This is to certify that the

thesis entitled

PART I: CRYSTAL STRUCTURE DETERMINATION OF A FLUORESCENT PROBE, 1-ANILINONAPHTHALENE-8-SULFONATE, ITS COMPLEX WITH α-CHYMOTRYPSIN AND THE NATURE OF pH DEPENDENCE ON THE COMPLEX. PART II: PROTEIN DERIVATIVE PHASE REFINEMENT BY DIFFERENCE ELECTRON DENSITY MODIFICATION

presented by

Lawrence D. Weber

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Chemistry

E. T. Adman

Major professor

Date 10/31/78.

PART I

CRYSTAL STRUCTURE DETERMINATION OF A FLUORESCENT PROBE, 1-ANILINONAPHTHALENE-8-SULFONATE, ITS COMPLEX WITH α-CHYMOTRYPSIN AND THE NATURE OF pH DEPENDENCE OF THE COMPLEX

PART II

PROTEIN DERIVATIVE PHASE REFINEMENT BY DIFFERENCE ELECTRON DENSITY MODIFICATION

By

Lawrence D. Weber

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1978

ABSTRACT

PART I

CRYSTAL STRUCTURE DETERMINATION OF A FLUORESCENT PROBE, 1-ANILINONAPHTHALENE-8-SULFONATE, ITS COMPLEX WITH α-CHYMOTRYPSIN AND THE NATURE OF pH DEPENDENCE OF THE COMPLEX

PART II

PROTEIN DERIVATIVE PHASE REFINEMENT BY DIFFERENCE ELECTRON DENSITY MODIFICATION

By

Lawrence D. Weber

The structure of the fluorescent probe 1-anilino-naphthalene-8-sulfonate (ANS) was determined by three dimensional X-ray crystallographic techniques. The molecule crystallized in space group $P2_1/c$ with four molecules in the unit cell and cell dimensions of a = 6.150Å, b = 9.544Å, c = 26.77Å, and β = 90.35°.

The structure was solved by use of the program MULTAN. The unit cell contained one ammonium counter ion and one water molecule per ANS. The carbon-nitrogen bonds in the latter were found to be shortened significantly from expected single-bonded values. Inspection of independent crystallographic structures of ANS (1,2) revealed that such shortening persists regardless of anilino geometry

or conformation. The naphthyl moiety of ANS demonstrated the geometric distortions typical of peri-substituted naphthalenes.

The 2.8Å resolution X-ray crystallographic structure of the complex between ANS and α-chymotrypsin (α-CHT) was studied at pH 3.6 and 6.6 with the difference Fourier method. At pH 3.6, the ANS binds on the surface of the enzyme close to the amino terminus of the A-chain and interacts intimately with the disulfide bond of Cys 1-122. The ANS is exposed to polar residues and possibly some ordered water molecules. Thus, the fluorescence enhancement at low pH (3) does not reflect a hydrophobic site but rather a polar one. It has been demonstrated that the ANS binding site in the crystal is the same as that in solution (3). The ANS and its protein environment remain practically the same at pH 6.6. The structural changes which occur attendent to α-CHT pH 5.4 conformer formation (4) are essentially the same as those which take place in the ANS-α-CHT complex at pH 6.6. Substituent-protein contacts and protein structural changes were summarized by use of the difference diagonal plot representation (DDP). A "multiplicity" characterization of the DDP contacts was developed to supplement this technique. The pH fluorescence dependence observed in solution (3) may be due to a proton-disulfide interaction which prevents the disulfide group from quenching the ANS fluorescence at low pH.

Another likely mechanism is an increase in the degree of mobility of water molecules as the pH is increased. The results suggest the need for caution in the interpretation of spectral characteristics of fluorescent probes that are interacting with biomacromolecules.

The difference Fourier method used in protein crystallography relies on the phases of the native protein to approximate those of a protein-derivative structure. A difference electron density modification (DDM) procedure was developed to refine the phases of a derivatized protein from the native starting set. The DDM procedure is an iterative sequence of alternating difference density modification and Fourier inversion calculations. Greater computational efficiency was achieved by use of a Cooley-Tukey Fast Fourier Transform (5). The criteria for effective difference density modifications were delineated. Primary among these is that, in contrast to electron density modification (6,7), the modifying function must take into account that the majority of the difference density is frequently below background.

REFERENCES

1. Cody, V. and Hazel, J., J. Med. Chem., 20, 12 (1977).
2. Cody, V. and Hazel, J., Acta Cryst. B33, 3180 (1977).
3. Johnson, J. D., Ph.D. Thesis, Michigan State University, 1976.

4. Vandlen, R. L. and Tulinsky, A., Biochem., 12, 4193 (1973).

5. Ten Eyck, L. F., Acta Cryst., A29, 183 (1973).

6. Collins, D. M., Brice, M. D., LaCour, T. F. M. and Legg, M. J., "Crystallographic Computing", Ahmed, F. R., Huml, K., Sledlacek, B., ed., p. 330, Munksgaard, Copenhagen (1976).

7. Raghavan, N. V., Tulinsky, A., Acta Cryst. (1978) in press.

ACKNOWLEDGMENTS

To Dr. Alexander Tulinsky, for his guidance, support and encouragement throughout this study, the author expresses his sincere gratitude.

The author would like to thank Dr. N. V. Raghavan, and Dr. S. R. Ernst for their many helpful contributions to this study. For their many stimulating discussions, their collaboration in the fluorescent probe work, and for imbuing the author with a portion of their enthusiasm for fluorescence methods, the author wishes to extend his appreciation to Dr. M. A. El-Bayoumi and Dr. J. D. Johnson

To Dr. D. L. Ward, the author is thankful for the use of a data collection facility, many computer programs, and useful discussions. For many helpful discussions the author is indebted to Dr. B. A. Averill, to Dr. M. N. Liebman, especially for his suggestion of the difference diagonal plot, and to Dr. I. M. Mavridis, Dr. A. Mavridis, Dr. L. S. Hibbard and Mr. M. Frentrup. Thanks are extended to Dr. D. J. Duchamp for the use of a computer graphics facility. The author expresses his gratitude to Ms. C. Britton for her technical and general assistance.

Support of the National Science Foundation and the National Institutes of Health is gratefully acknowledged.

To the memory of Samuel S. Abrams

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

LIST OF FIGURES

PART I

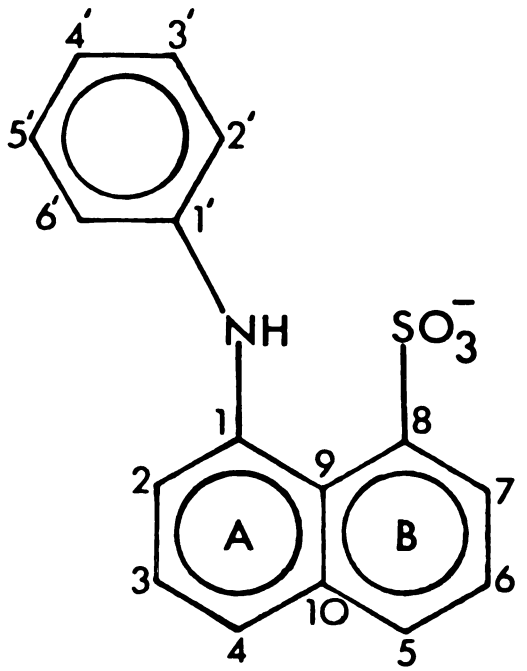
CRYSTAL STRUCTURE DETERMINATION OF A FLUORESCENT PROBE, 1-ANILINONAPHTHALENE-8-SULFONATE, ITS COMPLEX WITH α-CHYMOTRYPSIN AND THE NATURE OF pH DEPENDENCE OF THE COMPLEX

I. INTRODUCTION

A. Fluorescent Probes

The use of fluorescent probes of macromolecular structure has been motivated in part by the precision with which spectroscopic quantities can be measured. Fluorescence methods can also offer a distinct advantage in terms of time and cost, compared with other structure-probing techniques such as electron and nuclear magnetic resonance, and x-ray crystallography.

Although virtually non-fluorescent in water, naphthylanilines and related dyes undergo a dramatic fluorescence enhancement upon addition of bovine serum albumin or heat-denatured proteins, as was first reported by Weber and Laurence in 1954 (1). The emission enhancement of these dyes in non-polar solvents was also noted. Stryer subsequently found that 1-anilinonaphthalene-8-sulfonate (ANS) (I) (Figure I) adsorbed specifically to the non-polar heme-binding site of apomyoglobin and apohemoglobin, with a binding constant on the order of 10^5 M^{-1} (2). A two-hundred fold increase in the quantum yield of ANS fluorescence accompanied the binding, along with a 60 nm shift to higher energy (from green to blue) of the emission maximum. Noting the striking correlation between these results and the spectroscopic properties of such dyes in solvent systems of varying polarity, Stryer suggested that they could be



I

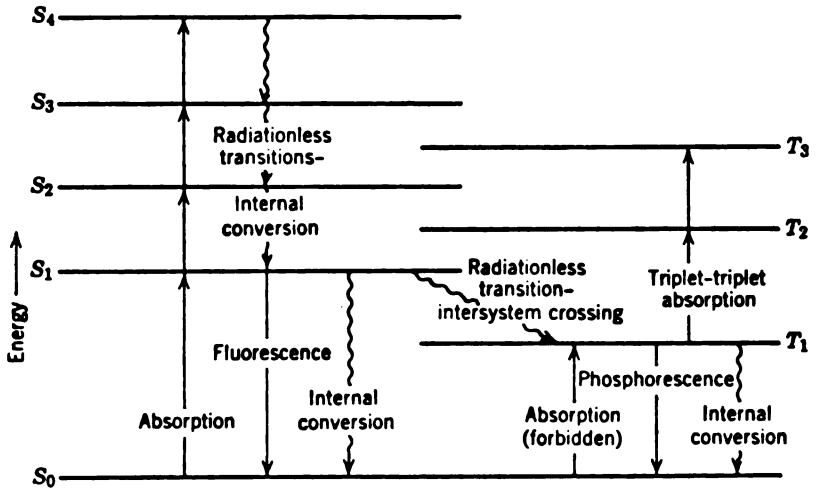
Figure 1. Skeletal formula of 1-anilinonaphthalene-8-sulfonate, with numbering scheme.

used as probes of non-polar sites of proteins. Thus, an inferential model was conceived that initiated use and elucidation of the properties of probe molecules (3-12).

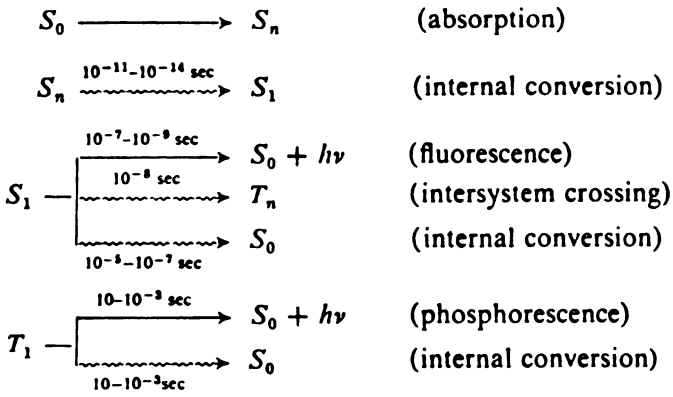
B. Review of Excited State Processes

The electronic energy levels of a hypothetical molecule are represented schematically in Figure 2a. Straight arrows indicate electronic state transitions in which radiation is absorbed or emitted, and wavy arrows indicate non-radiative transitions. Figure 2a illustrates the processes of absorption, fluorescence, phosphorescence, and deactivation. These processes, together with approximate time scales, are summarized in Figure 2b. The times indicate the high probability that environmental fluorescence sensitivity will be determined by the interactions which involve states S_1 and T_1, since these are by far the longest lived excited states. An electron spin selection rule $\Delta S = 0$, taken into account in Figure 2, where S is the electron spin angular momentum quantum number, applies rigorously to electronic transitions in conjugated molecular structures that contain atoms of a mass lighter than or comparable to bromine (13).

An excited singlet state which can decay only through an emissive process will show first order fluorescence decay behavior:



A



B

Figure 2. (a) Schematic electronic state energy level diagram: S is singlet and T is triplet. The S_0 state is the ground state and the subscript numbers identify individual states.

(b) Summary of processes in a, including approximate time ranges.

(Figure 2a,b used by permission of publisher (13).)

$$I = I_o e^{-t/\tau_F^o} \qquad (1)$$

where I is the fluorescence emission intensity at time t, I_o is the intensity at the onset of decay (at t = 0), and τ_F^o is the intrinsic or "natural" lifetime of the excited state. The natural lifetime is thus the time required for I to fall to 1/e times its value at t = 0, and we can define the fluorescence decay constant k_F as

$$k_F = \frac{1}{\tau_F^o} . \qquad (2)$$

In the event that other deactivation processes are operative, the observed mean lifetime, τ_F, of the excited state is

$$\tau_F = \left(\frac{k_F}{k_F+k_{IS}+k_{IC}}\right)\tau_F^o \qquad (3)$$

where k_{IS} is the rate constant for intersystem crossing and k_{IC} that for internal conversion, both of which are defined in a manner analogous to k_F.

Another useful measure of emission efficiency is the quantum yield, Φ_F, which can be shown to relate to the emission lifetime. The quantum yield of fluorescence is defined as

$$\Phi_F = \frac{\text{number of fluorescence quanta emitted}}{\text{number of quanta absorbed to a singlet excited state}}, \qquad (4)$$

and is equal to the intrinsic fluorescence quantum yield, Φ_F^o, when external quenching processes (e.g., deactivation by collision) are absent. For excited singlet state emission,

$$\Phi_F^o = \frac{k_F}{k_F+k_{IC}+k_{IS}} , \qquad (5)$$

so that

$$\tau_F = \Phi_F^o \tau_F^o \quad . \qquad (6)$$

Since Φ_F^o takes all internal deactivation processes into account, a longer observed mean fluorescence lifetime necessarily implies more effective competition by the fluorescence process as a deactivation pathway.

C. Solvent Polarity and its Effect on the Fluorescence Properties of N-Arylaminonaphthalene Sulfonates

Investigation of the spectral properties of fluorescent probe molecules has been aided by establishment of spectroscopically based empirical measures of solvent polarity. Prominent among these are the Z (14, 15) and $E_T(30)$ (16) parameters.

The value of the Z parameter of a solvent, expressed in kcal/mole, is the energy of the charge transfer band (the lowest energy absorption) of the solute 1-ethyl-4-

carbomethoxypyridinium iodide (II). The transition is represented in Figure 3, and is essentially of the dipole annihilation type. This description pertains due to the relative orientation of the molecular dipole moment, which is perpendicular to the 6-membered ring (between atoms of the ion pair) in the ground state, but parallel to the ring (and smaller) in the excited state. The equilibrium configuration between solvent molecules and the ground state of the solute defines the Franck-Condon (or "cybotactic") environment of the excited state. In this case it is an environment possessing a net dipole moment perpendicular to that of the excited solute. Such a circumstance leads to the absence of a net dipole interaction (in the ideal limit of point charges) and a destabilization of the excited solute relative to its energy in an environment of unorganized solvent (14).

The energy of II in the ground state will also vary significantly, depending on the characteristics of its solvent shell. Thus the higher a solvent is in polarity, the more effectively it can solvate the ion pair, and this will produce an absorption band at higher energy.

The $E_T(30)$ parameter is measured in the same manner, but utilizing the transition energies of pyridinium-phenolbetaine (derivative No. 30 (17)) and its methylated form (betaine), the latter for solvents of the lowest polarity. The relationship between Z and $E_T(30)$ is

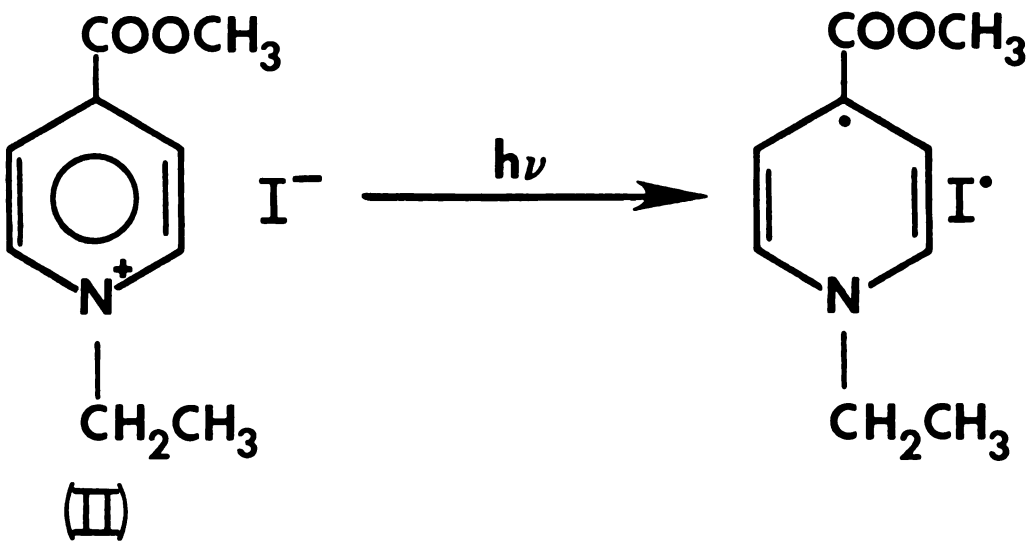


Figure 3. Lowest energy electronic absorption transition in 1-ethyl-4-carbomethoxypyridinium iodide.

fairly linear over the range of solvent polarities between that of benzene and water (15). Values for Z and $E_T(30)$ are given in Table I for some selected solvents.

An analogous solvent-solute interaction sequence for initial and final electronic states can be operative for the fluorescence process if we impose the constraint that the lifetime of the excited state is longer than that needed for equilibration of surrounding solvent molecules (solvent "relaxation").

A graphical representation of the behavior of ANS emission maximum in response to solvent polarity is presented in Figure 4a, the latter as measured by the Z parameter. An identical plot for 1-anilinonaphthalene-5-sulfonate (1,5-ANS) is given in Figure 4b. The degree of qualitative agreement between these plots is typical for the isomers of ANS. The data used in plotting Figure 4 are listed in Table II. It is clear that a red shift of the ANS emission maximum is observed in more polar solvents, and that this is accompanied by a dramatic quenching of fluorescence intensity (small Φ_F).

A theoretical formulation, attributed to Lippert (18) and Mataga, et al. (19,20) describes the electrostatic effects of permanent and induced dipole-dipole interaction between solvent and solute, under the constraint of "rapid" solvent relaxation mentioned earlier. These authors have successfully employed it to relate the energies of the

Table I. Z^{*} and $E_T(30)^{**}$ Values for Some Selected Solvents.

Solvent	z(kcal/mole)	$E_T(30)$(kcal)mole)
Water	94.6	63.1
Methanol	83.6	55.5
Ethanol	79.6	51.9
1-propanol	78.3	50.7
1-butanol	77.7	50.2
t-butanol	71.3	43.9
Pyridine	64.0	40.2
1,4-Dioxane	----	36.0
Benzene	54.	34.5

*Reference 15.

**Reference 16.

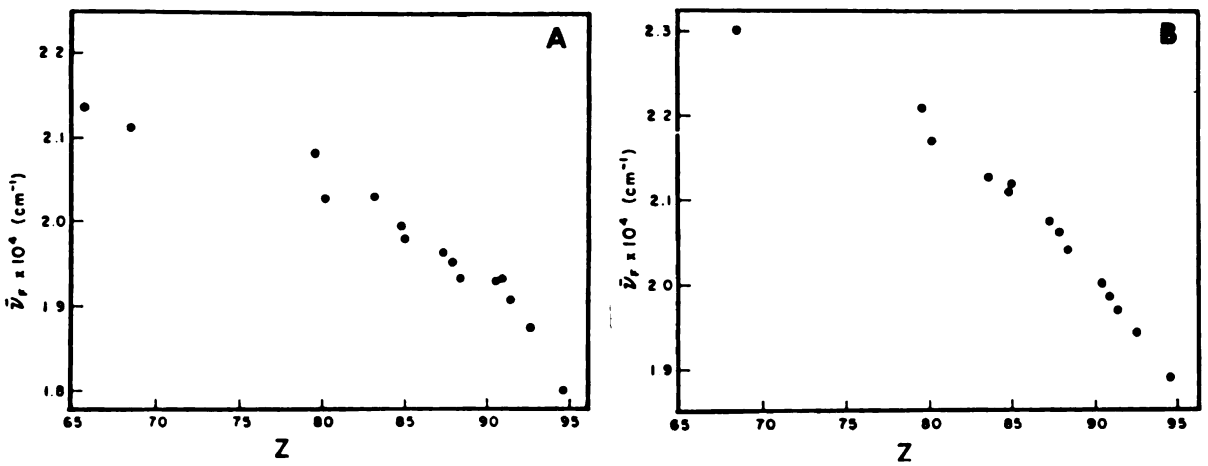


Figure 4. Plots of the transition energy of fluorescence as a function of the empirical solvent polarity scale, Z, for ANS(a) and 1,5-ANS (b) (9). the Data for ANS are taken from Table II.

Table II. Fluorescence of ANS in Various Solvents.

Solvent (%)	Quantum Yield (ϕ)	ANS $\bar{\nu}_F$ ($\times 10^4$ cm^{-1})
Water	0.0032	1.802
Ethanol(20)	0.0072	1.876
Ethanol(40)	0.020	1.931
Ethanol(60)	0.050	1.953
Ethanol(80)	0.11	1.996
Ethanol	0.40	2.083
Dioxane(20)	0.012	1.908
Dioxane(40)	0.038	1.934
Dioxane(60)	0.099	1.980
Dioxane(80)	0.23	2.028
Dioxane	0.57	2.118
Methanol(50)	0.029	1.934
Methanol(80)	0.077	1.964
Methanol	0.17	2.032
Acetone	0.39	2.137
Dimethylformamide	0.39	2.114
Ethylene glycol	0.12	1.968

excited (pre-emission) state of fluorescent molecules with that of the ground state, and thereby to explain fluorescence properties. This theory has also been successfully applied to ANS isomers and derivatives (11,21). For ANS-like molecules the Lippert-Mataga expression is

$$\bar{\nu}_A - \bar{\nu}_F \approx \frac{2}{hc_0 a^3}\left(\frac{\varepsilon-1}{2\varepsilon+1} - \frac{n^2-1}{2n^2+1}\right)(\vec{\mu}_e - \vec{\mu}_g)^2 + \text{constant} \quad (7)$$

where the constant, a sum of additional terms arising from different kinds of induced dipole interaction is, at maximum, only 10^{-3} as large as the lead term (20). In Equation (7), $\bar{\nu}_A$ and $\bar{\nu}_F$ are the wavenumber values (cm^{-1}) for the lowest energy singlet absorption maximum and fluorescence maximum respectively, so that their difference represents the non-radiative energy loss of the emitted photon. The explicitly written lead term, which accounts for the orientation effect between the permanent dipole of the solute and those of solvent molecule, contains h, Plancks constant; c_0, the speed of light in a vacuum; a, the Onsager cavity radius of the solute (22); ε, the dielectric constant of the solvent; n, the refractive index of the solvent; $\vec{\mu}_e$, the dipole moment of first excited singlet state of the solute; and $\vec{\mu}_g$, the dipole moment of the solute ground state.

A plot of $\bar{\nu}_A - \bar{\nu}_F$ *vs* $\left(\frac{\varepsilon-1}{2\varepsilon+1} - \frac{n^2-1}{2n^2+1}\right)$ can be used to obtain a value for the apparent dipole change accompanying the

emission process. Such graphical analysis for 2,6-ANS and its derivatives, performed with spectral data measured in ethanol-water mixtures, yielded plots of high linearity (11). A value $(\mu_e-\mu_g)$ = 40D was reported for 2,6-ANS, a higher apparent dipole moment in the excited state.

A simple explanation for sensitivity of emission maximum to conditions of solvent polarity can now be presented. The constraint requiring relatively short solvent relaxation times is amply satisfied in many organic solvents. Dielectric relaxation times measured for primary alcohols, attributed to hydrogen bond breaking followed by ROH rotation, were found to range from 0.4 ns in n-propanol to 2.0 ns in n-decanol at 20°C (23). In contrast, the observed mean lifetime for ANS in n-propanol is 10.2 ns (24).

Thus, the emitting fluorophore is in transition from a solvent equilibrated condition to a Franck Condon or cybotactic region. The excited state is more polar than the ground state, and undergoes a relative stabilization in polar solvents. The relative destabilization of the ground state, which is experiencing the Franck Condon environment, is also greater in more polar solvents and the result is a red shift of fluorescence maximum. ANS-like molecules will fluoresce with anomalously high energy and fluorescence intensity in polar media such as glycerol (11) or ice (25), where solvent relaxation is retarded.

The spectroscopic response of 2-anilinonaphthalene-6-sulfonate (2,6-ANS) to solvent polarity and that of two of its derivatives, as measured by the $E_T(30)$ parameter, is depicted in Figure 5 (26). The behavior of 2,6-ANS is readily seen to manifest distinct high and low polarity sensitivity regions (26,27). The onset of high sensitivity is displaced to lower polarity with a methyl substituent (Me-2,6-ANS*), an electron donor. With methoxy, a stronger effective donor (due to conjugation effects), only a high sensitivity is observed.

If $\bar{\nu}_F$ in Figure 5 is expressed in kcal/mole the high sensitivity slope is near unity at approximately 0.7. The $E_T(30)$ parameter is based on an intramolecular dipole annihilation transition and, in consideration of the substituent effect, intramolecular charge transfer is postulated to give rise to the highly polar (11) excited singlet (labeled S_{1ct} in Figure 5). The fact that the fluorescence of 2-aminonaphthalene-6-sulfonate is nearly solvent insensitive (3) underscores the importance of the aryl substituent. It has been suggested (27) that a relatively planar molecular configuration must be achieved to facilitate electronic communication between the phenyl and naphthyl ring systems. Also, due to the insensitivity of the absorption spectrum to the aryl substituent when

---

*Also referred to in the literature as 2-p-toluidinylnaphthalene-6-sulfonate (TNS).

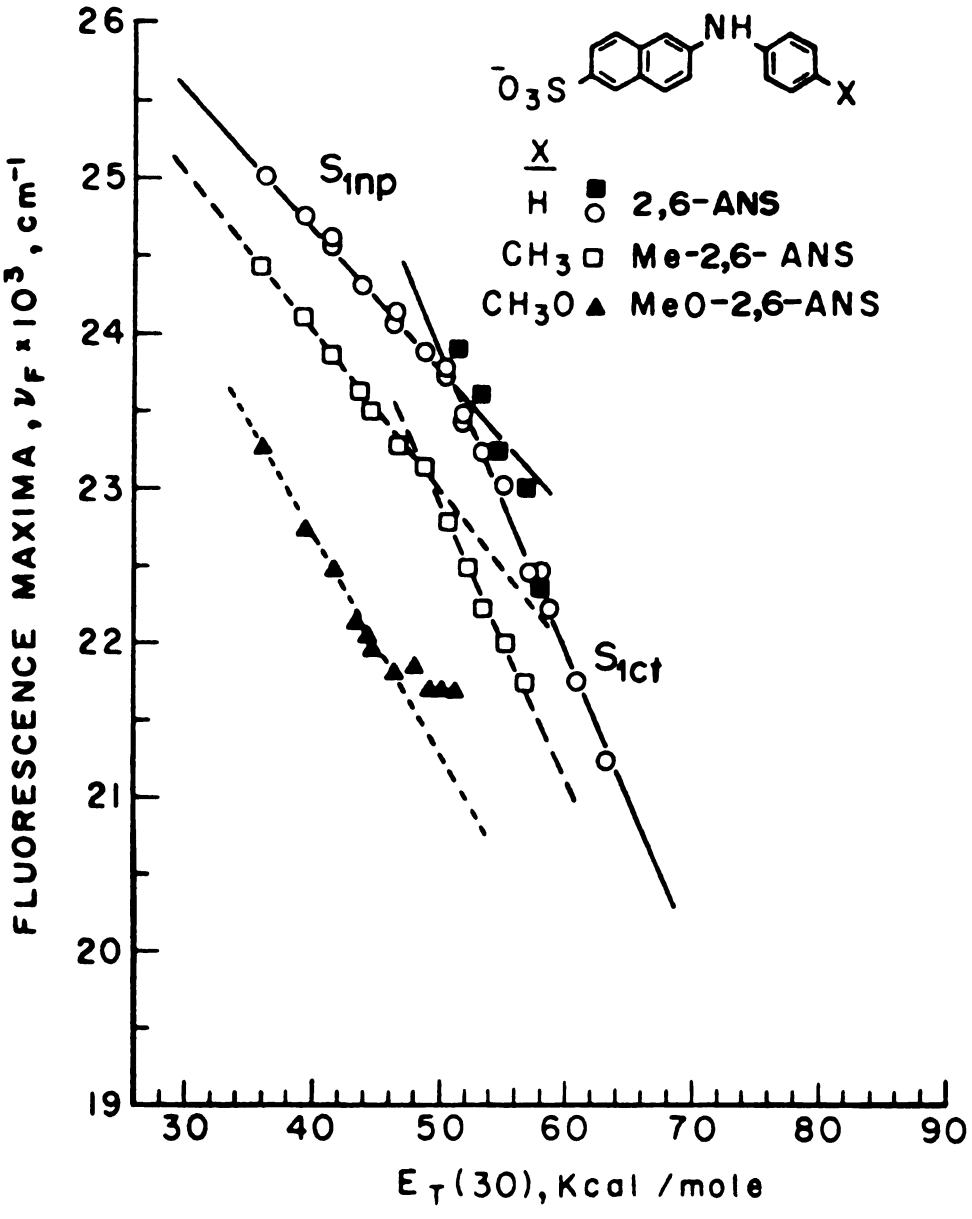


Figure 5. Plot of fluorescence emission maxima (in cm^{-1}) versus the solvent polarity standard $E_T(30)$, for 2,6-ANS and two derivatives labeled in the Figure (25). The notations S_{1np} and S_{1ct} are explained in text. All measurements were performed in dioxane-water mixtures except for those indicated by filled rectangles for which the solvent mixture was ethanol-water. These latter data were not included in the correlation lines.

measured in ethanol: 1) the ground state configuration is relatively non-planar (sterically favored to begin with), and 2) the Franck-Condon excited state is non-planar (labeled S_{1np} in Figure 5) as well. It is from the latter state, S_{1np}, that emission in the low sensitivity (low polarity) solvent region is assigned, in this two state - non-planar singlet, charge transfer singlet - emission model.

These conclusions follow from a line of reasoning similar to that of Seliskar and Brand (11), who also suggested the involvement of an intramolecular charge transfer process, although they proposed that the Franck-Condon 1st excited state was itself of a charge transfer nature. They did not observe a two-fold polarity response as their measurements were carried out in ethanol-water mixtures, which possess a minimum $E_T(30)$ value of 51.9 (Table I).

However, complete characterization of the emitting states of N-arylaminonaphthalene sulfonates has not yet been achieved. Indeed, wavelength dependent fluorescence lifetime studies (28) have suggested that the two-emitting state model may be inadequate, although no specific additional states have been proposed. Interestingly, H' NMR studies (29) of ANS that were conducted in D_2O and deuteromethanol have revealed a selective relative deshielding in the latter solvent, affecting the chemical

shifts of H_2, $H_{2'}$, and $H_{6'}$ (Figure 1). This deshielding has been interpreted as the result of a greater degree of intramolecular hydrogen bonding in the less polar solvent (a possibility unique to the 1,8 isomer), and indicative of a more planar molecular conformation. Although results of this study contrast with the report of a planar charge-transfer excited state in media of higher polarity (Figure 5), they apply in a direct way only to the ground state.

Two hypotheses have been advanced to explain the dramatic quenching observed for the fluorescence of ANS-like molecules in polar solvents. Brand (11,30) is the chief proponent of the widely accepted intersystem crossing hypothesis. This hypothesis is based on the relationship between the dipole moments of states S_0, S_1 and T_1. It has been shown that in aqueous solution the pk values of T_1 in 2-aminonaphthalene and N,N-dimethyl-2-aminonaphthalene lie between those corresponding to S_0 and S_1 (31), this implying an intermediate dipole moment value. Thus, by a mechanism of solvent relaxation, the energy difference between the S_1 and T_1 states will diminish in polar solvents, facilitating a greater efficiency for intersystem crossing (32) and subsequent non-radiative decay.

A second hypothesis has been advocated by Koscwer and coworkers (27). Pulsed laser excitation and chemical techniques were used to identify the several absorbing species born

of the excitation of 2,6-ANS derivatives in water-dioxane mixtures. In the high polarity range (as in Figure 5), the absorbance attributed to T_1 (510 mm) was found to decrease in a fashion similar to that of S_1, this implying that the former is not the intersystem crossing product of the latter. On the basis of additional studies of fluorescence yield and emission maximum in solvents of varying polarity and viscosity, it was postulated that in the high polarity ranges a charge transfer reaction $S_{1ct} \rightarrow S_0$ takes place, at a much lower rate than $S_{1np} \rightarrow S_{1ct}$, but which is favored in higher polarity-lower viscosity media, this accounting for the observed quenching.

D. Fluorescence of ANS in the Presence of α-CHT

Fluorescence properties of ANS dissolved in water are given in Table IL In addition to the quenching effect of water, stronger quenching is observed for 1-anilinonaphthalene-7-sulfonate (1,7-ANS) when protonation occurs at pH 1.5. Otherwise Φ_F and $\bar{\nu}_F$ are invarient between pH 1.5 and pH 12.0 (9). Quenching is also observed in methanolic ANS solutions that have been acidified below pH 3.0 with dry HCl gas (29).

When bovine α-chymotrypsin (α-CHT) is added to aqueous solutions of ANS at pH 3.6, a dramatic enhancement of ANS fluorescence intensity and a blue shift in $\bar{\nu}_F$ results. At

this pH, an emission maximum of 2.05 x 10^4 cm^{-1} is observed (24) which, referring to Figure 4a, yields an apparent Z value (15) suggestive of a mean environment similar to that of 1-butanol. Further, a significant sensitivity is observed as the ANS, α-CHT solution is adjusted toward neutral pH values. This behavior is summarized in Figures 6 and 7 (24,33).

The binding of salicylate derivatives to liver alcohol dehydrogenase has been investigated X-ray crystallographically (34). Accompanying fluorescence studies demonstrated that ANS competes with salicylate derivatives for the same binding site, located in the hydrophobic adenosine-binding pocket. However, no direct crystallographic evidence for ANS binding was presented.

The goal of this study was to shed more light on the binding of ANS to proteins, the subsequent fluorescence enhancement, and the mechanism by which fluorescence variability is produced in response to pH-induced or other conformational changes. The molecular environment of ANS has been mapped in crystals of bovine α-chymotrypsin (α-CHT) at 2.8 Å resolution, at two pH values. Conditions of pH were chosen such that the fluorescence properties of ANS would vary dramatically. Also taken into account was the fact that pH 4.0 and 7.0 correspond to discrete values (among several, higher and lower) at which conformational changes occur in crystalline α-CHT (35). Thus, derivative

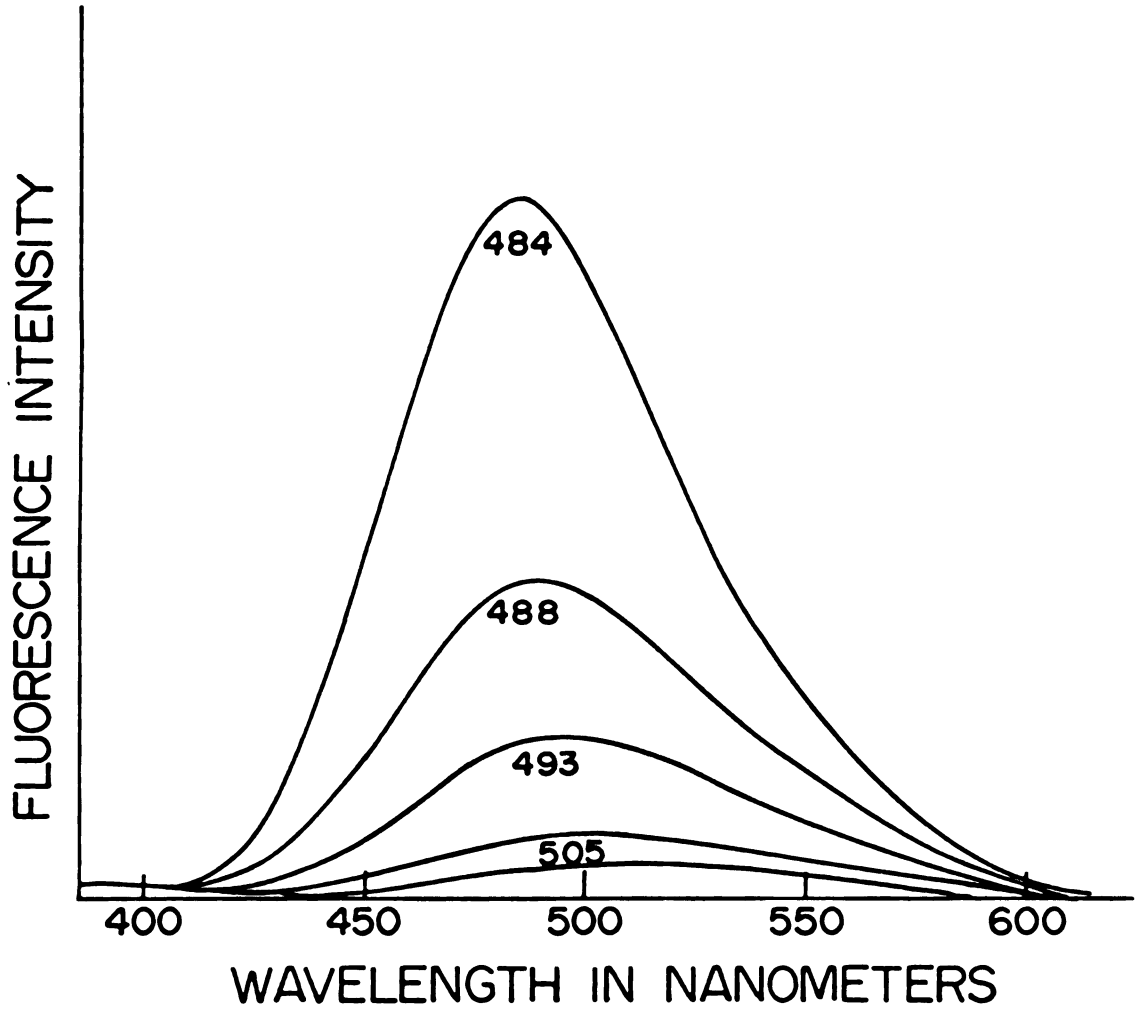


Figure 6. Fluorescence spectra of 10^{-4} M α-CHT in the presence of 2×10^{-5} M ANS (24) at various pH values: (2.4, 3.6, 4.75, 7.0 and 8.0). Successive red shifts of the emission maxima occur as the pH value is increased. Since the spectra were recorded at various sensitivities, relative intensities are not reported.

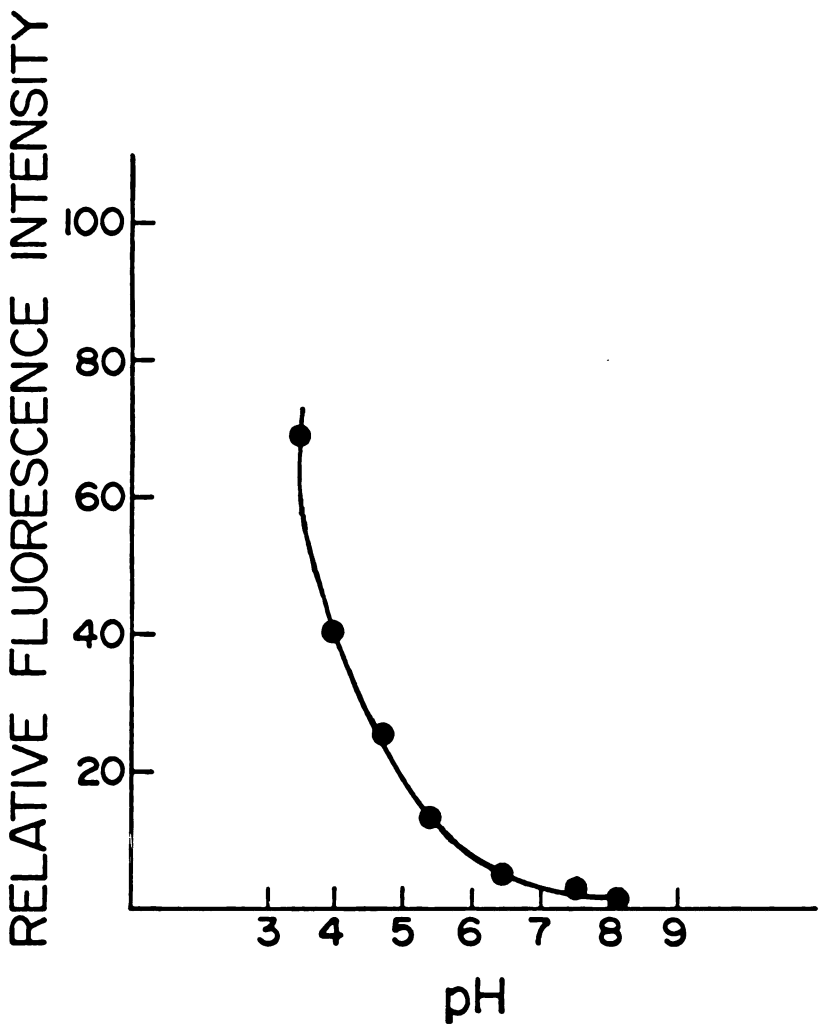


Figure 7. Relative fluorescence intensity of 2×10^{-5} M ANS in the presence of 10^{-4} M α-CHT (24); fluorescence intensities were measured at the emission maxima.

structures were investigated at pH 3.6 and 6.6.

E. Structural Studies of ANS

The crystal structure of ANS itself was also of interest due to the possible conformation dependence of its fluorescence (27,29). In addition, ANS is a member of the class of *peri*-substituted naphthalenes, which are known to exhibit a distorted naphthyl ring geometry arising from intramolecular steric effects (36,37).

Cody and Hazel have determined the structure of the ammonium (38) and magnesium (39) salts of ANS. Both crystallized in the triclinic system. In recrystallizing ammonium ANS for the protein work the crystals were found to be monoclinic. Since this suggested an additional conformation for ANS, the crystal and molecular structures of this crystal form were determined.

II. STRUCTURE DETERMINATION OF ANS

A. Experimental

Crystal Preparation and Intensity Data Collection

Ammonium ANS, purchased from the Sigma Chemical Company (St. Louis, MO), was recrystallized by evaporation of an aqueous solution at room temperature. A single crystal of approximate dimensions 0.7 x 0.3 x 0.15 mm was selected and mounted on a glass fiber. Weissenberg and precession photographs indicated that the crystals were monoclinic, space group $P2_1/c$.

Diffraction measurements were performed at 20°C using a Picker FACSI computer-interfaced diffractometer with Mo k_α radiation and a graphite monochrometer located between the x-ray source and the crystal. A least squares procedure was used to obtain precise unit cell parameters by fitting the setting angles of twelve reflections with 2θ greater than 35°. These data and other characteristics of the crystals are given in Table III.

A set of 3073 reflections was measured for $2\theta \leq 50°$, and of these, 2793 were unique. Intensity data were measured by means of the method of θ-2θ scans. A 20 sec background count was performed at the beginning and end of each θ-2θ range. Three independent axial reflections, $33° < 2\theta < 38°$, were chosen as intensity standards. Periodic measurement of these reflections during data collection indicated no significant change in intensity.

Table III. Crystal Data of ANS.

System Space Group	Monoclinic $P2_1/c$
a	6.150(a)Å
b	9.544(2)Å
c	26.774(7)Å
β	90.35(2)°
Z	4
μ_{Mo}	1.818 cm^{-1}
D_{cal}	1.413 g cm^{-1}

Both background-corrected intensities and their standard derivations contained corrections for Lorentz, polarization, and absorption effects. Values for the absorption correction, after the method of DeMeulenaer and Tompa (40), ranged between 1.022 and 1.074 (average 1.036), as calculated by using the atomic absorption coefficients of Cromer and Liberman (41). The standard deviations of the structure amplitudes, $\sigma(|F|)$, were calculated as in Wei and Ward (42) with application of an instrumental instability factor of 0.02. A weight of $1/\sigma^2$ was applied to structure factors during the least squares refinement.

B. Structure Solution and Refinement

The crystal structure was solved by use of the program MULTAN (43). The data set sufficiently over determined the structure so that only the 2077 reflections for which $|F| > 3[\sigma(|F|)]$ were included in the refinement.

The positions of the 21 skeletal atoms of ANS were located in the first E-map. After several cycles of full-matrix least squares refinement with isotropic thermal parameters ($R = 0.18$, $R = \frac{\Sigma||F_o|-|F_c||}{\Sigma|F_o|}$), a

difference map calculation showed two peaks which could be possible counter-ion positions. Although it was anticipated that one of these was the ammonium counter-ion and one a water molecule, both were favorably placed for hydrogen bonding with sulfonyl oxygen atoms. Each was included in the refinement as a nitrogen atom with an isotropic thermal parameter. After several additional cycles of isotropic and anisotropic refinement (R = 0.074) the 12 skeletal hydrogen atoms were located in a difference density and were included in subsequent calculations at theoretical hydrogen positions (for which an sp^2 amilino nitrogen geometry was assumed). A difference map was calculated again after additional refinement (R = 0.054, average parameter shift $\approx 0.01\sigma$, where σ is the standard deviation of the parameter). At this stage, the isotropic B values for the two possible counter ions still agreed within 2% (4.6 $Å^2$). However, the difference map revealed one of these to be located in an approximately spherical positive density. This atom was included as an oxygen, and a cycle of anisotropic refinement was performed in which the parameters of these two atoms alone were varied. The R-factor decreased to 0.049. Positions of the remaining six H atoms were now obtained in a new difference map, and two cycles of refinement, alternately varying H atom and non-H atom parameters, resulted in a final R-value

of 0.035. The final shifts in non-H atom parameters averaged 0.050σ.

C. Results

The final atomic coordinates and anisotropic temperature factors of the ANS, its ammonium counter ion and water of hydration are listed in Table IV. The numbering system used is shown in Figure 8(a). The coordinates of the hydrogen atoms and their isotropic thermal parameters are listed in Table V. Bond distances and angles of ANS are given in Figure 8 along with the errors in these quantities based on the standard deviations of the atomic coordinates of the final cycle of refinement. Inter-atomic distances and angles pertaining to hydrogen bonding are listed in Table VI.

D. Discussion

From an examination of the final bond lengths and angles, shown in Figure 8, it is evident that the ANS structure demonstrates distortions typical of peri-substituted naphthalenes and other ANS-like molecules. Notable among these are the in-plane deviations of the bond angles about C(1), C(8), C(9), C(1') and the anilino nitrogen.

Table IV. Fractional coordinates and thermal parameters ($Å^2$) for ANS non-hydrogen atoms. Thermal parameters are of the form $[-1/4\Sigma_i \Sigma_j (a^*_i a^*_j h_i h_j B_{ij})]$ where a^*_i is a reciprocal cell edge and h_i is a Miller index.

Atom	X	Y	Z	B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}
S	.6587(1)	.09201(1)	.17108(2)	3.97(3)	3.06(2)	3.14(2)	.28(2)	.07(2)	.49(2)
O(1)	.8829(3)	.0487(2)	.16459(6)	3.94(8)	4.23(8)	4.60(8)	.80(6)	-.11(6)	1.14(6)
O(2)	.6367(3)	.2318(2)	.19186(6)	5.48(9)	3.81(8)	3.63(7)	.52(7)	-.51(7)	-.63(6)
O(3)	.5378(3)	-.0107(2)	.20017(6)	5.21(9)	4.71(8)	4.07(8)	-.11(7)	.67(7)	1.48(7)
C(1)	.7620(4)	.2790(2)	.06575(8)	3.8(1)	3.3(1)	3.5(1)	-.14(9)	.10(9)	.37(8)
C(2)	.7874(5)	.3580(3)	.0233(1)	6.1(1)	4.8(1)	4.3(1)	-1.3(1)	.3(1)	1.0(1)
C(3)	.6420(6)	.3531(3)	-.0164(1)	9.4(2)	5.2(1)	3.1(1)	-.6(2)	.1(1)	1.1(1)
C(4)	.4710(5)	.2659(3)	-.01449(9)	7.1(2)	5.0(1)	3.5(1)	-.0(1)	-1.2(1)	-.2(1)
C(5)	.2605(4)	.0848(3)	.0268(1)	5.1(1)	6.0(1)	4.4(1)	-.4(1)	-1.3(1)	-.7(1)
C(6)	.2214(5)	-.0011(3)	.0654(1)	5.2(1)	6.2(2)	5.4(1)	-2.2(1)	-.6(1)	-.6(1)
C(7)	.3547(4)	.0058(3)	.10751(9)	4.8(1)	4.4(1)	4.2(1)	-1.0(1)	.3(1)	.0(1)
C(8)	.5284(3)	.0950(2)	.11098(8)	3.4(1)	2.55(9)	3.25(9)	.10(8)	.22(8)	-.19(8)
C(9)	.5817(3)	.1849(2)	.06950(8)	3.6(1)	2.78(9)	3.04(9)	.40(8)	.22(8)	-.15(8)
C(10)	.4370(4)	.1795(2)	.02728(8)	4.8(1)	3.7(1)	3.3(1)	.2(1)	-.41(9)	-.32(9)
N	.9171(3)	.2916(2)	.10349(8)	4.3(1)	2.97(8)	5.1(1)	-.62(8)	-.95(8)	1.08(8)
C'(1)	1.0082(3)	.4162(2)	.11995(8)	3.3(1)	3.3(1)	3.28(9)	-.37(9)	.41(8)	.57(8)
C'(2)	1.2109(4)	.4131(3)	.14340(9)	3.6(1)	4.1(1)	4.4(1)	.1(1)	-.01(9)	.4(1)
C'(3)	1.3006(5)	.5326(3)	.1630(1)	5.0(1)	5.8(2)	4.7(1)	-1.6(1)	-.5(1)	.1(1)
C'(4)	1.1958(6)	.6576(3)	.1597(1)	8.4(2)	4.5(1)	4.2(1)	-2.1(1)	-.3(1)	.1(1)
C'(5)	.9973(6)	.6644(3)	.1360(1)	8.8(2)	3.1(1)	4.3(1)	.6(1)	1.2(1)	.6(1)
C'(6)	.9013(4)	.5451(3)	.11633(9)	4.8(1)	3.8(1)	4.1(1)	.6(1)	.3(1)	.90(9)
N*	.8716(3)	.4587(2)	.24337(8)	4.5(1)	5.1(1)	4.5(1)	.02(9)	.15(8)	.61(8)
O_W	.7370(3)	.7476(2)	.24022(7)	7.4(1)	6.3(1)	5.6(1)	2.1(1)	1.56(9)	1.38(9)

Table V. Fractional Coordinates and Thermal Parameters ($Å^2$) for ANS Hydrogen Atoms.

Atom	X	Y	Z	B_{iso}
H(2)	.917(4)	.413(2)	.0216(8)	6.2(6)
H(3)	.661(4)	.413(3)	-0.436(9)	7.7(6)
H(4)	.355(4)	.246(2)	.4583(8)	6.5(6)
H(5)	.173(4)	.079(2)	-.0038(9)	6.1(6)
H(6)	.102(4)	-.064(2)	.0661(9)	7.3(6)
H(7)	.319(4)	-.057(2)	.1352(8)	5.7(5)
H(N)	.960(3)	.215(2)	.1185(7)	5.2(5)
H'(2)	1.283(3)	.330(2)	.1436(7)	5.5(5)
H'(3)	1.433(4)	.528(3)	.1799(9)	6.2(6)
H'(4)	1.274(4)	.731(2)	.1721(8)	7.5(6)
H'(5)	.917(4)	.744(2)	.1327(8)	5.5(6)
H'(6)	.761(3)	.545(2)	.0997(8)	4.9(5)
H_N*(1)	.794(3)	.375(2)	.2242(8)	9.2(5)
H_N*(2)	.874(3)	.539(2)	.2259(8)	9.0(5)
H_N*(3)	.983(4)	.420(2)	.2437(9)	8.9(5)
H_N*(4)	.767(4)	.459(2)	.2648(8)	10.1(5)
H_W(1)	.676(4)	.810(3)	.2229(9)	8.9(6)
H_W(2)	.625(4)	.737(3)	.262 (1)	9.6(6)

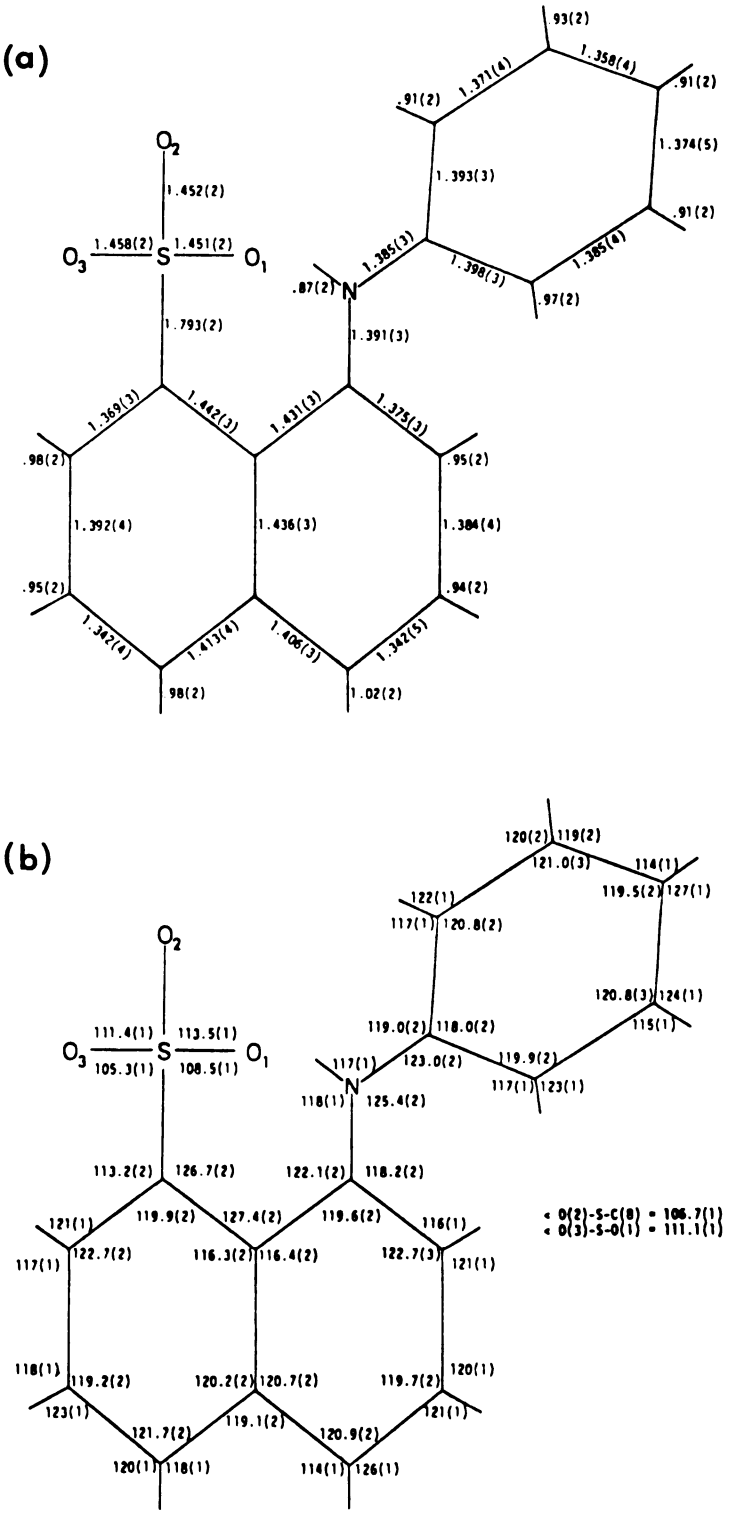


Figure 8. (a) Bond lengths of ANS(Å). Standard deviations are in parentheses.

(b) Bond angles (deg.) of ANS. Standard derivations in parenthesis.

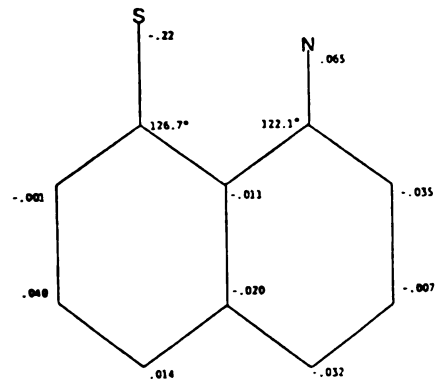
Table VI. Geometric Parameters of the Hydrogen Bonding Scheme in ANS. Atoms Constituting the Ammonium Ion are Designated with an Asterisk.

Distances (Å)		Angles (deg)	
N-O(1)	2.942	N-H(N)-O(1)	148
N*-O(2)	2.95	N*-H*(1)-O(2)	176
N^*-O_W	3.09	$N^*-H^*(2)-O_W$	133
$O_W-O(3)$	2.90	$O_W-H_W(1)-O(3)$	163

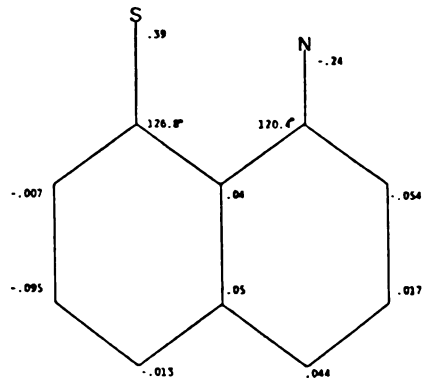
A puckering of the naphthyl ring is accomplished by a twist about the C(9)-C(10) bond, and it is accompanied by a splaying of the peri-substituents out of the mean plane of the ring system. Another consequence of this deformation is that the normals to the planes formed by 6-membered portions of the naphthalene fragment tilt by about 2.5° from their parallel orientations. Similar effects have been noted in previous studies (37,38,39). Distortions of the naphthyl portions of ANS are summarized in Figure 9 for several different ANS structural determinations. In contrast, Cruickshank and Sparks (45) found naphthalene to be planar to within 0.010Å. The sum of the in-plane expansions of the S-C(8)-C(9) and N-C(1)-C(9) angles is seen to bear an inverse relationship to the amount of splaying, the latter measured by the sum of the N and S distances normal to the least squares plane.

Theoretical considerations anticipate unequal carbon-carbon bond lengths in naphthalene (44). This has been investigated in a number of experimental (45) and theoretical (46) approaches. Bond lengths calculated in these independent investigations are in agreement within one crystallographic e.s.d. (.005Å) reported in the former study.

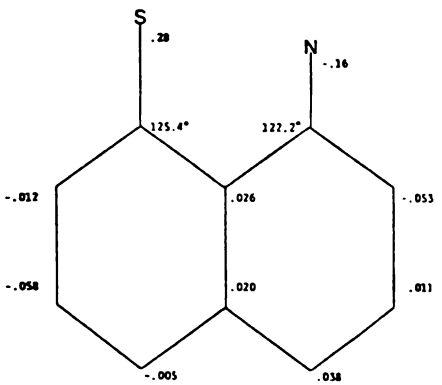
In peri-derivatives, the ring bonds involving C(1)



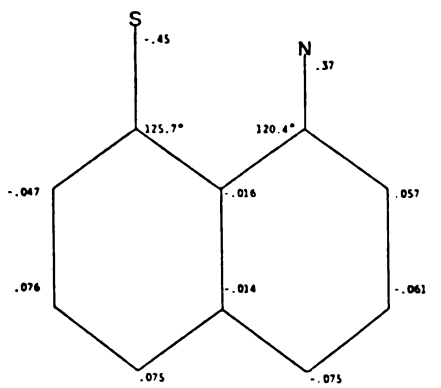
ANS
S=8.8°
D= .29Å
A=2.5°



ANS(1) (38)
S=7.2°
D= .63Å
A=5.6°



ANS Mg (39)
S=7.6°
D= .44Å
A=3.5°



ANS(2) (38)
S=6.1°
D= .82Å
A=6.5°

Figure 9. Deviations (Å) of naphthyl carbon atoms from best least squares plane of several ANS structure determinations. Additional parameters are:

S = [(<S-C(8)-C(9)) + (<N-C(1)-C(9))-240°]

D = combined distances of S and N atoms from best least squares plane.

A = angle between normals to least squares planes of 6-membered rings (A and D are directly correlated).

and C(8) have been found to be elongated relative to naphthalene, and those involving C(4) and C(5) have been found to be shortened (47). The average difference between the C(1) and C(4) and the C(5) and C(8) sets in this work are 0.029 Å and 0.028 Å, respectively. This agrees well with other ANS determinations (38,39), where the common molecule has a potential for extended π overlap between the naphthyl and phenyl rings through the amilino nitrogen atom. However, it also agrees closely with 1,8-di(bromomethyl)naphthalene, for which an average difference of 0.025 Å was reported by Robert, Sherfinski, Marsh and Roberts (37), and thus confirms their suggestion that the major cause of the effect is steric rather than electronic.

Comparing the anilino geometry, C(1')-N bond length, and phenyl ring orientation of the present work with the same parameters of other determinations (Table VII), one arrives at a view consistent with the assertion of Cody and Hazel (39) that a more coplanar orientation enhances conjugation between the anilino group and phenyl ring, which results in a shorter C(1')-N bond. However, inspection of the planarity of the anilino linkage with respect to the naphthalene plane does not give similar correlations with either the C(1)-N bond length, naphthyl planarity, <N-C(1)-C(9) or <S-C(8)-C(9). In all four molecular structures the C(1')-N and C(1)-N bond lengths are significantly shortened with respect to single-bonded

Table VII. Conformational Parameters of ANS Observed in Several X-ray Crystallographic Structure Determinations.

Parameter	Structure			
	ANS_{MSU}	ANS(1)*	ANS(2)*	ANS-Mg**
C(1')-N	1.385Å	1.415	1.386	1.388
C(1)-N	1.391Å	1.409	1.393	1.404
C(8)-S	1.793Å	1.788	1.796	1.786
C(1')-N-C(1)	125.4°	123.4	127.3	126.1
C(2')-C(1')-N-C(1)	-156°	112	171	174
C(1')-N-C(1)-C(2)	42.5°	21.8	-46.0	-54
H-N-C(1')-C(2')	24.5°	-12	-27	-21
H-N-C(1)-C(9)	40.2°	-36	-22	-35
Σ bond angles of N	360.0°	335.7	357.6	358.7
C(9)-C(8)-S-O(1)	-64°	70	77	664

* Reference (38).

**Reference (39).

values, which would be 1.470 .005 Å for a pure $C(sp^2)$-$N(sp^2)$ case (48) and would be slightly longer for a pure $C(sp^2)$-$N(sp^3)$ case (due to the greater p orbital character of the nitrogen atomic orbitals). Thus, if there are conformation-linked variations in conjugation within ANS, their effects on structure are once again outweighed by those of steric origin. The foregoing is consistent with the view that out-of-plane deformations in aromatic systems are energetically more favorable, and produce less of an effect on π orbital overlap than has been generally acknowledged (49).

Penzer (23) measured chemical shifts of ANS ring protons by NMR experiments in deuteromethanol and D_2O. He found ring current deshielding to be selectively operative in the former medium, wherein ANS is highly fluorescent. It was concluded that this implied a more planar conformation in deuteromethanol. A conformation-linked fluorescence dependence was suggested which was governed by solvent-enhanced favorability of intramolecular hydrogen bonding in non-polar media. However, a comparison of the ANS torsion angles C(1')-N-C(1)-C(2) and C(2')-C(1')-N-C(1) (Table VII) indicates that as the coplanarity measured by the former increases, that measured by the latter decreases. This observation implies that the NMR effects recorded by Penzer (29) may not be the direct response of variable intramolecular hydrogen bonding.

Cody and Hazel (38,39) pointed out that such hydrogen bonding was observed in all known crystallographic conformations. This is true in the present study as well, and it is consistent with their assertion that it is the peri-group geometry that determines such hydrogen bonding, which persists in solution.

Based on a solution spectroscopic study of 2-anilinonaphthalene-6-sulfonate (2,6-ANS), Kosower and coworkers (15) suggested an opposite conformation dependence which depends on excited state intramolecular charge transfer. Although the peri-interaction and intra-molecular hydrogen bonding ability is absent in 2,6-ANS, the reported polarity dependence of its fluorescence in solution can be seen as qualitatively similar to that of ANS (Figure 4a (1), Reference 15). Two key requirements for the formation of an intramolecular charge transfer complex are: the ability of the phenyl ring to communicate electronically with the naphthyl - for which a planar conformation in the excited state has been presumed necessary - and the ability of the environment to stabilize the charge separation. Quenching in polar media was proposed to occur via a second excited state charge transfer pathway that leads to the ground state. (The latter conformational considerations oppose those of Penzer (23), but the solution NMR based arguments apply only to the ground state.) It is evident from the crystallographic results that coplanarity is not a

requirement for significant anilino-naphthyl π orbital overlap in ANS, and this suggests a less sensitive dependence of the spectral properties on conformation than has been supposed. The ground state conformation is indeed likely to relate closely to that of the excited state in the crystal, but the fluorescence properties of ANS so constrained may not generally relate to its solution properties.

Camerman and Jensen (47) have reported the molecular structure of 2-p-toluidinyl-6-naphthalene sulfonate (TNS) and have compared the fluorescence emission spectrum of these crystals with that of a hydrated form for which the unit cell parameters indicated two water molecules per molecule of TNS. The anhydrous structure exhibited contracted bond lengths about the anilino nitrogen, comparable to that in ANS and a fluorescence emission maximum similar to that of TNS in non-polar solvents. The emission spectrum of the hydrated form was similar to that of TNS in water, and since the crystals lost water readily, it was suggested that the lattice water was able to reorient about the excited state fluorophore. However, no data concerning fluorescence lifetimes or the relative emission intensities were given.

The anhydrous crystal form presents a more rigid environment to the TNS molecule. The fact that a significant π overlap between the ring systems was observed with a

relatively high energy emission maximum does not disqualify an intramolecular charge transfer quenching mechanism in solution.

An inspection of the packing of ANS in the unit cell reveals an extended array of hydrogen-bonded atoms (Table Hydrogen bonding takes place between the ammonium ion and water molecule within the asymmetric unit and between these and two of the sulfonyl oxygens of different ANS molecules. This scheme gives rise to a distorted helical arrangement parallel to the $\vec{b}$ axis.

III. STRUCTURE OF THE ANS-α-CHT COMPLEX, AND THE NATURE OF ITS pH DEPENDENCE

A. Experimental

1. Crystal Preparation

Single crystals of native α-CHT were grown at pH 3.6, room temperature (20°-25°C), according to a method outlined previously (51,52). The α-CHT was obtained from Worthington Biochemical Corporation (Freehold, NJ) as the three-times crystallized, salt-free lyophilized reagent. When the diamond-shaped plates of crystalline α-CHT reached an appropriate size (long axis ≈ 0.8 to 1.3 mm) the mother liquor was removed and replaced with a solution of distilled water which was 75% saturated in ammonium sulfate and adjusted to pH 3.6 by addition of sulfuric acid. Test tubes containing crystals and storage solution were placed in a bath cooled by circulating tap water which maintained a temperature roughly between 14° and 18°C. Stored in this manner, crystals of α-CHT will retain their diffraction quality for several years.

A crystal was removed and sealed in a Debye-Scherrer capillary, in contact with its storage solution (53). Then, axial x-ray diffraction intensities were scanned to allow comparison with those of native α-CHT. Once the identity of this crystal as native α-CHT was verified, along with its diffraction quality, half the volume of

storage solution was replaced with storage solution that had been saturated with ANS (ammonium salt, Sigma Chemical Company, St. Louis, MO, Lot 28B-0870). This replacement procedure was repeated twice a day for three days. On day four the entire volume of solution was replaced. On day five some crystals of ANS were added to insure saturation, the pH was verified and the tube was stored. On standing for several days, the protein crystals became highly fluorescent.

After one month of soaking, the axial intensity distributions of a selected protein crystal showed significant changes with respect to those of the native enzyme, and this suggested that complexation was taking place. In addition to these changes, the unit cell parameters showed that the volume of the derivative crystal increased slightly (Table VIII). A three dimensional set of intensity data was measured at 2.8 Å resolution using this crystal.

Six months later, when the structure analysis was nearly complete the pH 6.6 experiment was being contemplated, and the diffraction pattern of an ANS-α-CHT crystal from the original soaking tube was examined. Surprisingly, the unit cell dimensions had increased significantly (Table VIII). The axial intensities showed additional but small changes beyond those observed previously. Another 2.8 Å resolution three dimensional data set at pH 3.6 was therefore

Table VIII. Unit Cell Parameters of Native and ANS-α-CHT pH Conformers.**

	a(Å)	b(Å)	c(Å)	β(°)	$Vx10^3(Å^3)$
Native, pH 3.6	49.24(7)	67.20(10)	65.94(9)	101.76(6)	213.60(10)
ANS-α-CHT, pH 3.6	49.32(4)	67.38(4)	66.02(6)	101.80(4)	214.80(5)
ANS-α-CHT, pH 3.6*	49.52(5)	67.55(6)	66.04(7)	101.88(4)	216.20(6)
ANS-α-CHT, pH 6.6	49.24(5)	67.31(6)	65.62(7)	102.07(6)	212.70(7)
Native, pH 5.4	49.13(5)	67.83(7)	65.81(7)	101.92(6)	214.60(7)

* six month soak.

**Least squares fit (54,55).

collected.

Finally, the pH of the six-month-soaked crystals was increased several times daily by replacement of 0.2 ml (of a total of 3.0 ml) of soaking solution, with a solution 75% saturated in ammonium sulfate, saturated in ANS, and at pH 6.6. After about one week, the pH of the mother liquor soaking solution had attained a value of 6.6. At that time, all of the storage solution was exchanged. During this procedure the protein crystals were observed to undergo a degree of cracking, but this did not seriously affect their diffraction quality (56).

2. Data Collection

All x-ray crystallographic measurements performed on the α-CHT derivatives utilized an automated Picker Nuclear Corp. FACSI four-circle diffractometer, controlled by a Digital Equipment Corp. (DEC) 4K PDP-8 computer coupled to a DEC 32K disc File, and equipped with an AMPEX TMZ Digital Tape Memory System for magnetic tape data storage. The tape system operated with an Eclectic Corp. 640 interface. Physical modifications of the diffractometer include a rapidly operating solenoid-actuated balanced filter system, placed between the x-ray source and the crystal in order to reduce sample exposure. Also, a 60 cm helium filled tube intervenes between the crystal and detector, thus improving the resolving power of the instrument and the

signal-to-noise ratio of intensity measurements.

The associated software package, based originally on that supplied by Picker (57), calculates unit cell parameters and refines crystal orientation by a least squares procedure after the method of Busing and Levy (54). Substantial modification of the software (55) has introduced a "wondering" count-six-drop-two ω step scan intensity measurement protocol, after Wyckoff and co-workers (58), and the capability of periodically revising the crystal orientation matrix, as required by small movements of the protein crystal in its wet capillary mount.

Each three dimensional data collection was preceded and followed by axial intensity distribution measurements (run-outs) and by a 6 Å resolution (h0ℓ) intensity data collection. The initial run-outs were part of the crystal selection procedure. The final run-outs and (h0ℓ) data collection were used to corroborate the effects of radiation induced crystal damage, otherwise monitored automatically every one hundred reflections (every 1/65 of the data collection) by measuring three intensity standards. The initial h0ℓ data sets (centrosymmetric in $P2_1$) were converted to two-dimensional electron densities during the data reduction procedure, under the assumption that the signs of the protein structure factors were unaltered by experimental perturbation of the native structure. Such Fourier maps were used to determine an approximate

absolute scale factor for each derivative by comparing them with that of the native enzyme, which was scaled previously (59).

Twinning of α-CHT crystals necessitates measurement of the relative scattering strength of the coexistent lattices. As described elsewhere in detail (59), twinning occurs about the $\vec{c}^*$ direction, with the (0kℓ) reflections unresolvable. The latter are corrected during data reduction by removal of their intensity component which has originated from the more weakly scattering lattice (referred to as the twin). A ratio of diffraction intensities is calculated between the more strongly scattering lattice (the crystal) and the twin by measuring the intensity of the (600) reflection (readily identified and resolved) for each prior to data collection. The position of $(600)_{crystal}$ is retained to define the direction of $\vec{a}^*$ in data collection. The unit cell dimensions and space group of the crystal and twin are the same.

Absorption characteristics of α-CHT crystals are somewhat complicated by the additional presence of solvent, which can vary in amount and composition, and by the presence of the glass capillary, both of which are bathed by the x-ray beam. A semi-empirical absorption correction technique (60) has been adopted. Our protein crystals are mounted with $\vec{b}^*$ along the spindle axis (ϕ) of the diffractometer, so that the ϕ dependence of the intensity of an

(0k0) reflection is only a function of the absorption characteristics of the crystal and mount. The method is based on the assumption that the observed absorption coefficient of a crystal in a given orientation will be the mean of that along the path of the incident and diffracted beam. The 2θ dependence is obtained by measuring absorption characteristics for a series of (0k0) reflections, over a range of θ values (>180°) which include the quadrant of reciprocal space intended for data collection. During data reduction, in the case of general reflections, the orientation of the crystal with respect to the incident and diffracted beam is related geometrically to a 2θ dependent (0k0) geometry.

Lack of balance between the Cuβ (Ni) and $Cu\alpha_1\alpha_2$ (Co) filters results from the fact that part of the background radiation is composed of incoherently scattered radiation. For each crystal, prior to data collection, the lack of balance was measured as a function of 2θ in a background region of reciprocal space, and has the form

$$I_{LB} = I_{Ni}(2\theta) - I_{Co}(2\theta) \ . \qquad (8)$$

It is used to augment the background count for each reflection, $I_{Co}(hk\ell)$, during data reduction.

In order to extract the most useful information in the most efficient manner, only about 6200 reflections are

measured out of 9900 significant reflections at 2.8Å resolution (92% of the total). The 6200 that are measured correspond to reflections for which the figure of merit (61,62) is greater or equal to 0.7. In this way, the most reliable phase angle data set can be obtained from one crystal, which will usually experience a decay of its intensity standards by no more than 10%.

3. Data Reduction

A set of structure amplitudes is derived from the measured intensity data, I (hkℓ), as

$$|F_{hk\ell}|^2 \alpha\, I\ (hk\ell)\ . \qquad (9)$$

The proportionality constant is composed of several crystallographic and instrumental corrections, applied during data reduction. More explicitly the relationship is:

$$|F_{hk\ell}|^2 = \left(\frac{1}{\sin 2\theta}\right)\left(\frac{1+\cos^2(2\theta)}{2}\right)\left(\frac{I_{std}^{t=0}}{I_{std}^{t=hk\ell}}\right)\left(\frac{I(hk\ell)_{max}}{\frac{1}{2}[I(\phi_{inc})+I(\phi_{refl})]}\right).$$

$$(k^2\exp(-2B\sin^2\theta/\lambda^2))[I_{Ni}(hk\ell)-(I_{Co}(hk\ell)+I_{LB}(2\theta))]. \qquad (10)$$

The first factor is the Lorentz correction, which takes the finite size of reciprocal lattice points into account.

Its form is determined by the geometry of the diffraction experiment. The second corrects for preferential diffraction of the component of the incident beam polarized parallel to the reflecting planes. The third factor is the decay correction, where $I_{std}^{t=0}$ is the average intensity of three intensity monitors at the beginning of data collection (measured approximately once per hour during a 65 hour data collection), and $I_{std}^{t=hk\ell}$ is their interpolated intensity at the time of measurement of (hkℓ). The fourth is the absorption correction (62), where $I(hk\ell)_{max}$ is the maximum intensity of an 0k0 reflection of similar 2θ as the general reflection being considered. The inc and refl refer to incident and reflected respectively, and for the four circle diffractometer have the form

$$I(\phi_{inc}) = I(\phi_{hk\ell} - \varepsilon_{hk\ell}) \qquad (11a)$$

$$I(\phi_{refl}) = I(\phi_{hk\ell} + \delta_{hk\ell}) \qquad (11b)$$

where $\varepsilon = \delta = \sin^{-1}(\sin\theta\cos\chi)$. The fifth factor contains the scale factor k^2 and an exponential temperature normalization factor. The latter is determined graphically by averaging the reduced $|F|^2$ in ranges of 2θ and fitting the distribution to that of native α-CHT. The k^2 is adjusted simultaneously and includes a factor from the initial (h0ℓ) electron density scaling. The final factor

written in Equation (10) is I(hkℓ), as corrected for background and lack of balance. An additional factor is added to correct 0kℓ reflections for twinning. A summary of the data reduction parameters used herein is given in Table IX.

4. Difference Electron Densities

Reduced sets of derivative structure amplitudes were converted to "best" difference electron densities (61,62) for comparison with the native α-CHT structure. The best phase for a reflection is that represented by the centroid of a Gaussian phase-value probability function plotted in complex space ("Argand diagram") (61,62). An assumption is made that the errors arising from the use of a native phase model in conjunction with a figure of merit weighting scheme are small and readily characterized (63,64).

We define $|\Delta F_{hk\ell}|$ as

$$|\Delta F_{hk\ell}| = |F^{ND}_{hk\ell}| - |F^{N}_{hk\ell}|, \qquad (12)$$

where $|F^{N}_{hk\ell}|$ designates a native protein structure amplitude, and $|F^{ND}_{hk\ell}|$ a derivative structure amplitude. The best difference electron density is

$$\Delta\rho_{BEST}(x,y,z)=\frac{1}{V}\sum_{hk\ell} m(hk\ell)|\Delta F_{hk\ell}|\exp(-2\pi i(hx+ky+\ell z) + i\alpha(hk\ell)_{BEST}) \quad , \qquad (13)$$

Table IX. Data Reduction Parameters for Several α-CHT Derivatives.

Derivative	Twin Ratio $\frac{I_{xtal}(600)}{I_{twin}(600)}$	Absorp. Max/Min	k^2	$B(Å^2)$	Decay %	Decay Refl. Range
ANS-α-CHT pH 3.6	13.9	1.55	1.26	0.00	6.4	0-6408
ANS-α-CHT pH 3.6 (long soak)	2.35	2.00	2.36	1.40	4.1 -6.4	0-2626 2627-6428
ANS-α-CHT pH = 6.6	1.29	1.76	1.14	0.300	8.9 0.0	0-2070 2071-6235

where x, y, and z are fractional unit cell coordinates, m(hkℓ) is the figure of merit, $\alpha(hk\ell)_{BEST}$ is the "best" native phase angle, and V the unit cell volume.

The rms error of the difference electron density ($\sigma(\Delta\rho)$), calculated by the method of Henderson and Moffat (63), with correction (64), ranges from about ±0.02-0.04 $e\text{Å}^{-3}$ in this study. However, previous work on the pH 5.4 α-CHT conformer in this laboratory (65) indicates that 0.05 $e\text{Å}^{-3}$ is a more realistic level for $\sigma(\Delta\rho)$. We have adopted a significance level for $\Delta\rho_{BEST}$ features at approximately three times 0.05 $e\text{Å}^{-3}$. An empirical and more conservative calculation with several of the derivatives now under study, of

$$\sigma(\Delta\rho) = \frac{1}{V}\left[\left(\sum_{xyz}[\Delta\rho(xyz)]^2/N\right]^{1/2}, \qquad (14)\right.$$

has confirmed the latter as an appropriate significance level.

B. Results

1. ANS Substitution at pH 3.6

The "best" difference map displayed two prominent saucer-shaped regions of electron density reaching 0.40 $e\text{Å}^{-3}$ in height ($\sim 8\sigma(\Delta\rho)$). The map was drawn to a scale of 2 cm Å^{-1} to correspond to a Kendrew model of α-CHT, and the saucer-shaped regions were readily fitted with models of ANS. An additional corraboration of the excellent overall fit of the molecular model of ANS was provided by comparison with that of p-toluenesulfonyl fluoride which tosylates α-CHT at Ser 195 and establishes an unambiguous sulfonyl location (52).

Fluorescence measurements have indicated that the ANS binding site of α-CHT is identical in the crystal to that in solution at pH 3.6 (24). The saucer-shaped difference electron densities are located in the vicinity of the amino terminus of the A-chain of α-CHT (Cys 1-122) and are situated between dimer molecules related by an approximate local "inter-dimer" 2-fold rotation axis (dyad B of Figure 10). In fact, from only a cursory inspection, it was not clear whether the binding was intramolecular (e.g., 1 and 1' or 2 and 2', Figure 10) or intermolecular (1-2' and 1'-2). However, the anions $PtCl_4^=$ and $PtI_4^=$ are known to bind intramolecularly in this region near Cys 1-122 (59,66), a moiety with which

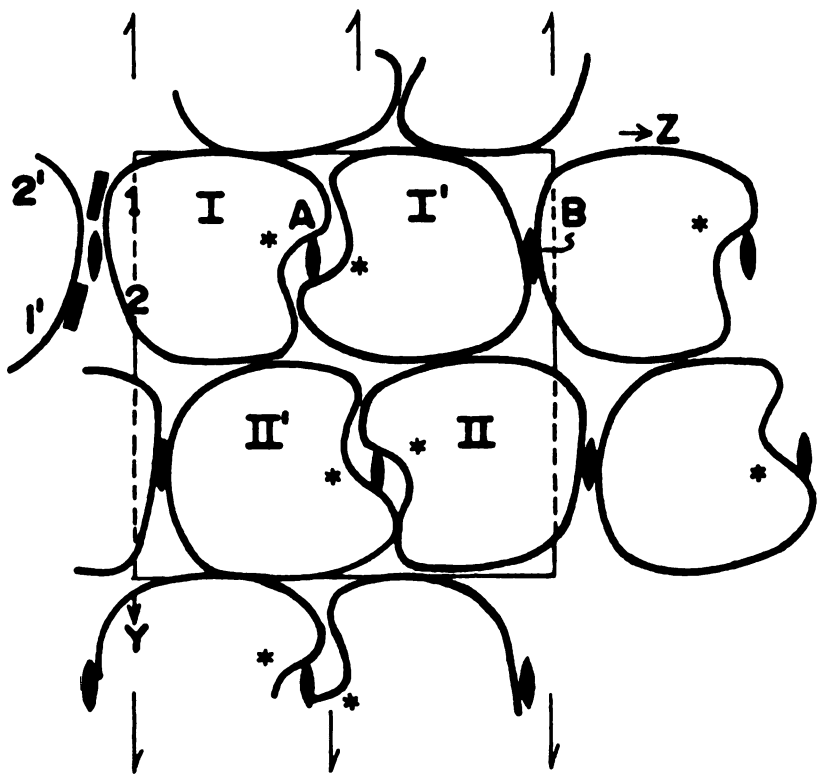


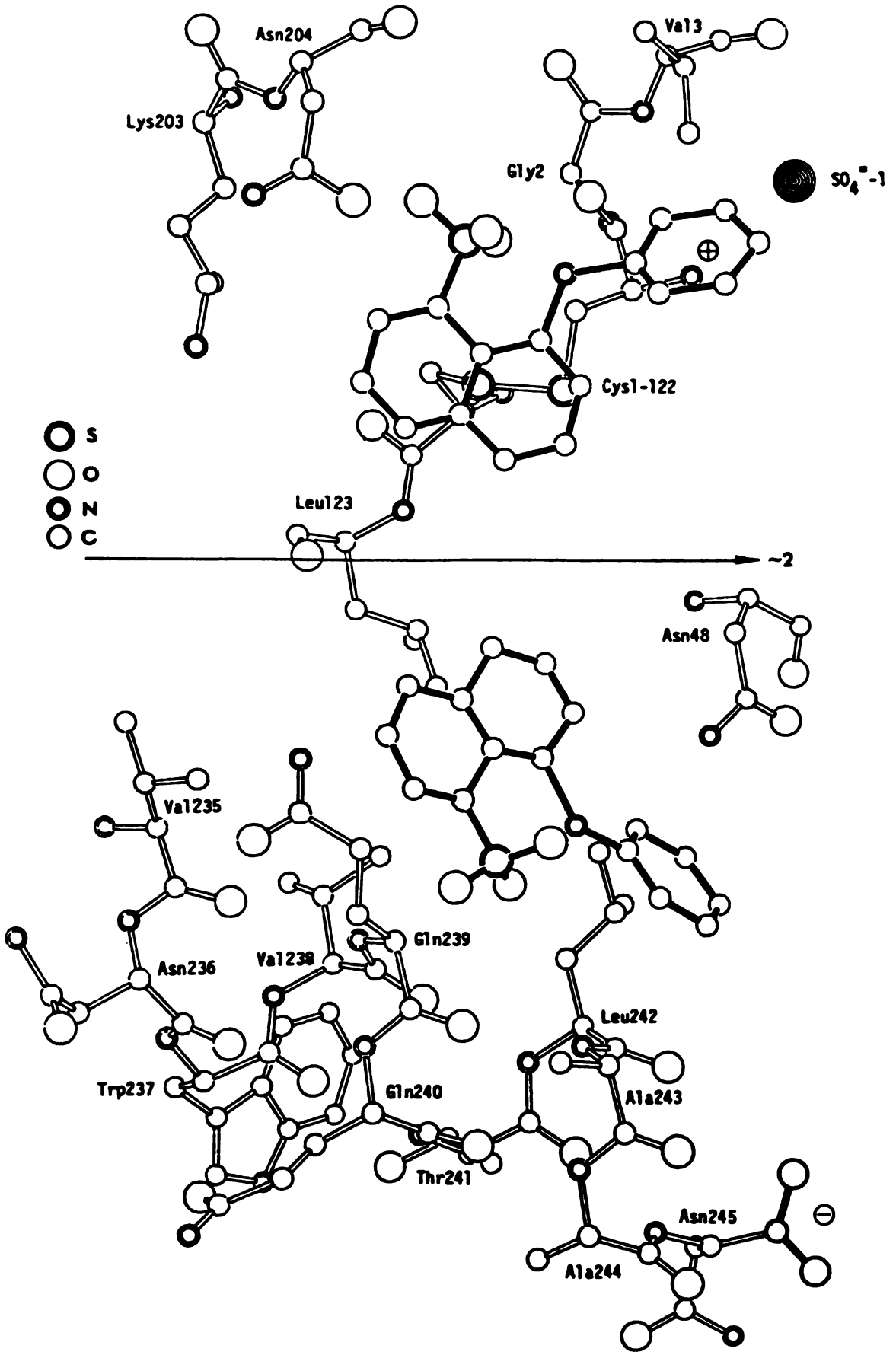
Figure 10. Schematic packing diagram of α-CHT viewed down the a* direction. Molecules I and I' form an asymmetric unit and are related by non-crystallographic 2-fold axes A and B; asymmetric units related by crystallographic 2-fold screw axes shown appropriately parallel to y-axis; asterisks denote active-site regions near center of dimer; 1 and 1' denote intramolecular ANS binding sites; 2 and 2' denote close ANS intermolecular contacts.

the negatively charged ANS makes van der Waals contact. In addition to the ANS density, there were some other smaller regions corresponding to small changes in the structure of α-CHT in the vicinity of the ANS binding. Otherwise, the remainder of the difference map was generally below background level ($\sim$0.05 eÅ^{-3}).

2. ANS Binding Site

There are only two general close contacts between the ANS and protein molecules in the crystal, and they remain the same following a change in pH from 3.6 to 6.6. The most extensive (nine contacts less than 4.4Å) and striking of these involves the disulfide bridge of Cys 1-122 (environment 1 or 1', Figure 10) which is also the amino terminus of the A-chain of α-CHT. The other close ANS-protein contact (six contacts less than 4.4Å) occurs in the vicinity of the side chain of Gln 239' which is located in a 2-fold-related neighboring molecule (molecule I') (environment 2' or 2, Figure 10). The Gln 239 is part of an α-helix formed by residues Val 235-Asn 245 at the carboxyl terminus of the C-chain. A drawing of the structure of the environment of ANS in the ANS-α-CHT complex viewed through the ANS toward the Cys 1-122 disulfide bridge is shown above the local 2-fold axis in Figure 11, while that viewed in the opposite direction, toward the α-helix of a neighboring molecule, is shown

Figure 11. Drawing of vicinity of ANS binding viewed down C-direction of crystal. ANS molecules shaded; local interdimer 2-fold axis shown appropriately; lower part of drawing constitutes remaining environment of upper ANS and vice versa; ANS molecules and local 2-fold axis in a plane perpendicular to view; ANS sulfonate and phenyl rotational orientations arbitrary but reasonable (see text).


Asn204
Lys203
Val3
Gly2
$SO_4^{=}$-1
Cys1-122
S
O
N
C
Leu123
-2
Asn48
Val235
Gln239
Asn236
Val238
Leu242
Trp237
Gln240
Ala243
Thr241
Asn245
Ala244

below the local 2-fold axis. The ANS molecules and the inter-dimer local 2-fold axis are essentially located in a plane perpendicular to the view in Figure 11.

The coordinates of the independent ANS molecules of ANS-α-CHT, pH 3.6 are listed in Table X. These were measured visually from the Kendrew model of ANS-α-CHT by use of a grid drawn on clear plexiglass which was suspended perpendicular to the yz plane (Figure 10); and also with the aid of a Richards optical comparator (67).

Close ANS-protein contacts are listed in Table XI, from which the close napthyl ring approach of ANS to the Cys 1-122 disulfide can be noted, particularly with respect to ring A (Figures 1, 11). Another prominent feature of this environment is the ion pair consisting of the quaternary N-terminal nitrogen atom of Cys 1-122 and its sulfate counter ion separated by about 3.2Å (Table XI, Figure 3), the latter having been located previously by a sulfate-selenate exchange experiment (68). Since ANS is a negative ion at pH 3.6, it is somewhat surprising that the binding of ANS in the same general vicinity did not significantly disturb the ion pair. Interestingly, however, the sulfonate group of ANS fits into an alternating charge array involving Asp 128-CO_2^-, Lys 203-NH_3^+, ANS-SO_3^-, Cys 1-NH_4^+, $SO_4^=$-1 (see Figure 11 also). The charge centers are separated by approximately 5.0, 7.5, 5.8 and 3.2Å, respectively. These distances appear to be large for strong interactions, but this may be due to

Table X. Coordinates of Bound ANS Molecules*

	Site 1			Site 1'		
Atom**	x	y	z	x	y	z
1	0.741	0.224	-0.012	0.747	0.363	0.006
2	0.757	0.239	-0.017	0.764	0.348	0.016
3	0.749	0.258	-0.022	0.753	0.331	0.022
4	0.721	0.264	-0.025	0.726	0.326	0.010
5	0.677	0.253	-0.025	0.682	0.339	-0.012
6	0.659	0.238	-0.019	0.668	0.355	-0.027
7	0.667	0.218	-0.016	0.680	0.373	-0.029
8	0.694	0.214	-0.015	0.706	0.377	-0.015
9	0.713	0.229	-0.018	0.721	0.360	-0.007
10	0.702	0.249	-0.021	0.709	0.343	-0.001
1'	0.780	0.199	-0.023	0.787	0.393	0.013
4'	0.834	0.192	-0.027	0.800	0.409	0.006
N	0.751	0.203	-0.015	0.760	0.385	0.000
S	0.707	0.192	-0.014	0.720	0.397	-0.016

Coordinates are fractional referred to the crystallographic axes; origin along y defined as in (59); local 2-fold axis approximately parallel to a-axis at y=0.291.

**Numbering as in I.

Table XI. Close Protein Contacts with ANS.*

Site 1			Site 2'		
Protein Atom	ANS	d**(A)	Protein Atom	ANS	d**(A)
Cys 1 - NH_3^+	N	4.2	Asn 48' - C_γ	4'	4.4
Cys 1 - S	2	3.0	Leu 123' - C_δ	10	4.4
Cys 1 - $C_{\gamma 2}$	1'	4.4	Gln 239' - C_δ	6	3.2
Cys 122 - S	9	2.9	Gln 239' - C_δ	7	3.1
Cys 122 - S	S	3.8	Leu 242' - C_γ	1'	4.0
Lys 203 - C_ε	6	4.1	Leu 242' - C_γ	4'	3.4
Lys 203 - N_ε	6	4.0			
Asn 204 - C_γ	S	4.4			
Asn 204 - O_γ	S	3.5			

Other Important Contacts

Protein Atom	ANS	d**(A)
Cys 1 - NH_3^+	S	5.9
$SO_4^=$ - 1	4'	4.9
Cys 1 - NH_3^+-$SO_4^=$-1		3.2

*See Figure 12 for exact nature of site designations.

**Estimated error $\pm$(0.6-0.7)Å.

the fact that hydrogen atoms intervene between the charge centers and limit closest approaches. More importantly, the charged array serves to stress the generally polar nature of the ANS binding region 1 and 1'. Other polar features in the binding locale include the side chains of Lys 203, Asn 204 and the side chain of Gln 239' from another molecule. The only non-polar side chains in the region are Leu 242', which makes a minor contact with the phenyl group of ANS, and Val 3 and Ala 206, both of which are relatively small sterically.

From the opposite side, the ANS makes a close contact with the carboxyamide side chain of Gln 239' and less importantly, with a methyl group of Leu 123' and C_γ of Asn 48', all of a neighboring molecule (below local 2-fold of Figure 11). Of the six contacts (Table XI), only three involve the naphthalene portion of ANS. The ANS here appears to span the planar surface of a hemispherical non-polar cavity created by the side chains of Ile 47', Leu 123', Val 235', Val 238' and Leu 242'. These residues are part of a larger internal spherical non-polar cavity near the center of the molecule (69) between the two cylinders (70) or folding domains of α-CHT. Part of the carboxyl terminal α-helix contributes prominently to this region (Figure 11). However, most of these non-polar features are quite distant to the ANS, especially relative to the proximity of the latter to the disulfide of Cys 1-122. The C_δ atom of the side chain of Gln 239'

is about 3.1Å and 3.2Å, respectively, from C_7 and C_6 of the naphthalene ring of ANS. Finally, all the close contacts in this region are far removed from the chemically active anilino and sulfonate groups of ANS (6.0-8.0Å).

The local 2-fold relationship between the difference densities corresponding to the independent ANS molecules is excellent, both in position and peak height, so that the fit of a model of ANS in each case leads essentially to an identical molecular orientation and conformation (Figure 11). Since the density corresponding to the phenyl group of ANS possessed axial symmetry about the N-C_1' direction, it was not possible to assign a rotational conformation and thus determine the coordinates of the C_2', C_3', C_5' and C_6' atoms of the ring.* However, since a rotation about the C_1-N bond displaces the phenyl ring, it was possible to assign C_1' coordinates and to calculate a C_1'-N-C_1-C_2 torsion angle. Parameters which characterize the conformation of ANS can be compared among: (1) The high resolution ANS structure determination reported in the previous chapter (ammonium salt, monoclinic form), (2) another high resolution ANS structure determination (ammonium salt, triclinic crystal form) with two molecules per asymmetric unit (38) and, (3) with parameters

*The phenyl-naphthyl dihedral angle in Figure 11 was chosen to optimize the van der Waals contacts within the ANS molecule.

obtained from the ANS-α-CHT complex. The torsional angle and the angle between the planes of the phenyl and napthyl rings of the foregoing are listed in Table XII from which it can be seen that the C_1'-N-C_1-C_2 torsion angle of ANS in the complex is similar to that in ANS (1) (38) but different from that of ANS_{MSU} and ANS(2). In any conformation, as observed by space filling models, the anilino hydrogen and the sulfonate oxygen atoms are sterically crowded. This precludes classical linear hydrogen bonding although it does not preclude an interaction among these atoms.*

While the changes in protein structure accompanying ANS binding will be discussed in detail below, it can be mentioned that the difference map of ANS-α-CHT, pH 3.6 shows a pronounced density gradient associated with the side chain of Gln 239'. Upon ANS binding, there is a rotation of the side chain of Gln 239' about the C_δ and C_γ atoms such that the polar end of the side chain is removed from the close proximity of the napthyl ring of ANS. This shift is of the order of 1Å and may be restrained by a possible interaction of the side chain with that of Lys 203 of another molecule. The side chains of Lys 203 and Lys 203' also move away from the fluorescence probe but to a lesser degree. Changes in structure on

---

*The rotational orientation of the sulfonate group about the S-C_β bond is arbitrary in Figure 11.

Table XII. Comparison of Conformations of ANS.

Parameter	ANS-α-CHT	ANS$^{*}_{MSU}$	ANS(1)	ANS(2)
$C_{1'}$-N-C_1-C_2 **	∿10°	43°	22°	-46°
α ***	?	61°	63°	127°

*Chapter II.

**Defined as zero when $C_{1'}$ is coplanar with N, C_1 and C_2.

***Angle between planes of phenyl and napthyl rings.

the opposite side of the ANS, near Cys 1-122, are restricted to those apparently involving solvent and do not always show local 2-fold equivalence. The region leading away from the quaternary terminal nitrogen along the local 2-fold axis (Figure 11) is generally composed of solvent. From all appearances, the ANS probably approaches the protein binding sites from this direction in the crystal.

3. Independent Evidence for the ANS Binding Site

Since the ANS binding site lies in an intermolecular region of the crystal, deduction of its intramolecular binding mode was based on the closeness and extent of van der Waals contacts. (Such a binding mode was also dramatically illustrated by the difference diagonal plot as outlined below.) Although these clearly showed that ANS contacts the protein more intimately in the Cys 1-122 region than in environment 2 or 2' (Figure 10), nonetheless, the possibility existed that the binding site in solution was associated with the more non-polar α-helical region, even though the van der Waals contacts there were fewer. This is especially so because complicated non-polar or hydrophobic interactions are still not understood well theoretically or in detail, and in any case, the entropic contribution to the free energy of complex formation is not known. In order to establish the ANS binding site more definitively, use was made of the fact

that $PtCl_4^=$ and $PtI_4^=$ are known to bind near Cys 1-122 in crystals of α-CHT (59,66). $PtI_4^=$ was found to strongly quench ANS fluorescence in solution studies of ANS-α-CHT, and this was accomplished via a mechanism of competitive binding (24,71). If the fluorescent probes were binding near the non-polar α-helical region of the protein its fluorescence would have been essentially unaltered by the presence of $PtI_4^=$.

4. Changes in Protein Structure of ANS-α-CHT, pH 3.6; Use of the Difference Diagonal Plot

The ANS binding and subsequent perturbation of the protein structure may be conveniently examined by means of a difference diagonal plot (DDP) (72,73). This plot is an extension of the well-known α-carbon diagonal distance plot (74-76) and is a way of expressing a difference electron density in a diagonal distance plot format. An NxN matrix is formed with elements r_{ij} that are the sum of moduli of vectors given by

$$r(k)_{ij} = |\vec{C}_i - \vec{\delta}_k| + |\vec{C}_j - \vec{\delta}_k|, \qquad (15)$$

where $\vec{C}_i$, $\vec{C}_j$ and $\vec{\delta}_k$ are coordinate vectors of the i^{th} and j^{th} α-carbon atoms of α-CHT and the k^{th} peak of the difference density, respectively, and N is the number of amino acid residues of α-CHT. This is then contoured at

some convenient distance or distances. Moreover, since $r_{ij} = r_{ji}$, only half of the DDP is unique so that the entire plot may be used to represent the changes occurring in both molecules of α-CHT dimer by plotting the respective interactions above (molecule I', Figure 12) and below (molecule I, Figure 12) the diagonal. We have found that 9.0Å was a meaningful lower limit for representing the r_{ij} of protein structural changes attending ANS binding but that a level of 12Å was better to represent the ANS molecule contacts since the latter binds on the surface of the protein.

Early in this study it was recognized that information may be lost in the DDP when several difference peaks satisfy $r_{ij} < 9$Å or generate the same ij block (Figure 12). The number of peaks generating a given block will thus be referred to as the "multiplicity" of the block. The DDP calculation was modified to produce a table listing the multiplicity and all of its contributing difference density peaks for each block, as well as to produce a "multiplicity plot" in which the multiplicities were plotted. The multiplicities can be used in identifying regions of the protein which are most susceptible to structural change under varied chemical conditions. They can also be used along with the DDP as a more sensitive index of asymmetry in the α-CHT dimer.

An examination of the multiplicities of the DDP's of

the α-CHT derivatives investigated in this study revealed that only a few of these were greater than or equal to two. They are discussed appropriately in text as they were not numerous enough to warrant presentation of a multiplicity plot in each case. However, the plot can be valuable and has been used in representing structural changes of partially denatured derivatives of α-CHT (73), where the multiplicities were of higher order and more numerous.

The DDP corresponding to contacts made by the ANS substitution and the difference peaks accompanying the substitution is shown in Figure 12 where broken lines are used to enclose the latter.* In an effort to approximate the shape and the volume of the saucer-like difference density of ANS substitution for the DDP calculation, the ANS was represented in terms of eight appropriately spaced points around its perimeter. The coordinates of these are given in Table XIII. In Figure 12, it can be seen that the substitution shows excellent local two-fold symmetry and that the principal ANS-protein interactions (e.g., Cys 1-122, Lys 203 locale and Gln 239') are clearly revealed. Moreover, of the 40 interactions (ij blocks) occurring below the diagonal due to the ANS (molecule I), all but nine are generated by the N-terminal binding

---

*The DDP maps are reproduced from a line printer plot (10 characters per inch, 8 lines per inch). This results in a slight distortion and hence a rectangular shape for the DDP.

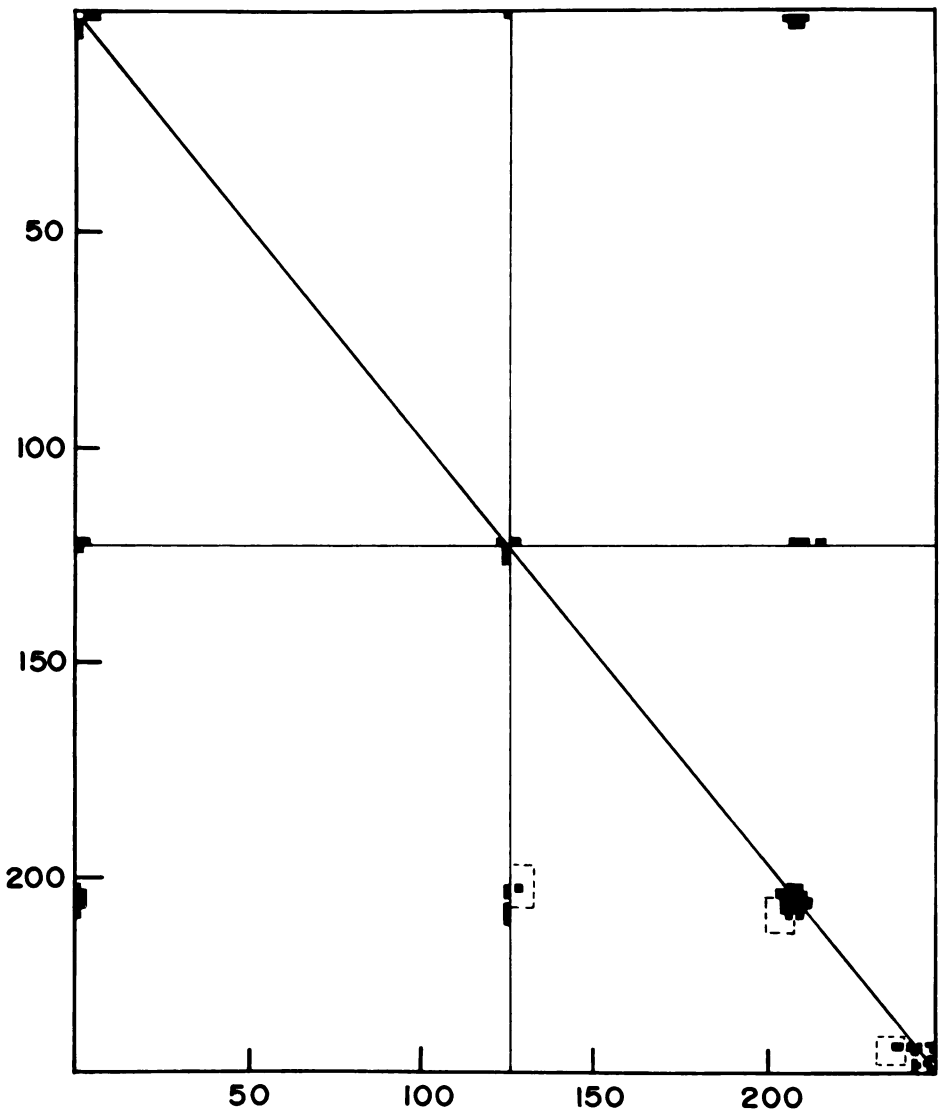


Figure 12. Diagonal plot of difference density between ANS-α-CHT, pH 3.6 and α-CHT, pH 3.6. Substitution proper blocked at $\leq$ 12Å; 8 changes ($|\Delta\rho| > 0.17 e Å^{-3}$) accompanying substitution blocked at $\leq$ 9Å and enclosed by broken lines; interactions with molecule I below and with molecule I' above diagonal.

Table XIII. Coordinates of Points Chosen to Represent the Finite Volume of the ANS Substitution.

Point	x	y	z
		Site 1	
1	0.711	0.208	0.000
2	0.724	0.179	-0.015
3	0.750	0.198	-0.015
4	0.842	0.189	-0.023
5	0.776	0.198	-0.023
6	0.737	0.274	-0.038
7	0.750	0.255	-0.015
8	0.711	0.245	0.000
		Site 1'	
1	0.711	0.396	-0.031
2	0.737	0.415	-0.015
3	0.829	0.415	0.023
4	0.789	0.377	0.023
5	0.750	0.330	0.015
6	0.763	0.406	0.000
7	0.737	0.330	-0.015
8	0.711	0.349	-0.031

environment; with molecule I', the corresponding numbers are 34 and five. The nine and five-block features are associated with the α-helical C-terminal region of the enzyme. Thus, it is clear that, in addition to being generally closer, the ANS maintains a more extensive contact with the protein in the disulfide region and that contact with the helical region is relatively minor and simply represents an intermolecular contact arising from packing of molecules in the crystal.

The ANS difference map contained 8 small peaks in the vicinity of binding in addition to the substitution itself. These ranged from 0.17 eA^{-3} ($\sim 3\sigma$) to 0.25 eA^{-3}, and are listed in Table XIV. Their effect on the DDP calculation was to give rise to those features enclosed by broken lines in Figure 12. These features occur exclusively in cylinder 2 (70) or the B folding domain (sequence > 122) of α-CHT (72). Moreover, since they are close to the diagonal of the DDP, they are necessarily localized near these particular main chain positions. As can be seen in Figure 12, only three regions are involved (near Pro 124, Lys 203; Lys 203-Asn 204; and Gln 239), and all are located in the vicinity of the ANS binding site. These features also show a complete lack of local 2-fold symmetry, but this may be partially due to their nature as side-chain structural perturbations which occur closest to atoms not directly accounted for by the DDP. Thus, a large gradient

Table XIV. Difference Electron Density Peaks $\geq |0.17e\text{Å}^{-3}|$ in the Vicinity of ANS Binding, Other Than the ANS Substitution.*

Peak		Peak Height	x	y	z
1.	a.	$0.18e\text{Å}^{-3}$	0.632	0.387	0.012
	b.**	0.23	0.645	0.377	0.008
	c.	0.25	0.658	0.377	0.012
	d.	0.23	0.671	0.377	0.015
2.	a.	-0.20	0.618	0.368	-0.031
	b.	-0.24	0.632	0.368	-0.031
	c.	-0.21	0.645	0.373	-0.027
3.	a.	-0.18	0.618	0.132	0.023
	b.	-0.22	0.632	0.132	0.023
	c.	-0.18	0.645	0.127	0.027
4.	a.	0.19	0.645	0.217	-0.062
	b.	0.21	0.658	0.217	-0.062
5.	a.	0.19	0.605	0.217	0.008
6.	a.	-0.18	0.724	0.212	-0.069
7.	a.	-0.18	0.868	0.193	-0.046
	b.	-0.18	0.882	0.193	-0.050
8.	a.	-0.18	0.645	0.203	-0.023

*Coordinate system as defined in Table X.

**Extended peaks are represented in three-dimensions by several point densities.

such as is found at the side-chain of Gln 239 on the surface of the protein may just satisfy the r_{ij} distance condition in one of the α-CHT monomers, but not in the other. This is precisely the case observed here. The asymmetric feature in the Lys 203 region is generated by two peaks individually and in combination with two of the resulting blocks possessing a multiplicity of 2. Due to the Asn 204 position being at the pivot of a hair-pin turn in the C-chain, wherein the carbon atoms of consecutive amino acids are more closely spaced, it is not unusual for the two difference peaks to record 16 interactions. On the other hand, the difference peaks at Lys 203' did not reach the significance level.

The foregoing changes correspond to small movement in the sidechain Lys 203 and the peptide chain between Lys 203-Asn 204. The movement of Gln 239' away from the ANS alters the environment of Lys 203, which in turn moves toward Asn 204, thus causing it and the main chain in that region to undergo small shifts in position ($\sim$1.0Å).

A compact, approximately spherical difference density was observed with a peak height of 0.25 $e\text{Å}^{-3}$, and with coordinates virtually identical to those of UO_2^{+2} substitution. The binding of the UO_2^{+2} to α-CHT was characterized during heavy atom multiple isomorphous replacement (MIR) phase determination of the native enzyme (59). The 0.25 $e\text{Å}^{-3}$ feature was present only in the (ANS-α-CHT, pH 3.6 - α-CHT, pH 3.6) "best" difference map. It occupies a region

exhibiting a striking lack of two-fold symmetry that results in one uranyl binding site per asymmetric unit, and which has been observed to manifest a high sensitivity to perturbations of the native protein (73,77). This location is approximately 25Å distant from the ANS binding site 1 (Figure 10). At first, possible sources of metal contamination were considered in order to account for the peak, which existed only in the most isomorphous of the derivatives under study, ANS-α-CHT, pH 3.6; the geometry of the binding site is disrupted above pH 5.4 (77). However, new phases for native α-CHT became available in the later stages of this study and form the basis of the work presented in Chapter V. The new phases were the 2.8Å subset of the 1.8Å α-CHT refinement-extension structure obtained (78) by using electron density modification techniques (79). After convergence, the average phase change for the 2.8Å subset was $\pm 28°$. This is on the order of the expected error of the starting MIR set, which is $\pm 40°$. It should be noted that the new phases are not associated with a figure of merit weighting scheme, the latter of which scales the peak heights in the "best" difference map approximately by a factor of $\langle m^2 \rangle$ (63), or 0.79 in the present study. Thus it was noteworthy when substitution of the new phases into the calculation induced a 25% decrease in the uranyl site difference-peak height. After four cycles of <u>difference</u> electron

density modification phase refinement (Chapter VI), the uranyl site difference-peak height was reduced below background.

These observations underscore the need for care in the interpretation of protein derivative structural changes in this site that are located using the original phase model. Several smaller features were also suppressed by the refinement procedure, but the major features associated directly with ANS binding were enhanced.

Lastly, other difference peaks in the ANS binding region, but on the borderline of significance, would indicate small changes in the location of solvent molecules, some changes being symmetric, some not. Otherwise, no significant changes appear elsewhere in the protein.

5. Changes in the ANS Binding Site at pH 6.6

The difference density corresponding to ANS in the ANS-α-CHT, pH 6.6 map is nearly identical to that observed at pH 3.6 and only a few small changes were observed in the binding environment upon comparison of difference Fourier syntheses of ANS-α-CHT, pH 3.6 with α-CHT, pH 3.6 and ANS-α-CHT, pH 6.6 with α-CHT, pH 3.6.

The previously discussed Gln 239 difference gradient underwent a diminution at pH 6.6. There were also some smaller changes involving the Lys 203 peaks. Since the

α-CHT pH 5.4 conformer (77) falls in the same range of structural stability as that of a pH 6.6 structure without the presence of ANS (35) and since examination of the (α-CHT, pH 5.4 - α-CHT, pH 3.6) "best" difference map showed no significant peaks in the probe binding region, one concludes that it was not the presence of the ANS in conjunction with the higher pH which depressed the Gln 239 changes originally observed at pH 3.6. Moreover, the pK_a values for the key polar side-chains in the binding region are much higher than 6.6, so that their behavior in these derivatives shcuild not be influenced by any significant change in extent of ionization. As for the ANS molecule itself, appreciable zwitterion formation is not observed above pH 3.0 in methanol (29), nor between pH 1.0-2.0 in water (71). Therefore, one concludes that alteration in the appearance of the difference peaks accompanying the pH change involving protein moieties closest to the ANS probably results from a poorer phasing approximation for the higher pH conformer.

Of some interest is a feature of difference density located about 9A from the ANS in the non-polar cavity of environment 2' (Figure 10). This feature resolved into three peaks in a line roughly parallel to the crystallographic y-axis in the (ANS-α-CHT, pH 6.6 - ANS-α-CHT, pH 3.6) (Map 1, Table XV) difference map. A summary of the behavior of the region in various difference maps is

given in Table XV. In two of the maps, where the feature is at a maximum of 0.16 eA^{-3}, it is only slightly above background. From inspection of Table XV, it appears these peaks are a product of both the pH increase and the ANS binding. One possible source of their origin may be a greater accessibility of the cavity to solvent with increased pH. However, even if this were the case, the additional molecules of solvent in this part of the protein would not be close enough to the bound ANS molecule to affect quenching as observed in solution by a mechanism of solvent relaxation. If it were postulated that binding in environment 2' could alter the fluorescence, it would also be assumed that better solvent exclusion is provided at low pH in the crystal than is found in solution. This would by the same reasoning result in a longer fluorescence lifetime in the former medium, which is not observed.

Finally, a number of small difference density peaks are located in the solvent region of the ANS binding site. However, with one exception, they hover around the significance level. The largest is a peak -0.18$e\text{Å}^{-3}$, appearing in map 2, Table XV, and it has neither a positive counterpart nor a local 2-fold equivalent. This feature also appears at background level in map 3. It is situated approximately 2Å distant from C_4' of ANS along the direction of the local 2-fold axis (Figure 11), and may have

Table XV. Peak Heights of Non-Polar Cavity Difference Density in Various Difference Maps.

	Difference Map	Peak Height ($e\text{Å}^{-3}$)
1.	(ANS-α-CHT, pH 6.6 - ANS-α-CHT, pH 3.6)	+0.20*
2.	(ANS-α-CHT, pH 6.6 - α-CHT pH, 5.4)	+0.16**
3.	(ANS-α-CHT, pH 6.6 - α-CHT, pH 3.6)	+0.16**
4.	(ANS-α-CHT, pH 3.6 - α-CHT, pH 3.6)	Background
5.	(α-CHT, pH 5.4 - α-CHT, pH 3.6)	Background

*Three resolved peaks.

**One broad maximum.

been induced by the combination of pH change and ANS binding. This peak is not associated with the Cys 1-NH_3^+ - $SO_4^=$ ion pair, and may signify a small asymmetric solvent loss or readjustment.

In summary, we have observed only very small alterations in or related to the ANS binding site region which accompany a change in pH from 3.6 to 6.6, and which are unique to the protein-probe complex. These do not seem directly related to the fluorescence properties of the probe, either from the point of view of solvent accessibility or of the conformation of the ground state of ANS.

6. Changes in Protein Structure of ANS-α-CHT, pH 6.6

Another difference in structure between ANS-α-CHT, pH 6.6 and α-CHT, pH 5.4 is a peak of -0.18 $e\text{Å}^{-3}$ located near dyad A in a solvent region of the dimer interface. The peak is mid-way between the side chains of Leu 143 and Leu 143' which are separated by about 8Å. A DDP of the most significant peaks ($\geq|0.18\ e\text{Å}^{-3}|$) between ANS-α-CHT, pH 6.6 and α-CHT, pH 3.6 (map 3, Table XV*) is shown in Figure 13, and the peaks themselves are listed in Table XVI. Since the structure of α-CHT, pH 5.4 (77) is an

*This map is practically the same as the superposition of maps 2 and 5 of Table XV.

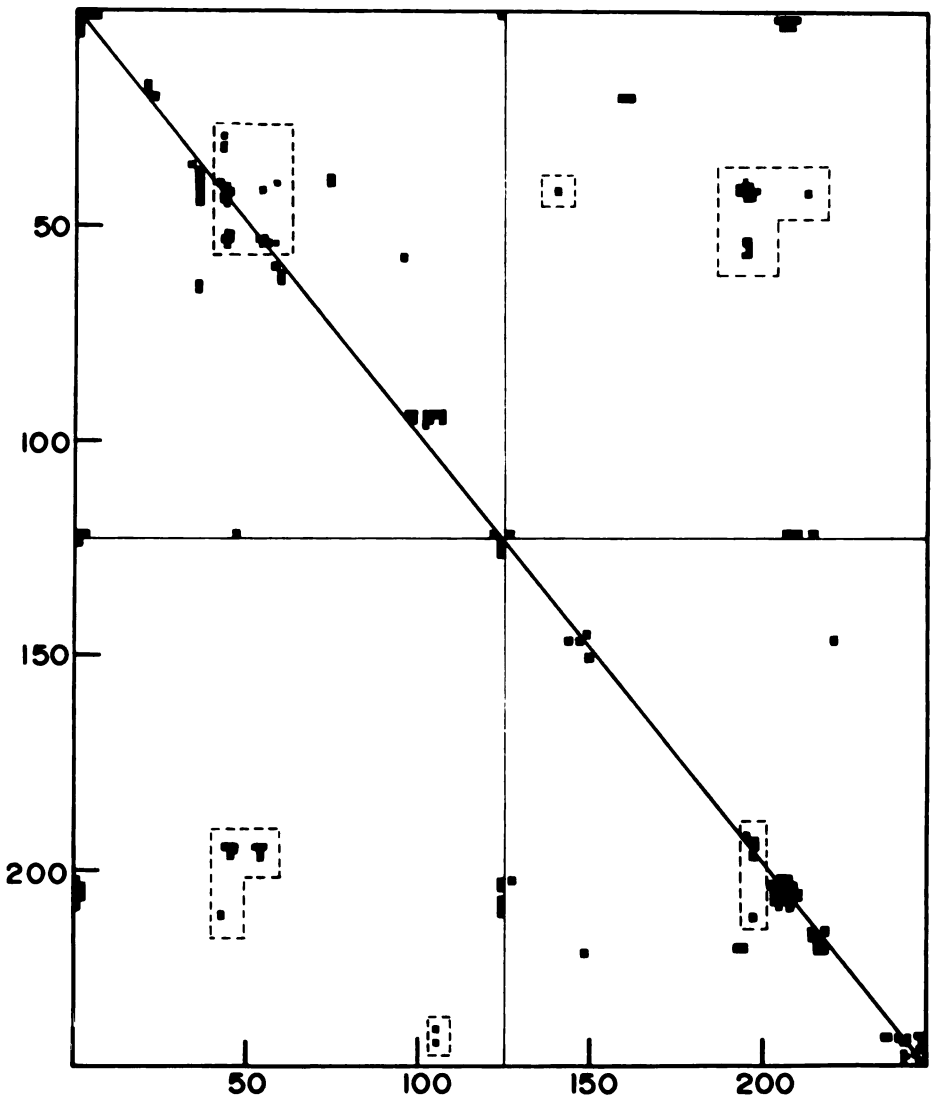


Figure 13. DDP between ANS-α-CHT, pH 6.6 and α-CHT pH 3.6. Substitution proper same as Figure 12 blocked at ≤ 12Å; 29 changes ($|\Delta\rho| > 0.18 e Å^{-3}$) accompanying substitution blocked at ≤ 9Å; interactions not common to pH 5.4 conformer enclosed by broken lines; molecule I below, molecule I' above diagonal.

Table XVI. Difference Electron Density Peaks $\geq |0.18 e\text{Å}^{-3}|$ Representing Protein Structural Changes in ANS-α-CHT, pH 6.6.***

Peak	Peak Height	x	y	z
1.	$0.18 e\text{Å}^{-3}$	0.0	0.300	0.650
2. a.	*	0.053	0.067	0.417
b.****	*	0.066	0.067	0.417
3.	-0.18	0.079	-0.033	0.55
4.	-0.18	0.211	-0.033	0.617
5. a.	*	0.276	0.067	0.367
b.	*	0.289	0.067	0.367
6. a.	-0.18	0.316	-0.083	0.583
b.	-0.18	0.329	-0.083	0.600
7. a.	0.18	0.368	-0.067	0.600
b.	0.18	0.382	-0.067	0.600
8. a.	0.20	0.553	0.267	0.467
b.	0.22	0.566	0.267	0.467
9.	0.18	0.566	0.117	0.633
10.a.	0.18	0.618	0.434	0.400
b.	0.18	0.632	0.434	0.400
11	*	0.671	0.175	0.538
12.	-0.18	0.671	0.417	0.400
13.	*	0.684	0.175	0.542
14.a	*	0.711	0.567	0.633
b.	*	0.724	0.567	0.633
15.	-0.18	0.763	0.333	0.483
16.	-0.18	0.789	0.467	0.383
17	0.22	0.868	0.325	0.600
18.	0.18	0.868	0.083	0.883
19.a.	0.24	0.882	0.317	0.608
b.	0.18	0.895	0.317	0.617
20.	-0.18	0.921	0.333	0.583
21.	*	0.934	0.133	0.583
22.	-0.18	0.934	0.400	0.450
23.	*	0.947	0.133	0.583
24.**	0.18	0.724	0.300	0.133
25.a.	0.20	0.724	0.367	0.133
b.	0.20	0.737	0.367	0.133
26.a.	0.18	0.724	0.467	0.133
b.	0.18	0.737	0.467	0.133
27.a.	0.18	0.776	0.300	0.667
b.	0.20	0.763	0.300	0.667
28.a	*	0.724	0.233	0.667
b.	-0.18	0.737	0.250	0.617
29.	-0.18	0.730	0.300	0.250

*Peak $< -0.20 e\text{Å}^{-3}$. **Peaks 24-29 are unique to Map 1, Table XV. ***Coordinate system as defined in Table X. ****Extended peaks are represented in three-dimensions by several point densities.

excellent model for the structure of α-CHT in the protein-probe complex at pH 6.6, the interactions generated by the ANS binding (map 2, Table XV) are enclosed by broken lines in Figure 13 and specifically represent the interactions of the non-polar cavity peaks, a negative peak near Gly 196, and their local two-fold counterparts. From Figure 13 it can be seen that at pH 6.6, much of the structural change in the protein is asymmetric.

The features enclosed near the diagonal in the A folding domain (sequence < 122) result from contributions of all the difference peaks just described. They represent contacts with a segment of chain from Gln 30-Ala 55. This portion of the backbone forms three strands of beta sheet by winding back on itself twice forming two discrete segments. The junction of both beta strand segments is near Ser 45. The sheet segments run from Gln 30-Ser 45 then form another strand from Ser 45 through Thr 54. This conformation brings consecutive amino acids in this region close together into a structure dependent on interchain hydrogen bonding, the sidechains of which define part of the nonpolar cavity in the Ser 45-Thr 54 region. As an illustration of the size of this nonpolar cavity, and the location of the difference peaks, the largest and most developed positive peak of the triplet in monomer I contributes no interactions to the DDP.

The features fairly symmetrically placed in the

off-diagonal corners result from difference peak contacts with the beta sheet segment just described as it passes near the active site region. The feature in molecule I is generated solely by contacts with the negative difference peak near Gly 196, whereas the one in molecule I' arises from contacts with both positive and negative peaks. The former contains 12 blocks, and the latter 25. Of the 25, 15 arise from the positive peak of molecule I' and 16 from contributions of the negative peak (there are six blocks with a multiplicity of 2). Moreover, the interactions of molecule I' arising from the negative difference peak do not have one-to-one correspondence with molecule I. Thus, approximate symmetry in the DDP can entail a large degree of asymmetry in the detailed steric interaction of a difference peak. In summary, from Figure 13 it appears that the A folding domain of α-CHT is somewhat more sensitive to pH change in the ANS-α-CHT complex than the B domain.

7. Comments on the DDP

Features in the ΔpH DDP may also be used to point out some of the limitations of the technique as presently constituted. Specifically, the alterations in protein structure which accompany the pH change are, in the majority, surface changes, changes involving amino acid side-chains, and changes located in the dimer interface

(77). The DDP calculation is currently restricted to the consideration of α-carbon-difference peak contacts thus limiting its use for summarizing side-chain movements. This is especially so on the surface of the protein where there are necessarily less contacts.* Furthermore, the calculation does not include r_{ij} between the different monomers in the dimer interface region. Therefore, a number of changes, which have been characterized by detailed analysis of the pH 5.4 conformer of α-CHT, are underrepresented in the DDP.

Conversely, the DDP accurately represents any lack of major perturbation in tertiary structure, and the involvement of certain amino acids in regions of important changes. In the latter category, one observes features in the Tyr 146, and Tyr 146', the Ala 149, Asp 64, and Phe 39 regions which have been described and discussed in detail previously (77). A difference peak which generates interactions in the Ser 217-Ser 218 region also generates a block at Ser 217'-Ser 218', thus demonstrating a close contact between monomer polypeptide chains (69). Along similar lines, the DDP calculation between pH conformers (5.4 - 3.6, map 5, Table XV) contains a total of 51 blocks, of which 47 have a multiplicity of one, two

---

*Since solvent does not enter into consideration, its volume is excluded from the calculation thus generally giving rise to fewer r_{ij} satisfying the 9.0Å limit.

of two, and two have a multiplicity of three. All of the higher multiplicity interactions involve Tyr 146 and Tyr 146', which indicates the sensitivity of this carboxylic B chain terminus to pH change in both monomers.

C. Discussion

1. Comparison Between ANS-α-CHT and Solution Studies of ANS Fluorescence

The crystallographic and fluorescence results (33) present several interesting contrasts to solution studies of N-arylaminonaphthalene sulfonates and other fluorescent molecule. For instance, the naphthalene ring system of α-CHT-bound ANS is within van der Waals contact of the Cys 1-122 disulfide, yet ANS fluoresces intensely at low pH. Disulfides are known to quench tryptophan fluorescence in proteins, and have been shown to quench indole emission in solution by collision (80). Further, fluorescence enhancement of the probe-protein complex in solution has been observed to increase continuously with decreasing pH, along with a modest blue shift in ANS emission maximum, to the lowest pH measurement, pH 2.4 (33). Yet fluorescence of ANS alone in solution is unaffected in the pH range we have studied, although quenched below pH 1.5 in water (71) and below pH 3. in anhydrous methanol (29). These contrasting properties are present

in the protein complex even though solvent accessibility to the probe binding site has been amply demonstrated by the accessibility of the site to ANS itself, and by the effectiveness of $PtI_4^=$ quenching.

Although the wavelength dependence of emission decay has not been studied, the excellent fit of a single exponential to the fluorescence decay curve at 490 nm (24) appears to indicate a single emitting state of ANS at pH 3.6 in the protein. Because of the nature of the protein environment that has been observed (which combines charged groups, the disulfide and water accessibility), this would imply either that the relaxation time for local dipoles is very long on the fluorescence time scale, or that specific interaction with the protein dominates fluorescence characteristics of the probe or both. The nature of the pH effect provides further evidence for a specific interaction. Plausibility is also given by the fact that the relatively hydrophobic environment 2' (Figure 10), which is encountered by the probe only in crystals of α-CHT, has had no measureable effect on its emission at low pH. In either case, this result contrasts measurement of 2-anilinonaphthalene-6-sulfonate (2,6-ANS) fluorescence in a viscous medium (glycerol) where more complex decay kinetics were observed (81).

Fluorescence lifetime data, required for direct comparison of the ANS fluorescence properties of ANS-α-CHT

in solution and in crystalline environments at pH 6.6 is presently unavailable. However, the solution (33) and crystallographic studies agree that affinity of the ANS for its binding site is not significantly different at pH 6.6 than it is at pH 3.6. Crystallographically, the region of the protein in which ANS binds at the higher pH was observed to be the same as at the lower. Thus, a quenching mechanism requiring separation of ANS is ruled out, and the high pH ANS-binding region in solution is very likely the one encountered at low pH. Surprisingly, the nature of the ANS-binding region was observed to change very little with pH change in crystals of α-CHT, and no definite structure-linked basis for the fluorescence dependence can be assigned. However, two possible mechanisms for the pH dependence can be proposed, the first involving direct proton-disulfide interaction at low pH, and the second involving the nature of the interaction between the fluorophore and locally ordered solvent. The justification and shortcomings of each are discussed below.

2. Disulfide Protonation Mechanism for Dependence of Fluorescence

The pH dependence of fluorescence in the ANS-α-CHT complex is most dramatic at low pH and a hypothetical proton-moderated quenching interaction between the Cys 1-122 disulfide and ANS would have most sensitivity in

this region. Proton-disulfide interaction is certainly feasible under these conditions. Whereas the covalent radius of sulfur is 1.04Å (44), S-S bond lengths reported crystallographically in organic disulfides are almost invariably shorter than 2.08Å. In tetragonal L-cystine, a S-S bond length of 2.043(6)Å is reported (82) which, with use of a standard formulation (44), indicates a double bond character of about 14%. Raman spectra have indicated a significant increase in S-S stretching frequency as the R-S-S-R dihedral angle approaches 90° (83) and these results have correlated well with CNDO/2 calculations in which S-S double bond character was varied with dihedral angle (85). The C-S-S-C dihedral angle in tetragonal L-cysteine was 74° (82) and that of Cys 1-122 is nearly 90° (Figure 11). Furthermore, a model ascribing partial π-bond character to the S-S bond has been successfully used in HMO calculations to predict inductive effects upon disulfide ionization potentials in alkyl disulfides (85). Thus, the Cys 1-122 disulfide may be expected to undergo a degree of induced polarization in proximity to the N-terminal amino group, which bears a positive charge at acidic pH values, and to experience a corresponding enhancement in the basicity of one or both of its sulfur atoms. Precedent for such an interaction has been set by proton interaction with sulfur atoms such as in OH...S hydrogen bonding, inferred from

a review of crystallographic data (86). Such an interaction has also been found to stabilize the eclipsed gas phase conformation of 2-mercaptoethanol intramolecularly (87). In fact, there is evidence that even CH...S hydrogen bonding may occur under appropriate conditions (88,89).

A recent survey of x-ray protein crystallographic studies (90) described the regions of eight globular proteins in which striking spatially defined "chains" of alternating sulfur and π-bonded atoms (of aromatic amino acid side chains) were found. In all the metaloproteins so categorized (five of the eight), the S-π "chains" were found to include the prosthetic group. For example, in rubredoxin the following S-π "chain" was located:

$$\text{Cys } 9 \xrightarrow{4.8\text{Å}} \text{Tyr } 11 \xrightarrow{3.6\text{Å}} \text{Cys } 39 \xrightarrow{4.2\text{Å}} \text{Phe } 49 \xrightarrow{3.8\text{Å}} \text{Cys } 6,$$

in which Cys 6, Cys 9 and Cys 39 are involved in the Fe-S_4 cluster. It was suggested that these "chains" might serve to help facilitate the flow of electrons. In fact, the quenching mechanism with R-S-S-R compounds (80) has been shown in model compound studies with tyrosine (91) and tryptophan (92) to involve an excited state intermolecular charge transfer from the fluorophore to the sulfur atoms. Even in the ground state, aromatic compounds have been found to interact with disulfides by forming in solution

a weak 1:1 complex (93) with an enthalpy of association on the order of -1 Kcal/mol (90), which is roughly one fourth that usually ascribed to hydrogen bond formation. If this kind of ground state interaction is operative in the binding of ANS to α-CHT it may enhance polarity at the disulfide and help facilitate the protonation. However, an important reservation to the foregoing presents itself in the work of Hoffmann and Hayon (94) who found that model disulfide compounds in aqueous solution react with solvated electrons (from a 2.3 MeV electron source), and subsequently with protons:

$$e^-_{aq} + RSSR \rightarrow (RSSR)^- \qquad (16)$$

$$(RSSR)^- + H^+_{aq} \rightarrow RS\cdot + RSH \ . \qquad (17)$$

Therefore, protonation of the neutral Cys 1-122 disulfide is required to effectively prevent $(RSSR)^-$ formation (i.e., ANS-disulfide charge transfer) in order to moderate quenching. The rate of reaction (16) was found to exhibit sensitivity to the ionization state of amino groups in symmetrical amino acid disulfides, but Hoffmann and Hayon (94) ultimately concluded that the rate of anion formation is largely governed at the actual site of electron attack. The degree and effect of RSSR' protonation could be tested in model compound studies by acid titration of asymmetrical

disulfides (e.g., x-CH_2-S-S-CH_2-CH_2-NH_2, x = CH_3, H, OH); and by examining aromatic-disulfide binding (84) as a function of pH, while simultaneously monitoring $(RSSR')^-$ formation by absorption spectroscopy (92). The latter binding experiment could also be carried out under conditions of aromatic photoexcitation. Spectroscopic detection of $(RSSR')^-$ in ANS-α-CHT might be feasible as well.

We have had further insights concerning the possible local high microviscosity contribution to fluorescence in the protein complex from the 1.8Å resolution electron density map of native α-CHT, pH 3.6 recently computed in our laboratory (78). An examination of the map in the vicinity of the ANS binding indicates a remarkable amount of ordered solvent in the region, in contrast with the 2.8Å resolution map where it appeared more diffuse and at a lower level. A large portion of this density with peak heights exceeding 2 $eÅ^{-3}$ is clearly associated as an apparent solvation shell with the N-terminal amino group and its sulfate counterion. A smaller disembodied portion of electron density is located 4.5Å from the disulfide bond. The ANS phenyl ring partially overlaps this density and displaces some of it. This may account for the difficulty we encountered assigning a rotational conformation to the phenyl ring. More importantly, however, it should be noted that this feature could create a high local viscosity for the ANS.

Proton exchange in ANS itself is thought to quench fluorescence at low pH by a vibrational coupling mechanism, but it is unknown whether this is directly linked to proton-nitrogen interaction per se or the accompanying flexibility of the probe molecule (29). Measurements of the pH dependence of fluorescence of ANS-like molecules in organic solvents of high viscosity are presently unavailable. However, studies of ANS fluorescence in concentrated aqueous $MgCl_2$ solutions, producing a viscous protic medium of high ionic strength, have demonstrated both an increase in quantum yield and a blue shift in ANS emission which increases with salt concentration (29). In view of all the foregoing, our data tend to support the notion that it is a combination of obstruction of the disulfide quenching interaction and a possibly highly viscous environment which are together responsible for the observed fluorescence in ANS-α-CHT at low pH.

3. pH Dependent Solvent Structure Mechanism for Fluorescence Dependence

The foregoing also engenders the basis for an alternately proposed mechanism of quenching regulation. One could postulate that, in the same manner in which an intramolecular charge transfer complex of 2,6-ANS cannot be stabilized and does not form in polar solutions of high viscosity (27), the ANS binding site presents a microviscosity to the probe at low pH which is high enough

to inhibit the likely ANS-disulfide charge transfer quenching mechanism.

Several pH ranges of stability for α-CHT have been observed X-ray crystallographically (35,77). Transitions, between some of these pH conformers take place in the pH range of interest, at pH 3.0, 3.9 and 7.0. These are accompanied by marked unit cell parameter changes. In solution, conformational changes appear to take place more gradually, in a more continuous way, as seen from the pH dependence of dimerization or activation in α-CHT (96-98). Perhaps such changes result in a gradual disordering of water in the ANS binding site, and a reduced microviscosity. Indeed, some small changes in solvent near the ANS are observed in difference maps (Table XV), although they appear near background level. The manner in which and the extent that protein molecules affect surrounding solvent is not fully understood. It appears that protein molecules in aqueous solution are to some degree capable of imposing their own rotational relaxation times on water molecules removed from the protein surface as far as the thickness of several hydration shells (98). Although there was no conformational or ionic change in the ANS binding region within the pH range of this study, it may be possible for pH induced changes in other parts of the α-CHT molecule to have an effect on the binding region's solvent structure.

Although this is an attractive alternative, it too must be considered with reservation. Recent Raman results (99) and energy modeling calculations (100) have indicated that surface point charges and hydrophobic regions of proteins provide the dominating influence in determining the ordering of surface solvent. In this regard it would be useful to firmly establish the differences, if any, which exist between the solution and crystal ANS-α-CHT fluorescence pH dependence. The pH dependence of the intensity of ANS fluorescence in crystals of α-CHT could be examined for discontinuities at pH values of known conformational transitions (35). The fluorescence quenching in solution as a function of pH (24) appears to be independent of the α-CHT dimerization equilibrium, the latter reported by Aune and Timasheff (97).

The range of emission maximum ($\bar{\nu}_F$) encountered for the complexed probe over the limits of our pH measurements (24) is only about two-fifths the relative variation found to be associated with behavior in mixed solvents which induces a variation in quantum yield similar to that we have observed. The actual $\bar{\nu}_F$ values do not seem to be associated with the polar end of the solvent polarity dependence plot (biphasic) where emission maximum has greater sensitivity, and resembles that of a dipole annihilation transition (27). This is consistent

with the view engendered by the crystallographic results that the environment of the probe does not change dramatically from pH 3.6 to 6.6 and that the pH dependence of fluorescence arises from a subtle effect which can block the quenching reaction.

Although it would be speculative to propose specific characteristics for the higher pH 8.8 nm emitting state at this time, the electron donating and withdrawing substituent effect noted for 2,6-ANS fluorescence in mixed solvents (27) can be compared with that expected to result from the present logical alternatives. In comparison with the solution studies, a quenching scheme involving an intramolecular charge transfer process (ictp) in ANS-α-CHT prior to or possibly as a requirement for disulfide quenching would show a greater or equal substitution effect. This would be due to the operation of an additional quenching pathway in the complex as the $S_{ict} \rightarrow S_o$, for example, and disulfide quenching pathways may compete. A quenching scheme involving non-interactive competition for the excited state between an ictp and $(RSSR)^-$ formation would exhibit an equal or lesser substituent effect. If the observed pH dependent fluorescence properties arise in the absence of an ictp, they would be insensitive to substituent. Additional work will be required to determine whether the ANS emitting state of ANS-α-CHT which appears at the higher pH (24) possesses

any intra- or inter-molecular charge transfer character. The chemical relationship of the latter (8.8 ns lifetime state) to the 12.0 ns lifetime state, persistent at both pH values (24), might also be advantageously studied with wavelength dependent lifetime data (28).

4. Conformation of ANS

On the basis of solution studies it has been assumed that a planar conformation in the excited state is a requirement for intramolecular charge transfer in ANS isomers and derivatives (27). NMR experiments conducted in polar and non-polar media have indicated a more planar ANS conformation in the latter (29), in which ANS fluoresces intensely. These results do not necessarily conflict, as the NMR work applies only to the ground state. Thus, the issue of a conformation dependence for ANS fluorescence properties has been the subject of some discussion. The x-ray crystallographic structural evidence, a part of this study discussed in detail in Chapter II, has shown that π-overlap between the phenyl and naphthyl ring systems in ANS does not depend on their coplanarity, and this has suggested a much lesser spectroscopic dependence on conformation. Although models indicate a sterically induced 40° minimum for the dihedral angle between the ring systems in ANS, structural evidence has also indicated a significant distortion of the naphthalene ring in

this and other peri-substituted naphthalenes (Chapter II, 37,39,102) which may allow for the possibility of greater coplanarity in solution media. (The saucer-like shape of the ANS difference density may be a result of such distortion, or may be induced by ground state ANS-disulfide electronic orbital interaction, or both in combination.) In the case of ANS-α-CHT, the $C_{1'}$-N-C_1-C_2 torsion angle reported in Table XII indicates the probe may not assume a highly planar conformation (Chapter II). Whatever the specific nature of the ANS quenching reaction, no measurable effect involving the $C_{1'}$-N-C_1-C_2 torsion angle was observed as the pH was raised to 6.6.

5. Secondary Binding Site of N-formyl Tryptophan

In a previous study of this laboratory (72) it was noted that a single non-active site substitution of N-formyltryptophan (form-Trp) is found in crystals of α-CHT at pH 3.6. It is appropriate to report new details of the substitution at this time. This secondary site is intramolecular and borders that of the ANS site. At one extremity, to which the six-membered portion of the indole ring was assigned, the difference electron density of form-Trp is in contact with dyad B and necessarily has no local 2-fold equivalent.* The indole ring is the

*Such a 2-fold counterpart would make an unacceptably close van der Waals contact with the form-Trp.

most well-defined portion of the difference density, which reaches a maximum peak height of 0.23 $e\text{Å}^{-3}$. The substitution is more interior to monomer I than that of ANS, binding in environment 1 (Figure 10) only, with the indole ring tilted slightly inward from dyad B to the depth of the disulfide. From the perspective of Figure 3, the form-Trp is located to the left of the ANS, at a depth placing it between the disulfide and δ carbon of Lys 203. The non-indole portion of the substitution electron density is less well defined, but it appears that the mean position of the oxygen atoms is in the same general vicinity as the ANS sulfonyl.

The relative peak heights of the difference electron density indicate the form-Trp has a lower affinity for this binding site than has ANS. The fact that the fluorescence lifetime of ANS is not influenced by form-Trp in solution (53) also indicates that the ANS binding constant is much larger, or possibly that the form-Trp site does not exist in solution.

PART II

PROTEIN DERIVATIVE PHASE REFINEMENT BY DIFFERENCE ELECTRON DENSITY MODIFICATION

IV. INTRODUCTION

A. Advantages and Disadvantages in Using the Difference Fourier

X-ray crystallographic determination of protein structure proceeds almost inevitably by the well-characterized method of multiple isomorphous replacement (MIR). With a new structure in hand, the crystallographer - molecular biologist wishes to add to the reservoir of knowledge being accumulated concerning the biochemistry of proteins by making further use of this graphic and frequently definitive tool. Fortunately, it is not necessary to solve derivatives of protein structures each as an independent determination. The standard practice is to utilize difference Fourier maps (Equations (12,13)) in studying substitutions and protein structural perturbations such as those which attend ligand or competitive inhibitor binding as well as chemically induced changes in conformation.

The power of the difference Fourier method is that the same set of phases may be used to approximate any number of related structures, each of which requires the measurement of only a set of structure amplitudes. Further, the difference electron density is extremely sensitive to small structural changes and can reveal more subtle features than those manifest in an electron density

calculated with the same native protein phase angles. Fourier series termination ripples, which contribute to the uncertainty of atomic positions in the protein electron density, are virtually absent in the difference density due to cancellation. The three contributing factors to the rms error level, $\sigma(\Delta\rho)$, in the "best" difference density are (63): 1) underestimation by $|\Delta F_{hk\ell}|$ of the unmeasurable derivative structure amplitudes, $|F^{D}_{hk\ell}|$, 2) the experimental errors in measurement of $|\Delta F_{hk\ell}|$, and 3) the experimental errors in the native phase angles. It can be shown that these sources of error combine in a manner which produces a $\sigma(\Delta\rho)$ that is proportional to $<|\Delta F_{hk\ell}|^{2}>^{1/2}$ (63 , 64), and that the rms error level in the corresponding "best" native protein electron density is proportional to $<|F^{N}_{hk\ell}|>^{1/2}$ (62). In general, $\sigma(\Delta\rho)$ is much lower than $\sigma(\rho)$.

The major shortcoming of difference Fourier analysis is that the maps are limited in non-centrosymmetric cases (such as α-CHT) to peak heights which average approximately one half the value, or less, of that expected if $F^{D}_{hk\ell}$ were known exactly. This was shown by Luzzati (102) who derived expressions for the relative Fourier map peak enhancement and suppression for atoms included and not included (unknown) respectively in the phase model of a partially known structure. The entire difference density is in the latter category. In the

centric case, all phases are constrained to 0 or π and an unknown atom may approach its true peak height. Some of Luzzati's results are illustrated in Figure 14. Independent of this consideration, Henderson and Moffat (63) have shown that if m can be assumed to be the same for all reflections, the "best" difference density will have peak heights of m^2 relative to their values in the absence of errors in the native phases. The present study contains a good approximation to such a condition, as the 2.8Å resolution data sets are those for which $0.7 \leq m \leq 1.0$. Thus, the average peak height in the "best" difference Fourier will have a value of $1/2\langle m^2 \rangle$ relative to the theoretical value. For α-CHT derivative data sets in this study $1/2\langle m^2 \rangle \approx 0.39$.

B. Phase Refinement by Electron Density Modification With a View Toward Difference Electron Density Phase Refinement

Difference densities have been frequently multiplied by a factor of two (103). However, this is only a convenience, which also multiplies the errors by the same factor. In the present study, the prospects of developing an independent method for obtaining protein derivative phase angles have been investigated. Such a set of phases would naturally have the following effects on the derivative difference map:

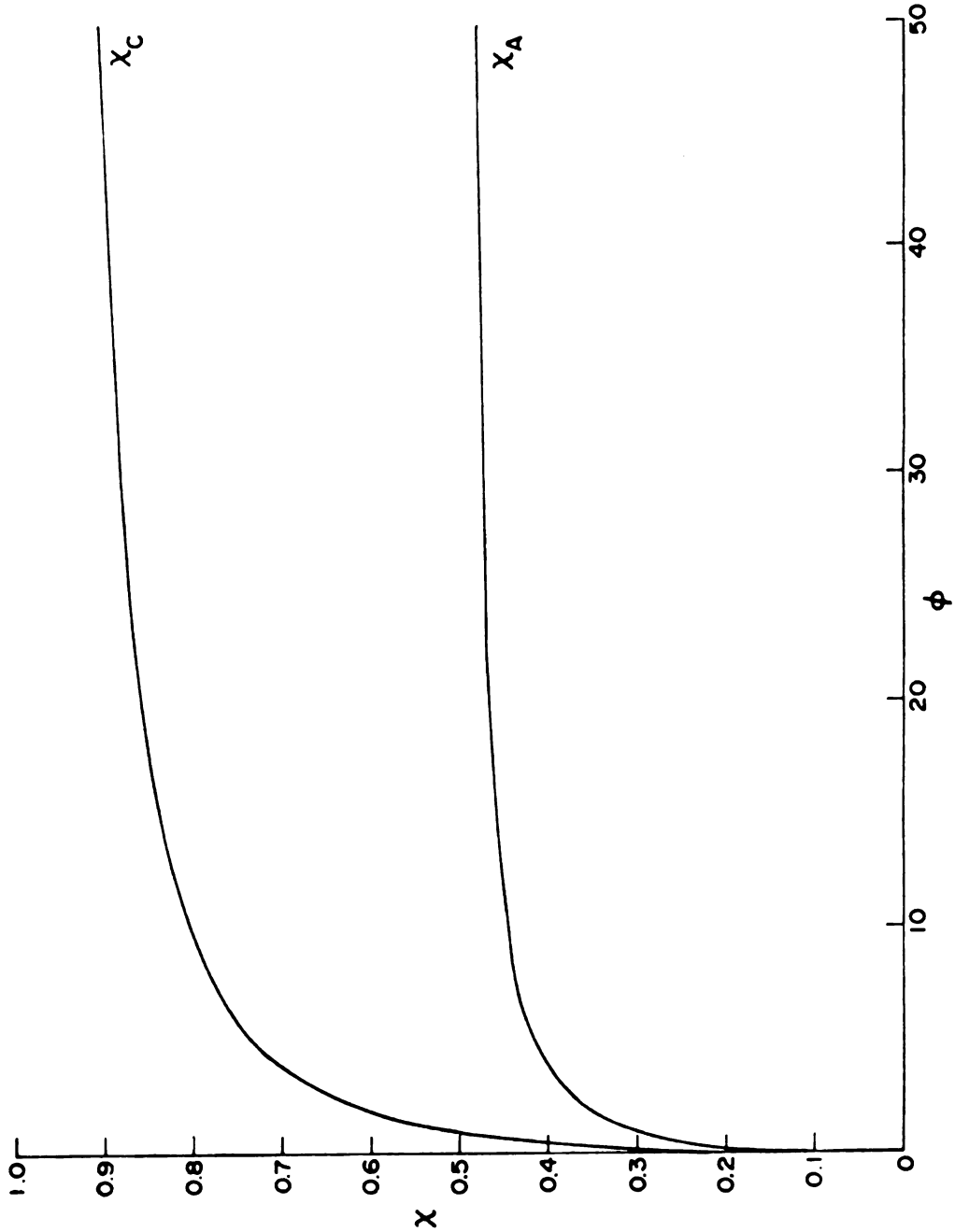


Figure 14. Average relative peak height of unknown atom in a Fourier map, as a function of its relative scattering power, ϕ, in centric case, χ_C; and acentric case χ_A. At $0 \rightarrow \infty$, $\chi_C \rightarrow 1.00$ and $\chi_A \rightarrow 0.5$ (63).

1) Enhance peak heights.
2) Eliminate spurious peaks resulting from artifacts in the native phase set.
3) Improve the signal-to-noise ratio.

As greater sophistication in computer hardware and programming techniques has rapidly evolved in the past years, so has the variety of protein structure refinement techniques. Considering that complicated difference densities frequently require 'interpretation', a method of improvement was sought which would render such maps more reliable without the use of model structures. Least squares and potential energy refinement were avoided in this study, although much work is being conducted along these lines in refining protein structures. Instead, a technique which would treat the electron density or crystallographic data set as a whole was pursued, and the method of choice was electron density modification.

Refinement techniques have been developed which use the principle of density modification and operate in real or reciprocal space. The earliest such method in the latter category was that of Sayre (104), who postulated that the electron density function and its square were very nearly alike in structures of like atoms, and acceptably like in practice for structures containing only C, N, O and H. Bochner and Chandrasekharan had shown that as a general result of Fourier theory the Fourier transform of the convolution of two

functions is the product of their transforms (105). This led Sayre to represent the squared structure as the "self-convolution" of its structure factors and led to the following relationships:

$$\sum_{pqr} F(p,q,r)F(h-p,k-q,\ell-r) = VS(h,k,\ell)F(hk\ell),$$

$$\text{for all } h,k,\ell. \qquad (16)$$

In these equations, the left hand side is the convolution of the structure with itself, V is the volume of the unit cell and S is a function to account for the change in atomic shape on squaring. Sayre later amplified this method in a least squares phase refinement and extension technique for protein structure (106,107). Phase refinement is achieved when (h,k,ℓ) in Equation (16) is within the set (p,q,r), and phase extension when it is not. Not withstanding its promise in protein crystallography, this method possesses what are presently considered prohibitively large computational requirements. However, it is worth noting that it constitutes the first application of direct methods to protein phase extension (106,107).

Barrett and Zwick (108) applied the squaring modification of Sayre (104) in cycles of a real space phase extension of myoglobin, from 3.0Å to 2.0Å resolution. Prior to squaring the electron density they set its

negative regions to zero. This noise elimination procedure cannot be accomplished by reciprocal space methods. The ρ_E map was used in their study instead of the conventional ρ_F in order to better satisfy the equivalence requirement between the normal and squared structure. The ρ_F is the Fourier transform of the $F_{hk\ell}$, and ρ_E is that of the coefficients $E_{hk\ell}$, where

$$E^2_{hk\ell} = \frac{F^2_{hk\ell}}{\sum_j^N f_{j_{hk\ell}}} \quad \text{(triclinic)}, \tag{17}$$

and $f_{j_{hk\ell}}$ is the atomic scattering factor of the j^{th} of N atoms and contains the θ (or (hkℓ)) dependent thermal factor. Thus, the E values depend only on the arrangement and atomic number of the atoms in the unit cell, as the thermal effects have been removed. This makes ρ_E sharper than the corresponding ρ_F. With the myoglobin phases extended and refined to 2Å resolution, the most reliable subset of E values gave $\langle|\alpha_{cal}-\alpha_{MIR}|\rangle = 78°$. Although a random change in phase angle predicts $\langle|\alpha_{cal}-\alpha_{MIR}|\rangle = 90°$, $\rho_{E_{cal}}$ was found to favorably compare with $\rho_{e_{MIR}}$ (108).

One of the difficulties in the application of this method is the problem of calculating E values for use in protein structure refinement. This is generally difficult because a frequently large fraction of the unit

cell volume is composed of solvent. In addition, ρ_E exhibits more severe series termination effects than does ρ_F (108). In the case of difference density refinement, a simple squaring procedure would be inapplicable because the negative peaks are equally important. Although this difficulty can be overcome in real space, the difference density fundamentally violates the central axiom of the squaring method: The structure is to be composed of identical atoms.

Hoppe and Gassmann (109) outlined the criteria by which the electron density of a partially known structure may be judged improved by a modification process:

1) Reduction of the enhanced maxima at known atomic sites.
2) Enhancement of the reduced maxima at unknown atomic sites.
3) Reduction of the reduced maxima at incorrectly placed atomic sites.
4) Reduction of background.

It is clear that 2, 3 and 4 are directly analogous to the goals of difference density improvement of protein derivative structures. A modification or density weighting function developed from a Taylor series was designed by Hoppe and Gassmann (109) to accomplish the four objectives. The form of this modification, in contrast to more intuitively

designed functions, was suited in this way for reciprocal space conversion, since the latter calculations require an analytical relationship between structure factors. In real space the relationship is of the following kind

$$\rho_{MOD} = a\rho + b\rho^2 + c\rho^3, \qquad (18)$$

where ρ and ρ_{MOD} are the observed and modified electron densities respectively, and a, b and c are constants. These terms correspond to the following convolutions (see Equation (16) also) in reciprocal space

$$F_h = aF_h + bS_h \sum_{h'} F_{h'}F_{h-h'} + cS'_h \sum_{h'}\sum_{h''} F_{h''}F_{h'-h''}F_{h-h'}, \qquad (19)$$

where S_h is the scattering vector and h, h' and h" represent the indices (hkℓ), (h'k'ℓ') and (h"k"ℓ") respectively. The procedure is the same in both real and reciprocal space refinement: calculate an 'improved' phase angle set (by inverse Fourier transformation in the former case), and combine the new phases with the experimental structure amplitudes to form F_h for the next cycle. A simple linear modification, designed to promote more rapid convergence in real space calculations was also proposed (110). Our results show that extreme care must be exercised in DDM in order to avoid overly severe modifications.

Ito and Shibuya successfully applied real space difference density modification to refine the lone pair electron distribution in a two-dimensional projection of $NaNO_2$ (111). The modification was accomplished by zeroing large regions of the difference map such that ρ_{MOD} corresponded only to selected peaks which were otherwise not modified. Refinement proceeded by cycles of alternating difference density modification and Fourier inversion, for which the difference density at cycle zero was the free atom X-ray structure (containing electron deformation information in $|F_{obs}|$) minus the neutron spherical atom structure. In each cycle the phase of F_{obs} was corrected by adding F_c to F_{obs}, extracting the new phase and applying it to $|F_{obs}|$ for subsequent calculation of the modified difference map. After ten cycles, excellent agreement was obtained with the difference density of the X-ray structure calculated using sp^2 scattering factors for the N and O atoms, minus the same neutron structure as above. Convergence had been essentially achieved at cycle 5, for which the average phase change was only 15% that of cycle 1. The average change after cycle 10 was 8% that of cycle 1.

This result is encouraging, although such a modification cannot be generally utilized in protein derivative density refinement. This is due to the complexity of protein derivative difference maps. Individual selection

of significant difference peaks would be subjective and difficult. The process of choosing which peaks are significant and which are not would introduce an unacceptable level of bias. However, the studies cited in the preceding discussions have demonstrated that there is considerable freedom in choosing a useful modification function. On the basis of test phase refinement calculations with an artificial 17 carbon atom single-bonded structure, Collins (112) demonstrated the utility of a weighting function applied to the electron density in the equal atom case given by:

$$\rho_{MOD} = \begin{cases} 3\rho_o^2 - 2\rho_o^3; & \rho_o > 0 \\ 0; & \rho_o \leq 0 \end{cases} \qquad (19)$$

In this modification, ρ_o is normalized to a maximum value of one. The relationship is plotted in Figure 15a. However, real structures rarely qualify for even a quasi-equal atom classification. In order to prevent the strongest electron density features from dominating the phases at the expense of lower lying detail, Collins and coworkers (113) proposed a tangential addition to Equation (19), which is shown in Figure 15b. By the choice of the factor χ (Figure 15b), which scales the function of Equation (19), an appropriate decrease in modulation of the higher electron density can be obtained. This latter function, with appropriate values for χ, has been used

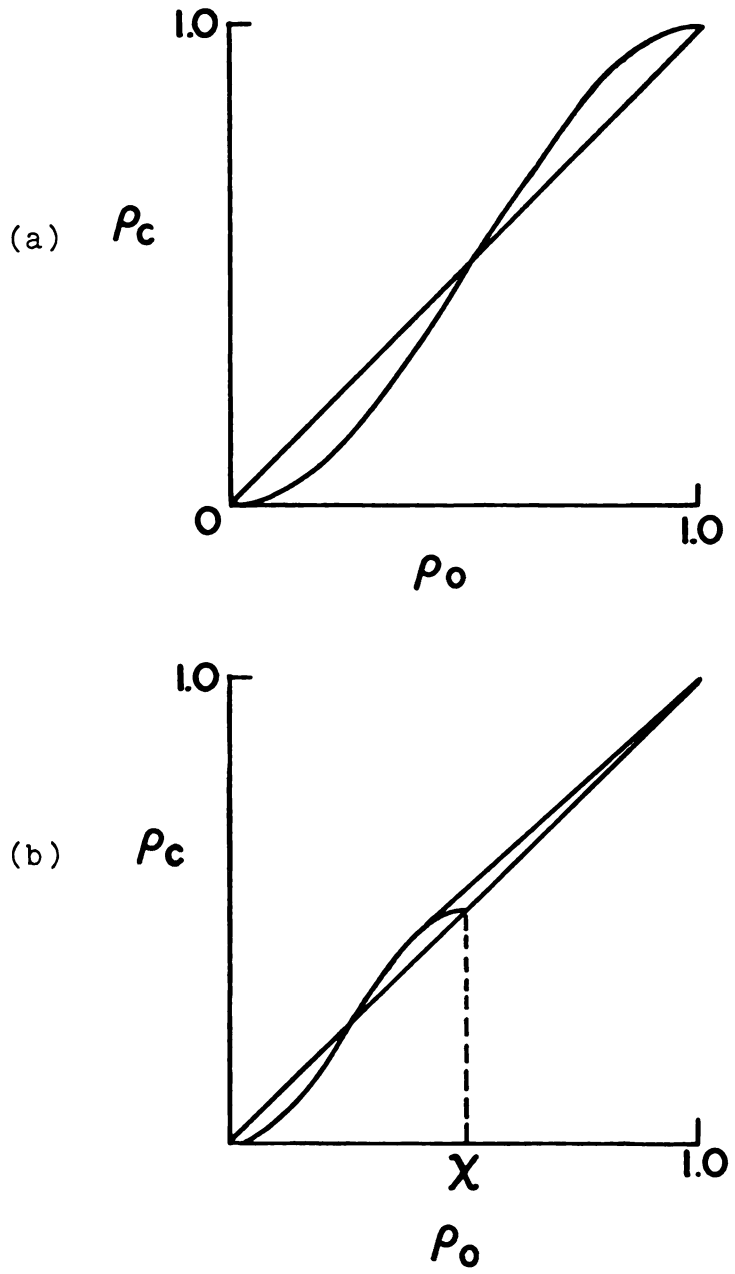


Figure 15. (a) Graphical representation of the function given in Equation (19), after Collins (112). The symbol ρ_o represents the observed electron density, and is normalized to unity. The symbol ρ_c represents the calculated electron density (ρ_{MOD}).

(b) Same as (a), but with tangential addition (113) explained in text.

with success in several protein structural refinements (78,113) and has suggested itself as a suitable modification function to meet the unique requirements of protein difference density refinement.

V. EXPERIMENTAL

A. The Fast Fourier Transform Algorithm

Practicality has been realized in iterative cycles of electron density modification with large structures by application of the Cooley-Tukey (114) fast Fourier transform algorithm (FFT). The economies of the FFT result from the manner in which the Fourier series is factored.

A factorization which is in wide-spread use for conventional calculation of crystallographic Fourier transforms is that of Beevers and Lipson (for example, see (115)). Their method represents a drastic reduction in the number of arithmetic operations required for the evaluation of the three dimensional transform:

$$\rho(xyz) = \frac{2}{V} \sum_{hk} F(hk\ell)\exp[-2\pi i(hx+k\ell y+\ell z)], \quad (20)$$

where Friedel's rule is assumed to operate and the sums are only over the positive indices. In practice, Equations like (20) are evaluated by considering a set of F(hkℓ) belonging to a limited sphere in reciprocal space, and although ρ(x,y,z) is a continuous function, it is calculated at the points of a regular network corresponding to division of the crystallographic directions x, y, and z into N_x, N_y, and N_z equal increments respectively. A 'trivial' evaluation of Equation (20) would involve $N_x \cdot N_y \cdot N_z \cdot NF$

operations, where NF is the size of the set F(hkℓ), and one operation is defined as a complex product followed by a complex sum. For example, if a grid of 76 x 100 x 100 were chosen for a map of an α-CHT derivative associated with a data set of 6300 reflections, a total of 4.8×10^9 operations would be required. However, the Beevers-Lipson factorization results in three one-dimensional transforms, evaluated serially (115,116):

$$T(h,k,z) = \sum_{\ell} F(hk\ell)\exp(-2\pi i\ell z) \qquad (21)$$

$$U(x,k,z) = \sum_{h} T(hkz)\exp(-2\pi ihx) \qquad (22)$$

$$R(x,y,z) = \sum_{k} U(xkz)\exp(-2\pi iky), \qquad (23)$$

where $\rho(x,y,z) = \frac{2}{V} R(x,y,z)$. The sums are only over positive indices, and the only condition is that Friedel's rule holds (115). Equations (21), (22) and (23) each represent a one-dimensional transform for every h and k value, k and z value, and every x and z value, respectively. The number of operations is determined by $(N_kN_x)(N_\ell N_z + (N_kN_z)(N_hN_x) + (N_zN_x)(N_kN_y)$. Since in practical cases $N_h < N_x$, $N_k < N_y$ and $N_\ell < N_z$, this will always be less than $(N_xN_yN_z)(N_x + N_y + N_z)$ which for the example is 2.1×10^8. Thus, an improvement by a factor of 20 is achieved.

After completion of each summation over ℓ in Equation (21), those values of F(hkℓ) no longer require storage

and are replaced by T(hkz). Such a replacement operates through each succeeding step and results in a major reduction of computer storage requirements. In 'trivial' calculations, with Equation (20), the entire sets of $F(hk\ell)$ and $\rho(x,y,z)$ need to be stored at all times. Finally, both the 'trivial' method and the Beevers-Lipson method benefit from fewer operations and lower storage requirements when additional crystallographic symmetry relates classes of $F(hk\ell)$.

A further factorization of the Beevers-Lipson type of one-dimensional transform is utilized by the Cooley-Tukey FFT. By way of example, one of the elementary one-dimensional transforms calculated in Equation (21) can be represented as

$$T(z) = \sum_{\ell=0}^{N_\ell - 1} F(\ell)\exp(-2\pi i \ell z). \qquad (24)$$

The number of operations required is $N_\ell N_z < N_z^2$. Note that in the case of the inverse transform, $\rho \rightarrow F$, the $F(\ell)$ can be replaced with $\rho(z)$, the exponential with its complex conjugate, and T(z) with $T(\ell)$. Defining $j \equiv zN_z$, let us re-write equation (24) in the Cooley-Tukey notation:

$$T(z) = \sum_{\ell=0}^{N-1} A(\ell)W^{j\ell} \qquad (25)$$

where $A(\ell) = F(\ell)$ and $W = \exp(-2\pi i/N_z)$. If N_z is equal to a power of two, j and ℓ can be represented by binary numbers. For the general case where $N_z = 2^M$, we can write

$$j = j_{M-1}2^{M-1} + j_{M-2}\,2^{M-2} + \ldots + j_1 2 + j_0 \qquad (26)$$

$$\ell = \ell_{M-1}2^{M-1} + \ell_{M-2}2^{M-2} + \ldots + \ell_1 2 + \ell_0. \qquad (27)$$

Equation (25) can now be re-written:

$$T(j)=T(j_{M=1},j_{M-2},\ldots,j_0)=\sum_{\ell_0=0}^{1}\sum_{\ell_1=0}^{1}\cdots\sum_{\ell_{M-1}=0}^{1}A(\ell_{M-1},\ell_{M-2},\ldots,\ell_0)x$$

$$W^{(j_{M-1}2^{M-1}+j_{M-2}2^{M-2}+\ldots+j_0)(\ell_{M-1}2^{M-1}+\ell_{M-2}2^{M-2}+\ldots+\ell_0)}. \qquad (28)$$

It can be shown (117) that the following recursive relationships result:

$$A_1(j_0,\ell_{M-2},\ell_{M-3},\ldots,\ell_0) =$$

$$\sum_{\ell_{M-1}=0}^{1}A(\ell_{M-2},\ell_{M-1}\cdots\ell_0)W^{j_0\ell_{M-1}2^{M-1}} \qquad (29)$$

$$A_2(j_0,j_1,\ell_{M-3},\ldots,\ell_0) =$$

$$\sum_{\ell_{M-2}=0}^{1}A_1(j_0,\ell_{M-2},\ell_{M-3},\ldots,\ell_0)W^{(j_1 2^1+j_0)\ell_{M-2}2^{M-2}} \qquad (30)$$

$$A_M(j_0,j_1,\ldots,j_{M-1}) =$$

$$\sum_{\ell_0=0}^{1} A_{M-1}(j_0,j_1,\ldots j_{M-2},\ell_0)W^{(j_{M-1}2^{M-1}+\ldots+j_0)\ell_0} . \quad (31)$$

Finally,

$$T(j_{M=1},j_{M-2},\ldots,j_0) = A(j_0,j_1,\ldots,j_{M-1}). \quad (32)$$

A reordering operation is necessary after completion of the calculation due to the bit reversal. This set of recursive equations corresponds to the original Cooley-Tukey FFT formulations and requires on the order of $N_\ell \log_2 N_\ell$ operations for completion, in contrast to N_ℓ^2 for the Beevers-Lipson type transform. For comparison with the example given previously, $100^2 = 10^4$ and $100 \log_2 100 = 664$, or an improvement of about a factor of 15 over the Beevers-Lipson type transform. In fact, the FFT is rapid enough so that when N_ℓ is not an exact power of 2, the extra terms may be added as zeroes without a prohibitive increase in the cost of calculation.

B. Determining the Appropriate Grid Size for FFT Calculations

It was mentioned previously that in practice the choice of N_x, N_y and N_z is such that each is greater than N_h, N_k,

and N_ℓ respectively. The reason for this follows from Shannon's sampling theorem (118): If a function f(t) contains no frequencies higher than ωcps, it is completely determined by giving its ordinates at a series of points spaced 1/2 ω seconds apart. Thus, in order to analytically reproduce a set of $F_{hk\ell}$ in the case $\rho_c = \rho_o$, the conditions $N_x \geq 2N_h$, $N_y \geq 2 N_k$ and $N_z \geq 2N_\ell$ must be satisfied, where N_h, N_k and N_ℓ refer to the maximum value observed for the associated indices.

The sampling requirements associated with non-trivial modifications of ρ_o increase with the degree of the modification. For example, when $\rho_c = \rho_o^2$, the sampling requirement doubles in each direction with respect to Shannon's criterion (119). In the present study some sacrifice in computational efficiency was opted for in favor of maximum reliability and flexibility in the choice of trial modification adjustments. It was desired to eliminate as completely as practical any possible computational artifacts. Thus, even though a milder modification scheme than $\rho_c = \rho_o^2$ was used, assignments of $N_x = 76$, $N_y = N_z = 100$ were made, which are approximately five times the maximum index in the respective directions. Even with such a fine grid, the cost of the FFT calculation was about 1/3 to 1/2 that of conventional Fourier transformation using a coarser grid of size $N_x = 76$, $N_y = N_z = 60$.

As a test of the programs, one cycle of difference

density modification was performed on ANS-α-CHT, pH 3.6, with the trivial modification $\Delta\rho_c = \Delta\rho_o$. The calculated structure factors were identical to the set from which $\Delta\rho_o$ was calculated, and the scale factor between the sets was 1.000. The structure factors calculated for reflections not included in $\Delta\rho_o$ were identically zero.

C. Difference Electron Density Modification Applied to α-CHT Derivatives

The difference electron density modification procedure (DDM) is an iterative sequence of alternating difference density modification and Fourier inversion calculations. The j^{th} cycle consists of the following steps:

$$\Delta\rho_o^{(j-1)} = F\{F_{ND}^{(j-1)} - F_N\}, \tag{33}$$

$$\Delta\rho_c^{(j-1)} = M\Delta\rho_o^{(j-1)}, \tag{34}$$

$$\Delta F_D^{(j)} = F^{-1}(\Delta\rho_c^{(j-1)}), \tag{35}$$

$$\alpha_{ND}^{(j)} = \tan^{-1}\left(\frac{B_N + B_D^{(j)}}{A_N + A_D^{(j)}}\right), \tag{36}$$

$$F_{ND}^{(j)} = |F_{ND}|\exp(i\alpha_{ND}^{(j)}). \tag{37}$$

The difference electron density is calculated in

Equation (33), where F is the FFT. For cycle 1, $F_{ND}^{(0)} = |F_{ND}|\exp(i\alpha_N)$, where $|F_{ND}|$ is the measured protein derivative structure amplitude. In Equation (34), $\Delta\rho_o^{(j-1)}$ is modified by function M, and this is followed by FFT inversion (F^{-1}) in Equation (35). A new phase angle, $\alpha_{ND}^{(j)}$, is obtained for the protein-derivative structure by using Equation (36), and a new protein-derivative structure factor, $F_{ND}^{(j)}$, is formed as in Equation (37). The native phase angles used to calculate A_N, B_N and F_N are the 2.8Å resolution subset of the 1.8Å resolution native α-CHT set obtained by EDM refinement-extension (78). The protein-derivative amplitudes, $|F_{ND}|$, are scaled to those of the native structure, $|F_N|$, as described in Chapter III, A.

The modification function is shown in Figure 16 and is based on the EDM function of Collins and coworkers (112, 113). The latter is shown in Figure 15 and is given by Equation (19). Since the difference density is of interest in the present case, the modification is performed on both positive and negative densities, and an additional background cutoff level, ψ, has been introduced. The X parameter is calculated on a relative basis. Thus its value in absolute units may change from cycle to cycle even if X is unaltered. The ψ parameter is designed for setting to zero those features in the difference map which are at such a low level that they are judged to be almost certainly random noise. However, its value must be assigned

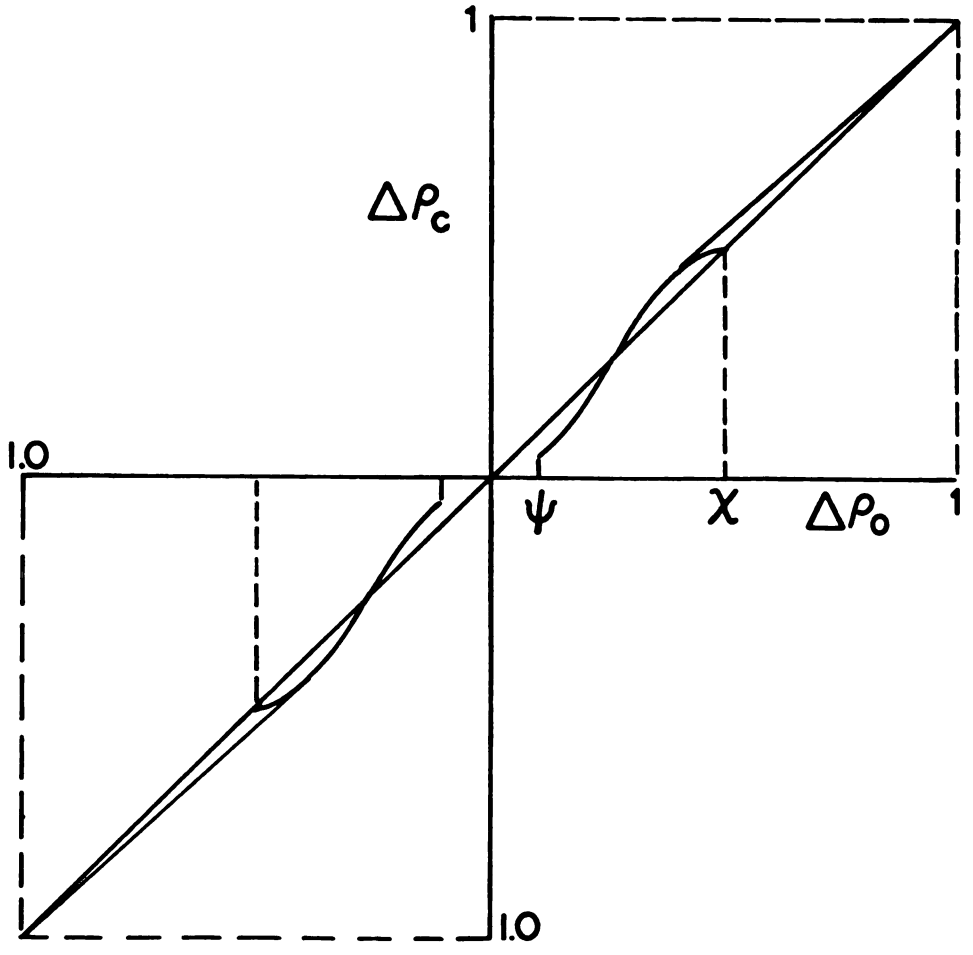


Figure 16. Modification of $\Delta\rho_o$ as applied to protein derivative phase refinement.

with care so that it does not produce too discontinuous of a modification. As pointed out by Collins and co-workers (113), it was Lanczos (120) who observed:"a function which is jumpy and highly irregular will have a Fourier series which is very slowly convergent". Thus, such a discontinuity may cause the need for higher frequency terms in the series, terms which are not provided for by the observed data, and this is tentamount to introducing series termination errors. However, when ψ is applied at low levels, it causes only a small discontinuity in the difference density, which is already close to zero. In practice, it was found that when χ is near 0.5, ψ could be assigned at a level between one and two times $\sigma(\Delta\rho_o^{(0)})$ with apparent safety. In contrast to χ, ψ was evaluated in absolute units so that it would remain constant from cycle to cycle in the absence of external intervention.

The goal of this study was to refine the phases $\alpha_{ND}^{(j)}$. The manner in which they are formed (according to Equation (36)) is shown geometrically in Figure 17. The vectors F_N and $\Delta F_D^{(j)}$ are fixed, as is the amplitude of F_{ND}, so that Equation (36) gives an $\alpha_{ND}^{(j)}$ corresponding to the extrapolation of $|F_{ND}|$ to the "head" of $\Delta F_D^{(j)}$. It is essential to form F_{ND} in this manner or subsequent $\Delta\rho_o^{(j)}$ would merely echo the choice of modification function. As shown in Figure 17, $\Delta\rho_0^{(j-1)}$ (Equation (33)) is the Fourier transform of the coefficients $\Delta F_f^{(j)}$.

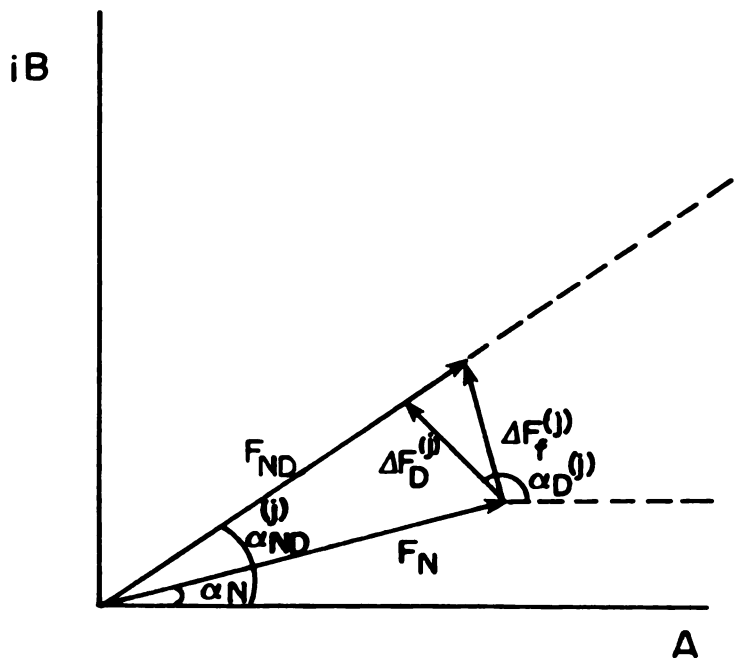


Figure 17. Hypothetical arrangement of F_N, $|F_{ND}|$, $\alpha_{ND}^{(j)}$, and $\Delta F_D^{(j)}$ in phase space. Once $|F_{ND}|$, $\alpha_{ND}^{(j)}$ and F_N are specified, $\Delta F_f^{(j)}$ follows as a geometrical consequence.

When calculating a set of structure factors $\Delta F_D^{(j)}$, according to sequential application of Equations (34) and (35), one must apply a scale factor in order to normalize $\Delta F_D^{(j)}$ relative to the observed data. A scale factor of the form

$$k^{(j)} = \frac{\Sigma(|\Delta F_D^{(j)}||\Delta F_f^{(\ell)}|)}{\Sigma(|\Delta F_D^{(j)}|^2} \qquad (38)$$

was calculated for this purpose during each cycle. In some trials, ℓ was set equal to zero, where $|\Delta F_f^{(0)}| = ||F_{ND}| - |F_N||$, so that changes in the difference maps between cycles would be only due to changes in $\alpha_D^{(j)}$ (Figure 17). With $\ell=0$, $\Delta F_D^{(j)}$ is placed on the same scale in each cycle. In other trials, an additional scale factor, k_{HA}, was subsequently applied to $\Delta F_D^{(j)}$ in order to compensate for the fact that $|\Delta F_f^{(0)}|$ underestimates the true $|\Delta F_D|$. Values for k_{HA} were derived from the observed peak height enhancement occurring with convergence of least squares refinement of heavy atom positions in MIR protein structure determinations (59,121). A k_{HA} was applied only at $\ell=0$ so that $\Delta F_D^{(j)}$ was again placed on the same scale in each cycle. Finally, in yet other trial cycles, ℓ was set to j-1. This allowed the scaling to be responsive to increases of $|\Delta F_f^{(j)}|$ accompanying improved phasing of F_{ND}.

D. Data Processing

All calculations were performed on a CDC 6500 computer. A space group specific FFT program (122) was originally modified by Dr. N. V. Raghavan to perform the native α-CHT EDM extension of phases. The program was further modified as a part of the present study to perform the DDM refinement. Prior to refinement of an α-CHT derivative, the observed derivative data were combined with the native data by use of an efficient SORTMERG program written by Dr. S. R. Ernst of this laboratory.

The data files carried between cycles were divided into blocks of 100 binary records with a record counter as a separate record before each block. Each of the blocked records may be referred to generically as:

h k ℓ AO AN AC BO BN BC FO FN FC IFM

where the suffixes O, N and C refer to observed, native and calculated respectively; F refers to a structure amplitude; IFM is the figure of merit (not included in the other variables); and h, k, ℓ, IFM are to be read as integers. "Observed" in the DDM study refers to $F_{ND}^{(0)}$ "native" to F_N, and "calculated" to $\Delta F_D^{(j)}$. The input file to cycle one (SORTMERG) had only 10 variables per binary record because AC, BC, and FC had to be determined.

In cycle one, in the step indicated by Equation (35),

FO is converted to $||F_{ND}| - |F_N||$, and so remains during subsequent cycles. This enables the user to compare $||F_{ND}| - |F_N||$ and $|\Delta F_D^{(j)}|$ directly in the outputs of the FFT calculations and still retrieve $|F_{ND}|$ from AO and BO. At the user's option, FO can be updated to $|\Delta F_f^{(j)}|$ during each cycle. The FFT calculations do not use FO.

The scale factor in Equation (38) is calculated and applied at the same time as F^{-1}. Several smaller programs have been created to apply an independent scale factor, to calculate refinement statistics, to reformat the data for further iterations (calculate $F_{ND}^{(j)}$) and also for input to the XRAY70 Fourier program. A more detailed description of this software is given in Appendix A. Reflections for which $|\Delta F_D^{(j)}|/|\Delta F_f^{(\ell)}|$ failed to reach a cut-off limit, usually 0.1 to 0.15, were discarded between cycles. This practice never resulted in the loss of more than 5% of the data in a given series of refinement cycles.

D. Trial Refinements

DDM refinements were performed with difference maps of ANS-α-CHT, pH 3.6, ANS-α-CHT, pH 6.6, and α-CHT, pH 5.4. The first is an excellent isomorphous substitution of α-CHT pH 3.6, while the second is a nearly isomorphous substitution of the third. With the first, the effects of DDM refinement could be observed in a derivative

with only small protein structural changes while the effects of larger changes could be observed with the derivatives that possess complex but similar difference maps, and in which the starting phase model is much poorer. Several parameters were inspected from cycle to cycle. These were $k^{(j)}$; R_A defined:

$$R_A = \frac{\Sigma|\,|F_{ND}| - |F_N|\,| - |\Delta F_D^{(j)}|\,|}{\Sigma|\,|F_{ND}| - |F_N|\,|} ; \qquad (38)$$

R_{AS}, when an external scale factor, k_{HA} was applied to $\Delta F_D^{(j)}$ as it appears in Equation (35):

$$R_{AS} = \frac{\Sigma|k_{HA}||F_{ND}| - |F_N|\,| - |\Delta F_D^{(j)}|\,|}{\Sigma\, k_{HA}||F_{ND}| - |F_N|} ; \qquad (39)$$

R_B defined:

$$R_B = \frac{\Sigma|\,|F_{ND}| - |F_N + F_D^{(j)}|\,|}{\Sigma|F_{ND}|} , \qquad (40)$$

which is related to the average lack of closure, l.c.:

$$\text{l.c.} = \frac{\Sigma|\,|F_{ND}| - |F_N + \Delta F_D^{(j)}|\,|}{NR} , \qquad (41)$$

in which NR is the total number of reflections; and the average change in the magnitude of the phase angle $<|\Delta\alpha|>$:

$$\langle|\Delta\alpha|\rangle = \frac{\Sigma|\alpha_{ND}^{(j)} - \alpha_N|}{NR} . \qquad (42)$$

The absolute values of these parameters are not necessarily meaningful in such a refinement (113), but convergence can be assessed by the manner in which they change.

A chronological summary of the DDM refinements is given in Table XVII. Initially convergence was expected to be achieved rapidly, similar to the manner of the native protein structure in EDM refinement at 2.8Å resolution (78). When this did not occur, the independent scale factor k_{HA} was introduced. This was determined by comparing the relative structure amplitudes of heavy atom derivatives, in α-CHT (59) and in KDPG aldolase (121), before (as $\Delta F_f^{(0)}$) and after least squares refinement. A negligible scattering angle dependence was observed and the resultant average ratio in the α-CHT MIR study was 1.73 and that in aldolase was 1.40.

Application of these factors (see Table XVII) resulted in faster convergence and a dramatic enhancement of difference peaks. From a visual inspection of the ANS-α-CHT, pH 3.6 difference maps, it was not at all clear whether a genuine improvement was realized in signal-to-noise ratio or whether it remained unchanged.

An independent method for determining peak significance was employed. One cycle of DDM was performed in which

Table XVII. Chronology of DDM Refinements.

Series Designation	Cycle	χ	ψ	k_{HA}
A ANS-α-CHT, pH 3.6	1	.5	2.5	1.00
	2	.5	2.5	1.00
B ANS-α-CHT, pH 3.6	1	(non-mod as test)		
C ANS-α-CHT, pH 3.6	1	.5	5.0	1.73
	2	.5	5.0	1.73
	3	.5	5.0	1.73
	4	.5	5.0	1.73
D ANS-α-CHT, pH 3.6	1	.5	5.0	1.40
	2	.5	5.0	1.40
	3	.5	5.0	1.40
	4	.5	5.0	1.40
E α-CHT, pH 5.4	1	.5	2.5	1.40
	2	.5	3.0	1.40
	3	.5	4.0	1.40
	4	.5	4.0	1.40
F ANS-α-CHT, pH 3.6	1	(Coor. Zero)		1.40
	2			1.40
G ANS-α-CHT, pH 6.6	1	0	8.0	1.00
	2	0	10.0	1.00
H ANS-α-CHT, pH 6.6	1	.5	2.5	1.00
	2	.5	2.5	1.00
	3	.5	2.5	1.00
	4	.5	2.5	1.00
	5	.5	2.5	1.00

Table XVII. Continued.

Series Designation	Cycle	χ	ψ	k_{HA}
I* ANS-α-CHT, pH 3.6	1	.5	2.5	1.00
	2	.5	2.5	1.00
	3	.5	2.5	1.00
	4	.5	2.5	1.00
J** ANS-α-CHT, pH 3.6	1	.5	0	1.00

*$\ell = j-1$.

**Modified as by Figure 18.

the modification function was a coordinate dependent zeroing of $\Delta\rho_o^{(j-1)}$ so that only the ANS substitution and its immediate environment were preserved. An independent scale factor, $k_{HA} = 1.40$, was applied, as in the immediately preceeding cycles of DDM (Table XVII). A product difference map was formed via an element by element multiplication of the zeroed DDM difference map with that of cycle 1 series D. Chemically significant features in the $\Delta\rho_o^{(1)}$ formed by the zeroing modification should recover a peak height above noise, as in the heavy atom method of structural solution (for example, see (123)). A cut-off was determined by calculating the ratio of ANS peak height increase over four cycles in series D, and multiplying that value times the noise level in $\Delta\rho_o^{(0)}$ (ANS-α-CHT, pH 3.6). Peaks in the product map corresponding to peaks in the cycle 4 map at or below the cut-off level varied unpredictably in height. It was therefore concluded that application of the scale factor, k_{HA}, produced too severe a modification, and this lead to an artifically "improved" difference map.

This example, further experience in which $\chi=0$ and $\psi \approx 4\sigma(\Delta\rho)$ were examined (Table XVII) and the results of Ito and Shibuya (111), in which five refinement cycles were necessary for reasonable convergence with a much more accurate starting phase set (but with a mild modification), have caused the philosophy of this study to evolve toward one of milder modification and more cycling.

VI. RESULTS

The statistics of the DDM refinements are summarized in Table XVIII. It is customary to judge convergence on the basis of an approach by such parameters to a constant and reasonable value. The criteria for reasonability are not always apparent prior to a refinement. For example, the R_A parameter is a measure of the variation of $|\Delta F_D^{(j)}|$ from its values at some reference point, in this case, cycle 0. Thus, although R_A (or R_{AS}) is expected to increase and level off during refinement, its final value depends both on the characteristics of the modification function, and on the manner in which these relate to the difference map features arizing from each data set, $|F_{ND}|$. On the other hand, $k^{(j)}$ can be expected to approach unit with convergence as the modification exhausts itself by suppressing low-lying features to zero after many cycles.

The average figure of merit for the data sets under study is 0.87, which corresponds to an average error of $\pm 30°$ in the α_N phase angles. Thus, an inspection of the progress of $\Delta\alpha$ in the series A refinement caused concern over whether a significant phase change was taking place. An additional scale factor of $k_{HA} = 1.73$ was then applied in series C, and ψ was increased to about $2\sigma(\Delta\rho)$ (Table XVII). The change in ψ had no effect on $k^{(1)}$. Apparent

Table XVIII. Statistical Results of DDM Refinements.

Series Deriv.	Cy	$k^{(j)}(k_{HA})$	$\Delta\alpha$*	$R_{A(s)}$	R_B	l.c.	$\Delta\rho_{max}$**
A. ANS-α-CHT, pH 3.6	1	1.35(1.00)	3.2°	16%	2.2%	2.4e	0.48eÅ^{-3}
	2	1.35(1.00)	5.8	21	2.7	2.9	0.54
B. ANS-α-CHT, pH 3.6	1	1.00(1.00)	($\Delta\rho_c^{(1)} = \Delta\rho_o^{(0)}$ as a test)				
C. ANS-α-CHT, pH 3.6	1	1.35(1.73)	8.9	***	9.5	10.3	0.56
	2	1.45(1.73)	16.7	42	7.7	8.5	0.80
	3.	1.07(1.73)	19.1	49	7.3	7.9	1.00
	4	0.95(1.73)	19.2	53	7.6	8.2	1.14
D. ANS-α-CHT, pH 3.6	1	1.35(1.40)	5.3	43.8	5.5	6.0	0.50
	2	1.79(1.40)	11.9	44.0	6.9	7.4	0.76
	3	1.27(1.40	13.9	51.0	7.1	7.1	0.90
	4	1.13(1.40)	13.9	55.0	7.6	8.3	0.98
E. α-CHT, pH 5.4	1	1.15(1.40)	4.0	13.0	5.6	6.3	0.32
	2	1.18(1.40)	7.7	21.0	5.0	5.6	0.38
	3	1.18(1.40)	11.7	35.0	4.8	5.3	0.48
	4	1.01(1.40)	13.9	40.0	4.4	4.9	0.52
F. ANS-α-CHT, pH 3.6	1	3.37(1.40)	10.6	69.0	13.0	14.3	1.10
	2	1.25(1.40)	10.6	70.0	12.0	13.5	1.20
G. ANS-α-CHT, pH 6.6	1	3.06(1.00)	10.8	52.0	9.1	10.1	0.50
	2	2.08(1.00)	12.7	62.0	11.0	12.3	0.62
H. ANS-α-CHT, pH 6.6	1	1.12(1.00)	3.0	11.0	1.8	2.0	0.40
	2	1.13(1.00)	5.7	13.8	1.95	2.2	NC
	3	1.12(1.00)	7.75	17.8	2.2	2.4	NC
	4	1.10(1.00)	9.3	22.0	2.5	2.8	NC
	5	1.08(1.00)	10.6	26.0	2.8	3.1	0.60

Table XVIII. Continued.

Series Deriv.	Cy	$k^{(j)}(k_{HA})$	$\Delta\alpha$*	$R_{A(s)}$	R_B	l.c.	$\Delta\rho_{max}$**
I. ANS-α-CHT, pH 3.6	1	1.35(1.00)	3.2°	16.0%	2.2%	2.4e	NC
	2	1.37(1.00)	5.8	22.0	2.7	2.9	NC
	3	1.38(1.00)	8.3	28.0	3.1	3.3	NC
	4	1.39(1.00)	10.0	35.0	3.5	3.8	0.68
J. ANS-α-CHT, pH 3.6	1	0.89(1.00)	0.11	0.6	0.0	0.1	0.40

* Equal to $<|\Delta\alpha|>$ in Equation 42.

** Cycle 0 values are: ANS-α-CHT, pH 3.6, $\Delta\rho_{max} = 0.42e\text{Å}^{-3}$
ANS-α-CHT, pH 6.6, $\Delta\rho_{max} = 0.42e\text{Å}^{-3}$
α-CHT, pH 5.4, $\Delta\rho_{max} = 0.24e\text{Å}^{-3}$

*** NC = Not calculated.

convergence resulted after four cycles, and a more significant $\Delta\alpha$ was observed along with an enhancement of the ANS substitution peak by a factor of 2.7 with respect to its $\Delta\rho_o^{(0)}$ value. The $\Delta\rho_o^{(0)}$ did not contain a figure of merit weighting scheme, and its ANS substitution peak height underwent only a 5% enhancement with respect to the "best" difference density.

The resulting difference map calculated with the coefficients $|F_{ND}|\exp(i\alpha_{ND}^{(4)})$ also demonstrated a remarkable enhancement of features previously regarded as below the limit of significance. Their enhancement was most frequently on the order of a factor slightly greater than two. They were features lying just above the ψ cut-off as they appeared in cycle 0. Either these peaks were real and their presence in the refined difference map indicated that the "isomorphous" ANS-α-CHT, pH 3.6 derivative involved much more complex and subtle protein structural changes than was supposed, or their enhancement was an artifact of an overly severe modification scheme which forced peak heights to increase without enhancing the signal to noise ratio. Although the greater enhancement factor of 2.7 had been observed for the ANS substitution peak height, this was somewhat suspect as only a factor of 2.0 had been expected.

Refinement series D and E were then performed with a smaller $k_{HA} = 1.40$. The α-CHT, pH 5.4 difference map had

many more peaks at cycle 0 than that of ANS-α-CHT, pH 3.6. They were of more uniform height in the former, and distributed throughout the asymmetric unit. Although α_N is a more inferior initial phase model for α-CHT, pH 5.4 than for ANS-α-CHT, pH 3.6, the former derivative structure is somewhat more analogous in its density distribution to a native structure, of the kind for which EDM was successfully used previously. Thus, series E was calculated with a derivative containing more difference electron density in the asymmetric unit, in the hope that it might lead to more obvious results. Since the maximum peak height in $\Delta\rho_o^{(0)}$ of α-CHT, pH 5.4 was much lower than that of the ANS-α-CHT, pH 3.6 substitution (Table XVIII), and because this leads to a corresponding lowering of the values in absolute units affected by a given X assignment, the value of ψ was initialized at a lower level in series E.

The final difference map of series D resembled that of series C, although it appeared somewhat less complex, and that of series E was far more complex than either. A total of 107 difference peaks of magnitude greater than $0.30e\text{Å}^{-3}$ were recorded for the final series D difference map. The level $0.30e\text{Å}^{-3}$ is on the order of twice the significance level that was adopted in "best" difference map analysis. In order to establish the noise level in the refined difference map more definitively, the non-ANS

peaks in the cycle 4 difference maps of series D were compared in detail with those of the product map (Chapter V, E) as well as with those resulting from cycle 2 series F, a follow up F^{-1} of the $|F_{ND}|\exp(i\alpha_{ND}^{(1)})$ map of non-ANS region zeroing. The ANS substitution peak height ($\Delta\rho_{max}$) increased by a factor of 2.3 from cycle 0 to cycle 4 of series D (Table XVIII). This multiple of the significance level, $0.17e\text{Å}^{-3}$, in the "best" difference map is approximately $0.40e\text{Å}^{-3}$. The coordinates and heights of peaks greater than or equal to $0.40e\text{Å}^{-3}$ in the cycle 4 series D difference map are listed in Table XIX. Also listed are the heights of corresponding peaks in the product map, and the cycle 2 series F difference map. A similar listing for a number of randomly selected peaks of height less than $|0.40|e\text{Å}^{-3}$ is presented in Table XX. From Tables XIX and XX, it appears that those peaks corresponding to cycle four series D peaks which are at a level of about $|0.40|e\text{Å}^{-3}$ or lower possess a much greater degree of peak height variability - and some of these decrease to what is clearly background. These observations do not support an assertion that the series D refinement produced an enhancement of the signal-to-noise ratio in the difference map, and suggest that $k_{HA} = 1.40$ type of scaling is also too severe. However, they are consistent with the choice of significance level assigned for the "best" difference maps.

Table XIX. Non-Substitution Δρ Peaks, ANS-α-CHT, pH 3.6.

Interpolated Coordinates*			Peak Heights		
X	Y	Z	Cy 4 Series D	(0.1xCy 1 Series DxCy 1 Series F)	Cy 2 Series F
0.724	0.217	0.933	-0.68eÅ^{-3}	-1.22eÅ^{-3}	-0.72eÅ^{-3}**
0.658	0.383	0.017	0.58	1.08	0.60**
0.026	0.267	0.750	-0.52	-0.58	-0.34
0.645	0.383	0.967	-0.52	-0.72	-0.60**
0.645	0.233	0.017	-0.50	-0.44	-0.42
0.316	0.317	0.733	-0.50	-0.26	-0.20
0.618	0.167	0.067	0.46	-0.26	0.22
0.934	0.467	0.450	-0.44	-0.20	-0.18
0.671	0.167	0.417	0.42	0.28	0.28
0.658	0.217	0.933	0.42	0.72	0.56**
0.526	0.283	0.733	-0.42	-0.22	-0.22
0.461	0.417	0.933	-0.42	0.14	<|0.16|****
0.040	0.100	0.750	-0.40	-0.14	<|0.16|
0.461	0.217	0.667	0.40	0.24	0.24
0.632	0.217	0.967	-0.40	-0.48	-0.40
0.336	0.283	0.417	-0.40	<|0.14|***	<|0.16|
0.250	0.317	0.300	0.40	<|0.14|	<|0.16|
0.592	0.433	0.217	-0.40	-0.14	-0.18
0.217	0.450	0.650	-0.40	-0.18	-0.26

*Coordinates system as defined in Table X. **Located in non-zero enclave. ***Values less than |0.14|eÅ^{-3} were not printed. ****Values less than |0.16|eÅ^{-3} were not printed.

Table XX. Non-Substitution Δρ Peaks, ANS-α-CHT, pH 3.6; Random Selection.

Peak Heights		
Cy 4 Series D	0.1 x Cy 1 Series D x Cy 1 Series	Cy 2 Series F
$0.38 e\text{Å}^{-3}$	0.22	0.22
0.38	0.12	<|0.16|*
-0.38	-0.22	-0.18
-0.36	-0.12	<|0.16|
-0.34	-0.14	<|0.16|
-0.34	-0.16	-0.16
-0.32	0.0	<|0.16|
0.30	0.16	0.16
-0.30	0.0	<|0.16|

*Values less than |0.16| were not printed.

Although a relative reduction of noise was not observed, the refinement procedure was not producing merely a scaling up of the map. For instance, not all of the 8 non-ANS peaks in the "best" difference map are equally enhanced by or even survive the refinement procedure. The EDM phase refinement from which the starting α_N were obtained also gave rise to a change in peak heights of cycle 0 which did not appear to correspond simply to the removal of the figure of merit. These effects must result from a variability of the phase changes. Maximum peak heights for the eight features, for which coordinates are given in Table XIV, are listed in Table XXI for the "best", the cycle 0, and the cycle 4 series D difference maps. Even more striking is the behavior of the spherical positive difference peak in ANS-α-CHT, pH 3.6, which appears in the UO_2^{+2} binding site (Chapter III,B4). The UO_2^{+2} substitution was observed during MIR structural solution of α-CHT (59). The spherical peak is observed at $0.25e\text{Å}^{-3}$ in the "best" ANS-α-CHT, pH 3.6 difference map, but decreases to $0.18e\text{Å}^{-3}$ in the cycle 0 map calculation using EDM refined native phases. Although still of significant peak height in cycle 0, subsequent DDM refinement in series D leads to further peak height suppression. In cycle 1, Equation (34), the modification function enhances all density greater than 0.105 and less than $0.42e\text{Å}^{-3}$. Therefore, not only is the refined map not

Table XXI. Peak Heights for the 8 Non-Substitution Features (Table XIII) of ANS-α-CHT, pH 3.6 as They Appear in Two Difference Maps Calculated Using Refined Phases.

	"Best"	Cy 0 Series D	Cy 4 Series D
1.	0.25 $e\text{Å}^{-3}$	0.26$e\text{Å}^{-3}$	0.54$e\text{Å}^{-3}$
2.	-0.24	<-0.18	-0.34
3.	-0.22	<|0.12|*	<|0.16|**
4.	0.20	0.20	0.40
5.	0.19	0.18	0.26
6.	-0.18	<-0.18	-0.34
7.	-0.18	<10.121	<|0.16|
8.	-0.17	-0.16	-0.40

*Values less than |0.12| were not printed.

**Values less than |0.16| were not printed.

simply scaled up but it is also not merely adopting the characteristics of the modification function. The significance of the difference density in the UO_2^{+2} site is called into question. This result was repeated, although somewhat less dramatically, in a later refinement, series I, which employed a relatively mild modification (Table XVII), but for which ℓ of Equation (38) equaled j-1. Following 4 cycles, Series I, the UO_2^{+2} site peak remained at $0.18e\text{Å}^{-3}$ in spite of significant enhancement of other features.

Ito and Shibuya's 'modification' method of difference peak selection (111) cannot be applied to protein DDM without the risk of introducing considerable bias because the features to be 'modified' must be preselected. An effort was made to utilize an analogous technique, but one which was designed to avoid this pitfall by having the modification function itself perform the selection of peaks. In series G the modification consisted of only applying a ψ-like cut-off to the observed difference map, such that ψ was set slightly below the well-established significance level. It was also anticipated that this procedure would constitute a mild but effective DDM. However, it was found to produce an extremely large scale factor, $k^{(j)}$ (Table XVIII), and this indicated that it was modifying the map too severely by suppressing a very large number of low lying points.

All of the foregoing served to engender a philosophy of using milder modifications and more cycling. Such an approach was exercized in series H which employed a χ value of 0.5 and a ψ value of $0.05 e\text{Å}^{-3}$. If χ had been reduced much further, convergence would have been approached so slowly that the procedure would not be practical. Considering the observed sensitivity of the scale factor to large values of ψ, the value of $0.05 e\text{Å}^{-3}$ ($\nu\sigma(\Delta\rho)$) was chosen to ensure the mildest possible modification and to still reject completely a certain level of noise. The results can be qualitatively summarized on the basis of a sampling of the behavior of certain categories of difference peaks. The largest peak, the ANS substitution, underwent an enhancement of a factor of 1.50 after five cycles. The largest peak due to pH change was enhanced by a factor of 1.93. A random selection of several other peaks regarded as background at the level of $0.14 e\text{Å}^{-3}$ in cycle 0 showed an average enhancement by a factor of 1.50.

At this time, another mild modification scheme was also being tested with ANS-α-CHT, pH 3.6, in series I. For this series, ℓ of Equation (38) was set to j-1. Although convergence had not been reached after 4 cycles, a 17% increase in $\langle|\Delta F_D^{(4)}|\rangle$ was observed relative to $||F_{ND}| - |F_N||$ and the large scale factor which resulted from this procedure did not decrease.

Extension of series H and I by additional cycles of refinement to achieve convergence was terminated in the light of observations which implied that the modification function which was being employed may not be optimally suited for DDM. In the EDM refinements reported to date (78,79), the scale factor, $k^{(j)}$, has never been larger than 1.20 and is typically 1.10 in the early cycles. In the present study, a consistently large $k^{(j)}$ has been related to over-modifying and its associated liabilities. Thus, the relatively large $k^{(j)}$ in the majority of the refinements has been a matter of constant concern. It had been noted, for example, by comparison of series G and H, that a dramatic sensitivity in $k^{(j)}$ resulted when ψ exceeded a value such that its effect was independent of the portion of the modification curve (given in Figure 16) which is nearest zero. Further, even with ψ at $0.05 e\text{Å}^{-3}$, the scale factors in the ANS-α-CHT, pH 3.6 series were consistently larger than those involving α-CHT derivatives in the pH 5.4-6.6 range.*

A simple empirical test of the origin of the larger scale factors was devised based on the analytical nature of the FFT^{-1} calculation. This was in the form of a modification which enhanced $\Delta\rho_o^{(j-1)}$ everywhere, and which

---

*In series including $k_{HA} > 1.00$, or for which $\ell = j-1$, no additional effect on $k^{(j)}$ is induced in cycle 1.

is represented in Figure 18. It was constructed by deleting the suppression domains (Figure 16) and substituting a second tangent, from the origin to χ. The modification in Figure 18 corresponds to a χ value of 0.5. One cycle with this modification yielded a value for $k^{(1)}$ of 0.89. In the α-CHT derivatives under study, the ratio $\sigma(\Delta\rho)/\Delta\rho_{max}$ ranged from approximately 0.125 to 0.21. Thus, oin contrast to the EDM studies, the great majority of sampling grid points in the difference electron density are subject to the most severe suppression effects of the modification function (Figure 16). A reduction in χ, which would help alleviate the excessively high scale factors, would slow convergence to an extent which would make the procedure impractical. The function plotted in Figure 18 was designed only for the purpose of testing a pure enhancement that would be mild near zero. Interestingly, although this function was elsewhere comparable to that in Figure 16 for $\chi = 0.5$, it gave rise to a negligible $\Delta\alpha$. These results suggest that the $\Delta\rho$ possesses different modification requirements than does the protein electron density itself.

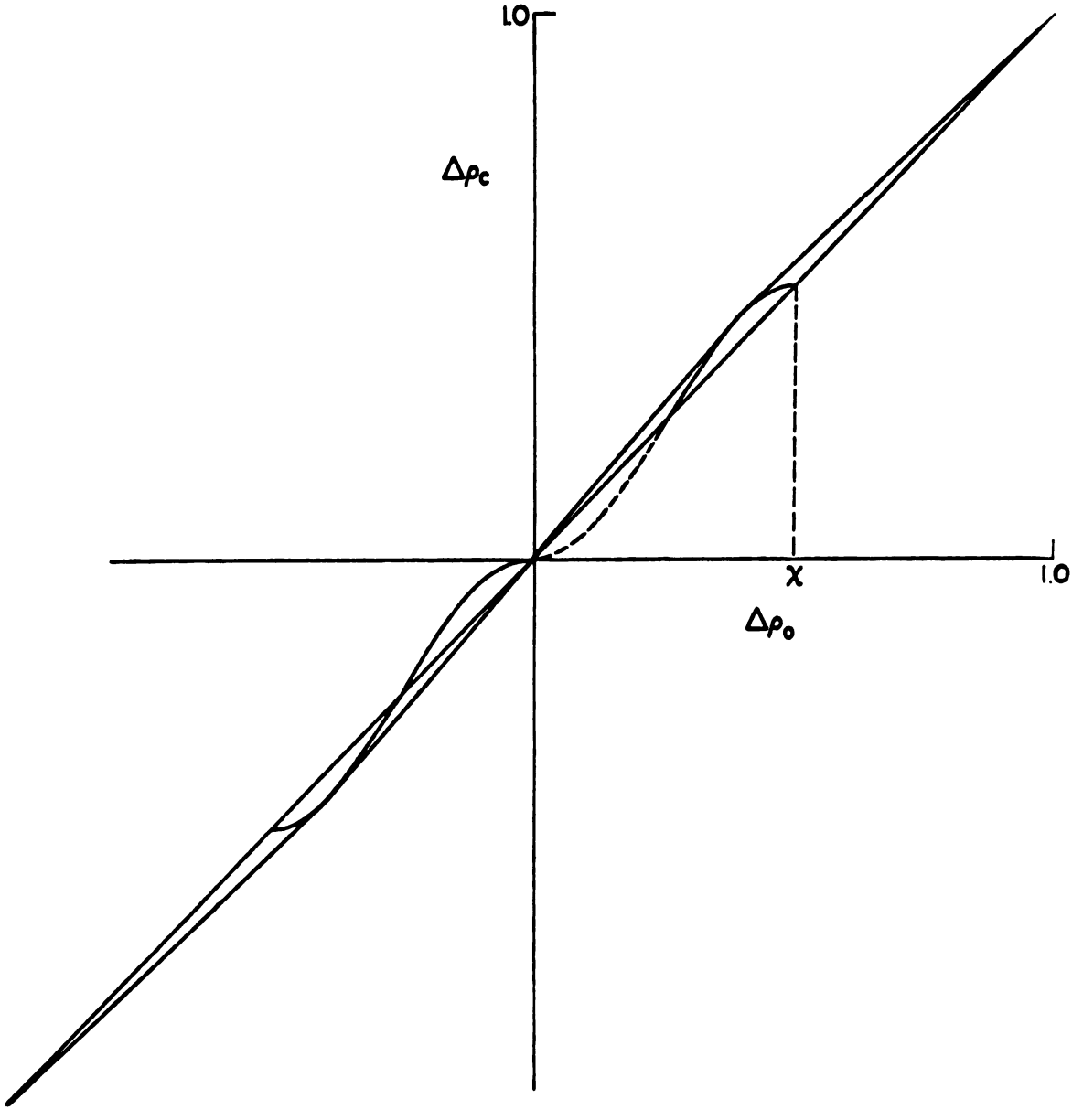


Figure 18. Modification function (tangential line segments) used to provide an exclusively enhancing modification.

IV. DISCUSSION

In principle, electron density modification techniques lead to a correct result because structural information is present in both intensity and phase data. Even if the modified phases were initially to arize in a somewhat arbitrary way, they would ultimately lead to preferential survival or enhancement of correct features by their combination with the measured structure amplitudes. This will favor phase changes which take place in the right directions in subsequent cycles. Thus arises the flexibility in choice of modification function. Judging by the results of Barrett and Zwick (108), it is possible to obtain an essentially accurate result in phase extension of protein structures even when the extended phases appear to be somewhat random. In their work with myoglobin (108), a comparison of the final extended phase set with an MIR set for identical diffraction data gave $\langle|\alpha_{cal}-\alpha_{MIR}|\rangle$ = 78°, where 90° is expected for random results. Even so, the refined electron density compared very favorably with the MIR map (108). This suggests that for a DDM refinement, where the native protein phase model is a good approximation to α_{ND}, but differs from it randomly, noticeable improvement might be obtained by changes in $\alpha_{ND}^{(j)}$ which are merely in the right direction, if not completely accurate in magnitude.

Phase refinement within a fixed resolution data set should produce markedly improved electron densities as a result of relatively small phase corrections. In this kind of experiment the phases are of greater importance in Fourier synthesis than the structure amplitudes. This was shown by Srinivasan (124), who found an impressive agreement between the Fourier transform of the structure factors of L-tyrosine and that calculated by replacing the L-tyrosine structure amplitudes with those derived from an unrelated, arbitrary atomic configuration.

Unlike the EDM procedure, the DDM experiment begins with a data set in which both the phases and structure amplitudes are deficient. Thus, peak height enhancement in the protein derivative difference map occurs not only from phasing, but from the development of $|\Delta F_f^{(j)}|$ as well. Also in contrast with EDM is the fact that the phases which are calculated from the modified difference map are not directly those quantities which are the desired result of refinement. Further, geometrical considerations (Figure 17) dictate an interdependence between the refined phases, $\alpha_{ND}^{(j)}$, and amplitudes $|\Delta F_f^{(j)}|$, which together influence subsequent $\alpha_D^{(j)}$. In DDM refinement a geometrical limit is imposed on $\Delta\alpha$, and this is why a larger $\Delta\alpha$ was observed at convergence in series C than in series D, the latter of which had the smaller k_{HA}.

With application of k_{HA}, all of the $|\Delta F_D^{(j)}|$ are scaled

uniformly regardless of the difference between α_N and α_{ND}. A parallel effect follows in $|\Delta F_f^{(j)}|$ and $\Delta\alpha$. Although the k_{HA} may somewhat incorrectly exaggerate $\Delta\alpha$ from cycle to cycle, the starting α_N are random with respect to α_D in the first place, so that this does not result in a trend toward suppression due to inferior phasing. On the other hand, any improvement of $\alpha_{ND}^{(j)}$ has been diminished. It appears that this may be the origin of the ambiguities reported in the results of series C, D, and E.

An identical effect will be produced when $k^{(j)}$ becomes excessively large. The value of $k^{(j)}$ was shown to be very sensitive to the manner in which the modification function suppressed the difference map features nearest zero, due to the overwhelming preponderance of the latter. Thus, those $|\Delta F_D^{(j)}|$ which were heavily influenced by features of a level that escaped suppression became unrepresentatively scaled by $k^{(j)}$, and this led to a significant peak-height enhancement, independent of phase refinement.

The influence which is exercised on $k^{(j)}$ by k_{HA}, when applied, or by setting $\ell=j-1$, is less straight forward. Frequently, the closer the external scaling is to 1.00, the more persistent or extreme are values of $k^{(j)}$ that are apparently excessive. It seems that this behavior results from reciprocating cycles of suppression and restoration of the low lying features - restoration

by the large cumulative scale on $|\Delta F_D^{(j)}|$ - each capable of undoing the other.

It should be noted that $k^{(j)}$ values which appear excessively large are not automatically damaging provided they result in response to the overall modification rather than primarily to its action on only one kind of peak. Such a response that is originating from background peaks in the difference map probably represents a worst case.

No significant phase change was observed in ANS-α-CHT, pH 3.6, series J, after one cycle of the enhancement-only modification. This was true regardless of the fact that the extent of modification exercised on peaks above the significance level was comparable to that in previous series, and that the procedure generated a $k^{(j)} = 0.89$. The value of R_A, 0.6%, indicated a negligible average variability in the magnitudes of $|\Delta F_D^{(1)}|$ with respect to $|\Delta F_D^{(0)}|$. It appears that the results of the modification were partitioned only into $|\Delta F_D^{(1)}|$. This is probably due to the mildness of the modification near zero, and the relative paucity of grid points near $\chi \approx 0.5$. It means that a relatively mild modification which is all positive or all negative may have an overall effect similar to scaling the difference map. It also indicates that there are not enough features above the significance level in the ANS-α-CHT, pH 3.6 difference map to produce

an effective perturbation of the native protein phases under a modification procedure which uses a function similar to that given in Figure 16. The difference map produced as a result of the enhancement-only modification was nearly identical to that of cycle 0.

In contrast to series J, the modification used in several other series was found to produce significant changes between α_N and α_{ND}. One such series is represented by cycle 5 series H, and a comparison of the differences between α_N and α_D is presented in Appendix B for this cycle and cycle 1 series J.

Two possible changes in approach have been indicated for DDM from this study: 1) Devise a new function which more severely modifies the significant features of the difference map (above $\sim 3\sigma(\Delta\rho)$) and modifies the lowest lying features to a much lesser extent or not at all, and 2) Refine derivatives with more significant 'scattering matter' in the difference map than has ANS-α-CHT, pH 3.6.

It might be advantageous to employ a modification which suppresses part of the map in order to avoid the scaler effect noted in series J. For this purpose, Hoppe and Gassmann's Taylor series-like function, Equation (18), may be found to fit the conditions of approach (1) above, with the threshold for ρ_o suppression, T (109), set between 0 and $\sim\sigma(\Delta\rho)$. In such a function, with $T > 0$,

suppression is still much milder than that shown in Figure 16. The linear function suggested by Hoppe, Gassmann and Zechmeister (110) would probably create an over-modification in the lowest lying features of the difference map if set in such a way to adequately modify the significant features. A double linear function such as that shown in Figure 18, may be of value for work with difference maps that are richer in features than that of ANS-α-CHT, pH 3.6. The severity of this function can be altered by adjusting the nominal value of χ. If the lower placed slope is re-oriented to cut the line described by the function $\Delta\rho_o = \Delta\rho_c$, a mild suppression can be achieved, but this would also entail zeroing those values of $\Delta\rho_o$ between the intersection of the lower slope with the ρ_o axis and zero. The refinement procedure is flexible enough to accommodate a very wide variety of modification functions. With an appropriate M, an $\ell=j-1$ scaling may still be of interest.

One measure of the relative amount of 'scattering matter' in a given difference map is the number of independent significant difference peaks. Where eight such peaks other than the ANS substitution were found in ANS-α-CHT, pH 3.6, there were 29 found in ANS-α-CHT, pH 6.6. An even greater number has been found in denaturant derivatives, where 60 significant peaks were caused by equilibrating α-CHT, pH 3.6 crystals with 2.0M guanidine hydrochloride and over 160 peaks caused by treatment

with 3.0M urea (73). Although the modifications that were employed did not appear to be sensitive to the 'scattering matter' in the ANS-α-CHT, pH 3.6 difference map, including the ANS substitution peaks, other derivative difference maps containing many more significant features are in existance. The ANS-α-CHT, pH 6.6 and α-CHT pH 5.4 derivatives may yet provide definitive refinement results in the absence of over-suppression near zero.

REFERENCES

REFERENCES

1. Weber, G. and Laurence, D. J. R., Proc. Biochem. J., 56, XXXI (1954).

2. Stryer, L., J. Mol. Biol. 13, 482 (1965).

3. McClure, W. O. and Edelman, G. M., Biochem. 5, 1908 (1966).

4. Haugland, R. P. and Stryer, L. Conformation of Biopolymers, ed. G. N. Ramachandran, 1, 321, Academic Press, New York (1967).

5. McClure, W. O. and Edelman, G. M., Biochem., 6, 559 (1967).

6. McClure, W. O. and Edelman, G. M., Biochem. 6, 567 (1967).

7. Cory, R. P., Becker, R. R., Rosenbluth, R. and Isenberg, I., J. Am. Chem. Soc. 90 (1968).

8. Edelman, G. M. and McClure, W. O., Accounts Chem. Res. 1, 65 (1968).

9. Turner, D. C. and Brand, L., Biochem. 7, 3381 (1968).

10. Seliskar, C. J. and Brand, L., J. Am. Chem. Soc. 93, 5405 (1971).

11. Seliskar, C. J. and Brand, L. J. Am. Chem. Soc., 93, 5414 (1971).

12. Brand, L. and Gohlke, J. R., Ann. Rev. Biochem, 41, 843 (1972)

13. Becker, R. S., "Theory and Interpretation of Fluorescence and Phosphorescence", Wiley Interscience, New York (1969).

14. Kosower, E. M., J. Am. Chem. Soc. 80, 3253 (1958).

15. Kosower, E. M., "An Introduction to Physical Organic Chemistry", Wiley, New York, (1968).

16. Reichardt, C. and Dimroth, K., Fortscher. Chem. Forsch. 11, 1 (1968).

17. Dimroth, K. Reichardt, C., Siepmann, T. and Bohlmann, F., Ann. Chem., 661, 1 (1963).

18. Lippert, E., Z. Naturforsch, A10, 541 (1955).

19. Mataga, N., Kaifu, Y., Masao, K., J. Chem. Soc. Jap., 28 (1955).

20. Mataga, N., Kaifu, Y., Masao, K., J. Chem. Soc. Jap. 29, 465 (1956).

21. Greene, F. C., Biochem., 14, 747 (1975).

22. Onsager, L., J. Am. Chem. Soc., 58, 1486 (1936).

23. Garg, S. K., Smyth, C. P., J. Phys. Chem., 69, 1294 (1965).

24. Johnson, J. D., Ph.D. Thesis, Michigan State University, 1976.

25. Seliskar, C. J. and Brand, L., Science, 171, 799 (1971).

26. Kosower, E. M. and Tanizawa, K., Chem. Phys. Lett., 16, 419 (1972).

27. Kosower, E. M., Dodiuk, H., Tanizawa, K., Ottolenghi, M. and Orbach, N., J. Am. Chem. Soc., 97, 2167 (1975).

28. DeToma, R. P., Easter, J. H. and Brand, L., J. Am. Chem. Soc., 98, 5001 (1976).

29. Penzer, G. R., Eur, J. Biochem., 25, 218 (1972).

30. Brand, L., Seliskar, C. J., Turner, D. C. Probes Struct. Funct., 1, 17 (1971).

31. Jackson, G., and Porter, G., Proc. Roy. Soc., A260 13 (1961).

32. Jortner, J., Pure Appl. Chem., 27, 389 (1971).

33. Johnson, J. D., El-Bayoumi, M. A., Weber, L. D., Tulinsky, A., Biochem. (1978), in press.

34. Einarson, R., Eklund, H., Zeppezauer, E., Boiwe, T. and Brändén, C., Eur. J. Biochem., 49, 41 (1974).

35. Mavridis, A., Tulinsky, A. and Liebman, N. M., Biochem., 13, 3661 (1974).

36. Balasubramaniyan, V., Chem. Rev., 66, 567 (1966).

37. Robert, J. B., Sherfinski, J. S., Marsh, R. E., and Roberts, J. D., J. Org. Chem., 39, 1152 (1974).

38. Cody, V. and Hazel, J., J. Med. Chem., 20, 12 (1977).

39. Cody, V. and Hazel, J., Acta. Cryst., B33, 3180 (1977).

40. DeMeulenaer, J. and Tompa, H., Acta. Cryst. 19, 1014 (1965).

41. Cromer, D. T. and Lieberman, D., J. Chem. Phys. 53, 1891 (1970).

42. Wei, K-T and Ward, D. L. Acta. Cryst., B32, 2768 (1976).

43. Germaine, G., Main, P. and Woolfson, M. M., Acta Cryst. A27, 368 (1971).

44. Pauling, L. "The Nature of the Chemical Bond", Cornell Univ. Press, New York (1960).

45. Cruickshank, D. W. J. and Sparks, R. A., Proc. Roy. Soc., A258, 270 (1960).

46. Buss, V. and Försterling, H. D., Tet. 29, 3001 (1973).

47. Einspahr, H. Robert, J.-B., Marsh, R. E. and Roberts, J. D., Acta Cryst. B29, 1611 (1973).

48. Camerman, A., Can. J. Chem., 48, 179 (1970).

49. Wynberg, H., Nieuwpoort, W. C. and Jonkman, H. T., Tet. Lett. 46, 4623 (1973).

50. Camerman, A. and Jensen, L. H., J. Am. Chem. Soc., 92, 4200 (1970).

51. Kunitz, M. and Northrop, J. H., J. Gen. Physiol. 18, 433 (1935).

52. Liebman, M. N., Ph. D. Dissertation, Michigan State University, 1977.

53. King, M. V. Acta Cryst., 7, 601 (1954).

54. Busing, W. R. and Levy, H. A., Acta Cryst., 22, 457 (1967).

55. Vandlen, R. L. and Tulinsky, A., Acta Cryst., B27, 437 (1971).

56. Mavridis, A., Ph.D. Dissertation, Michigan State University, 1976.

57. FACSI Programming Manual (1968). Cat. No. 6294, T55-543, Picker Instruments.

58. Wyckoff, H. W., Doscher, M., Tsernoglu, D., Inagami, T., Johnson, L. N., Hardman, K. D., Allenwell, N. M., Kelly, D. M. and Richards, F. M., J. MOl. Biol., 27, 563 (1967).

59. Tulinsky, A., Mani, N. V., Morimoto, C. N., Vandlen, R. L., Acta. Cryst., B29, 1309 (1973).

60. North, A. C. T., Phillips, D. C. and Mathews, F. S., Acta Cryst., A24, 351 (1968).

61. Blow, D. M. and Crick, F. H. C., Acta Cryst., 12, 794 (1959).

62. Dickerson, R. E., Kendrew, J. C. and Strandberg, B. E., Acta Cryst., 14, 1188 (1961).

63. Henderson, R. and Moffat, J. K., Acta Cryst., B27, 1414 (1971).

64. Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. and Tjian, R., J. Mol. Biol. 88, 349 (1974).

65. Vandlen, R. L. and Tulinsky, A., Biochem., 12, 4193 (1973).

66. Sigler, P. B., Blow, D. M., Matthews and Henderson, R., J. Mol. Biol., 35, 143 (1968).

67. Richards, F. M., J. Mol. Biol., 37, 225 (1968).

68. Tulinsky, A. and Wright, L. H., J. Mol. Biol. 81, 47 (1973).

69. Tulinsky, A., Vandlen, R. L., Morimoto, C. N., Mani, N. V. and Wright, L. H., Biochem., 12, 4185 (1973).

70. Birktoft, J. J. and Blow, D. M., J. Mol. Biol., 68, 187 (1972).

71. Weber, L. D., Tulinsky, A., Johnson, J. D., and El-Bayoumi, M. A., (1978), Biochemistry, in press.

72. Tulinsky, A., Mavridis, I. M., Mann, R. F., J. Biol. Chem., 253, 1074 (1978).

73. Hibbard, L. H. and Tulinsky, A., Biochem., (1978), in press.

74. Nishikawa, K., Ooi, T., Isagai, Y. and Saito, N., J. Phys. Soc. Japan, 32, 1331 (1972).

75. Rossmann, M. G. and Liljas, A., J. Mol. Biol., 85, 177, (1974).

76. Kuntz, I. D., J. Am. Chem. Soc., 97, 4362 (1975).

77. Vandlen, R. L., Tulinsky, A., Biochem., 12, 4193 (1973).

78. Raghavan, N. V., Tulinsky, A., Acta Cryst. (1978) in press.

79. Collins, D. M., Brice, M. D., T. F. LaCour, M. J. Legg, Crystallographic Computing, ed. F. R. Ahmed, K. Huml, B. Sedlacek, Munksgaard, Copenhagen, 1976.

80. Cowgill, R. W., Biochim. Biophys. Acta., 207, 556, (1970).

81. DeToma, R. P. Easter, J. H. and Brand, L., J. Am. Chem. Soc., 98, 5001 (1976).

82. Chaney, M. O., and Steinrauf, L. K., Acta Cryst., B30, 711 (1974).

83. VanWart, H. E., Lewis, A., Scheraga, H. A. and Saeva, F. D., Proc. Nat. Acad. Sci., USA, 70, 2619 (1973).

84. Boyd, D. B., Int. J. Quant. Chem., Symp. Quant. Biol. 1, 13 (1974).

85. Levitt, L. S. and Parkanyi, C., Int. J. Sulfur Chem., 8, 329 (1973).

86. Srinivasan, R. and Chacko, K. K., "Conformation of Biopolymers", G. N. Ramachandran, ed., Academic Press, New York, New York (1967).

87. Sung, E. M. and Harmony, M. D., J. Am. Chem. Soc. 99, 5603 (1977).

88. VanWart, H. E., Shipman, L. L. and Scheraga, H. A., J. Phys. Chem., 79, 1436 (1975).

89. Raghavan, N. V. and Seff, K. Acta Cryst., B33, 386 (1977).

90. Morgan, R. S., Tatsch, C. E., Gushard, R. H., McAdon, J. M. and Warme, P. K., Int. J. Pept. Prot. Res., 11, 209 (1978).

91. Arian, S., Benjamini, M., Feitelson, J. and Stein, G., Photochem. Photobiol. 12, 481 (1970).

92. Bent, D. V. and Hayon, E. J., J. Am. Chem. Soc., 97, 2612 (1975).

93. Bodner, B. L., Jackman, L. M., and Morgan, R. S., Abstracts Annual Meeting Biophysical Society, Biophys. J., 17, 57a (1977).

94. Hoffman, M. Z. and Hayon, E., J. Am. Chem. Soc., 94, 7950 (1972).

95. Stoesz, J. D., Ph.D. Dissertation, University of Minnesota, 1977.

96. Hess, G. P., McConn, J., Ku, E. and McConkey, G., Phil. Trans. Roy. Soc. Lond., B257, 89 (1970).

97. Aune, K. C. and Timasheff, S. N., Biochem., 10, 1609 (1971).

98. Koenig, S. H., Hallenga, K. and Shporer, M., Proc. Nat. Acade. Sci., USA, 72, 2667 (1975).

99. Cavatorta, F., Fontana, M. P. and Vecli, A., J. Chem. Phys., 65, 3635 (1976).

100. Hagler, A. T. and Moult, J., Pept. Proc. Am. Pept. Symp. 5th, 586 (1977).

101. Clough, R. L. and Roberts, J. D., J. Am. Chem. Soc., 98, 1018, (1976).

102. Luzzati, V., Acta Cryst. 6, 142 (1953).

103. Blundell, T. L. and Johnson, L. N., "Protein Crystallography", Academic Press, New York (1976).

104. Sayre, D., Acta. Cryst., 5, 60 (1952).

105. Bochner, S. and Chandrasekaran, K., "Fourier Transforms", Princeton University Press, Princeton (1949).

106. Sayre, D., Acta. Cryst., A28, 210 (1972).

107. Sayre, D., Acta Cryst., A30, 180 (1974).

108. Barrett, A. N. and Zwick, M., Acta Cryst., A27, 6 (1971).

109. Hoppe, W. and Gassmann, J., Acta Cryst. B24, 97 (1968).

110. Hoppe, W., Gassmann, J. and Zechmiester, K., "Crystallographic Computing", Ahmed, F. R., ed., p. 26, Mumksgaard, Copenhagen (1970).

111. Ito, T. and Shibuya, I., Acta Cryst. A31, 71 (1977).

112. Collins, D. M., Acta Cryst., A31, 388 (1975).

113. Collins, D. M., Brice, M. D., laCour, T. F. M. and Legg, M. J., "Crystallographic Computing", Ahmed, F. R., Huml, K., Sledlacek, B., ed., p. 330, Munksgaard, Copenhagen (1976).

114. Cooley, J. W. and Tukey, J. W., Math. Comput. 12, 297 (1965).

115. Stout, G. H. and Jensen, L. H., "X-ray Structure Determination", MacMillan, New York (1976).

116. Immirzi, "Crystallographic Computing", Ahmed, F. R., Huml, K., Sledlacek, B. ed, p. 400, Mumksgaard, Copenhagen (1976).

117. Bergland, G. D., I.E.E.E. Spectrum, 6, 41 (1969).

118. Shannon, C. E., Proc. Inst. Radio. Eng., 37, 10 (1949).

119. Gold, B. and Rader, C. M., "Digital Processing of Signals", McGraw-Hill, New York (1969).

120. Lanczos, C., "Discourse on Fourier Series", Oliver and Boyd, Edinburgh (1966).

121. Mavridis, I. and Tulinsky, A., Biochem., 15, 4410 (1976).

122. Ten Eyck, L. F., Acta Cryst., A29, 183 (1973).

123. Ramachandran, G. N., "Advanced Methods of Crystallography", Ramachandran, G. N. ed., p. 25, Academic Press, New York (1964).

124. Srinivasan, R., Proc. Indian Acad. Sci., A53, 252 (1961).

125. Cruickshank, D. W. J., Acta Cryst. 2, 65 (1949).

126. Cruickshank, D. W. J., Acta Cryst., 3, 72 (1950).

APPENDICES

APPENDIX A

COMPUTATIONAL DETAILS OF DDM REFINEMENT AT MSU

General

Each refinement begins with scaling of the derivative data, and normalizing its average apparent isotropic temperature factor with the native data set of 1977 (2.8Å resolution sub-set of 1.8Å resolution refinement-extension subset). Such a procedure was described in Chapter III, A. The data are then merged with the native structure amplitudes and phases by use of the program SORTMERG. The source listing of SORTMERG is located on CDC APLIB tape 1261, and contains a data card description. The native data are file FORMJ on CDC APLIB tape 7345. It is best to merge the derivative data reduction file as this avoids complications involving the figure of merit, which is not to be included.

In the step indicated by Equation (33), MAPTYP=4 is specified for calculation of the first difference map. An additional code, MAPTYP=7, was added for use in this step during subsequent cycles. Additional data parameters, BCOF and IDFR, have also been added. BCOF specifies the ψ parameters in "mini-map" units (50 mini-map units per $1.0 e\text{Å}^{-3}$); IDFR=0 specifies EDM; IDFR=1 specifies DDM as

in Figure 16; IDFR=2 specifies that ψ will be applied as the only modification in DDM; IDFR $\neq$ 0, 1, or 2 specifies a coordinate dependent zeroing of the map, for which the non-zeroed region is determined in the x direction as possessing a coordinate greater than or equal to IXLOW but smaller than or equal to IXHI. This last criterion is applied analogously in the y (IYLOW, IYHI) and z (IZLOW, IZHI) directions. When IDFR$\neq$0, 1, or 2, diagnostic statements are printed to help the user confirm the correct choice of coordinate limits. In any case, the maximum electron density is taken as the largest absolute magnitude. Dummy values should be provided for parameters not used in a given calculation. The derivative data file input to the F calculation should be retained because it is required in the F^{-1}, along with the modified Fourier map.

The data parameter IDENTL has been added to the F^{-1} program. The user always specifies IDENTL=1 for cycle 1. This causes FO to be replaced with $||F_{ND}| - |F_N||$. In subsequent cycles, when an ℓ=0 scaling (Equation (38)) is desired, the user inputs IDENTL=2. With IDENTL=2, FO is not altered. When an ℓ=j-1 scaling is desired, IDENTL=3 is input and FO is set to FC ($\Delta F_f^{(j-1)}$). This version of the F^{-1} program is called DIFSF3 and is located on CDC APLIB tape 2475. Also on this tape is DIFOUR7, the F program.

Subsequent to F^{-1} calculations, the data are processed

by the following programs. A modular approach was chosen to economically facilitate changes in the DDM procedure during its developmental stages. Ultimately, greater efficiency would result from combining these various functions as options in the existing FFT calculations. This would also conveniently enable calculation of two or more cycles of refinement without user intervention. The programs are written on CDC APLIB tape 2475.

Programs to Process Data Between Cycles

1. Name: 4SCL

   Data Processing Function: Apply a scale constant (k_{HA}) to FC, AC, and BC. Re-write scaled data file on TAPE2.

   Calculate Statistics: <FO>, <FC> (scaled), and <FN> over all reflections for which AN, BN, AC, and BC are not zero; $\langle \alpha_N - \alpha_c \rangle$, and $\langle |\alpha_N - \alpha_c| \rangle$.

   Input Data: a. F^{-1} output file, as TAPE1.

   b. One data Card:

   1. SCL, INDIC (F10.5, I5). The SCL is a scale constant. For INDIC=1, all the data are printed. Only the statistics are printed when INDIC≠1.

2. Name: APT214

   Data Processing Function: For reflections where AO, BO, AN, and BN are not zero, and for which (FC/FO)>0.15,

calculate $\alpha_{ND}^{(j)}$ and update FC, AC, and BC by forming the coefficients $(|F_{ND}|\alpha_{ND}^{(j)} - |F_N|\alpha_N)$. Re-write revised data file to TAPE4 for next cycle of refinement. Write Miller indices and new FC, AC, and BC to TAPE8 for input to the XRAY70 system Fourier program under FORMAT (3I5,3F10.2).

Calculate Statistics: R_A, R_{AS}, R_B, l.c. (see equations (38) to (41)) and $\langle|\alpha_N - \alpha_{ND}^{(j)}\rangle$. APT214 is programmed to calculate R_{AS} based on the scale factor written in cards 67 and 72.

Input Data: F^{-1} output file, or 4SCL output file as TAPE1.

3. Name: ALFDIF

Data Processing Function: None.

Calculate Statistics: See APT214. In addition, ALFDIF calculates a $\sigma(\Delta\rho) = \frac{4}{V}(\Sigma(|F_{ND}| - |F_N + FC|)^2)^{1/2}$ After Cruickshank (125,126).

Input Data: See APT214.

4. Name: FFRS

Data Processing Function: None

Calculate Statistics: Relative and Absolute distributions of FO and FC. Programmed to list FO and FC for which (FC/FO) > 4.5.

Input Data: F^{-1}, 4SCL, or APT214 output files.

5. Name: PRMAPS

Data Processing Function: Calculate product of the Fourier maps, element by element, and output in F^{-1} format. Map grids must be the same. Product elements are divided by 10 and preserve the sign of elements of map read as TAPE1. Product map is written to output, and in binary to TAPE3.

Input Data: a. Two F^{-1} output files (TAPE3 or TAPE4) as TAPE1 and TAPE2.

b. Two data cards

1. ITITLE (20A4)
2. NX, NY, NZ, MCT, NLINES, NCOL (6I5).

The grid dimensions are given by NX, NY, and NZ. A background suppression for the absolute value of positive and negative product map features is given in map units by MCT. The maximum number of lines and columns desired in printing are given by NLINES and NCOL respectively.

APPENDIX B

DETAILS OF CALCULATION OF PHASE CHANGE BETWEEN NATIVE PROTEIN (α_N) AND DERIVATIVE (α_D)

Calculations of the parameter $\Delta\alpha' = \langle \alpha_N - \alpha_D \rangle$ for cycle 1 series J, and cycle 5 series H, gave values of -0.64° and 1.1°, respectively. These data sets were chosen in order to exemplify the two extremes of negligible and substantial α_{ND} refinement. The values were expected to be near zero in both cases due to the random relationship between α_N and α_D. For the same two refinements, calculations of $\alpha'' = \langle |\alpha_N - \alpha_D| \rangle$ yielded 103° and 118°, respectively. The phase angles are processed in the computer programs as values ranging from 0° to 180° and 0° to -180°, according to crystallographic convention.* This is why a random change in phase yields $\Delta\alpha'' = 90°$. The values of $\Delta\alpha''$ which are reported above do not imply that α_N and α_D are correlated. If $|F_{ND}|$ is smaller than $|F_N|$, the negative sign in $\Delta F_f^{(0)}$ is stored by shifting the phases, $\alpha_{ND}^{(0)}$, by 180° with respect to α_N. If the modification is not too severe, the shifted phases that go into the difference map calculation are returned by the F^{-1} very nearly 180° away from α_N. Inspection of the individual

---

*This convention arises because of Friedel's Law, one consequence of which is that $\alpha(hk\ell) = -\alpha(\bar{h}\bar{k}\bar{\ell})$.

($\alpha_N - \alpha_D$) values in cycle 1 series J, revealed that the great majority were within one degree of 0° or 180°. In calculations of $\Delta\alpha'$ and $\Delta\alpha''$, if individual pairs of α_N & α_D are of opposite sign, one of them must be corrected by 360° in order to yield accurate results. The phase differences so calculated will range from 0° to 180° and 0° to -180°, and can have an average magnitude > 90° while $\Delta\alpha' = 0$, and without implying that α_N and α_D are correlated. The effect of a phase change of 180° is the same as that of 0° in calculation of $\alpha_{ND}^{(j)}$ for refinement.

In a derivative such as ANS-α-CHT, pH 3.6, which has few structural changes, many of the $|\alpha_N - \alpha_D|$ values near 180° arise from negative differences ($|F_{ND}| - |F_N|$) which involve structure amplitudes that are equal within the error. For ANS-α-CHT, pH 6.6, the number of significant negative ($|F_{ND}| - |F_N|$) was greater, and so $\Delta\alpha''$ was larger. The relative variation of $\Delta\alpha''$ from 90° between derivatives may serve as an indirect measure of the amount of 'scattering matter' in their difference maps.