STRUCTURAL INVESTIGATIONS OF CREATINE KINASE

Andrew Nigel Burgess BSc.

Glasgow University

Scotland

A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy, at the University of Glasgow, Chemistry Department.

May, 1987

Summary

X-ray crystallographic and other structural studies have been carried out on creatine kinase isolated from rabbit skeletal muscle. As a consequence of the exploration of a wide range of crystallisation conditions, a previously unreported crystal form of creatine kinase has been characterised. This crystal form is monoclinic C2 with three subunits in the asymmetric unit and cell dimensions a = 248 Å, b = 149 Å, c = 52 Å, β = 90°. The notable feature of this new crystal form is the remarkably strong local symmetry displayed extending to a resolution of at least 7 Å. This local symmetry gives rise to an apparent 6-fold axis and mirror planes are observable at 60° intervals in the direction of the principal axes of the pseudo-cell. From the analysis of this diffraction data it is concluded that one third of the dimers lie on a crystallographic two-fold axis and therefore comprise indentical subunits. The remaining two-thirds comprise subunits indistinguishable to a resolution of at least 7 Å.

In addition to this study, heavy atom projection data have been collected for two derivatives of another closely related crystal form of creatine kinase and used to derive some phase information.

Lastly, studies on creatine kinase employing isoelectric focussing have produced the finding that creatine kinase exists as a random pairing of two non-identical subunits. The additional finding that the two subunits can be made

equivalent by the reaction of a reactive thiol essential for catalytic activity implies a special role for this residue.

A.N.Burgess

Acknowledgements

I wish, firstly, to thank the late Dr I.D.A.Swan, for the help and understanding he provided as supervisor throughout the course of the practical work described in this thesis. I wish, also, to thank both Professor G. Sim and Dr L. Sawyer for their encouragement and advice.

To my parents

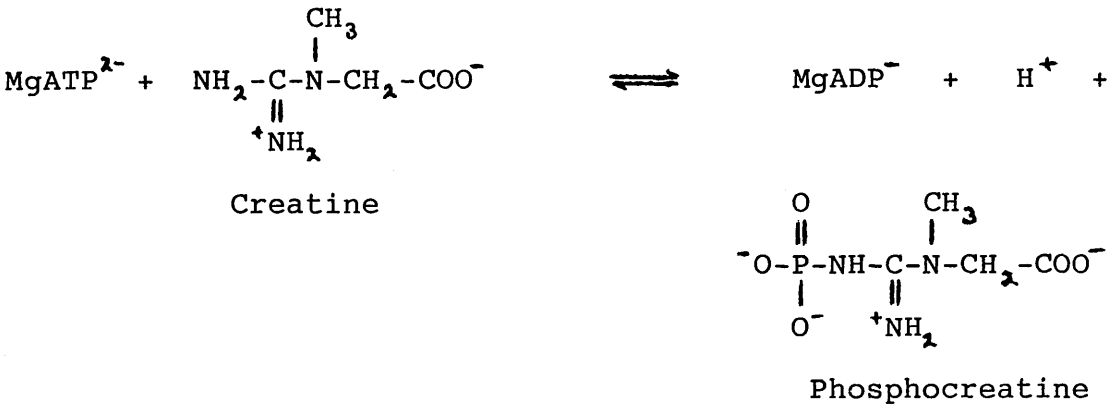
CONTENTS

CHAPTER I

CREATINE KINASE

1.1 Introduction

Creatine kinase (E.C.2.7.3.2), hereafter referred to as CK, catalyses the following reaction :



Creatine

Phosphocreatine

CK and its product phosphocreatine are generally found in tissues with contractile properties or nervous activity (1). CK isolated from vertebrates occurs as three forms distinguishable by their electrophoretic mobility (2,3). It has been shown that the species of intermediate mobility is a hybrid of the other two and that CK is, in fact, a dimer with an active site present in each of its subunits. The three types of CK are designated muscle type (MM), hybrid (MB) and brain type (BB) to indicate their major tissue of origin. The work carried out for this thesis concerns CK isolated from rabbit skeletal muscle and is thus of the MM variety.

Studies have been made of the distribution of the enzyme types during the development of the rat (4). It was shown that in the early stages of development, skeletal muscle, heart and brain tissues all contained only the BB type active enzyme. As development proceeded, skeletal muscle

displayed a continuous shift from BB to MB and finally MM whereas brain tissue kept the BB form throughout. Heart tissue followed the trend displayed by skeletal muscle but kept, in addition, a proportion of the MB hybrid form.
The concentration of CK in the tissue seems to be related to the level of activity and considerable interest has centred on CK levels in cases of muscular dystrophy or atrophy where CK is lost rapidly after the onset of the disease (5).
CK of all types from all species seems to be characterised by the possession of a reactive thiol per subunit essential for catalytic activity (1). It is interesting to note that CK isolated from dystrophic mouse muscle consisted of dissimilar subunits in that only one exhibited a reactive thiol (6).

1.2 Primary Structure

The cDNA derived amino acid sequence of rabbit skeletal muscle CK has recently been reported and is reproduced in Figure 1.1 (7). The calculated molecular weight of the subunit is 42,932 compared to c.41,000 by other methods (1). The sequence is shown to contain four cysteine residues per subunit with residue 282 being identified as the reactive essential thiol. It is interesting to note that this residue is adjacent to a proline residue, the possible significance of which will be discussed in Chapter VIII.

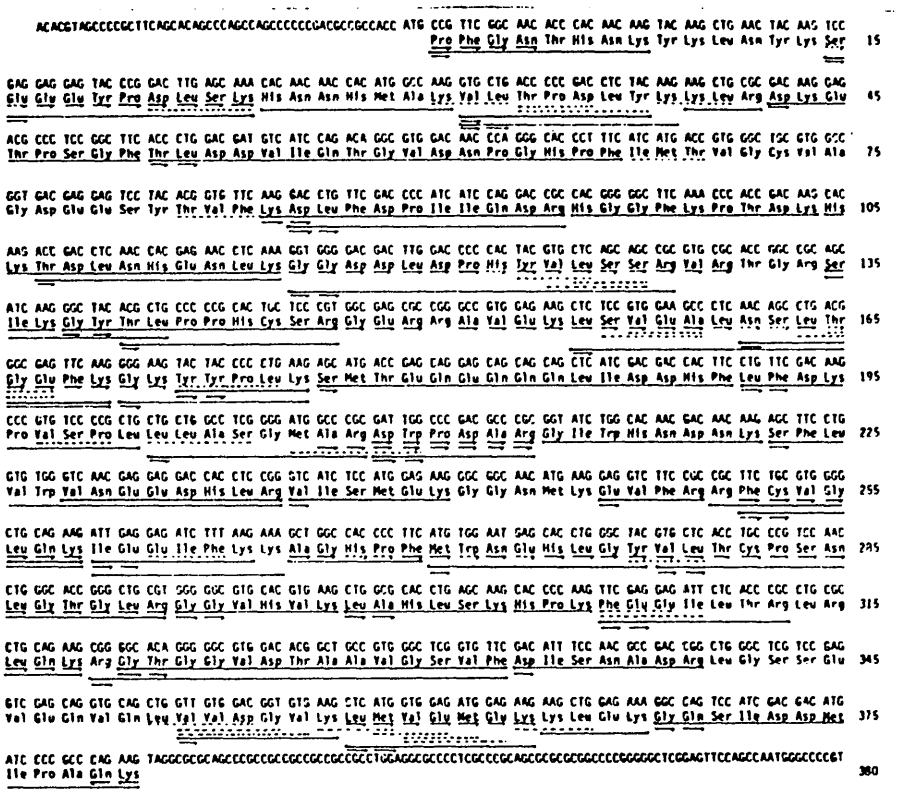


Figure 1.1 : The cDNA derived amino acid sequence of rabbit skeletal muscle creatine kinase reproduced from reference 7.

1.3 Higher Order Structure

A compact globular structure containing 25-30% α-helix and less than 15% β-pleated sheet has been indicated for CK by optical rotary dispersion studies (8). The view of a compact structure is supported by the resistance of the native enzyme to tryptic digestion. The rate of digestion is only 4% of that of the denatured enzyme (9).

Sedimentation studies (10) reveal that the subunits are probably prolate with an axial ratio of 4.4. It has further been demonstrated that the subunits lay side by side rather than end to end so maintaining a similar axial ratio in the dimerised state.

1.4 The Transition State Analogue

A much exploited finding has been the discovery that a remarkable decrease in the reactivity of the essential thiol could be brought about by the addition of certain anions to the enzyme in the presence of $MgADP^-$ and creatine. Small, planar anions, notably nitrate and formate, are particularly effective in this regard. It has been proposed that these anions can occupy the position normally occupied by the γ-phosphoryl group of $MgATP^{2-}$. In this way a complex is formed which resembles the supposed transition state of the transphosphorylation process (11).

The enzyme remains 80% active if Mg^{2+} is replaced by the paramagnetic ion Mn^{2+}. It has also been found that modification of the essential thiol, while giving rise to loss of activity, does not affect substrate binding. By

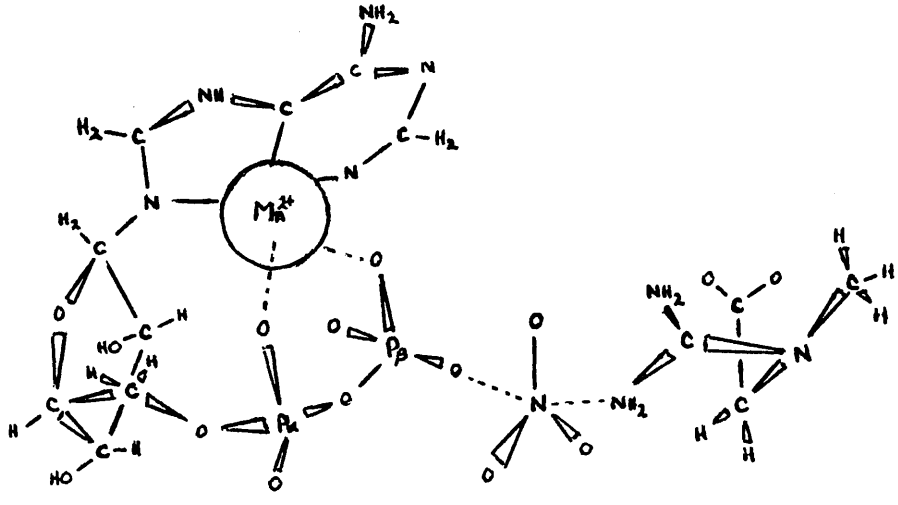


Figure 1.2 : The transition state analogue complex of creatine kinase as proposed by Reed et. al. (12)

attaching a suitable spin label to the thiol in conjunction with the replacement of Mg^{2+} by Mn^{2+}, resonance studies have enabled a detailed map of the geometry of the nitrate transition state analogue complex to be built up (12). The map proposed is shown in Figure 1.2.

It has been found that CK antibodies raised to the native protein will not bind in the presence of creatine plus $MgATP^{2-}$ but will if either substrate is present alone. From this it is inferred that CK undergoes a conformational change on substrate binding (13). This is a problem with kinases in general as regards the study of substrate binding by crystallographic methods (14). The result is that it is probably necessary to grow new crystals in the presence of the substrates rather than simply allowing them to diffuse into native crystals.

1.5 Subunit Identity In MM-Type CK

Up to now, normal MM-type CK has been described in terms of a dimer with identical subunits. There is, however, an increasing body of evidence to suggest that this is not necessarily the case.

High resolution SDS gel electrophoresis of rabbit muscle (MM-type) CK indicates the presence of two closely spaced bands (15). Using proteins of known molecular weight as calibrants, the molecular weights of these bands have been assigned values of 42,000 and 40,300. This corresponds to a difference in length of about fifteen amino acids.

In an attempt to ascertain whether this was evidence of

inherent subunit heterogeniety or the result of partial proteolysis, similar measurements were made at various stages of the preparation . A doublet was clearly visible at all stages including immediately after homogenisation of the muscle.

In this same work, the inactivation of CK by proteinase K was studied. It was found that CK was less resistant to attack by this protease than by trypsin. $MgADP^-$ in the presence of creatine slowed the rate of proteolysis by about 50%. On the addition of nitrate ions to form the transition state analogue complex, the rate was decreased by 96% and was thus attributed to a large conformational change in CK. Further analysis of this data revealed that normally the proteolysis is first order with respect to CK subunit concentration. However, in the case of the much slower proteolysis of the CK transition state analogue complex, the reaction is biphasic. The authors concluded that binding of the complex to the dimeric enzyme causes differences in conformation between the two subunits to be observed.

CK can be readily dissociated into subunits by agents such as urea or guanidinium chloride (10). Reassociation of the subunits by removal of the denaturant leads to substantial recovery of activity (16). In a study by Bickerstaff and Price (17), CK was immobilised on CNBr-activated Sepharose via one subunit only. The subunits were then dissociated to yield bound monomeric CK. It was found that the activity of this species was nearly as high as that of the dimer. Thus,

the quaternary structure of CK is not a major factor with regard to its catalytic function. This forms an interesting comparison with arginine kinase which performs a similar function in invertebrates. Arginine kinase from lobster is monomeric, of similar molecular weight to the CK subunit and has sequence similarities around its reactive thiol (18). A common ancestor might therefore be suspected for these two enzymes.

An interesting additional finding arising from the work of Bickerstaff and Price is that while acetate ions have an activating effect of about 20% on the CK monomer activity, the effect is some 70-90% with respect to the dimer. Furthermore, their study of the reaction kinetics of the reactive thiol with various derivatives of iodoacetamide, revealed a biphasic behaviour.

Finally, the work of Nevinskii et al (19) has also provided evidence of subunit heterogeniety in rabbit muscle CK. In this study, CK was dissociated using urea and applied to a Sepharose column from which two peaks eluted. The CK in these two fractions exhibited different pH maxima for the phosphorylation reaction.

References to Chapter I

1) Watts, D.C (1973) 'The Enzymes' , 8, 383

2) Burger, A., Eppengerger, M., Wiesmann, U. and Richterich, R. (1963) Helv. Physiol. Acta, 21, C6

3) Burger, A., Richterich, R. and Aebi, H. (1964) Biochem. Z., 339,305

4) Eppenberger, H.M., Eppengerger, M.E., Richterich, R. and Aebi, H. (1964) Develop. Biol., 10, 1

5) Richterich, R. (1966) 'Progressive Muskel Dystrophie', Kuhn, E. ed, Springer

6) Hooton, B.T. amd Watts, D.C. (1966) Biochem. J., 96, 81

7) Putney, S., Herliky, W., Royal, N., Pang, H., Aposhian, H.V., Pickering, L., Belagaje, R., Briemann, K., Page, D., Kuby, S. and Schimmel, P. (1984) J. Bio. Chem., 259, N23, 14317

8) Orial-Andit, C., Landon, M., Robin, Y. and Van Thoai, N. (1969) Biochem., 5, 144

9) Lui, N.S.T. and Cunningham, L. (1966) Biochem, 5, 144

10) Yue, R.H., Palmieri, R.A., Olson, O.E. and Kuby, S.A. (1967) Biochem, 6 3204

11) Milner-White, E.J. and Watts, D.C. (1971) Biochem. J., 122, 727

12) Reed, G.H. and Cohn, M (1972) J. Biol. Chem., 247, 3073

13) Samuels, A.J. (1961) Biophys. J., 1, 437

14) Blake, C.C.F. (1978) Nature, 271, 707

15) Williamson, J., Greene, J., Cherif, S. and Milner-White, E.J. (1977) Biochem. J., 167, 731

16) Dawson, D.M., Eppenberger, H.M. and Kaplan, N.O. (1967) J. Biol. Chem., 242, 210

17) Bickerstaff, G.F. and Price, N.C. (1978) Int. J. Biochem., 9, 1

18) Morrison, J.F. (1973) 'The Enzymes', 8, 457

19) Nevinskii, G.A., Ankilova, V.N., Lavrik, O.N., Mkrtchyau, Z.S., Nersesova, L.S. and Akopyan, Zh.I. (1983) Biochem., 48, Part 2, 299 (Translation of Biokhimiya)

CHAPTER II

CRYSTALLISATION

2.1 Introduction

Many advances have been made in the last few years in the methods employed in the crystallisation of proteins. The precipitating agents available comprise a variety of inorganic salts and organic solvents in additon to polyethylene glycol (PEG) which has been used recently to great effect (1). The various techniques in use rely, in the main, on the gradual attainment of these agents to a precipitating concentration by either vapour diffusion or dialysis. Alternatively, the varying solubility of proteins with respect to pH or temperature may be exploited. A comprehensive review of crystallisation techniques employed in protein crystallography has been given by McPherson (2). Although it is a simple matter to cause a protein to precipitate, causing it to do so in the form of large single crystals is less straightforward. Indeed, at present it is impossible to predict which combination of conditions will induce crystal formation in a new protein. Therefore, initially it is necessary to investigate a wide range of conditions in the hope that one or more will prove successful.

2.1.1 Electrostatic Theory

The use of inorganic salts to achieve protein precipitation has traditionally been the favoured method of crystallisation although present trends seem to indicate other methods are now taking precedence. The action of inorganic salts in solution to modify protein solubility

can be described by the electrostatic model of Debye-Huckel (3). At low salt concentrations, the protein molecule is surrounded by an atmosphere of ions comprising an excess of those ions of opposite charge to the net charge on the protein at that pH. The screening effected by this ionic atmosphere at low salt concentrations decreases the electrostatic free energy of the protein resulting in an increase in solubility. This precess is known as 'salting-in' and may be described by

$$\log S - \log S_o = \frac{Az_+z_-\sqrt{\mu}}{1 + aB\sqrt{\mu}} \qquad \ldots\ldots\ldots\ldots\ldots \textcircled{1}$$

where S is the solubility of the protein at a given ionic strength, μ ; S_o is the solubility at zero ionic strength ; A and B relate to temperature and dielectric constant and μ is given by

$$\mu = \frac{1}{2} \cdot \sum c_i z_i^2$$

The denominator in expression ① , $1 + aB\sqrt{\mu}$, relates to the average diameter, a, of the ions.

However, as the salt concentration increases, competition for water molecules starts to occur. This results in a decrease in solubility of the protein with increasing ionic strength. Expression ① must therefore be modified to account for this such that

$$\log S - \log S_o = \frac{Az_+z_-\sqrt{\mu}}{1 + aB\sqrt{\mu}} - K_s \mu$$

where K_s is the so-called 'salting-out' constant and is a parameter characteristic of the particular salt.

In practice, the favoured salt is ammonium sulphate. This is less due to its value of K_s, which is not particularly high, than it is to its own great solubility.

2.1.2 Organic Solvents

The effect of adding an organic solvent to water is to lower the dielectric constant of the medium. This increases the coulombic interaction between unlike charges on the protein molecules so encouraging aggregation and final precipitation. This may be described by

$$\ln(S/S_o) = (A/RT)(1/D_o - 1/D)$$

where D_o is the dielectric constant of water ; D is the dielectric constant of the mixture and A is a constant.

The practice of employing organic solvents has proved successful with many proteins and also with nucleic acids. Ethanol has been used in many circumstances but due to evaporation tends to lead to difficulties. As a result, the modern approach is to use 2-methyl-2,4-pentanediol (MPD) as an alternative precipitant. It is less volatile than ethanol and seems to be less harsh than many other organic solvents. Acetone and dimethyl sulphoxide have also been employed.

2.1.3 Polyethylene Glycol

Jansen and Ruelius (4) were the first to report the production of protein crystals from PEG precipitations. Shortly afterwards, in 1975, the structure of deoxyhaemoglobin A was published in which the crystals were reported as having been grown from PEG solutions (5).

In a theoretical treatment of precipitation of proteins by PEG, it was found that the action could be described by the equation

$$\log S = K - \beta w$$

where S is the solubility of the protein in the presence of PEG at a concentration, w ; K is a constant and β refers to the volume excluded by the polymer network (6). This is identical in form to the 'salting-out' expression although its derivation is based on a principle of steric exclusion. It is also interesting to note that PEG has the characteristic, like ammonium sulphate, of stabilising proteins in solution against denaturation. It possibly does this by immobilising the water molecules in the vicinity of the protein so encouraging 'hydrotactoid' formation (7).

2.1.4 Hydrophobic Interactions

Kauzmann (8) has suggested that the presence of non-polar groups on the protein surface which come into contact with water leads to an unfavourable free energy of interaction.

When a protein precipitates from solution, some of these interactions are replaced by protein-protein contacts. Thus protein solubility must be affected by these hydrophobic interactions. This has been related to the anomalously high value of the surface tension of water (9). In so-called cavity theory, one considers the free energy required to form a cavity in the solvent to accommodate the solute molecule. This parameter is given by

$$\Delta G_c = \sigma A$$

where A is the surface area of the cavity and σ is the surface tension of the solvent.

This idea has been developed to describe the effects of inorganic salts on protein solubility (10). The derived expression accounts for the 'salting-in' and 'salting-out' phenomena observed but appears to have limited applicability.

2.2 Techniques Employed in Protein Crystallisation

The preceding discussion has centred on the thermodynamics of protein solubility and precipitation. Successful crystal growth, however, depends also on kinetic considerations. Various methods have been developed to approach the point of precipitation in a slow, controlled manner in order to encourage crystal growth in preference to the production of amorphous precipitate. Two popular methods are those of

vapour diffusion and equilibrium dialysis.

2.2.1 Vapour Diffusion

Simply, conditions of supersaturation leading to precipitation can be attained by solvent evaporation. This process is, however, difficult to control and can lead to problems of drying out if the evaporation is not arrested.

In practice it is preferable to set up a protein solution containing a precipitating agent some 10% below the concentration required. This solution is then placed in equilibrium in a closed environment with a solution of higher concentration of precipitating agent. Thus, by a process of vapour diffusion, the point of protein precipitation may be slowly approached.

This principle may be applied in practice in a number of ways, the most popular being either the use of special vapour diffusion dishes or by the so-called hanging drop method (11).

2.2.2 Equilibrium Dialysis

As an alternative to vapour diffusion, a precipitating concentration of the agent employed can be attained by diffusion through a semi-permeable membrane. Such membranes allow the passage of small molecules through them but exclude large molecules such as proteins. A convenient adaptation of this method to the use of small quantities of solution is the microdiffusion method of Zeppezaur (12). In this instance, the protein solution is contained within a

thick walled capillary over the end of which is secured the dialysis membrane. The capillary is then immersed in a precipitating solution and allowed to equilibrate slowly in the familiar fashion.

This method has the added advantage that the capillaries may be reused or the smaller crystals redissolved in order to develop crystal size.

2.3 Crystallisation of CK

2.3.1. Existing Crystal Forms

CK was first crystallised by Kuby et al (13) in the form of microcrystalline needles. These were grown from ethanol-water mixtures in the final stage of the preparation. However, their size and fragility made them unsuitable for X-ray analysis.

In 1973, McPherson (14) published a preliminary crystallographic investigation of CK in which he reported three novel crystal forms, all grown from MPD. One of the forms reported was particularly suitable for X-ray studies. It was in an orthorhombic space group (either $P2_12_12_1$ or $P2_12_12_1$) with apparently one subunit in the asymmetric unit. The other two forms reported were monoclinic, in $P2_1$ and C2, each with four subunits in the asymmetric unit.

However, there were evidently problems encountered in reproducing these results and by 1975, when this project commenced, there had been no further publications on CK from that source. Repeated attempts in this laboratory to produce

the orthorhombic crystals described failed, yielding only micro-needles. However, the monoclinic C2 crystals grown in the presence of methyl mercury acetate by McPherson were reproduced in conditions of 60% MPD, 20 mM Tris/acetate pH 8.2 by the method of micro-dialysis. The crystals were of poorly defined external morphology giving the appearance of aggregates of similarly oriented needles as depicted in Figure 2.1. The cell dimensions were found to be in close agreement with those given by McPherson viz.

$$a = 238\ \text{Å}$$
$$b = 164\ \text{Å}$$
$$c = 52\ \text{Å}$$
$$\beta = 91°$$

All three principle projections were characterised and are depicted in Figures 2.2, 2.3 and 2.4. It can be seen from these 9° precession photographs that diffraction was observable at least to a resolution of 5 Å. However, the crystals were found to be very susceptible to damage by X-ray exposure. Even with the employment of crystal cooling during X-ray exposure, useful lifetimes were limited to approximately fifteen hours.

The enzyme employed in all the crystallisation work and other studies described in this thesis, was prepared as described in Appendix I.

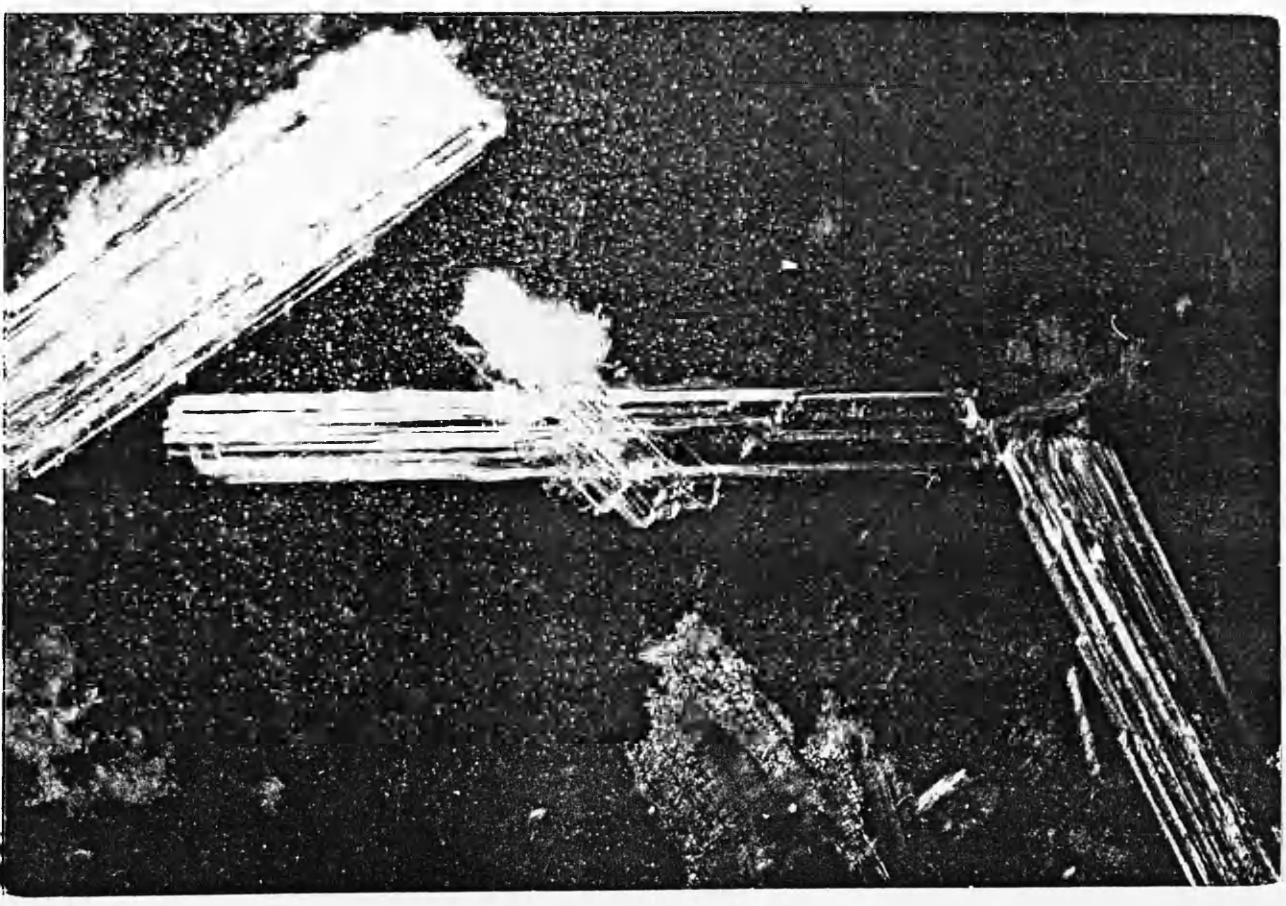


Figure 2.1 : Monoclinic crystal form of creatine kinase (magnification X 50).

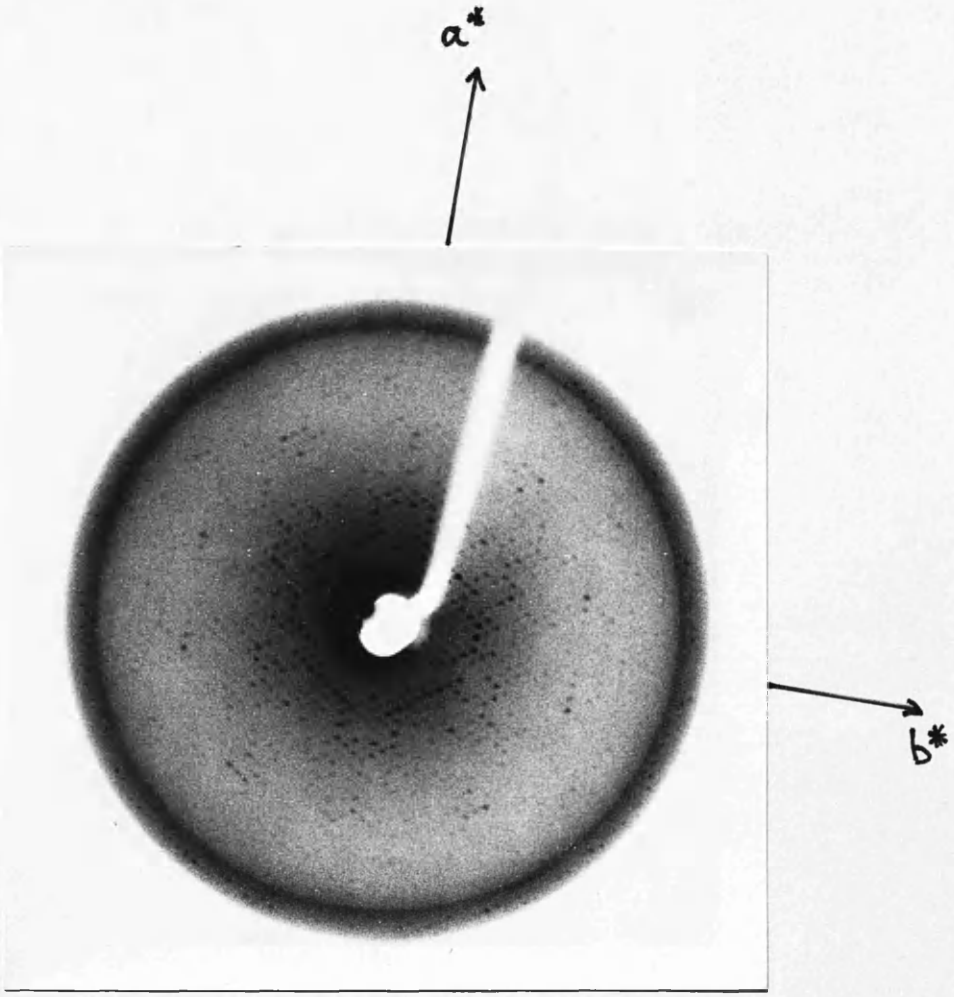


Figure 2.2 : hk0 9° precession photograph of the C2 crystal form of creatine kinase previously discovered by McPherson (14) referred to in text as K2. Exposure 12 hours. Crystal to film distance 100 mm, annulus width 2mm.

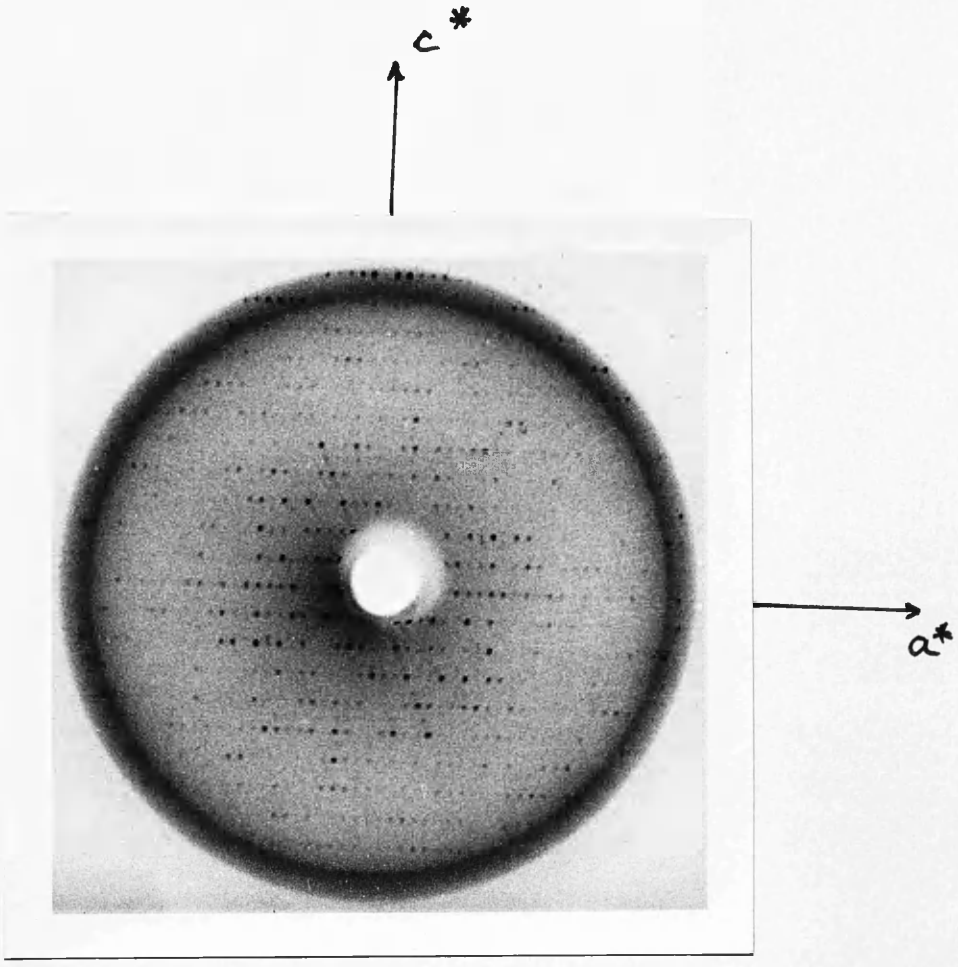


Figure 2.3 : h01 9° precession photograph of the C2 crystal form of creatine kinase previously discovered by McPherson (14) referred to in text as K2. Exposure 12 hours. Crystal to film distance 100 mm, annulus width 2 mm.

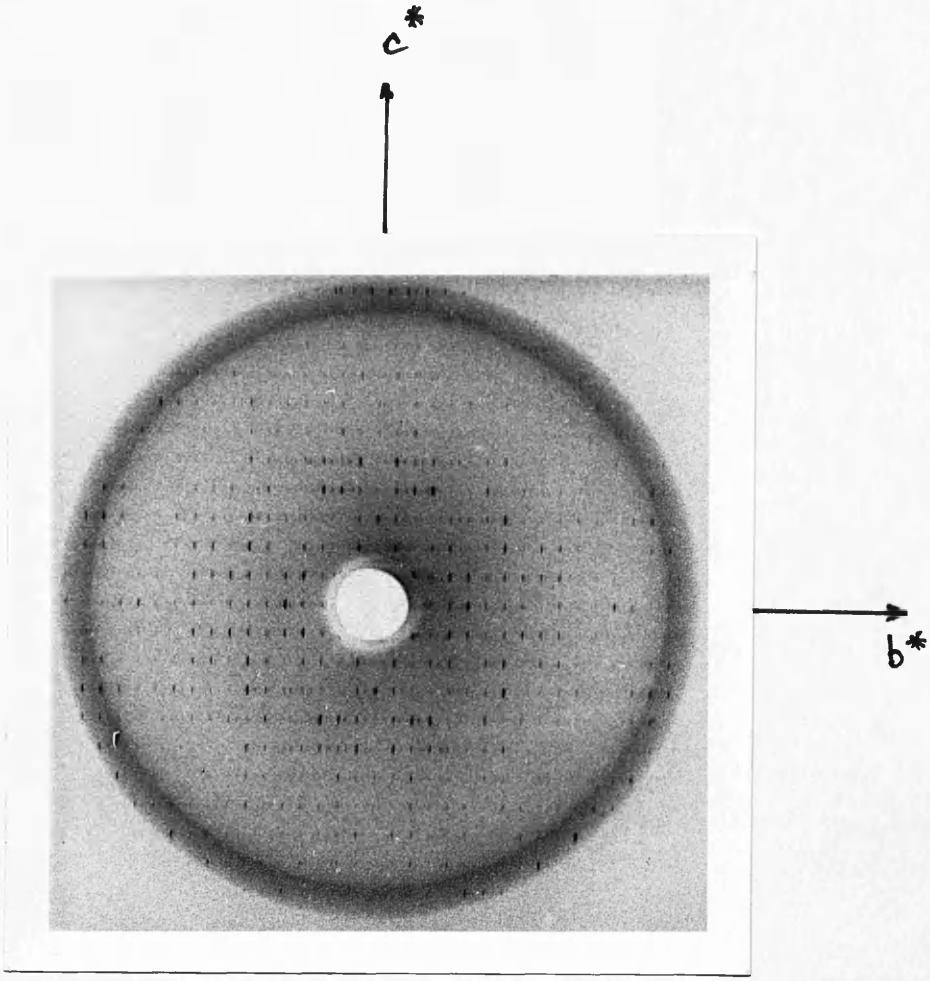


Figure 2.4 : 0kl 9° precession photograph of the C2 crystal form of creatine kinase previously discovered by McPherson (14) referred to in text as K2. Exposure 12 hours. Crystal to film distance 100 mm, annulus width 2 mm.

2.3.2 The Identification of a Novel Crystal Form of CK

In crystallisations under the conditions quoted above, it was discovered that a second, novel crystal form of CK was present in very small numbers. These crystals are indistinguishable in appearance from the previously characterised form. Systematic absences according to h+k odd revealed that the space group of this new crystal form is also C2 with cell parameters given by

$$a = 248\ \text{Å}$$
$$b = 149\ \text{Å}$$
$$c = 52\ \text{Å}$$
$$\beta = 90°$$

and is thus very similar to the other C2 form.

For the sake of brevity, these two monoclinic crystal forms of CK will hereinafter be referred to as K1, for the new variety and K2 for the previously characterised form.

The extreme rarity and similar susceptibility of the K1 form to radiation damage has meant that only the hk0 and h0l zones have been characterised. Precession photographs of these projections are depicted in Figures 2.5 and 2.6 respectively.

Comparison of the diffraction patterns of the two crystal forms reveals little or no similarity in the hk0 projections. However, the h0l zones do display a pronounced similarity indicating that the two crystal forms may be closely related as suggested by their cell dimensions.

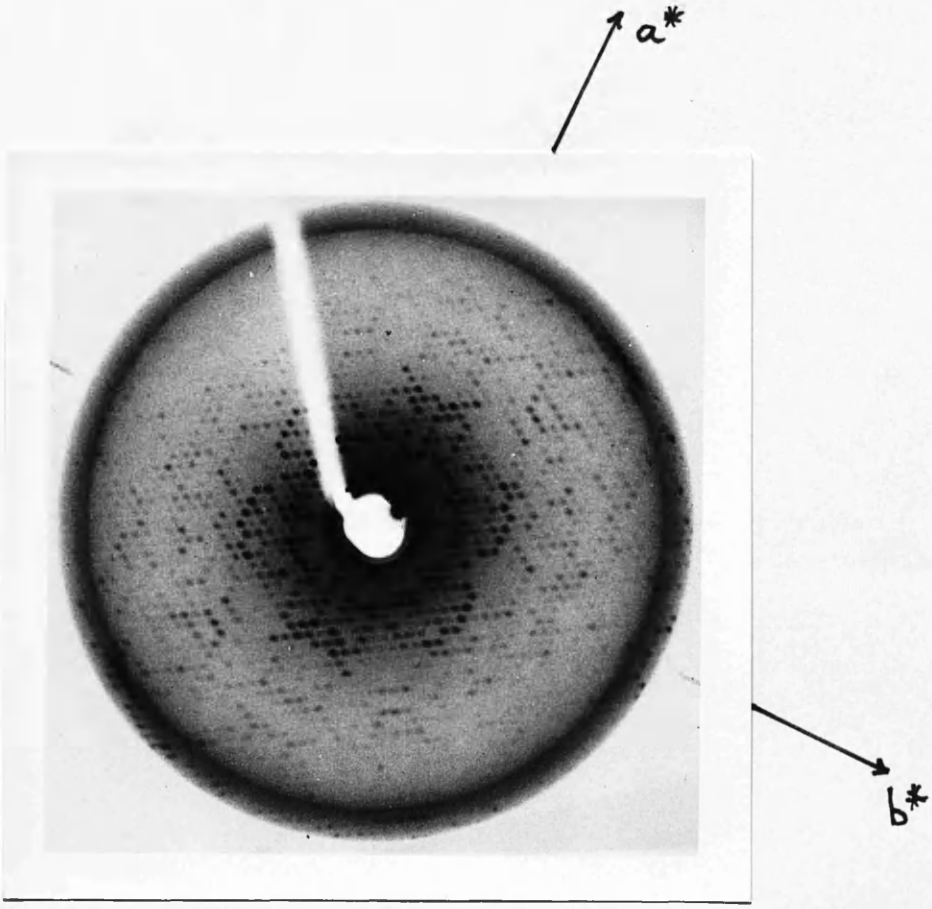


Figure 2.5 : hk0 9° precession photograph of the new C2 crystal form of creatine kinase referred to in text as K1. Exposure 12 hours. Crystal to film distance 100 mm, annulus width 2 mm.

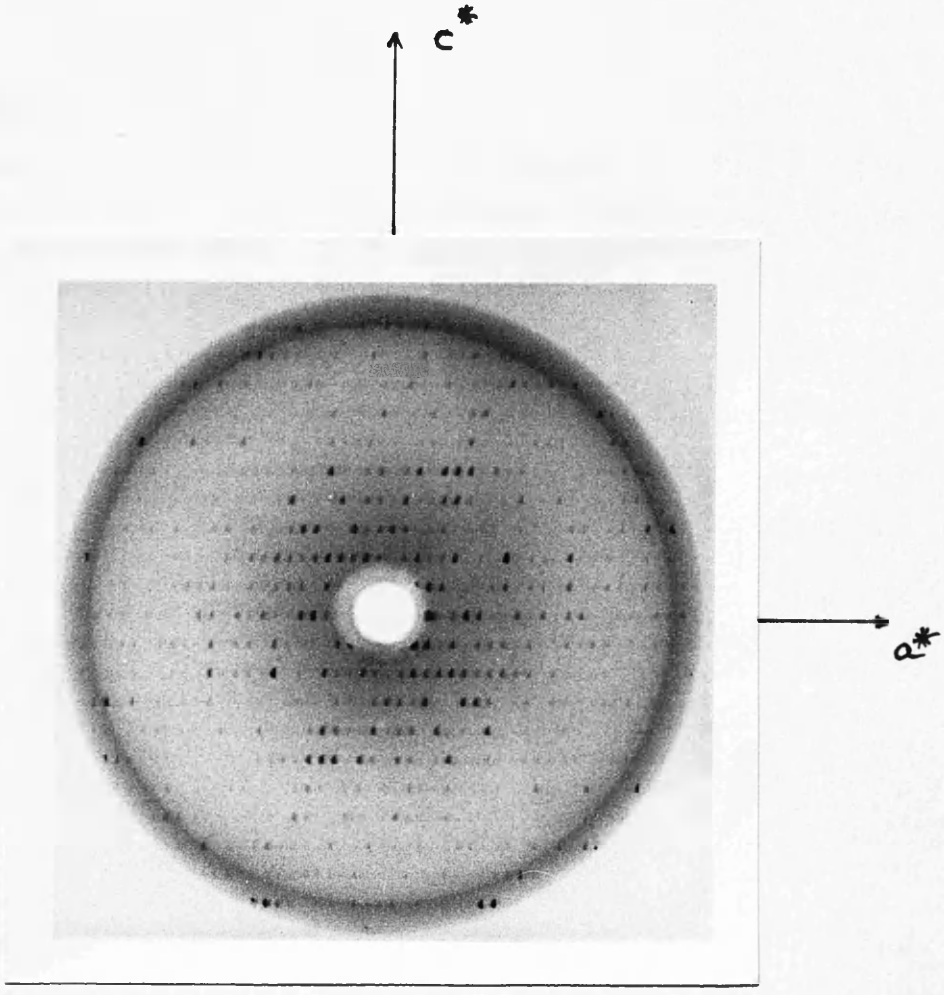


Figure 2.6 : h01 9° precession photograph of the new C2 crystal form of creatine kinase referred to in text as K1. Exposure 12 hours. Crystal to film distance 100 mm, annulus width 2 mm.

2.3.3 The Solvent Content of K1 and K2

Matthews (15), in a review of 116 reported crystal forms, has found that the ratio of the volume of the crystallographic asymmetric unit to the weight of protein in the unit generally falls in the range 1.68 to 3.53. In this way, it is possible to estimate the number of subunits in the asymmetric unit by calculating this ratio, Vm, for the various possibilities. Furthermore, assuming a value of 0.74 cm^3/g for the partial specific volume of protein in the crystal, it is possible to estimate the fractional volume occupied by solvent. This is given by

$$V_{solv} = 1 - \frac{1.23}{V_m}$$

Thus for the K1 and K2 crystal forms of CK, the following possibilities may be considered :

Type	No. of Subunits per asymmetric unit	V_m ($Å^3$/Dalton)	V_{solv}
K1	3	3.73	67%
	4	2.80	56%
	5	2.24	45%
	6	1.86	34%
K2	3	3.94	69%
	4	2.95	58%
	5	2.36	48%
	6	1.97	37%

The extreme fragility of the crystals and relatively high molecular weight of the protein would tend to suggest the higher solvent content possibilities.

2.3.4 Discussion

It will be observed that the conditions quoted both by McPherson and in this thesis for the production of the monoclinic crystal forms of CK, include the presence of acetate ions. As stated in Chapter I, it has been found that acetate ions are suspected to induce a conformational change in the dimer of CK (16). It was interesting therefore to assess the importance of the presence of acetate ions in the crystallisation conditions. Conditions employing Tris buffer with hydrochloric acid or succinic acid in place of acetic acid still produced crystals but these were of poor quality and no observable diffraction could be obtained from them.

Exploration of the conditions under which the monoclinic crystals would form, revealed reasonable flexibility. The buffer concentration could range from as low as 10 mM to as high as 50 mM though at concentrations in excess of 35 mM the crystals would not grow as large. Likewise, pH values ranged from 7.6 to 9.0.

Generally, the favoured method of crystallisation was micro-dialysis. For this, a slight variation of the method of Zeppezauer was employed. The method normally requires the thick-walled capillaries to be stood upright on small plastic feet. It was found, however, that this encouraged crystal growth to aggregate at the bottom of the capillary near the membrane making visualisation and retrieval of the crystals difficult. The simple variation of dispensing with the feet and laying the capillaries on their sides, resolved this difficulty and is to be recommended. It was found that

the monoclinic crystals could equally well be grown by vapour diffusion either using vapour diffusion dishes or by the hanging drop method.

It is interesting to note in passing that lower buffer concentrations produced better crystals at higher pH and vice versa suggestive of an inverse relationship between the two. Figure 2.7 depicts a grid of crystallisations that was investigated using triethanolamine/thioglycolic acid as buffer. The relationship is clearly suggested in this example.

It was also found that the protein could be crystallised using a completely different buffer. This could be accomplished if care was taken to ensure that the ionic strength was the same as that of the stated conditions. With CK, this generally gave rise to poorer crystals or no crystals at all but suggests that it is worth investigating a range of buffers once the pH and ionic strength required for crystal formation have been defined.

It is clear that with crystallisations involving MPD as precipitant, the ionic strength is a critical factor and a range of buffer concentrations should be investigated in any initial study.

It will be appreciated that neither of the two monoclinic C2 crystal forms that have been described is ideal for X-ray studies on the enzyme. The likely presence of multiple subunits in the asymmetric unit is unattractive. Furthermore, the only centric projection in this space

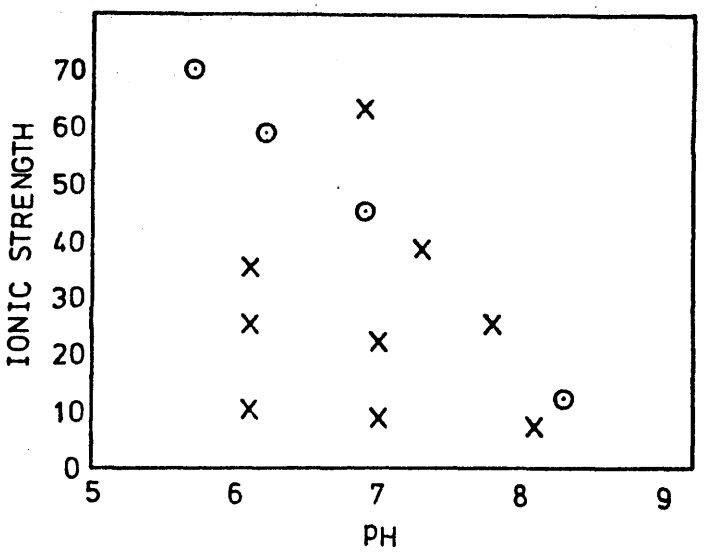


Figure 2.7 : Results from a grid of crystallisations conditions attempted for creatine kinase using TEA/thioglycolic acid buffer. ⊙ - crystals formed, X - amorphous precipitate.

group, h0l, is obtained at right angles to the needle axis and is difficult to locate due to the extremely poorly defined morphology of the crystals. Centric projections are useful in the determination of heavy atom positions (Chapter VI). The poor radiation resistance of these crystals coupled with the above considerations thus render a structure determination of CK based on either K1 or K2 an unattractive prospect although it would be more feasible using the high intensity synchrotron radiation source that is now available at Daresbury (17).

Consequently, a wide range of conditions was explored in search of an alternative crystal form. A full summary of these conditions is set out in Appendix II. Unfortunately, very little success was achieved and no other crystal forms suitable for X-ray analysis were discovered.

This study used up the initial batch of CK and another preparation was performed to replenish supplies. This second batch failed to produce the K1 and K2 crystal varieties which had formed so easily before. Subsequent preparations produced enzyme of varying ability to produce satisfactory crystals and none ever matched the ability of the first preparation.

One is led to assume that a similar fate befell the orthorhombic crystals of McPherson. Although the enzyme appears each time to be identical to previous batches by electrophoresis, there is evidently some inconsistency in the preparation or inherent in the protein itself. This matter is taken up more fully in Chapter VII.

2.3.5 Data Collection

Crystals were mounted and sealed in thin-walled quartz X-ray capillaries in equilibrium with a small amount of the liquor of crystallisation. In general, the crystals were mounted such that the needle axis of the crystal was oriented perpendicular to the capillary axis. In this way the hk0 zone could be readily located.

Data collection was achieved by means of screened precession photography. For this purpose, Ilford Industrial G film was used throughout. The majority of the data were obtained using a Marconi-Elliot GX20 rotating anode X-ray generator operating at 2 kW (40 kV, 50 mA), λ =1.54 Å (Cu K_α , Ni filtered).

Films were scanned using a Joyce-Loeble flat bed photodensitometer. Equivalent reflections were measured and averaged. Table 2.1 shows the details of the native data collected together with an estimate of the errors. These data were entered for computer analysis using punched cards. Subsequent processing was carried out on the Glasgow University ICL 2976 using the XRAY72 program suite.

Table 2.1 : Native Data Sets

Data Set	No. of unique reflections	Rsym	$(\sin\theta/\lambda)_{max}$
K2HK0	121	0.062	0.0854
K20KL	92	0.047	0.0918
K2H0L	187	0.054	0.0970
K1HKO	159	0.087	0.0841
K1H0L	189	0.052	0.0961

$$Rsym = \frac{\sum_h \sum_i (\bar{I}(h) - I(h)_i)}{\sum_h \sum_i I(h)_i}$$

2.4 Postscript to Chapter II

Since the work for this thesis was completed, two new reports concerning the crystallisation of CK have appeared in the literature. The first of these reports (18) describes the crystallisation of two crystal forms of bovine heart CK. The authors describe a process of what is assumed to be partial precipitation using MPD followed by clarification by centrifugation at 10,000 g for 30 minutes. There then follows a careful two stage seeding process to produce crystals suitable for diffraction.

The second, more recent work (19), concerns the crystallisation of rabbit skeletal muscle CK. In this report, isoelectric focussing was used as a final purification stage in order to obtain good crystals. This purification method is described in Chapter VII of this thesis where predictions of its likely success are discussed.

References for Chapter II

1) McPherson, A. (1976) J. Biol. Chem., 251, 6300

2) McPherson, A. (1976) 'Methods of Biochemical Analysis', 23, 249 (Academic Press, N.Y.)

3) Arakawa, T. and Timasheff, S.N. (1985) 'Methods in Enzymology', 114, 49

4) Jansenn, F.W. and Ruelius, H.W. (1968) Biochem. Biophys. Acta, 151, 330

5) Ward, K.B., Wishner, B.C., Lattman, E.E., Love, W.E. (1975) J. Mol. Biol., 98, 161

6) Juckes, I.R.M. (1971) Biochem. Biophys. Acta, 229, 535

7) Klotz, I.M. (1965) Fed. Proc., 24(2), part III suppl 15, S-24

8) Kauzmann, W (1959) Adv. Protein Chem., 14, 1

9) Sinanoglu, O. and Abdulner, S. (1965) Fed. Proc. Fed. Am. Soc. Exp. Biol., 24, 512

10) Melander, W. amd Horvath, C. (1977) Arch. Biochem. Biophys., 183, 200

11) Blundell, T.L. and Johnson, L.N. (1976) 'Protein Crystallography', p75

12) Zeppezauer, M. (1971) Adv. Enz., 22, 253

13) Kuby, S.A., Noda, L., Lardy, H.A. (1954) J. Biol. Chem., 209, 191

14) McPherson, A. (1973) J. Mol. Biol., 81, 83

15) Matthews, B.W. (1968) J. Mol. Biol, 33, 491

16) Bickerstaff, G.F., Price, N.C. (1976) Soc. Trans., 4, 1061

17) Helliwell, J.R. (1985) J. Mol. Struc., 130, 63

18) Gilliland, G.L. Sjolin, L, Olsson, G. (1983) J. Mol. Biol., 170, 791

19) Hershenson, S., Helmers, N., Desmeules, P., Stroud, R. (1986) J. Mol. Biol., 261, 3732

CHAPTER III

SUBUNIT IDENTITY

3.1 K1 Pseudo-symmetry

Closer inspection of the K1 hk0 diffraction pattern (Figure 2.5) reveals that mirror planes are observable with corresponding spikes of higher intensity at 60° intervals. In this way, the photograph displays 6m planar point group symmetry. The mirror planes extra to those characteristic of the C2 space group arise from the projection of the local, molecular two-fold axis onto the x-y plane.

Multi-subunit proteins can crystallise in such a way as to incorporate the molecular symmetry into the crystallographic system giving rise to only one subunit in the asymmetric unit. This, however, is not always the case and, indeed, the Vm values of both the monoclinic forms of CK indicate the presence of more than one subunit in the asymmetric unit.

The presence of molecular symmetry axes not coincident with any crystallographic symmetry axes can give rise to certain features in the diffraction pattern. Reflections in reciprocal space along the direction of a local (non-crystallographic) axis will be of higher intensity. This is because the atoms related by such an axis will diffract in phase. Thus, spikes of higher intensity can sometimes be observed indicating the position of the molecular two-fold or higher symmetry axes.

Such is the case with the rarer K1 monoclinic crystal form of CK. The spikes of higher intensity at 60° intervals in the hk0 diffraction pattern are highly indicative of non-crystallographic symmetry and suggest the presence of a pseudo-cell.

Due to Friedel's law, it is impossible to say whether the point group is actually 6m or 31m. It should be noted that there is no suggestion of any enhancement of any particular classes of reflections in the diffraction pattern. This would have implied the presence of screw axes in the pseudo-cell.

Thus, every molecule is related to a second molecule by a rotation of either 60° or 120° . Hexagonal and trigonal lattices are commonly distinguished by upper level precession photographs. Unfortunately, the extreme rarity of these crystals has so far precluded such an investigation. However, the two crystal systems can also be distinguished by principal projections perpendicular to the major symmetry axis. In this case, a trigonal lattice will display only a centre of symmetry while a hexagonal system will exhibit mm symmetry. The h01 zone of the K1 crytal form shown in Figure 2.6 would seem to indicate the former.

It is interesting to note that the ratio of a to b in the monoclinic cell is 1.67, close to the value of $\sqrt{3}$ which would be observed in a perfect trigonal system. As a consequence, the relative positions of the local symmetry axes very nearly satisfy simultaneously the requirements of both crystal systems.

On this basis the monoclinic cell very probably contains twelve repeating units arranged as in Figure 3.1. If these repeating units were CK dimers, the cell would contain tetramers for which there is no other supporting evidence.

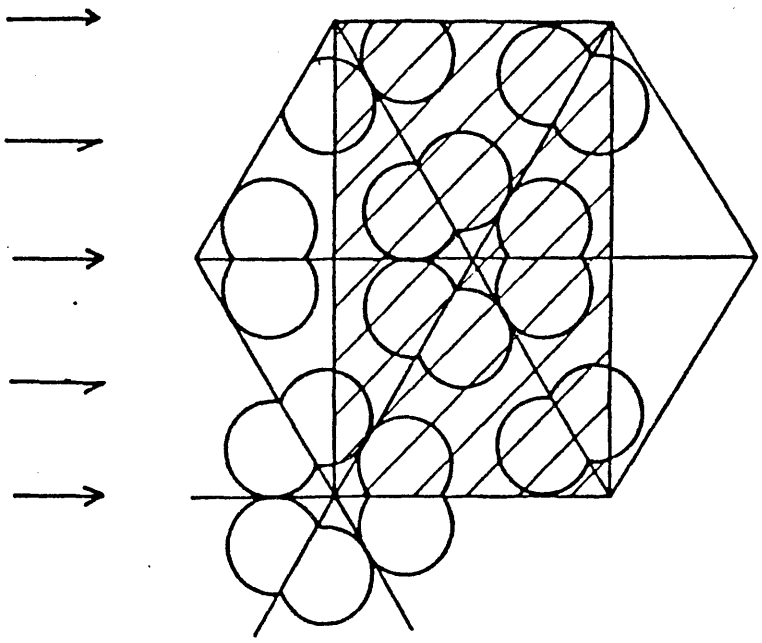


Figure 3.1 : Packing arrangement of creatine kinase in pseudo-trigonal cell. Monoclinic cell represented by shaded area. Trigonal cell corresponds to any pair of adjacent equilateral triangles.

The repeating unit is thus likely to be a single CK subunit. This means that there are three subunits in the monoclinic asymmetric unit giving rise to a Vm of 3.73 $Å^3$/Dalton which is within the range observed for proteins of this molecular weight (1).

The conclusions of this study relate to the matter covered in Chapter I regarding the question of subunit homogeneity. As was mentioned, various studies have cast doubt on whether the subunits of rabbit skeletal muscle CK are identical. The results of this X-ray investigation show that in this crystal form, one third of the dimers lie on a crystallographic two-fold axis and are thus composed of identical subunits. The remaining dimers give rise to local two-fold axes observable as mirror planes in the hk0 diffraction pattern. These mirror planes are observable to a resolution of c.7Å. One can, therefore, conclude that the subunits of these dimers are indistinguishable at this resolution and that if any differences exist, they do not affect the general folding of the polypeptide chain (2).

The preceding analysis reveals that the breakdown of non-crystallographic symmetry at high resolution stems, at least in part, from imperfect correspondence of the local symmetry axes with crystallographic axes. An example of the effect of such misalignment is the case of prealbumin, a protein with identical subunits. In this case, deviation of the molecular symmetry axes by only five degrees from the crystallographic directions is sufficient to destroy the pseudo-symmetry beyond 6Å resolution (3).

3.2 The Rotation Function

3.2.1 Theory

Two molecules in an asymmetric unit differing only in orientation and position may be related to one another by a rotation and a translation ie

$$X_2 = [C]\ X_1 + d$$

where [C] represents the rotation matrix. The best way to relate the orientations of the two molecules is by comparison of their self Patterson functions. A Patterson function contains peaks corresponding to both intramolecular vectors and intermolecular vectors (Section 4.1.7). The former occur within a radius from the origin represented by the maximum intramolecular distance possible. Maximum superimposition of these vectors upon a certain degree of rotation corresponds to the relative orientation of the molecules themselves. This maximum superimposition is determined by the use of the rotation function due to Rossman and Blow (4, 5, 6) viz:

$$R = \int_U P_2(X_2)\ P_1(X_1)\ dX$$

where R is the rotation function, P_1 is the model Patterson function and P_2 is the rotated Patterson function. This product is calculated at intervals of rotation within a volume, U, and consequently maximised. The full three

dimensional rotation is expressed by means of Eulerian angles as depicted in Figure 3.2.

Rotation of vectors in reciprocal space requires that points corresponding to non-integral reciprocal lattice vectors need to be considered. These are calculated by interpolation.

If the above expression for the rotation function is expanded, we obtain:

$$R = \frac{1}{V^2} \sum_P \sum_H |F_P|^2 |F_H|^2 \int_a \cos[2\pi(h+h')X_i]dX_i$$

where h' is a non-integral reciprocal lattice vector. The integral is an interferance function with value

$$\frac{U}{V} |G_{hh'}|$$

The function G is such that all terms where $|h+h'|$ is less than 1.5 can be neglected. The result of this is that one need only consider a cube in reciprocal space containing 27 lattice points. This is an important feature in what is a lengthy calculation demanding much computer time. In fact, in the application of the rotation function to the CK monoclinic data, a larger two-dimensional matrix was employed. This was because of the systematic absences present in the hk0 projection.

The rotation function used in this work was written by P.Cradwick and amended for my use by Dr.C.Gilmore.

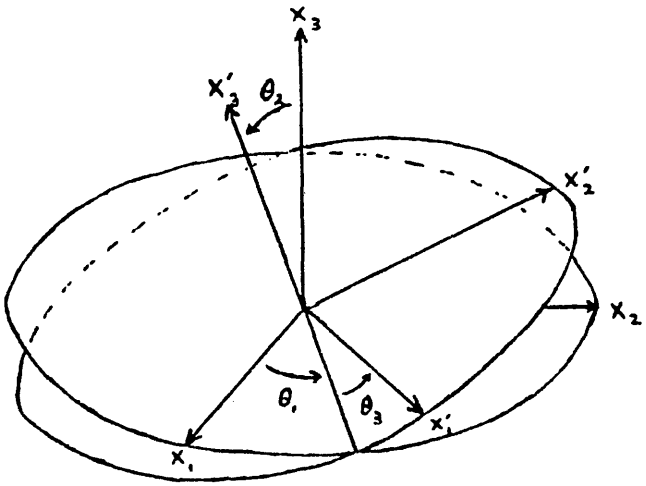


Figure 3.2 : Diagrammatic representation of the Eulerian angles employed in rotation function calculations. Reproduced from reference 4.

3.2.2 Application to CK Data

Figure 3.3 shows the result of the rotation function when applied to the hk0 projection data of the K1 crystal form of CK. As mentioned, due to the systematic absences present in this projection, 5x5 latice points were employed for interpolation and low angle data corresponding to a resolution less than 16.5 Å was omitted in order to sharpen the overlap. The unit cell thickness of both monoclinic crystal forms indicate that the diameter of the CK molecule is of the order of 50 Å. Thus, to ensure the exclusion of any intermolecular vectors, the initial sphere of rotation was taken to have a radius of 30 Å. Under these conditions, it can be seen that a peak of intensity c.94% of the origin peak was obtained at a rotation of about 61° compared to a background of approximately 80%. This shows that the pseudo-symmetry exhibited in the diffraction pattern is very pronounced and confirms the predictions made purely by inspection of the photograph.

Figure 3.4 shows how the rotation function plot is modified upon alteration of the radius of rotation. The larger radius of 80 Å causes a decrease in the relative intensity of the peak at 60° and gives rise to a noisier background presumably due to the inclusion of intermolecular vectors. It was discovered that the rotation function peak at 60° could be produced with the minimum of contamination from other correlations by reducing the resolution of the data. The exceptionally clean profile depicted in Figure 3.5 was obtained using data only in the range 10-33 Å, radius = 50Å.

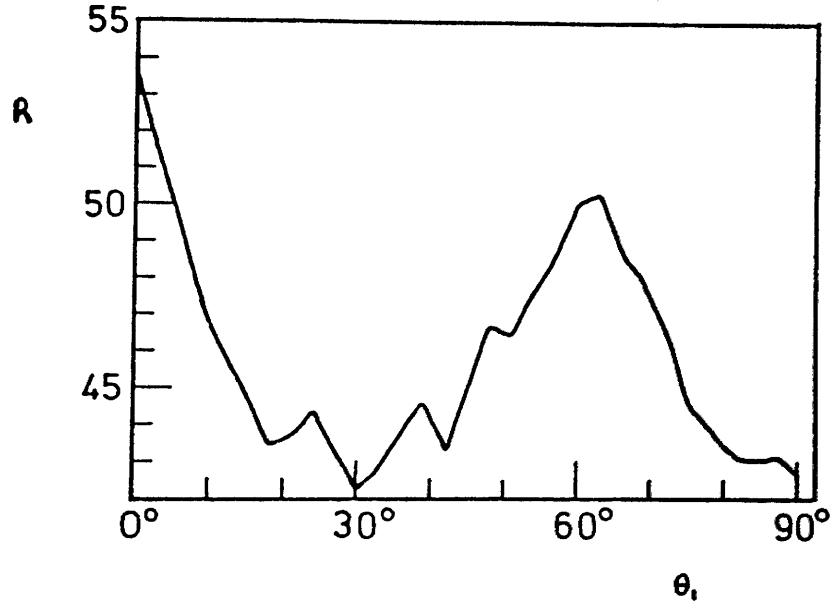


Figure 3.3 : Rotation function profile calculated from hk0 data of K1 crystal form. All observed data between 6 Å and 16.5 Å resolution. Radius of sphere, 30 Å.

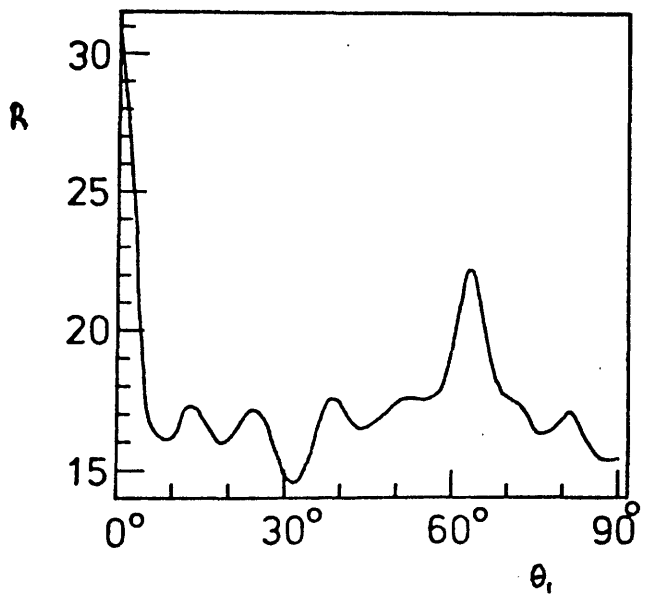


Figure 3.4 : Rotation function profile calculated from hk0 data of K1 crystal form. All observed data between 6 Å and 16.5 Å resolution. Radius of sphere, 80 Å.

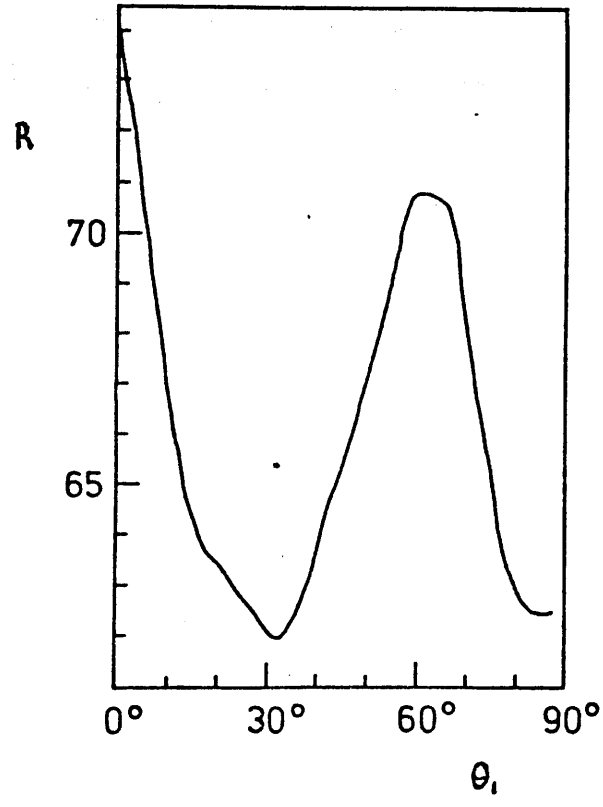


Figure 3.5 : Rotation function profile calculated from hk0 data of K1 crystal form. All observed data between 10 Å and 33 Å resolution. Radius of sphere, 50 Å.

Rotation function studies on the other monoclinic form, K2, were also carried out. The diffraction pattern of this form lacks the distinctive spiking of the K1 variety and this is echoed by the rotation function profile (Figure 3.6). There are several peaks in this profile, the largest of which is at about 69°. However, it is much less prominent than the 60° peak in the K1 analysis and three-dimensional data will be required before a more conclusive analysis of this crystal form can be accomplished.

As the two crystal forms both grow under the same conditions and also have similar morphology and cell dimensions, a relationship between the two was possible. While there is evidence for this from the h0l projections, the hk0 projections reveal no apparent similarity. Therefore, a cross rotation function was calculated in an attempt to show up any relationship not apparent by simple inspection. The profile resulting from this calculation (Figure 3.7) suggests that there is, indeed, a similarity between the two crystal forms but that it cannot be adequately revealed by projection data alone.

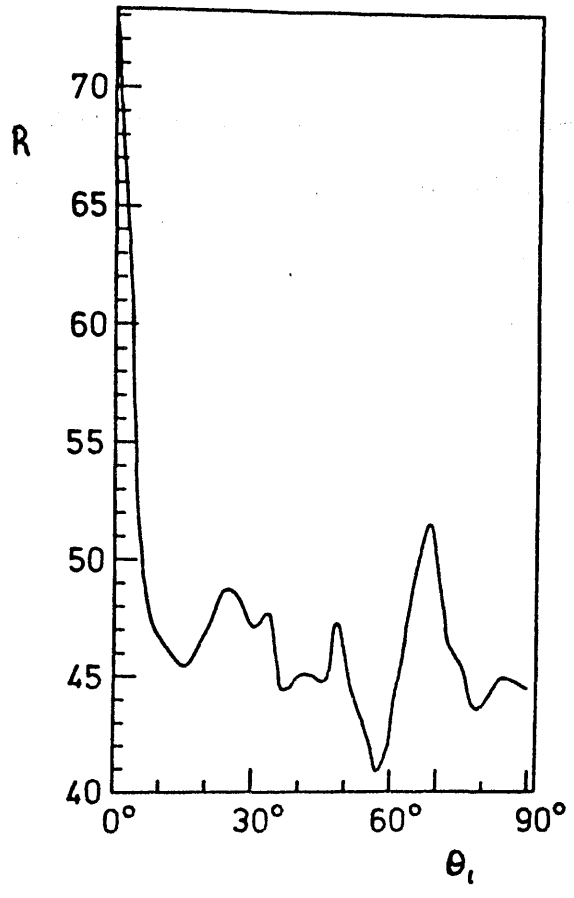


Figure 3.6 : Rotation function profile calculated from hk0 data of K2 crystal form. All observed data between 6 Å and 16.5 Å resolution. Radius of sphere, 30 Å.

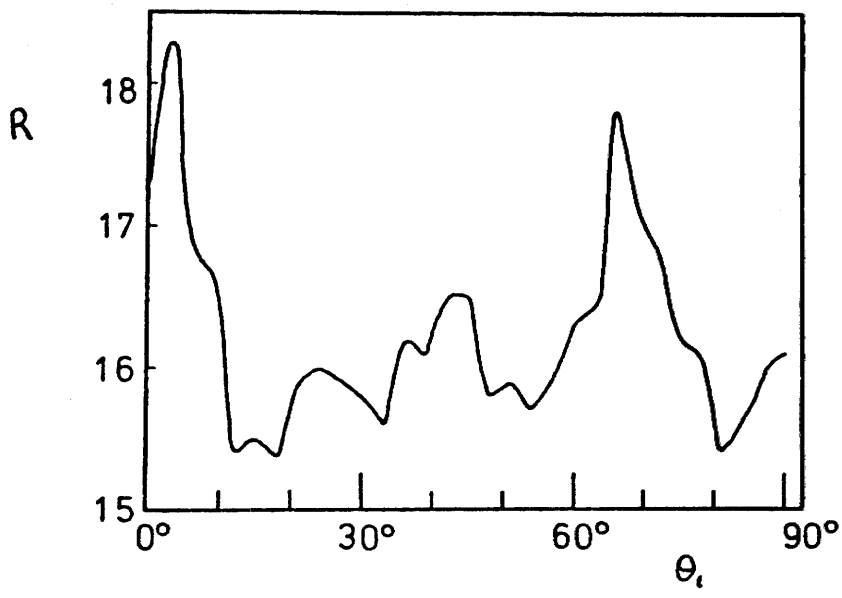


Figure 3.7 : Cross Rotation function profile calculated from hk0 data of K1 and K2 crystal forms. All observed data between 6 Å and 16.5 Å resolution. Radius of sphere, 30 Å.

3.3 Uses of Non-crystallographic Symmetry

One conclusion that arises from this work is the possible suitability of the K1 crystal form for further structural analysis. As has been pointed out, the weight of protein in the asymmetric unit is not ideal for high resolution structural analysis although such a study would be made more feasible using high intensity synchrotron radiation. However, the high pseudo-symmetry of the K1 crystal form makes the situation slightly more appealing. This is due to the possibility of using this pseudo-symmetry to either improve phase information or possibly even derive it directly.

Tobacco mosaic virus coat protein which has seventeen-fold molecular symmetry was phased at 7 Å resolution (7) in a manner comparable to direct methods by making use of the high symmetry of the system. The phases obtained were in good agreement with those obtained by isomorphous replacement at 9.5 Å resolution (8). However, the use of pseudo-symmetry to predict phases is a technique that is still in its infancy.

Nevertheless, when used in conjunction with isomorphous replacement, it has proved a powerful method. In the case of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), the use of non-crystallographic symmetry was an important factor in the structural analysis. GAPDH crystallises with one tetramer in the asymmetric unit. The electron density over the four subunits was averaged and back transformed to give improved

phase information (9). A similar technique would be feasible with the monoclinic forms of CK.

References for Chapter III

1) Matthews, B.W. (1976) 'The Proteins', 3, Acedemic Press, N.Y.

2) Burgess, A.N., Liddell, J.M., Cook, W., Tweedlie, R.M. and Swan, I.D.A. (1978) J. Mol. Biol., 123, 691

3) Blake, C.C.F. , Swan, I.D.A., Rerat, C., Berthou, J., Laurent, A. and Rerat, B. (1971) J. Mol. Biol., 61, 217

4) Rossmann, M.G. and Blake, D.M. (1962) Acta. Cryst., 15, 24

5) Rossmann, M.G., ed. (1972) 'The Molecular Replacement Method,' Int. Ser. Rev. Ser., No 13.

6) Lattman, E. (1985) Methods in Enzymology, 115, 55

7) Jack, A. (1973) Acta. Cryst., A29, 545

8) Gilbert, P.F.C. and Klug, A. (1974) J. Mol. Biol., 86, 193

9) Beuhner, M., Ford, G.C., Moras, D., Olsen, K.W. and Rossmann, M.G (1974) J. Mol. Biol.,82, 563

CHAPTER IV

ANALYSIS OF NATIVE DIFFRACTION DATA

4.1 Theory

4.1.1 The Atomic Scattering Factor

Consider an atom with electron density at a distance $\bar{r}$ from the centre of the atom give by $p(\bar{r})$. A wave impinging on this atom is scattered in a way dependent on the phase difference between the scattering at the centre of the atom and that at a point $\bar{r}$. If S is defined as the difference between the incoming and scattered waves, $(\bar{s}_o - \bar{s})$ and $|\bar{s}_o| = |\bar{s}| = 1/\lambda$, then the phase difference may be given as $2\pi\bar{r}.\bar{S}$. Thus the total wave scattered by the atom is given by

$$f(\bar{S}) = \int_V p(\bar{r})\exp(2\pi i\bar{r}.\bar{S})dV$$

where V is the volume of the atom (1). $f(S)$ is the atomic scattering factor and, due to the spherical symmetry of an atom, is a real quantity.

4.1.2 Temperature Factor

Above Absolute zero, atoms are in constant thermal motion. The smearing of electron density caused by this motion does not, however, affect the sharpness of the diffraction pattern. Rather, the scattered intensities are reduced by a factor

$$\exp[-B(\sin^2(\theta)/\lambda^2]$$

where $B = 8\pi^2\bar{\mu}^2/3$

where $\bar{\mu}$ represents the mean displacement of the atoms (2). Hence the scattered intensity falls off as a function of angle to a degree defined by the mobility of the atoms in the structure. This has a governing effect on the ultimate resolution attainable in a structure determination.

4.1.3 Scattering by a Molecule

To consider the scattering of an assembly of atoms relative to a fixed origin, one must first describe the scattering of a single atom at a point, r_1 , removed from the origin. This is given simply by

$$\bar{f}_1 = \int_V p(\bar{r})\exp(2\pi i\bar{r}_1 + \bar{r}).\bar{S})dV$$

The scattering from an assembly of atoms is just their vector sum i.e.

$$\bar{G}(\bar{S}) = \bar{f}_1 + \bar{f}_2 + \ldots\ldots\ldots + \bar{f}_N$$

or

$$\bar{G}(\bar{S}) = \sum_{j=1}^{N} f_j \exp(2\pi i\bar{r}_j.\bar{S})$$

where

$$f_j = \int_{V_{atom}} p(\bar{r})\exp(2\pi i\bar{r}.\bar{S})dV$$

4.1.4 Scattering by a Crystal

It has been shown that the scattering of an assembly of atoms can be described by the vector sum of the scattering by its component atoms. In a corresponding fashion, the

scattering from a crystal can be described by the vector sum of its component parts i.e. the scattering from each unit cell. The wave scattered by each unit cell is phase shifted relative to its neighbours by an amount dependent on the cell dimensions and the wavelength. As a result, the total scattered intensity is small. However, when the phase difference between the waves scattered by two unit cells is equal to an integral multiple of 2π, constructive interference occurs. When this condition is summed over all unit cells, an observable intensity is obtained.

Thus, if the unit cell is defined by the vectors $\bar{a},\bar{b},\bar{c}$, the condition for diffraction can be expressed by the Laue equations i.e.

$$\bar{a}.\bar{S} = h$$

$$\bar{b}.\bar{S} = k$$

$$\bar{c}.\bar{S} = l$$

where h,k and l are integers. Thus, the diffraction pattern of a lattice is also a lattice.

If in the previous expressions concerning the position of atom j we now say

$$\bar{r}_j = \bar{a}x_j + \bar{b}y_j + \bar{c}z_j$$

so that $$\bar{r}_j.\bar{S} = x_j\,\bar{a}.\bar{S} + y_j\,\bar{b}.\bar{S} + z_j\,\bar{c}.\bar{S}$$

which from Laue's equations

$$= hx_j + ky_j + lx_j$$

then we obtain

$$\overline{F}(hkl) = \sum_{j=1}^{N} f_j \exp 2\pi i(hx_j + ky_j + lz_j)$$

This is the structure factor equation and represents $\overline{G}(\overline{S})$, the molecular transform, sampled at integral points h,k,l. This expression may be rewritten

$$\overline{F}(hkl) = A + iB$$

where $A = F(hkl)\cos\alpha(hkl)$

and $B = F(hkl)\sin\alpha(hkl)$

The intensities recorded from diffraction data are given by

$$I(hkl) = \overline{F}(hkl).\overline{F}^{*}(hkl)$$
$$= F^{2}(hkl)$$

Thus, although the amplitude of the structure factor can be recovered, the phase angle, α, is lost. This is the classical phase problem of crystallography.

Note that in centrosymmetric structures or centrosymmetric projections, the phase angle is either 0 or π and the structure is real.

4.1.5 The Electron Density Equation

The electron density of the structure is related to the structure factors by

$$p(xyz) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \bar{F}(hkl)\exp -2\pi i(hx+ky+lz)$$

i.e. the diffraction pattern is the Fourier transform of the electron density and vice versa. In the case of projection data eg. hk0

$$p(xy) = \frac{1}{A} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \bar{F}(hk0)\exp -2\pi i(hx+ky)$$

4.1.6 Friedel's Law

Friedels law states that even if the structure lacks a centre of symmetry, its diffraction pattern will always appear centrosymmetric i.e.

$$I(hkl) = I(\overline{hkl})$$

The exception to this arises in the case where anomalous scatterers are present. This then may be used to derive phase information (see Chapter 6)

4.1.7 The Patterson Function

It has been described that the electron density of the structure is the Fourier transform of the diffraction pattern. However, to perform this transform, phase

information is required. The Patterson function (3) is the Fourier transform of F (hkl) and may always be calculated. It is given by

$$P(uvw) = \int_V p(xyz)\ p(x+u,\ y+v,\ z+w)dV$$

i.e. it corresponds to the product of the electron densities at two points separated by a vector $\bar{u} = (uvw)$, summed over the unit cell.

If we make the following definitions

$$\bar{x} = (xyz)$$

and $$\bar{h} = (hkl)$$

then $$p(\bar{x}) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \bar{F}_{\bar{h}} \exp -2\pi i \bar{h}.\bar{x}$$

$$p(\bar{x}+\bar{u}) = \frac{1}{V} \sum F_{\bar{h}} \exp -2\pi i \bar{h}.(\bar{x}+\bar{u})$$

then $$P(\bar{u}) = \frac{1}{V^2} \sum\sum \bar{F}_{\bar{h}}.\ \bar{F}_{\bar{h}} \exp -2\pi i \bar{h}.\bar{u} \int_V \exp -2\pi i(\bar{h}+\bar{h}').\bar{x}\ dV$$

$$= \frac{1}{V} \sum \bar{F}_{\bar{h}} \exp -2\pi i \bar{h}.\bar{u}$$

By Friedel's law $F_h = F_{\bar{h}}$, hence

$$P(uvw) = \frac{2}{V} \sum\sum\sum \overline{F}\,(hkl)\cos 2\pi(hu+kv+lw)$$

This function will assume a large value when $\overline{u}$ corresponds to an interatomic vector. The value according to vector $\overline{u}_{12}$ is equal to that due to $\overline{u}_{21}$ and, therefore, Patterson functions contain a centre of symmetry. As a result, the space group of the Patterson synthesis is different to that of the real structure and may be derived from the latter by the imposition of a centre of symmetry and the removal of translational symmetry.

For a structure containing N atoms, the Patterson map will contain N(N-1)+1 vectors and can thus be very complex. However, it has been used to solve relatively simple structures and is useful in the determination of heavy atom positions (Chapter VI).

4.2 Analysis of CK Native Diffraction Data

4.2.1 CK Native Pattersons

As has been stated, the complexity of Patterson syntheses is such that in the case of native protein data, they are rarely interpretable. However, the conspicuous pseudo-symmetry displayed by the K1 crystal form of CK prompted a belief that the Patterson maps may contain interpretable features. As a result, native self Pattersons were calculated from the projection data available from both crystal forms. The maps obtained are depicted in Figures 4.1 to 4.5.

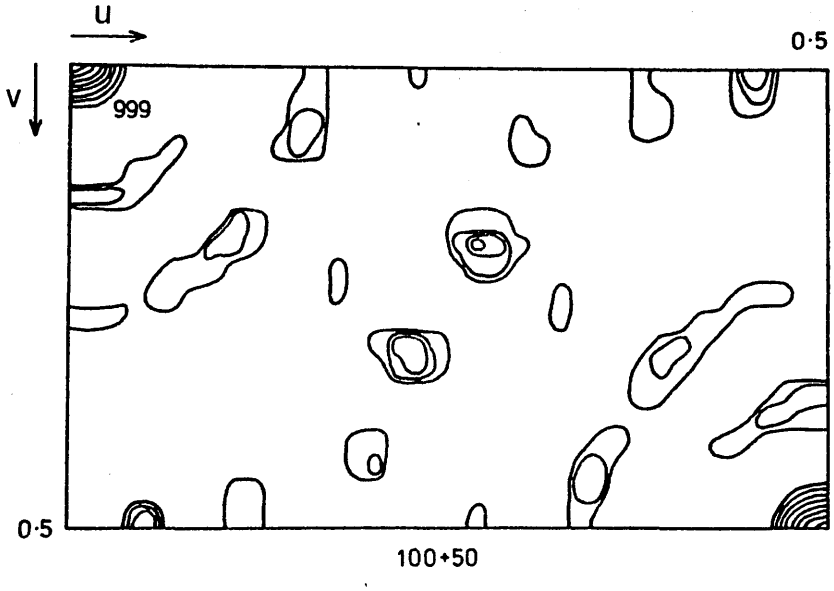


Figure 4.1 : K1 hk0 native Patterson (sin θ/λ = 0.01 - 0.03) (The legend on the x-axis, 100+50, refers to contouring in steps of 50 starting at a minimum of 100 relative to an origin peak of 999. A similar notation is used throughout this thesis for all Patterson maps).

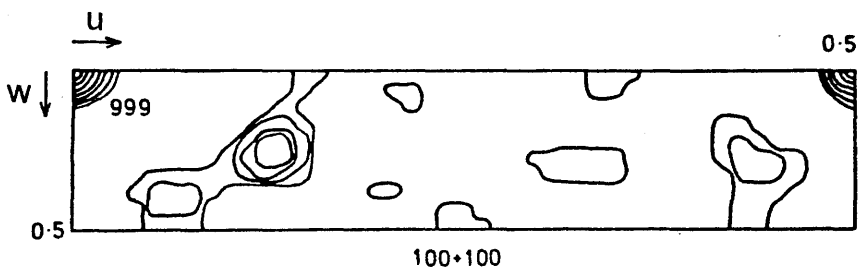


Figure 4.2 : K1 h0l native Patterson (sin θ/λ = 0.01 - 0.03)

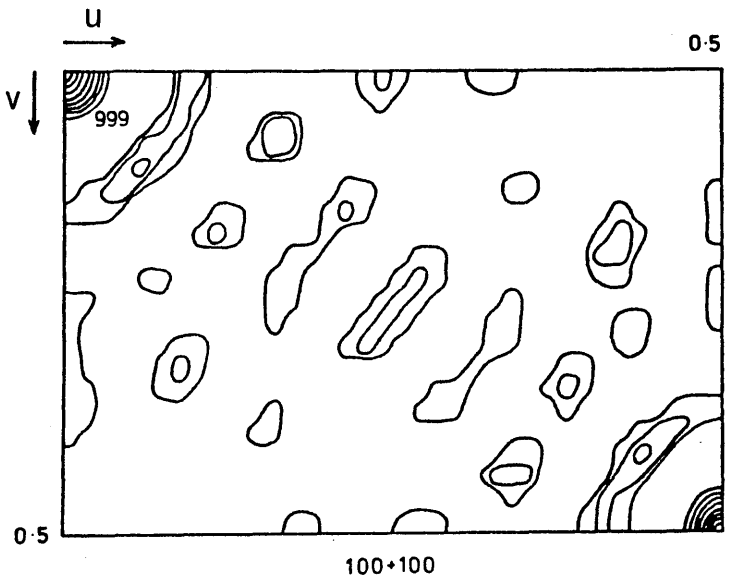


Figure 4.3 : K2 hk0 native Patterson (sin θ/λ = 0.01 - 0.03)

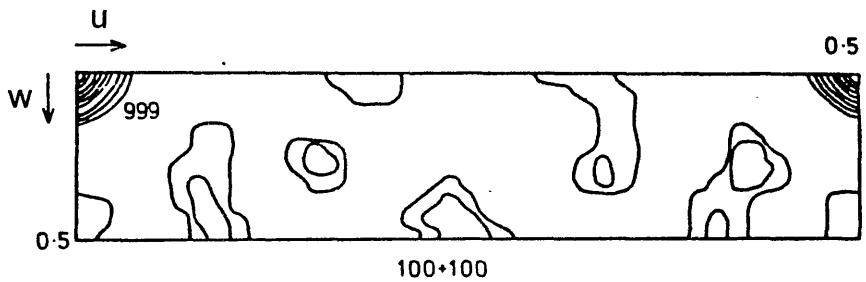


Figure 4.4 : K2 h0l native Patterson (sin θ/λ = 0.01 - 0.03)

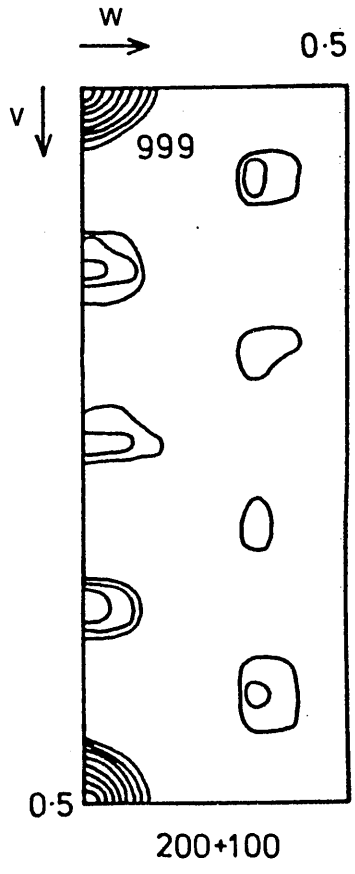


Figure 4.5 : K2 0kl native Patterson (sin Θ/λ = 0.01 - 0.03)

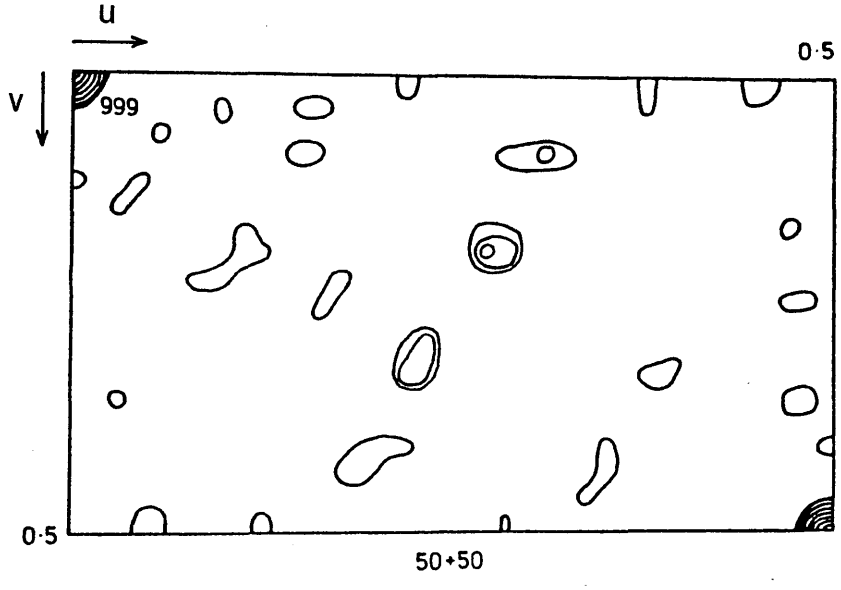


Figure 4.6 : K1 hk0 native Patterson (sin θ/λ = 0.01 - 0.09)

A qualitative comparison of the hk0 and h0l Pattersons for the two crystal forms reveals many similarities. Thus the suspicion that the two crystal forms are closely related is further reinforced.

These Pattersons were calculated using low resolution data only ($\sin\theta/\lambda$ = 0.01 - 0.03). Figure 4.6 depicts the hk0 Patterson for the K1 crystal form incorporating higher resolution data ($\sin\theta/\lambda$ = 0.01 - 0.09). The persistence of many of the features reflects the high degree of pseudo-order present in this crystal form.

4.2.2 Detailed Analysis of Native Pattersons

The equivalent positions in the C2 space group are

$$x \quad , \quad y \quad , \quad z$$

$$\bar{x} \quad , \quad \bar{y} \quad , \quad \bar{z}$$

$$\tfrac{1}{2} + x \quad , \quad \tfrac{1}{2} + y \quad , \quad z$$

and

$$\tfrac{1}{2} - x \quad , \quad \tfrac{1}{2} + y \quad , \quad \bar{z}$$

As has been shown, a Patterson synthesis gives peaks corresponding to vectors between atoms. These fall into two categories viz. self-vectors and cross-vectors. From the equivalent positions it can be shown (Appendix III) that self vectors appear at

$$2x \quad , \quad 0 \quad , \quad 2z$$

$$\tfrac{1}{2} \quad , \quad \tfrac{1}{2} \quad , \quad 0$$

and

$$\tfrac{1}{2} \pm 2x \quad , \quad \tfrac{1}{2} \quad , \quad 2z$$

with cross vectors appearing at

$$x_1 + x_2 \;,\; y_1 - y_2 \;,\; z_1 + z_2$$

and

$$x_1 - x_2 \;,\; y_1 - y_2 \;,\; z_1 - z_2$$

Thus, it is possible to obtain the coordinates of the scattering centres giving rise to the Patterson peaks. At the resolution employed in these Patterson analyses, the scattering centres referred to correspond to the subunit positions of the enzyme in the unit cell.

From the above, it can be seen that for hk0 Pattersons in this space group, self-vectors should be apparent on the y=0 line. Cross vectors should appear in pairs of the same y-coordinate. In this way, the hk0 self Patterson of the K1 crystal form displays the correct general features.

It should be noted that in this projection, the Patterson maps contain a centre of symmetry at x=0.25, y=0.25.

Thus, taking the pair of cross vectors at 0.27, 0.19 (254) and 0.11, 0.19 (192) as a starting point we have,

$$x_1 + x_2 = 0.27$$

and

$$x_1 - x_2 = 0.11$$

therefore

$$x_1 = 0.19$$

and

$$x = 0.08$$

Likewise, taking 0.31, 0.08 (148) and 0.15, 0.08 (182) we have,

$$x_3 + x_4 = 0.31$$

and

$$x_3 - x_4 = 0.15$$

therefore

$$x_3 = 0.08 \ (= x_2)$$

and

$$x_4 = 0.23$$

Thus, the two prominent pairs of cross-vectors can be accounted for by three sites. If we define the site at $x = 0.08$ to be at $y = 0$ (the origin in y is not defined in C2), we obtain sites at

site 1 = 0.08, 0.0

site 2 = 0.19, 0.19

and site 3 = 0.23, 0.08

Inspection of the $y = 0$ line of the Patterson reveals the corresponding self vectors (s.v.) for these sites i.e.

site 1 s.v. at 0.16, 0.0 (117)

site 2 s.v. at 0.38, 0.0 (126)

site 3 s.v. at 0.46, 0.0 (249)

where in all cases, the number in brackets indicates the peak height relative to an origin peak of 999.

In this way, the majority of the most prominent features of the map are accounted for and good self-consistency is clearly apparent. It should be emphasised at this point that the sites given are not uniquely defined and any or all may actually be at a position given by $\frac{1}{2}$-x, $\frac{1}{2}$-y.

It is now necessary to cross check this result with the corresponding h0l Patterson (Figure 4.2). A treatment equivalent to that detailed in the hk0 Patterson analysis, provides good correspondence between the two syntheses adding confidence to the interpretation given. It is possible from this projection also to obtain the z-coordinates of the sites yielding the following positions :

site 1 = 0.08, 0.00 , 0.125

site 2 = 0.19, 0.19 , 0.375

site 3 = 0.23, 0.08 , 0.875

or their equivalents

A similar treatment of the K2 native Pattersons is much less convincing due to the rather diffuse or elongated nature of many of the peaks. Therefore, while it is clear that the two crystal forms are closely related, it is still not possible to quantify this relationship.

4.2.3 The Arrangement of Subunits in the Unit Cell

Reference must now be made to the conclusion of Chapter III. It should be possible to relate the subunit positions chosen from the Patterson analysis to the arrangement required for pseudo-symmetry to become apparent.

Figure 4.7 depicts the x,y projection of the unit cell and the three subunit positions chosen from the Patterson analysis plus their symmetry related equivalents.

It would seem that there appears to be little correspondence with the arrangement predicted in Chapter III (Figure 3.1). If it is assumed that the three subunit positions are as widely spaced as possible so as to fully occupy the asymmetric unit, then

site 1 = 0.08, 0.0

site 2 = 0.19, 0.19

site 3 = 0.27, 0.42 (equivalent of original choice)

would seem a reasonable choice.

This arrangement is depicted in Figure 4.8 where several asymmetric units are included to reveal the super-symmetry. Hence, six-fold symmetry is clearly apparent. However, hexagonal pseudo-symmetry was originally discarded in favour of trigonal pseudo-symmetry due to the h0l diffraction pattern. As was stated in Chapter III, for hexagonal symmetry to apply, the h0l zone should display mm symmetry.

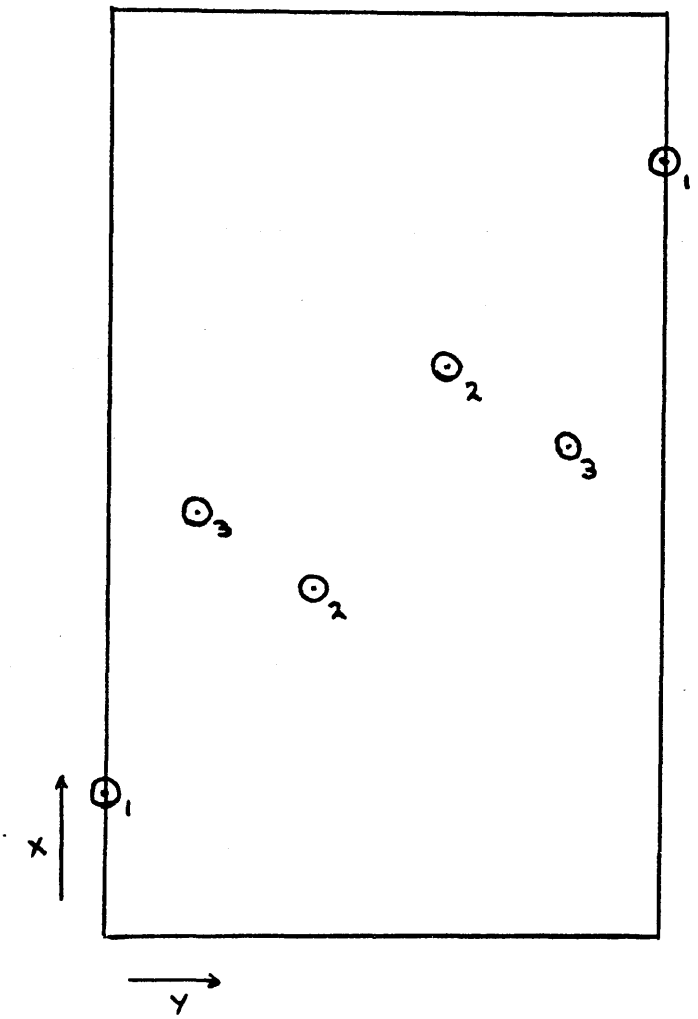


Figure 4.7 : x,y positions of subunits in asymmetric unit of K1 as inferred from native Patterson analysis. Equivalent positions are included.

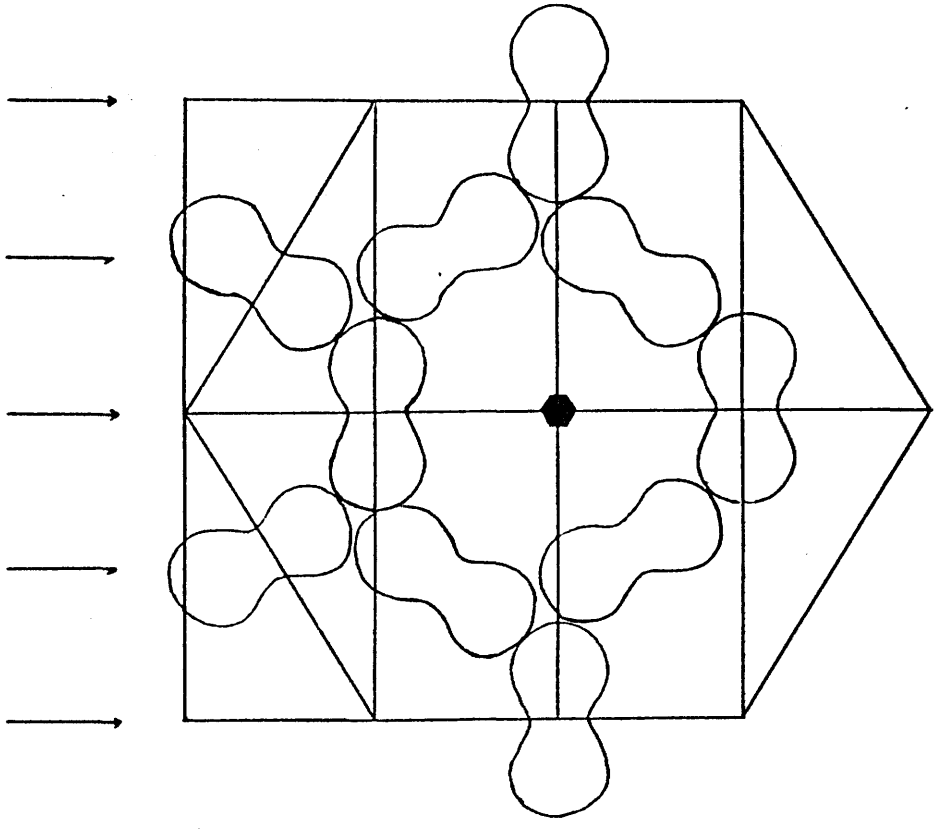


Figure 4.8 : Packing arrangement of creatine kinase subunits in pseudo-hexagonal cell. (cf. Figure 3.1)

Close inspection of the K1 h0l diffraction pattern (Figure 2.6) reveals a line of reflections of high intensity at c.20° to the c axis (vertical in the Figure). This line can similarly be seen in the K2 h0l pattern (Figure 2.3). However, unlike the K2 pattern, the K1 h0l pattern displays an inclination to higher intensity in the reflections related to this line by mirror symmetry. This only constitutes the merest suggestion of mm symmetry. However, this might be anticipated. While the arrangement of subunits in the x,y plane displays fairly good hexagonal symmetry in plan view, when viewed 'edge on' the symmetry is liable to be much less pronounced. This is due to variations in the z-coordinates of the subunits destroying the mirror symmetry. It would be rash to settle on a final interpretation of the pseudo-symmetry of the K1 crystal form. The alternative explanation presented in this chapter does not alter the conclusions regarding subunit homogeniety of Chapter III. It is also clear that K1 is a most unusual crystal form and would be suitable to use for further structural analysis of the enzyme could it be grown in sufficient numbers.

References for Chapter IV

1) Blundell. T.L. and Johnson, L.N. (1976) 'Protein Crystallography', Academic Press.

2) Debye, P. (1914) Ann. Phys., 43, 49

3) Patterson, A.L. (1934) Phys. Rev., 46, 372

CHAPTER V

PREPARATION OF DERIVATIVES

5.1 Introduction

Protein crystals contain channels of mother liquor which act as pathways for the diffusion of additives. Heavy metal salts diffused into a protein crystal can bind to the surface of the protein in such a way as not to alter the conformation of the protein but only to displace some of the mother liquor. The binding may be non-specific and occur through chelation or may be specific. In practice, suitable heavy atom derivatives tend to be found by a process of trial and error although guidelines exist to make the search less arbitrary.

In proteins, the potential ligands capable of binding heavy metals may be regarded as falling into two categories viz. hard and soft. Hard ligands have little tendency to delocalise electron density and take part in electrostatic interactions with metal ions in preference to forming covalent bonds. Such species include terminal carboxylates, hydroxyls of serine and threonine, glutamate and aspartate. Soft ligands, on the other hand, are polarisable and form covalent bonds. The amino acids which fall into this category are methionine, cysteine, cystine and histidine.

Metal ions may be similarly classified according to the type of reaction they tend to undergo. Hence, class (a) metals are defined as those which tend to react preferentially with hard ligands and class (b) as those which react with soft ligands. Generally speaking, class (a) metals are found on the left hand side of the periodic table i.e. the alkali metals, alkaline earths, lanthanides, some of the actinides

and the transition metals up to VA or VIA. The class (b) metals are found to the right hand side of the transition series and are often anionic.

The class (a) metals which have proved most useful have been uranium and samarium although other lanthanides have also been used. As regards class (b) metals, notable success has been achieved with a wide range of mercurials. Silver, platinum, palladium and gold have also frequently been employed as have metals of an intermediate character such as thallium and lead.

5.2 Preparation of CK Derivatives

The higher symmetry present in the much rarer K1 crystal form of CK makes it the more suitable variety with which to contemplate further structural analysis. However, the two crystal forms are indistinguishable by appearance alone and hence, derivatives were prepared without prior knowledge of the crystal variety. It should be added that the extremely poor stability of these crystals with regard to X-ray radiation meant that pre-screening of the crystals by diffraction was not possible.

The fleeting availability of suitable crystals for diffraction studies severely curtailed the screening of heavy atom derivatives. A table of the conditions explored is presented in Table 5.1

Table 5.1 Conditions Employed in Heavy Atom Derivative Search

Derivative	Concentration	Soaking Time	Result
$Hg(CH_3COO)_2$	5 mM	48 hrs	No diffraction
MNP	0.5 mM	48 hrs	No diffraction
MNP	0.04mM	24 Hrs	Good diffraction
K_2PtCl_4	4 mM	48 hrs	fragmentation
K_2PtCl_4	0.04 mM	48 hrs	No diffraction
$Pb(CH_3COO)_2$	0.04 mM	48 hrs	Poor diffraction
Glutaraldehyde	0.5% v/v	24 hrs	Poor diffraction
Glutaraldehyde	0.05% v/v	24 hrs	Good diffraction
Glutaraldehyde + MNP	0.05% v/v 0.04 mM	24 hrs	Good diffraction
$Sm(NO_3)_3$	0.05 mM	24 hrs	Good diffraction
$Sm(NO_3)_3$	0.08 mM	48 hrs	Good diffraction

It had been reported that CK could be titrated with two moles of the thiol specific reagent 2-chloromercuri-4-nitrophenol (MNP) without loss of enzyme activity (1). This observation suggested that MNP could therefore prove a useful heavy atom derivative for X-ray analysis. CK titrated with MNP in the manner described in reference (1) crystallised to produce bright yellow crystals of the K2 form. The diffraction was however, very poor and soaking methods were employed as an alternative. Immersion of native CK crystals in a 0.04mM solution of MNP in the mother liquor for 24 hours gave rise to bright yellow crystals as before. These crystals diffracted well showing observable intensity changes according to the K2 crystal form.

The only other derivative attempted from which satisfactory diffraction could be obtained was samarium nitrate (0.08 mM, 24 hours soak). The data collected from this derivative proved also to be of the K2 crystal variety.

An observation that was made on attempting to prepare derivatives from these crystals was that they were extremely susceptible to fragmentation on exposure to higher concentrations of heavy metal salt than those quoted above. This may be related to the low ionic strength of the crystallisation conditions. There is some evidence that crosslinking the crystals with low concentrations of glutaraldehyde may help to alleviate this difficulty. In all the quoted conditions of derivatisation, the results have to be tempered with the knowledge of considerable variation in crystal quality. Thus, in a batch of 100 crystals, there may

be one of the K1 form and twenty of the K2 form, the remainder diffracting too poorly to be of any use.

Attempts to diffuse the transition state analogue complex into the crytals led to no observable diffraction. The disordering inferred is to be expected for the reasons given in chapter I (2).

Due to the morphology of the crystals, approximate location of the centric projection (h01) was impossible by this method. The large difference in cell dimensions in the directions of the principle axes also gave rise to poor definition of the position of the b axis by Laue photography. These difficulties coupled with the short X-ray exposure lifetime of the crystals led to the decision to screen any derivatives by inspection of the hk0 diffraction pattern. This could be easily obtained by viewing down the needle axis. Having determined some promising derivatives in this way, the intention was then to examine the h01 projections and carry out heavy atom refinements using this centric data. However, the inability to obtain crystals of suitable diffracting quality from subsequent batches of CK defeated these plans. Consequently, only the acentric projection data (hk0) of two potential derivatives were obtained. These were MNP and samarium nitrate (SAM). For the same reasons, there was no opportunity to optimise the derivatisation conditions employed.

The cell dimensions obtained for these two derivatives are given in Table 5.2 where the native cell dimensions are

repeated for comparison. From this it can be seen that both derivatives are reasonably isomorphous with regard to a and b, at least.

Further details of the heavy atom derivative data are shown in Table 5.3. In the case of the SAM data, where anomalous scattering gives rise to the non-equivalence of Friedel pairs, Rsym is calculated for the pairs hkl=$\bar{h}\bar{k}\bar{l}$ and h$\bar{k}$l=$\overline{hkl}$. Anomalous data was not, however, employed and for the purposes of subsequent data analysis, the two data sets were averaged.

Table 5.2 Cell Dimensions of Derivatives

	a (Å)	b (Å)
Native	238.3	164.3
MNP	237.2	164.8
SAM	237.2	163.5

Table 5.3 Summary of Heavy Atom Data Sets

Data Set	No. of unique reflections	Rsym	$(\sin\theta/\lambda)_{max}$	k_{PH}	$\lvert\Delta\bar{F}\rvert/\lvert\bar{F}\rvert$
MNP	104	0.087	0.0959	1.023	0.0731
SAM	142	0.093 0.083	0.0959	1.036	0.0977

$$Rsym = \frac{\sum_h \sum_i (\bar{I}(h) - I(h)_i)}{\sum_h \sum_i I(h)_i}$$

Two values are quoted for SAM due the non-equivalence of Friedel pairs arising from anomalous scattering.

$$k_{PH} = \frac{\sum F^2_{PH}}{\sum F^2_P}$$

$\lvert\Delta\bar{F}\rvert$ is the mean absolute difference in amplitude between the native and derivative data sets and $\lvert\bar{F}\rvert$ is the mean amplitude of the native data set.

References for Chapter V

1) Quiocho, F.E. and Thomson, J.W. (1973) Proc. Nat. Acad. Sci. USA, 70, 2858

2) Blake, C.C.F. (1978) Nature, 271, 707

CHAPTER VI

TREATMENT OF HEAVY ATOM DERIVATIVE DATA

6.1 The Heavy Atom Method

The nature of the crystallographic phase problem was described in section 4.1. The protein crystallographer's main tool in tackling this problem is isomorphous replacement (1). The basis of this method is to introduce an additional atom into the crystal lattice without significant disturbance of the original atomic positions. Under these circumstances, it is possible to observe intensity changes in the diffraction pattern relating to changes in the structure factor amplitudes. These changes will be particularly apparent in the case of an atom with high X-ray scattering power i.e. a heavy atom. The resulting structure factor may be represented as

$$\overline{F}_{PH} = \overline{F}_{P} + \overline{F}_{H}$$

or diagrammatically as in Figure 6.1. If the positions of the heavy atom or heavy atoms in the unit cell can be determined then $\overline{F}_{H}$ can be calculated and information regarding the phase angle, α_{P}, derived. Thus,

$$\alpha_{P} = \alpha_{H} \pm \cos^{-1}\left(\frac{\overline{F}_{PH} - \overline{F}_{P} - \overline{F}_{H}}{2\overline{F}_{P}\,\overline{F}_{H}}\right) = \alpha_{H} \pm \alpha'$$

This phase ambiguity can be graphically illustrated by means of a Harker diagram (2) as shown in Figure 6.2.

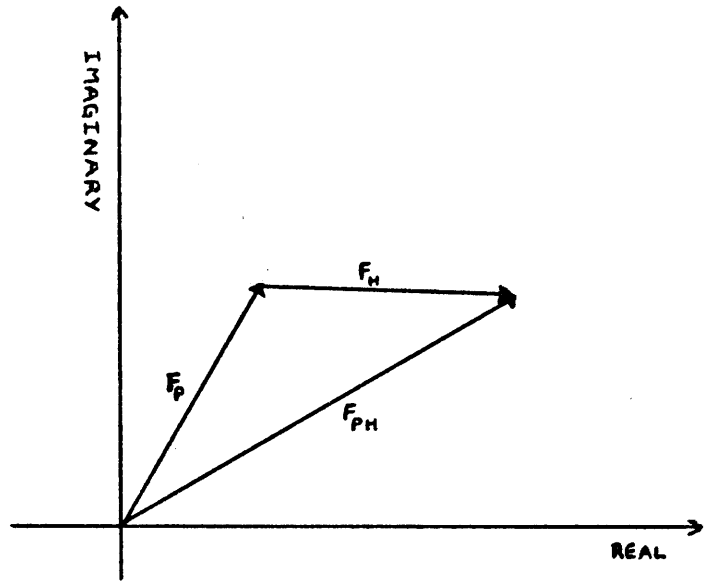


Figure 6.1 : Vector diagramme showing structure factor relationship between native and derivative.

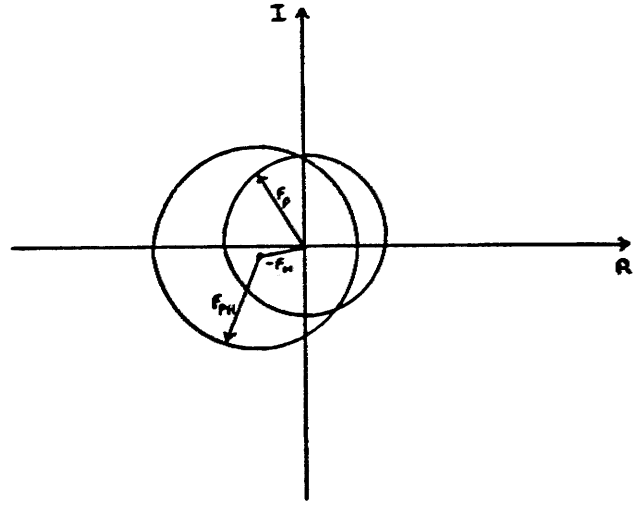


Figure 6.2 : Harker diagramme illustrating phase ambiguity of single derivative in an acentric zone.

Only when $\overline{F}_P$ and $\overline{F}_H$ are colinear is there no ambiguity as

$$\frac{\overline{F}_{PH}^2 - \overline{F}_P^2 - \overline{F}_H^2}{2\,\overline{F}_P\,\overline{F}_H} = 1$$

and hence $\alpha' = 0$.

This special case of colinearity exists when the structure is centrosymmetric and the phases can only assume the values 0 or π. As proteins are made up of L-amino acids only, it is not possible for them to crystallise in centrosymmetric space groups. However, centrosymmetric projections do occur in non-centrosymmetric space groups when viewing down an evenfold rotation axis. Thus in the example of the C2 CK crystal forms described in this thesis, one centric zone is present. If b is taken as the unique axis, this corresponds to the h0l zone.

Outside the special case of centric projections, the phase ambiguity must be resolved by employing more than one derivative. This approach of multiple isomorphous replacement (MIR) can be illustrated for the case of two derivatives by a Harker diagramme as in Figure 6.3.

One could argue that this method should be more accurately described as isomorphous addition rather than replacement. While this is strictly true, the addition of a heavy atom compound at some site on the protein is likely to displace some water molecules.

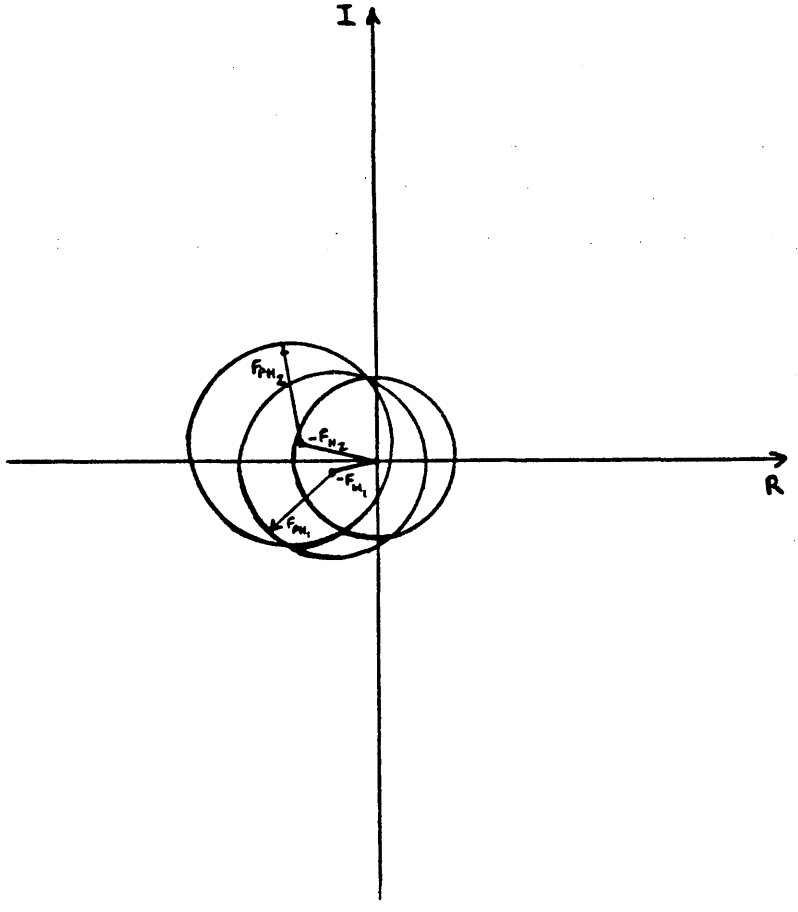


Figure 6.3 : Harker diagramme illustrating the use of multiple isomorphous replacement to resolve phase ambiguity.

Phase information may also be obtained by the use of anomalous scattering (3). This occurs when the X-ray wavelength is in the proximity of an absorption edge of an atom giving rise to an imaginary component of the atomic scattering factor. In this instance there is a breakdown in Friedel's law (4). The observed difference in intensity of centrosymmetrically related pairs of reflections increases with the non-colinearity of $\overline{F}_P$ and $\overline{F}_H$. No anomalous differences occur when the structure is centrosymmetric or in centric zones of non-centrosymmetric space groups. Thus, the techniques of MIR and anomalous scattering are, to some degree, complementary and may be used in conjunction to good effect.

For the case of CuK_α radiation ($\lambda = 1.54$ Å), samarium and gadolinium represent good anomalous scatterers.

6.2 Isomorphous Differences Pattersons

In 6.1 it was shown how, if the heavy atom positions in the unit cell can be determined, phase information may be derived. To determine these heavy atom positions, the technique of isomorphous differences Patterson synthesis is employed.

According to the description of Patterson methods given in Chapter IV, a Patterson function with coefficients

$$(\overline{F}_{PH} - \overline{F}_P)^2$$

will give rise to a map indicating vectors between heavy atoms (5). In fact, other contributions such as protein-

protein interactions will be present to some extent. These other interactions give rise to a noisy background. In the case of multiple heavy atom sites on a large protein, such as CK, the resulting difference Patterson can be very complex.

A Patterson synthesis employing anomalous differences as coefficients also gives rise to a heavy atom vector map although such maps tend to be rather noisy (6). In general, when anomalous differences are available, combinations of these and the isomorphous differences are used (7,8,9).

6.3 Difference Fourier Techniques

If, by either MIR or anomalous differences, some estimate of the protein phases is available, then a Fourier synthesis employing the coefficients

$$m(\bar{F}_{PH} - \bar{F}_{P})\exp(i\alpha_{P})$$

may be calculated (10). These coefficients give the vector contribution of the heavy atom in the direction of the protein structure specified by α_P. The coefficients are weighted by the figure of merit, m. This parameter represents the mean value of the cosine of the error in the phase angle. In this way a good estimate of α_P would correspond to $m > 0.8$.

Difference Fouriers of this sort have to be used with considerable care, however. Fouriers are much more sensitive

to phase information than they are to intensities (11). Thus, if a difference Fourier is calculated for the same derivative which gave rise to the phase information, the resulting map is severely biased. This problem is overcome by using the phase information from other derivatives to produce so-called cross-phase Fouriers (12).

As a check, it is important that at all stages of the determination of multiple heavy atom sites, each new site is confirmed by supporting evidence from the difference Patterson synthesis for that derivative.

Another type of Fourier analysis which may prove useful in the determination of multiple site derivatives is the double difference Fourier synthesis. In this case, the coefficients employed are given by

$$m(F_{PHobs} - F_{PHcalc})\exp(i \alpha_P)$$

A map of this type has the property of showing up additional minor sites (13).

6.4 Estimation of $\overline{F}_H$

In order to calculate phases in the manner described, it is necessary to obtain a good estimate of $\overline{F}_H$. In the case of centric zones, where $\overline{F}_{PH}$ and $\overline{F}_P$ are colinear, $\overline{F}_H$ may be simply given as

$$F_H = \left| F_{PH} \pm F_P \right|$$

Usually, the difference is assumed although in the case of very small values of $\overline{F}_{PH}$ and $\overline{F}_P$, 'crossing over' can occur. The situation is more complex in acentric zones. If the deviation from colinearity is great, the estimate used for centric data will be a poor one. North (7) pointed out that the complementarity of isomorphous differences and anomalous differences can be used to estimate F_H. Simply,

$$F_H \simeq \left(\Delta F^2_{iso} + \Delta F^2_{ano} \right)^{\frac{1}{2}}$$

The exact equation given by Singh and Ramaseshan (14) reduces to

$$F^2_{HLE} = F^2_M + F^2_P - 2 \left[F^2_M F^2_P - \left(\frac{k^2}{4} \right) \Delta I^2 \right]^{\frac{1}{2}}$$

where $$F^2_M = \tfrac{1}{2} \left[F^2_{PH}(+) + F^2_{PH}(-) \right]$$

and $$k = \frac{F_H}{F''_H}$$

and HLE stands for heavy atom lower estimate.
If no anomalous differences are available, this expression reduces to that for centric zones.

6.5 Refinement

The heavy atom positions revealed by difference Patterson analysis and cross phase Fouriers will not give rise to the best phase information obtainable. This is because their exact positions need to be optimised in conjunction with their respective occupancies and temperature factors. The way this is accomplished is by least squares refinement of the differences between F_H calculated and F_H observed. Thus, in the Hart technique (15) the quantity minimised is

$$E = \sum [(F_{Hcalc} - F_{Hobs})]^2$$

$$= \sum_n [z_j f_{oj} t_j \cos 2\pi(hx_j + lz_j) \pm F_P \pm KF_{PH}]^2$$

where n refers to one h0l reflection

z_j is an effective atomic number

f_{oj} is a unitary scattering factor for heavy atom j

t_j is the temperature factor

x_j and z_j are the heavy atom coordinates

and K is the scaling factor

The parameters are varied one or two units in each direction according to the stage of the refinement. This trial and error method avoids the pitfalls of settling in a local

minimum which can occur in conventional least squares methods.

As a monitor of the success of the refinement procedure, the convention is to calculate the reliability index,

$$R(F_H) = \frac{\sum |(F_{Hobs} - F_{Hcalc})|}{F_{Hobs}}$$

If anomalous differences are available then $R(F_{HLE})$ is calculated in an equivalent fashion.

6.6 CK Heavy Atom Derivatives

6.6.1 Data Processing

Data collection for the CK heavy atom derivatives was achieved by screened precession photography as described for the native data sets in Chapter II. However, in the case of the heavy atom work, a PDP 11/45 minicomputer was used for the data processing. For this stage of the analysis the Oxford program suite was employed which is composed of the following programs.

FLP - Corrects data for Lorentz polarisation.

RASCAL - Creates a heavy atom file for one derivative, scaled to the native over an optional number of equal sin θ ranges.

MERGE - Adds a new heavy atom derivative to an existing Heavy File. Both files will have been created using either RASCAL or MERGE.

The format of the Heavy File is hkl, $\sin^2\theta$, phase, scattering factor, figure of merit, F_P and then the heavy atom data. For each derivative, there is a column for F_{PH}, Fobs and Fobs - Fcalc. At this time the heavy files did not contain anomalous data.

EXTRAC - Selects appropriate data from a Heavy File or a Phase File and re-formats the data for input to FORDAP.

PANG, ACE - These are refinement programs in which each parameter is varied by ±1 or ±2 increments. Corresponding values of F_H are calculated and are used to build up error sums of the form $(\text{Fobs} - \text{Fcalc})^2$ for the shifted as well as the unshifted parameters. After a pass through all the data, the error sums are compared with the starting sum to determine whether any of the parameters are to be adjusted. Increments are increased or decreased according to which summation is smallest for a given parameter. ACE proceeds by the same method to refine acentric data.

PHASE - Phases are calculated by the method of Blow and Crick (16).

FORDAP - Calculates Pattersons and Fouriers.

In the refinement of the CK heavy atom data, the coordinates of the heavy atom sites were adjusted in increments of

0.002Å while the temperature factors were kept fixed at 6.0 Å². The occupancy and scale factors were allowed to vary in increments of 0.05 and 0.02 respectively. Six cycles of refinement were usually sufficient at each stage and in all cases, data in the resolution range 7.5-15 Å were employed.

6.6.2 Location of Initial Sites

As discussed in Chapter IV, the general positions of the vectors obtained in a Patterson synthesis in the C2 space group are given by self vectors at

$$\pm 2x,\ 0,\ 2\bar{z}$$

with cross vectors at

$$x'+x'',\ y'-y'',\ z'+z''$$

$$x'-x'',\ y'-y'',\ z'-z''$$

The difference Patterson syntheses for the MNP and SAM derivatives are depicted in Figures 6.4 and 6.5 respectively. Both maps exhibit large self vectors on the y=0 line as expected, however, the positions of the main self vector for each derivative are very close. In fact, the two maps display significant similarities indicating that they may share common or near common sites, at least in this projection.

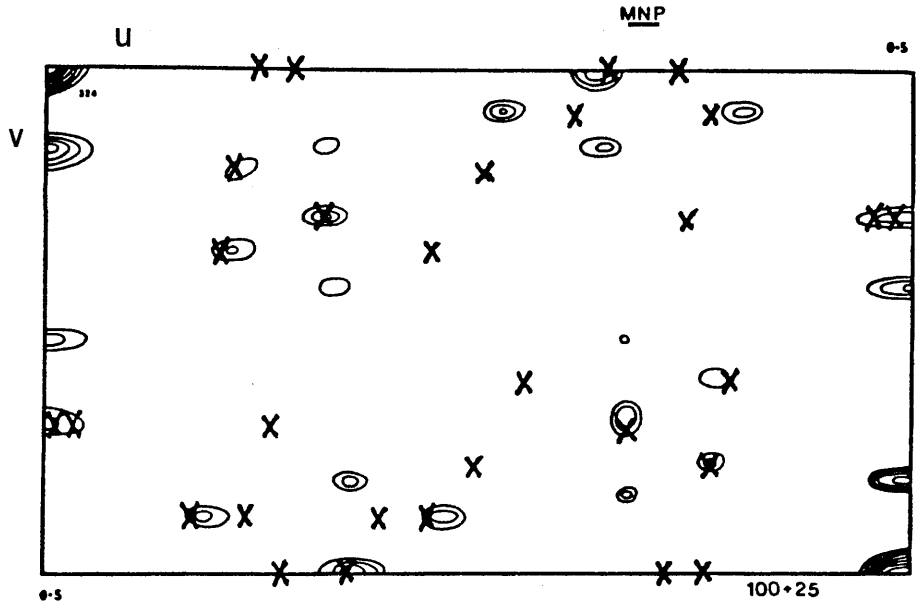
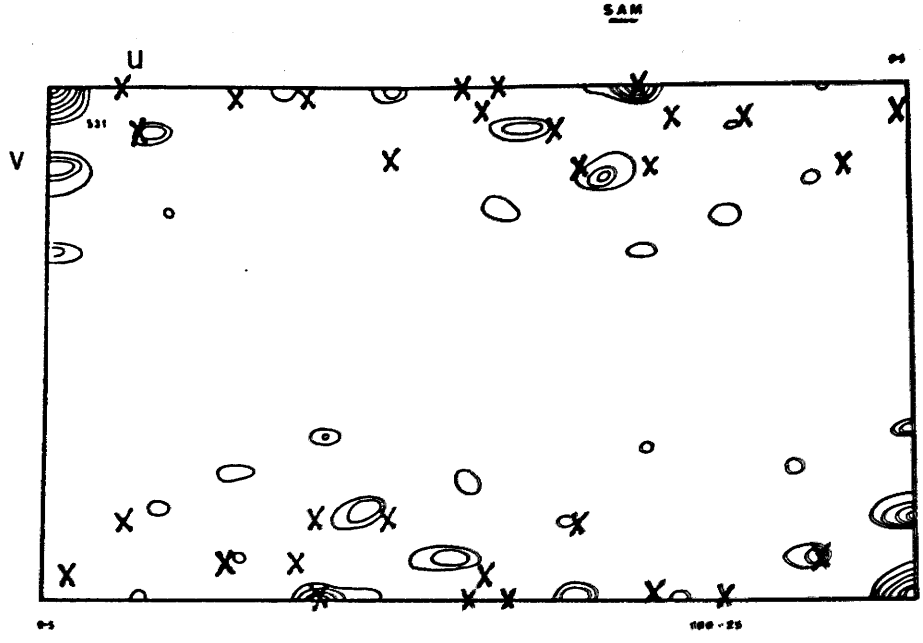


Figure 6.4 : MNP hk0 difference Patterson. The crosses refer to the positions of the cross vectors and self vectors relating to the heavy atom sites obtained after the refinement process.



Figures 6.5 : SAM hk0 difference Patterson. The crosses refer to the positions of the cross vectors and self vectors relating to the heavy atom sites obtained after the refinement process.

To obtain a starting set of phases, the structure was considered to be centrosymmetric. This can be justified in the following way. The structure factors for C2 are given by

$$A = 4\cos^2 2\pi \frac{h+k}{4} \cos 2\pi(hx+lz)\cos 2\pi ky$$

$$B = 4\cos^2 2\pi \frac{h+k}{4} \cos 2\pi(hx+lz)\sin 2\pi ky$$

Now, $$F = \sqrt{(A^2 + B^2)}$$

and $$\cos^2 2\pi ky + \sin^2 2\pi ky = 1$$

Therefore, y has no effect on F. This means that the origin with respect to y can be shifted so as to create a centrosymmetric arrangement so long as there is only one heavy atom site. In this way a starting set of phases can be calculated for each derivative based on their main sites as indicated by the difference Pattersons.

The method employed for the analysis of difference Pattersons is equivalent to that described in Chapter IV for self Pattersons and will not be repeated here. MNP site 1 was chosen from the difference Patterson to be at 0.16, 0. The parameters yielded upon refinement are given in Table 6.1. The centric phases produced were used to compute a SAM cross-phase difference Fourier. The resulting map shown in Figure 6.6 shows a large number of peaks corresponding to possible sites for both derivatives. This kind of ghosting

Table 6.1 : MNP 1 Site Parameters

x	y	occupancy
0.1650	0.0000	0.5600

Mean figures of merit

N	Centric	N	Acentric
4	0.44	61	0.24

N = number of reflections.

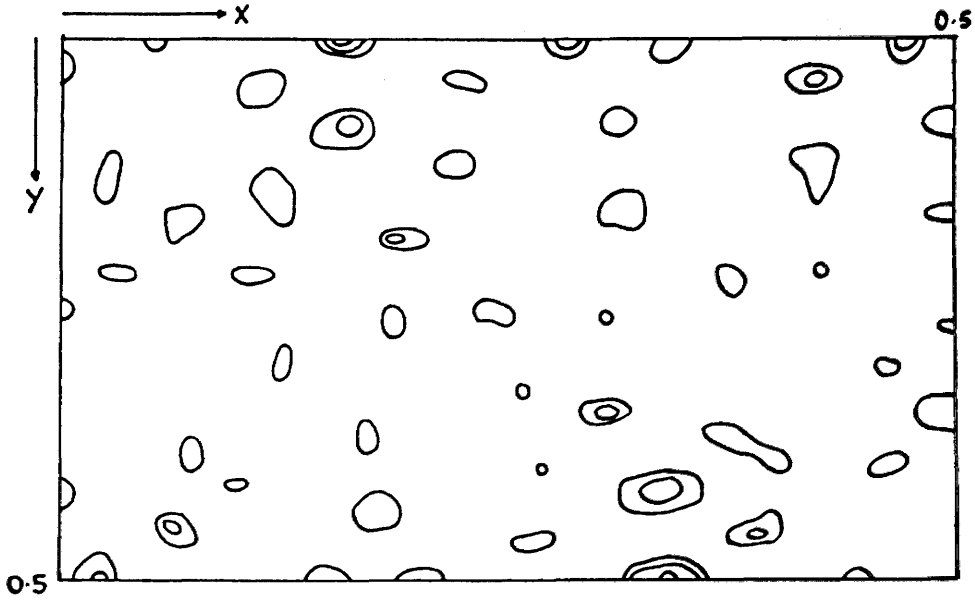


Figure 6.6 : SAM cross-phased Fourier using MNP site 1.

is common in cross-phase difference Fouriers but is unusually apparent in this case. A strong centre of symmetry is also present due to the use of centric phases.

This was considered to be an unsatisfactory approach to the SAM data and SAM site 1 was chosen from the SAM difference Patterson and refined in the same manner as MNP. Thus, refinement of site 1 at 0.17, 0 yielded the parameters given in Table 6.2. On comparison, it can be seen that at this stage, SAM appears the better of the two derivatives.

6.6.3 The Correlation of Origins

In the C2 space group, the origin is not defined with respect to the y axis. Consequently, in the interpretation of hk0 difference Pattersons, the origin with respect to y is fixed arbitrarily on the basis of the first heavy atom site chosen. Thus, to relate the information available from two different derivatives, their sites must be expressed with respect to the same origin.

When no phase information is available, correlation can be achieved by Patterson functions relating the two derivatives. Rossmann (17) suggested a synthesis using the coefficients

$$(F_{PH_2} - F_{PH_1})^2$$

Table 6.2 : SAM 1 Site Parameters

x	y	occupancy
0.1750	0.0000	0.7250

Mean figures of merit

N	Centric	N	Acentric
4	0.51	61	0.41

N = number of reflections.

This gives positive peaks corresponding to self vectors for each of the derivatives and negative peaks for cross vectors. In the case of the CK MNP and SAM derivative data, this synthesis gave rise to a map which was uninterpretable. This will probably often be the case where multiple site derivatives are concerned.

A much more readily interpreted map was supplied using the coefficients proposed by Steinrauf (18)

$$(F_{PH_1} - F_P)(F_{PH_2} - F_P)$$

This gives rise to a map in which cross vectors are represented by positive peaks. Self vectors within each derivative are suppressed so simplifying the interpretation.

The result of the Steinrauf correlation function applied to the CK heavy atom data is displayed in Figure 6.7. The presence of a large peak at 0,0 and a second large peak at 0.32,0 indicates common sites, as suspected. It is possible that this commonality exists only in this projection. However, the result is that the refinement process is hampered due to the similarity in the phase information supplied by the two derivatives. The abnormally high appearence of ghosting in the cross-phase difference Fourier is thus accounted for.

It can be seen from the complexity of the difference Pattersons and from the correlation function that both

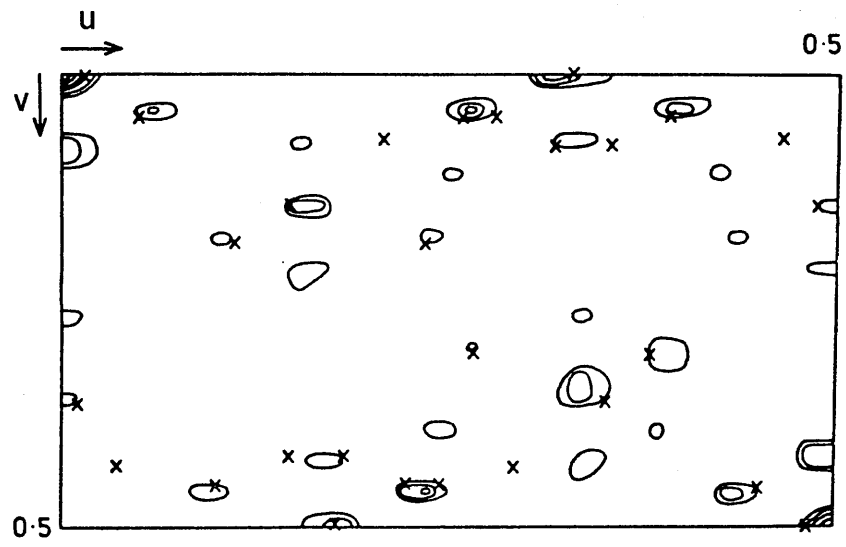


Figure 6.7 : MNP, SAM correlation function map (Steinrauf). The crosses refer to the cross vectors between MNP site 1 and all SAM sites and SAM site 1 and all MNP sites obtained after the refinement process.

derivatives contain multiple sites. Bearing in mind the difficulties already mentioned, it was decided that the refinement process for the two derivatives be conducted in tandem. In this way, the phases were calculated on the combined derivative data at each stage of the refinement. In the following account of the refinement procedure, the phase information obtained upon the location of each new site is reflected in the name given to that phase set. For example, phase set M2S2 corresponds to the phases obtained upon refinement of two MNP sites and two SAM sites.

6.6.4 M1S1 Phases

The combined phase set M1S1 was calculated from parameters obtained on refinement of the MNP site 0.16, 0 and the SAM site 0.17, 0. The details of this refinement are presented in Table 6.3. The M1S1 phases were then used to calculate difference and double difference Fouriers for each of the derivatives. In general, it was found that the double difference Fouriers were of greater use in the location of subsequent sites. It is these double difference Fouriers that will therefore be shown at each stage of the analysis. Comparison of the M1S1 phased SAM and MNP double difference Fouriers (Figures 6.8 and 6.9) indicates the presence of many common sites due to ghosting.

Table 6.3 : M1S1 Phase Set

Derivative	x	y	occupancy
MNP	0.1655	0.0000	0.5625
SAM	0.1739	0.0000	0.6825

Mean Figures of Merit

N	Centric	N	Acentric
4	0.55	61	0.45

N = Number of reflections.

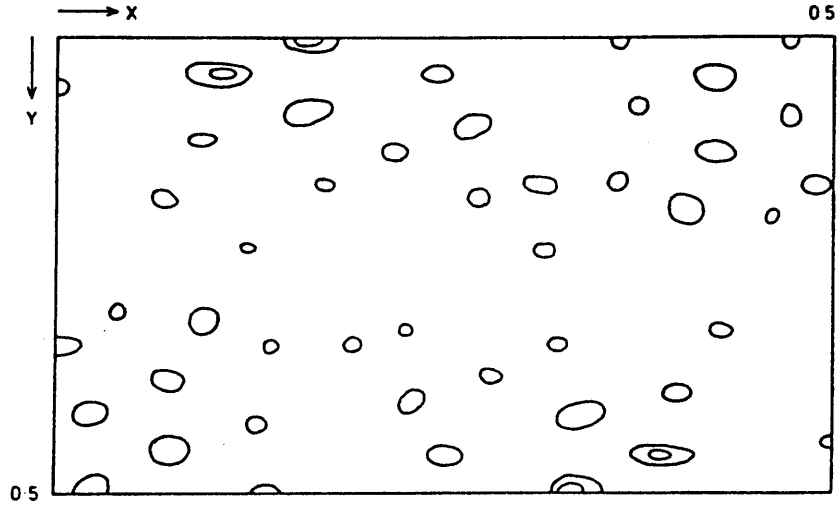


Figure 6.8 : SAM double difference Fourier based on M1S1 phases.

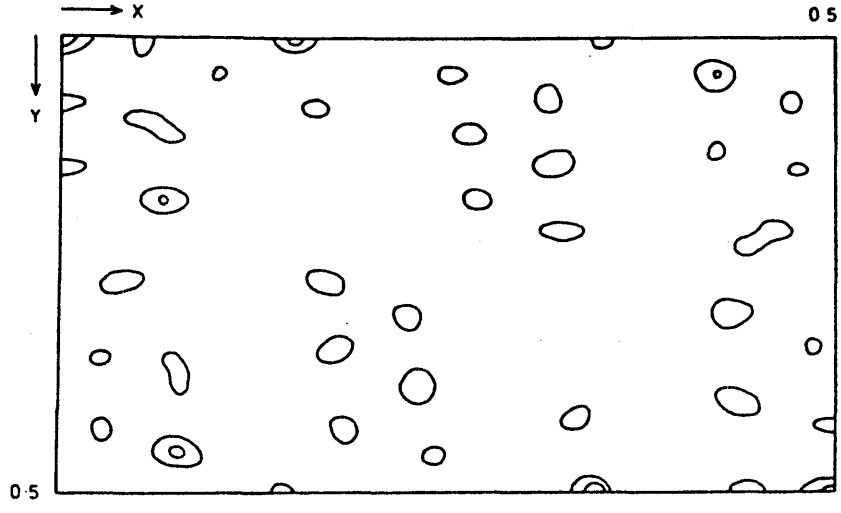


Figure 6.9 : MNP double difference Fourier based on M1S1 phases.

A strong centre of symmetry is also apparent as is to be expected at this early stage of the refinement. Considerable care was therefore exercised in the choice of the second sites for each derivative. This was accomplished by reference to the difference Pattersons from which strict agreement was sought.

By this method, MNP site 2 was chosen at 0.432, 0.048. Refinement of MNP sites 1 and 2 led to a residual of 0.623. Correspondingly, SAM site 2 was located at 0.115, 0.038 and refinement led to a residual of 0.554.

6.6.5 M2S2 Phases

The parameters for this phase set are given in Table 6.4. It can be seen that the mean figure of merit for acentric phases has improved from 0.45 to 0.59. As before, double difference Fouriers were calculated for the two derivatives (Figures 6.10 and 6.11). It can be seen that while the two maps now have fewer common sites than with the M1S1 phase set, the centre of symmetry still persists in them both. As a result, care was taken to compare the performance of each possible new site with its centric equivalent. This was done by comparison of the residuals obtained after refinement. Thus the choice of 0.065, 0.185 as the third MNP site gave a residual of 0.518 whereas the equivalent site at 0.435, 0.315 led to a residual of 0.589. No other site gave rise to a residual as low as 0.518 and so, upon confirmation from the MNP difference Patterson, 0.065, 0.185 was taken as MNP site 3.

Table 6.4 : M2S2 Phase Set

Derivative	x	y	occupancy
MNP	0.1643	-0.0050	0.5175
	0.4308	0.0492	0.5300
SAM	0.1739	0.0000	0.6650
	0.1155	0.0399	0.3800

Mean Figures of Merit

Centric	Acentric
0.57	0.59

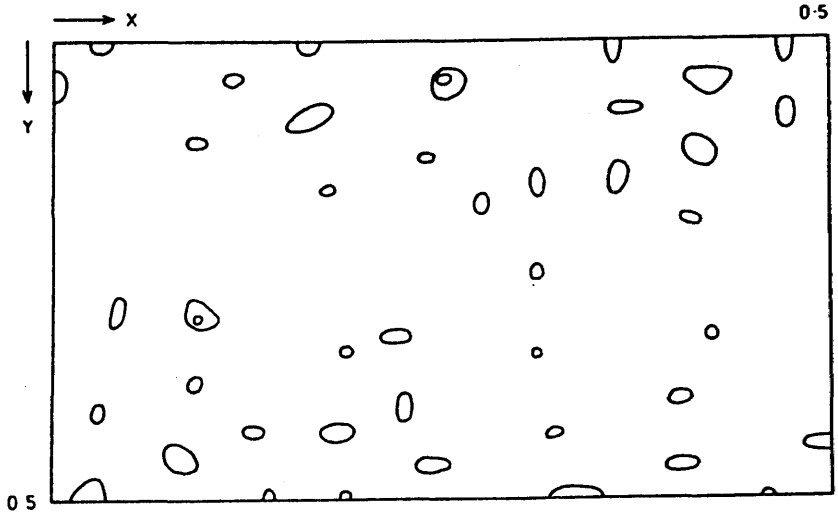


Figure 6.10 : SAM double difference Fourier based on M2S2 phases.

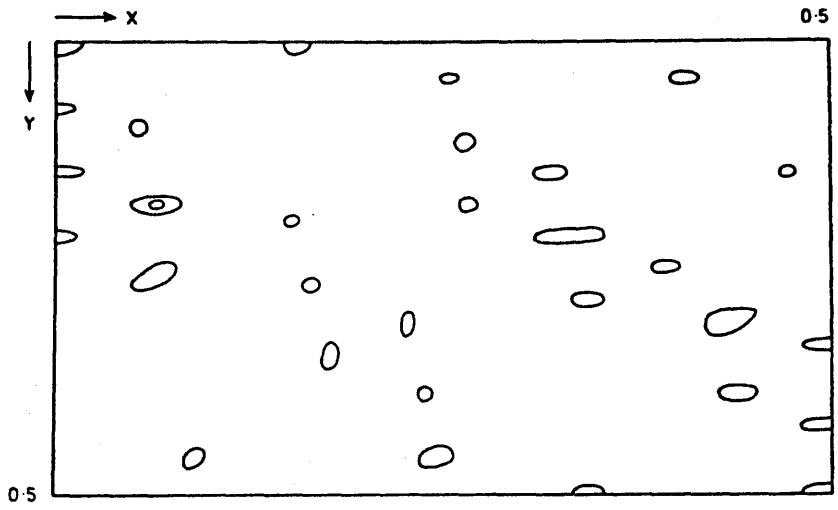


Figure 6.11 : MNP double difference Fourier based on M2S2 phases.

The M2S2 SAM double difference Fourier displays a greater number of possible minor sites than the MNP map. Refinements were carried out for every site that tallied with the SAM difference Patterson plus their centric equivalents. On this basis, the third SAM site was chosen to be at 0.479, 0.076. The residual obtained was 0.479 compared to 0.533 for its centric equivalent.

It will be noted that the site chosen is not the largest peak on the M2S2 SAM map. This is actually located at 0.26, 0.03. However, refinement of this site led to a residual of 0.516. It was observed that alternative sites also showed less discrimination from their centric equivalents on refinement. For instance, the equivalent position of the site mentioned above led to a residual of 0.535 compared to 0.516. It was felt that the poorer discrimination between equivalent sites as monitored by the residuals was an indication of their possibly spurious origin.

6.6.6 M3S3 Phases

As before, the new MNP and SAM sites were included in the phase calculation to produce a new phase set. The parameters for this M3S3 phase set are given in Table 6.5. It will be observed that the acentric figure of merit has further risen from 0.59 to 0.66.

Double difference Fouriers calculated for the two derivatives are shown in Figures 6.12 and 6.13. As can be seen, the centre of symmetry is now much less prominent in

Table 6.5 : M3S3 Phase Set

Derivative	x	y	occupancy
MNP	0.1637	-0.0050	0.5125
	0.4295	0.0467	0.5075
	0.0637	0.1850	0.3250
SAM	0.1734	0.0000	0.7350
	0.1165	0.0419	0.4150
	0.4794	0.0766	0.3141

Mean Figures of Merit

Centric	Acentric
0.67	0.66

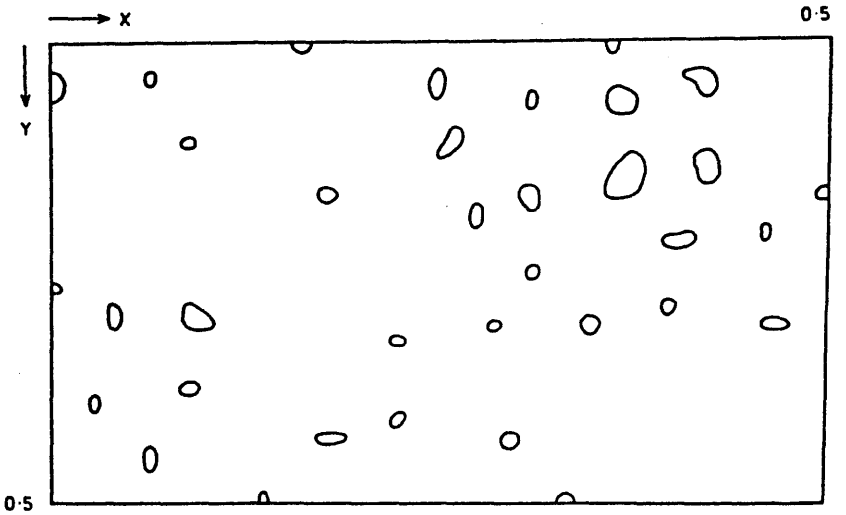


Figure 6.12 : SAM double difference Fourier based on M3S3 phases.

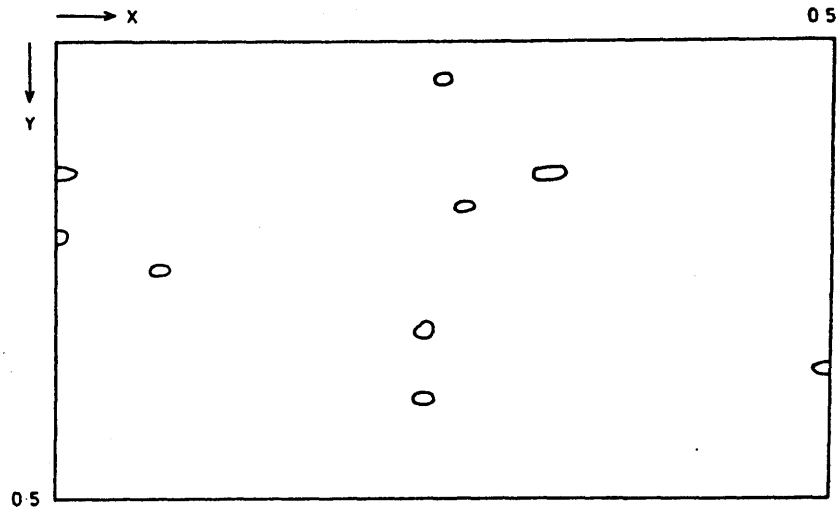


Figure 6.13 : MNP double difference Fourier based on M3S3 phases.

both maps. The MNP double difference Fourier shows considerably fewer features with the new phases while the SAM map is less close to a final solution. As a result, a new phase set was calculated based on an additional MNP site only. This site was chosen as 0.319, 0.143

6.6.7 M4S3 Phases

The details of this phase set are given in Table 6.6. A SAM double difference Fourier using these phases is shown in Figure 6.14. It can be seen that the improved phases have caused the enhancement of one peak in the SAM map relative to the previous phase set. This site, on refinement, led to a residual of 0.420. However, a residual of 0.403 was obtained for its centric equivalent. Several possible alternative sites gave residuals of between 0.4 and 0.45 and a decision on the the position of the fourth SAM site was eventually made on the basis of that which corresponded most favourably with that SAM difference Patterson.

In this way, SAM site 4 was located at 0.371, 0.069. The residual obtained was 0.416 compared to 0.455 for its centric equivalent.

Table 6.6 : M4S3 Phase Set

Derivative	x	y	occupancy
MNP	0.1620	-0.0050	0.4525
	0.4285	0.0450	0.4825
	0.0611	0.1859	0.4500
	0.3192	0.1428	0.3075
SAM	0.1734	0.0000	0.8087
	0.1005	0.0479	0.3325
	0.4770	0.0806	0.3153

Mean Figures of Merit

Centric	Acentric
0.75	0.72

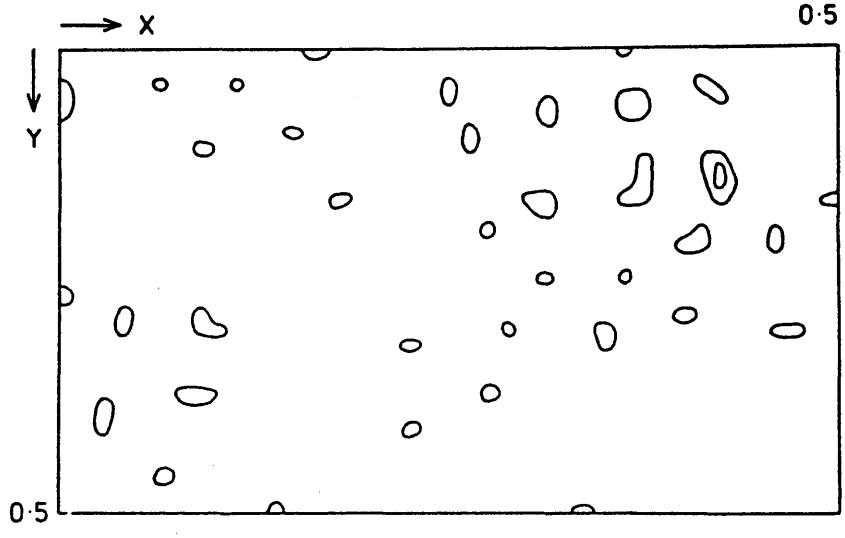


Figure 6.14 : SAM double difference Fourier based on M4S3 phases.

6.6.8 M4S4 Phases

The parameters for this final phase set are given in Table 6.7. The acentric figure of merit is 0.76.

An MNP double difference Fourier based on these phases (Figure 6.15) is virtually featureless and it was felt that a reasonable solution of this derivative had been obtained. The SAM map (Figure 6.16) on the other hand, still indicates the possible presence of minor sites. However, it was felt that little useful phase information could be derived from any further analysis of this data.

By way of final confirmation of the sites obtained from the two derivatives, cross vectors between them were checked against the correlation function map (Figure 6.7). The crosses on the map correspond to vectors between MNP site 1 and all four SAM sites and also vectors between SAM site 1 and all four MNP sites. It can be seen that these vectors account for the majority of the peaks on the map. None of the vectors fall in negative regions of the map. Cross vectors between minor sites for both derivatives also agree, in the main, with the correlation function map. These vectors have been omitted for clarity.

The individual difference Pattersons (Figures 6.4 and 6.5) also bear crosses indicating all the cross vectors between sites obtained for each respective derivative. Again the majority of the peaks are accounted for and no crosses fall on negative regions of the map.

Table 6.7 : M4S4 Phase Set

Derivative	x	y	occupancy
MNP	0.1620	-0.0050	0.4525
	0.4285	0.0450	0.4825
	0.0611	0.1859	0.4500
	0.3192	0.1428	0.3075
SAM	0.1734	0.0000	0.6837
	0.1185	0.0449	0.4725
	0.4795	0.0756	0.3953
	0.3709	0.0686	0.3203

Mean Figures of Merit

Centric	Acentric
0.75	0.76

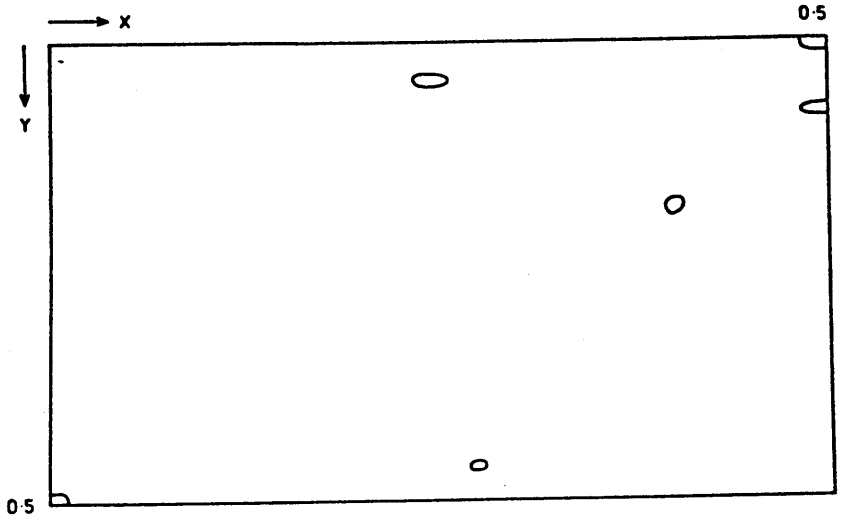


Figure 6.15 : MNP double difference Fourier based on M4S4 phases.

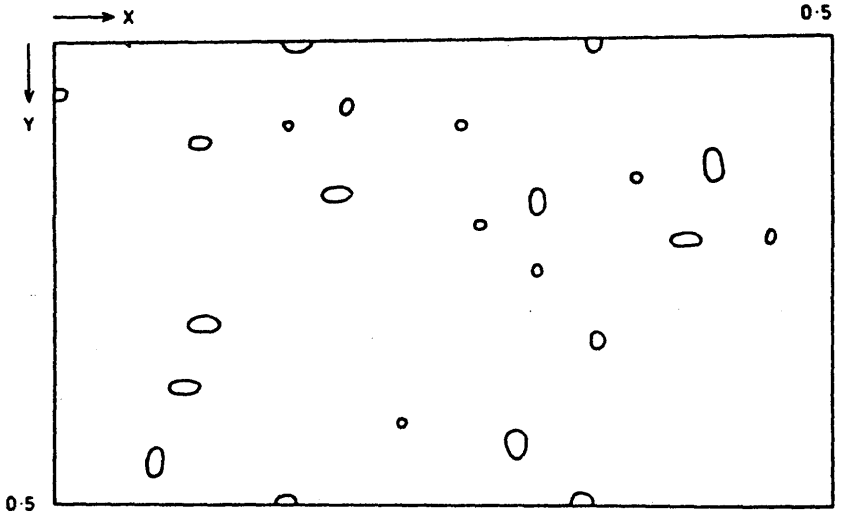


Figure 6.16 : SAM double difference Fourier based on M4S4 phases.

As a summary, the progress of the refinement process as monitored by the residual obtained is tabulated in Table 6.8. The phases calculated at each stage of the refinement are given in Appendix IV.

The refinement procedure detailed in this Chapter was the most successful of several variations attempted. It is described in some detail in order to highlight the difficulties encountered with respect to choice of minor sites and the breaking of the initial centre of symmetry.

In conclusion, it is felt that MNP could prove a useful derivative. SAM, on the other hand, is less encouraging though may prove useful at low resolution. The problems of dealing with acentric projection data and the unfortunate correspondance of the main sites for both MNP and SAM are considerable. As a result, the final outcome of the refinement process described is open to question especially with regard to minor sites.

What can be said is that both derivatives are worthy of consideration for any future structure analysis if the crystallisations can be more successfully repeated. In this case there is no doubt that the derivatives would benefit from analysis in a centric projection or better still, from the use of three dimensional data.

Table 6.8 : Refinement Progess

Derivative	no. of Sites	Residual
MNP	1	0.679
	2	0.623
	3	0.518
	4	0.367
SAM	1	0.607
	2	0.554
	3	0.479
	4	0.416

6.7 CK Native Electron Density Map

Using the M4S4 phases, a native electron density map was calculated for the K2 hk0 projection of CK. Figure 6.17 depicts one asymmetric unit of this map. Electron density was calculated at intervals of 2.96 Å in a and 2.73 Å in b and the map was contoured in equal but arbitrary intervals. Dotted contours represent negative regions of the map. As before, all data in the range 7.5 to 15 Å were employed.

The map reflects the poor quality of the phase information available with a degree of 'rippling' being apparent. This has the effect of producing negative regions such as in the top left hand corner of the map where two large peaks give rise to a negative region between them.

At this resolution, using only projection data, the only detail that one can hope possibly to derive is the boundary of the individual molecules. However, even this low level detail is not readily apparent from Figure 6.17. A slightly better impression may be gained from inspection of Figure 6.18 where an electron density map of the complete unit cell is reproduced.

In order to assist in the interpretation of this map, it is helpful, firstly, to estimate the likely size of a CK subunit. Assuming a partial molar volume for CK of 0.74 cm^3/g and a subunit molecular weight of 43,000 Daltons, this yields a molecular volume of 5.30×10^4 $Å^3$. Thus, assuming the CK subunit to be spherical, this gives rise to an equivalent diameter of approximately 47 Å. As this figure is very close to the unit cell thickness of both crystal

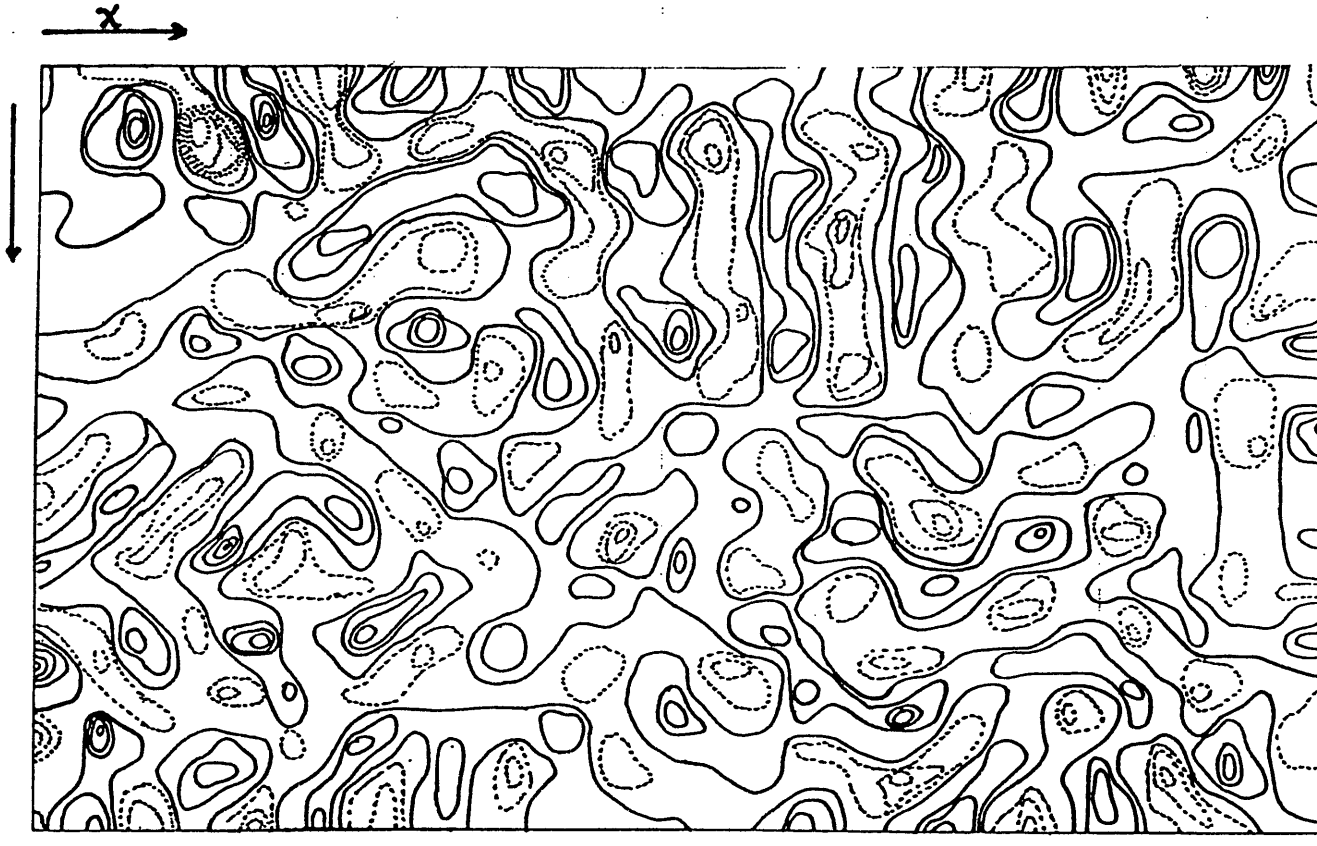


Figure 6.17 : Electron density map of one asymmetric unit of the hk0 projection of the K2 form of CK calculated using the M4S4 phase set. The map was calculated at intervals of 2.96 Å in a and 2.73 Å in b. Contouring is in equal but arbitrary intervals. Dotted contours indicate negative regions of the map.

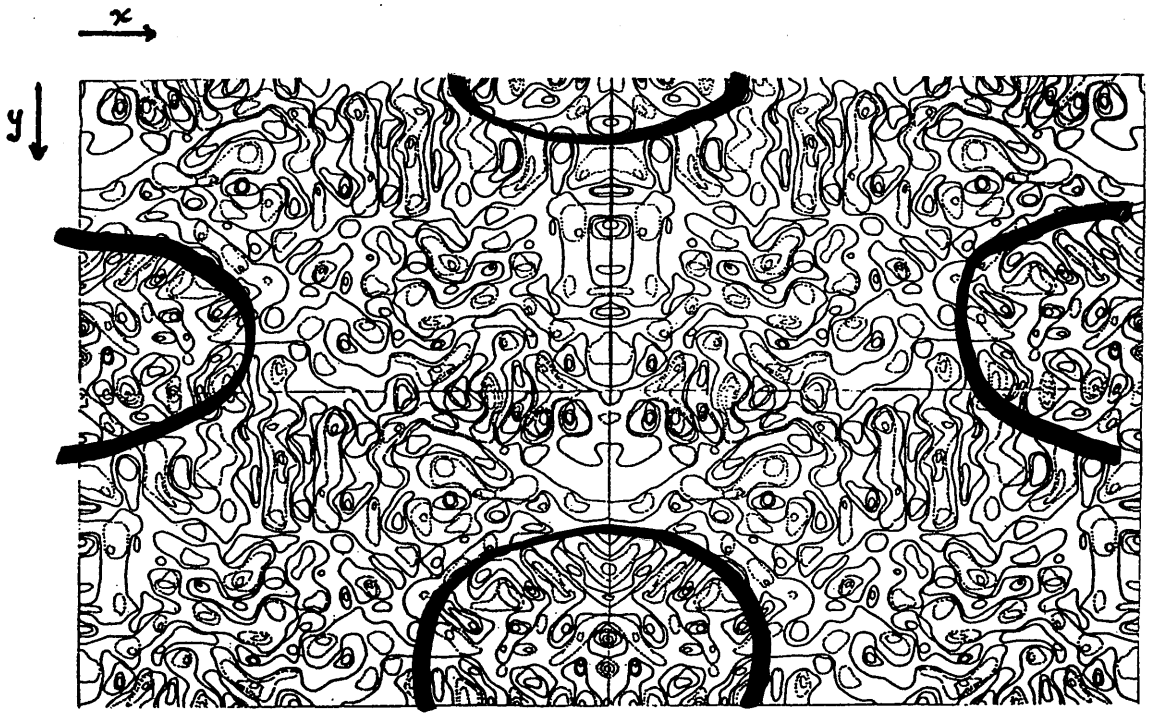


Figure 6.18 : Electron density map of the complete unit cell of hk0 projection of the K2 form of CK calculated using the M4S4 phase set. The map was calculated at intervals of 2.96 Å in a and 2.73 Å in b. Contouring is in equal but arbitrary intervals. Dotted contours indicate negative regions of the map.

forms, it seems reasonable to assume that the unit cell is only one molecule thick and that the projection view may be interpretable, at least in principle.

In order to aid further in the interpretation of the map it may be helpful to consider the closely related K1 form for which a subunit arrangement has been proposed. Firstly, consider the cell dimensions, viz.

K1 $a = 248, b = 149, c = 52$ Å, $B = 90°$, $Vol = 1.92 \times 10^6$ Å^3

K2 $a = 238, b = 164, c = 52$ Å, $B = 91°$, $Vol = 2.02 \times 10^6$ Å^3

As the cell dimensions and cell volumes are very similar, let us assume that both crystal forms contain the same number of subunits per asymmetric unit. Even if neither of the proposed arrangements of subunits in the K1 crystal form is correct, it is likely that any pseudo-trigonal or hexagonal form would contain three rather than four subunits in the asymmetric unit. Therefore, it will be assumed that the K2 crystal form also contains three subunits per asymmetric unit.

Further comparison of the cell dimensions of the two crystal forms reveals that with respect to K1, K2 is slightly shorter and slightly wider. This difference could be accommodated by a slight variation of the pseudo-hexagonal arrangement described in Chapter IV. Consider the dimers not lying on the crystallographic 2-fold axis. It would only require, in these dimers, a rotation of a few degrees in the x-y plane such that the molecular axes were to become more

closely parallel to the a axis, to destroy the pseudo-symmetry and produce the required change in cell dimensions to yield the K2 form.

Following such a scheme, it follows that one third of the dimers in the K2 form also lie on the crystallographic two-fold axis. The transparent overlay to Figure 6.18 indicates such an area of relatively high electron density, surrounded by a relatively featureless region attributable to solvent. If this interpretation is correct, the dimensions of the CK subunit are approximately given as 41 x 62 x 52 Å ; the third dimension being given by the unit cell thickness.

The position of the remaining molecules is not well defined. Clearly, there is a general region of high electron density running diagonally across the centre of each asymmetric unit. However, it is not possible to define the molecular boundary in this instance.

References for Chapter VI

1) Green, D.W., Ingram, V.M., Perutz, M.F. (1954) Proc. Royal Soc., A225, 287

2) Harker, D. (1956) Acta. Cryst., 9, 1

3) Pepinsky, R. and Okaya, Y. (1956) Proc. Nat. Acad. Sci. USA, 42, 286

4) Bijvoet, J.M. (1949) Proc. Koninkl. Ned. Acad. Wetenschap (B), 52, 313

5) Perutz, M.F. (1956) Acta. Cryst., 9, 867

6) Rossmann, M.G. (1961) Acta. Cryst., 14, 383

7) Phillips, D.C. (1966) Adv. Res. Diff. Meth., 2, 75

8) Karth, G. and Parasarathy, R. (1965) Acta. Cryst., 18, 745

9) Matthews, B.W. (1966) Acta. Cryst., 20, 230

10) Stryer, L., Kendrew, J.C., Watson, H.C. (1964) J. Mol. Biol., 8, 96

11) Srinivasan, R. (1961) Acta. Cryst., 14, 607

12) Dickerson, R.E., Kopka, M.L., Varnum, J.C., Weinzierl, J.E. (1967) Acta. Cryst., 23, 511

13) Blundell, T.L. and Johnson, L.N. (1976) 'Protein Crystallography', Acad. Press, p357

14) Singh, A.K. and Ramaseshan, S. (1966) Acta. Cryst., 21, 279

15) Hart, R.G. (1961) Acta. Cryst., 14, 1194

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

17) Rossmann, M.G. (1960) Acta. Cryst., 13, 221

18) Steinrauf, L.K. (1963) Acta. Cryst., 16, 317

CHAPTER VII

BIOCHEMICAL STUDIES

7.1 Preparation

The observation that different batches of CK exhibited varying abilities to crystallise in the monoclinic forms that have been described led to suspicions concerning the reproducibility of the preparation. It was felt that a possible failing in the classical preparation (1) was the omission of any protease inhibitors. As a consequence, the preparation was modified to include the addition of phenylmethanesulphonyl fluoride (1mM) at each stage. This inhibitor has been found to be very effective in the inactivation of serine proteases and has been used to greatly increase yields in the arom complex preparation (2). Preparations of CK so modified produced enzyme of comparable activity to that obtained without the addition of any protease inhibitor. Similarly, there was no observable increase in enzyme purity as monitored by SDS gel electrophoresis.

Crystallisations were set up using enzyme prepared in this manner under the conditions described for the production of K1 and K2 type crystals. These crystallisations yielded crystals of similar external morphology to that previously described. However, the diffraction was very poor.

7.2 Peptide Mapping

Williamson et al (3) claimed that high resolution SDS gel electrophoresis studies had revealed that rabbit muscle CK was composed of two peptides differing in length by

approximately 15 amino acid residues.

It is known that CK contains 4 thiols (4). Therefore tryptic digestion should yield a maximum of 4 cystic peptides. Indeed, inspection of the sequence of CK shows that precisely 4 cystic peptides should be produced. However, any heterogeniety existing in the enzyme would give rise to greater than 4 peptides.

CK was denatured using 8M urea and then reacted to completion with C-14 radioactive (1μCi/μmole) labelled iodo-acetamide (20°C for 90 minutes). After removal of excess reagent by dialysis, the carboxymethylated enzyme was freeze-dried. Trypsin (1% w/w) was added to the redissolved solid and digestion was taken to completion by incubation at 37°C for 4 hours. Paper chromatography employing the Waley-Watson eluent (n-butanol : pyridine : glacial acetic acid : water in the ratio 30 : 20 : 6 : 24 , pH 6.5) was used to separate the peptides and the resulting autoradiograph is shown in figure 7.1. The four spots representing the four cystic peptides expected occur in the ratio 2 : 2 : 1.5 : 1 according to intensity.

The X-ray studies of Chapter IV show that to a resolution of approximately 6 Å, the enzyme which crystallises in the pseudo-trigonal form has identical subunits. However, the crystallisation may have selectively purified one component of the enzyme system giving only a partial view. The peptide mapping results shown here indicate that there is, furthermore, no evidence for subunit heterogeniety in the enzyme in solution. It should, however be added that while

the relative intensity of the spots on autoradiographs of this type are not a reliable guide to their corresponding abundance, it is possible that the lower intensity of one of the spots may indicate the partial loss of one cystic peptide from CK. This experiment does not therefore exclude the possiblility that a terminal cystic peptide is partially lost due to proteolysis during the preparation so giving rise to subunit heterogeniety.

7.3 Isoelectric Focussing

Proteins are made from a stock of twenty amino acids. The side chains of these amino acids range from non-polar, non-ionisable to basic and acidic varieties. Thus, many of these side chains can enter into interactions involving the partial or complete transfer of protons. The pK values of these side chains vary from c.3.86 for aspartic acid to c.12.48 for arginine in the case of free amino acids. In proteins, the pK's of the amino acid side chains may be affected by their local environment. Generally, at an alkaline pH, carboxylic acid residues are negatively charged and basic residues are neutral. Thus, the protein is likely to have a net negative charge. Conversely, at acidic pH's the carboxylic acid residues are uncharged due to the capture of one proton and basic residues are positive for the same reason. Thus, the protein is likely to have a net positive charge. At a particular intermediate pH, the net charge on the protein will be zero. This is defined as the isoelectric point, pI, of the protein.

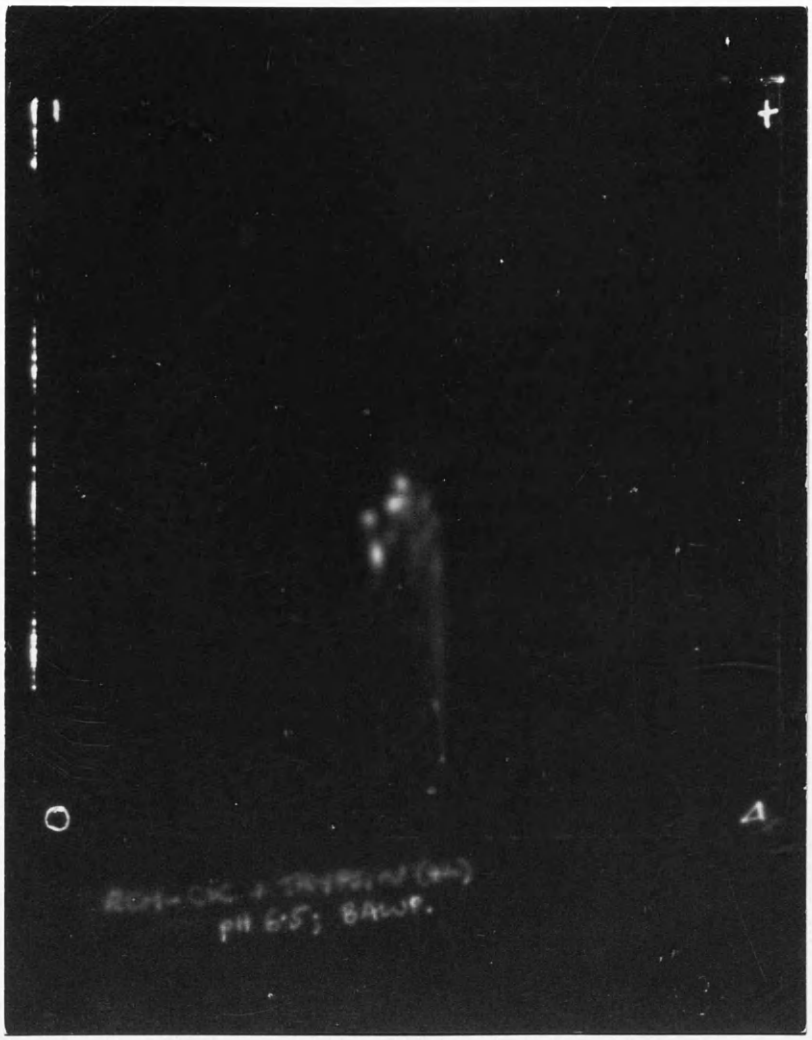


Figure 7.1 : Autoradiograph of the tryptic digest of creatine kinase.

Isoelectric focussing is a method for separating proteins on the basis of their isoelectric points. The way this is done is to carry out electrophoresis in the presence of a linear, stable pH gradient. This gradient is arranged such that the pH is alkaline at the cathode and acidic at the anode. Thus, a protein introduced at some point on the gradient above its isoelectric point will have a net negative charge and will consequently migrate towards the anode. However, as it migrates it moves down the pH gradient until it achieves its pI at which point the net charge on the protein is zero and migration ceases. A similar argument can be applied to a protein entering the system below its pI. Hence, the protein is focussed at a point corresponding to its isoelectric point and the problems of diffusion encountered with normal electrophoresis are eliminated.

The way this pH gradient is achieved is to use a mixture of so-called ampholytes. These ampholytes comprise a mixture of aliphatic polyamino-polycarboxylic acids with a wide range of pI's. At a pH of 7, such a mixture of ampholytes will contain some species that are negatively charged and some which are positively charged for the reasons already discussed. The application of an electric field causes these species to migrate in a corresponding fashion. As an addition, the anode is surrounded by a dilute acid and the cathode by a dilute alkaline solution. Thus, the carrier ampholyte with the lowest pI and hence the most negatively charged molecule migrates towards the anode and as it does so passes into an area of lower pH provided by the dilute

acid surrounding the electrode. The other ampholytes settle in a corresponding fashion. The carrier ampholyte, due to its high buffering capacity, now creates a local pH equal to its own pI. Hence, the gradient is formed.

7.3.1 Method

1. A glass plate of approximate dimensions 20cm x 12cm is washed with distilled water and then dryed under ultra-violet light.
2. The plate is then dipped into a solution of 0.5% w/w gelatin, 0.1% w/w chrome alum and dried under UV light. This step is repeated.
3. A convenient quantity (8.5 mls) of polyacrylamide solution is made using 5.59 mls distilled water, 0.434 mls Ampholine, 0.31 mls of solution A (1 ml N N N'N' tetramethlyethylenediamine (TMED) in 150 mls water) and 1.3 mls of solution B (100g acrylamide, 2.7g bisacrylamide dissolved in water to 300 mls and filtered). This mixture is cooled on ice and deaereated under vacuum.
4. 8.5 mgs of ammonium persulphate is added and the solution is poured onto the glass plate, taking care to remove any bubbles, and allowed to set.
5. A cell comprising two graphite electrodes spaced about 10 cm apart is used. The anode is wetted with a small amount of TMED solution and the cathode is wetted with a small amount of dilute phosphoric acid.
6. The gel coated glass plate is placed gel side down onto

the electrodes and a potential of 200 V is applied for about 1 hour. During this time, the initially high current drops and a certain amount of heat is generated.

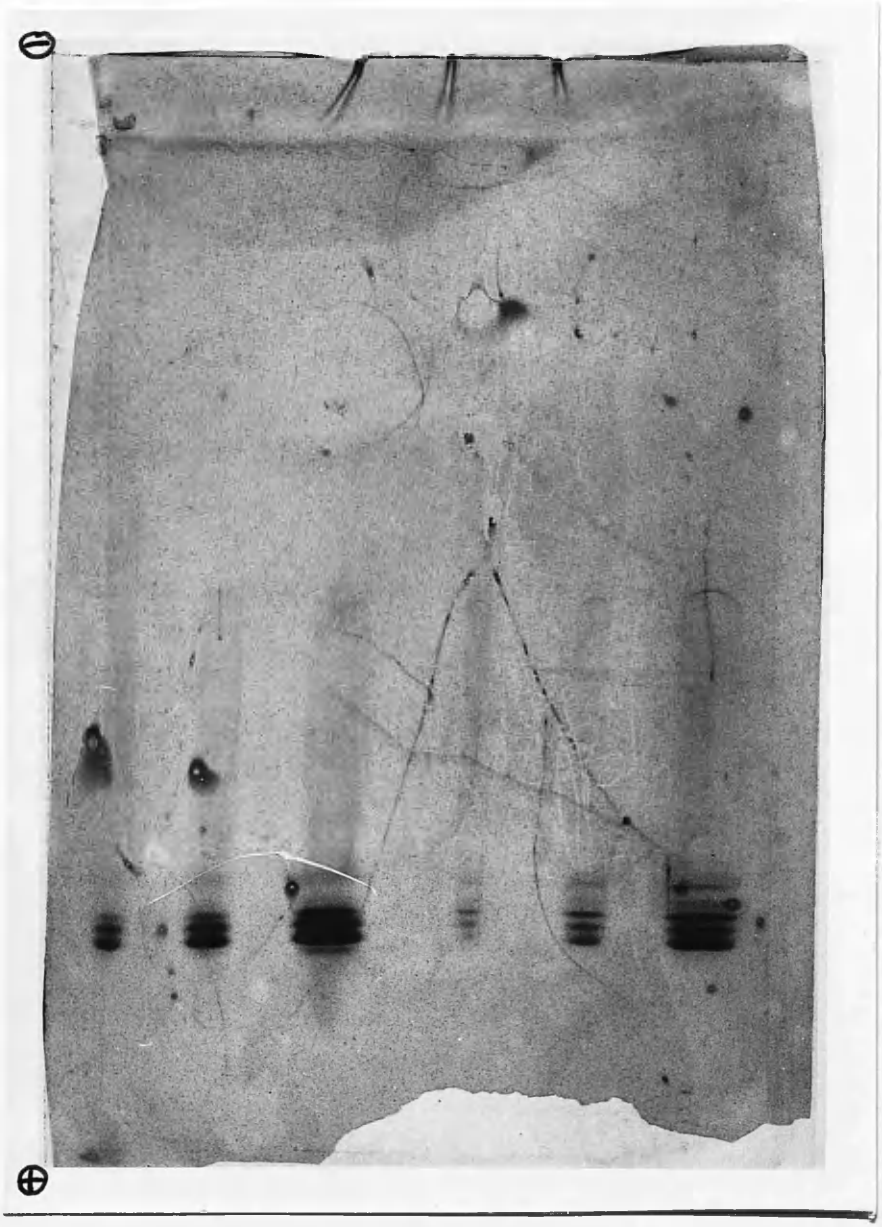
7. The plate is removed and protein is applied at a point approximately halfway between the electrodes in quantities of $\leqslant$ 25 μg in $\leqslant$ 5 μl of water. The plate is then replaced and focussed for approximately 18 hours in a cold room.

8. The resulting plate is stained using 0.1% coomassie blue (G250) in methanol : acetic acid : water in the ratio 9 : 2 : 9 for 30 minutes and then destained in a similar solution in the ratio 6 : 1 : 13.

7.3.2 Results Obtained With CK

Figure 7.2 depicts the pattern obtained after staining with Coomassie blue for CK electrofocussed employing ampholytes in the pH range 3.5-10 . As can be seen, the result on electrofocussing is a triplet with an approximate intensity ratio of 1:2:1. This is the classic pattern obtained for the random dimerisation of two non-identical subunits giving rise to one homodimer and two heterodimers.

Hooton (5) claimed that the electrophoretic mobility of CK on starch gels (pH 8.4) was reduced on carboxymethylation of the protein using iodoacetamide. However, no change in mobility was observed when the essential thiol was blocked using iodoacetate. It was decided to repeat this experiment using isoelectric focussing to monitor the resulting change. Figure 7.3 depicts the result obtained. As can be seen, both blocking agents give rise to a more homogeneous product.



(1) (2) (3) (4) (5) (6)

Figure 7.2 : Pattern obtained after isoelectric focussing of varying loadings of native creatine kinase (1% solution). Coomassie blue staining, ampholyte range 3.5-10. (1) 1 ul (2) 3 ul (3) 10 ul (4) 1 ul (5) 3 ul (6) 10 ul

(a) (b) (c)

Figure 7.3 : Isoelectric focussing result for creatine kinase (a) native (b) blocked using iodoacetamide (c) blocked using iodoacetate. All loadings 3 ul of 1% solution.

Thus, one is forced to the conclusion that the heterogeniety of the two subunits of CK is intimately related the essential thiol.

The pI of a protein can be determined by isoelectric focussing using a special surface pH electrode. Unfortunately, one of these was not available. However, the pH gradient formed by the ampholyte mixture used is claimed by the manufacture (LKB) to be reasonably linear and thus the pI of CK can be estimated from its position on the plate. In this way, it was determined that the native triplet centred on a pH of c.8.1. This is quite close to the pKa of cysteine in peptides which is c.8.33. Thus, the essential thiol might be present in two distinct forms i.e. $-S^-$ and -SH. Blocking the thiol with iodoacetamide gives rise to a non-ionisable, neutral moiety. As the position of CK blocked in this way corresponds to one of the 'homodimer' bands in the native pattern, it is tempting to suggest that this native band corresponds to a homodimer with the essential thiols unionised.

It should be pointed out that this thiol is unlikely to be base paired with a histidine residue as has been proposed (6). There are three forms that the thiol can be in viz :

S^- ; S----H----N ; SH

Blocking the thiol with iodoacetamide would disrupt the base paired possibility and consequently probably give rise to a species with a different pI from any of the native moieties. Blocking the enzyme with iodoacetate gives rise to an

ionisable side group and is thus quite a different case. The pKa of free thioacetic acid is c.3.33 and the move to lower pH of the isoelectric point of CK modified in this way is thus to be anticipated.

7.4 Summary

It has been shown by isoelectric focussing that native CK consists of three species. One interpretation of this is that there are two forms of the CK subunit and that the dimer comprises a random pairing of these two forms. The existance of two differing subunits is in direct agreement with the high resolution SDS gel electrophoresis results of Williamson (3). The random pairing proposal permits the production of two homodimers. The presence of at least one homodimer is supported by the crystallographic evidence already discussed. The fact that the particular homodimer which gives rise to crystals with higher symmetry is present to an extent of only 25% of the total may go some way to explaining the difficulty in producing good crystals.

The result which is more difficult to explain is the finding that the subunit heterogeniety can be apparently removed by blocking the essential thiol. The finding of Williamson that the two subunits differ in molecular weight by the equivalent of 15 amino acids could mean either that the subunits are completely different or that the lower molecular weight species has simply lost a portion of the chain due to proteolytic action, possibly during the

preparation. The latter is the more likely option and is not necessarily in conflict with the IEF results.

References for Chapter VII

1) Kuby, S.A., Noda, L., Lardy, H.A. (1954) J. Biol. Chem., 209, 191

2) Lumsden, J., Coggins, J.R. (1977) Biochem. J., 161, 599

3) Williamson, J., Greene, J., Cherif, S., Milner-White, E.J. (1977) Biochem. J., 167, 731

4) Watts, D.C. (1973) 'The Enzymes', 8, 383

5) Hooton, B.T. (1968) Biochem, 7, 2063

6) Watts, D.C. and Rabin, B.R. (1962) Biochem. J., 85, 507

CHAPTER VIII

DISCUSSION

8.1 The Active Site of CK

From the myriad of biochemical studies that have been made of CK, a picture can be built up of the active site of the enzyme. Certain residues are known to be essential for CK activity and these will be considered one at a time.

8.1.1 Arginine

Treatment of rabbit muscle CK with either butanedione in borate buffer or phenylglyoxyl in veronal buffer causes inactivation of the enzyme (1). The inactivation caused is proportional to the extent of reaction and results indicate that there is one essential arginyl residue per subunit. The modified enzyme is unable to bind nucleotide substrates. Further to this, $MgATP^{2-}$ and $AgADP^{-}$ both provide substantial protection against inactivation by these reagents. It is thus assumed that arginine is involved in the binding of the nucleotide substrate, probably through interaction with the charged oligophosphate moiety of the nucleotide. Arginine has been found to play a similar role in Staph. Nuclease (2), lactate dehydrogenase (3) and horse liver alcohol dehydrogenase (4).

8.1.2 Cysteine

The presence of an essential cysteine in the CK subunit has been described in chapter I. However, the results regarding inhibition of CK by blocking the reactive thiol are less clear cut than was originally the case. Initial findings on rabbit muscle CK were that complete inhibition could be

achieved using either iodoacetamide or iodoactetate (5). This, however, is not the case when the experiment is repeated using chicken heart type III enzyme where only c.70% inhibition is obtained (6). A similar result is obtained with this enzyme using the less reactive reagent, iodomethane. More recently, studies on rabbit muscle CK have shown that if the thiol is blocked by a CH_3S- group using methanethiolsulphonate, a residual activity of c.20% remains (7).

Thus, it would appear that the precise role of the 'essential' thiol is more complex than was originally thought. The authors of this more recent work point out that an obvious difference between the CH_3S- blocking group and the more common thiol reagents is one of hydrophobicity. They argue that the difference in size between CH_3S- and H_2NCOCH_3 is not sufficent to account for the different effects of the two blocking groups. They conclude that it is the hydrogen bond forming capability of H_2NCOCH_3 that destroys CK activity.

The amino acid sequence surrounding the active thiol is Ala-Gly-Pro-His-Phe-Met-Asp-Gly-Tyr-Val-Leu-Thr-Cys-Pro-Ser-Asn-Leu-Gly-Thr-Gly-Leu-Arg (8). If the rules of Chou and Fassman (9) are applied to this sequence, it is found that the cysteine occurs at the beginning of a β-bend between two regions of β-sheet (7). If one considers such a position as a kind of 'molecular hinge' , it can be appreciated that the role of the reactive thiol may be that it is a key factor in

the conformation change that occurs on substrate binding. The authors argue that this conformational change could result in the thiol becoming buried in a more hydrophobic environment so accounting for the results obtained with different blocking groups.

8.1.3 Histidine

95-100% inhibition of CK can be effected by the reaction of 2 moles of diethylpyrocarbonate with one mole of CK (10). The binding properties of the enzyme and associated conformation changes do not appear to be affected, however (11). Thus, it can be inferred that one histidine per subunit is necessary for catalytic activity but that it is not directly involved in substrate binding.

8.1.4 Lysine

Evidence for one essential lysine per subunit has been obtained by the reaction of CK with p-nitrophenyl acetate (12). It was, at first, thought that inactivation by this reagent was due to modification of the essential cysteine. However, it was later shown (13,14) that the reaction was with lysine. It was also found that the transition state analogue complex afforded protection against inhibition by this reagent (12). However, this protection was due to a rise in the pK of the lysine by 0.85 pH units and not due to actual physical blocking. This rise in the pK was explained to result from the conformation change that occurs on binding of the complex.

Further evidence for the importance of a lysine in CK has been supplied by nuclear Overhauser effect (NOE) studies (15). In this work, the transition state analogue complex was employed once again and NOE measurements of the formate proton were performed. Results indicated that the formate was bound to a lysyl residue as evidenced by the chemical shift of the INDOR peak and by the lack of an NOE in the dansylated enzyme. It was also noted that a large NOE was observed only when all the other substrates were present indicating the existance of a unique conformation corresponding to the transition state.

8.1.5 Tyrosine

Tyrosine is involved in substrate binding in arginine kinase and by analogy has been assumed to play a similar role in CK. However, experiments to support this (16) have produced less convincing evidence than was obtained for arginine kinase. While the conversion of four to five tyrosyl residues to monoiodotyrosines renders the enzyme unable to bind its nucleotide substrates, a change in the reactivity of some non-essential thiols was observed. There were no gross conformational changes in the modified enzyme but the change in reactivity of the non-essential thiols does indicate small changes. Hence any conclusions that can be made concerning the role of the tyrosine residues must await further investigation.

8.2 The Mechanism of CK

The possibility that the essential thiol of CK may be pair bonded to histidine has been discussed. Watts and Rabin (17) have suggested a mechanism for CK based on this pair bonding in which a circular flow of electrons is permitted (figure 8.1). The evidence available at that time certainly supported such a notion but the recent results obtained with the CH_3S- blocked enzyme would tend to cast doubt on the viability of this mechanism. However, there is some doubt that, in cases where residual activity persists after reaction with thiol blocking agents, total reaction to the extent of two thiols per molecule has been achieved (18).

Another representation of the active site has been proposed by McLaughlin et al (19). In this there is no suggestion of a mechanism as such (Figure 8.2). It appears from this proposal that phosphorylation occurs cis to the N-methyl of creatine rather than trans as is known to be the case (20). Neither is there any firm evidence to suggest that Mg^{2+} is bound to the enzyme nor that histidine is involved in binding in any way. The interesting feature of this plan is the way the terminal phosphoryl is bound to the enzyme by a lysine and an arginine. The notion is attractive and the evidence is quite good (15,1). The position of the cysteine and what its function may be are not discussed.

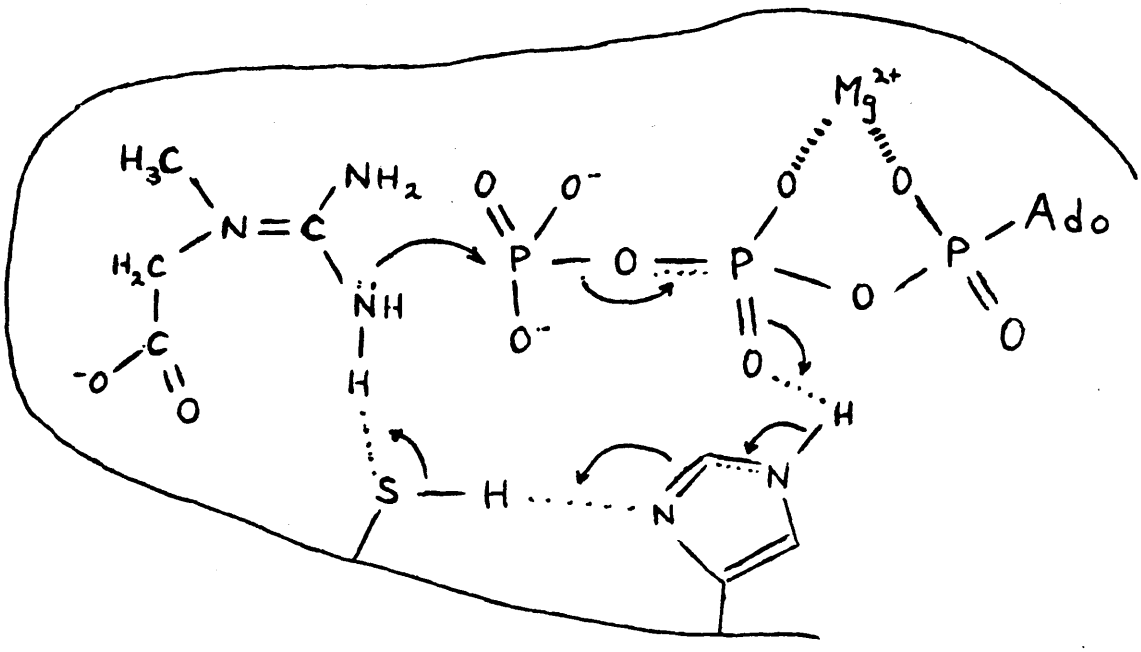


Figure 8.1 : Creatine kinase mechanism as proposed by Watts and Rabin (17).

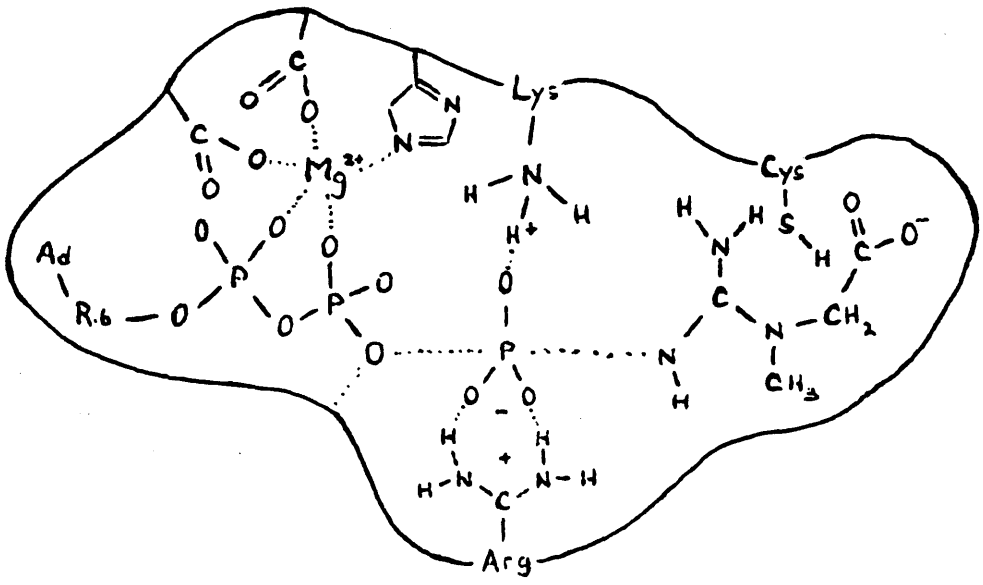


Figure 8.2 : Configuration of substrates in the active site of creatine kinase as proposed by McLaughlin et al (19)

Thus, from the evidence available and a critical inspection of the ideas that have been proposed to date, some impressions of the active site can be gained. The mechanism proposed by Watts involving a histidine was examined further using space filling models. It was found that the thiol is not necessary to span the distance required for hydrogen transfer from creatine to ADP. The transfer process could therefore occur using histidine alone so accounting for the inability of methanethiolsulphonate totally to inactivate the enzyme. The role of the cysteine is more likely to be that it plays an important part in the conformational changes that are so necessary for kinase activity.

As regards substrate binding, a tyrosine residue is possibly involved in nucleotide binding by a 'face to face' type of interaction. The binding of the γ-phosphoryl group is likely to be that which was proposed by McLaughlin et al.

The manner in which creatine is bound is not clear although one can envisage a small hydrophobic pocket in which the N-methyl group sits so aligning the molecule for the transfer process. It is known that the N-methyl of creatine is important in this regard (21).

The conformational changes that facilitate both the initial binding of the substrates and the subsequent trans-phosphorylation process must remain the subject for further study in CK and kinases in general.

8.3 Other Kinases

The structures of five other kinases have now been determined. It is therefore possible to examine these structures to determine if there are any similarities that would enable some tentative predictions as regards the structure of CK to be made.

8.3.1 Phosphoglycerate Kinase

The structure of horse muscle phosphoglycerate kinase (PGK) has been determined to 2.5 Å resolution by Blake et al. (22,23). Also, the structure of PGK from yeast has been determined (24). The two enzymes have been found to be structurally homologous and only PGK from horse muscle will be examined in detail.

It has been found that the enzyme folds in such a way as to produce two domains, roughly equal in size, separated by a deep cleft. Although both ends of the polypeptide chain are located in the same domain, the two domains can be described as the N-terminal and C-terminal halves of the chain and are connected by a neck region comprising two strands of the chain.

The domains themselves are composed of a central β-sheet of six parallel strands flanked by α-helices. This is very similar to the NAD binding site in lactate, 3-glyceraldehyde phosphate, alcohol and malate dehydrogenases and may provide evidence for a common ancestral gene (25). However, an alternative proposal of convergent rather than divergent evolution has also been proposed to account for the

similarity (26).

Substrate binding studies have successfully revealed the nucleotide binding site as a narrow slot in the inner face of the C-terminal domain. The γ-phosphate of ATP appears not to interact with any residues and occupies an exposed position. It is thought that the phosphoglycerate site is located in the N-terminal domain some 12 Å away and that PGK undergoes a hinge-bending mechanism to bring the two substrates together (27). Evidence for the large change in tertiary structure implied by this action is supported by small angle X-ray scattering experiments (28) and further crystallographic studies (29).

8.3.2 Hexokinase

Like PGK, hexokinase monomer folds in such a way as to form two lobes separated by a deep cleft (30). However, in the case of hexokinase, the two lobes are quite dissimilar in that one is composed largely of α-helix while the other contains substantial β-structure in addition to α-helix (26). The β-pleated sheet forms an 11-stranded barrel-like structure with two α-helices in the centre and therefore appears to be quite unlike PGK. However, the nucleotide binding site thought to be involved in the phosphate transfer does bear similarity to the NAD binding site of lactate dehydrogenase except that in hexokinase, two of the six β-strands are anti-parallel to the other four.

8.3.3 Pyruvate Kinase

The 6 Å resolution study of cat muscle pyruvate kinase (PK) (31), reveals that the monomer of this tetrameric enzyme is organised into two domains of very different size. While these domains are separated by a cleft, the subunit is not so distinctly bilobed as either PGK or hexokinase. Difference Fouriers on the ADP-PK binary complex reveal the active site to be situated between the two lobes of the subunit.

When the study was extended to higher resolution (32,33), the subunit was found to consist of three domains (A, B and C). Domain A displays striking 8-fold near-symmetry, being composed of eight consecutive β-strand/α-helix units arranged cylindrically. Domain B may be described as a closed anti-parallel β-structure while Domain C comprises a 5-stranded β-sheet interconnected by α-helices. Part of this feature resembles the first mononucleotide binding fold (mbf) of lactate dehydrogenase (34) defined as one half of the NAD binding domain. Thus, the mbf comprises three parallel strands of β-sheet and two interconnecting α-helices.

8.3.4 Phosphofructokinase

The structure of phosphofructokinase (PFK) from Bacillus stearothermophilus has been determined to 2.4 Å resolution (35). The subunit of this tetrameric enzyme is also bilobal. Each domain has a central β-sheet sandwiched between α-helices. The β-sheet in the first domain has

seven strands of which the central five are parallel while that of the second domain comprises four parallel strands. Thus, while PFK displays the familiar bilobal structure with domains rich in β-sheet, the topological arrangement of these sheets is different from PGK.

8.3.5 Adenylate Kinase

The 3 Å resolution structure of adenylate kinase from porcine muscle (36) reveals the familiar arrangement of two domains separated by a deep cleft. Domain A is composed of three helical rods enclosing a small hydrophobic core. Domain B contains a central 5-stranded β-sheet with accompanying α-helices.

8.3.6 Structural Predictions for CK

It seems very likely from the preceding survey that the CK subunit is bilobal in nature. It also seems probable that one of these domains comprises a central β-sheet flanked by α-helices. The number of strands in this sheet is likely to be at least five, however the order in which the strands occur cannot be predicted with any certainty. It is also probable that this domain provides the nucleotide binding site and that this is situated on the inner face of the domain near the base of the cleft. The conformational change that occurs on binding of the transition state analogue complex of CK suggests that the enzyme's mode of action may be similar to PGK and that it is also a 'hinge-bending'

enzyme.

In Section 8.1.2 reference was made to the result obtained on applying the rules of Chou and Fassman (9) to the amino acid sequence of a fragment of CK containing the essential cysteine (7). Appendix V presents the results obtained using the full CK sequence with a number of different prediction algorithms (see Appendix for details). This confirms the prediction that the essential cysteine (residue 282) is situated at the end of a section of β-sheet, followed by a reverse turn. While the overall prediction gives some indication of alternating α/β topology, the evidence is not strong. There is, however, some indication of a domain structure. The sequence 134-224 appears to contain little or no evidence for β-sheet but fairly strong indications of two lengthy sections of α-helix separated by reverse turns. Indeed, it is possible that this arrangement may continue right up to residue 277 giving, in total, a pattern of 5 or 6 α-helices and 6 reverse turns. Evidence for β-sheet appears to be confined mainly to those sequences before and after this central domain.

However, despite the increasing sophistication of such prediction methods, it would be unwise to attempt to draw firm conclusions at this stage. Instead, it is hoped that a full crystallographic structural analysis of CK will be achieved in the near future and that any relationships with other kinases that may exist will then be revealed.

References for Chapter VIII

1) Borders, C.L. and Riordan, J.F. (1975) Biochem, 14, 4699

2) Arnone, A., Bier, C.J., Cotton, F.A., Day, V.W., Hazen, A. (1971) J. Biol. Chem., 246, 2302

3) Adams, M.J., Beuhner, M., Chandrasekhar, K., Ford, G.C., Hackert, M.L., Liljas, A., Rossmann, M.G., Smiley, I.E, Alison. W.S., Everse, J., Kaplan, N.O. Taylor, S.S. (1973) Proc. Nat. Acad. Sci. USA, 70, 1968

4) Eklund, H. Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Branden, C.I. (1974) FEBS. Lett., 44, 200

5) Mahowold, T.A., Nordstrom, E.A., Kuby, S.A. (1962) J. Biol. Chem., 237, 1535

6) Hooton, B.T. (1968) Biochem, 7, 2063

7) Maggio, E.T., Kenyon, G.L., Markham, G.D., Reed, G.H. (1977) J. Biol. Chem., 252, 1202

8) Thomson, A.R., Eveleigh, J., Miles, J.B. (1964) Nature, 203, 267

9) Chou, P.Y., Fasman, G.D. (1974) Biochem, 13, 211

10) Pradel, L.A., Kassab, R. (1968) Biochem. Biophys. Acta, 167, 317

11) Roustan, C., Pradel, L.A., Kassab, R., Fattoum, A., Van Thoai, N. (1970) Biochem. Biophys. Acta, 206, 369

12) Watts, D.C. (1963) Biochem. J., 89, 220

13) Clark, J.R., Cunningham, L.W. (1965) Biochem, 4, 2637

14) Kassab, R., Roustan, C., Pradel, L.A. (1968) Biochem. Biophys. Acta, 167, 308

15) James, E. and Cohn, M. (1974) J. Biol. Chem., 249, 2599

16) Fattoum, A., Kassab, R., Pradel, L.A. (1975) Biochem. Biophys. Acta, 405, 324

17) Watts, D.C. and Rabin, B.R. (1962) Biochem. J., 85, 507

18) Watts, D.C. (1973) 'The Enzymes', 8, 383

19) McLaughlin, A.C., Leigh, J.S., Cohn, M. (1976) J. Biol. Chem, 251, 2777

20) Rowley, G.L., Greenleaf, A.L., Kenyon (1971) J. Am. Chem. Soc., 93, 5542

21) James, E., Morrison, J.F. (1966) Biochem. Biophys. Acta, 128, 327

22) Blake, C.C.F. and Evans, P.R.(1974) J. Mol. Biol., 84, 585

23) Banks, R.D., Blake, C.C.F., Evans, P.R., Haser, R., Rice, D.W., Hardy, G.W., Merret, M. and Phillips, A.W. (1979) Nature, 279, 773

24) Bryant, T.N., Watson, P.L. and Wendell, P.L. (1974) Nature, 247, 14

25) Rossmann, M.G., Moras, D. and Olsen, K.W. (1974) Nature, 250, 194

26) Fletterick, R.J., Bates, D.J. and Steitz, T.A. (1975) Proc. Nat. Acad. Sci. USA, 72, 38

27) Blake, C.C.F., Rice, D.W. and Cohen, F.E. (1986) Int. J. Peptide Res., 27, 443

28) Pickover, C.A., McKay, D.B., Engleman, D.M. and Steitz, T.A. (1979) J. Biol. Chem., 254, 11323

29) Rice, D.W. and Blake, C.C.F. (1984) J. Mol. Biol., 175,

219

30) Steitz, T.A., Fletterick, R.J., Anderson, W.F. and Anderson, C.M. (1976) J. Mol. Biol., 104, 197

31) Stammers, D.K. and Muirhead, H. (1975) J. Mol. Biol., 95, 213

32) Stammers, D.K. and Muirhead, H. (1977) J. Mol. Biol., 112, 309

33) Stuart, D.I., Levine, M., Muirhead, H. and Stammers, D.K. (1979) J. Mol. Biol., 134, 109

34) Levine, M., Muirhead, H., Stammers, D.K. and Stuart, D.I. (1978) Nature, 271, 626

35) Evans, P.R. and Hudson, P.J. (1979) J. Mol. Biol., 279, 500

36) Schulz, G.E., Elzinga, M., Marx, F. and Schirmer, R.H. (1974) Nature, 250, 120

Appendix I : Preparation of Rabbit Muscle Creatine Kinase

A large adult rabbit which has been starved overnight is killed by a nembutal injection in the ear. The rabbit is then skinned and the muscle removed from the back, legs and arms. The thinly sliced muscle is then packed in ice to cool for 20 minutes. After mincing, the muscle is weighed and homogenised in a blender using 2 litres of cold 0.01M KCl, 1mM EDTA and 0.2M sucrose solution per kg of muscle, taking care to maintain a low temperature. The homogenate is stirred at 0°C for 15 minutes and then centrifuged for 20 minutes at 6000 r.p.m.

The supernatant is decanted and solid NH_4Cl is added to a concentration of 0.1M with stirring. Then, while continuously stirring, this solution is made pH 9.0 by the careful addition of 5M ammonia and stirred for a further 30 minutes at 0°C. 1.5 volumes of cold 95% ethanol (AR grade) are then added slowly maintaining the temperature at 0°C. This mixture is then stirred for 2 hours at 20°C and then centrifuged for 15 minutes at 5000 r.p.m.

2.0M $MgSO_4$ is added to the supernatant with stirring to a final concentration of 0.03M. Then cold 95% ethanol is also added. The volume of ethanol added is equal to 1.5 times the volume of $MgSO_4$ used. After stirring for 30 minutes at 20°C the precipitate is collected by centrifugation for 15 minutes at 5000 r.p.m.

The precipitated enzyme is then resuspended and extracted at

0°C with 0.07M Mg Acetate, pH 9.0, in volumes equal to 6% and 4% of the volume of the supernatant obtained after very first centrifugation. Stirring is necessary to help redissolve the precipitated enzyme. The mixtures are then centrifuged at 12000 r.p.m. for 20 minutes and the final precipitate discarded.

Cold ethanol is added to the combined supernatant with stirring until a concentration of 36% (volume percent of 95% ethanol) is achieved. After 30 minutes at 0±0.5°C the precipitate is removed by centrifugation at 12000 r.p.m. Ethanol is then added to the clear supernatant to a concentration of 50% and after 30 minutes, the precipitate is collected as before. This precipitate is redissolved in distilled water at 0°C and very carefully, using 1M ammonia solution, the pH is adjusted to 9.0. This solution is then dialysed against 5-10 litres of distilled water, 1mM EDTA made to pH 9.5 with ammonia, overnight. Any precipitate appearing at this stage is removed and the final solution is thrice dialysed overnight against a similiar solution but with 10 mM EDTA. The protein concentration is measured by absorption at 280 nm using a factor of 1.07 to express protein concentration in mg/ml.

In order to store the enzyme, it was made 50% with glycerol and kept at -20°C.

The activity of the enzyme was assessed by the pH-stat method of Milner-White and Watts (1). The specific activity of the enzyme so measured was in the range 0.13 to 0.15 μmol of product released/min per μg of enzyme in the forward

direction at 30°C, pH 9.0.

Two preparations were carried out. The first of these was identical to the procedure outlined above while the second was modified by the addition of protease inhibitor as described in Chapter VII.

Other batches of CK were also studied during the course of this work. This enzyme, prepared by essentially the same route as the above, was kindly supplied by Dr E J Milner-White.

1) Milner-White, E.J. and Watts, D.C. (1971) Biochem. J., 122, 727

Appendix II : Crystallisation Conditions Explored

Key to abbreviations :

Ac	-	Acetic acid
ATP	-	Adenosine-triphosphate
CD	-	Capillary dialysis
DS	-	Davies-Segal method (Methods In Enz., 22, 266)
HD	-	Hanging drop vapour diffusion
MPD	-	2-Methylpentane-2,4-diol
Oxal	-	Oxalic acid
PEG	-	Polyethylene glycol
Ppt	-	Amorphous precipitate
Succ	-	Succinic acid
TEA	-	Triethanolamine
TG	-	Thioglycolic acid
Tris	-	Tris(hydroxymethyl)methylamine
TSA	-	Transition state analogue solution
VD	-	Vapour diffusion

Crystallisation conditions explored using MPD as precipitatant :

MPD Conc	pH	Buffer	Additives	Method	Result
50%	5.7	70mM TEA/TG		CD	Thin Plates
"	6.1	10mM "		"	Ppt
"	"	25mM "		"	"
"	"	35mM "		"	"
"	6.2	60mM "		"	Thin Plates
"	6.9	20mM "		"	Ppt
"	"	50mM "		"	Thin Plates

"	"	70mM "		"	Ppt
"	7.0	10mM Tris/Ac		"	Needles
"	"	10mM TEA/TG		"	Ppt
"	"	20mM TEA/Ac		"	Needles
"	"	25mM TEA/TG		"	Ppt
"	"	50mM TEA/Ac		"	"
"	7.1	10mM Tris/Ac		VD	Needles
"	"	"	2% TSA	"	"
"	7.3	50mM TEA/TG		CD	Ppt
"	7.5	10mM Tris/Ac		"	"
"	"	10mM TEA/Ac		"	Needles
"	7.7	10mM Tris/Ac	2% TSA	VD	"
"	7.8	"	1% TSA	"	Small Rods
"	"	50mM TEA/TG		CD	Ppt
"	8.0	10mM Tris/Ac		"	"
"	8.1	"		DS	Rods
"	"	"		VD	Monoclinic
"	"	20mM TEA/TG		CD	Ppt
"	"	20mM Tris/TG		"	"
"	8.2	20mM TEA/Ac		"	Needles
"	"	10mM Tris/Ac	50mM Cysteine	DS	"
"	"	40mM "	2mM ATP	"	Ppt
"	"	"	3mM "	"	"
"	"	"	4mM "	"	"
"	"	"	5mM "	"	"
"	8.3	10mM Tris/Ac		VD	Monoclinic
"	"	50mM TEA/TG		CD	Ppt

"	8.5	10mM Tris/Ac	2% TSA	VD	Needles
"	8.7	"		DS	Monoclinic
"	"	"		VD	"
55%	5.7	40mM Tris/Ac		CD	Needles+Ppt
"	"	70mM TEA/TG		"	Needles
"	6.0	35mM Tris/Ac		"	"
"	6.5	30mM "		"	"
"	6.6	25mM "		"	"
"	6.8	20mM "		"	"
"	7.2	50mM "		"	"
"	7.4	50mM TEA/TG		"	"
"	7.6	20mM Tris/Ac		"	Small Rods
"	7.9	"		"	Monoclinic
"	8.2	"	2mM ATP	"	Ppt
"	"	"	4mM "	"	"
"	"	"	8mM "	"	"
"	"	30mM Tris/Ac	2mM "	"	"
"	"	"	4mM "	"	"
"	"	"	8mM "	"	"
"	"	40mM Tris/Ac	2mM "	"	"
"	"	"	4mM "	"	"
"	"	"	8mM "	"	"
"	"	"	2mM Creatine	HD	"
"	"	"	4mM "	"	"
"	"	"	6mM "	"	"
"	"	"	8mM "	"	Needles
"	"	"	10mM "	"	"

"	"	"	12mM	"	"
"	8.4	5mM Tris/Ac		CD	Needles+Ppt
"	"	10mM	"	"	Needles
"	"	15mM	"	"	Needles+Rods
"	8.6	10mM	"	"	Rods+Ppt
"	8.8	"		"	Rods
"	9.0	"		"	Small Rods
60%	7.8	0.1mM Tris/TG		"	Needles
"	8.1	3mM Tris/Ac		"	"
"	"	10mM	" 20mM ME,1mM NaAc	"	Rods
"	8.2	20mM Tris/Ac		"	Monoclinic
"	"	" Tris/Succ		"	Large xtals
"	"	" Tris/Oxal		"	Needles+Ppt
"	"	" Tris/HCl		"	Needles
"	"	" Tris/Ac	20mM Ca(Ac)	"	"
"	8.3	20mM Cysteine/KOH		HD	Ppt

The thin plates, needles and rods referred to above failed to yield any observable diffraction. Seeding methods employed in an attempt to develop crystal size were unsuccessful. The term 'monoclinic' is used above to denote the crystal forms K1 and/or K2

Crystallisation conditions explored using other organic solvents :

Precipitant	pH	Buffer	Method	Result
40% Dioxane	8.0	20mM Tris/Cl	CD	Ppt
60% Acetonitrile	8.2	20mM Tris/Ac	"	"
" t-Butanol	"	"	"	Needles+Ppt
" Butyrolactone	"	"	"	Ppt
" Dimethylformamide	"	"	"	"
" Dimethylsulphoxide	"	"	"	"
" 1,3 Propanediol	"	"	"	"
30% Methanol	8.5	"	VD	"

Crystallisation conditions explored using polyethylene glycol as precipitant :

Precipitant	pH	Buffer	Method	Result
2% PEG 6000	8.0	10mM Tris/Ac	HD	Ppt
" 20000	"	"	"	"
" 6000	9.0	20mM Glycine/NH_4OH	"	"
4% PEG 6000	8.0	10mM Tris/Ac	"	"
" 20000	"	"	"	"
6% PEG 6000	"	"	"	"
" 20000	"	"	"	"
25% PEG 6000	6.5	20mM Tris/Ac	"	"
"	7.0	"	"	"
"	7.5	"	"	"
"	8.0	"	"	"
"	9.0	"	"	"
33% PEG 6000	9.0	20mM Glycine/NH_4OH	CD	"

Crystallisation conditions explored using inorganic salts as precipitant :

Precipitant		pH	Buffer	Method	Result
40%	$(NH_4)_2 SO_4$	8.0	20mM Tris/HCl	HD	Ppt
50%	"	"	"	"	"
52%	"	7.3	KOH	"	"
53%	"	7.5	"	"	"
55%	"	7.0	20mM Phos/KOH	CD	"
"	"	7.5	"	"	"
"	"	8.0	"	"	"
60%	"	6.0	20mM Tris/Ac	"	"
"	"	6.5	"	"	"
"	"	7.0	"	"	"
"	"	7.4	"	"	"
"	"	8.0	20mM Tris/HCl	"	"
"	"	"	20mM Phos/KOH	"	"
"	"	8.5	"	"	"
65%	"	"	"	"	"
"	"	9.0	"	"	"
70%	"	8.0	"	"	"
"	"	"	20mM Tris/HCl	HD	"
"	"	9.0	"	CD	"
20%	NH_4OOCCH_3	7.0	10mM Tris/Ac	CD	Ppt
40%	"	7.5	"	"	"
70%	"	8.0	"	"	"
78%	"	8.5	"	"	"

50%	NH_4NO_3	8.0	20mM Tris/HCL	HD	Ppt
70%	"	"	"	"	"
90%	"	"	"	"	"
50%	Na Citrate	8.0	20mM Tris/HCl	"	"
70%	"	"	"	"	"
90%	"	"	"	"	"
50%	Li_2SO_4	8.0	20mM Tris/HCl	"	"
70%	"	"	"	"	"
90%	"	"	"	"	"
50%	Zn Ac_2	8.0	20mM Tris/HCl	"	"
70%	"	"	"	"	"
90%	"	"	"	"	"
10%	LiCl	8.0	20mM Tris/HCl	"	"
20%	"	"	"	"	"
30%	"	"	"	"	"
40%	"	"	"	"	"
50%	"	"	"	"	"
60%	"	"	"	"	"
70%	"	"	"	"	"
80%	"	"	"	"	"
90%	"	"	"	"	"

Appendix III : C2 Patterson Symmetry

Equivalent positions in C2 are given by :

$x, \quad y, \quad z$	$\bar{x}, \quad y, \quad \bar{z}$	$\frac{1}{2}+x, \frac{1}{2}+y, \quad z$	$\frac{1}{2}-x, \frac{1}{2}+y, \quad \bar{z}$

Therefore, self-vectors are found at :

$0, \quad 0, \quad 0$	$2\bar{x}, \quad 0, \quad 2\bar{z}$	$\frac{1}{2}, \quad \frac{1}{2}, \quad 0$	$\frac{1}{2}-2x, \quad \frac{1}{2}, \quad 2\bar{z}$
$2x, \quad 0, \quad 2\bar{z}$	$0, \quad 0, \quad 0$	$\frac{1}{2}+2x, \quad \frac{1}{2}, \quad 2z$	$\frac{1}{2}, \quad \frac{1}{2}, \quad 0$
$\frac{1}{2}, \quad \frac{1}{2}, \quad 0$	$\frac{1}{2}-2x, \quad \frac{1}{2}, \quad 2\bar{z}$	$0, \quad 0, \quad 0$	$2\bar{x}, \quad 0, \quad 2\bar{z}$
$\frac{1}{2}+2x, \quad \frac{1}{2}, \quad 2\bar{z}$	$\frac{1}{2}, \quad \frac{1}{2}, \quad 0$	$2x, \quad 0, \quad 2\bar{z}$	$0, \quad 0, \quad 0$

producing the unique self-vectors :

$\pm 2x, \quad 0, \quad 2\bar{z}$		
$\frac{1}{2}, \quad \frac{1}{2}, \quad 0$		
$\frac{1}{2} \pm 2x, \quad \frac{1}{2}, \quad \pm 2z$		

For two scattering centres at (x, y, z) and (x', y', z'), cross-vectors are given by :

$x - x'$	$y - y'$	$z - z'$
$-x - x'$	$y - y'$	$-z - z'$
$\frac{1}{2}+x - x'$	$\frac{1}{2}+y - y'$	$z - z'$
$\frac{1}{2}-x - x'$	$\frac{1}{2}+y - y'$	$-z - z'$

$x + x'$	$y - y'$	$z + z'$
$-x + x'$	$y - y'$	$-z + z'$
$\frac{1}{2}+x + x'$	$\frac{1}{2}+y - y'$	$z + z'$
$\frac{1}{2}-x + x'$	$\frac{1}{2}+y - y'$	$-z + z'$
$x-\frac{1}{2} - x'$	$y-\frac{1}{2} - y'$	$z - z'$
$-x-\frac{1}{2}- x'$	$y-\frac{1}{2} - y'$	$-z - z'$
$x - x'$	$y - y'$	$z - z'$
$-x - x'$	$y - y'$	$-z - z'$
$x-\frac{1}{2} + x'$	$y-\frac{1}{2} - y'$	$z + z'$
$-x-\frac{1}{2}+ x'$	$y-\frac{1}{2} - y'$	$-z + z'$
$x + x'$	$y - y'$	$z + z'$
$-x + x'$	$y - y'$	$-z + z'$

Therefore the unique cross vectors in the asymmetric unit are given by :

$x + x'$	$y - y'$	$z + z'$
$x - x'$	$y - y'$	$z - z'$

In this way, cross vectors should appear in pairs of the same y-coordinate. Furthermore, in the hk0 zone there will will be a centre of symmetry at $\frac{1}{4}$, $\frac{1}{4}$.

Appendix IV : K2 Phases

The format of the data presented in this appendix is as described in Chapter VI viz. hkl, $\sin^2\Theta$, phase, scattering factor, figure of merit, F_P and then the heavy atom data. For each derivative there is a column for F_{PH}, Fobs and Fobs - Fcalc.

M1S1 Phases

H	K	L	S2	ALP	SCAT	FIG	FP	FPH(I)	DELF(I)	DODIF(I)			
~~0~~	~~12~~	~~0~~	~~0.0043~~	~~3.14~~	~~87.27~~	~~0.18~~	~~1353~~	~~1314~~	~~-6~~	~~-27~~	~~0~~	0	0
0	14	0	0.0043	3.14	87.01	0.54	1844	1548	-158	-56	1890	24	148
0	20	0	0.0088	3.14	86.02	0.61	1757	0	0	0	1589	-102	33
1	11	0	0.0027	6.28	87.38	0.72	1304	1690	278	208	1371	48	-29
2	12	0	0.0032	3.14	87.26	0.68	1962	2061	67	3	2193	157	65
2	14	0	0.0044	3.14	87.00	0.34	1508	1620	38	6	1546	13	-32
3	11	0	0.0028	3.14	87.36	0.51	2019	2039	10	-88	2202	93	-25
5	11	0	0.0029	3.14	87.33	0.42	1445	0	0	0	1331	-47	19
5	15	0	0.0052	6.28	86.81	0.80	1493	0	0	0	1875	304	179
5	17	0	0.0066	3.14	86.50	0.06	1625	0	0	0	1612	0	8
5	21	0	0.0100	6.28	85.77	0.36	1479	0	0	0	1582	37	-16
6	12	0	0.0036	6.28	87.19	0.92	1721	2004	260	84	2293	525	319
6	14	0	0.0047	6.28	86.93	0.93	1324	1955	588	411	1752	399	192
9	11	0	0.0035	3.14	87.19	0.94	1248	1787	505	325	1802	519	319
9	13	0	0.0046	3.14	86.96	0.81	2150	2160	8	-145	2598	363	191
10	10	0	0.0033	6.28	87.25	0.57	2085	1863	-125	-64	1984	-56	-47
10	12	0	0.0042	3.14	87.03	0.01	1979	1968	0	-1	2081	1	1
12	12	0	0.0047	6.28	86.93	0.62	2112	2219	66	-50	2279	104	-18
13	7	0	0.0029	6.28	87.34	0.17	1373	1412	6	-12	1249	-20	-17
13	13	0	0.0055	6.28	86.75	0.01	2176	2206	0	-1	2348	2	2
14	6	0	0.0029	6.28	87.35	0.58	1583	1701	68	113	1380	-117	6
14	8	0	0.0035	3.14	87.21	0.16	1490	1404	-13	-26	1566	12	-22
14	12	0	0.0052	6.28	86.81	0.52	1421	0	0	0	1291	-66	41
15	5	0	0.0029	3.14	87.33	0.23	1487	1539	12	-32	1525	8	-33
15	7	0	0.0034	6.28	87.21	0.77	1543	0	0	0	1251	-225	-85
15	9	0	0.0042	6.28	87.05	0.68	2679	2565	-76	51	2496	-122	-1
15	11	0	0.0050	3.14	86.86	0.27	1544	1658	31	-20	1532	-3	-51
15	13	0	0.0061	6.28	86.62	0.75	1664	0	0	0	1389	-206	-73
15	15	0	0.0073	3.14	86.35	0.54	1576	0	0	0	1739	88	-6
16	2	0	0.0028	3.14	87.36	0.09	1665	1577	-8	-19	1388	-26	-21
16	10	0	0.0049	6.28	86.89	0.41	1880	1578	-123	-76	1523	-146	-165
16	12	0	0.0059	6.28	86.67	0.56	1775	1575	-111	-48	1341	37	11
16	14	0	0.0070	3.14	86.42	0.08	1635	1623	0	-10	1539	-8	-4
17	3	0	0.0032	3.14	87.26	0.53	2753	2616	-71	-32	2652	-52	65
17	5	0	0.0036	3.14	87.18	0.84	1554	0	0	0	1208	-288	-101
17	7	0	0.0041	3.14	87.06	0.77	2051	2002	-37	20	1778	-210	-37
17	15	0	0.0080	6.28	86.20	0.50	1863	0	0	0	2001	69	-38
18	0	0	0.0034	3.14	87.22	0.48	1297	0	0	0	1244	-25	51
18	2	0	0.0035	3.14	87.20	0.40	1429	0	0	0	1323	-41	21
19	1	0	0.0038	3.14	87.13	0.65	2723	2300	-276	-199	2536	-122	-173
19	3	0	0.0040	3.14	87.09	0.35	2121	1986	-47	-6	2106	-5	-32
19	13	0	0.0088	3.14	86.31	0.33	1521	1465	-38	0	0	0	0
19	15	0	0.0088	3.14	86.04	0.04	1684	1753	3	7	1805	5	2
20	0	0	0.0042	0.00	87.04	1.00	1818	1659	-158	-88	1371	-446	-217
20	2	0	0.0043	3.14	87.02	0.31	2152	2067	-26	-47	2259	33	-37
20	4	0	0.0046	3.14	86.96	0.63	1509	1479	-18	-63	1709	139	-5
20	6	0	0.0050	6.28	86.86	0.64	1775	1718	-36	8	1400	-112	34
20	12	0	0.0074	6.28	86.34	0.77	1778	1790	9	62	1490	-222	-49
21	1	0	0.0047	6.28	86.94	0.10	2186	2152	-3	16	2186	1	14
21	3	0	0.0048	3.14	86.90	0.38	1393	1569	67	-4	1725	-25	-74
21	5	0	0.0052	6.28	86.82	0.59	1542	0	0	0	1345	-119	-41
22	0	0	0.0051	0.00	86.85	0.30	1604	0	0	0	1451	14	17
22	4	0	0.0054	3.14	86.77	0.16	1858	1766	-14	-34	1697	-26	-9
22	6	0	0.0059	3.14	86.67	0.17	1829	1721	-18	-38	1647	-31	-13
22	8	0	0.0073	[illegible]	86.5[illegible]	0.[illegible]3	1677	1752	17	-9	0	0	0
22	10	0	0.0050	3.14	86.56	0.02	1709	0	0	0	1683	-2	6
22	12	0	0.0083	3.14	86.15	0.46	1805	0	0	0	1630	-79	-32
23	3	0	0.0058	6.28	86.70	0.79	1703	1851	117	66	1999	274	54
23	5	0	0.0061	6.28	86.62	0.36	1782	1732	-17	-41	1893	[illegible]	-41
[illegible]	[illegible]	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

M2S2 Phases

H	K	L	S2	ALP	SCAT	FIG	FP	FPK(I)	DELF(I)	DODIF(I)		
0	12	0	0.0043	1.64	87.27	0.11	1353	1333	-2	15	0	0	0
0	14	0	0.0043	1.74	87.01	0.81	1844	1569	-223	-17	1884	33	82
0	20	0	0.0088	2.71	86.02	0.63	1757	0	0	0	1584	-108	68
1	11	0	0.0027	0.08	87.38	0.87	1304	1713	357	143	1367	55	39
2	12	0	0.0032	3.18	87.26	0.73	1962	2090	94	-36	2186	164	57
2	14	0	0.0044	3.72	87.00	0.52	1508	1643	70	-1	1541	17	-47
3	11	0	0.0028	3.29	87.36	0.63	2019	2068	31	-38	2195	111	10
5	11	0	0.0029	2.98	87.33	0.50	1445	0	0	0	1327	-58	74
5	15	0	0.0052	0.26	86.81	0.84	1493	0	0	0	1869	315	103
5	17	0	0.0066	3.61	86.50	0.11	1625	0	0	0	1607	-1	21
5	21	0	0.0102	0.80	85.77	0.34	1479	0	0	0	1579	83	18
6	12	0	0.0076	4.28	87.19	0.93	1721	2033	288	15	2286	523	279
6	14	0	0.0047	0.24	86.93	0.95	1324	1982	626	411	1747	402	159
9	11	0	0.0035	3.03	87.19	0.95	1248	1812	537	260	1797	522	210
9	13	0	0.0046	3.12	86.96	0.80	2150	2191	32	-159	2590	351	88
10	10	0	0.0035	5.68	87.25	0.54	2085	1890	-105	-78	1978	-57	-10
10	12	0	0.0042	2.76	87.03	0.29	1979	1995	4	-13	2075	28	2
12	12	0	0.0047	5.72	86.93	0.70	2112	2250	97	4	2272	112	-66
13	7	0	0.0029	0.77	87.34	0.57	1373	1431	33	0	1245	-72	-26
13	13	0	0.0055	0.11	86.75	0.51	2174	2237	31	-71	2341	83	26
14	6	0	0.0029	1.05	87.35	0.08	1503	1724	124	23	1376	-181	-20
14	8	0	0.0035	4.72	87.21	0.60	1490	1423	-39	27	1561	43	-13
14	12	0	0.0052	0.11	86.81	0.42	1421	0	0	0	1287	-55	-8
15	5	0	0.0029	3.85	87.33	0.33	1487	1560	24	-51	1520	10	-38
15	7	0	0.0034	0.08	87.21	0.79	1543	0	0	0	1247	-233	-95
15	9	0	0.0042	0.12	87.05	0.64	2679	2602	-49	-47	2489	-121	-14
15	11	0	0.0050	1.38	86.86	0.53	1544	1681	73	19	1527	-8	5
15	13	0	0.0061	6.27	86.62	0.74	1664	0	0	0	1385	-211	-91
15	15	0	0.0073	3.09	86.35	0.54	1576	0	0	0	1734	85	0
16	2	0	0.0028	3.63	87.36	0.74	1665	1600	-48	-22	1384	-209	-122
16	10	0	0.0049	5.82	86.89	0.85	1880	1601	-237	-28	1518	-308	-280
16	12	0	0.0059	0.18	86.67	0.66	1775	1598	-117	35	1836	40	57
16	14	0	0.0070	1.46	86.42	0.07	1635	1646	0	6	1534	-47	-5
17	3	0	0.0032	2.68	87.26	0.54	2753	2653	-53	-38	2644	-58	66
17	5	0	0.0036	3.58	87.18	0.86	1554	0	0	0	1204	-299	-55
17	7	0	0.0041	3.71	87.06	0.82	2051	2031	-16	10	1773	-228	-38
17	15	0	0.0080	5.72	86.20	0.45	1843	0	0	0	1995	59	0
18	0	0	0.0034	3.14	87.22	0.78	1997	0	0	0	1240	-43	166
18	2	0	0.0035	3.35	87.20	0.48	1439	0	0	0	1319	-52	72
19	1	0	0.0036	3.39	87.13	0.86	2723	2333	-336	-167	2528	-167	-194
19	3	0	0.0040	3.50	87.09	0.47	2121	2014	-50	31	2100	-9	-24
19	13	0	0.0088	2.16	86.31	0.37	1591	1486	-35	10	0	0	0
19	15	0	0.0088	4.46	86.04	0.52	1684	1778	49	-2	1800	60	32
20	0	0	0.0042	0.00	87.04	1.00	1918	1682	-135	38	1367	-450	-179
20	2	0	0.0043	2.23	87.02	0.74	2112	2096	-18	-33	2252	34	-20
20	4	0	0.0046	3.00	86.94	0.69	1509	1500	-5	-55	1724	148	-18
20	6	0	0.0050	0.31	86.86	0.71	1775	1743	-22	1	1595	-127	35
20	12	0	0.0074	0.09	86.34	0.89	1778	1815	30	-32	1486	-232	-95
21	1	0	0.0047	5.53	86.94	0.01	2186	2183	0	0	2189	0	1
21	3	0	0.0048	1.57	86.90	0.83	1393	1592	164	37	1301	-58	-8
21	5	0	0.0052	0.58	86.82	0.70	1548	0	0	0	1341	-144	-6
22	0	0	0.0051	3.14	86.85	0.06	1604	0	0	0	1646	2	1
22	4	0	0.0054	1.66	86.77	0.68	1858	1791	-45	37	1692	-113	-40
22	6	0	0.0059	1.94	86.67	0.74	1829	1746	-61	13	1642	-137	-28
22	8	0	0.0073	4.86	86.53	0.28	1677	1777	27	6	0	0	0
22	10	0	0.0058	2.79	86.53	0.17	1709	0	0	0	1678	-5	30
22	12	0	0.0082	3.07	86.15	0.65	1805	0	0	0	1625	-116	26
23	3	0	0.0058	0.08	86.70	0.85	1703	1878	148	36	1993	247	104
23	5	0	0.0061	5.72	86.62	0.53	1732	1757	-13	9	1987	55	-52
24	0	0	0.0066	0.00	86.43	0.85	2057	1960	-41	-71	2111	26	-27
24	12	0	0.0093	3.05	85.91	0.55	2039	2096	-79	37	2325	45	80
24	14	0	0.0101	3.19	85.69	0.58	1638	1539	-57	44	1564	-74	-28
25	3	0	0.0065	3.63	86.17	0.16	1487	1710	7	-17	1655	-1	7
25	11	0	0.0074	2.40	85.94	0.27	1744	1703	-17	17	0	0	0

M3S3 Phases

H	K	L	S2	ALP	SCAT	FIG	FF	FPH(I)	DELF(I)	DODIF(I)		
0	12	0	0.0043	2.43	87.27	0.04	1353	1327	-1	0	0	0	0
0	14	0	0.0043	1.35	87.01	0.92	1844	1563	-257	-4	1884	-37	2
0	20	0	0.0088	2.57	86.02	0.68	1757	0	0	0	1584	-117	93
1	11	0	0.0027	0.22	87.38	0.90	1304	1707	364	59	1367	57	48
2	12	0	0.0032	3.43	87.26	0.78	1942	2082	94	-16	2186	176	113
2	14	0	0.0044	3.42	87.00	0.58	1508	1636	74	-53	1541	19	-7
3	11	0	0.0028	2.89	87.36	0.66	2019	2060	27	-38	2195	115	-37
5	11	0	0.0029	3.31	87.43	0.55	1445	0	0	0	1327	-63	70
5	15	0	0.0052	0.10	86.81	0.87	1493	0	0	0	1869	328	153
5	17	0	0.0064	3.31	86.50	0.12	1625	0	0	0	1607	-2	25
5	21	0	0.0100	0.76	85.77	0.45	1479	0	0	0	1577	44	-40
6	12	0	0.0036	4.15	87.19	0.94	1721	2025	284	5	2286	533	197
6	14	0	0.0047	0.27	86.93	0.96	1394	1974	624	314	1747	407	71
9	11	0	0.0035	3.03	87.19	0.96	1248	1385	535	175	1797	527	150
8	17	0	0.0046	3.10	86.86	0.85	2150	2199	27	-111	2590	375	34
10	10	0	0.0033	5.70	87.05	0.80	2085	1982	-161	-55	1978	-84	-24
10	12	0	0.0042	3.36	87.03	0.30	1979	1987	2	-19	2075	29	7
10	14	0	0.0047	5.57	86.93	0.73	2110	2041	94	14	2272	117	-71
13	7	0	0.0029	0.90	87.34	0.73	1371	1426	38	12	1245	-92	-34
13	13	0	0.0055	0.21	86.75	0.55	2174	2328	29	-77	2341	91	11
14	6	0	0.0029	0.75	87.35	0.87	1583	1718	118	-22	1376	-180	9
14	8	0	0.0035	4.65	87.21	0.77	1490	1418	-55	25	1561	55	-8
14	12	0	0.0052	4.19	86.81	0.59	1421	0	0	0	1287	-78	9
15	5	0	0.0029	4.24	87.33	0.55	1487	1554	36	-74	1520	18	-36
15	7	0	0.0034	0.05	87.21	0.85	1543	0	0	0	1247	-252	-47
15	9	0	0.0042	0.35	87.05	0.79	2679	2591	-68	17	2489	-149	18
15	11	0	0.0050	0.98	86.86	0.57	1544	1475	25	-3	1527	-9	39
15	13	0	0.0061	6.22	86.62	0.82	1664	0	0	0	1385	-226	-29
15	17	0	0.0093	2.94	86.35	0.43	1576	0	0	0	1734	99	-10
16	2	0	0.0028	3.82	87.36	0.84	1665	1573	-59	-46	1384	-235	-177
16	10	0	0.0049	5.06	86.89	0.92	1880	1596	-262	-129	1518	-337	-208
16	12	0	0.0059	5.61	86.67	0.67	1775	1591	-122	29	1836	41	101
16	14	0	0.0070	1.41	86.42	0.30	1635	1639	1	37	1534	-29	-6
17	3	0	0.0032	3.53	87.26	0.48	2753	2642	-52	-47	2644	-51	29
17	5	0	0.0044	3.01	87.10	0.89	1554	0	0	0	1204	-318	-54
17	7	0	0.0041	4.86	87.04	0.83	2051	2023	-93	17	1773	-249	20
17	15	0	0.0080	4.01	86.30	0.57	1863	0	0	0	1995	75	-48
18	0	0	0.0034	3.13	86.42	0.84	1297	0	0	0	1240	-47	0
18	2	0	0.0035	3.14	87.20	0.52	1429	0	0	0	1319	-57	-19
18	8	0	0.0068	3.40	87.13	0.93	3723	3583	-369	-190	2528	-190	-138
18	10	0	0.0080	3.46	87.09	0.65	2121	2004	-74	34	2100	-13	-5
19	11	0	0.0038	1.66	86.81	0.44	1581	1480	-44	-2	0	0	0
19	13	0	0.0058	4.29	86.64	0.32	1664	1771	27	-28	1800	36	56
20	0	0	0.0062	0.00	86.54	1.00	1818	1676	-141	46	1367	-450	-47
20	2	0	0.0043	2.17	87.02	0.40	2152	2088	-25	-28	2252	40	-14
20	4	0	0.0046	3.28	86.96	0.65	1509	1494	-9	-73	1324	139	-37
20	6	0	0.0050	0.50	86.84	0.77	1775	1776	-29	28	1595	-138	0
20	12	0	0.0079	4.03	86.34	0.85	1778	1808	25	-51	1486	-247	-23
21	1	0	0.0047	4.18	86.94	0.01	2186	2174	0	0	2189	0	1
21	3	0	0.0048	1.20	86.90	0.89	1793	1585	172	32	1321	-63	-4
21	5	0	0.0052	0.76	86.82	0.75	1548	0	0	0	1341	-154	45
22	0	0	0.0061	5.14	86.65	0.50	1604	0	0	0	1446	21	-44
22	6	0	0.0091	1.76	86.22	0.80	1858	1794	-64	0	1692	-145	15
22	8	0	0.0109	1.44	86.62	0.93	1999	1737	-93	-7	1642	-172	-21
22	10	0	0.0127	5.34	86.53	0.64	1577	1770	60	-8	0	0	0
23	10	0	0.0138	0.03	0.12	0.19	1799	0	0	0	1678	-5	31

M4S3 Phases

0	14	0	0.0043	1.33	87.01	0.95	1844	1559	-269	-47	1867	21	16
0	20	0	0.0088	2.65	86.02	0.71	1757	0	0	0	1570	-132	45
1	11	0	0.0027	0.29	87.38	0.91	1304	1702	362	17	1354	45	64
2	12	0	0.0032	3.28	87.26	0.82	1962	2076	93	-19	2166	167	111
2	14	0	0.0044	3.43	87.00	0.43	1508	1632	53	-70	1527	8	-9
3	11	0	0.0028	2.75	87.36	0.56	2019	2054	19	-48	2175	87	-44
5	11	0	0.0029	3.39	87.33	0.59	1445	0	0	0	1316	-76	67
5	15	0	0.0052	0.13	86.81	0.86	1493	0	0	0	1852	308	153
5	17	0	0.0066	3.30	86.50	0.20	1625	0	0	0	1593	-6	34
5	21	0	0.0100	0.72	85.77	0.40	1479	0	0	0	1563	33	-35
6	12	0	0.0036	6.11	87.19	0.95	1721	2019	282	-5	2266	517	182
6	14	0	0.0047	0.24	86.93	0.97	1324	1969	627	233	1731	394	58
9	11	0	0.0035	3.07	87.19	0.97	1248	1800	533	123	1781	517	139
9	13	0	0.0046	3.08	86.96	0.88	2150	2176	22	-53	2567	366	23
10	10	0	0.0033	5.75	87.25	0.87	2085	1877	-179	-22	1961	-107	-49
10	12	0	0.0042	3.40	87.03	0.19	1979	1982	0	-18	2056	14	2
12	12	0	0.0047	5.42	86.93	0.86	2112	2235	105	-2	2252	120	-45
13	7	0	0.0029	0.98	87.34	0.84	1373	1422	41	1	1234	-116	-37
13	13	0	0.0055	0.78	86.75	0.43	2176	2222	19	-59	2319	61	7
14	6	0	0.0029	0.59	87.35	0.94	1583	1713	121	-20	1363	-206	-2
14	8	0	0.0035	4.84	87.21	0.69	1490	1414	-52	27	1547	39	18
14	12	0	0.0052	6.14	86.81	0.62	1421	0	0	0	1276	-89	4
15	5	0	0.0029	4.21	87.33	0.54	1487	1550	33	-81	1506	10	-27
15	7	0	0.0034	0.02	87.21	0.86	1543	0	0	0	1236	-264	-61
15	9	0	0.0042	0.31	87.05	0.84	2679	2584	-79	24	2467	-178	1
15	11	0	0.0050	0.73	86.86	0.62	1544	1670	77	-7	1513	-19	79
15	13	0	0.0061	0.02	86.62	0.84	1664	0	0	0	1372	-245	-106
15	15	0	0.0073	3.01	86.35	0.59	1576	0	0	0	1718	83	-21
16	2	0	0.0028	4.02	87.36	0.95	1665	1589	-72	-20	1371	-279	-213
16	10	0	0.0049	6.05	86.89	0.94	1880	1590	-273	-80	1504	-353	-284
16	12	0	0.0059	5.74	86.67	0.77	1775	1587	-145	16	1820	34	102
16	14	0	0.0070	1.62	86.42	0.46	1635	1635	0	26	1520	-52	-17
17	3	0	0.0032	2.17	87.26	0.91	2753	2635	-106	-80	2621	-120	-46
17	5	0	0.0036	3.82	87.18	0.90	1554	0	0	0	1193	-324	-72
17	7	0	0.0041	4.06	87.06	0.87	2051	2017	-29	-28	1757	-255	-106
17	15	0	0.0080	5.96	86.20	0.51	1863	0	0	0	1978	58	-39
18	0	0	0.0034	3.14	87.22	0.89	1297	0	0	0	1229	-60	132
18	2	0	0.0035	3.13	87.20	0.57	1429	0	0	0	1308	-68	63
19	1	0	0.0038	3.58	87.13	0.95	2723	2317	-385	-90	2505	-207	-166
19	3	0	0.0040	3.90	87.09	0.73	2121	2000	-88	-9	2081	-29	0
19	15	0	0.0088	5.70	86.04	0.40	1684	1766	32	-29	1784	40	60
20	0	0	0.0042	6.28	87.04	1.00	1818	1671	-147	-58	1354	-464	-58
20	2	0	0.0043	2.20	87.02	0.86	2152	2082	-59	4	2232	68	-9
20	4	0	0.0046	3.36	86.96	0.70	1509	1490	-13	-30	1708	139	-48
20	6	0	0.0050	0.58	86.86	0.77	1775	1731	-33	26	1581	-149	-13
20	12	0	0.0074	5.79	86.34	0.89	1778	1803	22	-49	1473	-271	-59
21	1	0	0.0047	0.46	86.94	0.15	2186	2168	-2	2	2169	-2	23
21	3	0	0.0048	1.24	86.90	0.90	1393	1581	169	31	1310	-74	9
21	5	0	0.0052	0.25	86.82	0.77	1548	0	0	0	1328	-169	31
22	0	0	0.0051	3.14	86.85	0.34	1604	0	0	0	1632	9	-33
22	4	0	0.0054	1.35	86.77	0.96	1858	1779	-75	4	1676	-174	3
22	6	0	0.0059	2.09	86.67	0.96	1829	1734	-90	11	1628	-192	29
22	10	0	0.0058	3.40	86.53	0.27	1709	0	0	0	1662	-12	42
22	12	0	0.0083	3.62	86.15	0.71	1805	0	0	0	1611	-137	-38
23	3	0	0.0058	0.29	86.70	0.79	1703	1865	128	-65	1976	215	31
23	5	0	0.0061	0.09	86.62	0.72	1782	1745	-26	24	1870	63	-48
24	0	0	0.0060	6.28	86.63	0.25	2053	1947	-26	0	2092	9	-7
24	12	0	0.0092	2.42	85.94	0.79	2239	2072	-132	8	2305	52	19
24	14	0	0.0104	2.58	85.69	0.76	1638	1529	-82	30	1492	-110	-84
25	3	0	0.0068	5.17	86.47	0.54	1687	1699	6	-16	1641	-24	77
27	5	0	0.0082	4.61	86.16	0.93	1736	1826	83	47	1521	-199	-77

M4S4 Phases

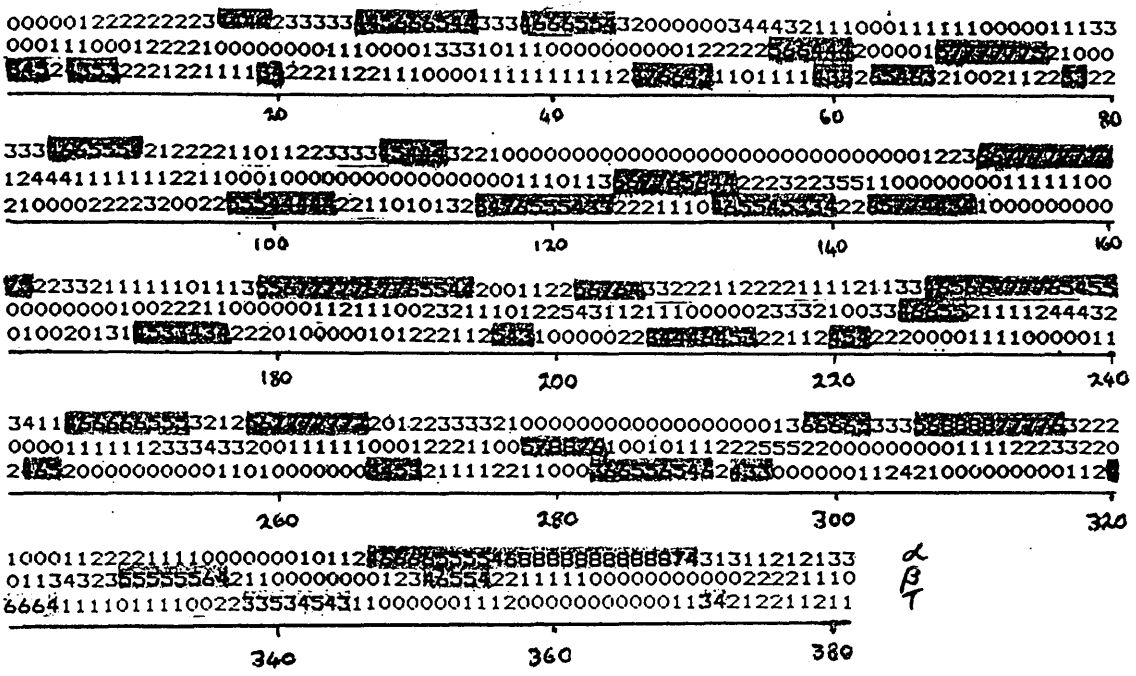
H	K	L	S2	ALP	SCAT	FIG	FP	FPH(I)	DELF(I)	DODIF(I)		
0	12	0	0.0043	1.15	87.27	0.02	1353	1350	0	0	0	0	0
0	14	0	0.0043	1.33	87.01	0.97	1844	1590	-246	11	1871	26	0
0	20	0	0.0088	2.96	86.02	0.76	1757	0	0	0	1573	-140	-26
1	11	0	0.0027	0.33	87.38	0.96	1304	1736	413	35	1357	51	44
2	12	0	0.0032	3.13	87.26	0.90	1962	2117	139	9	2171	188	152
2	14	0	0.0044	3.36	87.00	0.51	1508	1664	79	-75	1530	11	2
3	11	0	0.0028	3.03	87.36	0.68	2019	2095	51	-45	2180	110	-22
5	11	0	0.0029	3.23	87.33	0.60	1445	0	0	0	1318	-76	70
5	15	0	0.0052	0.24	86.81	0.90	1493	0	0	0	1856	326	133
5	17	0	0.0066	3.55	86.50	0.20	1625	0	0	0	1596	-5	36
5	21	0	0.0100	1.13	85.77	0.47	1479	0	0	0	1566	41	-21
6	12	0	0.0036	5.97	87.19	0.97	1721	2059	326	29	2270	531	189
6	14	0	0.0047	0.24	86.93	0.98	1324	2008	670	258	1735	402	50
9	11	0	0.0035	3.08	87.19	0.98	1248	1836	573	136	1784	523	132
9	13	0	0.0046	3.30	86.96	0.91	2150	2219	63	-4	2572	384	-14
10	10	0	0.0033	5.81	87.25	0.81	2085	1914	-137	40	1965	-96	-42
10	12	0	0.0042	3.47	87.03	0.47	1979	2021	20	-37	2060	38	7
12	12	0	0.0047	5.47	86.93	0.96	2112	2279	161	45	2256	139	12
13	7	0	0.0029	0.85	87.34	0.95	1373	1450	73	-1	1237	-128	-27
13	13	0	0.0055	0.50	86.75	0.57	2176	2266	51	-68	2324	84	-16
14	6	0	0.0029	0.63	87.33	0.96	1383	1747	157	-13	1566	206	-8
14	8	0	0.0035	4.83	87.21	0.74	1490	1442	-35	12	1550	44	7
14	12	0	0.0052	6.26	86.81	0.67	1421	0	0	0	1278	-95	-6
15	5	0	0.0029	4.57	87.33	0.60	1487	1581	56	-70	1509	13	-18
15	7	0	0.0034	6.25	87.21	0.88	1543	0	0	0	1239	-266	-138
15	9	0	0.0042	6.12	87.05	0.91	2679	2635	-39	6	2472	-187	-42
15	11	0	0.0050	0.83	86.86	0.87	1544	1703	137	23	1516	-23	40
15	13	0	0.0061	6.10	86.62	0.88	1664	0	0	0	1375	-254	-41
15	15	0	0.0073	3.17	86.35	0.57	1576	0	0	0	1722	82	-64
16	2	0	0.0028	3.92	87.36	0.98	1665	1620	-43	-12	1374	-284	-100
16	10	0	0.0049	6.26	86.89	0.95	1880	1621	-244	-59	1507	-352	-232
16	12	0	0.0059	5.36	86.67	0.84	1775	1618	-131	35	1823	40	29
16	14	0	0.0070	2.10	86.42	0.90	1635	1667	29	-2	1523	-99	-2
17	3	0	0.0032	2.41	87.26	0.91	2753	2687	-59	-38	2626	-114	-12
17	5	0	0.0036	3.85	87.18	0.92	1554	0	0	0	1196	-328	-73
17	7	0	0.0041	3.98	87.06	0.94	2051	2057	5	-2	1760	-273	-109
17	15	0	0.0080	5.77	86.20	0.58	1863	0	0	0	1982	69	-35
18	0	0	0.0034	3.14	87.22	0.83	1297	0	0	0	1232	-54	59
18	2	0	0.0035	2.84	87.20	0.62	1429	0	0	0	1310	-73	40
19	1	0	0.0038	3.67	87.13	0.93	2723	2363	-333	10	2510	-196	-57
19	3	0	0.0040	3.97	87.09	0.57	2121	2040	-46	2	2085	-20	53
19	13	0	0.0088	1.74	86.31	0.55	1581	1505	-41	34	0	0	0
19	15	0	0.0088	6.29	86.04	0.85	1684	1801	99	-30	1787	88	-2
20	0	0	0.0042	0.00	87.04	1.00	1818	1704	-113	76	1357	-460	84
20	2	0	0.0043	2.15	87.02	0.54	2152	2123	-15	8	2237	45	-14
20	4	0	0.0046	3.48	86.96	0.66	1509	1519	7	-14	1712	134	-77
20	6	0	0.0050	1.41	86.86	0.75	1775	1765	-7	44	1584	-142	0
20	12	0	0.0074	5.67	86.34	0.95	1778	1839	57	-11	1476	-286	0
21	1	0	0.0047	6.02	86.94	0.10	2186	2211	2	2	2174	-1	14
21	3	0	0.0048	1.07	86.90	0.94	1393	1612	206	79	1312	-75	-10
21	5	0	0.0052	0.19	86.82	0.79	1548	0	0	0	1331	-170	47
22	0	0	0.0051	3.14	86.85	0.41	1604	0	0	0	1635	12	-29
22	4	0	0.0054	1.60	86.77	0.97	1858	1814	-42	0	1680	-173	-2
22	6	0	0.0059	2.39	86.67	0.97	1829	1768	-58	4	1631	-191	19
22	8	0	0.0073	6.24	86.53	0.70	1677	1800	86	13	0	0	0
22	10	0	0.0058	3.33	86.53	0.28	1709	0	0	0	1666	-12	40
22	12	0	0.0083	3.49	86.15	0.77	1805	0	0	0	1614	-146	-41
23	3	0	0.0058	0.15	86.70	0.89	1703	1902	177	-21	1980	246	67
23	5	0	0.0061	0.43	86.62	0.84	1782	1779	-1	-5	1874	77	-60
24	0	0	0.0060	0.00	86.63	0.75	2053	1985	-50	40	2096	32	-89
24	12	0	0.0092	2.61	85.94	0.73	2239	2113	-91	36	2309	51	38
24	14	0	0.0104	2.38	85.69	0.91	1638	1559	-71	1	1495	-129	-37
25	3	0	0.0068	5.71	86.47	0.83	1687	1732	38	-56	1644	-35	57
25	11	0	0.0074	1.14	85.94	0.29	1761	1725	-10	32	0	0	0
27	5	0	0.0082	4.52	86.16	0.97	1736	1862	123	44	1524	-205	-61

Appendix V : The Predicted Secondary Structure of CK

Using the full amino acid sequence of rabbit muscle CK (1), the secondary structure was predicted using a combination of seven separate methods [Burgess et al. (2); Chou and Fassman (3), as modified by Lenstra et al. (4) and with the updated indices of Geisow and Roberts (5); Dufton and Hider (6); Garnier et al. (7); Lim (8); McLachlan (9) and Nagano (10)] as suggested by Eliopoulos et al. (11). The rules for the final prediction were as described by Sawyer el al. (12) i.e. (a) a residue is considered likely to adopt a helical or sheet conformation if at least four of the above methods agree that it is so; (b) a sheet or helix is deemed to exist only if there are at least four consecutive residues with that prediction; (c) a turn is deemed to exist if there are at least two consecutive residues predicted by at least three methods (d) all remaining residues are considered to be irregular.

The result so obtained for CK is depicted overleaf. The digits indicate the number of methods supporting that prediction. Sequences obeying the rules outlined above are highlighted accordingly. The sequence is analysed for prediction of α-helix (α), β-sheet (β) and reverse turns (T) as indicated. For a discussion of this result see Section 8.3.6.

I would like to thank Dr L. Sawyer for providing the facilities to carry out this work.



Secondary structure prediction for CK

1) Pickering, L., Pang, H., Biemann, k., Munro, H. and Schimmel, P. (1985) Proc. Nat. Acad. Sci. USA, 82, 2310

2) Burgess, A.W., Ponnuswamy, P.K. and Scheraga, H.A. (1974) Isr. J. Chem., 12, 239

3) Chou, P.Y. and Fassman, G.D. (1974) Biochem., 13, 222

4) Lenstra, J.A., Hofsteenge, J. and Beintema, J.J. (1977) J. Mol. Biol., 109, 185

5) Geisow, M.J. and Roberts, R.D.B. (1980) Int. J. Biol. Macromol., 2, 387

6) Dufton, M.J. and Hider, R.C. (1977) J. Mol. Biol., 115, 177

7) Garnier, M.J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol., 120, 97

8) Lim, V.I. (1974) J. Mol. Biol., 88, 873

9) McLachlan, A.D. (1977) Int. J. Quantum Chem., 12, Supp 1, 371

10) Nagano, K. (1973) J. Mol. Biol., 75, 401

11) Eliopoulos, E.E., Geddes, A.J., Brett, M., Papin, D.J.C. and Findlay, J.B.C. (1982) Int. J. Biol. Macromol., 4, 263

12) Sawyer, L., Fothergill-Gilmore, L.A. and Russell, G.A. (1986) Biochem. J., 236, 127