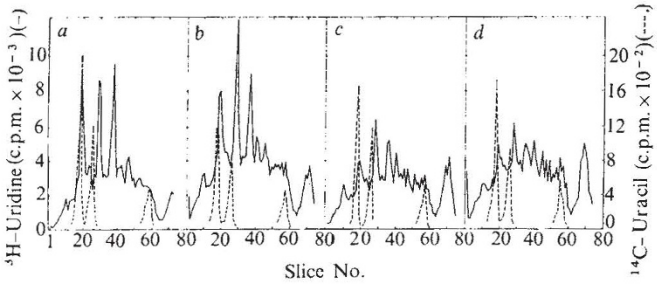

FIG. 6 Polyacrylamide-agarose gel electrophoretic pattern of T7 *am*193 mRNA after rifampicin treatment. RNA samples were as used in the hybridisation experiment shown in Fig. 1*b*. Electrophoresis was carried out as described in the legend for Fig. 5. *a*, T7 *am*193 RNA from ultraviolet-irradiated cells 12 min after infection (rifampicin was added at 10 min); *b*, 15 min; *c*, 20 min; *d*, 25 min; —, ^{3}H-labelled T7 *am*193 RNA; – – –, ^{14}C-labelled marker *E. coli* 23S, 16S, and 4S, from left to right.

This work was supported by a grant from the US Public Health Service. We thank Dr E. G. Minkley for his advice on electrophoresis of T7 mRNA, Drs F. W. Studier and W. C. Summers for stocks of T7 phages and bacteria, and Mr L. Coco for assistance.

YOSHIHIKO YAMADA
PATRICIA A. WHITAKER
DAI NAKADA

Department of Biochemistry,
University of Pittsburgh School of Medicine,
Pittsburgh, Pennsylvania 15261

Received October 8; revised December 3, 1973.

1 Summers, W. C., *J. molec. Biol.*, **51**, 671 (1970).
2 Marrs, B. L., and Yanofsky, C., *Nature new Biol.*, **234**, 168 (1971).
3 Studier, F. W., *Virology*, **39**, 562 (1969).
4 Minkley, E. G., and Pribnow, D., *J. molec. Biol.*, **77**, 255 (1973).
5 Adesnik, M., and Levinthal, C., *Cold Spring Harbor Symp. quant. Biol.*, **35**, 451 (1970).
6 Puga, A., Barrás, M.-T., Tessman, E. S., and Tessman, I., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 2171 (1973).
7 Hayashi, Y., and Hayashi, M., *Cold Spring Harbor Symp. quant. Biol.*, **35**, 171 (1970).
8 Jaenish, R., Jacob, E., and Hofschneider, P. H., *Nature*, **227**, 59 (1970).
9 Robertson, H. D., Webster, R. E., and Zinder, N. D., *J. biol. Chem.*, **243**, 82 (1968).
10 Dunn, J. J., and Studier, F. W., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 1559 (1973).
11 Hagen, F., and Young, E. T., *Virology*, **55**, 231 (1973).
12 Gillespie, D., and Spiegelman, S., *J. molec. Biol.*, **12**, 829 (1965).
13 Peacock, A. C., and Dingman, C. W., *Biochemistry*, **7**, 668 (1968).

Hydrophobic bonding and accessible surface area in proteins

THE hydrophobic bond is the term used by Kauzmann[1] to describe the gain in free energy on the transfer of non-polar residues from an aqueous environment to the interior of proteins. This has been accepted as one of the major forces involved in the folding of proteins. The exact origin of the energy of the hydrophobic bond is controversial[2], but empirical values have been derived for 10 protein residue side chains by Nozaki and Tanford[3] who measured the solubility of amino acids in the organic solvents ethanol and dioxane.

For an atomic group in a protein to interact with the solvent it must be able to form Van der Waals contracts with water molecules. The concept of accessible surface area has been introduced by Lee and Richards[4] to describe quantitatively the relationship of proteins to solvent. The accessible surface area of an atomic group is defined as the 'area on the surface of a sphere radius R, on each point of which the centre of a solvent molecule can be placed in contact with the atom without penetrating any other atoms of the molecule'[4]. R is the sum of the Van der Waals radii of the atom and the solvent molecule.

In Fig. 1, the hydrophobicity of each of the 10 protein residue side chains is plotted against its accessible surface area. To derive the hydrophobicity of a side chain Nozaki and Tanford subtracted the value for glycine from that of the amino acid. I calculated the accessible surface areas of protein residue, X, from the accessibilities given by Lee and Richards[4] for X in tripeptide Ala-X-Ala, where the main chain is in an extended β conformation and the side torsion angles antiplanar. The values for the side chains in Fig. 1 are the accessible surface area of the residue minus the accessible surface area of glycyl (82 Å^2).

From Fig. 1 we see that for the non-polar side chains Ala, Val, Leu and Phe there is a linear relationship between hydrophobicity and accessible surface area; the correlation coefficient is 0.998 and the slope equivalent to 22 calorie Å^{-2}. For the side chains with hydroxyl groups Ser, Thr and Tyr there is a similar linear relationship with a correlation coeffiicient of 0.998 and a slope equivalent to 26 calorie Å^{-2}. For residues with similar accessible surface areas those with a polar group are about 1 kilocalorie less hydrophobic than those with only non-polar groups. In aqueous solution the OH, NH, N and S groups will form hydrogen bonds with water. But on transfer to ethanol or dioxane these solute-solvent hydrogen bonds are significantly weakened[5] and there is therefore a loss of free energy (1-2 kilocalorie per polar group is a reasonable estimate). In proteins this loss does not occur: ~ 92% of polar groups buried in the interior of proteins form hydrogen bonds. Thus the hydrophobicity of polar side chains will be greater than the values in Fig. 1. A 'correction' of 1–2 kilocalorie per polar group makes the polar residues as hydrophobic as non-polar residues. (The histidine side chain has two polar groups.) This implies that, if hydrogen-bonded, oxygen and nitrogen atoms are hydrophobic; they could, for example, be favourably accommodated in the interior of membranes. This result is related to the finding of Lee and Richards[4] that when lysozyme, RNase S and myoglobin go from an extended chain to the folded conformation, the reduction in accessible surface area is almost the same for polar and non-polar atoms. Methonine is a little less, and tryptophan a little more, hydrophobic than this argument and the linear relationship illustrated in Fig. 1, would suggest.

In conclusion, we can say that as a good approximation the hydrophobicity of residues in proteins is 24 calorie per Å^2 of accessible surface area. This value is applicable to protein folding, subunit contacts and the interaction of small molecules with enzymes and immunoglobins. For example the formation of the $\alpha_1\beta_1$ contact in haemoglobin reduces in each monomer the surface area accessible to water by ~900 Å^2 (my unpublished results). This implies that the hydrophobic contribution to the free energy of dimer formation is 43 kilocalorie. The formation of the contacts that cause the reduction in accessible surface area when proteins fold or aggregate involves large conformational entropy losses[6,7] and appreciable steric strain (Szabo and Karplus, personal communication) which substantially reduces the figure for the free energy.

The value given here for the hydrophobicity of protein surfaces confirms Kauzmann's hypotheses[1] that the entropic term that stabilises protein aggregation and the binding of small molecules is mainly a result of hydrophobic bonding, and suggests that only relatively small contact areas are necessary for stability. In haemoglobin ~12% of the ac-

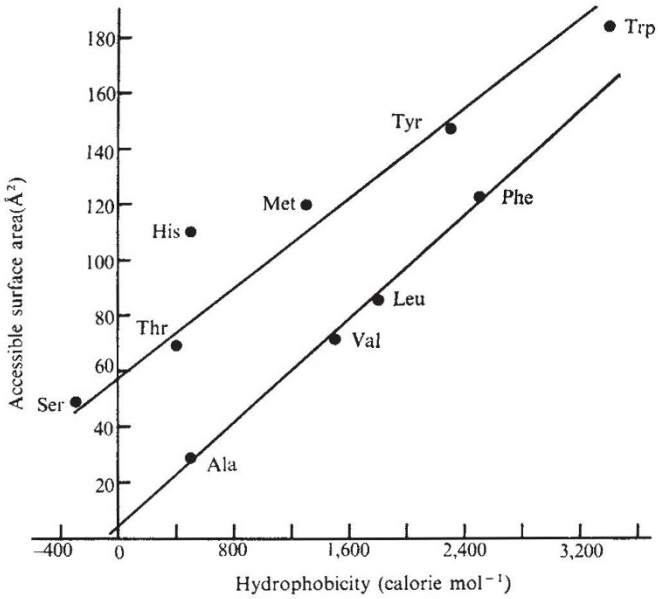

FIG. 1 Accessible surface areas of residue side chains (see text) plotted against hydrophobicity (free energy change for the transfer from 100% organic solvent to water[3]).

ccssible surface area of each monomer is covered by the $\alpha_1\beta_1$ contact and ~6% by the $\alpha_1\beta_2$ contact.

In a discussion of the general theory of hydrophobic bonding Hermann[8] deduced and found a linear relationship between the logarithm of the solubility of hydrocarbons and the surface area of the cavity they form in water. This latter term is equivalent to the accessible surface area of the hydrocarbons.

I thank my colleagues for critical discussion.

CYRUS CHOTHIA

Medical Research Council,
Laboratory of Molecular Biology,
Hills Road, Cambridge CB2 2QH

Received September 26; revised November 27, 1973.

1 Kauzmann, W., *Advances in Protein Chemistry*, **14**, 1 (1959).
2 Klapper, M. H., *Progress in Biorganic Chemistry*, **2**, 55 (1973).
3 Nozaki, Y., and Tanford, C., *J. biol. Chem.*, **246**, 2211 (1971).
4 Lee, B., and Richards, F. M., *J. molec. Biol.*, **55**, 379 (1971).
5 Klotz, I. M., and Franzen, J. S., *J. Am. chem. Soc.*, **84**, 3461 (1962).
6 Page, M. I., and Jencks, W. P., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 1678 (1971).
7 Kauzmann, W., in *The Mechanism of Enzyme Action* (Johns Hopkins University Press, Baltimore, 1954).
8 Hermann, R. B., *J. phys. Chem., Ithaca.*, **76**, 2754 (1972).

Insulin/proinsulin, a new crystalline complex

SINCE the first crystallisation of proinsulin in this laboratory[1], we have tried to grow crystals appropriate for detailed X-ray crystal structure analysis. None of the proinsulin crystal forms prepared proved ideal[2], and we attempted to cocrystallise insulin and proinsulin. Such crystals, if both components were in ordered array, could provide a basis for the determination of the structure of proinsulin by the use of rigid-body vector search methods[3,4] using the known stereochemistry of insulin[5].

Proinsulin is found in crystalline native insulin preparations[6]. It is common to find proteins from inhomogeneous fractions incorporated as impurities into crystals of the dominant component[7]. Very similar protein molecules, for example, β_1 and β_2 lactoglobulin[8] often form mixed crystals. These are defect solids, with all crystallographically equivalent positions occupied, but by molecules of different kinds. Steiner[9] has reported the preparation of what he seems to claim are mixed crystals of insulin and proinsulin of this type. Mixed crystals are not cocrystals if, by this we imply structures with each separate protein ordered by the full symmetry of the space group nor, because of disorder, are they useful for X-ray crystal structure analysis.

We report here the isolation and preliminary X-ray study of a 1:1 crystalline complex of insulin and proinsulin which seems to be a true cocrystalline form.

Cocrystallisation studies were made with both beef and pork hormone/prohormone pairs using the molar ratio 1:1 exclusively. Heterogeneous pairs were not studied. Separate insulin and proinsulin solutions of initial molarities 1.4 mM were prepared by dissolving each protein in 50 μl of a solution containing 0.05 M sodium citrate, 30 mg ml^{-1} sodium chloride, and, where used, 0.836 mg ml^{-1} of zinc chloride and 3 mg ml^{-1} of *m*-cresol. The final *p*H was adjusted to 7.25 by addition of 0.5 N HCl. Both proteins crystallise at this *p*H in the presence of *m*-cresol and zinc; insulin as diamond-shaped plates, proinsulin as tetragonal prisms with pyramidal termination. Hormone and prohormone solutions were warmed to 50° C with stirring, then mixed and kept at 50° C for a few minutes. Three vials of bovine material containing cresol and zinc were prepared. All gave a precipitate on cooling to room temperature. Protein-free crystallisation medium was added (with rewarming) to provide a series with final volumes of 130, 150 and 180 μl respectively. All vials were then kept at room temperature. Small, well-formed crystals began to appear after 2–4 d. The vials were then stored at 2° C. The crystals continued to grow for 1–2 weeks.

All preparations gave crystals. Only tiny, unidentified crystals, 50–60 μm long grew in the absence of *m*-cresol and/or zinc. The crystals in both pork and beef insulin—

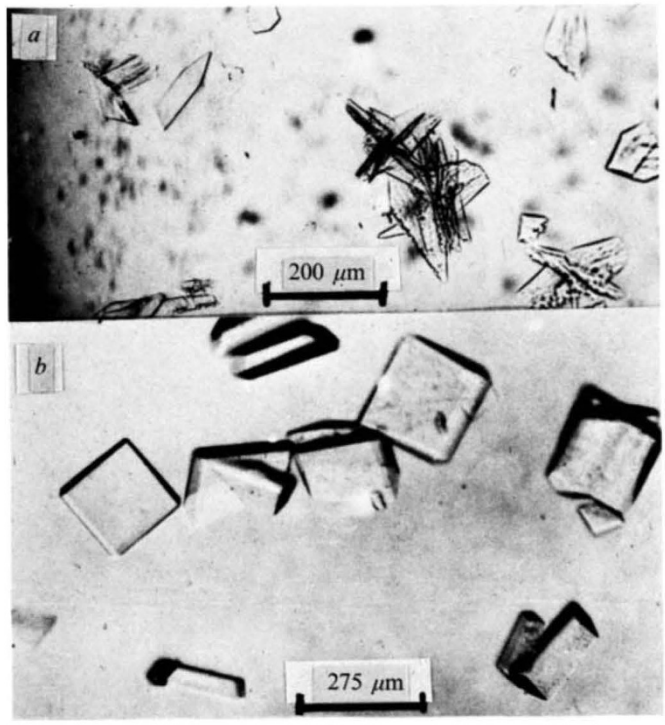

FIG. 1 *a*, 1:1 crystalline complex of porcine insulin/proinsulin; *b*, 1:1 crystalline complex of bovine insulin/proinsulin.