

MAD Phasing Strategies Explored with a Brominated Oligonucleotide Crystal at 1.65 Å Resolution

M. R. Peterson,^a S. J. Harrop,^a S. M. McSweeney,^{b†} G. A. Leonard,^a
A. W. Thompson,^{b†} W. N. Hunter^a and J. R. Helliwell^{a*}

^aDepartment of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK, and ^bCCLRC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, UK

(Received 9 July 1995; accepted 25 September 1995)

The crystal structure of a brominated oligonucleotide d(CGCG^{Br}CG), chemical formula C₁₁₄N₄₈O₆₈P₁₀Br₂, has been analysed by multiwavelength anomalous dispersion (MAD) methods. The oligonucleotide crystallizes in space group *P*2₁2₁2₁ with *a* = 17.97, *b* = 30.98, *c* = 44.85 Å, $\alpha = \beta = \gamma = 90^\circ$. Data to a resolution of 1.65 Å were collected at four wavelengths about the *K*-absorption edge of the bromine atom ($\lambda_1 = 0.9323$ Å, a reference wavelength at the long-wavelength side of the edge; $\lambda_2 = 0.9192$ Å, at the absorption-edge inflection point; $\lambda_3 = 0.9185$ Å, at the 'white line' absorption maximum; $\lambda_4 = 0.8983$ Å, a reference wavelength at the short-wavelength side) using synchrotron radiation at Station PX9.5, SRS, Daresbury. Multiwavelength data could be collected on a single crystal as the sample was radiation stable. Anomalous and dispersive Patterson maps were readily interpretable to give the bromine anomalous scatterer positions. Phase calculations to 1.65 Å resolution, using all four wavelengths, gave a figure of merit of 0.825 for 2454 reflections. The electron-density map was readily interpretable showing excellent connectivity for the sugar/phosphate backbone and each base was easily characterized. The two nucleotide strands paired up as expected in an antiparallel Watson–Crick-type manner. The structure was refined to 1.65 Å using all the data (*R* factor = 17.0% based on 3151 reflections, with a data-to-parameter ratio of 2.6). In addition to the four-wavelength analysis, a variety of other phasing strategies, and the associated quality of the resulting electron-density maps, were compared. These included use of either of the reference wavelength data sets in the two possible three-wavelength phasing combinations to assess their relative effectiveness. Moreover, the time dependence upon measuring the Bijvoet differences and its effect upon phasing was also investigated. Finally, the use of only two wavelengths, including Friedel pairs, is demonstrated (the theoretical minimum case); this is of particular interest when considering overall beam time needs and is clearly a feasible experimental strategy, as shown here.

Keywords: MAD phasing; two-wavelength phasing; *f'*, *f''* values; nucleic acid, brominated; crystal alignment.

1. Introduction

In structural crystallography the determination of phases is a key hurdle to be surmounted, and for biological macromolecules the multiple isomorphous replacement (MIR) method (Green, Ingram & Perutz, 1954) has been the most commonly used approach in giving an *ab initio* evaluation of the required phase angles. An alternative approach for the *de novo* phasing of biological macromolecular crystal structures is the use of multiwavelength anomalous dispersion (MAD) phasing which exploits wavelength-dependent scattering effects [for a variety of strategies, see for example, Hendrickson (1985, 1991), Kahn *et al.* (1985), Helliwell (1979, 1984, 1992) and Hai-Fu, Woolfson & Jia-Xing (1993)]. Essentially, MAD experiments can be thought of as *in situ* isomorphous replacements, with anomalous

differences generated by the variation in scattering factor which accompanies change of wavelength for a specific atom or atoms in the structure concerned. In parallel with the evolution of strategy, instrumentation has developed at synchrotron radiation sources whereby the combination of point focusing of the beam and the rapid tuning of the wavelength for protein crystallography have been combined. Daresbury SRS station PX9.5 (Brammer *et al.*, 1988; Thompson *et al.*, 1992; Deacon *et al.*, 1995) incorporates these capabilities. The use of station PX9.5 for anomalous-scattering-based phasing in macromolecular crystallography is illustrated in this study of a brominated nucleotide. With two bromine atoms in the 240 light atoms, sufficiently large anomalous scattering effects can be induced, and should thus produce interpretable electron-density maps without recourse to further phase-improvement procedures, such as solvent flattening, histogram matching *etc.* In the case based

† Present address: EMBL, F-38043 Grenoble CEDEX, France.

on all four wavelengths there are redundant measurements made and MAD phasing (Karle, 1967, 1980, 1989; Hendrickson, 1985) is possible. It is of interest to take other combinations of fewer wavelengths, ranging from three and then to two different wavelengths (Helliwell, 1979). The use of two wavelengths is the theoretical minimum case (Okaya & Pepinsky, 1956; Hoppe & Jakubowski, 1975). In this paper each case was analysed in terms of phasing statistics and map quality. Since two different reference wavelengths were taken, one on the short-wavelength side and one on the long-wavelength side of the Br K edge, two three-wavelength combinations can be evaluated. The practical importance of the time dependence on the measurement of Bijvoet mates (Nieh & Helliwell, 1995) was also investigated by comparing a data set from the aligned crystal with a data set from the crystal misaligned, both recorded at the 'white line' absorption maximum wavelength (λ_3). Finally, the brominated nucleotide structure was solved and refined at 1.65 Å resolution.

Since this paper explores a large variety of wavelength combinations, four through three down to two, some comment on terminology is in order. Multiwavelength anomalous dispersion or scattering [MAD (Hendrickson, 1991) or MAS (Hai-Fu *et al.*, 1993)] is usually taken to mean use of three, four or even five wavelengths. An apparatus and experimental scheme for making simultaneous measurements of anomalous dispersion data over a full band of wavelengths or energies (*e.g.* 100 eV or more) in order to facilitate sampling of all wavelengths in the vicinity of an absorption edge has been described (Arndt *et al.*, 1982). At the other extreme is the use of two wavelengths for a unique determination of the phase of a reflection, provided at least one anomalous as well as the dispersive difference is utilized; two is certainly a multiple of one. The generic term MAD therefore will be used here.

2. Data collection

CGCG^{Br}CG crystallizes in space group $P2_12_12_1$ with $a = 17.97$, $b = 30.98$, $c = 44.85$ Å, $\alpha = \beta = \gamma = 90^\circ$. The single crystal selected for the data collection had a hexagonal plate morphology of dimensions $0.2 \times 0.1 \times 0.01$ mm. A XANES (X-ray absorption near-edge structure) experiment was carried out on this crystal at the SRS, Daresbury, on station PX9.5 (Brammer *et al.*, 1988; Thompson *et al.*, 1992; Deacon *et al.*, 1995). The $\delta\lambda/\lambda$ for the beam was set at 4.4×10^{-4} , whereby the vertical divergence was reduced by the use of slits upstream of the focusing mirror. The observed spectrum (f'' curve), shown in Fig. 1, underwent a Kronig–Kramers transformation (Kronig & Kramers, 1928) using the program *KRAMIG* (Hendrickson, Smith, Phizackerley & Merritt, 1988), to yield an f' curve. These curves were then used to choose the wavelengths for data collection.

The four wavelengths were chosen to try and stimulate the largest values of f'' and $\Delta f'$ overall; these wavelengths were (1) a reference on the long-wavelength side of the

edge ($\lambda_1 = 0.9323$ Å); (2) at the absorption-edge inflection point ($\lambda_2 = 0.9192$ Å); (3) at the 'white line' absorption maximum ($\lambda_3 = 0.9185$ Å); (4) a reference on the short-wavelength side for the edge ($\lambda_4 = 0.8983$ Å). The choices of λ_2 and λ_3 follow what are known as f' (dip) and f'' (max.), respectively [see *e.g.* the tabulated values of f' and f'' in Templeton, Templeton & Phizackerley (1980) and Templeton, Templeton, Phillips & Hodgson (1980)]. For each wavelength the crystallographic data were collected in a total of 15 blocks whereby each block comprised two exposures, each involving a 4° rotation. Hence, the sequence of wavelengths was λ_1 through to λ_4 repeated 15 times. For each 4° sweep the total exposure time was 60 s. In total 120° of data were collected for each of the four wavelengths. A fifth data set was then collected on the same crystal immediately afterwards at the 'white line' (*i.e.* λ_3) but the crystal was misaligned by offsetting one of the goniometer head arcs by approximately 30° . Again 4° images were employed, this time for a total of 80 s per image (due to beam decay). This strategy of misaligning the crystal was so as to increase the number of independent reflections on each diffraction image and thereby facilitate the calculation of scale factors for the images. It also allowed reflections previously in the blind region to be measured and combined with the λ_3 data set. Moreover, these two data sets allowed us then to test the impact of measuring the Friedel anomalous differences at widely different times on the phasing quality.

Overall the amount of beam time required was three shifts of 8 h each. Essentially, one shift was used to record the fluorescence spectrum from the crystal, having centred the crystal, and tested its diffraction. This was then followed by two shifts to record all of the five crystallographic data sets, each of 120° angular range at the four wavelengths involved.

3. Data processing

The diffraction data images for the first four data sets were processed using *DENZO* (Otwinowski, 1992) and reduced using programs from the *CCP4* suite (Collaborative Computing Project, Number 4, 1994). The fifth data set was processed using the *MOSFLM* suite of programs, including *ABSCALE* for Lp corrections (Leslie, Brick & Wonacott, 1986) and thereafter reduced in the same way as the first four data sets. The data images for each of the λ_1 to λ_4 sets in turn were scaled using the fifth misaligned crystal data set. For the λ_3 set on its own there were 3054 independent reflections to 1.65 Å, whereas for $^5\lambda_3$ (the fifth set) there were 2988 independent reflections to 1.70 Å. The output from the fifth data set's *AGROVATA* was combined with each specific wavelength reflection file using *MTZUTILS* and the reflection files sorted using *SORTMTZ*. The reflections in turn were scaled using the *CCP4* program *ROTAVATA*. At the *AGROVATA* stage the reference fifth wavelength set was removed. The quality of the data for

all data sets is illustrated in Table 1. (The use of $^5\lambda_3$ for scaling was, in fact, not critical, *i.e.* MAD electron-density map quality was identical without its use.) The weak and negative intensities were made consistent with a Wilson distribution of structure-factor amplitudes using *TRUNCATE*. The computer programs *CAD* and *SCALEIT* (*CCP4*) were employed to combine the four data sets into one file and to put them on an overall common scale with respect to λ_2 , which was subsequently treated as the 'native'. *SCALEIT* provides useful estimates of the largest acceptable dispersive and absorptive differences between and within the different wavelength data sets. Owing to the sensitivity of Patterson methods to spurious

large differences it was important to reject any unacceptably large differences as outliers. Typically, a total of less than five reflections were rejected for each data set using these tests. The final *SCALEIT* statistics are shown in Table 2. Both dispersive and absorptive Patterson maps were then generated with *FFT*. The bromine sites could be readily identified using both the anomalous (*e.g.* $\lambda_3 F^+ - \lambda_3 F^-$) and dispersive (*i.e.* $F_{\lambda_4} - F_{\lambda_2}$) Patterson maps. From the three Harker sections in both maps, two consistent bromine sites could be easily found. The quality of these Patterson maps can be seen in Figs. 2(a) and 2(b). The positions of the two bromine sites were 0.3241, 0.2009, 0.0100 (site A) and 0.5010, 0.1807, 0.2310 (site B).

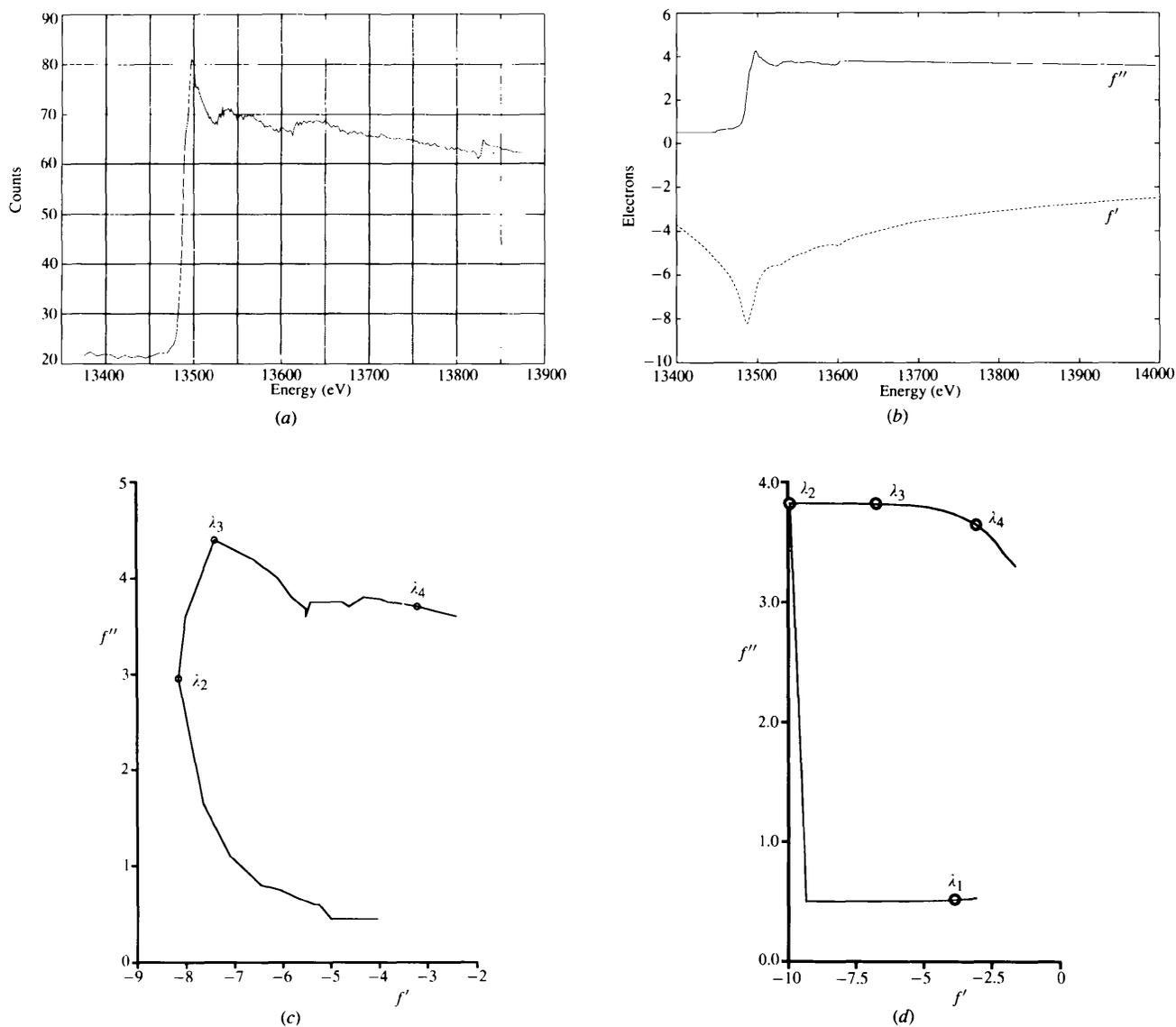


Figure 1

(a) X-ray fluorescence XANES spectrum from a single crystal of CGCG^{Br}CG recorded immediately before data collection. (b) Kronig-Kramers transformation of the fluorescence spectrum to give the f' (bottom) values from the spectral measurements, and f'' shown for comparison (top). (c) Plot of f'' versus f' obtained from the Kronig-Kramers transformation (b) illustrated by the solid line with the points supposedly accessed by the experiment indicated by circles [but see Table 2 and (d)]. (d) Plot of f'' versus f' from the Sasaki theoretical values illustrated by a solid line with the experimental positions (λ_1 to λ_4) marked by circles.

Table 1

Merging statistics for the five data sets from *AGROVATA* (Collaborative Computing Project, Number 4, 1994).

	λ_1	λ_2	λ_3	λ_4	$^5\lambda_3$
R_{merge} (%)	2.3	2.7	2.9	2.5	2.9
R_{anom} (%)	2.0	8.6	7.2	5.8	6.3
Total No. of reflections	11370	11598	11572	11653	11506
No. of independent reflections	3028	3052	3054	3056	2988
Completeness (%)	93.1	93.7	93.8	93.6	94.5
Multiplicity	3.8	3.8	3.8	3.8	3.9

Table 2

Scaling statistics between the native data set (λ_2) and the other data sets.

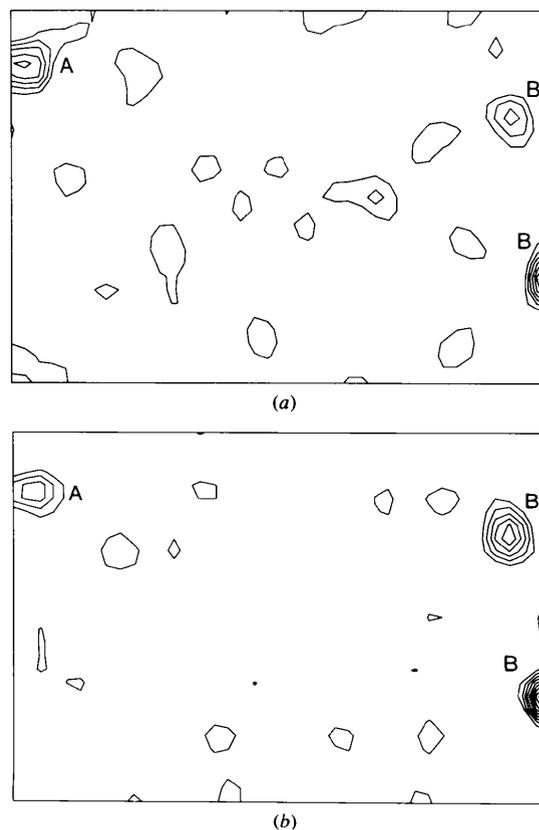
	MFID (%)	Mean anomalous difference	K_{emp}	K_{theor}
λ_1	5.7 (7.51)	2.7(1.28)	3.67	11.71
λ_2	0	11.5 (9.42)	0.0	0.0
λ_3	3.5 (3.93)	9.8 (9.42)	0.64	0.84
λ_4	6.9 (8.50)	7.8 (8.98)	1.48	1.89
$^5\lambda_3$	6.1 (3.93)	9.3 (9.42)	1.04	0.84

Notes: MFID is the mean fractional isomorphous difference (on F) observed for each wavelength with respect to the λ_2 data set. Calculated Δ_{iso} values are in parentheses, where $\Delta_{\text{iso}} = (N_A/N)^{1/2}(\Delta f'/f_{i1}) \times 100$. The mean anomalous difference on F , *i.e.* observed Friedel differences, is followed by calculated Δ_{anom} values in parentheses, where $\Delta_{\text{anom}} = 2(N_A/N)^{1/2}(\Delta f''/f_{i1}) \times 100$. K_{emp} is essentially MFID/0.5(mean anomalous difference). $K_{\text{theor}} = (\Delta f'\lambda_i - \lambda_2 f''\lambda_i)$, $i = 1$ to 4 ($\neq 2$).

4. Heavy-atom refinement, MAD phase calculation, electron-density map calculation and structure refinement

The refined positions of the two bromine atoms were used in *MLPHARE* on both hands (*i.e.* $x y z$ and $\bar{x} \bar{y} \bar{z}$). The figures of merit of the 'MAD' phases, using all four wavelengths (excluding the $^5\lambda_3$ data set), were 0.83/0.82 to 1.65 Å resolution for the acentric/centric data for both hands. *MLPHARE* treats the data sets collected at different wavelengths as isomorphous derivatives with one data set being chosen as the 'native'. The f' (dip) data set (λ_2) was used as the native to maintain a consistent positive dispersive difference between the other data sets. The native data set at the f' (dip) also has, theoretically, an f'' value of 3.823 electrons (Sasaki, 1989). The dispersive differences between λ_2 and the other data sets give rise to isomorphous differences, especially λ_1 and λ_4 , which are treated as apparent real occupancies of the anomalous scatterers. The initial estimates of the bromine positions were refined with iterative cycles of phase refinement. Each atom had its coordinates, temperature factors and occupancies (both real and anomalous) refined in *MLPHARE*. The real occupancies of the anomalous scatterers in the λ_2 native reference data set were fixed to zero throughout the iterative cycles. f' and f'' were added to the form-factor list, both being equal to one electron. Therefore, the real and anomalous occupancies corresponded to the number of electrons involved in the dispersive and absorptive differences, respectively, as the data sets were on a common scale from *SCALEIT*

[Table 3 gives the f' and f'' values at each wavelength extracted this way along with the theoretical values from Sasaki (1989)]. The MAD electron-density maps on both hands (Figs. 3*a* and 3*b*) were calculated at 1.65 Å resolution using the respective output phases from *MLPHARE*. The map calculated on the correct hand (Fig. 3*a*) shows the bases clearly and building of the model with *O* (Jones, Zou, Cowan & Kjeldgaard, 1991) could be easily started from the known heavy-atom positions. The connectivity of the phosphate backbone was of superb quality. The map calculated on the wrong hand was totally uninterpretable (Fig. 3*b*). The structure was refined to 1.65 Å resolution using the most complete data set, *i.e.* the combined set of the third and fifth, aligned and misaligned λ_3 data sets, respectively. After several cycles of slow cooling in *XPLOR* (Brünger, 1990) and restrained least-squares refinement in *NUCLSQ* (Hendrickson & Konnert, 1981), the final model obtained had 242 non-hydrogen atoms including 61 observed water molecules. This model gave an R factor of 17.0% for all the data (3151 reflections) with r.m.s. deviations of bond distances and bond angles of 0.010 Å and 0.015°, respectively. All conformational angles took up favourable values. Fig. 4 shows a portion of a $2F_o - F_c$ map in the final stages of refinement. The structure is

**Figure 2**

A Harker section ($U = 1/2$) of the Patterson maps (shown from 0 to 1/2 in V and W) calculated with (a) coefficients based upon anomalous differences recorded at λ_3 (note the use of the λ_2 coefficients enhanced the peaks); and (b) coefficients based upon dispersive differences between data sets λ_2 and λ_4 .

Table 3

Theoretical values of the anomalous scattering factors at the wavelengths (λ) used from Sasaki (1989) [corrected for an absorption-edge shift of 11 eV (to yield λ')] and experimentally derived from *MLPHARE*.

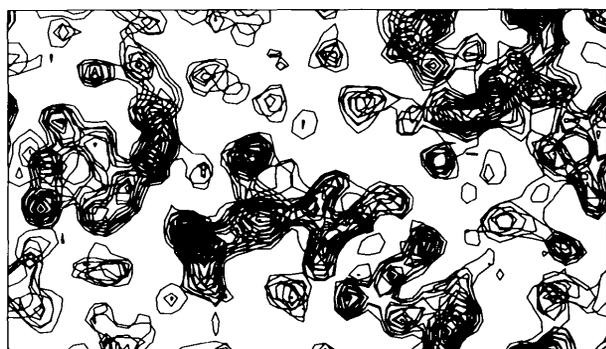
	λ	λ'	Theory (Sasaki, 1989)			Experimental (<i>MLPHARE</i>)			
			f'	f''	$\Delta f'$	$\Delta f'_A$	$\Delta f'_B$	$\Delta f''_A$	$\Delta f''_B$
λ_1	0.9323	0.9332	-3.83	0.515	6.12	5.64	5.50	0.45	0.45
λ_2	0.9192	0.9201	-9.95	3.823	0	0	0	5.68	3.55
λ_3	0.9185	0.9194	-6.76	3.817	3.19	3.05	2.31	3.83	3.17
λ_4	0.8983	0.8992	-3.08	3.644	6.15	6.58	6.37	2.99	2.72
$^5\lambda_3$	0.9185	0.9194	-6.76	3.817	3.19	1.96	2.70	2.55	2.40

MLPHARE has anomalous scattering factors for both bromine sites *A* and *B* as given by the subscript.

essentially identical to the non-brominated oligonucleotide structure of Wang *et al.* (1979); the r.m.s. deviation between the two structures is 0.438 Å excluding the Br atoms. The major difference between the two structures was the reorientation of the phosphate group on the fifth nucleotide. This sort of variation in nucleotide structures has been observed earlier (Egli, Williams, Gao & Rich, 1991). The largest displacements as compared with the Wang model were for P, O1P and O2P, being 1.678, 3.689 and 3.468 Å apart, respectively. No major reorientation was observed for the phosphate group on the corresponding nucleotide on the other strand. Excluding these particular structural differences the r.m.s. deviation between the two structures is 0.261 Å.

5. Phase information and electron-density map quality from various wavelength combinations

The description of the phase determination in the previous section describes the overdetermined case, utilizing all the anomalous and isomorphous effects in the measurement of each h , k , l (and Bijvoet mate) at all four wavelengths. In this section various other wavelength groupings are explored, some of which are of important historical significance. The following analysis can essentially be split into three categories involving data sets recorded at respectively four, three and two wavelengths in a variety of combinations to explore both theoretical strategies for phasing and experimental strategies of measurement (see Table 4 and Fig. 5).



(a)



(b)

Figure 3

A portion of the MAD (four-wavelength) electron-density map (shown from 0 to 1 in x and y for a 2.8 Å thick slab in z) calculated at 1.65 Å resolution, contoured at 0.5 r.m.s. intervals and commencing at 0.5 r.m.s. calculated (a) on the correct hand and (b) on the wrong hand.

5.1. λ_1 , λ_2 , λ_3 , λ_4

This combination of wavelengths is the four-wavelength overdetermined case where the f'' anomalous effects of each wavelength are all utilized along with the isomorphous effects between λ_2 and each of the other three wavelengths. The overall figure of merit was 0.8251 with 2454 reflections phased. The map was of excellent quality and structural moieties could be easily characterized as described in §4.

5.2. λ_1 , λ_2 (anomalous off), λ_3 , λ_4

The maximum anomalous signal, which occurred at λ_2 [consistent with the theoretical values of Sasaki (1989)] was turned off in this case to explore the possible benefit of simultaneous measurement of Bijvoet mates from a well aligned crystal (Nieh & Helliwell, 1995) on the λ_3 'white line' data sets (*i.e.* it should be compared with case 5.3 below). The overall figure of merit was 0.8045 with 2604 reflections phased. The map again was of high quality and again could be easily characterized.

5.3. λ_1 , λ_2 (anomalous off), $^5\lambda_3$, λ_4

This case allows a direct comparison with case 5.2 and illustrates the effects of time dependence of the misaligned f'' 'white line' set, $^5\lambda_3$ (Nieh & Helliwell, 1995). The overall figure of merit was 0.7797 with 2646 reflections phased, slightly lower than the previous example. However, the map quality utilizing the four wavelengths was not noticeably affected.

Table 4

Figures of merit for the various wavelength scenarios in §5 for acentric (*a*), centric (*c*) and overall cases (number of reflections in brackets).

Case	Description	Mean FOM (<i>a</i>)	Mean FOM (<i>c</i>)	Overall FOM
5.1	$\lambda_1, \lambda_2, \lambda_3, \lambda_4$	0.8262 (1818)	0.8221 (636)	0.8251 (2454)
5.2	λ_1, λ_2 (an off), λ_3, λ_4	0.7982 (1968)	0.8239 (636)	0.8045 (2604)
5.3	λ_1, λ_2 (an off), $^5\lambda_3, \lambda_4$	0.7726 (2010)	0.8019 (636)	0.7797 (2646)
5.4	$\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (an all off)	0.5418 (2399)	0.8297 (636)	0.6021 (3035)
5.5	$\lambda_1, \lambda_2, \lambda_3$	0.7332 (1804)	0.7196 (636)	0.7297 (2440)
5.6	$\lambda_2, \lambda_3, \lambda_4$	0.7858 (1733)	0.7667 (636)	0.7807 (2369)
5.7	λ_2, λ_4	0.7193 (1730)	0.6973 (636)	0.7134 (2366)
5.8	λ_2, λ_1	0.5877 (1818)	0.6273 (636)	0.5978 (2494)
5.9	λ_3, λ_4	0.6313 (1757)	0.5720 (635)	0.6156 (2392)
5.10	$^5\lambda_3, \lambda_4$	0.4217 (2304)	0.1957 (636)	0.3728 (2940)
5.11	λ_3, λ_4 (an off)	0.5153 (1940)	0.5650 (635)	0.5275 (2575)
5.12	$^5\lambda_3, \lambda_4$ (an off)	0.3413 (2314)	0.1846 (636)	0.3075 (2950)
5.13	λ_3, λ_1	0.3751 (1900)	0.3861 (635)	0.3779 (2535)

From the figure of merit, FOM, a mean phase error value can be derived as $\cos^{-1}(\text{FOM})$.

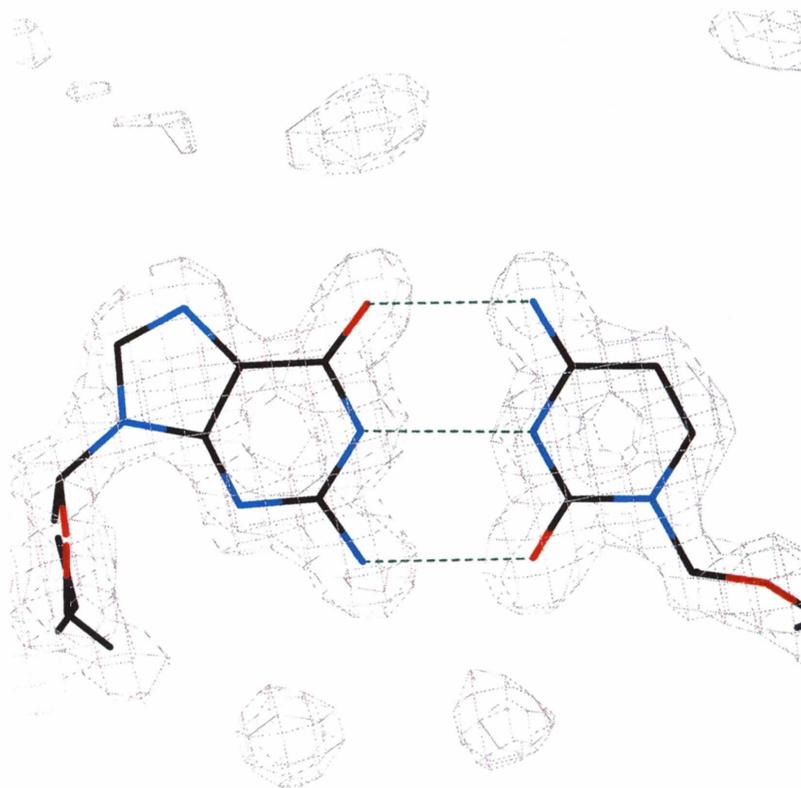
5.4. $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (anomalous off in all cases)

In all four data sets here the Friedel mate is switched off in each case. Hence, all the desired phase information has to be derived from structural amplitude changes for an *hkl* only (*i.e.* without $\bar{h}\bar{k}\bar{l}$ information) between λ_2 and the other respective wavelength data sets (Herzenberg & Lau, 1967). This combination could be described as a multiple isomorphous replacement (MIR) experiment with various degrees of bromine substitution along the heavy-atom vector. The phasing of the centric reflections was not affected as the figure of merit was still high (0.8297), but a

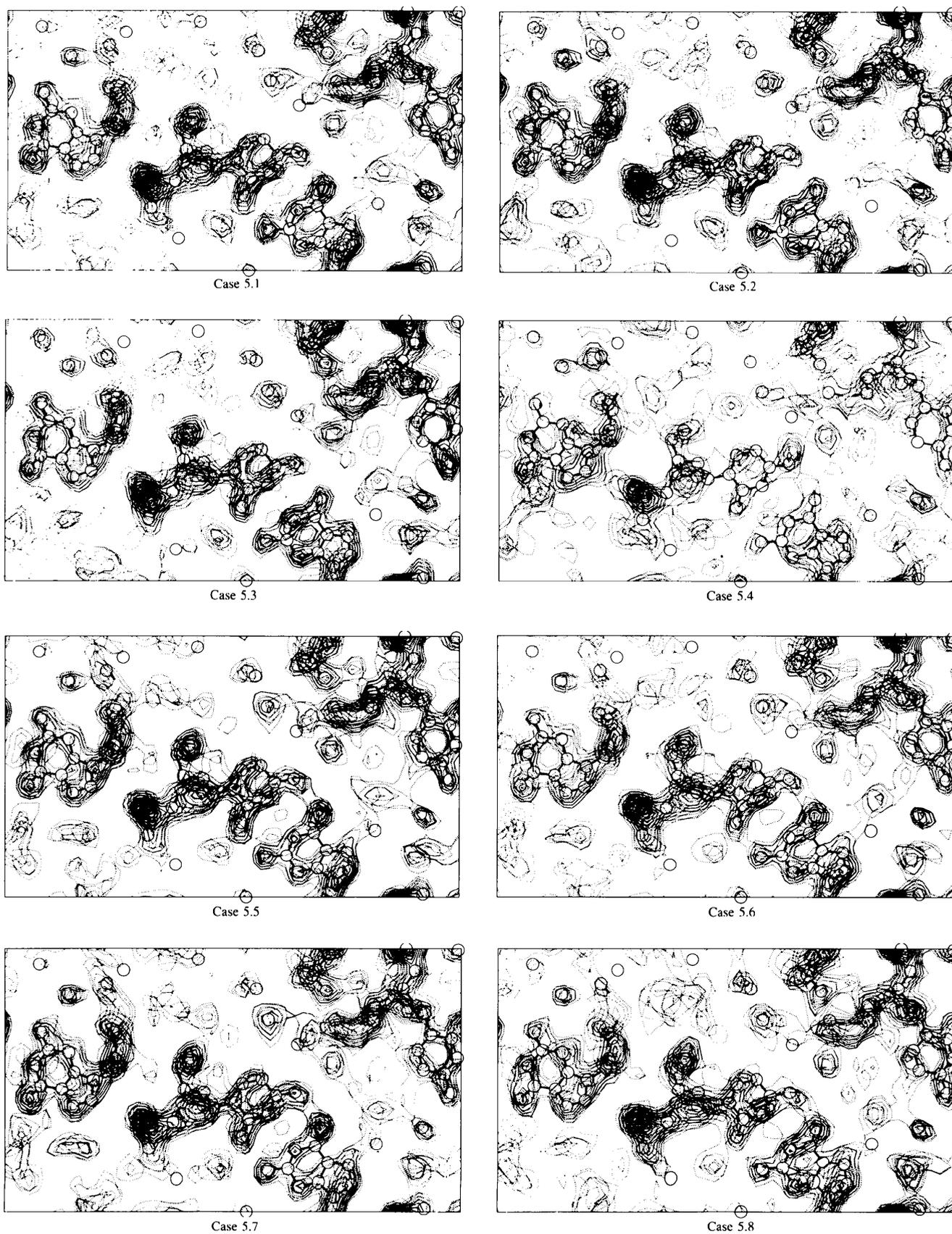
big fall in the figure of merit was observed for the acentric reflections to 0.5418 from 0.8262. That phasing was still achieved indicates that the centres of the phasing circles are still showing some degree of being separated and non-collinear. However, the quality of the map had severely deteriorated.

5.5. $\lambda_1, \lambda_2, \lambda_3$

This three-wavelength case, and the next, is to compare the two possible choices of reference wavelength. Sometimes, due to lack of synchrotron beam time and/or

**Figure 4**

The final $2F_o - F_c$ map contoured at 1.0 r.m.s. calculated at 1.65 Å resolution ($R = 17.0\%$).

**Figure 5**

Electron-density maps contoured at 0.5 r.m.s. intervals and commencing at 0.5 r.m.s., corresponding to the cases described in the text as 5.1 to 5.8 (as labelled), shown from 0 to 1 in x and y for a 2.8 Å thick slab in z . The final refined molecular model, including waters, is superimposed to assist comparison of the maps.

prolonged exposure times, it may be only possible to collect data at three wavelengths. The reference wavelength, λ_1 , has no anomalous signal (mean anomalous difference = 2.7%, see Table 2) as it is situated on the long-wavelength side of the Br *K* edge. The overall figure of merit was 0.7297 for 2440 reflections phased. The map, however, was of excellent quality and could be easily characterized.

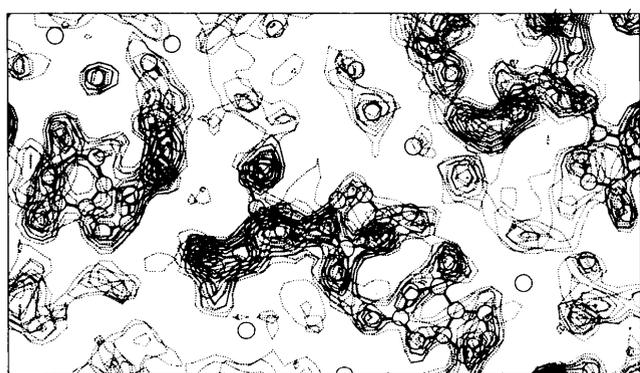
5.6. $\lambda_2, \lambda_3, \lambda_4$

The reference wavelength, λ_4 , has a good anomalous signal (mean anomalous difference = 7.8%, see Table 2) as it is situated on the short-wavelength side of the absorption

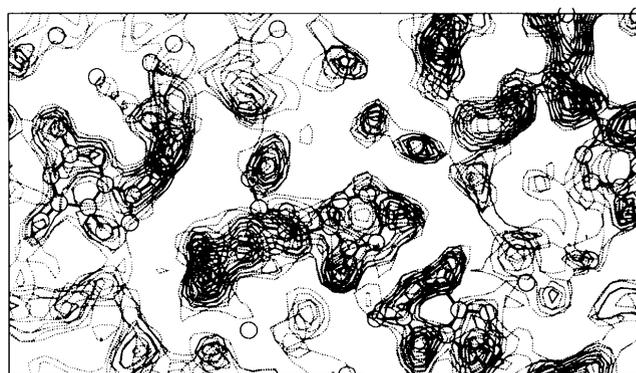
edge, unlike λ_1 . The overall figure of merit was certainly improved compared with case 5.5 above to 0.7807 for 2369 reflections phased. The map was of excellent quality and could be easily characterized, but was not noticeably superior to case 5.5.

5.7. λ_2, λ_4

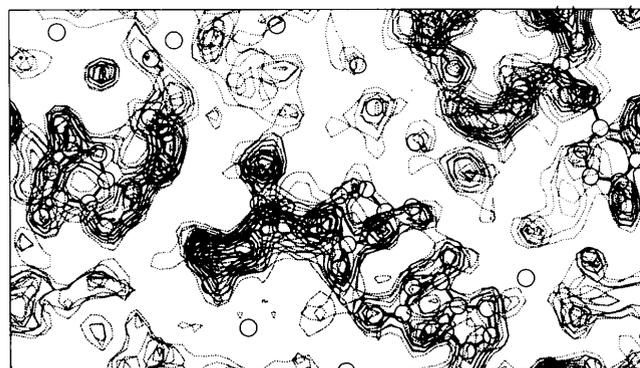
The theoretical minimum case involves two wavelengths. This, and cases 5.9 and 5.10, are akin to the 'two-short-wavelength method' of Hoppe & Jakubowski (1975). It is required that the centres of the phasing circles be well separated and non-collinear (Helliwell, 1984) and this is



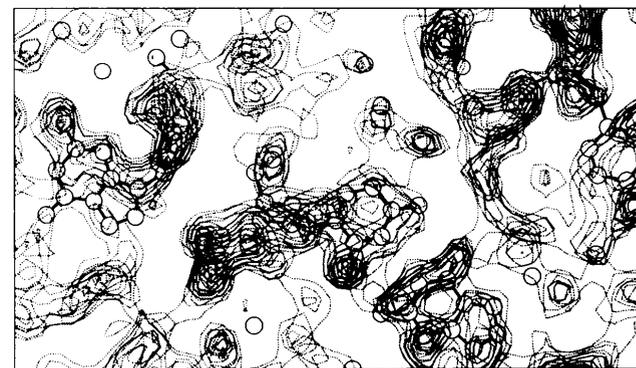
Case 5.9



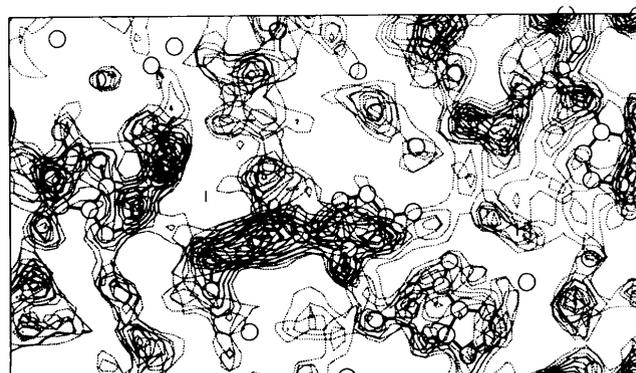
Case 5.10



Case 5.11



Case 5.12



Case 5.13

Figure 5 (cont.)

Electron-density maps contoured at 0.5 r.m.s. intervals and commencing at 0.5 r.m.s., corresponding to the cases described in the text as 5.9 to 5.13 (as labelled), shown from 0 to 1 in *x* and *y* for a 2.8 Å thick slab in *z*. The final refined molecular model, including waters, is superimposed to assist comparison of the maps.

achieved well here. It is important to utilize the maximum isomorphous and maximum anomalous differences possible. The λ_2, λ_4 pairing has the largest isomorphous difference (MFID = 6.9%, see Table 2), whilst λ_2 also has the maximum Friedel difference [as expected from the theory of the free atom (Sasaki, 1989; Hoppe & Jakubowski, 1975) (mean anomalous difference = 11.5%, see Table 2) but not as expected from the Kronig–Kramers transform (Fig. 1c)]. λ_4 , being the reference on the short-wavelength side of the edge, also has a substantial anomalous difference (mean anomalous difference = 7.8%, see Table 2). The overall figure of merit achieved was 0.7134 for 2366 reflections, and the electron-density map was of high quality and totally interpretable.*

5.8. λ_2, λ_1

Again a strong isomorphous signal was induced in this pairing (MFID = 5.7%, see Table 2), but the λ_1 reference has little anomalous signal (mean anomalous difference = 2.7%, see Table 2) as it is on the long-wavelength side of the Br *K* edge. This case (Helliwell, 1979) then (Fig. 6) is analogous to the case of SIROAS – single isomorphous replacement with optimized anomalous scattering (Baker *et al.*, 1990). It is in fact the ‘conventional two-wavelength method’ of Hoppe & Jakubowski (1975) and Okaya & Pepinsky (1956), as is case 5.13, but here only very small changes of wavelength are employed, because of the ease of tuning with synchrotron radiation, and which yield optimal

* Recently, we have seen that Friedman, Fischmann & Steitz (1995) used two wavelengths, in essence λ_2 and λ_4 , for Hg *L*_{III} (but with Bijvoet mates only at λ_4) for the structure determination of the *lac* repressor core protein (38 kDa).

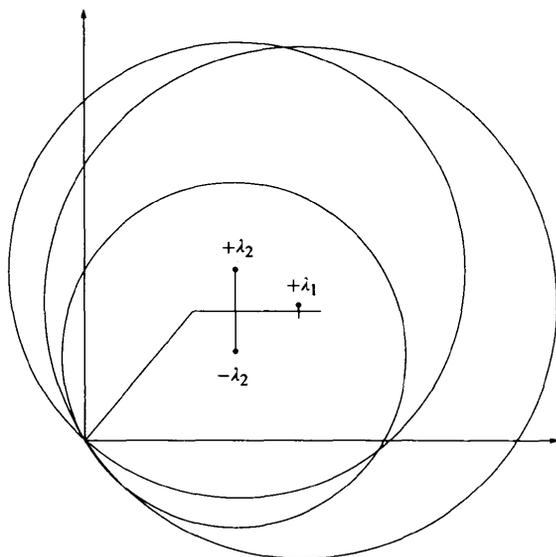


Figure 6

The centres for the phasing circles in this Harker diagram are based on the Bijvoet mates for λ_2 , where, from theory (consistent with Table 2) the f' (dip) is coincident with the f'' (max.), and one reference set where there is very little anomalous signal λ_1 (*i.e.* two wavelengths in total; case 5.8, see text). The centres are well separated and non-collinear and the unique phase is well resolved.

probing of the f', f'' values *etc.* The overall figure of merit was 0.5978 for 2494 reflections. The map was of poorer quality than for the λ_2, λ_4 case; however, it was still interpretable.

5.9. λ_3, λ_4

The biggest difference between this case and case 5.7 is that there is a smaller isomorphous difference between the λ_3 and λ_4 sets (MFID = 5.0%). Phasing is nearly entirely dependent on the anomalous signals within λ_3 and λ_4 (mean anomalous differences of 9.8 and 7.8%, respectively, see Table 2). The overall figure of merit was 0.6156 for 2392 reflections and the map was of similar quality to case 5.8.

5.10. $^5\lambda_3, \lambda_4$

This pairing (and the next examples) does illustrate the increasing importance of careful crystal alignment in measuring Bijvoet mates (Nieh & Helliwell, 1995) as the number of wavelengths used for phasing become less, unlike case 5.3. The overall figure of merit dropped to 0.3728 for 2940 reflections. The map quality had deteriorated and structure determination became more difficult in contrast to case 5.9.

5.11. λ_3, λ_4 (anomalous off)

The anomalous signal is switched off for λ_4 , so the total phasing power comes largely from the anomalous signal of λ_3 (mean anomalous difference = 9.8%, see Table 2) alone with the small isomorphous difference between λ_3 and λ_4 . The effect of the accuracy of the Bijvoet mate measurement from this well set crystal can be compared with case 5.12 (mis-set crystal). The overall figure of merit was 0.5275 for 2575 reflections. The map was of poorer quality.

5.12. $^5\lambda_3, \lambda_4$ (anomalous off)

$^5\lambda_3$ provides the anomalous signal here (mean anomalous difference = 9.3%, see Table 2), like λ_3 in case 5.11. The overall figure of merit for this case was 0.3075 for 2950 reflections. The map quality observed here was of extremely poor quality (worse than case 5.11) and in which basically no model could be built.

5.13. λ_3, λ_1

This wavelength pairing illustrates the importance of choice of reference wavelength in the phasing of this map. A small isomorphous difference is involved (MFID = 3.3%), along with little anomalous signal for the λ_1 reference set (mean anomalous difference = 2.7%, see Table 2). The overall figure of merit was 0.3779 for 2535 reflections, and the calculated map again was basically not interpretable, in contrast to case 5.9 which had a superior figure of merit.

6. Discussion and concluding remarks

Station PX9.5 at the SRS, Daresbury, enabled multiwavelength anomalous dispersion measurements to be made

quickly and efficiently. The brominated oligonucleotide d(CGCG^{Br}CG) was chosen as a test crystal for several reasons: it was radiation insensitive; it had a very good concentration of anomalous scatterers, *i.e.* two bromines in 240 light atoms; and the bromine edge was very near to the critical wavelength flux output of the SRS wiggler. It also diffracted strongly, due to the relatively small unit cell, in spite of the rather small crystal volume. The experimental strategies illustrated within this overview range from the theoretical minimum case through to the MAD overdetermined method (Okaya & Pepinsky, 1956; Herzenberg & Lau, 1967; Hoppe & Jakubowski, 1975; Karle, 1967, 1980, 1989; Hendrickson, 1985, 1991; Helliwell, 1979, 1984, 1992). Consideration of the experimental design becomes very important when the overall beam time is to be considered (Helliwell, 1979). Several specific comments and guidelines on experimental strategy can be given from this work.

The choice of wavelengths was made by reference to the fluorescence spectrum. However, although λ_3 was expected to have the largest Friedel anomalous difference, in fact that was the case for the λ_2 (f' dip) data set. The crystallographic data are actually entirely consistent with the tabulated theoretical variations of f' and f'' with wavelength [Sasaki (1989, p. 57); see also Hoppe & Jakubowski (1975, p. 445, Fig. 4) and Fig. 1d]. There is then a discrepancy of those with the Fig. 1(c) values, derived from the Kronig–Kramers transform. Note that the position of the absorption edge from the tabulated absorption coefficients (Sasaki, 1990) and of f' and f'' (Sasaki, 1989) agree. It might be possible that dichroism effects could explain this discrepancy. However, the values of f'' and $\Delta f'$ derived for each Br site (Table 3), which, if so, might show differing behaviour in terms of the wavelength that has the largest f'' value, actually agree *i.e.* λ_2 rather than λ_3 is the one with the largest f'' for each site. Dichroism then is unlikely to be the explanation. Hence, it appears that the Kronig–Kramers transform computer program used to derive the f' curve from the measured fluorescence spectrum has wrongly shifted the f' curve. This is important because in our X-ray diffraction data λ_2 alone yields the largest f'' value and the most negative f' value, as expected from theory, if not the Kronig–Kramers transform curve. Hence, the use of just two wavelengths, λ_2 with a reference wavelength, λ_4 , whilst being the theoretical minimum, also yielded the largest f'' and $\Delta f'$ differences in the diffraction data.

The case of crystallographic data collected at just two wavelengths, *i.e.* λ_2 and λ_4 alone, is then the optimized version of the Hoppe & Jakubowski (1975) 'two-short-wavelength method', in that $(\lambda_2 - \lambda_4)$ yields the largest $\Delta f'$ whilst λ_2 , as well as being the f' (dip) is simultaneously the f'' (max.) data set; f'' at λ_4 is also sizeable and very useful (whereas for the reference wavelength, λ_1 , it is not). Crystal alignment (Nieh & Helliwell, 1995) can be critical as shown in cases 5.11 and 5.12. The theoretical minimum (case 5.7) of use of just these two key wavelengths, λ_2 and λ_4 , is important then in situations where exposure times

per oscillation image need to be long to reduce random errors (*i.e.* to obtain good R_{merge} values), especially when crystal size may be small, unit cell large, beam flux low, or beam time is restricted. Future work will expand on these strategies, and comparisons, to a macromolecular crystal with a larger unit cell.

We are grateful to the staff at the Synchrotron Radiation Source at Daresbury for the provision of beam. SRS Station PX9.5 has been funded in large part by the Swedish Research Council (NFR) as well as the UK SERC (now EPSRC) and the UK MRC to whom we are extremely grateful. SJH has been funded under a research grant from SERC (now EPSRC) to JRH and WNH. The University of Manchester has provided a studentship to MRP. The Wellcome Trust and the University of Manchester have provided funding for an ESV molecular graphics system. The EPSRC and the University of Manchester have recently provided funds for a cluster of SG workstations.

References

- Arndt, U. W., Greenough, T. J., Helliwell, J. R., Howard, J. A. K., Rule, S. A. & Thompson, A. W. (1982). *Nature (London)*, **298**, 835–838.
- Baker, P. J., Farrants, G. W., Stillman, T. J., Britton, K. L., Helliwell, J. R. & Rice, D. W. (1990). *Acta Cryst.* **A46**, 721–725.
- Brammer, R. C., Helliwell, J. R., Lamb, W., Liljas, A., Moore, P. R., Thompson, A. W. & Rathbone, K. (1988). *Nucl. Instrum. Methods*, **A271**, 678–687.
- Brünger, A.-T. (1990). *X-PLOR Manual*. Version 2.1. Yale University Press.
- Collaborative Computing Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Deacon, A., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Leonard, G. A., Peterson, M. R., Hadener, A., Kalb (Gilboa), A. J., Allinson, N. M., Castelli, C., Moon, K., McSweeney, S. M., Gonzalez, A., Thompson, A. W., Ealick, S., Szebenyi, D. M. & Walter, R. (1995). *Rev. Sci. Instrum.* **66**(2), 1287–1292.
- Egli, M., Williams, L. D., Gao, Q. & Rich, A. (1991). *Biochemistry*, **13**, 4143–4158.
- Friedman, A. M., Fischmann, T. O. & Steitz, T. A. (1995). *Science*, **268**, 1721–1727.
- Green, D. W., Ingram, V. M. & Perutz, M. F. (1954). *Proc. R. Soc. London*, **225**, 287–302.
- Hai-Fu, F., Woolfson, M. M. & Jia-Xing, Y. (1993). *Proc. R. Soc. London Ser. A*, **442**, 13–32.
- Helliwell, J. R. (1979). *Daresbury Study Weekend*, DL/SCI/R13, pp. 1–6. Daresbury Laboratory, Warrington, UK.
- Helliwell, J. R. (1984). *Rep. Prog. Phys.* **47**, 1403–1497.
- Helliwell, J. R. (1992). *Macromolecular Crystallography with Synchrotron Radiation*. Cambridge University Press.
- Hendrickson, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.
- Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
- Hendrickson, W. A. & Konnert, J. H. (1981). *Biomolecular Structure, Conformation, Structure and Evolution*, Vol. 1, edited by R. Srinivasan, pp. 43–57. Oxford: Pergamon Press.
- Hendrickson, W. A., Smith, J. L., Phizackerley, R. P. & Merritt, E. A. (1988). *Proteins Struct. Funct. Genet.* **4**, 77–78.
- Herzenberg, A. & Lau, H. S. M. (1967). *Acta Cryst.* **22**, 24–28.

- Hoppe, W. & Jakubowski, U. (1975). *Anomalous Scattering*, edited by S. Ramaseshan & S. C. Abrahams, pp. 437–461. Copenhagen: Munksgaard.
- Jones, T. A., Zou, J. Y., Cowan, S. W & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Dideberg, O. & Wery, J. P. (1985). *FEBS Lett.* **179**, 133–137.
- Karle, J. (1967). *Appl. Opt.* **6**, 2132–2135.
- Karle, J. (1980). *Int. J. Quantum Chem. Quantum Biol. Symp.* **1**, 357–367.
- Karle, J. (1989). *Phys. Today*, **42**(6), 22–29.
- Kronig, R. de L. & Kramers, H. A. (1928). *Z. Phys.* **48**, 174–179.
- Leslie, A. G. W., Brick, P. & Wonacott, A. J. (1986). *CCP4 Newsl.* **18**, 33–39.
- Nieh, Y.-P. & Helliwell, J. R. (1995). *J. Synchrotron Rad.* **2**, 79–82.
- Okaya, Y. & Pepinsky, R. (1956). *Phys. Rev.* **103**, 1645–1647.
- Otwinowski, Z. (1992). *DENZO* and *SCALEPACK*. Personal communication.
- Sasaki, S. (1989). KEK Report 88-14. National Laboratory for High Energy Physics, Tsukuba 305, Japan.
- Sasaki, S. (1990). KEK Report 88-16. National Laboratory for High Energy Physics, Tsukuba 305, Japan.
- Templeton, D. H., Templeton, L. K., Phillips, J. C. & Hodgson, K. O. (1980). *Acta Cryst.* **A36**, 436–442.
- Templeton, L. K., Templeton, D. H. & Phizackerley, R. P. (1980). *J. Am. Chem. Soc.* **102**, 1185–1186.
- Thompson, A. W., Habash, J., Harrop, S., Helliwell, J. R., Nave, C., Atkinson, P., Hasnain, S. S., Moore, P. R., Harris, N., Kinder, S. & Buffey, S. (1992). *Rev. Sci. Instrum.* **63**, 1062–1064.
- Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawfoot, J. L., van Boom, J. H., van der Marel, G. & Rich A. (1979). *Nature (London)*, **282**, 680–686.