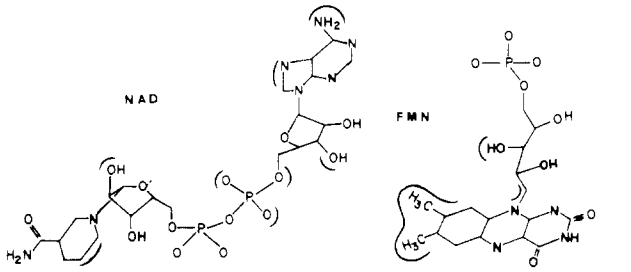

FIGURE 1: Accessibility of FMN in flavodoxin and of NAD in LDH. The coenzymes are drawn as they appear in their bound form. Chemical groups which remain accessible to solvent are outlined.

The correlation can be written (Janin, 1976; Teller, 1976):

$$A = 11.1M^{2/3}$$

which yields A = 6750 Å² for flavodoxin (M = 15 000), in excellent agreement with the value given above.

The formula predicts accessible surface areas of about 12 000 Å² for molecules with the molecular weight of the subunits of LDH and GPDH (36 000). The values calculated from the atomic coordinates are significantly larger than this: 17 700 Å² for apo-LDH, 16 900 Å² for holo-LDH, and 15 400 Å² for GPDH. However, most of this excess surface area must be buried in the subunit interfaces which are quite extensive in the two tetrameric dehydrogenases. Nearly the same value is obtained for apo- and for holo-LDH, and also for the "red" and "green" subunits of lobster GPDH, despite possible differences in conformation and random errors in the x-ray data. This indicates that the calculation of accessible surface areas from three-dimensional structures is not sensitive to localized conformation changes and to small atomic movements.

The Accessible Surface Area of the Coenzymes. We calculated the accessible surface area of FMN and of NAD in the conformation they have when bound to the proteins. Table I shows that this area is much smaller in the complexes than in the free nucleotides: the coenzymes are almost completely shielded from contact with the solvent and buried inside the protein. Figure 1 indicates which chemical groups remain accessible. In flavodoxin, the two methyl groups on the flavin ring of FMN are in contact with water. They provide most of its residual accessible surface area (103 Å² out of 145 Å²). The ribityl chain and the phosphate group are buried. For NAD bound to LDH or to GPDH, one edge of the adenine ring (with the N^6H_2 and the N^7 polar groups) provides nearly one-half of the surface area of the coenzyme which remains accessible to water. The nicotinamide and its attached ribose are buried in the active site.

The Protein Accessible Surface Area Buried by the Coenzymes. In order to calculate the protein surface area which is shielded from the solvent due to the presence of the coenzyme, we simply removed the latter from the sets of atomic coordinates used for the accessibility calculations. This technique provides a convenient way of finding which chemical groups on the protein side are in contact with the coenzyme: their accessibility changes when it is removed from the set of atomic coordinates while the rest of the molecule is not affected. Thus, Figure 2 shows that, in flavodoxin, about 15 residues are involved. The largest changes in accessible surface area affect Met-56 and Trp-90, the side chains of which are in the immediate vicinity of the flavin ring (Burnett et al., 1974; Mayhew & Ludwig, 1975). In LDH and in GPDH, NAD decreases the accessibility of about 25 residues. The total changes in accessible surface area are given in Table I.

With a few minor exceptions, the amino acids which lose

accessible surface area are also observed to make hydrogen bonds and/or van der Waals contacts with the coenzymes (Mayhew & Ludwig, 1975; Holbrook et al., 1975; Moras et al., 1975). They are clustered on the polypeptide chain and occur on the carboxy-terminal end of four strands of the parallel β-pleated sheet which forms the core of flavodoxin and of the first domain of LDH and GPDH (Rao & Rossmann, 1973; Rossmann et al., 1975; Moras et al., 1975).

The difference between the protein conformations of apo- and holo-LDH involves the folding of the loop formed by residues 98–120 to cover the active site and the readjustment of the surrounding residues (White et al., 1976). The accessible surface area of apo-LDH (17 700 Å^2) is larger than that of holo-LDH with NAD removed (17 300 Å^2). This difference, 400 Å^2 or 2.5% of the subunit accessible surface area, may not be significant if considered by itself, but, counting only residues in contact with the coenzyme, their accessible surface area is also larger by 160 Å^2, or 15%, in apo-LDH than in holo-LDH with NAD removed, and the surface of the loop 98–120 is larger by 80 Å^2. Thus the changes in the tertiary structure of LDH that occur when NAD is bound not only bury the coenzyme but also increase the amount of the protein surface that is buried within the molecule.

Discussion

Energy Needed to Bind the Coenzymes to the Proteins. When a molecule binds to a protein it loses its translational and rotational degrees of freedom. For the binding of a small molecule (molecular weight 100 to 1000), this loss is estimated to be 17–22 kcal/mol in the standard state (see Appendix). In addition the dissociation constants for the coenzymes considered here are lower than 5×10^{-4} M for NAD binding to LDH (Dalziel, 1975) and 10^{-8} M for FMN binding to flavodoxin (Mayhew & Ludwig, 1975). This corresponds to a free energy of dissociation of 5 and 11 kcal/mol. Thus the free energy required for the association of the proteins and their coenzymes is at least 27–33 kcal/mol. How do the various intermolecular interactions produce sufficient free energy to favor association? In the case of FMN bound to flavodoxin and of NAD bound to dehydrogenases, hydrogen bonds are made (Mayhew & Ludwig, 1975; Rossmann et al., 1975), and also possibly favorable van der Waals contacts. But the contribution of these interactions to the free energy of dissociation is limited, because the chemical groups involved certainly make interactions of similar strength with water or solute ions when the complexes are dissociated. The balance is difficult to estimate, but it can hardly exceed 10 kcal/mol in each case (for about 12 H bonds between FMN and flavodoxin or 10 H bonds between NAD and LDH). On the other hand, the reduction of surface area accessible to solvent which occurs when the complexes form is a source of favorable free energy: about 25 cal/mol for each square angström in the case of amino acids (Chothia, 1974, 1975) and alkanes in water (Hermann, 1972; Reynolds et al., 1974; Franks & Eagland, 1975). In protein–protein complexes which bury 1500 Å^2 or more, this has been shown to account for the stability of the association (Chothia & Janin, 1975; Janin & Chothia, 1976).

Assuming that a hydrophobic energy of 25 cal/mol per Å^2 applies to the nucleotide parts of the complexes studied here as well as to the protein parts, the surface area buried upon binding of FMN or NAD must be about 1100 Å^2 for LDH and 1250 Å^2 for flavodoxin in order to compensate for a loss of 20 to 22 kcal/mol in translational and rotational free energy and for the observed stability of the complexes. The size of these areas could be reduced if the many H bonds made with the coenzymes were significantly more stable than H bonds made with water, but other factors contribute to make them larger by playing against the formation of the complexes: the loss of internal degrees of freedom within the coenzymes and in amino acid side chains, and the possible release of cations bound to the coenzyme in solution. Also, the correlation established between hydrophobicity and the accessible surface of amino acids may not be valid for the more polar parts of the nucleotides, and especially for the phosphate groups, though these represent only about 20% of the accessible surface area of FMN and of NAD.

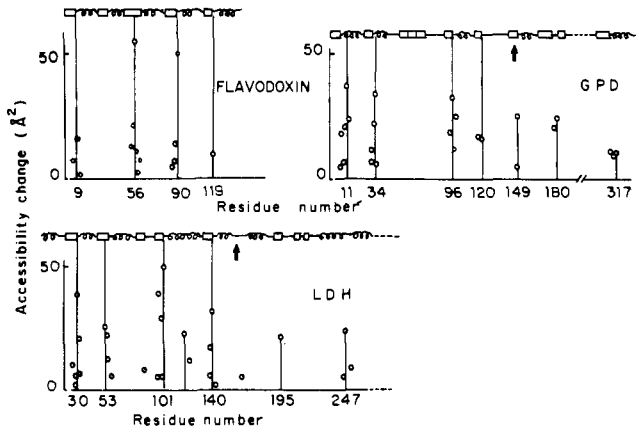

FIGURE 2: Accessibility change of amino acid residues. Accessible surface areas are calculated with and without the coenzyme present in the sets of atomic coordinates. The difference (square angstroms) is given for each amino acid residue (circles). Secondary structures are represented schematically as a succession of helices and of strands (boxed). They relate to the residue numbering as shown by the vertical lines. Arrows point to the limit between the NAD binding domain and the catalytic domain of LDH and of GPDH. Residues altering accessibility in flavodoxin are: Gly-8, Thr-9, Asn-11, Thr-12, Ser-54, Ala-55, Met-56, Gly-57*, Asp-58, Glu-59, Ser-87, Tyr-88*, Gly-89, Trp-90, Asn-119*; asterisked residues are not mentioned in Table 3 of Mayhew & Ludwig (1975), which lists residues in contact with FMN. Residues altering accessibility in LDH are: Val-27, Gly-28, Val-29, Gly-30, Ala-31, Val-32, Asp-53, Val-54, Met-55, Lys-58, Tyr-85, Thr-97, Ala-98, Gly-99, Ala-100*, Arg-101, Ile-119, Ile-123, Val-138, Ser-139, Asn-140, Val-142*, Ser-163*, His-195, Thr-246, Ser-247, Ile-250; asterisked residues are not mentioned in Table X of Holbrook et al. (1975) which lists contacts with NAD in LDH. The residue numbering used here is that of Holbrook et al. (1975), but the amino acid sequence has been revised according to Eventoff et al. (1977). Residues altering accessibility in GPDH are: Asp-7, Gly-8, Phe-9, Gly-10, Arg-11, Ile-12, Asn-31, Asp-32, Pro-33*, Phe-34, Ile-35*, Ser-95*, Thr-96, Gly-97, Phe-99, Ser-119, Ala-120, Ser-148, Cys-149, Thr-179*, Ala-180, Asn-313, Glu-314*, Tyr-317; asterisked residues are not mentioned in Table VI of Moras et al. (1975) which lists contacts with NAD in GPDH. The accessibility changes quoted here are averages of the values in the "green" and "red" subunits of lobster GPDH.

The Hydrophobic Contribution to the Binding of NAD by LDH. Table I shows NAD on binding to LDH buries 775 Å^2 of its accessible surface. But this "buried surface area" is the change in surface area that occurs upon binding assuming that the dissociated components have the same conformation as found in the complex. This is certainly not correct. NMR evidence suggests a folded conformation for the dinucleotides in water (Oppenheimer et al., 1971), as opposed to the open conformation of NAD in LDH or GPDH (Rossmann et al., 1975). The adenine and nicotinamide bases would be stacked, which reduces significantly their accessibility. Assuming that this reduction is 50%, the accessible surface area of free NAD is about 650 Å^2 rather than 820–850 Å^2 with the bases wide apart. If NAD binds to LDH in the holo conformation the protein loses 450 Å^2 (Table I). In the results section above we noted that, when the tertiary structure of LDH changes from that of apo to that of holo, 160 Å^2 of protein accessible surface becomes buried. Taking these figures together, 1200 Å^2 is a

good estimate of the total loss of accessible surface area occurring when one NAD molecule binds to LDH. This is sufficient to account for the observed stability of the complex.

Tight binding of the coenzyme is, however, not the only requirement for efficient functioning of the dehydrogenase. Turnover rates of about 250 s^{-1} achieved at NAD or NADH concentrations of 4×10^{-4} M imply that the binding reaction is faster than 10^6 M^{-1} s^{-1}. Experimental data indicate that the binding rate reaches 6×10^7 M^{-1} s^{-1} for NADH (Dalziel, 1975) close to the limit set by diffusion of the coenzyme. The binding site must therefore be very accessible in apo-LDH, in accordance with its more open conformation and the larger accessible surface area found for the apo-LDH structure. Thus, a loose complex may be formed very fast, followed by the conformation change, involving the movement of the loop of residues 98–120 to cover the NAD molecule (White et al., 1976). This not only increases the buried surface area, and so stabilizes the complex, but also removes the catalytic region from contact with water.

The Binding of FMN by Flavodoxin. When FMN binds to the flavodoxin in the conformation observed for the holo complex, FMN loses 490 Å^2 of accessible surface and the protein 250 Å^2, 740 Å^2 in all (Table I). The preferred conformation of FMN in water is not known, but one may also suppose that some of the surface of the flavin ring is covered part of the time by the ribityl chain, and that its accessible surface area is a little less than 640 Å^2, the value observed with the ribityl chain extended and pointing away from the flavin ring (Figure 1).

In the molecule two loops that connect strands of β-pleated sheet to α helices (residues 54 to 59 and 87 to 92) close round the coenzyme like a pair of jaws and make most of the contacts that occur between the protein and its coenzyme. It seems unlikely that these loops will retain, in absence of FMN, the conformation observed in holoprotein, and probable that the binding site of apoflavodoxin will have an open structure like that found in apo-LDH.

Thus though the many H bonds made with FMN, and especially with its phosphate group (Ludwig & Mayhew, 1975), may contribute significantly to the free energy of dissociation, the loss of accessible surface we find by comparing the structure of the holoprotein with and without FMN is 740 Å^2, and the increase in this figure due to probable conformational changes related to coenzyme binding, mean that hydrophobicity makes at least a major contribution to the free energy of association.

The function of flavodoxin does not require the dissociation of FMN. It does, however, require the aromatic part of the coenzyme molecule (the xylene moiety of the flavin ring) to interact with other protein components of the electron transfer chain. This explains why the region of the FMN molecule which remains accessible to water when it binds to flavodoxin is the hydrophobic ring.

The Anticooperative Binding of NAD to GPDH. If NAD binds to a GPDH monomer in the holo conformation, it loses 630 Å^2 of accessible surface area, and the protein loses 430 Å^2, 1060 Å^2 in total (Table I). This figure is less than that for NAD binding to the apo tetramer for two reasons. First, two residues from a neighboring subunit (Leu-187 and Pro-188) form additional contacts with NAD (Moras et al., 1975). Second, though the detailed structure of apo-GPDH is not known, Biesecker et al. (1977) describe a change in the conformation of the S loop (residues 178–201) which makes contact with the coenzyme in the holoprotein, leading to a more open structure in the apoprotein. These remarks apply to the *B. stearothermophilus* enzyme, but the situation is likely to be the same in the lobster enzyme. As in the case of LDH, this suggests that the figure obtained for the accessible surface area buried on NAD binding is a lower value; still, this figure (1060 Å^2 in GPDH) shows clearly that hydrophobicity is the major force stabilizing the association with NAD.

NAD binding to muscle GPDH shows strong anticooperativity (Conway & Koshland, 1968), with two tight binding sites (dissociation constant about 10^{-8} M in the rabbit muscle enzyme; Dalziel, 1975) and two sites with an affinity comparable to that of LDH (about 10^{-4} M), which binds NAD normally. This large difference is not easily explained by the geometry of the NAD sites, or by the coenzyme environment which is similar in all four subunits of holo-GPDH (Olsen et al., 1976; Biesecker et al., 1977) even though the active site may show asymmetry in ternary complexes (Garavito et al., 1977). But the NAD sites are close to the subunit interfaces formed by the molecule's twofold axis R, and therefore one site is very close to the other one related to it by this axis (Buehner et al., 1974). The S loop of one subunit is in contact with the coenzyme on the symmetry-related subunit as well as with that of its own subunit (Moras et al., 1975; Biesecker et al., 1977). A change in conformation of the S loop associated with NAD binding in one subunit may therefore affect the neighboring coenzyme site on the other subunit.

We propose that the binding of NAD at the first of the two sites related by the R axis causes structural changes that bury more accessible surface area than is buried by the changes that occur when the second site is filled. If the accessible surface buried on binding the first NAD molecule is 300 Å^2 greater than that buried on binding the second, hydrophobicity would give an extra 7.5 kcal/mol of binding energy and a dissociation constant lower by a factor of 10^5. The argument, of course, applies to both the pairs of sites close to the R axis. Thus this model gives two sites of high affinity and two of low and implies that the apo and the final holo structures are symmetric.

Conclusion

We have seen that hydrophobicity provides the major source of stabilization free energy in these protein–coenzyme complexes, and it may also be important in modulating their stability for the purpose of regulation. A ligand of molecular weight 1000 must remove more than 900 Å^2 of accessible surface from contact with water in order to compensate for the loss of translational/rotational degrees of freedom. For NAD and FMN this is done by the protein burying most of the coenzyme in its binding site: about 600 Å^2 of coenzyme surface area and a comparable amount of protein surface become inaccessible to water. Molecules smaller than NAD cannot provide such a large reduction in surface area. Indeed, chemically removing part of the NAD molecule has a drastic effect on its affinity for dehydrogenases, while replacing the adenine base with quite different aromatic groups of comparable size causes much smaller changes (Bränden et al., 1975). This and the great variety of their interactions may have forced the evolution of "coenzymes" as rather large and complex molecules, a limited portion of which is actually useful to their chemical function as reactive groups.

How, in general, do proteins bind small molecules? Electrostatic interactions obviously play a role in binding inorganic ions and molecules and organic phosphate compounds. However, we believe in the general importance of the mechanism described here, whereby the conformational changes that occur on binding almost bury the ligand and reduce the protein's accessible surface thus releasing a large hydrophobic free energy.

References

Adams, M. J., Ford, G. C., Liljas, A., & Rossmann, M. G. (1973) *Biochem. Biophys. Res. Commun. 53,* 46–51.

Biesecker, G., Harris, J. I., Thierry, J. C., Walker, J. E., & Wonacott, A. J. (1977) *Nature (London) 266,* 328–333.

Brändén, C. J., Jörnvall, H., Ecklund, H., & Furugren, B. (1975) *Enzymes, 3rd Ed. 11,* 103–190.

Buehner, M., Ford, G. C., Moras, D., Olsen, K. W., & Rossmann, M. G. (1974) *J. Mol. Biol. 90,* 25–49.

Burnett, R. M., Darling, G. D., Kendall, D. S., Lequesne, M. E., Mayhew, S. G., Smith, W. W., & Ludwig, M. L. (1974) *J. Biol. Chem. 249,* 4383–4392.

Chothia, C. H. (1974) *Nature (London) 248,* 338–339.

Chothia, C. H. (1975) *Nature (London) 254,* 304–308.

Chothia, C. H., & Janin, J. (1975) *Nature (London) 256,* 705–708.

Chothia, C. H., Wodak, S., & Janin, J. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 3793–3797.

Conway, A., & Koshland, D. E., Jr. (1968) *Biochemistry 7,* 4011–4023.

Dalziel, K. (1975) *Enzymes, 3rd. Ed. 11,* 1–42.

Eventoff, W., Rossmann, M. G., Taylor, S. S., Torff, H. J., Meyer, H., Keil, W., & Kiltz, H. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74,* 2677–2681.

Franks, F., & Eagland, D. (1975) *CRC Crit. Rev. Biochem. 3,* 165–219.

Garavito, R. M., Berger, D., & Rossmann, M. G. (1977) *Biochemistry 16,* 4393–4398.

Hermann, R. B. (1972) *J. Phys. Chem. 76,* 2754–2759.

Holbrook, J. J., Liljas, A., Steindel, S. J., & Rossmann, M. G. (1975) *Enzymes, 3rd Ed. 11,* 191–292.

Janin, J. (1976) *J. Mol. Biol. 105,* 13–14.

Janin, J., & Chothia, C. H. (1976) *J. Mol. Biol. 100,* 197–211.

Lee, B. K., & Richards, F. M. (1971) *J. Mol. Biol. 55,* 379–400.

Levitt, M. (1974) *J. Mol. Biol. 82,* 393–420.

Mayhew, S. G., & Ludwig, M. (1975) *Enzymes, 3rd Ed. 12,* 57–118.

Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C., & Rossmann, M. G. (1975) *J. Biol. Chem. 250,* 9137–9162.

Olsen, K. W., Garavito, R. M., Sabesan, M. N., & Rossmann, M. G. (1976) *J. Mol. Biol. 107,* 577–584.

Oppenheimer, N. H., Arnold, L. J., & Kaplan, N. O. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68,* 3200–3205.

Rao, S. T., & Rossmann, M. G. (1973) *J. Mol. Biol. 76,* 241–256.

Reynolds, J. A., Gilbert, D. B., & Tanford, C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71,* 2925–2927.

Richards, F. M. (1974) *J. Mol. Biol. 82,* 1–14.

Rossmann, M. G., Liljas, A., Brändén, C. J., & Banaszak, L. J. (1975) *Enzymes, 3rd Ed. 11,* 61–102.

Smiley, I. E., Koekoek, R., Adams, M. J., & Rossmann, M. G. (1971) *J. Mol. Biol. 55,* 467–475.

Teller, D. C. (1976) *Nature (London) 260,* 729–731.

White, J. L., Hackert, M. L., Buehner, M., Adams, M. J., Ford, G. C., Leutz, P. J., Smiley, I. E., Steindel, S. J., & Rossmann, M. G. (1976) *J. Mol. Biol. 102,* 759–779.

Appendix: Translational and Rotational Contribution to the Free Energy of Dissociation

The free energy of dissociation of a molecular complex is related to its dissociation constant and to the equilibrium concentrations of the chemical species A, B, and AB by:

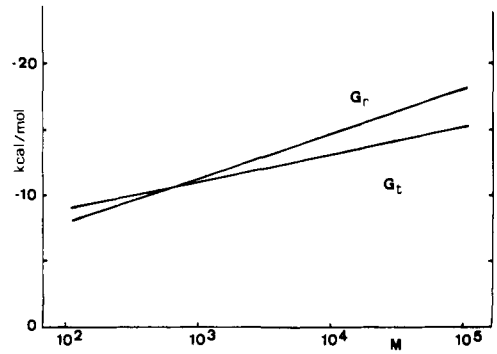

FIGURE 3: Calculated translational and rotational free energies. Free energies expressed as a function of molecular weight are calculated using eq 4 and 9, for spherical particles of partial specific volume $\bar{v} = 0.75$ cm^3/g in the standard state (1 mol per L) and at 300 K. They are negative. The translational and rotational enthalpies are 0.9 kcal/mol each, independent of molecular weight.

$$\Delta G_D = -RT \ln K_D = -RT \ln[A][B]/[AB] \quad (1)$$

For aqueous solutions, the concentrations and the value of K_D are expressed as molarities, the free energy of dissociation being zero if K_D is one molar. This does not imply that no free energy is involved, but that the balance of terms favoring and opposing the dissociation of AB is zero when A and B are molar. Among the terms favoring dissociation is the contribution of six degrees of translational and rotational freedom lost when a single molecule is formed from two. The corresponding free energy may be calculated from statistical thermodynamics, at least in the case of freely moving particles.

The translational contribution is derived from the partition function of a particle of mass m moving in a volume V:

$$Z_t = (2\pi mkT)^{3/2}V/h^3 \quad (2)$$

where k and h are Boltzmann's and Planck's constants (see for instance, eq 9-1 in Hill, 1960).

For N equivalent particles, the partition function is $Z_t{}^N/N!$ and the free energy:

$$G_t = -kT \ln(Z_t{}^N/N!) \quad (3)$$

At molar concentration, N is Avogadro's number and V is 1 L. Using Stirling's approximation for $N!$, we obtain

$$G_t = -RT(\ln \frac{Z_t}{N} + 1) = -RT \ln[\alpha(RT)^{3/2}M^{3/2}] \quad (4)$$

where $M = Nm$ is the molecular weight and $R = Nk$ is Rydberg's constant. The coefficient α evaluates to:

$$\alpha = 42.8V/N^4h^3 \quad (5)$$

The partition function for the rotational degrees of freedom on a polyatomic molecule is:

$$Z_r = 8\pi^2(2\pi kT)^{3/2}(I_1I_2I_3)^{1/2}/h^3 \quad (6)$$

where I_1, I_2, and I_3 are the principal moments of inertia of the molecule (see eq 9-7 in Hill, 1960). For N molecules, the partition function is now $Z_r{}^N$ and the corresponding free energy:

$$G_r = -RT \ln Z_r \quad (7)$$

The translational and rotational enthalpies may be derived from the temperature dependence of G_t and G_r, leading to the classical result: $H_t = H_r = \frac{3}{2}RT$. The corresponding entropies are then $(H - G)/T$.

Equation 4 shows that G_t varies linearly with the logarithm of the molecular weight. This is also true of G_r, at least for molecules of similar shape (Figure 3). For homogeneous

spheres of radius r and partial specific volume $\bar{v}$:

$$M = Nm = \frac{4\pi}{3}\frac{Nr^3}{\bar{v}}$$

$$I_1 = I_2 = I_3 = \frac{2}{3}I = \frac{8\pi}{15}\frac{r^5}{\bar{v}} \quad (8)$$

which, using eq 6 and 7 leads to:

$$G_r = -RT \ln[\beta(RT)^{3/2}M^{5/2}] \quad (9)$$

$$\beta = 75.1\bar{v}/N^4h^3 \quad (10)$$

For nonspherical or inhomogeneous particles, the molecular weight M may be replaced in eq 9 by M', the molecular weight of spherical particles having the same value as the $I_1I_2I_3$ product.

Then, the translational and rotational contributions to the free energy of a dissociation of the complex AB are simply:

$$\Delta G_t = G_t^A + G_t^B - G_t^{AB}$$

$$= -RT \ln\left[\alpha(RT)^{3/2}\left(\frac{M_AM_B}{M_A + M_B}\right)^{3/2}\right] \quad (11)$$

$$\Delta G_r = G_r^A + G_r^B - G_r^{AB}$$

$$= -RT \ln\left[\beta(RT)^{3/2}\left(\frac{M'_AM'_B}{M'_{AB}}\right)^{5/2}\right] \quad (12)$$

The free energy released is that of a particle of molecular weight $M_AM_B/(M_A + M_B)$, the moment of inertia of which corresponds to a sphere of molecular weight $M'_AM'_B/M'_{AB}$. The latter depends somewhat on the geometry of the complex. In the case where one of the molecules is much heavier than the other, eq 11 and 12 show that the free energy released is essentially that of the small molecule.

The above theory assumes that the molecules move and rotate in the absence of external forces; in other words, that they behave like perfect gases. Is that approximation reasonable for aqueous solutions? Interactions between solute molecules may be neglected, provided experimental data are extrapolated to infinite dilutions. Volume exclusion due to the solvent is easily taken into account in the calculation of G_t. Due to the loose packing of water, which leaves empty more than half of its total volume, this effect is small (about 0.4 kcal/mol). Much more important, interactions made with the solvent affect the partition functions Z_t and Z_r in a way which is difficult to estimate.

Still, the "perfect gas" theory has been shown to yield reasonable results in predicting rates of chemical reactions in solution (Page & Jencks, 1971), as it does for gases. Because the rates are very sensitive to the values of the free energies (of activation) involved, the agreement of calculated rates with experimental values indicates that the theory gives a correct estimate of the translational/rotational free energy (at least within a few kcal/mol).

References

Hill, T. L. (1960) *An Introduction to Statistical Thermodynamics,* Addison-Wesley, London.

Page, M. I., & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68,* 1678–1683.

Partial Purification and Characterization of a Neutral Protease Which Cleaves the N-Terminal Propeptides from Procollagen†

Leena Tuderman, Kari I. Kivirikko,§ and Darwin J. Prockop*

ABSTRACT: A rapid assay procedure was developed for cleavage of the N-terminal propeptides of procollagen. With the assay a neutral procollagen N-protease was purified about 300-fold from chick embryo tendon extract. The enzyme had an apparent molecular weight of 260 000 and a pH optimum of 7.4. Ca^{2+} was required for enzymic activity but this requirement was partially replaced by Mg^{2+} or Mn^{2+}. The enzyme was bound to concanavalin A–agarose and therefore was presumably a glycoprotein. The N-propeptides released from type I procollagen were of about 23 000 and 11 000 daltons as estimated by polyacrylamide gel electrophoresis in sodium dodecyl sulfate. The partially purified enzyme was also found to cleave type II procollagen and the N-propeptide obtained was about 18 000 daltons. Heat denaturation of either type I or type II procollagen decreased the rate at which the proteins were cleaved by the N-protease.

The three proα chains of procollagen contain propeptide extensions at both their N- and C-terminal ends, and these propeptides must be removed before the molecule can form stable collagen fibers (for reviews, see Martin et al., 1975; Grant & Jackson, 1976; Prockop et al., 1976; Bornstein & Traub, 1978). Several observations have been made as to how the propeptides are removed and the enzymes involved. Genetic defects in the conversion of type I procollagen to collagen occur in cattle (Lapière et al., 1971; Lenaers et al., 1971), sheep (Fjølstad & Helle, 1974; Schofield & Prockop, 1973), and man (Lichtenstein et al., 1973), and these genetic defects are characterized by the presence of pNcollagen in which the C-propeptides are missing but the N-propeptide of the two proα1 chains and at least part of the N-propeptide of the proα2

† From the Department of Biochemistry, College of Medicine and Dentistry of New Jersey-Rutgers Medical School, Piscataway, New Jersey 08854. *Received February 7, 1978.* This work was supported in part by research Grant AM-16,516 from the National Institutes of Health and by a grant from the Medical Research Council of the Academy of Finland. A preliminary abstract on the work has been presented at the 62nd Annual Meeting of the American Society of Biological Chemists, Atlanta, Georgia, June 1978.

§ Present address: Department of Medical Biochemistry, University of Oulu, Oulu, Finland.

0006-2960/78/0417-2948$01.00/0