Russian Chemical Reviews **63** (9) 781 – 796 (1994)
©1994 Russian Academy of Sciences and Turpion Ltd

UDC 577.15.035

Tritium planigraphy of biological systems

A V Shishkov, L A Baratova

Contents

I.	Introduction	781
II.	Fundamentals of the method	782
III.	Study of the accessible surface of proteins. Limits of application of the method	785
IV.	*Conformational transitions in proteins*	788
V.	Accessible surface and intramolecular dynamics of proteins	789
VI.	Modelling the three-dimensional structure of proteins	790
VII.	Three-dimensional organisation of supermolecular systems	791

Abstract. The transfer of the ideas and methods of nuclear physics and nuclear chemistry to the study of structural biology has led to the creation of a new method of studying the accessible surface and the three-dimensional structure of biological materials. We shall review the papers which have formed a basis for the tritium planigraphy method and the main directions in which the method has developed in its applications to molecular biology problems. The bibliography includes 129 references.

I. Introduction

Recent advances in the study of many biological phenomena at the molecular level have stimulated the search for and development of new approaches to the study of the structure of supermolecular systems, to the determination of the size and composition of the surface of the interacting components in the system, and to the study of the mechanism of specific protein–protein and protein–nucleic acid interactions.

High-resolution (1 – 2 Å) X-ray structural analysis is the only experimental method capable of giving wide-ranging data on the three-dimensional structure, and therefore also on the surface area of the object. However, its range of applications is limited to a relatively narrow choice of materials, available as highly ordered crystals. In addition to the purely technical problems associated with the growth of these crystals (especially in the case of large and labile complexes) there are difficulties in the interpretation of their complex diffraction patterns. Furthermore, demonstrating the identity in structure of a given compound in a crystalline sample and in solution (the natural working medium of biological systems) remains an unsolved problem.

Because of these difficulties methods such as NMR are still used in studies of relatively low-molecular peptides and proteins.[1]

A V Shishkov N N Semenov Institute of Chemical Physics, Russian Academy of Sciences, 4 Ul. Kosygina, 117334 Moscow, Russian Federation, Fax (7-095) 938 21 56
L A Baratova A N Belozersky Institute of Physicochemical Biology, M V Lomonosov Moscow State University, Vorob'evy Gory, 117400 Moscow, Russian Federation, Fax (7-095) 939 31 81

Received 8 June 1994
Uspekhi Khimii **63** (9) 825 – 841 (1994); translated by J I Carasso

In the case of large-scale objects (100 nm or more) structural information can be obtained by electron microscopy methods, which (in conjunction with the most recent data processing techniques) offer a resolution of the order of 20 Å, increasing to 5 Å in cryoelectron crystallographic studies.[2,3] This method allows the molecular architecture of complexes to be studied, together with the distribution of the component masses in the complex, the relative localisation of the molecules, and the quaternary structure of the protein (see, for example, Holzenburg et al.[4]).

The new scanning electron microscopy method, which offers the highest reported resolution (~1 Å), must also be mentioned among the methods of studying the surface, but it has not yet achieved the status of a generally usable method. So far its application to biological materials is mainly in general illustration of areas of interest, but we may well find that with some further development the method will provide information at the level of localisation of individual atoms.[5]

Traditional approaches, based on the introduction of various types of probe and of modifier, are widely used in studies of the surface topography and the structure of biological systems. The recently produced low-frequency ESR spectrometers have allowed whole cells and even fragments of tissue to be studied *in vivo*.[6] However, the structural information obtained in this way applies exclusively to local fragments of the object, or (in the most favourable case) to the nearest surroundings of the tracer species, and is totally unrelated to the structure of the sample as a whole. Perturbations which can seriously alter the native structure of the object can be induced by the introduction of probes of various kinds, and should not be ignored when this information is interpreted.

In all the studies listed above the use of modifying agents labelled with radioactive isotopes increases the sensitivity of the detection by a large factor. Successful studies of the selective iodination of surface proteins in cell membranes[7] and of intact hepatitis A virus particles[8] are good examples of this effect.

However, there is also another group of methods based on the use of radioactive indicators. In these methods structural data are obtained directly from the method used to introduce the label, which establishes the accessibility of the various groups in the three-dimensional structure to the particular test label.

A good example of this approach is the study of the surface of specific protein-transfer RNA complexes, in which the label was introduced into the H-8 position of the purine bases of RNA by exchange with HTO water.[9] Of course, there are many cases in

which both groups of methods can be applied: in these, the steric accessibility of the reactive group is established (from the yield of the labelled product) together with its structural surroundings (for example, from the spectral characteristics).

The tritium planigraphy method belongs essentially to the second group. It is characterised by lack of selectivity of the labelling. As will be shown below, it can be used to determine the steric accessibility of the hydrocarbon groups ($-CH$, $-CH_2$, $-CH_3$) to attack by a hot tritium atom under free flight conditions. The result of this attack is the replacement of ^1H hydrogen by its radioactive isotope tritium, which takes place in all the sterically accessible fragments of the object with approximately equal probability and does not affect their properties: the labelled product and the unlabelled initial molecule are chemically identical. The range of application of this method is quite wide, since practically all the biological molecules are derivatives of hydrocarbons. It should be pointed out that the small size of the tritium atom results in a higher level of structural discrimination than can be achieved by using larger probes or labels.

In this review we shall not discuss the various methods of introducing the tritium label: interested readers are referred to the monograph by Neiman et al.[10] We shall instead confine the discussion to the tritium planigraphy method, in which the hot tritium atoms are generated by thermal dissociation of molecular tritium.

II. Fundamentals of the method

The tritium planigraphy method is based on chemical reactions involving fast particles (the so-called hot-atom chemistry), which began to be studied in the 1950s and 1960s. The most important result of these studies was the concept of the excitation function of a chemical reaction, i.e. the dependence of the probability (or cross-section) of the reaction on the energy of the particles. It was shown that in many cases this function has a limiting character, and the cross-section of the reaction increases from zero over a relatively narrow range of energies. Obviously, if several reactions (with different thresholds) can occur together in a given system the process can be shifted in the required direction by varying the energy. Thus, if the particles X and Y can interact by three channels having activation energies E_1, E_2, and E_3, where $E_1 < E_2 < E_3$, all three products can be formed in the system when $E > E_3$. However, if $E_2 < E < E_3$ only two channels are available to the reaction, and lastly for $E_1 < E < E_2$ only one reaction is possible (Fig. 1).

The most important chemical process taking place during the interaction of hot hydrogen atoms with saturated hydrocarbons is the substitution reaction, which (according to the energy of the atoms) can proceed by the direct substitution mechanism (activation energy 1.5 eV)[11,12]

$$RH + \overset{*}{H} \longrightarrow RH + H, \quad (1)$$

or through a free radical formation stage

$$RH + \overset{*}{H} \longrightarrow \dot{R} + H_2, \quad (2)$$

$$\dot{R} + \dot{H} \longrightarrow RH, \quad (3)$$

where the asterisk denotes a hot atom and the dot denotes a free radical. The limiting step is the removal of the hydrogen atom, whose activation energy lies between 0.3 and 0.4 eV (depending on the bond type).[13] The recombination reaction (3) has a zero activation energy, and takes place with the participation of hot hydrogen atoms. The nonreactive collision is accompanied by loss of energy (thermalisation) of the hot atom:

$$RH + \overset{*}{H} \longrightarrow (RH)' + H, \quad (4)$$

where $(RH)'$ is an excited molecule. The probability of the collision and the amount of energy transferred in a single collision depend on the velocity of the incident particle, and also on the structural and energy characteristics of the partner molecule. When ordinary hydrogen is replaced by its isotope (tritium) the product of the substitution reaction is labelled. As a method of studying the accessible surface and the three-dimensional structure of the object tritium planigraphy is applicable only under conditions in which the chemical substitution reaction leading to the incorporation of the label occurs at the first collision between the hot atom and the hydrocarbon molecule. Obviously, this requires a source of atoms having energies which satisfy the condition $0.3 < E < 1.5$ eV and $E - \Delta E < 0.3$ eV, where ΔE is the loss of energy accompanying a nonreactive collision. In this case one nonreactive collision makes the subsequent reaction impossible, and only reaction (2) (hydrogen removal) is observed in the system.

The ideal variant would seem to be the presence of a source of monoenergetic particles having energies in the required range, but this is extremely difficult to realise in practice. The tritium planigraphy method relies on the atomisation of diatomic gases on heated metal filaments, first studied by Langmuir.[14] On the surface of an incandescent filament the gas molecules dissociate into atoms, which are emitted from the surface with an energy corresponding to the temperature of the filament. This method was used in studies[15,16] of the reactions between oxygen and hydrogen atoms.

At a filament temperature of 2000 K the average energy of the emitted atoms is 0.17 eV, and in the Maxwell spectrum of the atomic beam ~25% of the atoms have an energy sufficient to displace hydrogen ($E \geqslant 0.3$ eV).[17] The proportion of the particles having energies of 1.5 eV, and therefore able to replace hydrogen by the direct substitution mechanism [reaction (1)], is small (<3%), and therefore reaction (1) can be ignored.

Attempts to use the thermal activation method to prepare labelled compounds began in the middle 1960s. Moser et al.[18] bombarded films of sugars with tritium atoms, and obtained labelled products, but the high yield of decomposition products even then raised serious doubts as to the practical value of the method. For the same reason other similar studies (for example, early attempts to label whole granules of starch)[19] have not been developed further. It was found that the main decomposition mechanism was the use of excessively thick targets, which suppressed the removal of heat from the reaction zone. The surface of the target became heated by absorbing the luminous radiation of the filament, by heat transfer through the molecular gas, and also by the heat evolved by the reaction of the tritium atoms with the substance. It was subsequently shown that by using effective heat removal procedures almost 100% of the substance

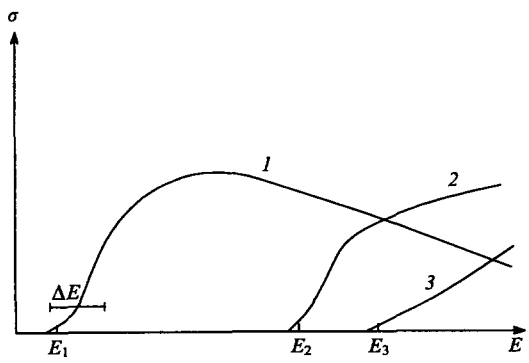


Figure 1. Dependence of the probability of a chemical reaction (σ) on the energy of the interacting particles (E) for three possible channels with thresholds E_1, E_2, and E_3: ΔE is the loss of energy for a single nonreactive collision.

could be preserved.[20] This required working with thin layers of organic materials and cooling the target with liquid nitrogen. Even after a long exposure the temperature of the target was not more than 10–15 K higher than the liquid nitrogen temperature.[21] A low target temperature (77 K) suppresses the mobility of the free radicals formed by hydrogen removal, and therefore the main path for their disappearance is the recombination with thermalised tritium atoms with formation of the labelled initial molecule [reaction (3)].

The possibility of studying the accessible surface and the three-dimensional structure of the macromolecules is due to the small penetration depth of the tritium atoms which are able to react with the target atoms. The direct measurement of the depth of penetration of the reactive tritium atoms is experimentally difficult. However, this quantity can be calculated by making certain assumptions, and it can also be obtained from some indirect results. The calculations show that for tritium hot-atom initial energies close to the threshold energy of the removal reaction even a single nonreactive collision is sufficient to transfer the particle to a region in which hydrogen removal becomes impossible. Fig. 2 shows the dependence of the fractional number of tritium atoms with an energy $E \geqslant 0.3$ eV upon the depth of penetration into a layer of organic compound. It can be seen that whereas at X = 0 (i.e. at the surface of the object or at the first collision of the tritium atom with the molecule) the fractional number of reactive atoms is 26% this value decreases to 0.4% when the atoms have penetrated within the layer to a depth of only 6 Å. In practice the free path of the reactive atoms (the thickness of the reaction layer) is even smaller, since their number in the beam passing through the target is lowered also by a chemical reaction whose cross-section is comparable to the collision cross-section in the case of atoms with an energy higher than the threshold value.[10] This means that if a macromolecule is used as the target the distribution of the label over its fragments becomes a measure of its steric accessibility for a collision with the primary bombarding particle, i.e. it characterises the surface area of the macromolecule accessible to the tritium atoms. The thickness of the reaction layer can also be estimated from the results of experiments on the introduction of a tritium label into adsorption layers of surface-active substances with hydrocarbon chains of different lengths

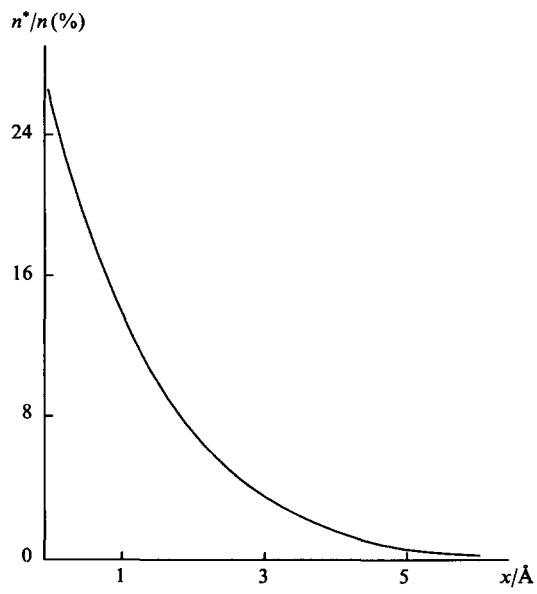


Figure 2. Decrease in the relative number of reactive particles during the passage through a layer of thickness x in a condensed organic substance (calculated for atoms with a temperature of 2000 K under conditions of elastic retardation, $E_{act} = 0.3$ eV).

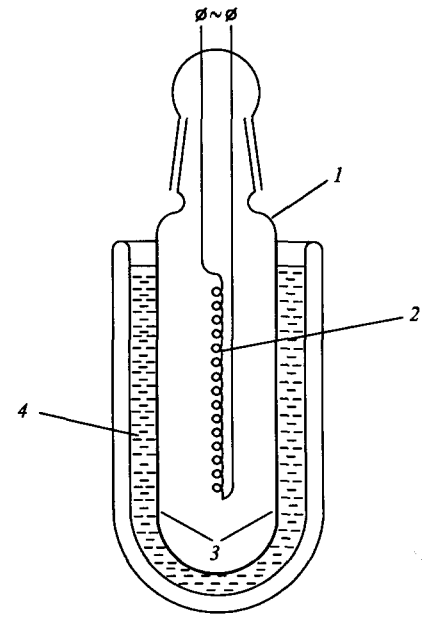


Figure 3. Glass reactor for tritium bombardment.

(from C_1 to C_{16}).[22] From the dependence of the yield of labelled product (incorporation of the label) on the length of the molecule and on its orientation at the interface it was concluded that a change in layer thickness by one methylene group, i.e. by ~3 Å, has a marked effect on the yield of labelled product. This length can be used as a quantitative characteristic of the thickness of the reaction layer. A similar value is suggested by experiments on the incorporation of tritium into myoglobin, which will be discussed below.

The apparatus used in all the experiments on tritium labelling is shown in Fig. 3. The central part of the apparatus (the reactor) is a cylindrical glass vessel (*1*) fitted with an axial tungsten wire (*2*) heated by an electric current. The target material (*3*) is deposited as a thin film on the side wall of the vessel, which is cooled in liquid nitrogen (*4*). The reactor is evacuated to a residual pressure of 10^{-5} Torr, and gaseous tritium is then admitted into the system. In the reaction vessels used in the majority of experiments (60 mm in diameter) the pressure of molecular tritium was 5×10^{-3} Torr. When the filament was heated to 2000 K the tritium was atomised on its surface, and the resulting atoms (formed under conditions of free flight to the target) were allowed to bombard the walls of the reactor. Experimental details and the sequence of operations have already been discussed in some depth,[10] and in this review we shall confine ourselves to examining the methods of preparation of the target and the general methodology adopted for analysing the distribution of the label.

The simplest method is the lyophilisation of a solution of the substance. However, if this method is chosen it is essential to ensure that the resulting film is firmly attached to the reactor walls over the whole of its internal surface. Furthermore it should not be forgotten that the lyophilic drying process may itself influence the structure of the substance under study.

In order to study the structure of the macromolecule in solution a target can be prepared by 'smearing' the experimental solution as a thin film by rapidly rotating the reactor immersed in liquid nitrogen. In many cases this method gives good results, but there are at least two aspects which limit its application. The first is the crystallisation of the water induced by the relatively low rate of cooling. When the vessel is immersed in liquid nitrogen the latter boils, forming a layer of gas bubbles close to the wall which slows down the cooling. This effect can be avoided by choosing other cooling agents (for example, liquid ethane), but working with this

gas has some undesirable features. The formation of a crystalline phase of ice must be avoided in work with materials containing 'internal' water, such as macromolecules and their complexes (including clusters) including more than 10 water molecules, liposomes, and whole cells.

The second limitation is due to the specific features of the method. Since the incorporation of the label is confined to a zone accessible to the first collision of the tritium atom incident from the gas phase, i.e. to the surface of the target, and in practical experiments the specific radioactivity is measured, it is essential to know, or at least to be able to control, the concentrations of the solute at the surface and in the bulk of the sample. In the method of target preparation described above a liquid solution exists for a certain time, giving a water–air interface. Therefore the bombardment by atomic tritium is experienced by an adsorption layer of molecules, whose parameters are determined by the adsorption process. The concentration of solute molecules in this layer is described by the adsorption isotherm. It is most convenient to work under conditions of complete saturation of the surface layer: these offer adequately reproducible results even when the surface concentration of the substance is not known. Calculations show that for solutions of macromolecules the time of formation of the saturated layer is a few minutes ($\sim 2-5$ min), which is of the same order as the time of preparation of the sample. The formation of a saturated layer is confirmed by the observed dependence of the time on the concentration of the solution.[22] In the case of amphiphilic substances the use of this method of target preparation can lead to the orientation of the molecules with respect to the phase separation boundary, and in this case the distribution of the label will represent the character of the orientation rather than the structural features of the object. In substances of low-molecular mass the orientation effects are especially clearly displayed. The orientation of macromolecules requires long times, and their amphiphilicity is weaker.

Both these difficulties (crystallisation of the water and the contribution from adsorption processes) can be avoided by rapidly freezing the solution using a 'spraying' technique. The solution is sprayed as very small droplets onto the inner surface of the reactor previously cooled in liquid nitrogen by using a specially designed nebuliser. The lifetime of each drop, and therefore also of the water–air interface, is small (equal to the time of flight from the nozzle of the nebuliser to the wall), and adsorption cannot take place. However, the rate of cooling of the drop incident on the reactor surface at 77 K is so high that the solution solidifies as a glass. Under these conditions we find a state of affairs similar to that which exists on a freshly cleaved solid surface. The concentration of molecules on the cleaved surface can be related to that in the bulk by the simple expression $C_s = C_v^{2/3}$. The dependence of the activity taken up by the sample on the concentration ($I = C_v^{0.67}$) is found to be consistent with the above relationship in experiments with both substances of low-molecular mass and solutions of proteins.[10,22,23].

Yusupov and Spirin[24] proposed an original method of target preparation which avoided the artefacts associated with the crystallisation of water and with the adsorption of the solute on the phase separation surface. A solution or suspension of the sample was introduced dropwise into liquid nitrogen. The very small particles of ice formed in this way were crushed directly in the liquid nitrogen, and this 'slurry' was deposited on the reactor walls. The liquid nitrogen quickly boiled off, and the ice particles were left firmly adhering to the reactor wall. The results obtained confirm the effectiveness of the proposed method, but its practical application requires special skills.

When considering the technique of the method we must particularly note the special factors which determine the reproducibility of the results obtained. Among customary factors, such as the precision of radioactivity measurement and of the determination of temperature of the wire, of the pressure of molecular tritium, of the exposure time, and also of the errors introduced by the chosen method of analysis, we should include the reproducibility of the total surface area of the target bombarded by the tritium atoms. When the techniques of solution spraying or evaporation from a suspension are used it is difficult to achieve good reproducibility of the total target area. To ensure a correct comparison of the results of independent experiments we must normalise the reported values with respect to the total flux of tritium atoms reacting with the target surface. Obviously, the latter quantity corresponds to the total radioactivity of the sample, including the activity of the sample itself and that of the solvent. If all the other standardisation conditions are observed (for example, the mass of the label attached to the C–H bond) the assignment of the radioactivity of the substance under study to the flux of tritium atoms should be constant for a given substance. An additional method (in some cases, the only possible method) of standardising the experiment is to use a different kind of monitor: organic substances placed at a fixed distance from the filament. Monitors of this type include various polymer films, cellulose or paper filters, etc.

In studies of proteins the transition from the amount of label included to the determination of the steric accessibility (or, in other words, of the accessible surface area) requires an allowance for the probability of the incorporation of tritium into aminoacid residues of different types under identical exposure conditions. These calibrating coefficients are needed also in all other cases of experimental samples consisting of components of different chemical nature. Several attempts have been made to determine the relative reactivities of aminoacid residues by comparing free aminoacids[25] and short peptides,[23] and determining the inclusion of the label in the aminoacid residues and in proteins with known three-dimensional structures in which the accessibility (exposure response) of the residues could be calculated from X-ray structural analysis results. Though the results of different workers, obtained for different systems, show some discrepancies they can be summarised as follows:

(1) in proteins we find a linear relationship between the amount of label incorporated and the accessible surface area calculated for the hydrocarbon fragments of the aminoacid residue;

(2) the nature of the neighbouring residue on the polypeptide chain does not affect the amount of label which is taken up (if normalised for unit area of accessible surface); this amount is also independent of the site on which the residue is located (the N or the C end of the peptide);

(3) the inclusion of the label in the residues of different aminoacids has almost equal probabilities when calculated for a single C–H bond or for unit area of accessible surface.

The results obtained by Gedrovich and Badun[23] deserve a special mention: these workers found that in the composition of di- and tripeptides the relative reactivities of three aminoacids (valine, leucine, and isoleucine) are several times higher than those of other aminoacids. This effect was ascribed to the presence (in the side group) of tertiary carbon atoms, from which the removal of hydrogen is substantially easier than in other configurations.[26] Table 1 shows the probabilities of including the label into the residues of aminoacids, established experimentally for the model protein lysozyme. We consider these results to be among the most reliable.

In conclusion we should dwell briefly on the general methodology used in the analysis of the distribution of the label in the experimental object. It varies according to the particular problem being studied, and it is determined by the required level of 'resolution'. Even the first stage [the determination of the overall amount of label remaining in the sample after removal of the labelled solvent, i.e. the sum of the labile (exchangeable) and of the stable labels (held in the C–H bond), and also the labelled water strongly bound to the surface of the macromolecule] allows us to determine the total area of accessible surface of the sample. Since the overall amount of label is usually 2–3 orders of magnitude

Table 1. Probability of including the label (K) in the aminoacid residues of lysozyme, normalised for unit area of accessible surface of the hydrocarbon fragment ($C_\alpha - H$, side groups).[45]

Aminoacid residue	K[a]	Aminoacid residue	K[a]
Lys	0.12	Pro	0.11
His	0.16	Ala	0.12
Arg	0.10	Val	0.11
Asx	0.10	Ile	0.10
Glx	0.20	Leu	0.08
Thr	0.10	Tyr	0.07
Ser	0.10	Phe	0.09
Gly	0.09		

[a] Standard deviation 20%.

greater than the amount of stable label in the C–H bonds the differences between the probabilities of interaction of tritium with aminoacids of different types are quite unimportant.

Preliminary estimates can be made, and a fairly wide range of problems whose solution requires a knowledge of the total aminoacid content of the solvent can be examined, by determining the label incorporated in aminoacids of different types. To this end the protein freed from the labile label (by dialysis, chromatography, electrophoresis, etc.) is hydrolysed, and the inclusion of the label in individual aminoacids is determined by analysing the hydrolysates and at the same time measuring their radioactivity.

Often we need to know the distribution of the label along the polypeptide chain of a protein. In this case the protein is fragmented into peptides, the mixture of peptides is separated, the peptides are identified and hydrolysed, and lastly the inclusion of label in each aminoacid is determined. Obviously, in an ideal situation, peptides in which each aminoacid occurs only once are produced by the fragmentation of the protein. In practice this state of affairs is very seldom met, but every effort is made to minimise the number of repeat occurrences. Lastly, a sequence of peptides with an analysis of the activity in each of the aminoacids removed is also possible.

Thus, by correctly preparing the target, labelling the sample, and choosing an appropriate strategy for analysis of the distribution of the label, the latter parameter can be used to construct a model of the three-dimensional structure. This possibility was first demonstrated in a study of labelled myoglobin.[27] According to the X-ray structural analysis results the N end of the protein fractions is an α-helix A, including the residues 3–18 (Fig. 4) and oriented so that part of its surface is directed outwards, and part is in the 'shadow' produced at its contact with the G and H helices where A connects them with a chain fragment.[28] The distribution of the label in the N end of the protein (aminoacid residues 1–20) shows that the label is incorporated only in the aminoacid residues pointing outwards, and not in those localised in the region of contact and therefore not accessible to the tritium. A reasonably good correlation was obtained in comparisons of the incorporation of the label with the area of the accessible surface in the same residues, calculated by the method of Lee and Richards[29] from X-ray structural analysis data. This experiment confirmed earlier reports of the dependence of the take-up of the label on the three-dimensional structure in peptides, and predicted the possibility of developing a new experimental method of studying the structure and the topography of the accessible surface of macromolecules and of supermolecular systems: the tritium planigraphy method.

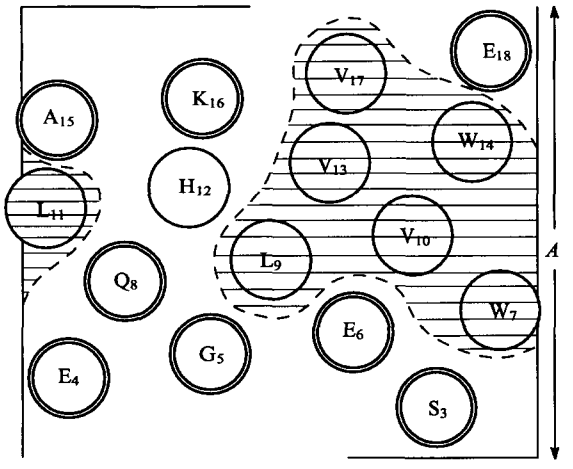


Figure 4. Schematic of the 'unfolding' of the surface of the α-helix region A of the whale myoglobin protein (residues 4–18). The one-letter code is used for the aminoacids. The cross-hatched regions are aminoacid residues sterically inaccessible to the tritium atoms (regions of 'shadow').

III. Study of the accessible surface of proteins. Limits of application of the method

The interaction of polymeric molecules takes place through groups of atoms localised on their surface. Methods proposed to describe the surface (Fig. 5) include determining the area of the macromolecule surface accessible to water molecules: this was found to be a very important concept, because this surface plays a fundamental role in the formation, stabilisation, and operation of the native structure of biopolymers. A new term was created by Lee and Richards[29] to describe it: the 'accessible surface'. It denotes the surface which can accommodate the centres of spherical test species (probes) with a radius of 1.4 Å (approximately equal to the radius of water molecules) in contact with the van der Waals surface of the macromolecule. The macromolecule itself is treated as an assembly of atoms approximating van der Waals spheres, and the coordinates of their centres are determined by high-resolution (1–2 Å) X-ray structural analysis. Lee and Richards obtained a quantitative description of the accessible surface by applying a specially developed computational algorithm, which they used to calculate the accessible surface area of three globular proteins: lysozyme, ribonuclease A, and myoglobin.[29] These calculations were subsequently applied to a series of other globular proteins, whose three-dimensional structure was thus determined with an acceptably high resolution.

The great functional importance of the surface of macromolecules, and its relationship with the characteristics of their three-dimensional structure, stimulated a large number of studies (mainly theoretical) of various aspects of the correlation of the area, composition, and topography of the surface with the globule

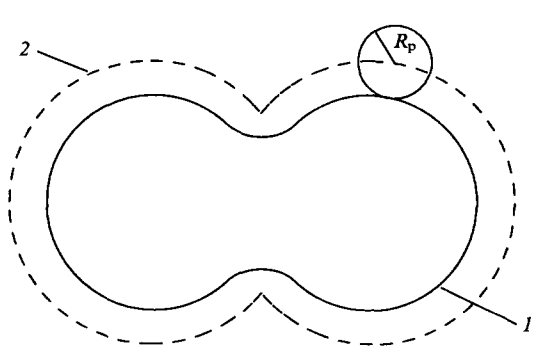


Figure 5. Description of the 'surface' by testing with a spherical probe of radius R_p in a two-centre system.[30] (*1*) contact surface; (*2*) accessible surface.

volume, the molecular mass of the protein, its packing density, and other parameters of the three-dimensional structure (see, for example, Refs 31–36).

In recent years the surface of macromolecules has been analysed increasingly by the molecular cartography method, which allows almost the whole of the protein surface to be 'viewed' simultaneously.[37] By using these charts Barlow and Thornton[38] studied the distribution of charged groups, and Fanning et al.[39] and Novotny et al.[40] studied the topography of the protein surface with the aim of localising the antigenic centres. An analysis of the surface features can be carried out by super-imposing three different charts.[41] This analysis of the nature of the accessible surface provided an unambiguous interpretation of the hydrophobic properties of proteins, calculated exclusively from their aminoacid composition. The concept of an accessible surface has proved very fruitful in the construction of models of the self-organisation of the polypeptide chain.[42–44]

Without belittling the value of the computational methods of determining the accessible surface we note that all the calculations are based on high-resolution X-ray structural analysis results. There is no doubt that the tritium planigraphy method now offers the fullest possible information on the accessible surface of the object. As stated above, this method relies on establishing the steric accessibility of the hydrocarbon groups (CH, CH_2, and CH_3) of all the component units (in the case of proteins, the aminoacid residues) in the substance under study, and provides information on its surface. As could have been expected, a simple relationship has been observed between the specific radioactivity (A) and accessible surface area calculated by the Lee and Richards algorithm (S): $A = kS$, where k is the probability of finding the label in the aminoacid residue of a given type, normalised for the accessible surface area. By using model proteins with known three-dimensional structures it was shown that this relationship applies for the majority of protein residues.[27,45] In most cases the results corresponded to those given by calculations of the accessible surface by the Lee and Richards algorithm.[29] However, some discrepancies were noted: these were associated with special features of the mechanism chosen to introduce the label and with the difference in size between the tritium atom and the water molecule. A new model, proposed to eliminate these contradictions, describes the accessibility to tritium atoms of fragments of the sample on the assumption that any scattering event must result in loss of reactivity.[46] A calculation by this model for the 20 N end residues of myoglobin gave a better agreement with previously reported[27] experimental data. The effective radius of the tritium atom (1.1 Å) was found, as expected, to be smaller than the radius assumed for a water molecule.

It should not be forgotten that the surface defined by the tritium planigraphy method differs in principle from the accessible surface in its generally accepted sense. In the surface accessible to the incorporation of the label there are no atoms not directly linked to hydrogen atoms (in the proteins these are the oxygen atoms of the carbonyl groups, =C=O, in disulfide bonds and in methionine they are the sulfur atoms). There are also no groups from which the label is readily removed by exchange with water (the −OH, −SH, $-NH_2$, =NH, and −COOH groups). The fragments which are sterically inaccessible to reactions with tritium atoms incident on the surface with a rectilinear trajectory are also not included. As a result, the surface determined by tritium planigraphy includes breaks associated with the regions containing the atoms and groups of atoms listed above, whereas the surface calculated by the Lee and Richards algorithm is free from discontinuities.

It is legitimate to ask to what extent and between what limits the tritium planigraphy data can be correlated with calculations of the accessible surface by the 'sliding' probe method. To answer this question the radius of the probe approximating the tritium atom, for which the calculation by the algorithm which assumes 'lining' of the van der Waals surface agrees most accurately with experiments on the incorporation of the label, had to be determined. It was also necessary to determine the limits between which these data can be extrapolated to test probes having other radii. These studies were carried out with the model protein lysozyme, for which high-resolution X-ray structural analysis data are available, and the accessibility profile of the aminoacids towards the introduction of the tritium label was established.[47]

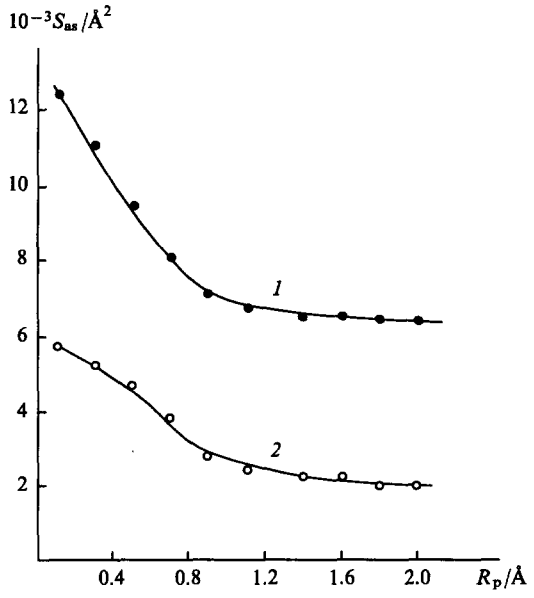


Figure 6. Dependence of area of accessible surface in a lysozyme globule upon the radius of the test probe (R_p). (*1*) Accessible surface of the globule (S^{Σ}_{as}), (*2*) accessible surface of the hydrocarbon fragments (S^{CH}_{as}).

By using a slightly modified algorithm (as compared with that proposed by Lee and Richards) the surface area of the protein accessible to probes with radii between 0.1 and 2.0 Å (and even 10.0 Å) was calculated for each aminoacid residue and for the protein globule as a whole. The calculation was carried out separately for each residue, including the 'labile' hydrogen-containing groups (S^{Σ}_{as}), and only for the hydrocarbon fragment (S^{CH}_{as}) of the aminoacid. Fig. 6 shows the dependence of the accessible surface area of the protein globule upon the radius of the test probe. It can be seen that for low radii (from 0.1 to 1.0 Å) the accessible surface decreases sharply (by ~40%) as the radius is increased, an effect ascribed to the smoothing of the surface topography. For probe radii greater than 1.0 Å the S^{Σ}_{as} value stays almost constant until it begins to increase again for the largest probes (10 Å) as a result of the increased radius of the 'sliding' trajectory. For a probe with a radius of 1.4 Å (the water molecule) the calculated value of S^{Σ}_{as} stays practically constant at the level of 6510 Å^2, which is close to the value of 6710 Å^2 obtained by Lee and Richards.[29] The best agreement between experiment and calculation is obtained by assuming a sphere with a radius of 0.9 Å as an approximation to the tritium atom (Fig. 7).

Fig. 8 shows experimental and calculated accessibility profiles for the peptide chain of the protein. Data on the inclusion of the label are presented as the ratio A/K, which is equivalent to the accessible surface area (in Å^2). It can be seen that on the whole the correlation between the two distributions is acceptably good. Accessibility calculations for probes having different radii showed that the results obtained for a tritium atom assumed to have a radius of 0.9 Å can be extrapolated with reasonable assurance to probes having radii of up to 2 Å for aminoacid residues with a high (> 50%) or a moderately high (20% – 50%) average level of accessibility if the ratio of the residue areas (in the composition of the protein molecule) to the area of a hypothetical fully unfolded

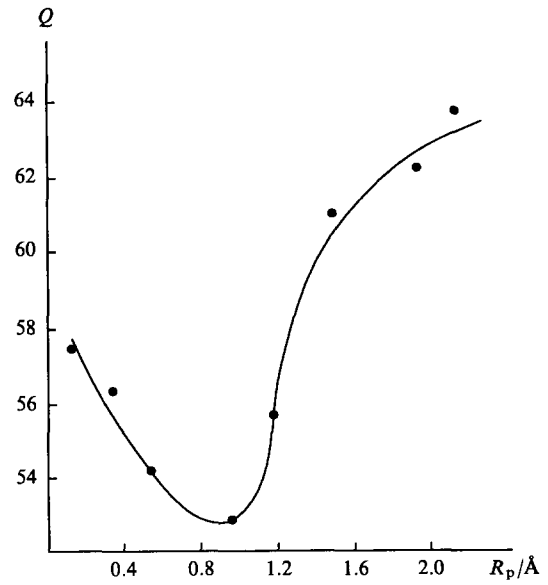


Figure 7. Deviation function Q, describing the discrepancy between calculated and experimental data for test probes having different radii. The minimum of the function corresponds to an effective radius of the tritium atom of R = 0.9 Å.

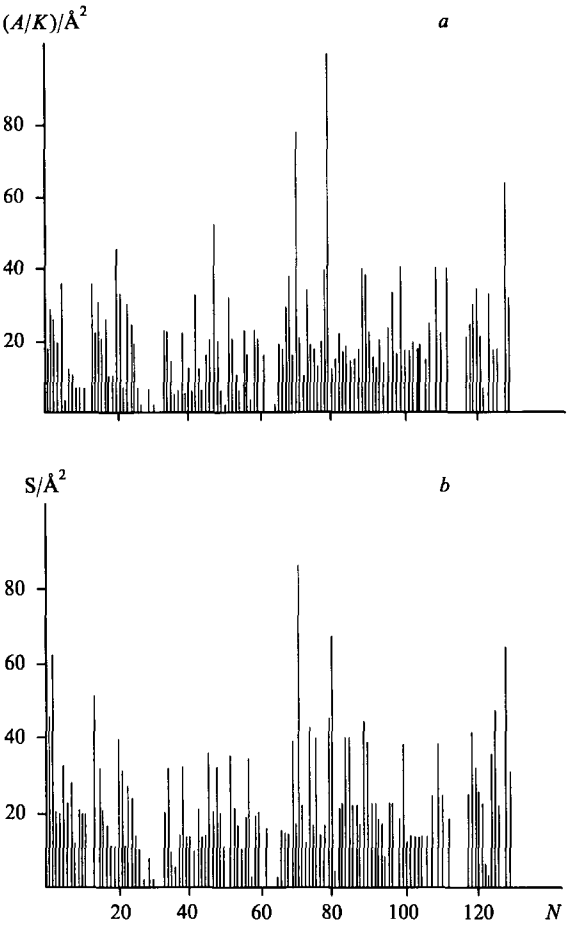


Figure 8. Accessibility profile of the polypeptide chain in lysozyme, obtained experimentally by introducing the label (*a*) and calculated for a test probe of radius 0.9 Å (*b*). A/K is the specific activity of the aminoacid residues, normalised for the probability of incorporating the label, S is the accessible surface area in Å², N is the number of the aminoacid residue in the protein sequence.

polypeptide chain for a probe with a radius of 1.4 Å is taken as a measure of the accessibility (the *trans*-peptide Ala–X–Ala is a suitable model compound). In the case of poorly accessible residues (< 20%) the data on the incorporation of the label should be interpreted with some caution. However, the proportion of these residues on the surface of the macromolecule is not large, and it cannot introduce a significant error into the final result. In the determination of the total area of the protein their contribution balances out.

In the same investigation the coefficients needed to convert the accessible surface area of the hydrogen fragment into the total area of the residue ($S_a^{CH} = \gamma S_{as}^{\Sigma}$) were determined. It was found that the parameter γ depends not only on the nature of the residue (which is not surprising) but also on its level of accessibility in the tertiary structure. However, at the level of the entire globule this quantity remains constant over a wide range of probe radii (0.4 ± 0.06). We note that according to Chothia[81] the fractional amount of accessible surface formed by nonpolar atoms (mainly, the hydrocarbon fragments) depends on the molecular mass of the protein: it decreases as the latter increases. In the work of Lee and Richards[29] this fraction was 0.392 for lysozyme, which corresponded to the value obtained by Bogacheva et al.[47]

Gedrovich et al.[55] suggested an approach which allows quantitative estimates of the total accessible surface area of the protein from the amount of label included in the aminoacid residue groups which are free to exchange and in the water molecules firmly bound to the surface (hydration shell). They examined a group of proteins having known three-dimensional structures (lysozyme, ribonuclease S, myoglobin, papain, subtilisin, and α-chymotrypsin), and obtained a linear dependence in $\lg I - \lg S$ coordinates. The straight line thus obtained can be treated as a calibrating plot, and used to determine the accessible surface area of globular proteins of unknown structure (Fig. 9). Calculations of this type carried out for the globular protein parvalbumin III found in the pike gave an accessible surface area of 5700 ± 900 Å². Using the empirical dependence of the accessible surface area on the molecular mass of the protein gave[34] an area of 5800 Å².

The results of experiments with proteins whose three-dimensional structure had been determined by high-resolution X-ray structural analysis allowed the method to be extended to objects on which structural information was inadequate or totally absent.

Thus an analysis of the specific labelling of aminoacids in the overall hydrolysate of the protein led to reasonable conclusions on various parameters of the periphery of human serum albumin,[45,49] for which available X-ray structural analysis data had a resolution of 6 Å which was insufficient for the calculation.[48] It was found that (as compared with that of lysozyme) the accessible surface of the albumin was enriched in hydrophobic aminoacid residues, which accounted for 63% of the total aminoacid content

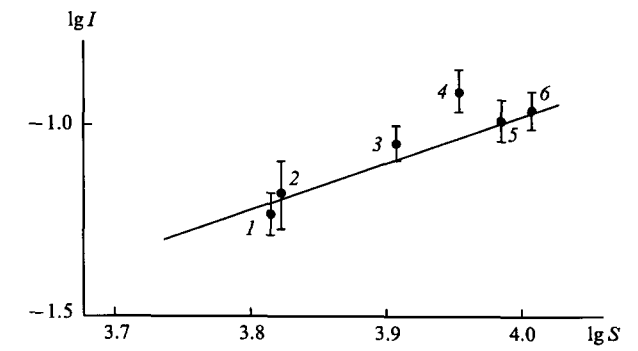


Figure 9. Dependence of the reduced radioactivity on the accessible surface area of proteins. (*1*) Lysozyme, (*2*) ribonuclease S, (*3*) myoglobin, (*4*) papain, (*5*) subtilisin, (*6*) α-chymotrypsin. The I value includes the label in the C–H bonds, in the labile groups, and in the strongly held water, normalised for the flow of tritium atoms.

of the surface (whereas in lysozyme the amount was 18%). Summing the accessible surface areas in the aminoacids and allowing for the number of residues of each type gives[45] a total area $S_{as}^{CH} = 19\,900$ Å^2. Using the ratio $S_{as}^{CH}/S_{as}^{\Sigma}$, whose value is 0.4 for lysozyme, gave a total area of 49 000 Å^2 for the protein. Using the empirical dependence of the accessible surface area of a fully unfolded polypeptide chain on the molecular mass gave a coefficient of 0.48 for the degree of unfolding of the polypeptide chain. By treating the albumin globule as an ellipsoid of revolution with axes of 140:40:40 Å important structural parameters of the surface were determined:[50] for example, the 'roughness coefficient' of the native protein was found to be 2.7. For the lysozyme molecule the corresponding unfolding and roughness coefficients were 0.68 and 1.7. From these results we may conclude that the packing is less compact (and the surface topography is more broken) in the polypeptide chain of human serum albumen than in lysozyme, in good agreement with the known physicochemical properties of this protein and above all with its strong hydrophobicity.[51] The openwork appearance of its three-dimensional structure is confirmed by many data on the capacity of albumin to bind low-molecular hydrophobic substances (for example, fatty acids), and is consistent with reports of the presence of several hydrophobic cavities in the molecule of this protein.[52] The rough packing of the albumin globule is also consistent with views on the flexibility of this molecule and on its tendency to undergo volume fluctuations.

Bobkova et al.[53,54] reported a comparative study of the surfaces accessible to label introduction in two important enzymes of protein biosynthesis (aminoacyl-tRNA-synthetase). The two enzymes (phenylalanine-tRNA-synthetases from *E. coli* and *Thermus thermophilus*) have complex quaternary structures. The similarity in structure (at least, in the region of the active centres of these enzymes) suggests that the thermophilic enzyme catalyses the aminoacylation reaction of tRNA from *E. coli*, whereas the enzyme from *E. coli* can catalyse the aminoacylation of the thermophilic tRNA. The enzyme from *Thermus thermophilus* was obtained in the form of crystals suitable for use in X-ray structural analysis, but attempts to prepare crystalline phenylalanine-tRNA-synthetase from *E. coli* were unsuccessful. In these circumstances the tritium planigraphy method is one of the few methods available for studying the surface structure of the enzyme from *E. coli* and for comparing it with that of its thermophilic analogue. Fig. 10 shows the diagrams defining the surface areas of the aminoacid residues for the two proteins, obtained from data on the specific activity of the aminoacids and of the total hydrolysates of the proteins. We note a fairly high level of hydrophobicity in the surface of the proteins (60% for the enzyme from *E. coli* and 67% for the thermophilic protein). Comparing the data for the degree of hydrophobicity and for the overall sum of hydrophobic residues in the two proteins shows that this distribution is approximately the same for the mesophilic and the thermophilic synthetase being studied. The proportionate increase in hydrophobicity on the surface and in the inner parts of the protein globule on passing from the mesophilic to the thermophilic enzyme is a very interesting effect. However, the possibility of identifying the difference in the aminoacid residues localised on the surface of the protein is even more important. The increased accessibility of the arginine residues in the identifiable composition of the aminoacyl-tRNA-synthetase led to an explanation[53,54] of the increased sensitivity to exposure of the charged residues on the globule surface in terms of the tendency of the thermophilic tRNA-synthetase to compensate the destabilising action of the high temperature by using additional electrostatic interactions with the substrate.

Similar studies had been previously carried out with the Fe- and MoFe-proteins of the azotobacter nitrogenase.[10] Without discussing these results in detail we shall only point out that they yet again confirm the possibility of obtaining valuable structural information from a relatively simple experiment.

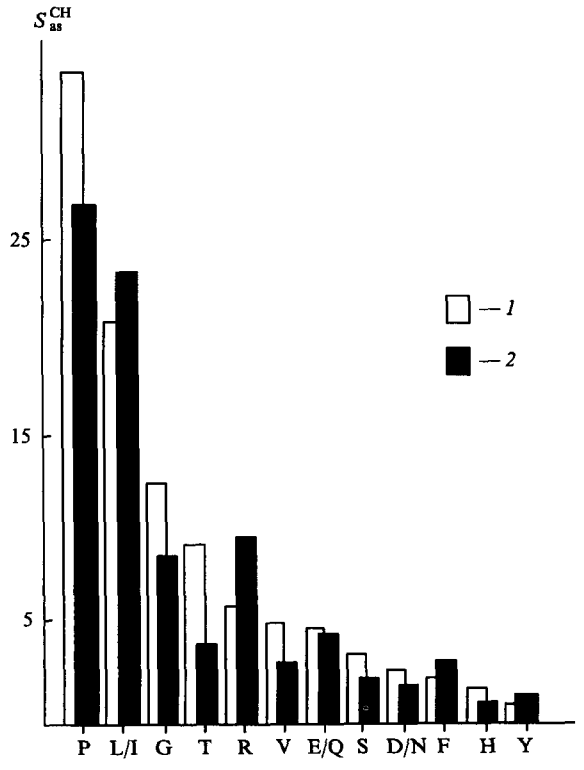


Figure 10. Aminoacid composition of the accessible surface phenylalanine–tRNA–synthetase from *E. coli* (*1*) and from *Thermus thermophilus* (*2*) expressed as a percentage of the total area of the globule.

IV. Conformational transitions in proteins

Obviously, any changes in the three-dimensional structure of a macromolecule must affect the topography of its accessible surface. This suggests very promising applications of the tritium planigraphy method in studies of conformational transitions. Volynskaya et al.[56,57] demonstrated this prospect in a study of the denaturation of lysozyme by 8 M urea with reduction of the disulfide bonds. Distribution of the label on the aminoacid residues in the denatured form differs markedly from the native distribution and approaches random distribution. The latter was calculated by using the tripeptide Ala–X–Ala as a model,[29] allowing for the probability of incorporating the label. Of course, this model can only be accepted as a crude approximation. In the analysis of the real structure of the denatured state we must allow for the mutual screening of the residues by their nearest neighbours, and also for the possible orientation of the polypeptide chain on the surface of the target. Nevertheless the calculated and the experimental distribution of the label in the fully denatured form are reasonably close.

We note that, under the influence of 8 M urea only, a state having a well-defined structure is formed without breaking the S–S bonds. It is characterised by an accessibility of the hydrophobic aminoacid residues greater than in the native material, coupled with a small change in the accessibility of the hydrophilic groups. This is interpreted[56,57] as a breakdown of the hydrophobic contacts although the globular structure as a whole is retained. The interpretation is consistent with independent data on the viscosity of the protein solutions in the presence of 8 M urea, on the circular dichroism (CD) spectra, and on optical rotary dispersion.[58,59]

The high sensitivity of tritium planigraphy led to the detection of intermediate states of the protein, formed as a result of weak denaturing interactions and not revealed by traditional

methods.[56] One of these states can be induced by adding 1 M lithium chloride and is characterised by an accessibility of the polar aminoacid residues substantially lower than in the native conformation, while the accessibility of the hydrophobic residues is almost unchanged. This transition can be treated as a 'compression' of the globule. Conversely, another state is characterised by an increased accessibility of all the types of aminoacid residue, and is induced by the addition of 1 M urea and guanidine·HCl. This transition is similar to an unfolding of the globule while retaining all the other structural characteristics (viscosity, CD spectra, etc.). Obviously, in this case we are dealing with a slight 'loosening' of the packing of the elements in the secondary structure without dissociation and without changes in the overall architecture of the globule. A similar type of intermediate state was detected also in ribonuclease.[60]

The conformational transitions observed in lysozyme have been compared with the changes in enzyme activity. By using the rate of lysis of *Micrococcus lysodecticus* as a test it was shown that the 'compression' and the 'loosening' of the lysozyme globule result in a decreased enzyme activity, the 'compression' effect being the stronger. Similar changes were observed in ribonuclease A.[60] Low salt concentrations (0.1–0.2 M) produced compression of the protein globule, accompanied by a rapid decrease in the enzyme activity. Under these conditions the intramolecular mobility of the protein, monitored as its mean-square deviation $\langle X^2 \rangle$ from Rayleigh scattering data for the Mössbauer radiation, is also substantially lowered. Unlike the addition of salts, low concentrations of urea produce a loosening of the protein globule also accompanied by a decrease in the enzyme activity. In this case complete loss of enzyme activity was observed only in the presence of 8 M urea.

The results of studies by tritium planigraphy of pH-dependent conformational transitions in proteins call for a detailed discussion. These changes often play a key role in the functioning of macromolecules, and in this case changes in the pH of the medium are one of the factors which control their biological activity.

Human serum albumin is one of the most important proteins in blood plasma, whose function is mainly to control the transport of various kinds of biologically active substances. It is known to undergo various pH-dependent conformational changes, of which the most striking are the N–F and the N–B transitions, taking place in acid (pH 3.5–4.5) and in alkaline media (pH 7.0–9.0) respectively.[61,62] In spite of extensive studies the nature of these structural transitions is not well understood.

A comparative study by tritium planigraphy of the accessible surface of a protein in the N, F, and B form produced some valuable structural information.[63,64] It was found that in the N–F transition the total accessible surface is increased 2-fold, in good agreement with the change in size of the molecule, determined by low-angle scattering and by viscosity measurements.[65,66] Broadening of the globule in the F form is indicated by an increase in the ratio of the molecular axes[67] from 3.3 to 5.8, and also by a 50% increase in the hydration of the molecules.[68] In the F form of the protein the hydrophobic residues are more accessible to labelling than in the N form. This effect is well correlated with the greater hydrophobicity of the protein in the F form as compared with the native form. An analysis of the distribution of the label in the B form of the protein indicates a greater accessibility of the aminoacid residues in this form than in the native state, but the scale of this increase and the number of residues undergoing a change in accessibility are much smaller than in the N–F transition.

Unfortunately in all the above studies the label incorporated in the sum of the aminoacids of various types was determined, and no attempt was made to analyse the distribution of tritium over the chain of the protein. This allowed the occurrence of a conformational transition to be established, and the most general characteristics of the states thus produced to be identified, but none of the information was helpful in identifying the detailed structure.

Interesting results were obtained in studies of pH-dependent conformational transitions in hemagglutinin, a functionally important protein of the virus membrane group. It is now firmly established that the fusion of the viral and of the endosomal membrane induced by pH changes (the most important contribution to the infection of the cell) takes place with the participation of this protein.[69] However, in spite of wide-ranging studies by biochemical, biophysical, genetic, and immunologic methods,[70] the molecular mechanisms and the intermediate states of the fusion of the membranes are still a problem to be solved. Glushakova et al.[71] proposed a sensitive experimental model for studying pH-dependent conformational changes in proteins leading to their interaction with the membrane.[71]

The model system consisted of liposomes with the ionophore nigericine attached to the lipid membrane and containing hemagglutinin in their bulk. Lowering the pH of the surrounding medium produced a corresponding change in the pH within the liposome and induced a transition of the protein to a conformation able to interact with the liposomal membrane. Tritium labelling was used to visualise the outer surface of the liposomes. By comparing the specific labelling of the protein and the distribution of radioactivity in the aminoacids in control samples (at a neutral pH) and in the experimental samples (at the fusion pH) of the liposomes the authors were able to explain the nature of the protein–membrane interaction. It was found that in an acid medium the specific activity of the protein is increased (the activities of the hydrophobic aminoacids Ile, Phe, and Tyr increased 2–3-fold). This indicates the introduction of a protein molecule (or a part of it) into the membrane, and possibly also that part of the molecule is exposed on the outer surface, and it confirms the widely held opinion that the hydrophobic N-end peptide in a small subunit of the protein (the so-called fusion peptide) is responsible for the interaction of the membranes. An analysis of the intramolecular distribution of the label along the protein chain from control and from experimental liposomes should unambiguously establish the detailed pattern of the lipid–protein interaction. The model can be treated as universal for studies of lipid–protein interactions of any type, since any protein (and even an entire virus) can be incorporated within the liposome.

The tritium planigraphy method has been used to study the pH-dependent conformational change of hemagglutinin in the composition of the whole virus, leading to its inactivation. So far attempts have been made to calculate the change in the conformation of the hemagglutinin after acid preincubation of the virus from the change in the antigenic properties of the protein, in its sensitivity to proteinases, and above all from changes in the exposure of its 'fusing peptide'. An attempt to study this phenomenon by tritium planigraphy was reported by Ksenofontov et al.,[72] who showed that on the one hand the small subunit of hemagglutinin in the virus inactivated in an acid medium is readily accessible, as shown by the approximate doubling of its specific labelling as compared with the active virus; on the other hand inactivation in an acid medium of the target without membrane does not produce a strong exposure of the 'fusion peptide'. These results led the authors to propose a new model of the conformation of hemagglutinin in the composition of the inactive virus.

V. Accessible surface and intramolecular dynamics of proteins

Several published papers suggest a relationship between the intramolecular dynamics of biopolymers and their functional properties.[73] Clearly, the intramolecular mobility (at least the large-scale segmental mobility) should be reflected by the structure of the accessible surface. However, only a few calculations aiming to demonstrate this relationship have been published so far. In particular, Levit[74,75] showed by a computer modelling experiment that in dynamic calculations the accessible surface of several aminoacid residues may differ from the static value by

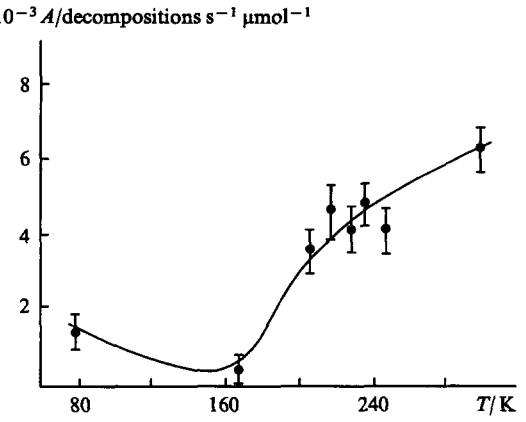


Figure 11. Temperature dependence of the inclusion of the label in lysozyme. *A* is the radioactivity of the protein normalised for the activity of the monitor.

50%–60%. No direct experimental data on the relationship between the accessible surface and the dynamic parameters have been published so far.

We know that the intramolecular mobility of biopolymers is substantially dependent on two parameters: the water content of the sample and its temperature.[76,77] This suggests the possibility of studying the relationship between the accessible surface and the molecular dynamics by the tritium planigraphy method by analysing the distribution of the label for different target temperatures and a constant humidity, or vice-versa.

A similar attempt was reported by Bogacheva et al.,[78] who chose the protein lysozyme as the model substance and determined the temperature dependence of the incorporation of the label in both the protein as a whole and the aminoacid residues of each type. The target temperature was varied between 77 K and room temperature.[79] In Fig. 11 we show the temperature dependence of the specific radioactivity of the protein, normalised for the flux of tritium atoms (*A*) in the temperature range 77–293 K. Two regions can be distinguished in the curve: a decreasing incorporation of the label at 77–163 K followed by an increase at 163–293 K. The increased level of labelling above 163 K cannot be explained in terms of the temperature dependence of the rate constant of the reaction leading to an increased incorporation of the label, since the temperature of the hot tritium atoms incident on the target surface (2000 K) is much higher than the temperature of the sample. An increase in temperature from 77 to 293 K can increase the incorporation of the label by not more than a factor of 1.2, but the observed increase in activity is 3-fold. According to data on the Rayleigh scattering of the Mössbauer radiation, and by using paramagnetic probes, it was shown[76,80] that an 'unfreezing' of the intramolecular mobility takes place at 170 to 190 K. This observation is well correlated with the observed[78] increase in the accessibility. The greater incorporation of the label (corresponding to an increased accessible surface area) upon lowering the temperature from 163 to 77 K is inexplicable in terms of dynamic effects. It may be due to a change in the three-dimensional structure of the protein as a result of the freezing of the internal water, firmly bound to the protein, which destroys the balance of forces stabilising the three-dimensional structure and leads to its breakdown.

The temperature dependence of the probability of incorporating the label was studied also at the level of the individual aminoacid residues. According to the features of this dependence all the residues can conveniently be divided into three groups. Most of the residues of the first group, whose accessibility increases in the range of 170–293 K, lie outside the region of enhanced mobility of the polypeptide chain. They are usually located on the N and C ends of the structured regions, near the sites of rotation of the polypeptide chain. The residues of the second type have a lower accessibility, independent of the temperature in this region. As a rule they are found in the middle of α-helix regions, in the immediate vicinity of the residues forming the inner region of the protein globule. Lastly 70% of the residues, which form the regions of enhanced mobility, are residues of the third type. Their accessibility changes only slightly in the range 160–270 K, but increases steeply at 293 K.

Unfortunately these results, like the results of the studies of conformational transitions discussed above, are averaged over the types of aminoacid and do not allow a detailed picture to be established. We are only at the first stage in the attempt to show that the accessible surface is a function of the molecular dynamics and that this relationship can be studied by applying the present method.

VI. Modelling the three-dimensional structure of proteins

As was stated above, X-ray structural analysis and the NMR method are laborious and have a number of fundamental limitations which seriously hinder the study of the three-dimensional structure of macromolecules. These shortcomings stimulated numerous attempts to develop theoretical predictions of the tertiary structure of proteins on the basis of data on the primary structure, of a statistical and evolutionary analysis of objects having a known structure, etc.[82–85] Unfortunately we must conclude that (in contrast to the methods of predicting the secondary structure) the theoretical modelling of the three-dimensional structure is still an unresolved problem, even at the level of predicting the general outline of the polypeptide chain.

Clearly, the most promising approaches are those which use, in addition to the theoretical predictions, parameters obtained directly from experiments (for example, data on the topography of the accessible surface). In fact, since an unambiguous linkage exists between the three-dimensional structure and the accessible surface, and since this linkage can be calculated from the coordinates of atoms, it is legitimate to formulate the inverse problem, namely to construct a model of the three-dimensional structure of the protein molecule from the accessibility of the constituent aminoacid residues. The algorithm for the solution of this problem can be defined in the following general form:

(1) determine the accessibility profile of the aminoacid residues of the protein molecule labelled with tritium;

(2) predict the elements of the secondary structure of the protein by using existing theoretical approaches;

(3) find a method of packing together the elements of the secondary structure while ensuring their surface localisation on the model of the chain fragments accessible to the tritium atoms, and (conversely) identify the inaccessible regions within the model of the globule.

Modelling greatly facilitates the acquisition of additional information: for example, data on the number and localisation of disulfide bonds, on the binding sites of ligands, substrates, and metals, on the localisation of individual functional groups, and antigenic mapping data. Taken as a whole this information may be sufficient for the choice of the optimum method of packing the elements of the secondary structure at all stages in the assembly of the model.

The first attempt to construct a model of the three-dimensional packing of a polypeptide chain was undertaken for the protein in the shell of the tobacco mosaic virus.[86] The simplicity of the object, containing only four α-helix portions, made it easy to construct a three-dimensional model satisfactorily consistent with the X-ray diffraction analysis data.[87]

More recently the scope of this approach was tested on the globular protein, pike parvalbumin, belonging to the family of Ca^{2+} binding proteins[88] (a relatively small protein containing 108 aminoacids). Two widely used methods were used to calculate the elements of the secondary structure: those of Chow and Fasman[89]

and Lim.[90] The five-predicted α-helices included 69 aminoacids, corresponding to 64% of the whole number of aminoacids. This feature greatly facilitated the modelling. Models of the three-dimensional structure of the protein were constructed by the method of Ptitsyn and Rashin,[91] which includes as an essential step a search for the contacts between neighbouring α-helices (using rigid cylinders as an approximation to the α-helices in this case). The stacking of the polypeptide chain was carried out by first searching for the contacts between the inner C and D helices (classified as inner from the standpoint of their accessibility to tritium labelling), and then successively adding to the resulting complex the helices E, F, B, and A. The choice of the relative orientation of the helices was helped by a knowledge of the calcium binding regions, which imposes strict limitations on the mutual disposition of the residues which coordinate the calcium ions.[92] The model of the three-dimensional structure of pike parvalbumin (Fig. 12) was found to be very similar to that of the homologous protein carp parvalbumin.[93]

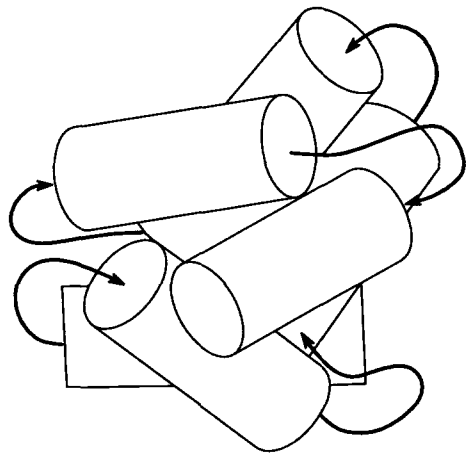


Figure 12. Model of the three-dimensional structure of the protein, pike parvalbumin III.

The similarity in structure of these two proteins is seen also in the similar topographies of their accessible surfaces, observed by comparing calculated data for carp parvalbumin with experimental data for pike parvalbumin.[87]

The results examined in this section cannot be treated as a complete concept, allowing three-dimensional models of proteins to be constructed from a combination of experimental data on the topography of the accessible surface and of theoretical predictions of the elements of the secondary structure. However, there are good reasons to hope that the further development of this approach and the application to it of the appropriate elements of computer modelling will provide researchers with a new and fruitful method of obtaining information on the three-dimensional structure of macromolecules.

VII. Three-dimensional organisation of supermolecular systems

The relationship between label distribution and structure of the accessible surface demonstrated above for proteins is seen even more clearly in studies of supermolecular systems, in which the mutual screening effects of the components play a substantial role. Depending on the experimental plan, information was obtained on the most general characteristics of the structure (general architecture of the complex, relative distribution of the components of the system in three dimensions, localisation of the contact surfaces, etc.) as well as on structural details such as determining the accessibility of the individual monomer units of the protein in the composition of the complex system.

As far as can be established the first attempt to use the tritium planigraphy method to study the three-dimensional structure of a complex supermolecular object was the tritiation of the ribosome from *E. coli*. As we know, a ribosome is a nucleoprotein complex which contains two molecules of highly polymeric RNA: 16S-RNA in the smaller moiety and 23S-RNA in the larger. The composition of the large 50S-subparticle also includes the low-molecular 5S-RNA. In the small 30S-subparticle, 21 proteins are associated with the 16S-RNA, and the large subparticle contains 32 different proteins.

By comparing the amounts of label incorporation in the proteins of the isolated subparticles of the 30S and 50S ribosomes and in the whole 70S ribosome, the proteins exposed on the surface could be identified and the absence of proteins on the contacting surface of the subparticles was demonstrated.[24,94-96] As was shown in separate experiments, under these conditions the RNA molecules retain their integrity and their ability to undergo specific interactions with proteins.[97] However, when considering the absence of proteins from the contacting surface of ribosomal subparticles, we should not forget that in practice the contact of two ribosomal subparticles is not very tight and cannot prevent the incorporation of tritium in the protein in the region between the subunits.

An interesting attempt to determine the accessibility of a polypeptide chain synthesised on the ribosome has been reported.[98,99] It has been suggested that the first units of the growing peptide are accommodated in a channel whose geometry allows the polypeptide chain to unfold into a helix. However, experimental approaches aimed at demonstrating the existence of the channel mostly relied on large macromolecules (protease and its antibody) as test reagents applied to the proteins synthesised by the ribosome. It was found that the small size of the tritium atom offers information with the required detail. The data on the inclusion of the label showed that a polypeptide chain already three or more residues long is exposed on the ribosome surface. Unfortunately these workers used a synthetic poly-(Y)-dependent cell-free translation system, i.e. they did not allow for the initiation of the translation in the form found in the translation of natural mRNAs. This may account for the discrepancy between the results and current views on the presence of a channel.

Another example of a nucleoprotein complex which has been studied by tritium planigraphy are the chromosomes from mouse fibroblast.[100] It was found that in the metaphase chromosomes both the chromosome proteins and the RNA are accessible to tritium. An analysis of the take-up of label by the protein showed that the 'scaffold' proteins are the most exposed, which suggests that they are located on the outer surface of the metaphase chromosomes.

A substantial volume of work has been devoted to applying the tritium planigraphy method to studies of the three-dimensional architecture of protein complexes having many subunits. Thus, Mironov and coworkers[101,102] tried to determine the relative accessibility of the protein subunits in the composition of cytochrome oxidase and in its complex with cytochrome C. Cytochrome oxidase is the terminal enzyme of the respiratory chain of living cells. It is a transmembrane protein, consisting (in the higher mammals) of 12 different subunits. Tritiation of the free enzyme and of the enzyme incorporated in the complex was carried out in order to establish, confirm, and add some detail to the existing model of the structural organisation of the subunits in cytochrome oxidase, and also to identify the regions of interaction of the enzyme with cytochrome C. Comparing the specific labelling values of the subunits showed that in the formation of the complex the cytochrome molecule is partially screened by the subunits I, II, VIc, and VIIa–VIIc of the protein, which agrees with previous reports.[103]

Applications of tritium planigraphy to the study of membrane structures (in particular to the study of the topography of transmembrane proteins) are of special interest because the presence of a lipid bilayer should lead to marked differences in

the accessibility to tritium between the exposed and the intramembrane parts of the polypeptide chain. In fact, as was shown in the experiments with liposome loaded with hemagglutinin discussed above, the bilayer lipid membrane weakens the flow of tritium atoms by a factor of 10.[72] The cell membrane almost completely screens the contents of the cell from the incident tritium atoms. Thus, in the case of erythrocytes it was observed that the level of incorporation of the label into the hemoglobin within the cell is more than two orders of magnitude lower than the same level in the free protein.[104]

The first example of the study of membrane structures by tritium planigraphy was applied to bacteriorhodopsin.[104-106] This is one of the best characterised membrane proteins, for which the type of three-dimensional organisation has been established in general traits and various exposed residues have been identified (see, for example, Ovchinnikov et al.[107]). This work allowed the prospects of the method to be assessed, and (in the case of favourable results) it also allowed additional experimental data, needed for checking and more detailed characterisation of existing models of the protein location in the membrane, to be obtained. A packing scheme (Fig. 13) was proposed by analysing the incorporation of label in the aminoacid residues of the protein in the membrane composition, and shown to agree satisfactorily with the model constructed from theoretical predictions.[108] As a result of this work the positions of the tryptophan residues were accurately established (positions 10 and 12), and shown to lie near the surface rather than within the membrane, as had been suggested earlier.[109,110] The new model allows for the partial screening of the fragments not included in the composition of the intramembrane α-helix segments.[111]

The fact that residues within the membrane, at least 5 Å distant from the surface (i.e. at a distance greater than the mean free path of the tritium atoms), nevertheless include some tritium can be attributed to the number of lipids in the purpuric membrane (10 molecules per molecule of protein[112]) being insufficient to make the inner part of the polypeptide chain wholly inaccessible to tritium.

Ostrovskii et al.[113] tritiated the intact cells and the protoplasts of *Micrococcus luteus* bacteria with the aim of localising the components of the membrane (potential antigens) exposed on the cell envelope. It was shown that over 90% of the non-exchanged tritium is included in cytoplasmatic membranes, and that 98% of this amount is associated with lipids, suggesting good prospects for studying the asymmetric distribution of lipids in the cell membrane. It was also found that in the case of protoplasts the tritium is incorporated preferentially in three membrane proteins, but in the case of cells in a single protein.

In most of the work discussed above the information on the structure of the supermolecular object was limited to the determination of the accessibility of whole molecular complexes (with the exception of bacteriorhodopsin). A more detailed study was carried out with the tobacco mosaic virus (TMV). Data for the virus from an X-ray diffraction analysis[114,115] with a resolution of 3.6 Å, and a detailed antigenic characterisation of the virus capsid,[116] show that this object can be used as the ideal model system.

Whole virus particles were exposed to tritium bombardment.[117-119] In Fig. 14 we show the distribution of the label over the aminoacid residues of the polypeptide chain of the protein.

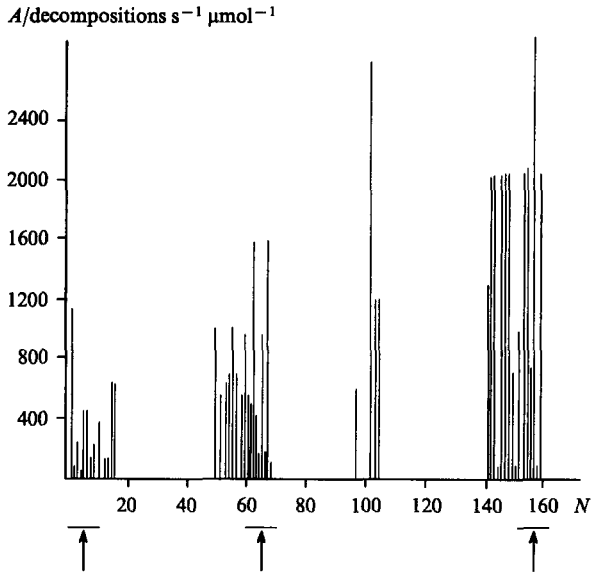


Figure 14. Experimental accessibility profile of the polypeptide chain in the protein of the tobacco mosaic virus shell in the virion composition. *N* is the number of the aminoacid residue in the polypeptide chain.

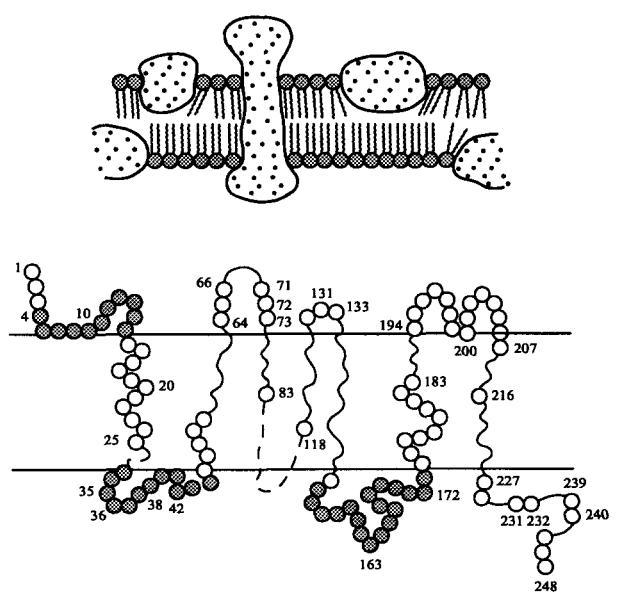


Figure 13. Schematic diagram of the 'placement' of the polypeptide chain of bacteriorhodopsin in the membrane. Filled circles denote aminoacid residues in which incorporation of the label has been observed. The sections 25-33, 66-71, 83-118, 207-231, 232-239, and 240-246 have not been analysed.

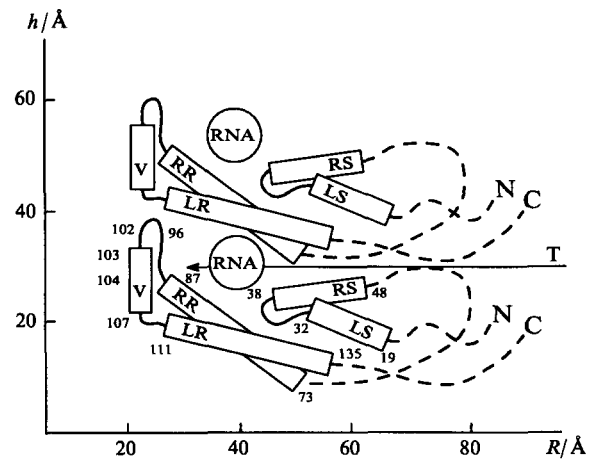


Figure 15. Schematic diagram of the placement of protein subunits in the TMV virion, demonstrating the possibility of penetration by the tritium atoms into the inner region of the virion.[115]

Relatively extensive regions can be observed, in which the aminoacid residues are not accessible to the tritium atoms. Fig. 15 shows a schematic view of the packing of the polypeptide chain for the protein coat of the virus. Comparing the experimental results on the distribution of the label in the protein with those given by this scheme shows that the N and C end fragments of the protein, located on the outer side of the virus particle, are the most accessible; the part of the protein not accessible to the labelling is found in the contact region between subunits. Surprisingly, at first sight the label was found to be included in residues 93–112 of the polypeptide chain, which (according to the X-ray structural analysis data) is remote from the surface of the virus and in contact with the RNA molecule. The authors suggested that channels are present at the 'joint' between three shell protein subunits, and that the tritium atoms can penetrate into the bulk of the virion and interact with the 93–112 residues and the virus RNA located on its 'bottom'. On the X-ray diffraction pattern these regions, 3–5 Å in diameter, appear as regions of low electron density.[114–116] It appears that channels of this type have been observed by Neiman et al.[120] in a study of the structure of the MS2-phage. According to their results 25% of the label is combined with RNA within the bacteriophage, which also suggests the presence of a relatively large number of 'pores' in the capsid of the phage between the protein molecules, through which the tritium atoms can interact with the internal regions of the phage.

It is of interest to compare the information on the surface of the virion obtained in this way with data on the localisation of the antigenic determinants. In Fig. 14 we have marked the regions 1–10, 62–68, and 153–158, which are the antigenic determinants of the protein in the composition of the virus.[116] It can be seen that they all fall among the fragments accessible to labelling by tritium. There can be no doubt that the tritium atoms are not the best indicator of the location of antigenic determinants. The small size (~1 Å) of the tritium atoms allows them to 'see' too much. However, a knowledge of the surface regions undoubtedly helps this work by narrowing the circle of possibilities.

The results obtained for the model object stimulated studies of a virus on whose structure (mainly the conformation of the subunits of the shell protein and their packing in the virion) very little was known. This was the X potato virus (XPV), one of the most widespread members of the potato virus group. Its flexible rod-like virion contains single-chain RNA, and the capsid of the virion is composed of identical protein subunits containing 236 aminoacid residues. When this work was begun data on the surface of the virion were scarce, and to some extent contradictory. Immunological studies indicated the localisation of the N-end region of the protein on the virion surface.[121] By predicting the elements of the secondary structure of the protein in the shell of XPV, using the three-dimensional structure of TMV as a first approximation, Sawyer et al.[122] were able to postulate a certain generality in the formation of the architecture of virions, namely the surface localisation of the N and C end regions in all the rod-like vegetable viruses. Subsequently Schukla et al.[123] suggested that the deflection of the polypeptide chains in proteins was a general phenomenon, leading to surface localisation of their N and C end regions in all the rod-like vegetable viruses.

In a comparative study by tritium planigraphy of intact XPV particles and of XPV degraded in situ (lacking the first 19 aminoacids through the action of the protease in the cell sap[126]) it was shown that the distribution of the label (Fig. 16*a*) differs

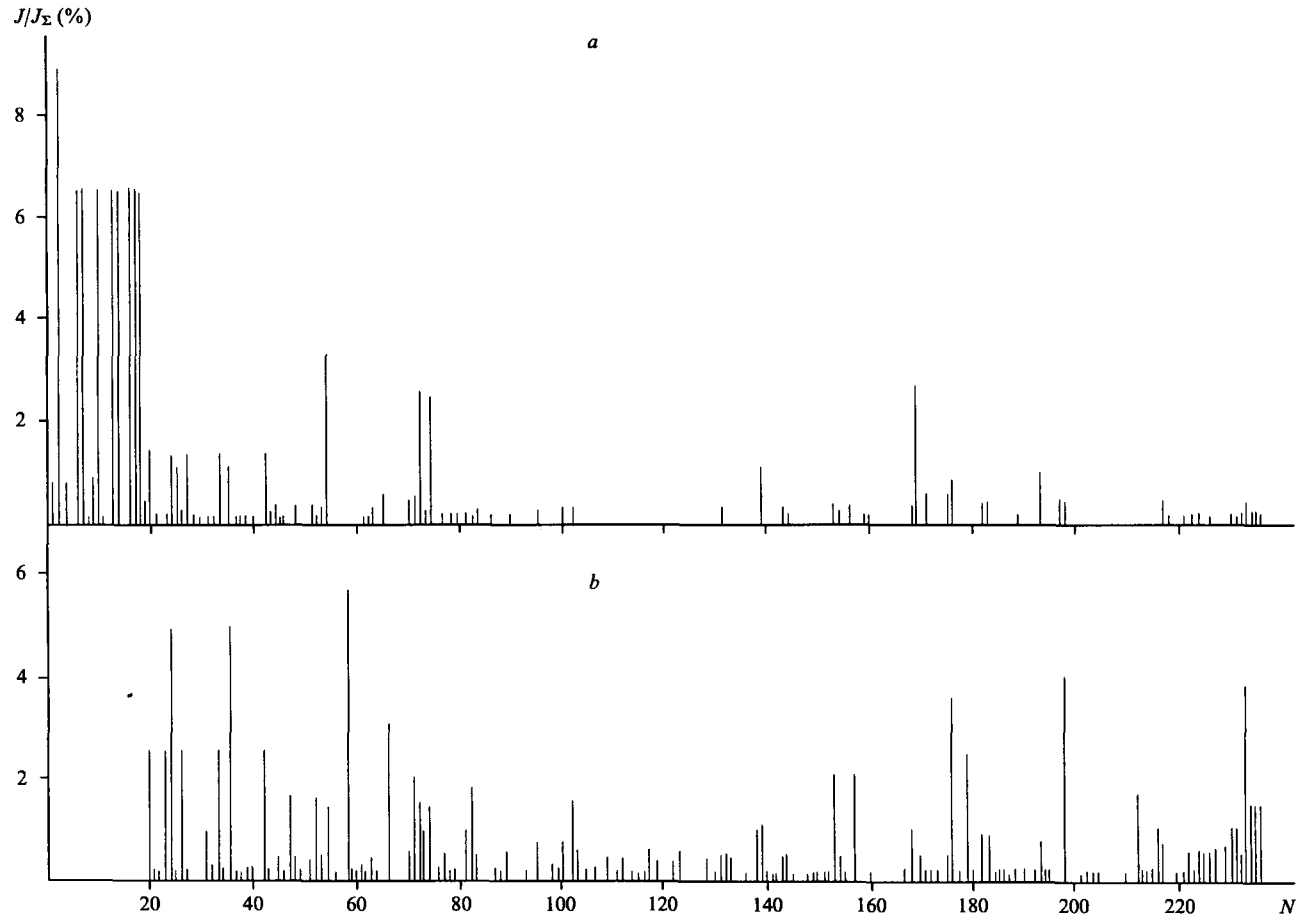


Figure 16. Experimental accessibility profile of the polypeptide chain in the X potato virus (XPV) in the native form (*a*) and partially degraded by proteases of the cell sap (*b*). J/J_Σ is the portion of the activity attributable to the respective aminoacid residue; N is the number of the aminoacid residue in the polypeptide chain.

from the distribution observed for TMV (see Fig. 14): the sharp boundaries between accessible and inaccessible regions have disappeared. This suggests that XPV has a much 'rougher' surface than TMV.[124,125] We note that 85% of the total amount of label incorporated by the preparation is located in the first 19 amino-acids in the intact virus: this shows that the entire N-end region of the protein is exposed on the virus surface. The absence of the first 19 end aminoacids (Fig. 16*b*) in the coat protein of the virus preparation degraded in situ produces a qualitative change in the labelling process. Evidently, the C-end region of the protein, which apparently had been previously 'covered' by the N-end part of the protein in the coat, has become accessible to the tritium atoms.

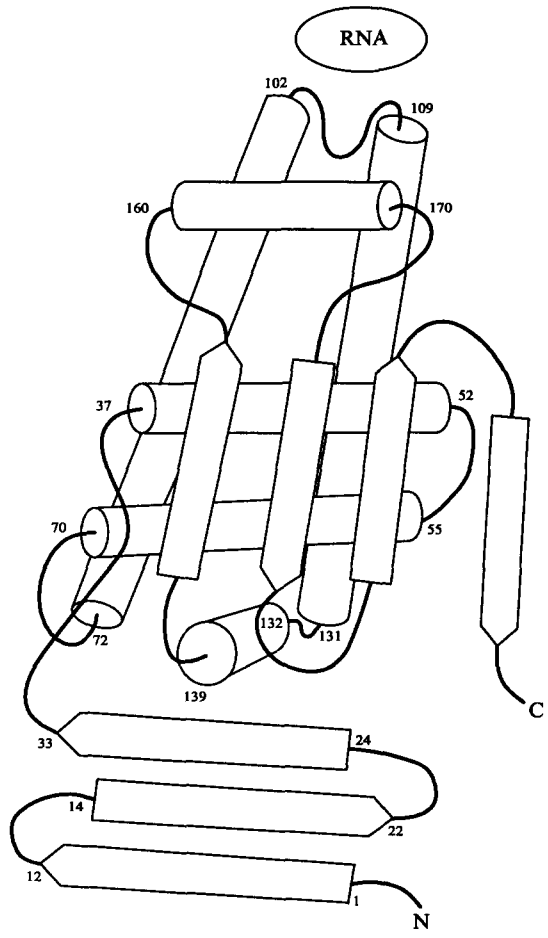


Figure 17. Model of the three-dimensional packing of the polypeptide chain in the protein of the XPV shell. The α-helices are shown as cylinders; the β structural regions are indicated by arrows. The numbers denote the boundaries of the structured regions.

On the basis of the result obtained by tritium planigraphy, of existing biochemical and immunological data, and also of the basic principles governing the packing of the predicted elements of the secondary structure in globular proteins Baratova et al.[125] proposed a model of the tertiary structure of the protein in the coat of XPV (Fig. 17). In this model one β group, containing the β structural regions 1–12, 14–22, and 24–33, lies alongside the molecule and forms the outer surface of the virus particles.

To conclude this review we shall briefly mention another group of problems which could be solved effectively by the application of tritium planigraphy. The concept of linear or conformational antigenic determinant proteins is widely adopted in modern immunology and immunochemistry. It is assumed that linear determinants are composed of several aminoacid residues, amounting to a substantial proportion of the residues in the protein sequence, whereas the conformational determinants contain aminoacid residues from different parts of the sequence, lying close together in the native structure of the protein globule.

However, recent work suggests that all the antigenic determinants are conformational.[127,128] Thus, the surface of the protein obtained from X-ray structural analysis data is always complex and has a very broken outline. Even a visual analysis of any structure shows that most of the residues on the protein surface have neighbours which are quite removed in the primary structure of the protein. Furthermore the most 'unbroken' parts are mostly loops and/or projecting regions. This explains why surface properties such as hydrophilicity, accessibility, and mobility can be used to predict possible antigenic determinants. This concept is well illustrated by the correlation between the accessibility to tritium atoms and the antigenic properties of these parts of the structure, observed in the TMV and XPV viruses and (earlier) in globular proteins. By describing the structure of the object in terms of data on its surface (provided by an analysis of the tritium labelling), and also by a comparative analysis of the surfaces of the free protein and of the protein complexes with a monoclonal antibody, it should be possible, we believe, to identify the conformational determinants. So far this possibility has been realised only at the level of the X-ray structural analysis of structures such as complexes with site-directed mutagenesis.[129]

References

1. G M Clore, M A Robien, A M Gronenborn *J. Mol. Biol.* **231** 82 (1993)
2. R Henderson, J M Baldwin, T M Ceska, F Zemlin, E Beckmann, K H Downing *J. Mol. Biol.* **213** 899 (1990)
3. D N Wang, W J Kuhlbrandt *J. Mol. Biol.* **217** 691 (1991)
4. A Holzenburg, F H Wilson, M E Finlow, R C Ford *Biochem. Soc. Trans.* **20** 591 (1992)
5. H Arakawa, K Umemura, A Ikai *Nature (London)* **358** 171 (1992)
6. S Colacicchi, M Ferrari, A Sotgin *Int. J. Biochem.* **24** 205 (1992)
7. J A Thompson, A L Lau, D D Cunningham *Biochemistry* **26** 743 (1987)
8. B H Robertson, V K Brown, B P Holloway, B Khanna, E Chan *Arch. Virol.* **104** 117 (1989)
9. P R Schimmel, H J P Schoemaker *Methods Enzymol.* **59** 332 (1979)
10. L A Neiman, V S Smolyakov, A V Shishkov *Obshchie Problemy Fiziko-khimicheskoi Biologii, T. 2 (Itogi Nauki i Tekhniki)* [Common Problems of Physicochemical Biology, Vol. 2 (Advances in Science and Engineering)] (Moscow: Izd. VINITI, 1985)
11. M Menzinger, R Wolfgang *J. Chem. Phys.* **50** 2991 (1969)
12. M Menzinger, R Wolfgang *J. Am. Chem. Soc.* **89** 5992 (1967)
13. V N Kondrat'ev *Konstanty Skorosti Gazofaznykh Reaktsii. Spravochnik* (Rate Constants of Gas-phase Reactions. Handbook) (Moscow: Nauka, 1970)
14. I Langmuir *J. Am. Chem. Soc.* **34** 1310 (1912)
15. R Klein, M D Sheer *J. Phys. Chem.* **62** 1011 (1958)
16. A N Ponomarev, V L Tal'roze *Dokl. Akad. Nauk SSSR* **130** 120 (1960)
17. L A Neiman, V S Smolyakov, L P Antropova, I A Gedrovich, I A Kashirin, A V Shishkov, in *I Vsesoyuznyi Biofizicheskii S'ezd, Moskva, 1982 (Tezisy Dokladov, Stendovye Soobshcheniya 1690-2318)* [The First All-Union Biophysical Congress, Moscow, 1982 (Abstracts of Reports, Posters 1690-2318)] p. 214
18. H C Moser, P Nordin, J K Senne *Int. J. Appl. Radiat. Isot.* **15** 557 (1964)
19. P Nordin, H C Moser, J K Senne *Biochem. J.* **96** 336 (1965)
20. A V Shishkov, E S Filatov, E F Simonov, M S Unukovich, V I Gol'danskii, A N Nesmeyanov *Dokl. Akad. Nauk SSSR* **228** 1237 (1976)
21. A K E Hagopian, R H A Johnsen *J. Chem. Soc., Chem. Commun.* 837 (1974)
22. A V Volynskaya, A Yu Skripkin, A V Shishkov, V I Gol'danskii *Dokl. Akad. Nauk SSSR* **266** 871 (1982)
23. A V Gedrovich, G A Badun *Mol. Biol. (Moscow)* **26** 558 (1992)
24. M M Yusupov, A S Spirin *Methods Enzymol.* **164** 426 (1988)

25. L A Baratova, V I Gol'danskii, Yu M Rumyantsev, M S Unukovich, A V Shishkov *Mol. Biol. (Moscow)* **16** 117 (1982)
26. A F Trotman-Dickenson, G S Milne *Natl. Stand. Res.-Natl. Bur. Stand.* **9** 41 (1967)
27. V I Gol'danskii, Yu M Rumyantsev, A V Shishkov, L A Baratova, L P Belyanova *Mol. Biol. (Moscow)* **16** 528 (1982)
28. S E V Philips *Nature (London)* **273** 247 (1978)
29. B K Lee, F M Richards *J. Mol. Biol.* **55** 379 (1971)
30. S A Islam, D L Weaver *Proteins: Structure, Function and Genetics* **10** 300 (1991)
31. E Stellwagen, M Wilgus *Nature (London)* **275** 342 (1978)
32. S H Bryant, S A Islam, D L Weaver *Proteins: Structure, Function and Genetics* **6** 418 (1989)
33. R E Gates *J. Mol. Biol.* **127** 345 (1979)
34. S Miller, J Janin, A M Lesk, C Chothia *J. Mol. Biol.* **196** 641 (1987)
35. J Janin, S Miller, C Chothia *J. Mol. Biol.* **204** 155 (1988)
36. C Chothia *Nature (London)* **254** 304 (1975)
37. D J Barlow, J M Thornton *Mol. Graphics* **4** 97 (1986)
38. D J Barlow, J M Thornton *Biopolymers* **25** 1717 (1986)
39. D W Fanning, J A Smith, G D Rose *Biopolymers* **25** 863 (1986)
40. D M Novotny, M Handschumacher, E Haber, R E Brucoletti, W B Carlson, D W Fanning, J A Smith, G D Rose *Proc. Natl. Acad. Sci. USA* **83** 226 (1986)
41. Yu Chirgadze, N Kurochkina, S Nikonov *Protein Eng.* **3** 105 (1989)
42. G Neweths, H A Scheraga *Q. Rev. Biophys.* **10** 239 (1977)
43. M Levit, C Chothia *Nature (London)* **261** 552 (1976)
44. C Chothia *Ann. Rev. Biochem.* **53** 537 (1984)
45. A V Volynskaya, A Yu Skripkin, A V Shishkov, E S Dzhafarov, Yu M Rumyantsev, V I Gol'danskii *Mol. Biol. (Moscow)* **19** 1294 (1985)
46. A B Ivlev, A V Gedrovich, A V Shishkov *Mol. Biol. (Moscow)* **26** 1047 (1992)
47. E N Bogacheva, A A Shishkov, A V Shishkov, L A Baratova *Mol. Biol.* (Moscow) (1994) (in the press)
48. D C Carter, He Hiao-min, S H Munson, F D Twigg, K M Gernett, M B Broom, T Y Miller *Science* **244** 1195 (1989)
49. A V Volynskaya, E A Kasumov, E N Bogacheva, A V Shishkov, V I Goldanskii *Eur. J. Biophys.* 1994 (in the press)
50. T Peters *Adv. Clin. Chem.* **13** 37 (1970)
51. E Keshavarz, S Nakai *Biochim. Biophys. Acta* **576** 269 (1979)
52. A V Volynskaya, V N Izmailova, V A Pchelin *Biofizika* **18** 210 (1973)
53. K E Bobkova, A V Gedrovich, V N Ankilova, O I Lavrik, L A Baratova, A V Shishkov *Biokhimiya* **55** 1570 (1990)
54. E V Bobkova, A V Gedrovich, V N Ankilova, O I Lavrik, L F Baratova, A V Shishkov *Biochem. Int.* **20** 1001 (1990)
55. A V Gedrovich, L A Baratova, E N Bogacheva, V N Medvedkin, A V Shishkov *Mol. Biol. (Moscow)* **27** 309 (1993)
56. A V Volynskaya, S A Murashova, A Yu Skripkin, A V Shishkov, V I Gol'danskii *Mol. Biol. (Moscow)* **23** 356 (1989)
57. A Yu Skripkin, A V Volynskaya, A V Shishkov, V I Gol'danskii *Mol. Biol. (Moscow)* **23** 365 (1989)
58. K Hamaguchi, K Rokkaku *J. Biochem.* **48** 358 (1960)
59. H Edelhoch, R F Steiner *Biochim. Biophys. Acta* **60** 359 (1962)
60. E A Kasumov, A V Volynskaya, A V Shishkov, V I Goldanskii *Stud. Biophys.* **136** 167 (1990)
61. Ch N Cornell, L J Kaplan *Biochemistry* **17** 1750; 1755 (1978)
62. J Welting, M M Weideman, A C-J Roomer, J H Perrin *Biochim. Biophys. Acta* **579** 469 (1979)
63. E S Dzhafarov *Mol. Biol. (Moscow)* **25** 1412 (1991)
64. E S Dzhafarov, L A Aliev, A Yu Skripkin, A V Volynskaya, A V Shishkov, L A Baratova, V I Gol'danskii *Izv. Akad. Nauk Azerb. SSR* (3) 102 (1985)
65. S F Sun *Arch. Biochem. Biophys.* **129** 411 (1969)
66. V Luzzati, J Witz, A Nicholaieff *J. Mol. Biol.* **3** 379 (1961)
67. M Ioly, in *A Physico-chemical Approach to the Denaturation of Proteins* (Translated into Russian; Moscow: Mir, 1968) p. 38
68. V A Kashpur, V Ya Maleev, T Yu Shchegoleva *Mol. Biol. (Moscow)* **10** 568 (1976)
69. J White, M Kielian, A Helenius *Q. Rev. Biophys.* **16** 151 (1983)
70. D Hoekstra, J W Kok *Biosci. Rep.* **9** 273 (1989)
71. S E Glushakova, A L Ksenofontov, N V Fedorova, L A Mazhul, O N Ageeva, L B Margolis, L A Baratova, A V Shishkov *Biosci. Rep.* **11** 131 (1991)
72. A L Ksenofontov, A L Danilov, S E Glushakova, L B Margolis, L A Baratova, A V Shishkov *Biochem. Int.* (1994) (in the press)
73. M Karplus, J A McCammon *CRC Crit. Rev. Biochem.* **9** 293 (1981)
74. M Levit *J. Mol. Biol.* **168** 595 (1983)
75. M Levit *J. Mol. Biol.* **168** 621 (1983)
76. E N Frolov, O V Belonogova, G I Likhteinshtein, in *Ravnovesnaya Dinamika Nativnoi Struktury Belka* (Equilibrium Dynamics of Native Protein Structure) (Pushchino: Sci. Centre Biol. Inv., 1977) p. 99
77. V I Goldanskii, Yu F Krupyanskii *Hyperfine Interact.* **39** 341 (1988)
78. E N Bogacheva, N D Zhukov, A V Shishkov *Mol. Biol. (Moscow)* **27** 1044 (1993)
79. A V Shishkov, L A Neiman, V S Smolyakov *Usp. Khim.* **53** 1125 (1984) [*Russ. Chem. Rev.* **53** 656 (1984)]
80. I V Kurinov, Yu F Krupyanskii, I P Suzdalev, V I Goldanskii *Hyperfine Interact.* **33** 223 (1987)
81. C Chothia *J. Mol. Biol.* **105** 1 (1976)
82. M E Snow *Proteins; Structure, Function and Genetics* **15** 183 (1993)
83. F E Cohen, S R Presnell, L M Gregoret, I D Kuntz, in *Protein Folding: Deciphering the Second Half of the Genetic Code* (Eds L Gierasch, J King) (Washington, DC/AAAS, 1990) p. 251
84. R Unger, J Moult *J. Mol. Biol.* **231** 75 (1993)
85. S Saitoh, T Nakai, K Nishikawa *Proteins: Structure, Function and Genetics* **15** 191 (1993)
86. A V Shishkov, Yu M Rumyantsev, M S Unukovich, V I Gol'danskii, E F Simonov, E S Filatov, L A Baratova, L P Belyanova, R K Ledneva, in *IV Vsesoyuznyi Simpozium po Khimii Belkov i Peptidov, Minsk, 1977 (Tezisy Dokladov)* (Abstracts of Reports at the IVth All-Union Symposium on the Chemistry of Proteins and Peptides, Minsk, 1977) p. 23
87. G J Stubbs, S G Warren, K S Holmes *Nature (London)* **267** 216 (1976)
88. A V Gedrovich, A V Shishkov, V I Goldanskii, L A Baratova, N I Grebenshchikov, A V Efimov *Eur. Biophys.* **19** 283 (1991)
89. P V Chow, G D Fasman *Annu. Rev. Biochem.* **47** 251 (1978)
90. V I Lim *J. Mol. Biol.* **88** 872 (1974)
91. O B Ptitsyn, A A Rashin *Biophys. Chem.* **3** 1 (1975)
92. R H Tufty, R H Kretsinger *Science* **187** 167 (1975)
93. P C Moews, R H Kretsinger *J. Mol. Biol.* **91** 210 (1975)
94. A V Gedrovich, M M Yusupov, A V Shishkov, V I Gol'danskii, A S Spirin *Dokl. Akad. Nauk SSSR* **267** 1255 (1982)
95. M M Yusupov, A S Spirin *Biokhimiya* **51** 1858 (1986)
96. M M Yusupov, A S Spirin *FEBS Lett.* **197** 229 (1986)
97. I A Kashirin, A V Gedrovich, A V Shishkov, V K Kagramanova, L A Baratova *Bioorg. Khim.* **9** 1531 (1983)
98. V A Kolb, A A Kommer, A S Spirin *Dokl. Akad. Nauk SSSR* **296** 1497 (1987)
99. V A Kolb, A A Kommer, A S Spirin *Mechanism and Control of Translation* (Netherlands: Noordwijkerhout, 1990)
100. N D Belyaev, V G Budker, V A Dubrovskaya, A A Kim, E V Kiseleva, V N Sidorov *FEBS Lett.* **297** 43 (1992)
101. I A Filatov, M A Kulish, A F Mironov, E V Nozhevnikova, L A Neiman *Bioorg. Khim.* **13** 623 (1987)
102. N G Kotlova, M A Kulish, A F Mironov *Bioorg. Khim.* **14** 1695 (1988)
103. B Kadenbach, A Stroh *FEBS Lett.* **173** 374 (1984)
104. L V Gordeeva, L A Baratova, L B Margolis, A V Shishkov *Biofizika* **34** 971 (1989)
105. A B Kuryatov, V V Shemyakin, L P Antropova, in *XVI Konferentsiya FEBS, Moskva, 1984 (Tezisy Dokladov)* (Abstracts of Reports at the XVIth FEBS Conference, Moscow, 1984) p. 335
106. A B Kuryatov, A S Arsen'ev, T N Alenycheva, L A Neiman, V I Tsetlin, in *Mezhdunarodnyi Simposium "Retinal'soderzhashchie Belki", Moskva, 1986 (Tezisy Dokladov)* (Abstracts of Reports at the International Symposium "Retinene-containing Proteins", Moscow, 1986) p. 76
107. Yu A Ovchinnikov, N G Abdulaev, R G Vasilov, I Yu Vturina, A B Kuryatov, A V Kiselev *FEBS Lett.* **179** 343 (1985)
108. J K Mohana Rao, P Argos *Biochim. Biophys. Acta.* **869** 197 (1986)
109. Yu A Ovchinnikov, N G Abdulaev, M Yu Feigina, A V Kiselev, N A Lobanov *FEBS Lett.* **100** 219 (1979)
110. D M Engelman, R Henderson, A D Mclachlan, B A Wallace *Proc. Natl. Acad. Sci. USA* **77** 2023 (1980)

111. T N Alenycheva, A B Kuryatov, L P Antropova, V V Shemyakin, L A Neiman, V I Tsetlin, V T Ivanov *Bioorg. Khim.* **13** 898 (1987)
112. R Henderson *J. Mol. Biol.* **93** 123 (1975)
113. D N Ostrovskii, V A Eremin, I G Zhukova, A A Masyagin, A V Shishkov *Mikrobiologiya* **57** 875 (1988)
114. K Namba, G Stubbs *Science* **231** 1401 (1986)
115. G Stubbs, S Warren, K Holmes *Nature (London)* **267** 216 (1977)
116. D Altschuh, D Hartman, J Reinbolt, M H V Van Regenmortel *Mol. Immunol.* **20** 271 (1983)
117. I A Kashirin, N I Grebenshchikov, L A Baratova, A V Shishkov, V I Gol'danskii *Mol. Biol. (Moscow)* **20** 1264 (1986)
118. I V Kashirin, N I Grebenshchikov, L A Baratova, A V Shishkov, V I Gol'danskii *Mol. Biol. (Moscow)* **21** 456 (1986)
119. V I Goldanskii, I A Kashirin, A V Shishkov, L A Baratova, N I Grebenshchikov *J. Mol. Biol.* **201** 567 (1988)
120. L A Neiman, L P Antropova, M A Zalesskaya, E I Budovskii *Bioorg. Khim.* **12** 1070 (1986)
121. J Sober, L V Jarvekulg, I E Toots, Yu L Radavsky, R Villems, M Saarma *J. Gen. Virol.* **69** 1799 (1988)
122. L Sawyer, P Tollin, H R Wilson *J. Gen. Virol.* **68** 1229 (1987)
123. D D Shukla, P M Strike, S L Tracy, K H Gough, C M Ward *J. Gen. Virol.* **69** 1497 (1988)
124. L A Baratova, N I Grebenshchikov, A V Shishkov, I A Kashirin, Yu L Radavsky, L Jarvekulg, M Saarma *J. Gen. Virol.* **73** 229 (1992)
125. L A Baratova, N I Grebenshchikov, E N Dobrov, A V Gedrovich, I A Kashirin, A V Shishkov, A V Efimov, L Jarvekulg, Yu L Radavsky, M Saarma *Virology* **188** 175 (1992)
126. M I Gol'shtein, N I Grebenshchikov, S V Kust, A S Kaftanova, E N Dobrov, I G Atabekov *Mol. Gen., Mikrobiol. Virusol.* (2) 9 (1990)
127. M H V Regenmortel *Hybridonia Technology in Agriculture Veterinary Research* (Eds N J Stern, H R Gramble) (New Jersey: Rowman and Allanheld, 1984) Ch. 3
128. D J Barlow, M S Edwards, J M Thornton *Nature (London)* **322** 747 (1986)
129. J M Nuss, P B Whitaker, G M Air *Proteins: Structure, Function and Genetics* **15** 121 (1993)