Bull. Mater. Sci., Vol. 17, No. 1, February 1994, pp. 1–18. © Printed in India.

Electron density maps of lysozyme calculated using synchrotron Laue data comprising singles and deconvoluted multiples

JOHN W CAMPBELL[1], ASHLEY DEACON[2], JARJIS HABASH[2], JOHN R HELLIWELL[2]*, SEAN McSWEENEY[1], HAO QUAN[1], JAMES RAFTERY[2] and EDWARD SNELL[2]

[1]Daresbury Laboratory, SERC, Daresbury, Warrington WA4 4AD, England

[2]Department of Chemistry, University of Manchester, Manchester M13 9PL, England

MS received 1 October 1993

Abstract. We have used lysozyme as a test case to illustrate the application of a new method of estimating the intensities of Laue multiples reflection data in protein crystallography. Hen egg-white lysozyme, an enzyme with a single polypeptide chain of 129 amino acids, crystallizes in space group $P4_32_12$ with cell parameters $a = b = 79{\cdot}1$ Å, $c = 37{\cdot}9$ Å. Laue image plate data were collected using synchrotron radiation on station 9·5 at Daresbury with a total exposure time of 0·95 seconds. The data processed were separated into two data sets comprising the singles and the combined singles and deconvoluted multiples. The method of deconvolution was that of Campbell and Hao (1993), which utilizes the intensity variation of the λ-curve. Electron density maps ($2F_o - F_c$) based on the two data sets are then compared. This comparison shows that the deconvoluted multiples do indeed contribute usefully to the continuity of the maps due to the improved completeness of the data. A number of map sections along the polypeptide chain, based on the two Laue data sets, are shown for comparison. These include Arg 5, His 15, Phe 38, Asp 52, Tyr 53, Pro 70, Trp 108 and the four disulphide bridges.

Keywords. Synchrotron radiation; Laue diffraction; hen egg-white lysozyme; image plate; X-PLOR; electron density; amino acids.

1. Introduction

The synchrotron Laue method has opened up the field of time-resolved macromolecular crystallography (Allinson *et al* 1992; Cruickshank *et al* 1992; Pai 1992; Blow 1993). The method is particularly apposite for time-resolved studies on timescales of seconds or subseconds so as to detect catalytic processes in enzyme crystals.

Reflecting planes in a stationary crystal, when illuminated by a white beam (wavelength range λ_{min} to λ_{max}) satisfy Bragg's law resulting in spots stimulated at different wavelengths. If reflections are members of the same harmonic series i.e. **h**, 2**h**, 3**h**, etc., then they will be coincident on the same spot, forming a so-called energy-overlap multiple-spot. The percentage of single-order spots depends on the ratio $\lambda_{max}/\lambda_{min}$. It rises from 83% when the ratio is infinite to 88% when the ratio is 3 and reaches 100% when this ratio tends to unity, i.e. the monochromatic limit (Cruickshank *et al* 1987).

Structure solution using standard Patterson and direct methods procedures with Laue singles data alone has proved successful in small molecule studies (Clucas *et al* 1988; Harding *et al* 1988; Gomez de Anderez *et al* 1989; M Helliwell *et al* 1989a).

The structure refinement of a cobalt derivative of concanavalin A in I222 space group at 2 Å resolution using Laue singles data has also been realized (Cassetta *et al*

*For correspondence

1993). Although the parallel refinement at 1·6 Å for the monochromatic data (Emmerich *et al* 1993) showed better continuity in the $2F_o - F_c$ electron density, the Laue maps were of good quality.

$F_o - F_c$ Fourier syntheses using Laue data have proved interpretable in protein structural studies even allowing single water molecules to be seen (Lindahl *et al* 1992; Singer *et al* 1993; Cameron *et al* 1993). Improvements to the analysis procedures, so as to include multiples, thereby increasing the total number of reflection data, have been made. One method for deconvoluting the multiple spots based on the use of the differential absorption by the multiple films in a film pack has been available for some time and has been applied (Helliwell *et al* 1989b; Duke *et al* 1992; Weisgerber and Helliwell 1993) using the program UNSCRAM (Zurek *et al* 1985). New approaches based on direct methods (Hao *et al* 1993) and via the wavelength normalization curve (Campbell and Hao 1993) have been introduced which allow unscrambling of multiplets even when only a single film or an image plate in one exposure is employed, in contrast to the earlier method of differential absorption.

Until the advent of the image plate or storage phosphor, film was the only detector used in Laue crystallography. Detectors such as the image plate have a high detective quantum efficiency and a wide dynamic range, which results in a great reduction in exposure time (Amemiya 1990) over film, especially for short wavelengths. We report here as an example of protein crystallography using Laue image plate data from lysozyme as a test case, and particularly involving the use of multiples deconvoluted by the new procedure of Campbell and Hao (1993).

2. Experimental

Lysozyme crystallizes in space group $P4_32_12$ with unit cell parameters $a = b = 79{\cdot}1$ Å, $c = 37{\cdot}9$ Å. A crystal of dimension $0{\cdot}14 \times 0{\cdot}27 \times 0{\cdot}79\ \mathrm{mm}^3$ was used to collect Laue data on station 9·5 at the SRS Daresbury Laboratory in the polychromatic mode (Brammer *et al* 1988). The SRS operated at 2 GeV and 230 mA with a 5 tesla wiggler. The wavelength band-pass was 0·5–2·2 Å. The crystal to detector distance used was 216·7 mm. The collimator size was 0·3 mm. Data were collected on a 90 mm radius Image Plate (Amemiya and Miyahara 1988) system of MAR Research, which uses on-line scanning (1200 × 1200 pixels and 0·150 mm pixel size).

Four images were collected with the crystal set at angles of − 15, 0, 15 and 40° about the spindle axis (see figure 1 for one example). The exposure time was 0·20 sec for the first exposure and 0·25 sec for each of the other three exposures, and the real time for data collection was less than 30 min. The time to clear the image for this device is 90 sec. The total time taken could have been reduced substantially if required.

3. Data reduction

The four synchrotron Laue images were processed using the Daresbury Laboratory suite (Helliwell *et al* 1989b) updated in 1991 for dealing with image plate data.

The computer programs SPOTIN and NEWLAUE were used to determine the orientation of the crystal. SPOTIN was used to select high angle nodals (about six spots) to be used by NEWLAUE to auto-index the image.

GENLAUE used the orientation angles to predict the diffraction pattern to 2·5 Å

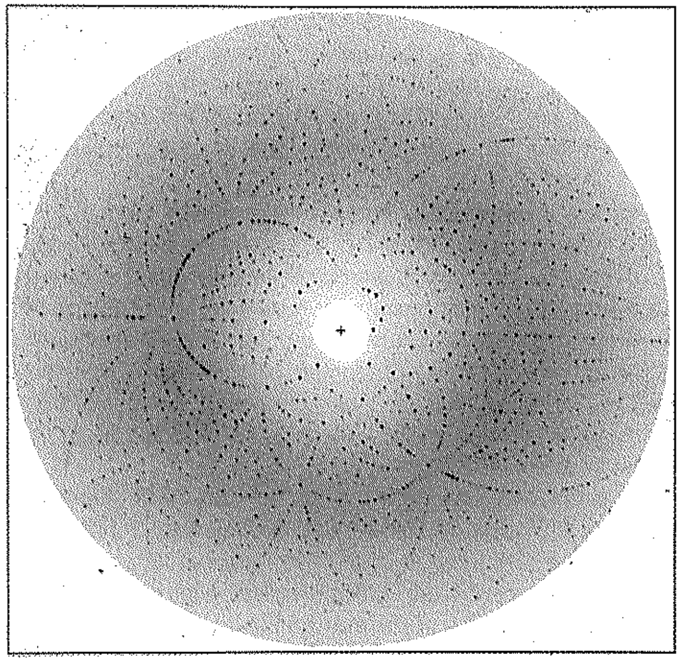


Figure 1. Synchrotron Laue diffraction pattern of lysozyme recorded on an image plate on station 9·5 at Daresbury (exposure time 0·25 sec).

resolution within the wavelength range 0·5–2·2 Å for each image. An average of 140 nodal spots were used to refine the misorientation angles. Spots were classified as spatially overlapped when their centres were within 0·4 mm of each other. These were small in number, comprising approximately a few hundred out of ~5000 for each image.

INTLAUE, with the profile fit option, was used to measure the intensity and standard deviation of each spot. Spatially overlapped spots were rejected from measurement. There were no overloaded spots.

The computer program LAUENORM was used for the wavelength normalization procedure in order to scale together intensities stimulated at different wavelengths. This procedure produces a λ-curve, which contains the combined information of the distribution of beam intensity with energy as well as other λ-dependent effects on the intensity of spots. This program relies on the intensity of symmetry equivalent reflections measured at different wavelengths to determine the scale (Campbell *et al* 1986), assuming there is no significant anomalous scatterer at the wavelengths measured.

For obtaining the scaled singles data set, spots stimulated by wavelengths within the range 0·49–1·6 Å were selected, and divided into 10 bins. Reflections stimulated by the shorter wavelengths are both weakly scattered and detected while spots stimulated at longer wavelengths also become weak due to absorption. The peak of the λ-curve, with this experimental arrangement (including the image plate), is at 1·1 Å. Reflections whose intensities were less than $1{\cdot}5\sigma$ were also rejected. The bulk of the accepted reflections were between wavelengths 0·49 and 1·16 Å with the most dense concentration occurring between 0·6 and 1·05 Å. This agrees with a recent suggestion that in Laue crystallography the wavelength range be limited to an octave, i.e. a factor of two, with 0·75–1·50 Å as the optimal choice (Sweet *et al* 1993), as well as our own work suggesting a range of 0·6–1·6 Å (Helliwell *et al* 1989b). By this means the intensity errors at both ends of wavelength range are reduced. A total of 7336 normalized reflection data were output with a final merging R-factor on intensity of 0·070.

For obtaining the deconvoluted multiples a new modified version of the LAUENORM program was used (Campbell and Hao 1993). Advantage can be taken of the λ-curve to deconvolute intensities of component reflections within multiple spots. This can be done when a multiple has been repeatedly measured with the crystal in more than one orientation. To deconvolute a double, two or more such orientations are needed, three or more for a triple, and so on. In the present case, the use of four crystal orientations and the presence of high symmetry enable the method to be used to deconvolute many of the multiples data. A detailed description of the procedure is given in Campbell and Hao (1993).

So as to increase the number of multiples which may be deconvoluted a wider wavelength range was now used, i.e. 0·485–1·9 Å, divided into 20 bins. Reflections whose intensities were less than 3σ were rejected. The number of multiples deconvoluted yielded 415 reflections (derived from 226 multiples). Figure 2 shows the number of multiples versus multiplicity; 85% of the multiple spots were doubles and 13% triples, with only two quadruples and one pentuple. The maximum multiplicity actually seen was 10. By this procedure, about 56% of the multiples found were deconvoluted, i.e. 226 out of 404 multiples.

These deconvoluted multiple reflections were scaled, on F, to an existing diffractometer data set to test their quality. There were 391 reflections common to the two data sets. The overall $\Delta F/F$ value was 0·229, ranging from 0·15 at 6·3 Å to 0·18 at 3·65 Å and rising to 0·43 at 2·5 Å. For comparison the singles alone data set was scaled to the diffractometer set. There were 2551 reflections in common. The

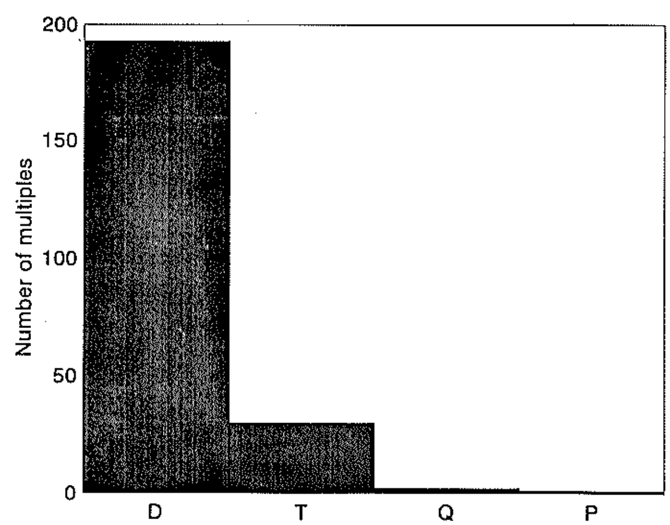


Figure 2. Plot of number of deconvoluted multiple spots vs multiplicity (D(doubles), T(triples), Q(quadruples) and P(pentuples)).

overall $\Delta F/F$ value was then 0·12, varying from 0·12 at 5·0 Å to 0·106 at 3·16 Å and rising to 0·152 at 2·5 Å.

The programs ROTAVATA/AGROVATA were used to merge the singles and deconvoluted multiple reflections and obtain a unique set of the data. In addition a second data file was scaled by separately using the singles alone to allow comparison. For the singles alone a total of 2573 reflections were obtained to 2·5 Å resolution with a merging R-factor on intensity of 0·068. The completeness of the data between $\infty - d_{min}$ was 58·4%, between $\infty - 2d_{min}$ 0%, and between $2d_{min} - d_{min}$ 68%. The value of zero between $\infty - 2d_{min}$ is unusual and was due primarily to a scanner limit, preventing image plate data being read near the centre of the image. A more typical value is 16%. The lower than usual overall completeness of the data (i.e. normally

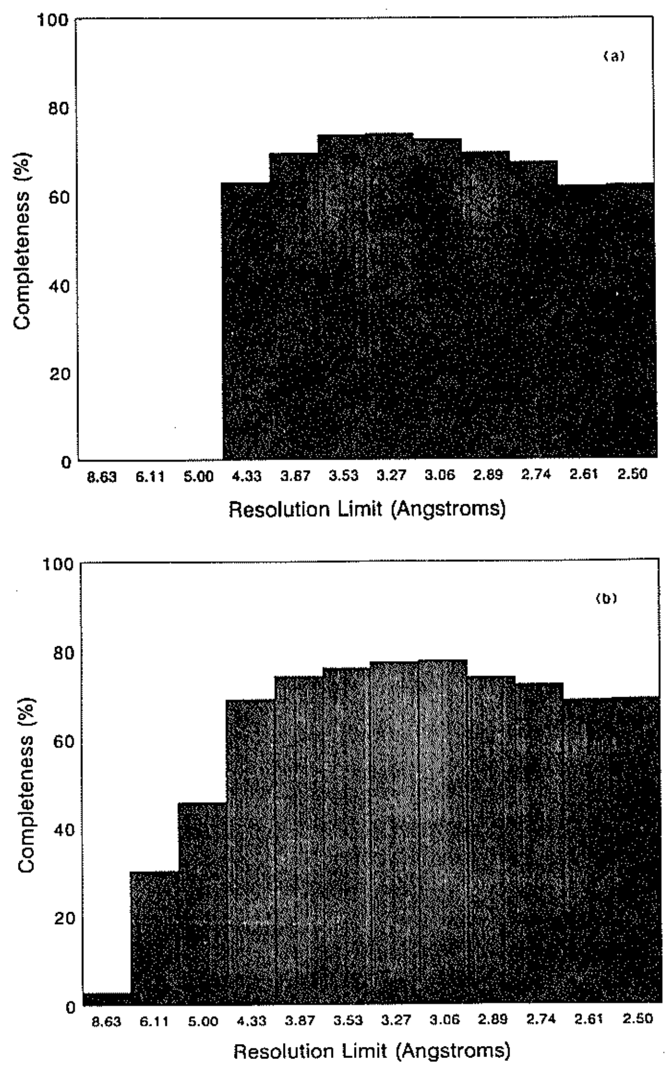


Figure 3. Plot of the completeness vs the resolution (d_{min}) for the data sets based on (a) singles alone and (b) combined singles and deconvoluted multiples.

70%) was also due to the scanning limit. 94·1% of all the data were greater than 3 standard deviations between ∞–d_{min}. For the combined data set comprising singles and deconvoluted multiples the data set contained a total of 2963 unique reflections with the merging R-factor on intensity of 0·081 to 2·5 Å. The completeness of the data between ∞–d_{min} was 67·3%, ∞–$2d_{\text{min}}$ 31%, and $2d_{\text{min}}$–d_{min} 73·3%. Figure 3 shows the completeness versus the resolution for each data set.

4. Structure refinement and calculation of $2F_o - F_c$ electron density maps

The original coordinates deposited at the Protein Data Bank by R Diamond, D C Phillips, C C F Blake and A C T North in February 1975 (6LYZ) were refined with the X-PLOR program suite (Brünger 1990) against the two data sets and used to generate the structure factors. The X-PLOR refinement included 40 cycles of positional refinement followed by 20 cycles of temperature (B-factor) refinement.

Using the singles data in the refinement the R-factor dropped from an initial value of 0·349 to 0·197. After four iterations of X-PLOR the R-factor converged to a final value of 0·144. The root mean square (r.m.s.) deviation for bond lengths = 0·014 Å, bond angles = 3·1° and torsion angles = 23·4°. In a separate calculation and using the combined singles and multiples data in the refinement the R-factor dropped from an initial value of 0·350 at the start of the refinement to 0·215. After four iterations of X-PLOR the R-factor converged to a final value of 0·163. The r.m.s. deviation for bond lengths = 0·016 Å, bond angles = 3·4° and torsion angles = 23·5°.

The $2F_o - F_c$ map calculations, for the singles data set and the combined singles and multiples data set, were done by using the CCP4 (1979) suite of programs. These maps were inspected on an ESV30 computer graphics workstation using the FRODO program (Jones 1978).

5. Results and discussion

By inspection of the $2F_o - F_c$ maps for the two data sets using the same contouring levels (1·0σ), there is a general consensus that the maps based on the full data set are better defined, along both the main chain and side chains.

In figure 4 we show Arg 5, His 15, Phe 38, Asp 52, Tyr 53, Pro 70 and Trp 108. The four disulphide bridges Cys 6–Cys 127, Cys 30–Cys 115, Cys 64–Cys 80, and Cys 76–Cys 94 are also shown for comparison.

In most cases it is evident from the maps that the deconvoluted multiples contribute to the improvement of the map features. Interestingly, in some cases there is no improvement. While the singles alone can yield maps adequate for interpretation, the combination of singles and deconvoluted multiples enhances the quality of the maps.

6. Conclusions

We have shown the usefulness of the procedure of Campbell and Hao (1993) for deconvoluting multiples in synchrotron Laue image plate data and including these data with wavelength normalized Laue singles data in the calculation of $(2F_o - F_c)$

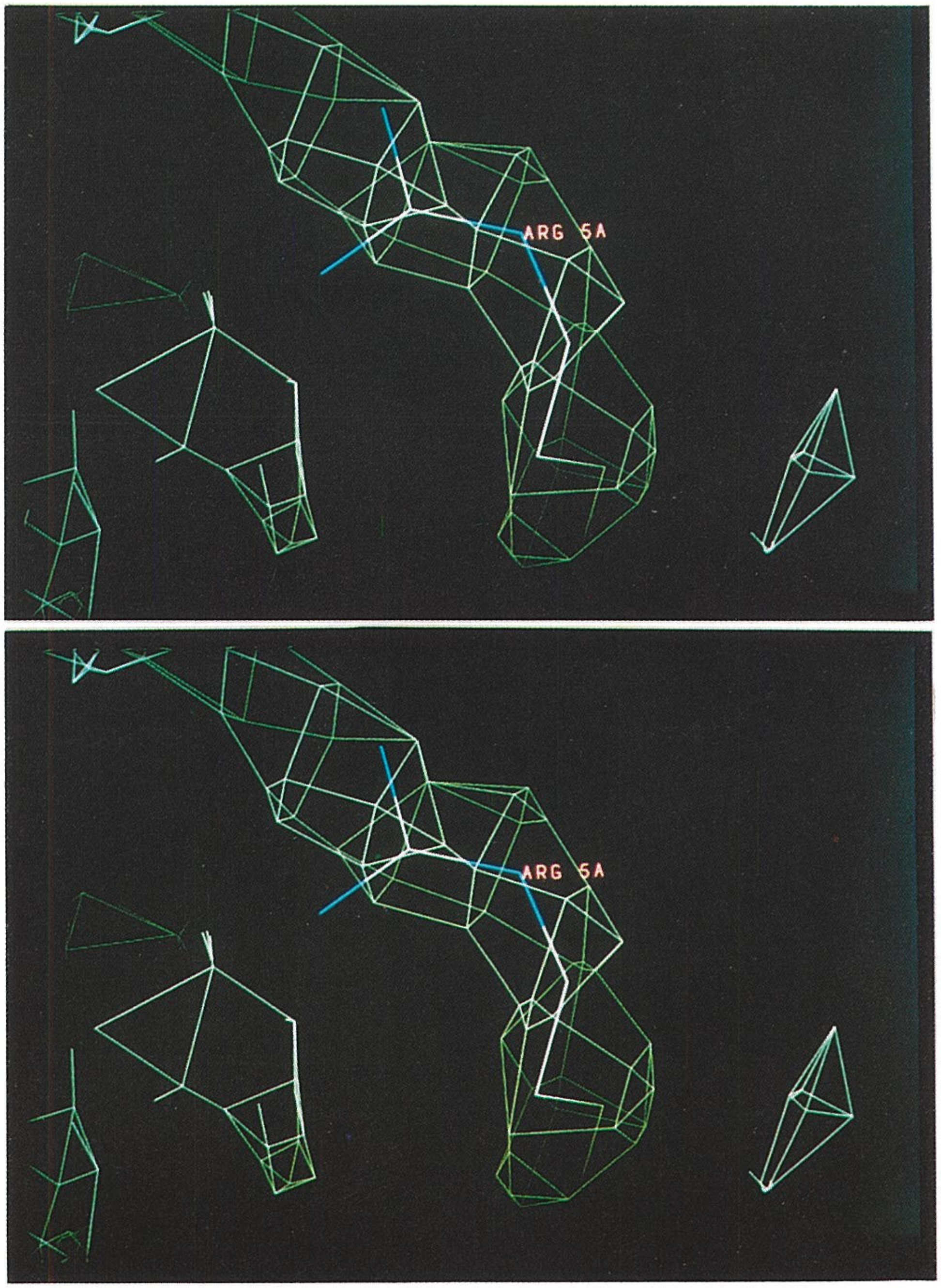


Figure 4. For caption, see p. 17.

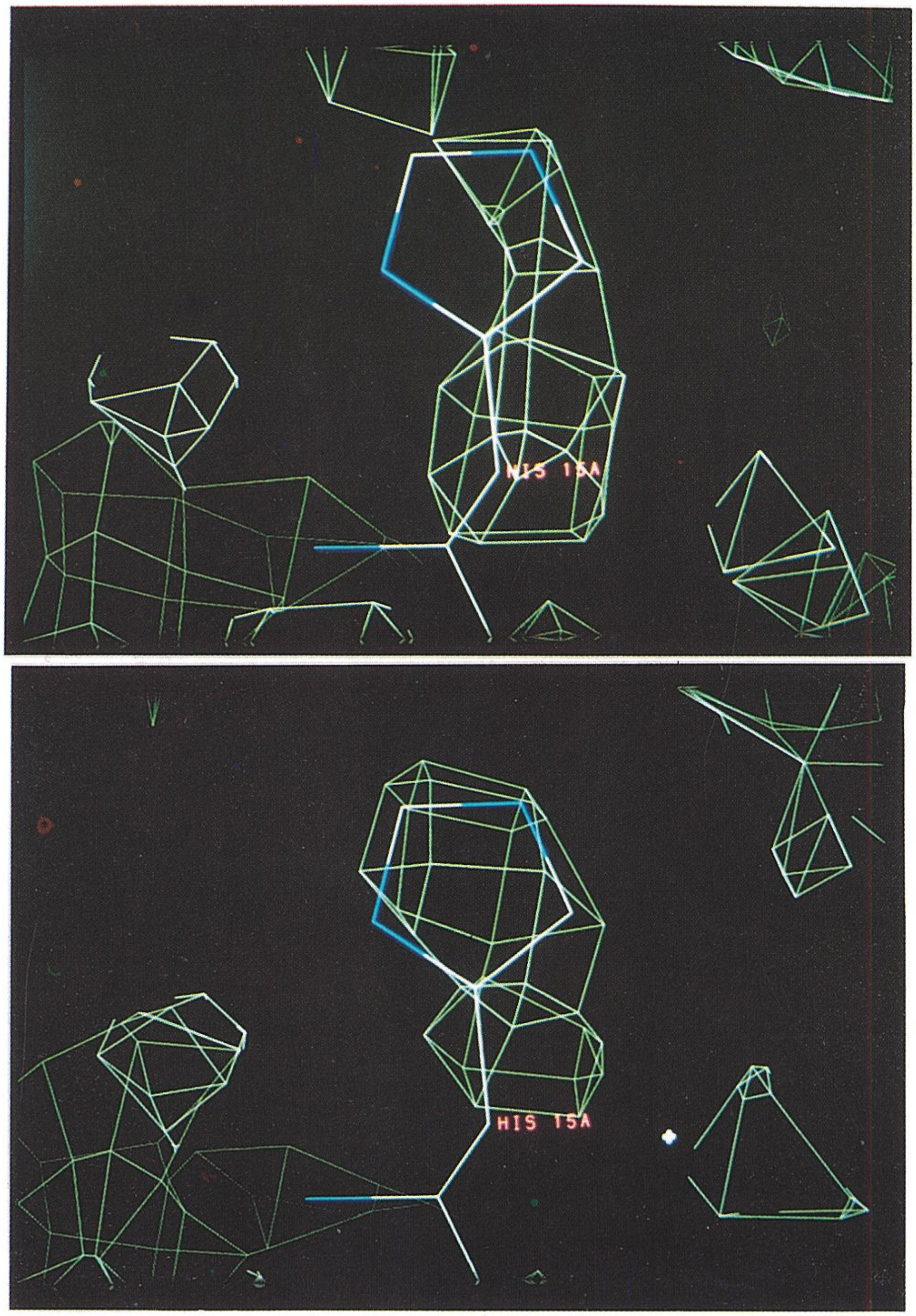


Figure 4. For caption, see p. 17.

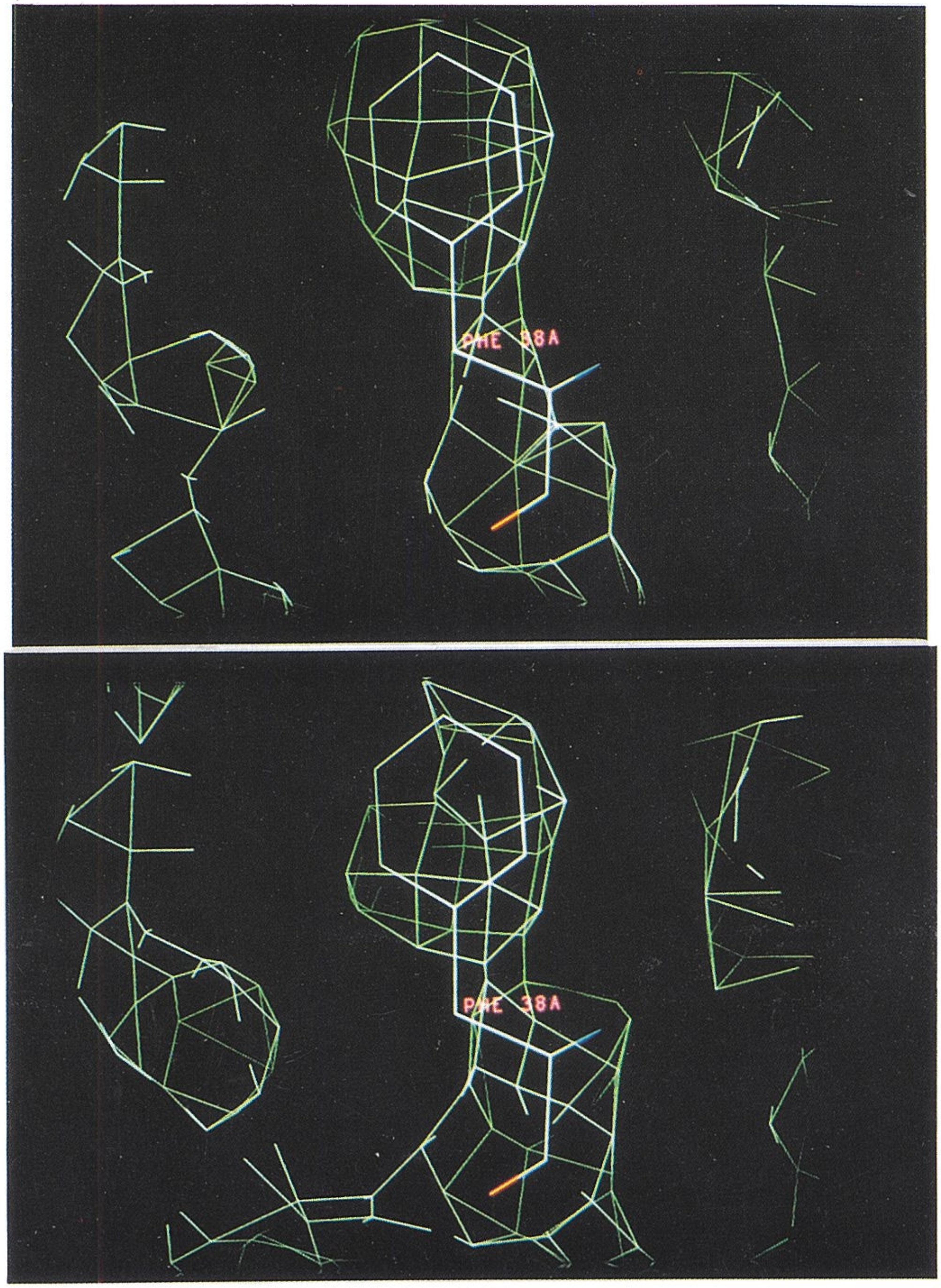


Figure 4. For caption, see p. 17.

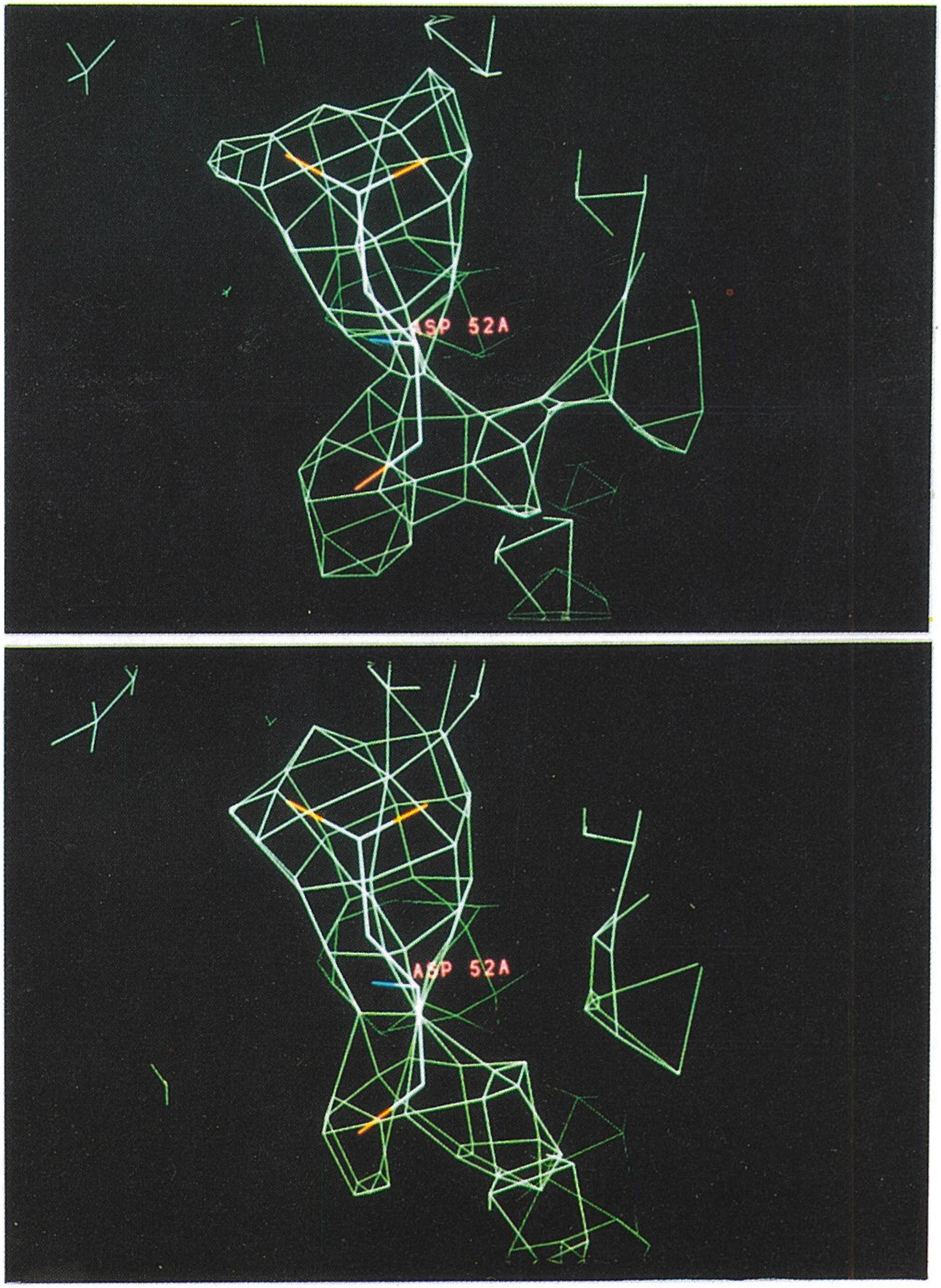


Figure 4. For caption, see p. 17.

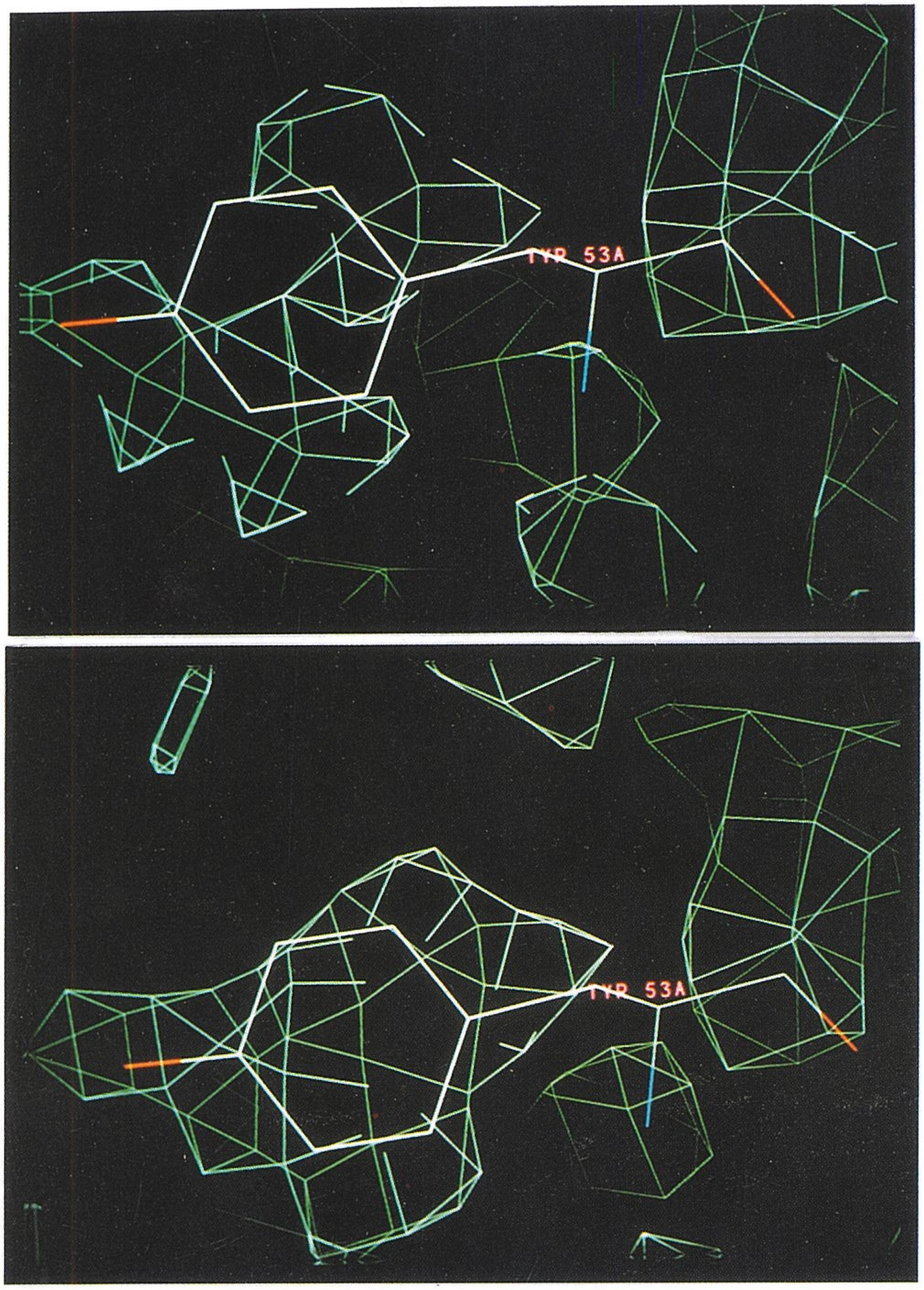


Figure 4. For caption, see p. 17.

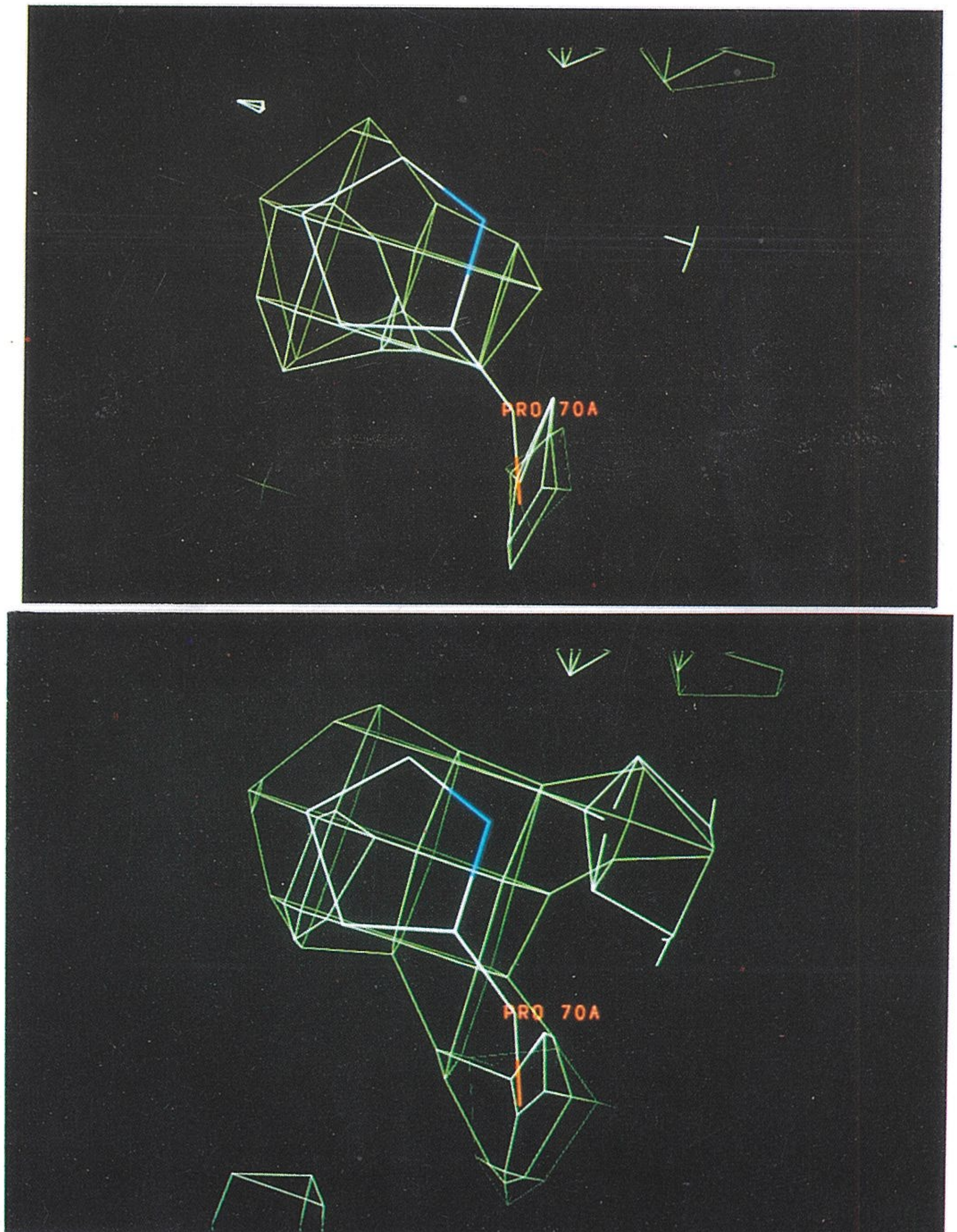


Figure 4. For caption, see p. 17.

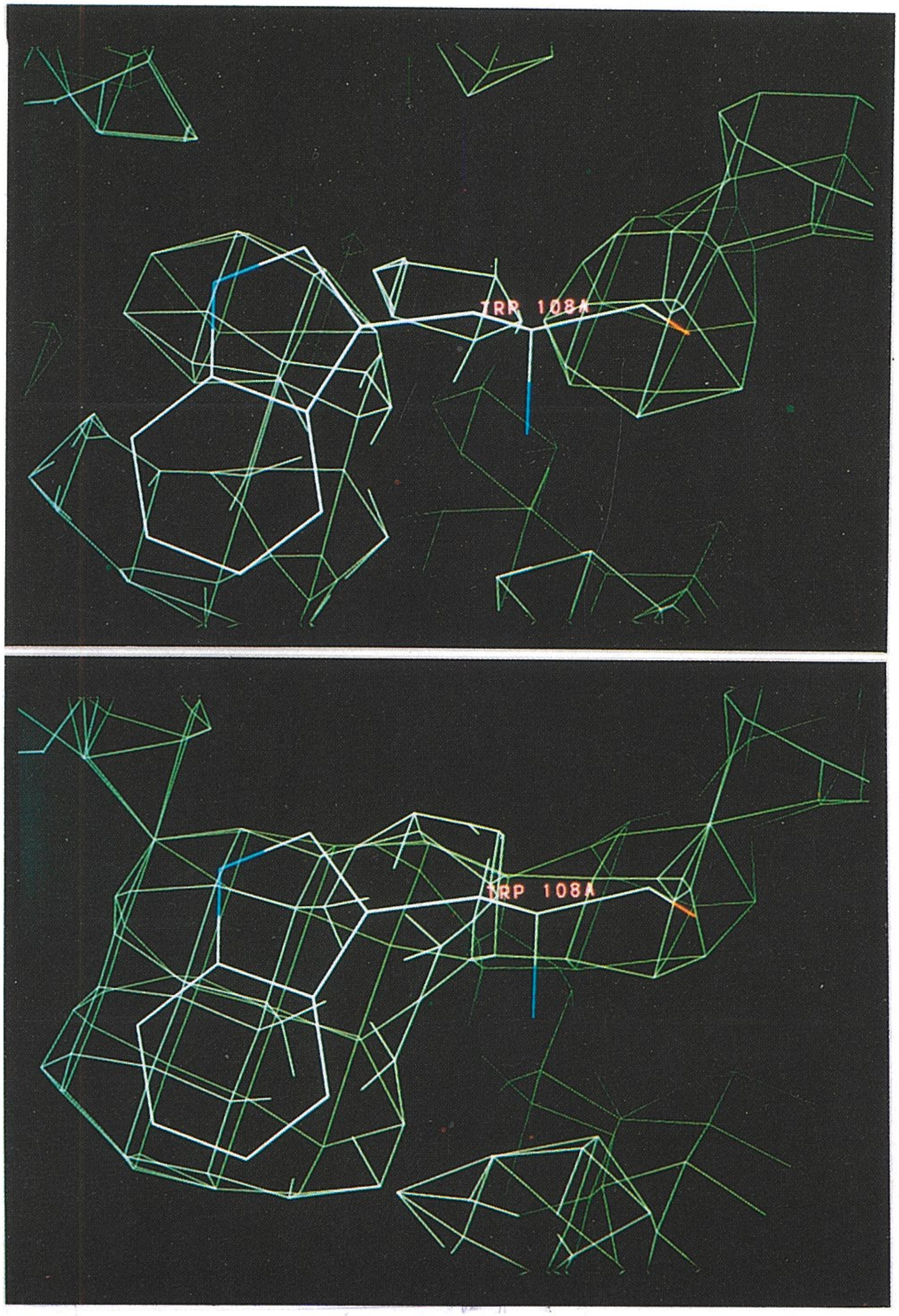


Figure 4. For caption, see p. 17.

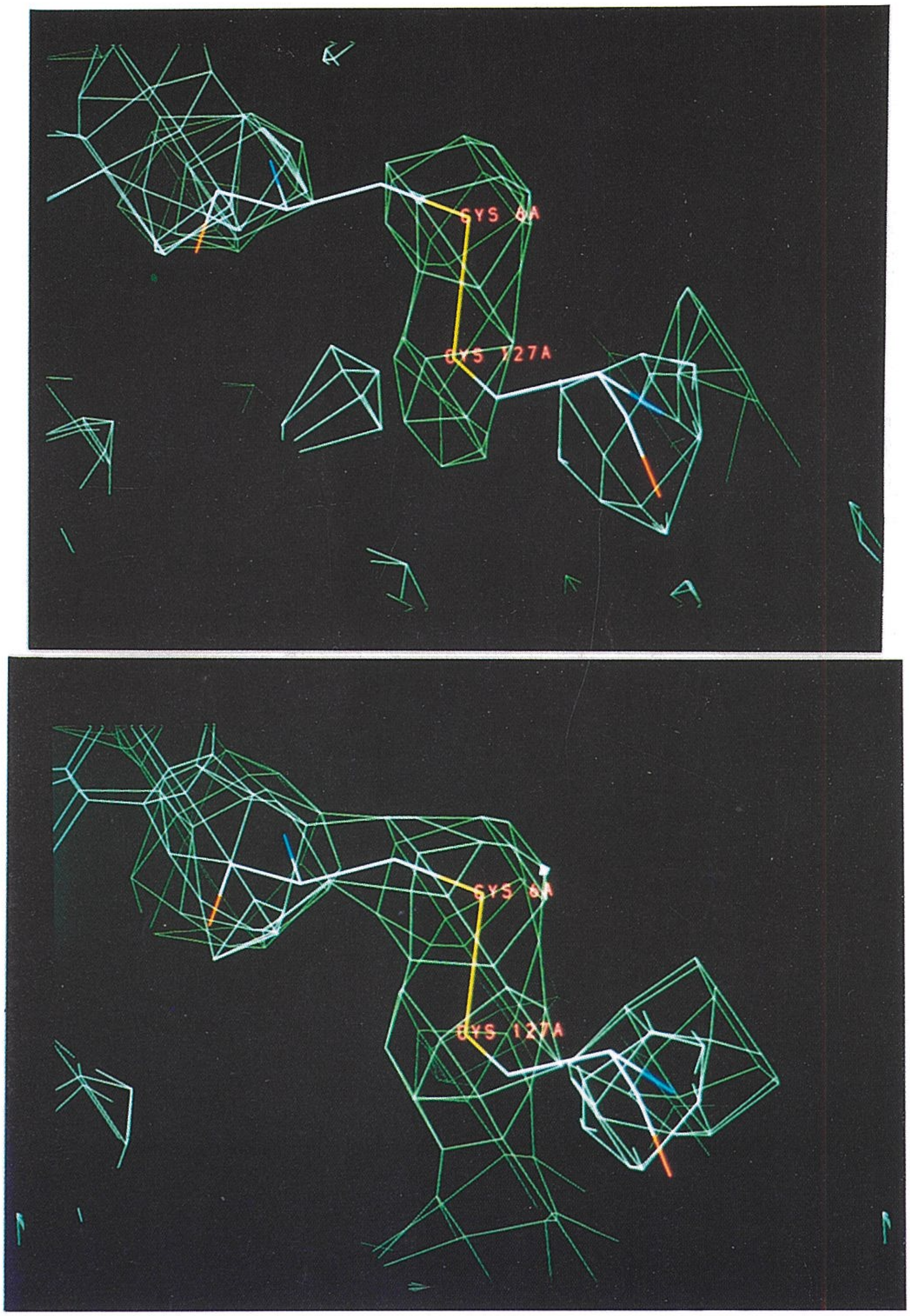


Figure 4. For caption, see p. 17.

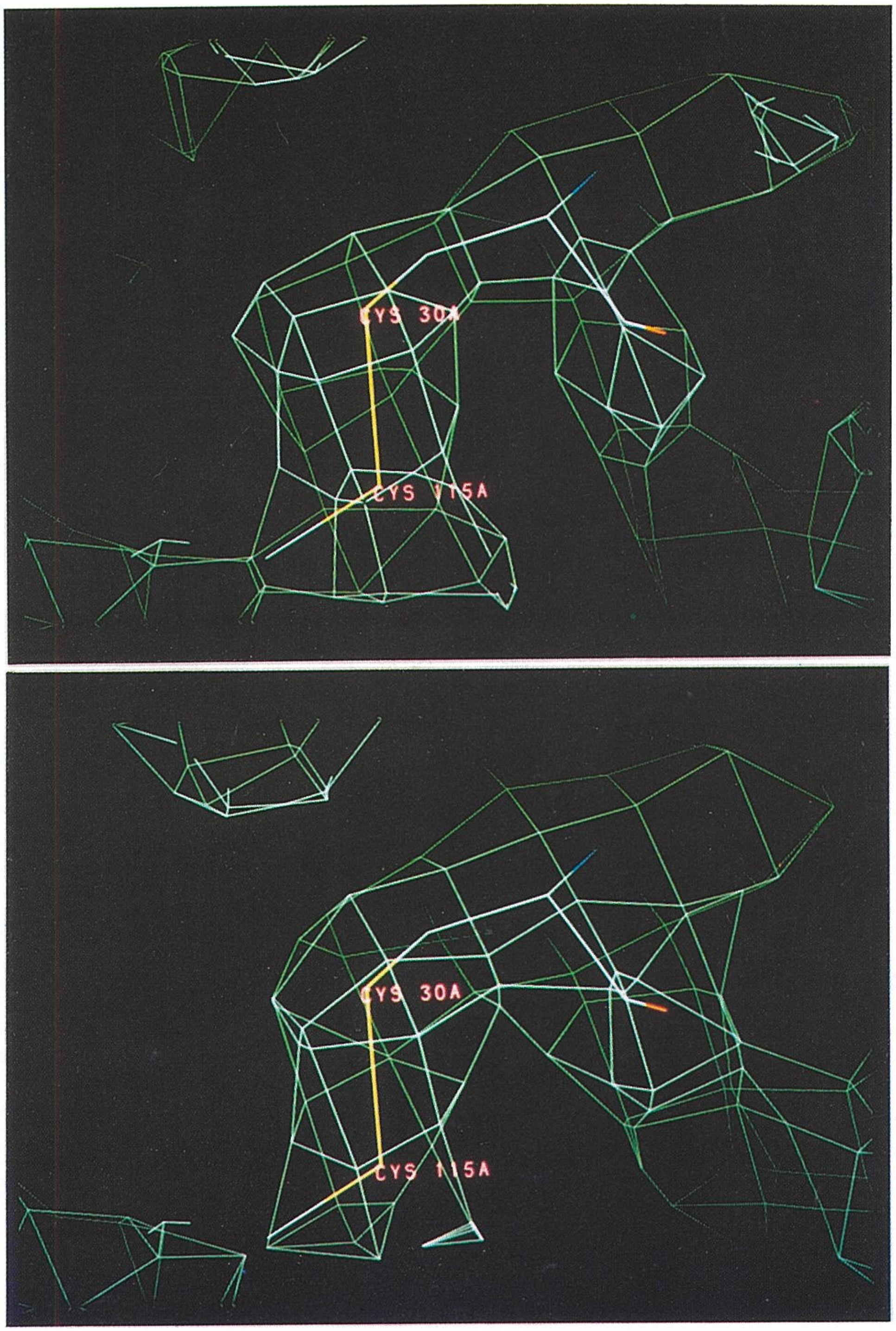


Figure 4. For caption, see p. 17.

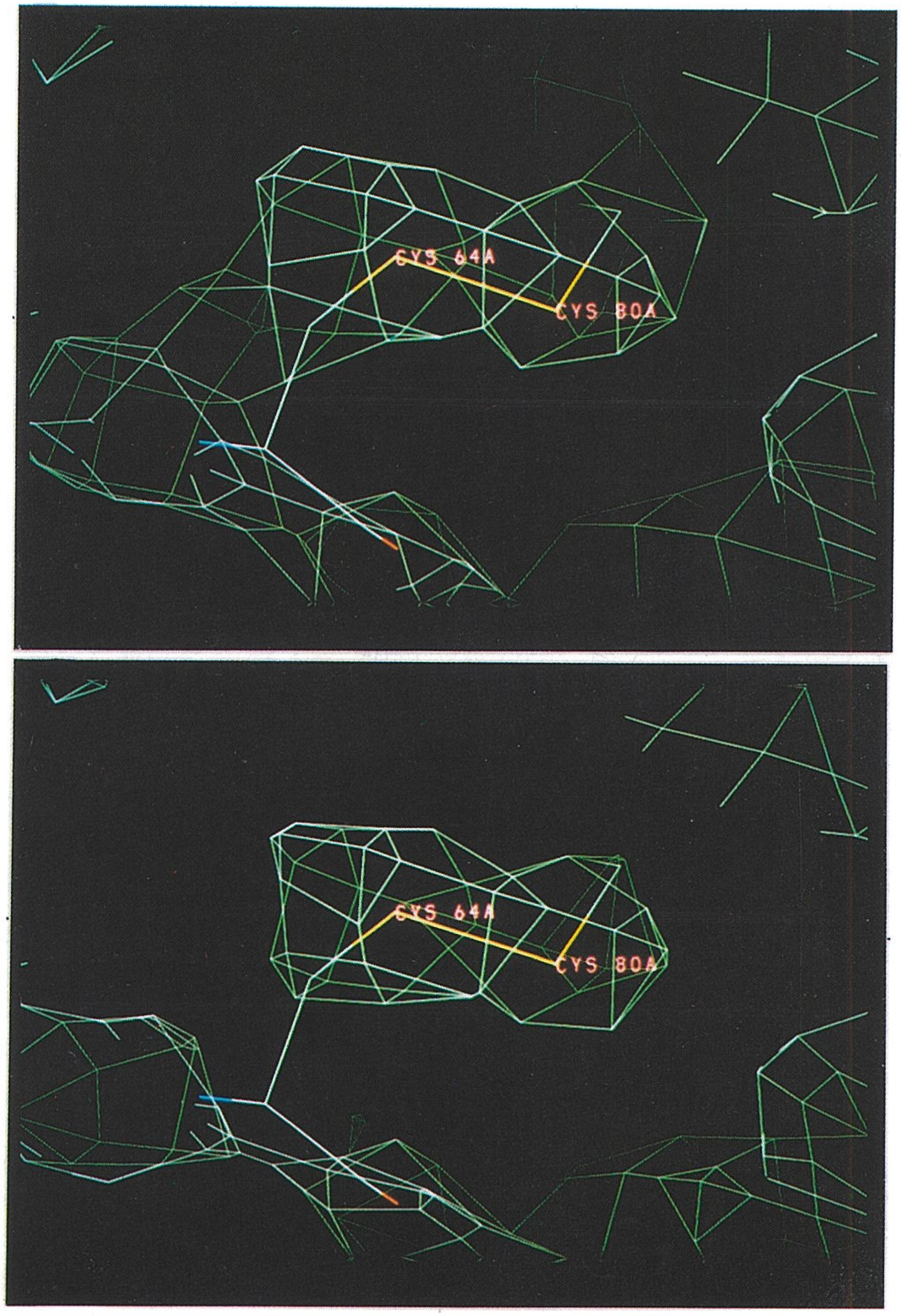
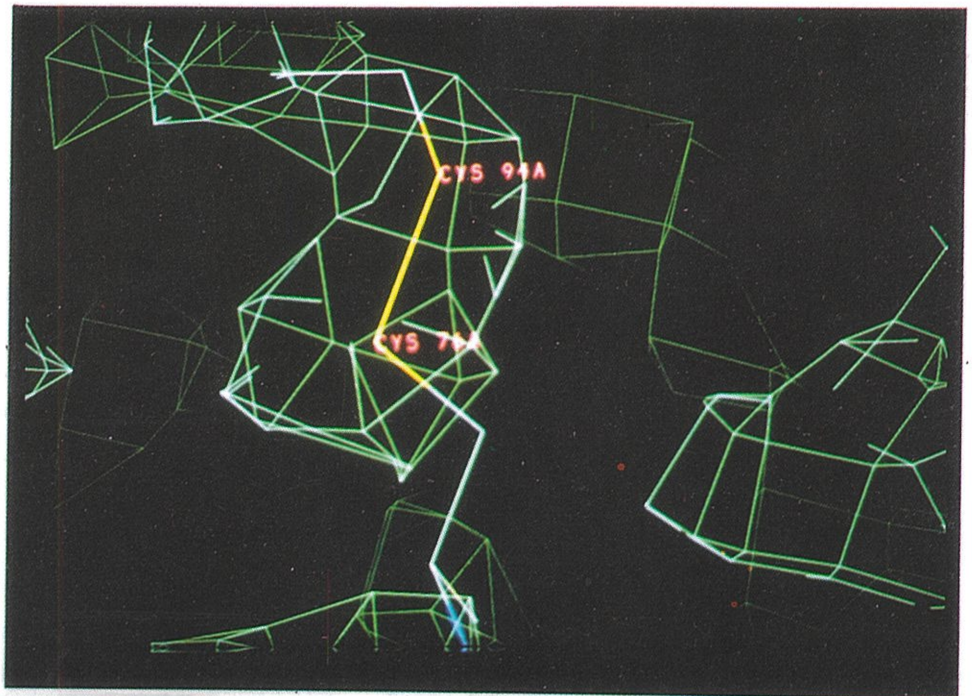


Figure 4. For caption, see p. 17.



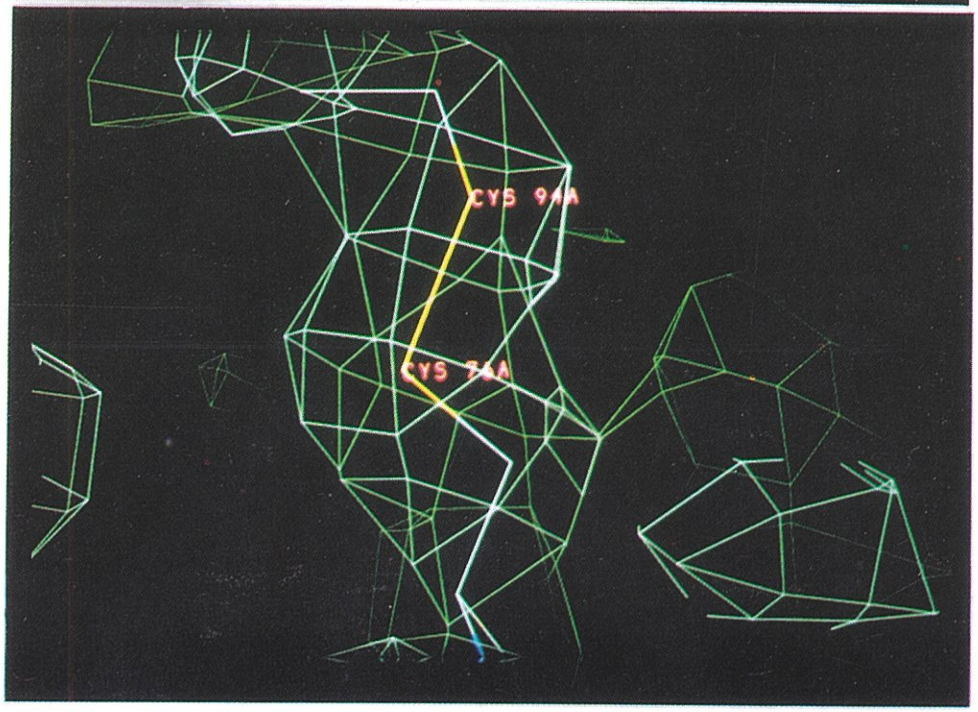


Figure 4. $2F_o - F_c$ electron density maps contoured from $1{\cdot}0\sigma$ for various amino acid residues: (Arg 5, His 15, Phe 38, Asp 52, Tyr 53, Pro 70, Trp 108 and disulphide bridges Cys 6–Cys 127, Cys 30–Cys 115, Cys 64–Cys 80 and Cys 76–Cys 94 (singles alone (top figures) and combined singles and deconvoluted multiples (bottom figures)).

electron density maps. Since this is a straightforward procedure, not necessitating any additional experimental data collection to realize improved completeness, it is a significant step forward for the synchrotron Laue method, where time may be an important constraint in data collection.

Acknowledgements

The SERC and the EC are thanked for research grant support and The Wellcome Trust for provision of funds towards an ESV30 molecular graphics system. The SERC, Daresbury Laboratory, is thanked for the provision of synchrotron radiation. The Swedish NFR is also thanked for financial support of the station 9·5 development, as well as SERC and MRC.

Dr W N Hunter and Dr G A Leonard are thanked for useful help and discussions.

References

Allinson N M, Carr P D, Colapietro M, Harding M M, Helliwell J R, Thompson A W and Weisgerber S 1992 *Phase Trans.* **39** 145

Amemiya Y 1990 *Synchrotron Radiation News* **3** 21

Amemiya Y and Miyahara J 1988 *Nature (London)* **336** 89

Blow D M 1993 *Curr. Biol.* **3** 204

Brammer R C, Helliwell J R, Lamb W, Liljas A, Moore P R, Thompson A W and Rathbone K 1988 *Nucl. Instrum. Methods* **A271** 678

Brünger A T 1990 *X-PLOR (version 2·2) Manual* (Yale Univ.: Howard Hughes Medical Institute)

Campbell J W and Hao Q 1993 *Acta Crystallogr.* **A49** 889

Campbell J W, Habash J, Helliwell J R and Moffat K 1986 *Information Quarterly for Protein Crystallography* Daresbury Laboratory Newsletter No. 18 p. 23

Cameron A D, Smerdon S, Wilkinson A, Habash J, Helliwell J R, Li T and Olson J 1993 *Biochemistry* **32** 13061

Cassetta A, Deacon A, Emmerich C, Habash J, Helliwell J R, McSweeney S, Snell E and Thompson A W 1993 *Proc. R. Soc. London Ser.* **A442** 177

CCP4 1979 The SERC (UK) collaborative computing project No. 4: a suite of programs for protein crystallography, distributed from Daresbury Laboratory, Warrington, UK

Clucas J A, Harding M M and Maginn S J 1988 *J. Chem. Soc. Chem. Commun.* 185

Cruickshank D W J, Helliwell J R and Moffat K 1987 *Acta Crystallogr.* **A43** 656

Cruickshank D W J, Helliwell J R and Johnson L N (eds) 1992 *Time-resolved macromolecular crystallography* (The Royal Society, London: Oxford University Press)

Duke E M H, Hadfield A, Walters S, Wakatsuki S, Bryan R K and Johnson L N 1992 in *Time-resolved macromolecular crystallography* (eds) D W J Cruickshank, J R Helliwell and L N Johnson (The Royal Society, London: Oxford University Press) p. 77

Emmerich C *et al* 1993 *Acta Crystallogr.* D (submitted)

Gomez de Anderez D, Helliwell M, Habash J, Dodson E J, Helliwell J R, Bailey P D and Gammon R E 1989 *Acta Crystallogr.* **B45** 482

Hao Q, Campbell J W, Harding M M and Helliwell J R 1993 *Acta Crystallogr.* **A49** 528

Harding M M, Maginn S J, Campbell J W, Clifton I and Machin P 1988 *Acta Crystallogr.* **B44** 142

Helliwell M, Gomez de Anderez D, Habash J, Helliwell J R and Vernon J 1989a *Acta Crystallogr.* **B45** 591

Helliwell J R *et al* 1989b *J. Appl. Crystallogr.* **22** 483

Jones T A 1978 *J. Appl. Crystallogr.* **11** 268

Lindahl M, Liljas A, Habash J, Harrop S and Helliwell J R 1992 *Acta Crystallogr.* **B48** 281

Pai E F 1992 *Curr. Biol.* **2** 821

Singer P T, Smalas A, Carty R P, Mangel W F and Sweet R M 1993 *Science* **259** 669

Sweet R M, Singer P T and Smalas A 1993 *Acta Crystallogr.* **D49** 305

Weisgerber S and Helliwell J R 1993 *J. Chem. Soc. Faraday Trans.* **89** 2667

Zurek S, Papiz M Z, Machin P A and Helliwell J R 1985 *Information Quarterly for Protein Crystallography* Daresbury Laboratory Newsletter No. 16 p. 37