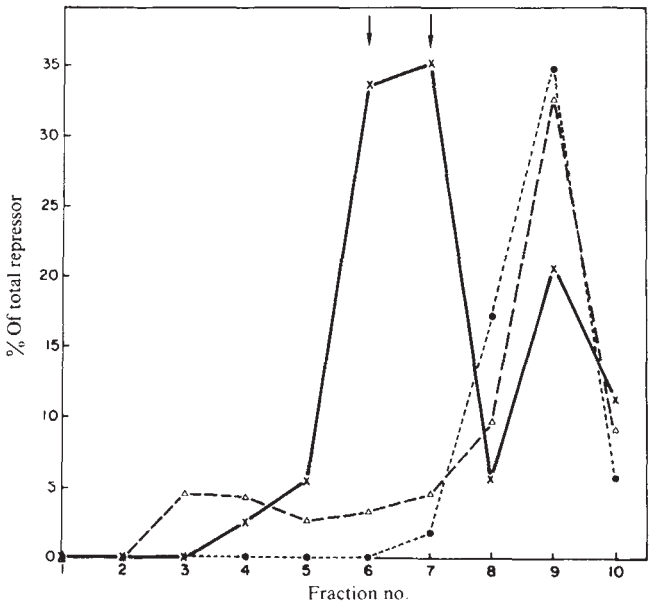

Fig. 2 Comparison of repressor binding to plasmid 25 and pMB9 DNA. A crude lysate of minicells was prepared as described in the legend to Fig. 1. As before, the top fraction of the high salt gradient was used as the source of repressor. Aliquots were mixed with buffer IV or buffer IV plus 2.5μg of linearised plasmid 25 or pMB9 DNA, and adjusted to 0.2 M NaCl. The binding mixture was incubated and run on gradients as described in Fig. 1 legend. The fractions collected were assayed for proteins and DNA as before. The repressor band in autoradiographs of the acrylamide gels were scanned using a Joyce, Loebl densitometer (Chromoscan Mk. II.). The peaks were cut out and weighed to determine the amount of repressor in each fraction. To normalise the amount of repressor in each fraction, samples of the binding mixtures before sedimentation were subjected to electrophoresis on 17.5% SDS-polyacrylamide gels. Autoradiographs of these gels were scanned for the repressor band and the peaks cut out and weighed. The amount of repressor in the control was multiplied by dilution factors to determine the amount of repressor in each binding mix. As an equal amount of the repressor preparation was used for each binding mixture, an average of the amount of repressor in the controls was used to normalise the amount of repressor in fractions of the gradients. The per cent of the total repressor recovered in fractions of gradients with *a*, repressor alone (●); *b*, repressor plus plasmid 25 DNA (×); and *c*, repressor plus pMB9 DNA (△) is shown. The vertical arrows indicate the position of DNA in the gradients. DNA of both plasmid 25 and pMB9 was digested with the restriction enzyme *Hin*dIII (Bethesda Research Labs) which cleaves both DNAs once, making linear molecules. Uncut plasmid 25 also shows binding of the repressor; however, two peaks of bound repressor are detected, corresponding to DNA in the monomer and dimer covalently closed circular forms. In these conditions the dimer DNA sediments further down the gradient than the monomer form. Plasmid 25 and pMB9 were purified from strain χ1274 (*E. coli* K12, Thi$^-$, Thr$^-$, Leu$^-$, LacY$^-$, MalA$^-$, MtL$^-$, Xyl$^-$, Ara$^-$, Gal$^-$, Str A, ton A, λ^R, azi, Min$^-$, su$^+$) according to the procedure of Clewell[13].

repressor found in the top fraction of the gradients did not decrease significantly with increasing amounts of DNA (Table 2). The amount of inactive repressor in any one preparation increased with the storage time of minicell preparations before use.

As shown in Table 2, both the amount of DNA-bound repressor and the overall amount of repressor recovered increase with increasing concentrations of plasmid 25 DNA, except at very high concentrations of DNA. Similarly, more repressor was recovered when pMB9 DNA was added to the same preparation, than was recovered after sedimenting the preparation alone. In this case, however, the repressor was not associated with pMB9 DNA, but was spread throughout the gradient (Fig. 2). This suggests that the active repressor proteins will bind to each other, if not bound to DNA, forming aggregates which sediment rapidly in the gradient. Thus, if the repressor extract is sedimented without added DNA, much of the active repressor goes to the bottom of the gradient.

This repressor–repressor interaction must be weaker than repressor–DNA binding, as any DNA added to the repressor seems to inhibit aggregate formation. As the repressor dissociates from non-operator-containing DNA (that is, pMB9) during sedimentation, however, it will begin to form repressor–repressor aggregates, and this repressor is detected throughout the gradient. Repressor binding to operator-containing DNA (plasmid 25 DNA) is tight, so that very little repressor dissociates from the DNA during sedimentation and no aggregation of the repressor occurs. This finding is consistent with the observation that there is a decrease in the amount of total repressor recovered in binding assays carried out at 0.3 M rather than 0.1 M NaCl. The decrease in recovery is more dramatic with D108 or pMB9 DNA than with Mu or plasmid 25 DNA.

Further characterisation of the nature of the binding of Mu repressor to Mu DNA (and repressor–repressor interaction) requires purification of significant amounts of repressor protein. This is feasible because large quantities of repressor can be made in strains carrying plasmid 25 (or similar recombinant plasmids), and crude DNA binding assays (like that described above) allow easy detection of active repressor during purification.

This work was supported by NSF grant PCM 74-0016 A01 and NIH grant GM 17612-07 to D.Z. Plasmid p5D7 was obtained from J. Broach.

DEBORAH KWOH
DAVID ZIPSER

Cold Spring Harbor Laboratory,
PO Box 100,
Cold Spring Harbor, New York 11724

Received 18 October; accepted 15 December 1978.

1. Howe, M. M & Bade, E. G. *Science* **190,** 624–633 (1975).
2. Bukhari, A. I. *A. Rev. Genet.* **10,** 389–411 (1976).
3. Abelson, J. *et al. Virology* **54,** 90–92 (1973).
4. Allet, B. & Bukhari, A. I. *J. molec. Biol.* **92,** 529–540 (1975).
5. Figurski, D., Meyer, R., Miller, D. & Helinski, R. *Gene* **1,** 109–119 (1976).
6. Magazin, M. & Allet, B. *Virology* **85,** 84–87 (1978).
7. Zipser, D., Moses, P., Kahmann, R. & Kamp, D. *Gene* **2,** 263–271 (1977).
8. Gilbert, W. & Muller-Hill, B. *Proc. natn. Acad. Sci. U.S.A.* **56,** 1891–1898 (1966).
9. Ptashne, M. *Nature* **214,** 232–234 (1967).
10. Hull, R. A., Gill, G. S. & Curtiss, R. III *J. Virol.* **27,** 513–518 (1978).
11. Ptashne, M. & Hopkins, N. *Proc. natn. Acad. Sci. U.S.A.* **60,** 1283–1287 (1968).
12. Kahmann, R., Kamp, D. & Zipser, D. *Mol. gen. Genet.* **149,** 323–328 (1976).
13. Clewell, D. B. *J. Bact.* **110,** 667–676 (1972).

Surface and inside volumes in globular proteins

MEASUREMENTS of the surface area accessible to solvent provide a convenient definition of the surface and the inside volumes in proteins of known X-ray structure. The study of the accessibility to solvent of amino acid residues in several proteins[1–3] has confirmed the early observation that polar residues are found mostly on the surface and non-polar residues mostly inside globular protein structures. But the accessibility shows systematic variations with the molecular weight, because of the change in surface to volume ratio. Experimental data[4] indicate that the accessible surface area A (in Å^2) of monomeric globular proteins follows the law[5,6]:

$$A = 11.1\, M^{2/3} \qquad (1)$$

which implies that the mean accessible surface area per residues decreases like $M^{-1/3}$ (where M is molecular weight) with increasing M, from about 68 Å^2 for proteins of 6,000 molecular weight to 38 Å^2 for proteins of 35,000 molecular weight. Thus, the accessibility to solvent is not a characteristic of the amino acids.

We made histograms of the accessible surface areas of residues in a sample of 22 proteins, and found that the number n_B of buried residues, defined as having less than a given accessible surface area A_m, varies with the total number n of residues in the

0028–0836/79/0277—0491$01.00

polypeptide chain according to:

$$n^{1/3} - n_B^{1/3} = k \quad (2)$$

where k depends only on A_m. Completely buried residues (zero accessible surface area) are few: 4 in pancreatic trypsin inhibitor (58 residues), 40 in carboxypeptidase A (307 residues). Better statistics are obtained by choosing $A_m = 10$ to 20 Å^2, which amounts to consider as buried residues which have one or two atoms in contact with the solvent. With $A_m = 20$ Å^2, the average value and standard deviation of k are 1.62 ± 0.13. The geometrical significance of equation (2) is apparent in Fig. 1: k represents the thickness of the surface layer made of residues in contact with the solvent. While the hard sphere model of globular proteins, first presented by Fisher[7], is in formal agreement with equations (1) and (2), the fit only implies that the gross geometrical properties of globular proteins (surface and volume) are those of a set of solid bodies which deviate in similar ways from a spherical shape. This is not true of subunits in oligomeric proteins: the subunits of tetrameric lactate dehydrogenase[4] and glyceraldehyde-3-phosphate dehydrogenase have much larger accessible surface areas and fewer buried residues than expected from equations (1) and (2).

Table 1 gives the amino acid composition of our sample of buried and accessible residues. The buried residues make up the protein inside; 56% are Leu, Ile, Val, Ala, or Gly, another 22% are Phe, Ser, Thr or Cys, and 22% are of the 11 other types. The protein surface has a much more regular amino acid composition than the inside: non-polar residues are by no means absent, even though Asn, Asp, Glu, Gln, Lys and Arg constitute 40% of the accessible residues. The ratio f of the molar fractions of each amino acid in the two samples is of particular interest. It represents the partition coefficient of this amino acid between the surface and inside volumes. Consequently, $\Delta G_t = RT \ln f$ may be taken as the free energy of transfer from the inside to the surface. Though the exact value of f depends on the choice of A_m made in defining buried residues, the effect is small: reducing A_m from 20 Å^2 to 10 Å^2 changes the free energies by no more than their statistical standard deviation of 0.1–0.2 kcal mol^{-1}. Thus f and ΔG_t are characteristic of the amino acids in globular proteins.

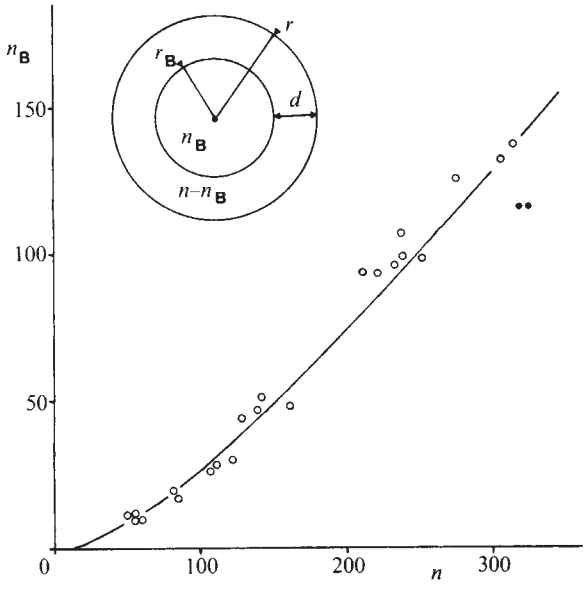

Fig. 1 Buried residues in globular proteins. The accessible surface areas of individual residues are calculated according to Lee and Richards[1] using a computer program of M. Levitt. The atomic coordinates are obtained from the Protein Data Bank, Cambridge, UK. The following 22 protein structures are used here and in Table 1: insulin, ferredoxin, rubredoxin, pancreatic trypsin inhibitor, high potential iron protein, cytochrome b_5, parvalbumin, cytochrome c_2, ribonuclease S, hen lysozyme, staphylococcus nuclease, flavodoxin, phage T4 lysozyme, papain, trypsin, concanavalin A, elastase, α-chymotrypsin, carbonic anhydrase B, substilisin, carboxypeptidase A, thermolysin. The number n_B of residues with less than $A_m = 20$ Å^2 of accessible surface area is plotted against the total number n of residues in the polypeptide chain (open circles). Filled circles represent the subunits of lactate dehydrogenase and glyceraldehyde -3-phosphate dehydrogenase. The insert represents the protein as made of two concentric spheres, the volumes of which are proportional to n and n_B; the radii are proportional to $n^{1/3}$ and $n_B^{1/3}$, and therefore k is proportional to the thickness d of the surface layer, made of $n - n_B$ accessible residues. For the 22 proteins: $A_m = 0$ Å^2, $k = 3.0 \pm 0.4$ (s.d.), $d = 9.6$ Å; $A_m = 10$ Å^2, $k = 2.0 \pm 0.2$, $d = 6.4$ Å, $A_m = 20$ Å^2, $k = 1.62 \pm 0.13$, $d = 5.2$ Å. The solid line is calculated from equation (2) with $k = 1.62$.

Table 1 Amino acid composition of the inside and surface

Residue	Molar fraction		Free energy (kcal mol^{-1})		
	Buried	Accessible	f	ΔG_t	ΔG_h
Leu	11.7	4.8	2.4	0.5	1.8
Val	12.9	4.5	2.9	0.6	1.5
Ile	8.6	2.8	3.1	0.7	
Phe	5.1	2.4	2.2	0.5	2.5
Cys	4.1	0.9	4.6	0.9	
Met	1.9	1.0	1.9	0.4	1.3
Ala	11.2	6.6	1.7	0.3	0.5
Gly	11.8	6.7	1.8	0.3	0.0
Trp	2.2	1.4	1.6	0.3	3.4
Ser	8.0	9.4	0.8	−0.1	−0.3
Thr	4.9	7.0	0.7	−0.2	0.4
His	2.0	2.5	0.8	−0.1	0.5
Tyr	2.6	5.1	0.5	−0.4	2.3
Pro	2.7	4.8	0.6	−0.3	
Asn	2.9	6.7	0.4	−0.5	
Asp	2.9	7.7	0.4	−0.6	
Gln	1.6	5.2	0.3	−0.7	
Glu	1.8	5.7	0.3	−0.7	
Arg	0.5	4.5	0.1	−1.4	
Lys	0.5	10.3	0.05	−1.8	

Molar fractions (%) of the 20 amino acids in samples of 2,001 buried residues (less than 20 Å^2 accessible surface area) and 3,220 accessible residues (more than 20 Å^2); f is the ratio of the buried/accessible molar fractions; $\Delta G_t = RT \ln f$. ΔG_h is the free energy of transfer from an organic solvent to water[8]. Cys residues with a free −SH group are few in the sample; our data are more representative of substituted −SH groups (disulphides, metal adducts or thioethers).

In a protein of n residues, the most probable fraction g of buried residues of one type, say valine, is easily calculated from the data of Table 1:

$$g = \frac{f n_B}{n - n_B + f n_B} \quad (3)$$

where n_B is given by equation (2). Thus, the fraction of valines which are buried increases from 0.4 to 0.7 when the chain length increases from 60 to 300 residues, while it remains less than 0.1 for lysines.

Chothia[3] has noted that the average accessibility of residues in proteins is not correlated to the free energies ΔG_h of transfer from an organic solvent to water, measured by Nozaki and Tanford[8]. This is seen clearly in Table 1: ΔG_t is governed by the presence or absence of oxygen and nitrogen atoms in the side chain[1,3]. These polar atoms must find hydrogen bond partners when the residue is buried, a stringent requirement except for Ser and Thr residues, the hydroxyl group of which can easily form a hydrogen bond to adjacent main chain carboxyl groups[9]. Proline appearas as polar; its imino nitrogen cannot participate in normal main chain hydrogen bonds and behaves as a part of the side chain.

JOËL JANIN

Service de Biochimie Cellulaire,
Département de Biochimie et Génétique Microbienne,
Institut Pasteur,
28, rue du Docteur Roux, 75724 Paris Cedex 15, France

Received 3 November; accepted 18 December 1978.

1. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
2. Shrake, A. & Rupley, J. A. *J. molec. Biol.* **79**, 351–372 (1973).
3. Chothia, C. H. *J. molec. Biol.* **105**, 12 (1976).
4. Chothia, C. H. *Nature* **254**, 304–308 (1975).
5. Janin, J. *J. molec. Biol.* **105**, 13–14 (1976).
6. Teller, D. C. *Nature* **260**, 729–731 (1976).
7. Fisher, H. *Proc. natn. Acad. Sci. U.S.A.* **51**, 1285–1286 (1964).
8. Nozaki, Y. & Tanford, C. *J. biol. Chem.* **246**, 2211–2217 (1971).
9. Janin, J., Wodak, S., Levitt, M. & Maigret, B. *J. molec. Biol.* **125**, 357–386 (1978).

Dimeric association of band 3 in the erythrocyte membrane demonstrated by protein diffusion measurements

BAND 3, a major constituent of the human erythrocyte membrane, comprises a class of membrane-spanning proteins which are implicated in anion and possibly other transport functions[1-4]. Various observations indicate that band 3 proteins may exist as dimers in the membrane[5-9]. The most direct indication has been obtained from chemical cross-linking experiments. The band 3 dimer is an early product of various cross-linking reactions[7], and the particular case of cross-linking by oxidative disulphide bridge formation results in almost quantitative dimerisation of band 3 (ref. 8). However, cross-linking experiments are in general inconclusive for they do not indicate whether a successful cross-linkage reflects a naturally occurring stable complex, or is the result of collisions between independent proteins which rapidly diffuse over small distances. In the recent elegant studies of Kiehm and Ji[9], rapid photochemical cross-linking was achieved by photolysing with a millisecond flash. Whilst these experiments undoubtedly provide the best evidence so far obtained for the existence of the band 3 dimer, they still leave some doubts concerning the possibility of collision-induced crosslinks. These can only be ruled out if the mean time between collisions is shown to be much longer than the lifetime of the flash-induced reactive species. The latter time was assumed to be short (of the order of 1 ms) on the basis of solution studies but was not measured in the experimental conditions. The collision frequency is a function of the local lateral diffusion coefficient. Band 3 diffusion over distances >1 μm seems to be slow in the erythrocyte membrane[10,11], but as discussed elsewhere[12], this may reflect restrictions to long range diffusion and does not rule out the possibility of more rapid local diffusion. If the local diffusion coefficient is in the range 10^{-9} to 10^{-10} cm^2 s^{-1}, the mean time between collisions is also of the order of 1 ms (ref. 9). A lateral diffusion coefficient in this range is consistent with the rotational diffusion of band 3 (ref. 12). Here we present a new physical approach to determining the self-association of band 3, based on the measurement of rotational diffusion. The rotational diffusion coefficient of an intrinsic membrane protein is strongly dependent on the diameter of its cross-section in the plane of the membrane. We show that cross-linking band 3 into dimers produces no observable decrease in the rotational diffusion coefficient of band 3 in the membrane. We therefore conclude that the dimer pre-exists in the membrane and is not a product of the cross-linking reaction following random collisions.

In the present experiments, the rotation of band 3 was measured by observing the flash-induced transient dichroism of a triplet probe, as previously described[12,13]. However, band 3 was labelled with eosin maleimide instead of the eosin isothiocyanate previously used; essentially similar results are obtained with either probe. Dimerisation of band 3 in the membrane was achieved by copper–phenanthroline-catalysed oxidation of cytoplasmic sulphhydryl residues to form disulphide bridges, as described by Steck[8].

Figure 1 shows the result of the cross-linking reaction as visualised by SDS-polyacrylamide gel electrophoresis. Band 3 has disappeared from its normal position in the gel, moving to a new band with a molecular weight of approximately 180,000, corresponding to the band 3 dimer. This result is identical to that reported by Steck[8].

The rotational mobility of band 3 in the cross-linked membranes was measured and compared with that of untreated controls. The transient dichroism measurements were analysed

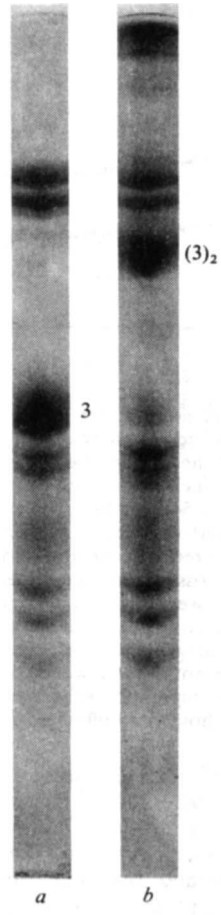

Fig. 1 Dimerisation of band 3 by oxidative disulphide bridge formation. Intact human erythrocytes were labelled with eosin-5-maleimide as described previously for other eosin derivatives[12,13,19], except that incubation was for 1 h at room temperature. These conditions yield approximately 1 molecule of bound eosin per band 3 monomer. Ghosts were prepared by hypotonic lysis[12] and given a final wash in 5 mM sodium phosphate, *p*H 8.0. They were allowed to react for 1 h at room temperature with one volume of *o*-phenanthroline (200 μM) and copper sulphate (40 μM) in the same buffer. The reaction was ended by the addition of 30 volumes 1 mM EDTA and 5 mM sodium phosphate, *p*H 8.0, and by a further wash in 5 mM sodium phosphate, *p*H 7.4. The control sample was subjected to the same procedures but without the addition of the copper–phenanthrolin reagent. Ghosts were then dissolved for electrophoresis and run on cylindrical 4% polyacrylamide gels in 0.2% SDS as described by Steck[8]. 3 and $(3)_2$ denote the positions of the band 3 monomer and dimer, respectively. Gel *a*, control; gel *b*, cross-linked ghosts.

by calculating the absorption anisotropy $r(t)$ given by

$$r(t) = \frac{A_{\parallel}(t) - A_{\perp}(t)}{A_{\parallel}(t) + 2A_{\perp}(t)} \quad (1)$$

where $A_{\parallel}(t)$ and $A_{\perp}(t)$ are, respectively, the absorbance changes at time t after the flash for light polarised parallel and perpendicular with respect to the polarisation of the exciting flash. Figure 2 shows the time dependence of r for dimerised and untreated band 3. It is clear that the shapes of the two curves are virtually identical.

As discussed elsewhere[12,14], the predicted form of $r(t)$ for a protein which rotates only about the axis normal to the plane of the membrane is

$$r(t) = A_1 \exp(-D_{\parallel}t) + A_2 \exp(-4D_{\parallel}t) + A_3 \quad (2)$$

where $D_{\parallel}$ is the rotational diffusion coefficient and A_1, A_2, A_3 are constants. To make a quantitative comparison of the cross-linked and control samples, we made a least-mean-square fit of

0028–0836/79/0277—0493$01.00