AN ESTIMATE OF INTRAPROTEIC ELECTROSTATIC FIELDS VALUES ORIGINATED BY THE PEPTIDE GROUPS IN α-CHYMOTRYPSIN.

Georges Johannin and Nicolas Kellershohn
Laboratoire d'Enzymologie physico-chimique et moleculaire
Université de Paris-Sud,Centre d'Orsay-91400,Orsay,France.

Received August 1,1972; accepted August 28,1972

SUMMARY: The effective concentration of peptide dipolar groups within the protein globular structure gives the basis of a new multipolar model,as a more adequate description of the protein core.The contribution of the peptide backbone to the intraproteic electrostatic field is estimated in different parts of the α-Chymotrypsin macromolecule,by applying the Coulomb's law.This contribution is large enough to be considered as a non-negligible part of the origin of both chemical and physical properties of a globular protein.

INTRODUCTION

Since many years,the well-known "oil-drop" model has become a sort of corner-stone in the field of protein physical-chemistry.First assumed from thermodynamic considerations dealing with the hydrophobic character of side-chains of neutral amino-acids,(1),this model later appeared as fairly substantiated by X-Ray diffraction studies,which have shown the inner and outer situation as being respectively prefered by the apolar and polar side-chains of amino-acids (2).

The purpose of the present note is 1/:to present some arguments leading to the conclusion that the protein interior is not analogous to a non-polar medium and may be more adequately described as a multipolar,low-dielectric constant medium.2/:to report an estimate of intraproteic electrostatic fields or potential values.3/:to briefly analyse the main consequences that can be expected from such a model.

A classical method which has been used to give informations on the polarity of the intraproteic medium consists in a comparison of some physical property of a probe,either bound at the protein core,or free in a series of solvents of different polarities.For instance,using spectrophotometric and fluorimetric measurements,this approach has been used to investigate the nature of the environment constituted by the active center of chymotrypsin and/or trypsin.The probes which have been tested in these cases were aromatic structures which can be divided into two classes:one class of compounds which are non-covalently bound at the active site:competitive inhibitors such as proflavin (3) or thionin (4) or chromophoric substrates (5),and a second class

of compounds which are covalently bound,for example via an ester bond,to the reactive serine:cinnamoyl (6),indoleacryloyl (7) and anthranyloyl (8) groups.

Our attention has been focussed on an apparent contradiction which arises from a comparison of the successive reports quoted above.Some of them (3,7) lead to the idea that the active site of α-chymotrypsin is similar to a highly non-polar solvent,the other (6,8) leading to a comparison of the same active site to a highly polar solvent.In fact,when we go back to the level of facts,it has to be noticed that behind this divergence of interpretation,the absorption spectra of all these aromatic structures are characterized bya common and rough physical effect:namely,a red-shift.

One of us (5) has recently shown that the extent of these red-shifts can, in fact,be reproduced or approximated with polar solvents as well as,or better than using apolar solvents.Among the series of solvents which have been tested in this work,N-Methylacetamide (N.M.A.) and N-Methylpropionamide (N.M.P.) which are characterized by a strong dipole moment attached to the CONH group (9) are suitable to mimic the polarity of both chymotryptic and tryptic active sites. Moreover,it has been shown in the same work that this ability to reproduce the protein environment is not limited to the active sites of these two enzymes since N.M.A. and N.M.P. are also able to induce in N-Acetyl-tryptophan-ethylester a shift to the red which is approximately similar to the one observed when tryptophan is included in a native protein structure (10).

We then made the hypothesis that at least part of the red-shifts which-as witnessed by an excessive amount of litterature to be reviewed here- are currently exhibited by any aromatic structure embedded in the protein macromolecule,whether associated to it or constitutive part of it,could be relevant to an electrostatic effect due to the proximity of strongly dipolar peptide groups.

This hypothesis was first encouraged by a crude calculation based on the overall dimensions of several proteins:lysozyme (11),chymotrypsin (12),and carboxypeptidase A (13);these calculations show,in agreement with Klapper's results (14),that the average volume per amino-acid,and therefore,the average domain which can be considered as being under the electrostatic influence of one CONH dipole is of the same order of magnitude as the molecular volume of liquid polar solvents such as N.M.A. and N.M.P..This only point seems to us of a nature to justify a substitution of the oil-drop model by a multipolar model,assuming that the protein interior cannot be considered as electrostatically neutral at the submacromolecular level.Then,it becomes necessary to know if the values of the electrostatic field which is set up by the dipolar peptide groups are large enough to play a part in certain physical and chemical properties of proteins.

PRINCIPLES AND METHODS OF CALCULATIONS

The calculations presented here were intended to give an estimate of intraproteic electrostatic fields (I.P.F.) and potentials (I.P.P.).They were carried out by assuming the following essential features:

-*Point charge approximation:* The center of each atom of every peptide group is supposed to bear a fractional elementary charge,whose value is chosen so as to reproduce the dipole moment of the entire group.As suggested by Ramachandran (15), these values were taken as equal to +0.4,-0.4,+0.3,-0.3,for the atoms C,O,N,H, respectively.In addition,all these fractional charges are supposed embedded in an isotropic medium,the dielectric constant of which lies between 1 and 3.

-*Protein cage concept:* We propose this term to designate any domain of the protein core,which is surrounded by,and under the electrostatic influence of neighbouring peptide groups:this domain may be,for example,the one which encloses the indole ring of a tryptophan unit,or even the whole active center of an enzyme.The important difference between a protein cage and a solvent cage mainly bears on the necessarily multipolar and quasi-rigid character of the former,with respect to the latter.

These assumptions,together with proteins X-rays diffraction data,allow one to calculate the approximate I.P.F. and I.P.P. values at different points located in different strategic protein cages,by applying the Coulomb's law.Given the importance of strong electrostatic fields in the absorption,emission,and dichroism properties of aromatic compounds (16),we have first calculated the I.P.F. values at different points of the indole rings of seven over the eight tryptophan units contained in α-chymotrypsin.The peptide groups which are immediately adjacent to the indole ring were only taken into account in this part of our calculations.Then,expecting a possible role of I.P.F. in the catalytic properties of this protein,we have calculated the I.P.P. values at several points of the charge relay system,presented by Blow et all (17) as being responsible of the enhanced nucleophilicity of the serine 195 γ-hydroxyle;in this case,all the peptide groups of the α-chymotrypsin molecule were taken into account.

Our calculations have made large use of the atomic coordinates and the related nomenclature for tosyl-α-chymotrypsin,as they heve been published by Birktoft et all (18).The coordinates of the γ-oxygen of Ser 195 are those given by Henderson (19) as being related to the native form of the enzyme.The H atom coordinates of every peptide group were calculated from the coordinates of the three other atoms pertaining to the same peptide group,assuming a perfect trans configuration,and standard dimensions,as given by Corey et Pauling (20).

Concerning the I.P.F. values all over the tryptophan indole ring system, one direction which is of particular interest,since it represents approximately the direction of the ground state dipole moment of this system,is defined by the

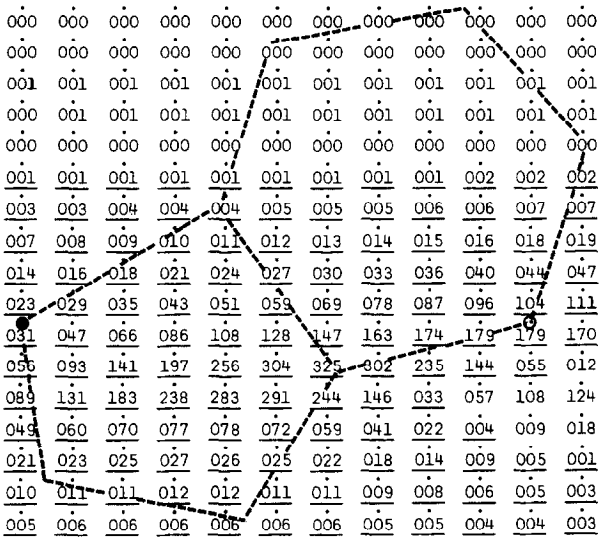

Fig. 1: Values, in mV/Angström, of the electrostatic field component along the direction defined by the couple of atoms NE1 (●) and CE3 (○) over the indole ring plane of tryptophan 172, in α-chymotrypsin. (Contribution from the two peptide groups pertaining to this amino-acid. The positive values are underlined).

atoms NE1 and CE3 of the indole ring (21). The component of the field vector along this direction was calculated at 21 x 13 points over the indole ring surface, through a summation of the projections of the vector field components relative to the tridimensional cartesian reference system defined by Birktoft et all.

All calculations were programmed for a Wang calculating system fitted with a 373 data storage unit and a 377 teletype control unit.

RESULTS AND DISCUSSION:

Fig 1, relative to tryptophan 172 is intended to give an idea of the order of magnitude and of the variations of the I.P.F. values which are originated by the peptide groups which belong to the same amino-acid. Assuming a dielectric constant equal to 1, the magnitude of the I.P.F. values lies between the extreme values $\pm$0.1 to $\pm$1.0 Volt/Angström. A strong variation of the field along distances as short as one angström unit with a reversal of sign can be noticed at some places. Although it is possible, according to Liptay's theory, (16), to predict the directions of the shifts induced on the absorption band corresponding to one given transition, from the knowledge of the relative orientations of excited and ground state dipole moments with respect to the external field, the information produced by our calculations is still not sufficient to start such an analysis,

at least for two reasons:on the one hand,the two adjacent peptide groups are not, by far,the only two to be determinant in originating the electric field amplitude on the indole ring.For tryptophan 172,for example,a calculation of interatomic distances show that the oxygen atom of the 168th,215th,216th,224th,225th,and 226th peptide groups are located at respectively 5.7,6.2,5.7,5.9,3.3,and 6.1 Angströms from the nitrogen atom NE1 of tryptophan 172,whereas the corresponding distances for the adjacent peptide groups are 5.8 and 6.5 Å.The other tryptophan similarly have several peptide groups in their vicinity.On the other hand,the plane on which the field component has been calculated,actually represents the nodal plane of the indole ring where the electron density is zero.Further calculations,taking into account the actual number of near peptide groups and a more realistic shape of the indole ring,are in progress in our laboratory.

Inasmuch as the I.P.F. might play a role in the catalytic properties of proteins,the charge relay system in α-chymotrypsin appears as a particularly interesting example to test this possibility.This system involves three amino-acid residues and extends over about eight angströms.The I.P.P. values were calculated at eight points along the whole system;the results are presented on Fig 2,in the form of a potential difference with respect to the oxygen atom OD2 of Asp 102,versus distance .This distance axis,going from OD2 (Asp 102) to the oxygen OG (Ser 195) through the two nitrogen atoms ND1 and NE2 (His 57) may be visualized as a reaction coordinate (22),for the net exchange of one proton between the Asp 102 carboxylate and the Ser 195 hydroxyle group.

Our results show that the charge relay system lies on a potential valley; each oxygen atom stands at the top of one of the sides of the valley,whereas the nitrogen ND2 of His 57 stands at the bottom of the valley.One of the key features of the charge relay system mechanism (23),consists in a release of one proton from the NE2 nitrogen atom to a water molecule,which passes through an ion-pair-wise canonical form where NE2 and OD2 respectively bear a positive and a negative charge.At pH 4,the pH where the crystallographic data were obtained,it can be noticed that this ion-pair formation is favoured by an electrostatic energy which is estimated to lie between 0.2 and 0.6 eV. according to which extreme value of the dielectric constant is taken into account.Such an electrostatic contribution from the peptide backbone can be of some help to understand how a carboxyle group,even though it is buried within a medium of low dielectric constant, can loose a proton relatively easily.

One point which might be baffling at first sight,concerns the low value of dielectric constant used throughout our calculations.In fact,this view is consistant with the fact that the dielectric constant of a medium,whether polar or not, as long as it is rigid,depends only upon polarizability (24).We would like to add to this argument,the empirical fact that the dielectric constant of a dry

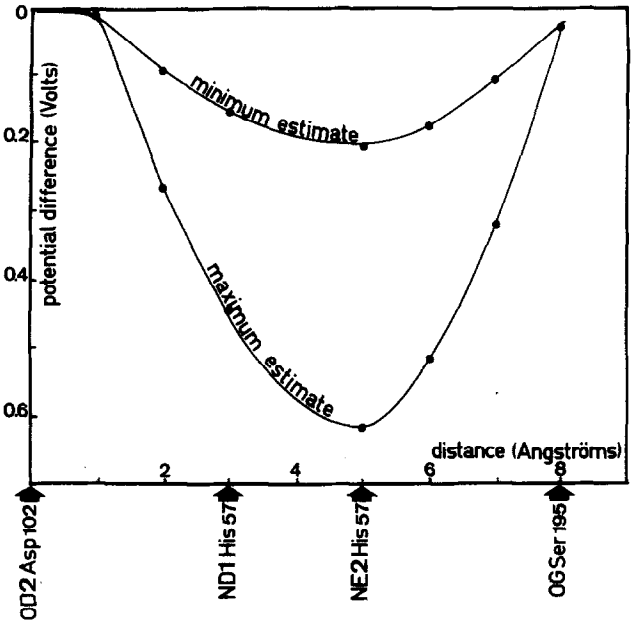

Fig 2: Potential valley along the charge relay system, originated by the peptide backbone of α-chymotrypsin. The potential of the atom OD2 of Asp 102 is taken as a reference. The minimum and maximum estimates respectively correspond to a dielectric constant equal to 3 and 1.

sample of hemoglobin was reported as equal to 3 (25), and that the refraction index of quite a number of proteins lies around 1.6 (26). It should also be kept in mind that as much as one accepts that an ionized or protonated group is strong enough to polarize a bond in a given direction, one should also accept that three or four dipolar CONH groups at the right place can result in the same effect.

CONCLUSION:

Up to now, the physical chemistry of globular proteins has been almost exclusively interpreted on the basis of the oil-drop model. Our feeling is that the new multipolar model might lead to new insights in this field, and we would like to summarize our working hypothesis in the three following statements:

1/: Being compact and highly concentrated in strongly dipolar peptide groups, the intraproteic medium must be considered as a multipolar medium.

2/: The multipolar distribution, within a low dielectric constant medium, gives rise to *high electrostatic fields* (about 10 to 100 MVolts/cm.) which can result in a perturbing effect, *at any point of the protein core, on every molecule or group of atoms which are either part of the protein structure, or associated to it.*

3/: This perturbing effect can be manifested and looked for under both aspects of perturbed *physical and chemical properties.*

The two last statements give a precise physical meaning to the idea according to which the peptide backbone, all around the active center of an enzyme, should be seen, not only as a catalytic and binding group holder, but also as a particular cage which possesses, by nature, as any other protein cage, the property to disturb its contents.

AKNOWLEDGMENTS: The work reported here was supported by the C.N.R.S.. We are indebted to Professor J.Yon and Dr. D.Thusius for a critical evaluation of this manuscript.

REFERENCES:

1. Kauzmann,W.,Adv. Protein Chem., 14,1,(1959)
2. Perutz,M.F., European J. Biochem.,8, 455 (1969)
3. Bernhard,S.A.,Lee,B.F.,Tashjian,Z.H., J. Mol. Biol.,18,405 (1966)
4. Glazer,A.N., J. Biol. Chem.,242,3326 (1967)
5. Johannin,G.,These de doctorat,Orsay,1972.To be published.
6. Bender,M.L.,Schonbaum,G.R.,Zerner,B.,J.Am.Chem.Soc.,84,2540 (1962)
7. Bernhard,S.A.,Lau,S.J.,Noller,H.,Biochemistry,4,1108 (1965)
8. Haugland,R.P.,Stryer,L.,in Conformation of Biopolymers,Vol I (Ramachandran,edr) p. 321 (1967)-Academic Press.
9. Riddick,J.A.,Bunger,W.B.,Techniques in Chemistry,Vol II (1970) Wiley-Interscience.
10. Wetlaufer,D.B.,Adv. Protein Chem.,17,304 (1962)
11. Blake,C.C.F.,Mair,G.A.,North,A.C.T.,Philips,D.C.,Sarma,V.R., Proc. Roy. Soc. Lond. B 167,365 (1967)
12. Blow,D.M.,in The Enzymes,3rd edition,Vol III (Boyer Edr.) (1970) Ac.Press.
13. Lipscomb,W.N.,Reeke,G.N.,Hartsuck,J.A.,Quiocho,F.A.,Bethge,P.H., Phil. Trans. Roy. Soc. Lond. B 257,177 (1970)
14. Klapper,M.A., Biochim. Biophys. Acta,229,557 (1971)
15. Ramachandran,G.N.,Sasisekharan,V.,Adv. Protein Chem. 23,284 (1968)
16. Liptay,W., Angew. Chem.,Internat. edit.,8,177 (1969)
17. Blow,D.M.,Birktoft,J.J.,Hartley,B.S.,Nature,221,337 (1969)
18. Birktoft,J.J.,Matthews,B.W.,Blow,D.M.,Biochem.Biophys.Res.Comm.,36,131 (1968)
19. Henderson,R.,J.Mol. Biol.,54,341 (1970)
20. Corey,R.B.,Pauling,L.,Proc. Roy. Soc. Lond. B 141,10 (1953)
21. Yearger,E.,Biophys. J.,8,1505 (1968)
22. Glasstone,S.,Laidler,K.J.,Eyring,H.,The theory of rate processes (1941)-Mc Graw Hill.
23. Birktoft,J.J.,Blow,D.M.,Henderson,R.,Steitz,T.A.,Phil. Trans. Roy. Soc., B 257,67 (1970)
24. Davies,M.,Some electrical and optical aspects of molecular behaviour (1965)-Pergamon Press.
25. Maricic,S.,Pifat,G.,Pravdic,V.,Biochim. Biophys. Acta,79,293 (1964)
26. Doty,P.,Geidusheck,E.P.,in The Proteins (Vol I,part A),Neurath & Bailey Edrs. p. 393 (1953),Academic Press.