Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

addenda and errata

Applications of single-wavelength anomalous dispersion at high and atomic resolution. Corrigendum

D. E. Brodersen,[a]* E. de La Fortelle,[b] C. Vonrhein,[b] G. Bricogne,[b,c] J. Nyborg[a] and M. Kjeldgaard[a]

[a]Department of Molecular Biology, University of Aarhus, Gustav Wieds Vej 10c, DK-8000 Aarhus C, Denmark, [b]MRC-LMB, Hills Road, Cambridge CB2 2QH, England, and [c]LURE, Bâtiment 209D, 91405 Orsay CEDEX, France.
Correspondence e-mail: deb@mb.au.dk

In the paper by Brodersen *et al.* [(2000), *Acta Cryst.* D**56**, 431–441] errors were published in values in the last set of fractional coordinates and in the molecular mass. The correct values are given in this article.

In the article by Brodersen *et al.* (2000) there is an error in the values for the last set of fractional coordinates. The last sentence of the second paragraph on p. 434 should read 'In a direct-methods run applying 256 phase permutations, the positions of the two Ho atoms were determined to be (0.0676, 0.1540, 0.1048) and (0.0200, 0.3433, 0.0856) in fractional coordinates of the unit cell'. The error occured because of a misreading of the log file from *SHELXS*.

Also, on p. 432 the molecular mass is listed incorrectly as 22.1 kDa. This should be given as 22.7 kDa.

References

Brodersen, D. E., de La Fortelle, E., Vonrhein, C., Bricogne, G., Nyborg, J. & Kjeldgaard, M. (2000). *Acta Cryst.* D**56**, 431–441.

  doi:10.1107/S0907444906024073

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Applications of single-wavelength anomalous dispersion at high and atomic resolution

D. E. Brodersen,[a] E. de La Fortelle,[b] C. Vonrhein,[b] G. Bricogne,[b,c] J. Nyborg[a] and M. Kjeldgaard[a]*

[a]Macromolecular Crystallography, IMSB, Aarhus University, Gustav Wieds Vej 10c, DK-8000 Aarhus C, Denmark, [b]MRC-LMB, Hills Road, Cambridge CB2 2QH, England, and [c]LURE, Bâtiment 209D, 91405 Orsay CEDEX, France

resent address: INSEAD, Boulevard de Constance, F-77305 Fontainebleau CEDEX, France.

Correspondence e-mail: mok@imsb.au.dk

Received 27 September 1999
Accepted 17 anuary 2000

Two e amples of the application of single-wavelength anomalous dispersion (AD) in macromolecular structure determination are described, both using the statistical phasing program *SHA* . For the holmium-substituted calcium-binding protein psoriasin (22.7 kDa), a set of accurate phases has been obtained to a resolution of 1.05 A without recourse to an atomic model of the molecule. The accuracy of the phases resulted in an electron-density map of a uality comparable to σ_A-weighted 2 − $_c$ maps derived from the final model refined with *SHELX* 7. Comparison of the refined and AD electron-density maps showed significant discrepancies resulting from the iterative refinement in reciprocal space. Additionally, it is shown that the structure of psoriasin can be determined from native data e tending to 2.0 A alone by e ploiting the minute anomalous signal from a bound inc ion.

1. ntroduction

In order to construct an atomic model of a macromolecule *e* by means of crystallography, it is necessary to turn the measured intensities of -rays diffracted from a crystal into an electron-density map of sufficient uality for interpretation. This involves estimation of the phase angles, which are unmeasurable in crystallography, and tracing of an appro - imate atomic model in the electron-density map that results from Fourier transformation. The model is subse uently improved in the refinement process by taking into account stereochemical information, as well as by minimi ing the discrepancy between measured structure factors and structure factors calculated from the model.

A techni ue for overcoming the problem of estimating the phases and obtaining an initial electron-density map which was pioneered by erut and co-workers, the multiple isomorphous replacement (MI) method, yields good inde-pendent phase information (Green *et al.*, 1 54). However, the uality of the phases tends to deteriorate rapidly with reso-lution owing to lack of isomorphism between native and heavy-atom substituted crystals. Conse uently, the initial electron-density map may contain gross errors that prevent its proper interpretation in terms of a model. eciprocal-space refinement of an incomplete structure model has been shown to involve the risk of model bias that can be difficult to escape later in the refinement (Hodel *et al.*, 1 2). Molecular replacement is another powerful method for determining macromolecular structures that uses phase information derived from a similar protein structure (ossmann & Blow, 1 62). However, here too the partial correctness of the initial

model has important conse uences for the progress of the subse uent refinement.

A special physical effect, namely the absorption of -rays by atoms at certain energies, also known as 'anomalous dispersion', can be used to yield independent phase information (Laue, 1 16 Coster *et al.*, 1 33 Honl, 1 33 Bijvoet, 1 4). The 'anomalous' signal, which is manifested in differences between Friedel-related re ections (the so-called Bijvoet differences), is usually small and re uires accurate data collection and processing. However, the power and energy tunability of modern synchrotron beamlines have had profound impact on its use in recent years, most notably in the conte t of the multiple-wavelength anomalous dispersion (MAD) method (Hendrickson, 1 1).

The traditional use of initial electron-density maps as a first poor sketch of the structure in which an appro imate and perhaps incomplete atomic model can be traced has recently been challenged on two fronts. n the e perimental side, the data collected at modern synchrotron-radiation beamlines have become much more accurate and complete (Helliwell, 1 2 Dauter *et al.*, 1 7) thanks to phosphor imaging-plate systems and CCD detectors, and the average time of data ac uisition has decreased. Moreover, the advent of cryogenic technology (Haas, 1 68 for a review, see odgers, 1 7) has enabled crystals to diffract significantly longer in a strong -ray beam with minimal radiation damage. Therefore, one can now collect complete and highly redundant data sets from a single crystal, which again results in precise measurement of both diffraction intensities and their standard deviations. Likewise, improvements have also been achieved with respect to the software used to process the diffraction data (twinowski, 1 3, 1 7 Leslie, 1 3 ugrath, 1 7) and new ma imum-likelihood-based heavy-atom refinement and phasing algorithms have made it possible to more efficiently take advantage of small signals (La Fortelle & Bricogne, 1 7).

The outcome is that under suitable conditions the initial electron-density map can now carry at least as much information as the final atomic model after refinement or the electron-density map derived from it. Burling and coworkers have shown that an e perimental electron-density map calculated from a multiple-wavelength e periment at 1.8 A resolution can provide useful information about multiple conformers of protein side chains and the organi ation of water molecules at the protein–solvent interface (Burling *et al.*, 1 6 ead, 1 7). However, several e periments indicate that in some cases the anomalous signal available at a *s le* wavelength can suffice for calculation of phases and thus structure determination (Hendrickson *et al.*, 1 85). By analogy to the multiple-wavelength counterpart MAD (Hendrickson, 1 1), phasing based on the Bijvoet differences only has been named single-wavelength anomalous dispersion or AD. In general, phase estimation in such cases results in an ambiguity similar to the single isomorphous derivative (I) case. In the original work by Hendrickson & Teeter (1 81) for the determination of the structure of crambin, the phase ambiguity was resolved by combining the bimodal phase distribution with a

able 1
Data-collection statistics.

Bound metal(s)	Holmium	Calcium, inc
pace group	$2_12_12_1$	4_32_12
esolution range (A)	37.5–1.05	1 . –2.05
avelength (A)	0. 0 1	0.862
No. of observations	700858 (160426 uni ue)	14331 (107 1 uni ue)
nit-cell parameters (A)		
a	46.753	51. 6
	61.602	51. 6
c	62.750	116.03
Temperature (K)	100	100
Completeness ( )	7. (84.0)	8. (100.0)
edundancy	1.8 (1.2)	7.0 (6.)
$_{sym}$ ( )	3.3 (10.8)	11.3 (42.3)
σ	26.0 (4.7)	15. (5.0)

Friedel mates separate. Values in parentheses represent the outermost resolution shell (1.07–1.05 A). Values in parentheses represent the outermost resolution shell (2.12–2.05 A).

im weight calculated from the known positions of the anomalous scatterers. This approach has been used in the determination of several other crystal structures, including trimeric haemerythrin (mith *et al.*, 1 83) and gramicidin A (allace, 1 86 allace *et al.*, 1 0). Combination of direct methods and single-wavelength anomalous scattering, known as iterative single-wavelength anomalous scattering (I A ang, 1 85) has been used to break the phase ambiguity for a number of structures including metallothionin (Furey *et al.*, 1 86), bovine neurophysin II in comple with a dipeptide (Chen *et al.*, 1 1), the N-terminal fragment of IF3 (Biou *et al.*, 1 5) and rusticyanin (Harvey *et al.*, 1 8).

hasing protocols using single-wavelength data all re uire a strong anomalous signal from heavy atoms such as iodine (neurophysin II), lead (IF3N), caesium (gramicidin A) or cadmium (metallothionin) and generally lead to relatively good electron-density maps. In cases where only a weak signal is available, however, other techni ues must be employed before tracing the chain, for e ample, molecular replacement (oyer *et al.*, 1 8), or data must be collected very close to the absorption edge of the anomalous scatterer (Harvey *et al.*, 1 8). However, Dauter and coworkers have recently demonstrated that by mimicking Cu α radiation at a synchrotron source, the anomalous signal produced by a large number of and Cl atoms in a lyso yme crystal can be used to solve the phase problem for data collected away from the absorption edge (Dauter *et al.*, 1).

The work presented here shows two e amples of the phase information that can be e tracted from single-wavelength diffraction e periments in the case of a 22.1 kDa dimeric calcium-binding protein. hen calcium is substituted by a strong anomalous scatterer, holmium, we show that diffraction data collected far from the absorption edge but to true atomic resolution can be used to calculate an initial electron-density map of outstanding uality. Comparison with the refined $2 - _c$ map has revealed notable differences resulting from the iterative refinement of the structure in reciprocal space. This observation has implications for our understanding of model-bias phenomena at this level of resolution. In

able 2
Analysis of the anomalous signal and phasing statistics.

esolution bin (A)	37.5–2. 4	2. 4–2.08	2.08–1.70	1.70–1.47	1.42–1.32	1.32–1.20	1.20–1.11	1.11–1.05	37.5–1.05
Number of re ections	4032	7172	176	10 3	11804	14068	14 35	12461	84587
⟨ ⟩	374.5	217.1	128.6	81.6	57.	43.	34.3	26.1	86.0
⟨σ( )⟩	4. 5	3.16	1.77	1.54	1.46	1.63	1. 6	2.37	2.07
⟨|Δ |⟩	15.42	12.68	8.8	6.88	5.47	4.58	3.87	3.2	6.36
⟨σ(Δ)⟩	6.11	4.53	3.01	2.82	2.73	3.07	3.64	4.27	3.53
⟨|Δ |⟩ ⟨σ(Δ)⟩	2.5	2.8	3.0	2.4	2.0	1.5	1.1	0.8	1.8
.m.s.(Δ) r.m.s.( )	4.8	7.1	7.	.5	10.8	11.8	12.8	14.0	10.7
C (phasing) ( )	5 .6	73.5	82.7	86.1	86.0	84.1	81.8	80.0	81.4
C (dens.mod.) ( )	84.0	2.	5.2	5.5	5.0	3.6	1.	88.0	2.6
Δφ (phasing)¶	43.0	32.2	27.1	25.8	26.0	28.8	32.4	34.3	30.24
Δφ (dens.mod.)	32.0	18.7	15.1	13.0	13.8	16.5	1 .5	26.4	18.44

Δ is the Bijvoet difference, which is defined as Δ $_{l}$ − $\overline{\ }_{l}$. Calculated as $[(1\ )\sum_{l}\Delta\ _{l}{}^{2}]^{1\,2}$ $[(1\ )\sum_{l}\ _{l}{}^{2}]^{1\,2}$. C is the average correlation between the set of phases calculated from the final refined model (φ_c) and the phases output directly from *SHA* (phasing) and after density modification (dens.mod.). ¶ Δφ is the average non-weighted difference between phases calculated from the final refined model (φ_c) and the phases output directly from *SHA* (phasing) and after density modification (dens.mod.).

another e periment with weaker diffracting crystals of the native protein, we demonstrate that for data collected far from the absorption edge and without fine energy tuning, the faint anomalous signal produced by a single inc ion can be used to calculate a good electron-density map to a resolution of 2.0 A. This shows that the structure of the protein and most likely other metal-binding proteins can be determined *a t* from a single native data set at this resolution without the use of a known model, heavy atoms or direct methods.

The AD-phased electron-density maps are especially useful because they provide information that is independent of any assumption based on an atomic model of the protein. The maps rely on no parameters other than those introduced to carry out heavy-atom refinement. In addition, the solvent fraction in the crystal and the assumption that there is a boundary between protein atoms and solvent atoms were used in the solvent- attening step. The independence of the electron-density maps from model hypotheses used in refining the protein model makes it possible to check the validity of the model and thus help start refinement closer to the optimal structure and reduce the risk of model bias.

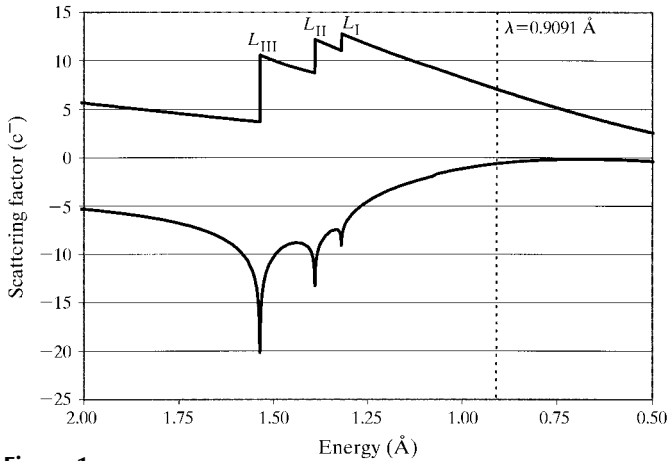

Figure 1
The three *L* absorption edges of the lanthanide holmium are located at 1.5362 (L_{III}), 1.3 03 (L_{II}) and 1.31 8 A (L_{I}). The wavelength used for the AD e periment with holmium-substituted psoriasin, 0. 0 1 A, is indicated with a vertical dashed line. At this point, the theoretical values of f' and f'' are −0.628 and 7.102 e⁻, respectively.

2. Methods

2.1. igh-resolution AD e periment

2.1.1. Data collection. Monochromatic diffraction data were collected to a resolution of 1.05 A at the 11 beamline, EMBL, Hamburg, Germany, for crystals of a holmium-substituted ($^{67}Ho^{3}$) preparation of the 22.7 kDa dimeric calcium-binding 100 protein psoriasin (100A7) as previously reported (Brodersen *et al.*, 1 8). In order to determine properly the intensities of the strongest re ections at this resolution, the data were collected in three consecutive runs (50–2.8 A crystal-to-detector distance, 440 mm Δφ 2° 15–1.7 A crystal-to-detector distance, 250 mm Δφ 2° 5.3–1.05 A crystal-to-detector distance, 120 mm Δφ 0.5°). The crystals belong to the space group $2_12_12_1$, with unit-cell parameters *a* 46.81, 61.63, *c* 62.77 A, and diffract e tremely well (1.0 A). The data were reduced and processed using the *E SCALE AC* suite of programs (twinowski & Minor, 1 7 twinowski, 1 3) and have an overall σ of 28.8 and an $_{sym}$ of 4.3 (see Table 1 for data-collection statistics). tructure-factor amplitudes were calculated

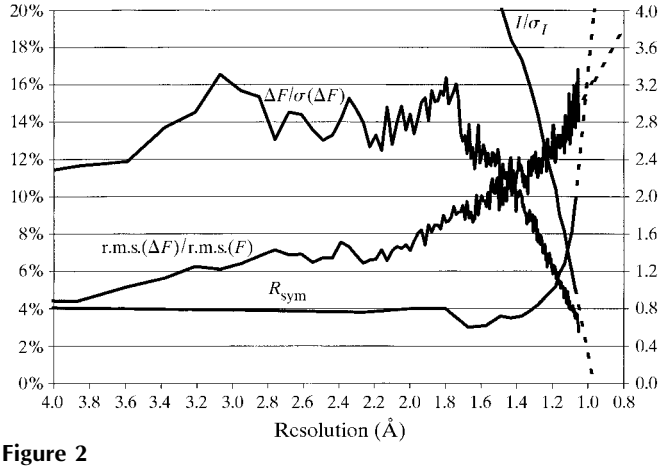

Figure 2
The breakdown with resolution of several crystallographic data- uality indicators. The $_{sym}$ and r.m.s.(Δ) r.m.s.( ) are measured in percentages on the left *y* a is, whereas σ and |Δ | σ(Δ) are absolute numbers (right *y* a is). The r.m.s.(Δ) r.m.s.( ) shows the average relative contribution of the anomalous signal in resolution bins and Δ σ(Δ) the signal-to-noise level of the anomalous differences.

able
efined heavy-atom parameters.

Ho1	
	0. 331
y	0.3457
	0.1048
iso	
ccupancy	1.72
Ho2	
	0. 803
y	0.1565
	0.0860
iso	
ccupancy	0. 6
cattering factors (e^-)	
′	−5.0
″	6.7

arameter was fi ed during refinement in *SHA* .

using the *CC* 4 program *CA E* (Collaborative Computational roject, Number 4, 1 4 French & ilson, 1 78).

2.1.2. eavy-atom detection and refinement. No special wavelength-stabili ation apparatus was used because the data were collected at a wavelength of 0. 0 1 A (13.6 keV), far from any of the holmium *L* absorption edges (L_1 1.31 8, L_{II} 1.3 03, L_{III} 1.5362 A). The data were thus collected at the high-energy side of the L_I edge, where both ′ and ″ are well tabulated and insensitive to small wavelength uctuations. The fact that ″ remains large in this range (∼7.1 e^-) gives rise to the significant anomalous differences shown in Tables 2 and 3.

From an anomalous difference atterson, it was known that the monomers of this dimeric protein only contained a single holmium ion each, a total of two ions per asymmetric unit. This information, along with the Δ s for uni ue re ections (37.5–1.05 A) where both Friedel mates had been recorded, was input into the program *SHELXS* (heldrick, 1 3, 1 7). In a direct-methods run applying 256 phase permutations, the positions of the two Ho atoms were determined to be (0.0676, 0.1540, 0.1048) and (0.200, 0.3433, 0.1548) in fractional coordinates of the unit cell.

2.1. . hasing. The positions and identities of the two heavy atoms along with the observed anomalous differences and structure factors were input to *SHA* (La Fortelle & Bricogne, 1 7) in a heavy-atom parameter refinement and phasing run (see Table 3 for a summary of the refined heavy-atom parameters). The *SHA* phasing resulted in an overall figure of merit (F M) of 82.6 . To further improve the uality of the initial map, the data were subjected to 70 cycles of solvent ipping using the program *S L* (Abrahams & Leslie, 1 6 Collaborative Computational roject, Number 4, 1 4) in a script implemented into the *SHA* interface in which the radius of the 'solvent sphere' is gradually reduced from an initial value of 5.0 A to the ma imum resolution of the data (1.05 A in this case). This crystal form is tightly packed and contains appro imately 40 solvent. Electron-density maps were calculated with and *EX E* (Collaborative Computational roject, Number 4, 1 4) and displayed in the program (Jones & Kjeldgaard, 1 7). σ_A-weighted 2 − $_c$ maps representing the final refined map from the *SHELX* refinement were calculated using the *SHELX* 7 utility program *SHELX* .

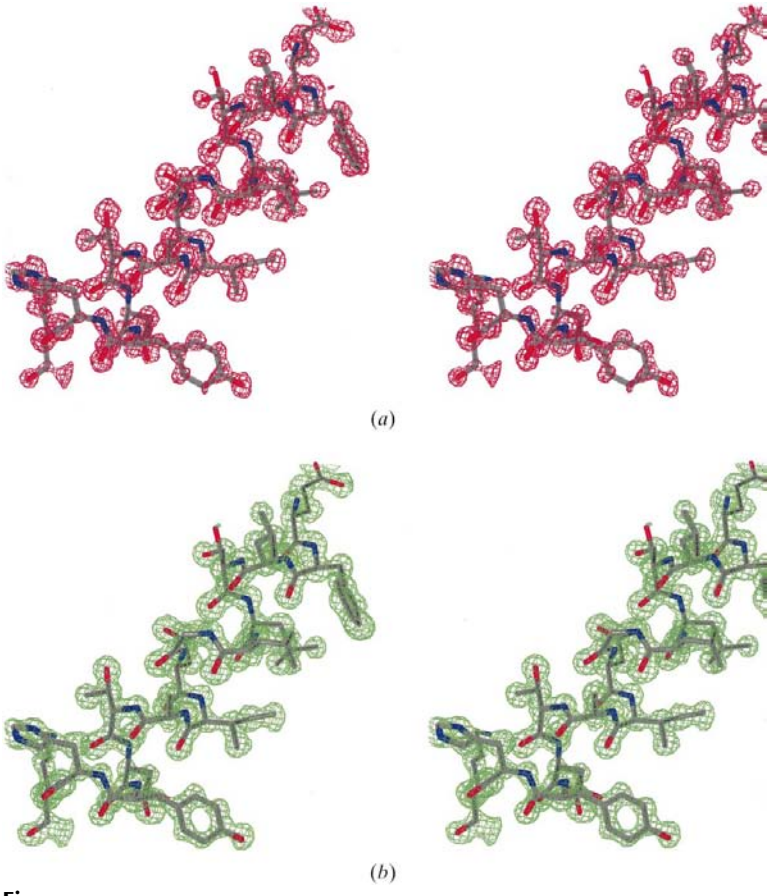

Figure
Two stereographs of the electron density resulting from the AD e periment with holmium–psoriasin. (*a*) Map calculated with phases direcly from *SHA* , *.e.* without any kind of density modification. Contouring is 1. r.m.s. (1.11 e A^{-3}). ( ) The final AD electron density after solvent attening in *S L* . Contouring is 1.6 r.m.s. (0. 3 e A^{-3}).

2.2. Direct phasing using native data

2.2.1. Data collection. Monochromatic -ray diffraction data were collected on a tetragonal crystal of the native protein to a resolution of 2.0 A using an -ray wavelength of 0.863 A (Table 1). Crystals of the native protein that appear in the presence of inc belong to the space group 4_32_12, with unit-cell parameters *a* 51. 55, *c* 116.033 A (Nols e *et al.*, 1 7 Brodersen *et al.*, 1). This data set is 8. complete for all data between 1 . and 2.05 A. The data set was processed with the *E SCALE AC* suite, taking into account the small but measurable anomalous signal

from the inc ion. tructure factors and anomalous differences were calculated using *CA E*. The orientation of the protein dimer within the unit cell was determined by molecular replacement using the previously determined structure of holmium-substituted psoriasin (Brodersen *et al.*, 1 8) and the program *A e* (Nava a, 1 0, 1 4). This showed that the protein crystalli es with a single monomer (100 residues) in the asymmetric unit, containing one bound calcium and one bound inc ion (Brodersen *et al.*, 1).

2.2.2. eavy-atom detection. The anomalous differences calculated with *CA E* were sorted based on the relative contribution of the observed structure factor. nly re ections for which $|\Delta| > 0.25$ were kept. hen calculating the anomalous atterson summation, the re ections were additionally subjected to an absolute value criterion in that only re ections which satisfied $|\Delta| > 50.0$ were used in . This resulted in 1577 re ections out of 11 213 (14) being rejected.

2.2. . hasing. The single position of the inc ion was used along with the structure-factor magnitudes and anomalous differences to calculate optimal comple structure factors and phase distributions using *SHA* (La Fortelle & Bricogne, 1 7). In the density-modification step, 130 cycles of solvent ipping were carried out using *S L* and the method of reducing the radius of the solvent sphere described above. The solvent radius was again reduced from an initial value of 5.0 A down to the ma imum resolution of the data, 2.0 A. This crystal form contains appro imately 60 solvent. Correlation coefficients were calculated with *A A* (Kleywegt & Jones, 1 6).

. Results

.1. igh-resolution AD e periment

ingle-wavelength anomalous diffraction (AD) data were collected to a resolution of 1.05 A at a synchrotron source for a crystal of the holmium-substituted calcium-binding 100-protein psoriasin (100A7) at a wavelength of 0. 0 1 A (13.640 keV Brodersen *et al.*, 1 8). For a data-collection summary, see Table 1. At this energy, the anomalous signal of the two Ho atoms bound to the dimeric protein is significant even though the wavelength is far from any of the *L* absorption edges of this lanthanide element (Fig. 1). Because of the very intense diffraction from these derivative crystals, the data set was collected in three consecutive runs corresponding to low, medium and high resolution (Table 1).

In cases where the temperature factors of the anomalous diffractors are much lower than the average temperature factor of the protein atoms, the e pected anomalous signal represented by differences between Friedel re ections will increase with resolution on a relative scale. To estimate the measurable signal, the anomalous differences must then be compared with the measurement noise, which also increases with resolution. Fig. 2 shows the resolution dependence of the relative anomalous signal r.m.s.(Δ) r.m.s.( ) and the significance of the anomalous signal $|\Delta| / \sigma(\Delta)$, along with the crystallographic symmetry factor and σ for the measured intensities. Even for the highest resolution bin, the increased measurement noise does not bury the anomalous signal. In addition, the anomalous signal as measured by $|\Delta| / \sigma|\Delta|$ remains higher than 1.0 for all but the last bin (Table 2). In other words, the anomalous signal is significant to at least 1.1 A resolution. $|\Delta| / \sigma|\Delta|$ for the whole resolution range is 1.8. In this case, the anomalous signal represents almost 15 of the intensities of the highest order re ections. Therefore, a systematic treatment of the Bijvoet differences at high resolution is of great value in the initial phase estimation. However, we believe that the overall data uality, rather than a high contribution, makes the anomalous signal a reliable source of information.

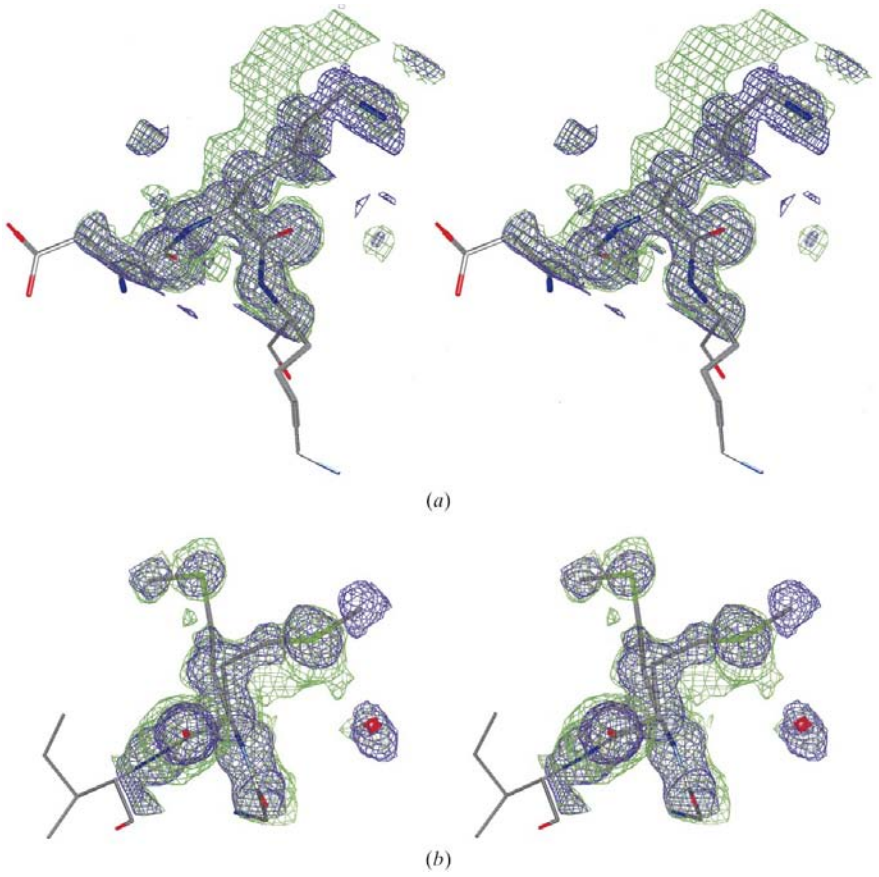

Figure
Two stereographs comparing the AD and refined electron density. (*a*) esidue LysA48 with refined 2 − $_c$ density from *SHELX* at 1.5 r.m.s. (0. 26 e A^{-3}, blue) and AD density at 1.0 r.m.s. (0.617 e A^{-3}, green). ( ) esidue MetA12 and a water molecule shown with the same maps as above.

.2. eavy-atom detection and phasing

ince holmium had been introduced to replace the calcium ions which are naturally bound in the protein, detection of the holmium positions in the atterson map was relatively straightforward owing to the e istence of only two clearly defined and high-occupancy heavy-atom sites in the structure. Both heavy-atom positions could be found either by direct methods using the program *SHELXS* (heldrick, 1 3 heldrick & chneider, 1 7), by manual interpretation of the anomalous atterson function or from atterson superposition methods (heldrick, 1 7).

nce the positions of the two Ho atoms in the asymmetric unit were located, parameters were refined and phasing was carried out using the statistical phasing program *SHA* (La Fortelle & Bricogne, 1 7). The parameters representing the occupancies and the atomic scattering factors are strongly correlated in a AD refinement. Therefore, the occupancy of the first heavy atom was fi ed at 1.0 and the $'$ value for holmium fi ed at an arbitrary value of $-5.0\ e^-$, while $''$ was refined along with the positional parameters, temperature factors and the occupancy for the other Ho atom. Table 3 shows the refined heavy-atom parameters. The imaginary part of the atomic scattering factor for holmium, $''$, refines to $6.7\ e^-$, very close to the corresponding theoretical value of $7.1\ e^-$.

The electron density could be further improved by use of solvent ipping attening as implemented in the program *S L* (Abrahams & Leslie, 1 6 Collaborative Computational roject, Number 4, 1 4). The procedure used, which gradually produces a finer mask, results in the determination of a final molecular envelope that closely fits the electron density of the molecule, including structural waters. ell defined spherical solvent molecules are thus strengthened rather than attened. The map calculated from the solvent- attened phases will be referred to as the AD map in the following sections.

The overall correlation between phases calculated from the final model refined in *SHELX* 7 and the phases output directly from *SHA* is 81.4 . Fig. 3(*a*) shows the electron density for one part of an α-heli of the protein, but the general uality of the map is similar even for regions outside stable secondary structure. In this map, which has not been subjected to any density modification, very fine details are visible such as atomic detail of the main chain and rigid aliphatic side chains, and static disorder (notice the multiple conformations of the serine residue in the lower left part of the figure). Table 2 shows the correlation between the phases output from *SHA* and phases calculated from the final refined model in resolution bins. Evidently, the correlation between the phase sets is highest between 1.7 and 1.3 A and lowest below 2. A, probably owing to poor modelling of bulk solvent in the refined structure.

After the density-modification step, the correlation between the two sets of phases is almost perfect (2.6). The highest resolution shell has a somewhat lower correlation, perhaps indicating that the molecular mask still needs refining. Fig. 3( ) illustrates the final electron-density map calculated from the measured structure-factor magnitudes and the phases output from *S L* . The map shows clear atomicity and resembles to a large e tent the final refined $2 - _c$ map from the refinement in *SHELX*.

. . Model bias

The model of the protein was originally traced in a map calculated from MAD phases e tending to 1. A, with the 1.05 A data set being collected on a different crystal (Brodersen *et al.*, 1 8). The high-resolution data set was then used for refinement of the model with the program *SHELX* 7 (heldrick, 1 3 heldrick & chneider, 1 7). The iterative least-s uares refinement resulted in an overall crystallographic factor of 10.4 for all data, with a cross-validation $_{free}$ value of 14.12 for a random 3 test set. The final refined model is used for all comparisons in this paper.

wing to the atomic resolution data, the observation-to-parameter ratio, even with anisotropic factors, approaches that seen in small-molecule -ray studies 160 437 observations (the Friedel mates being held separate) and 18 753 parameters, giving an overall ratio of 8.6. The traditional refinement included all the features e pected at this level of resolution, such as riding H atoms, full anisotropic -matri (AD) refinement and local static disorder. e believed initially that with the combination of a high observation-to-parameter ratio, e cellent data uality, a well established high-resolution refinement protocol and very good refinement statistics, the model could not be improved upon at the end of refinement. However, as we will show in the following, the refinement protocol allowed some bias to remain in the final high-resolution model.

The phases resulting from the solvent- attening procedure described in the previous section are very accurate to the highest resolution. The AD map thus offers the uni ue possibility of studying important structural features of a macromolecule at atomic resolution. Although the AD map (Fig. 3) fits the refined model almost perfectly, it also provides additional information in its fine detail. By comparing the AD map with the final σ_A-weighted $2 - _c$ map from the *SHELX* refinement, we have been able to locate several occurrences of model bias either resulting from the iterative comparison of calculated and measured structure factors or from phase errors present when the initial MAD map was calculated. This shows that interpretation of the electron density as an atomic model *al ays* affects the result, regardless of the degree of overdetermination of the set of parameters to be estimated (*.e.* positional coordinates and temperature factors).

. . mproved modelling of multiple conformers

tatic disorder is a common phenomenon and from high-resolution studies it is estimated that between 6 and 15 of all residues in a protein are likely to e ist in more than one single conformation (mith *et al.*, 1 6). In this case, 1 out of a total of 200 residues in the asymmetric unit (~10) were initially

modelled with double conformations. A high degree of disorder is often a cause of trouble in cases of average to poor data, relatively weak initial phases or low resolution, because the occupancies of the multiple conformers are less than unity and the density thus appears considerably weaker than for an average rigid side chain. Moreover, multiple conformations are often modelled towards the end of the refinement process, at which point the density will be strongly biased by phases calculated from the model. Global uality indicators such as $_{\rm free}$ are generally not sensitive enough to catch errors caused by such low-occupancy conformations. Independent phase information is thus indispensable when it comes to determining the number of discrete variations as well as their individual significance, since phases calculated from a model tend to distort the picture and favour the modelled conformation.

Fig. 4(*a*) shows an e ample of a lysine side chain (LysA48) in the structure of psoriasin which appears as a single discrete conformer in the refined σ_A-weighted 2 $-$ $_c$ map from *SHELX* (blue) even though the terminal atoms C$^{\gamma}$ C$^{\varepsilon}$ N$^{\zeta}$ show some signs of e ibility as judged from the density. This is also re ected in the values of the e uivalent isotropic factors for this residue, which appears near the end of an α-heli . hereas the main chain and the first part of the side chain (until C$^{\gamma}$) is relatively rigid ($_{\rm iso}$ 15 A^2), e ibility increases drastically near the terminus of the side chain ($_{\rm iso}$ of 32 A^2 for C$^{\varepsilon}$ and 42 A^2 for N$^{\xi}$). hen investigating the AD map (green density), it is evident that this apparent thermal mobility is caused by a distinct double conformation of most of this side chain. Not even a σ_A-weighted $-$ $_c$ difference density map (not shown) reveals the true identity of this second side-chain orientation. Modelling of this side chain as a double conformation and refining the structure again significantly reduced the isotropic factors for both conformations ($_{\rm iso}$ of 26 and 21 A^2 for the two C$^{\varepsilon}$ atoms and 37 and 22 A^2 for the two N$^{\xi}$ atoms).

Fig. 4( ) shows the final 2 $-$ $_c$ electron density around a methionine residue (MetA12, blue) along with the same region of the AD map (green). This methionine residue occurs in at least two different orientations which are clearly identified from the two positions of the dense atom. However, when comparing the model with the AD density (green), it is obvious that the C$^{\varepsilon}$ position for one of the conformations is incorrect. Modelling of the C$^{\varepsilon}$ position as pointing upwards earlier in the refinement has resulted in artefactual density appearing in this position (blue density). The peak in the density at the real position of the C$^{\varepsilon}$ (towards the right) was mistakenly modelled by a water molecule, which has then supposedly been pushed away by anti-bumping restraints. For neither of the cases mentioned here did the σ_A-weighted difference density (Fourier coefficients $-$ $_c$) calculated by *SHELX* reveal the true conformation of the side chain.

For more than 20 amino-acid side chains, discrepancies between the AD map and the refined map were found, thus involving more than 10 of the protein. ome of these cases involved residues where the refined electron density only revealed one out of two or more possible side-chain conformations. For other residues, the side chain had initially been modelled in an incorrect conformation and the real orientation did not appear in the 2 $-$ $_c$ map. Finally, for some of the residues in the psoriasin structure it was discovered in the comparison with the AD map that only the low-occupancy conformation had been modelled. These errors could have arisen from slight differences between the crystal used for the 1. A MAD phasing and that used for the 1.05 A

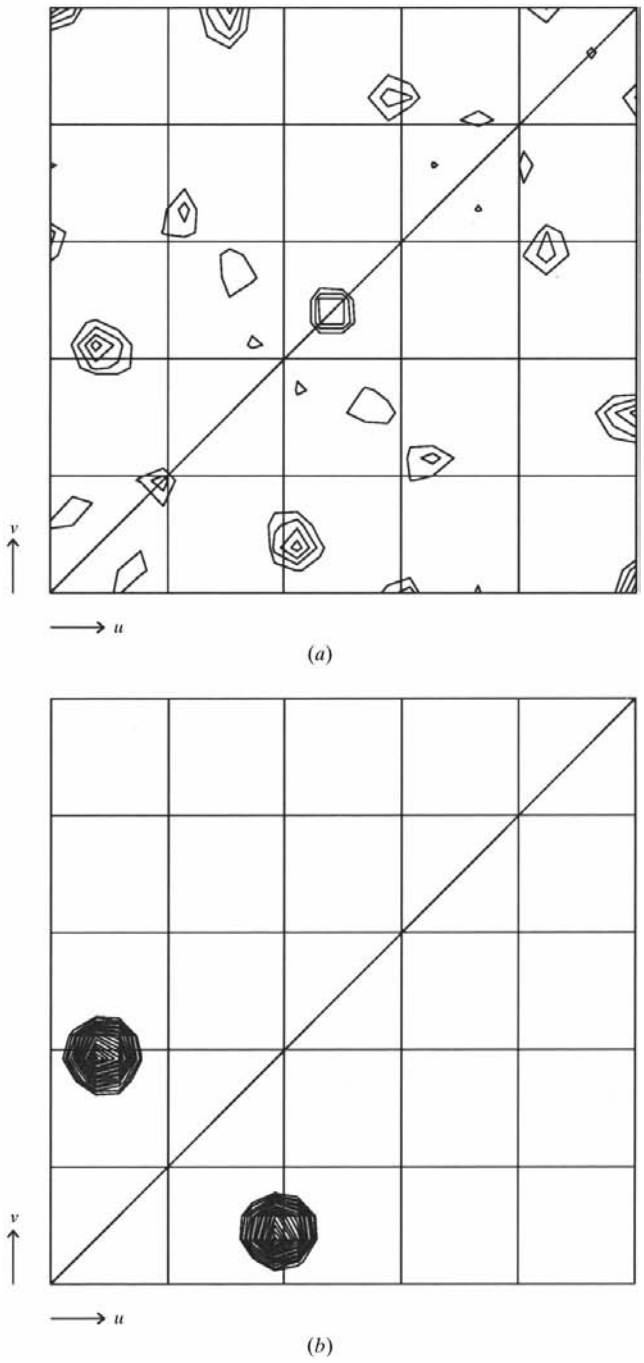

Figure
(*a*) The 1 4 Harker section of the filtered anomalous difference atterson function. Contouring is 1σ to 10σ in 0.5σ steps. ( ) The corresponding atterson section calculated from the refined position of the n atom.

refinement. They show, however, that even at atomic resolution, least-s uares-based refinement cannot always escape the in uence of the model-calculated phases.

. . Direct phasing using native data

The native calcium-bound psoriasin crystalli es in two different space groups depending on whether divalent inc is present or not and this has been shown to be the result of binding of inc to a high-affinity site within the protein (Brodersen *et al.*, 1). Crystals formed in the presence of inc belong to the tetragonal space group 4_32_12 and diffract to about 2.0 A. A native data set was originally collected at 0.863 A (14.368 keV) from these crystals with molecular replacement in mind (for a data-collection summary see Table 1). The wavelength used for data collection is located on the high-energy side of the inc absorption edge at 1.2837 A (.660 keV), but far from the edge, at a point where the value of $''$ is only 2.0 e^- or only about half of what would be observed close to the absorption edge.

Even though the structure was solved by molecular replacement, we have taken a thorough look at the anomalous atterson function to see whether the position of the n atom within the unit cell can be found without the use of an initial search model. n average, the anomalous signal strength as measured by $|\Delta|/\sigma(\Delta)$ is less than one and the anomalous differences are only clearly significant below 3 A resolution. At higher resolution (even though the anomalous contribution theoretically should increase relative to) the anomalous signal is buried in measurement noise.

Given that the determination of heavy-atom positions by inspection of the atterson function is one of the bottlenecks in phasing methods based on difference data [MI (A) MAD AD], the uestion is how to obtain the most from the available information. In this case, careful rejection of outliers turned out to be critical for interpretation of the anomalous atterson function. Fig. 5(*a*) shows the Harker section 1 4 of the anomalous atterson function calculated using all anomalous differences not e ceeding 25 of the value of the observed structure factor . This is a very reasonable criterion, in that the signal even for the strongest anomalous scatterers at the absorption edge rarely constitutes more than 25 of the structure-factor amplitude. In addition to this relative magnitude criterion, a limit on the absolute $|\Delta|$ was applied. The result of the filtering of the anomalous differences is a significant improvement in the signal-to-noise ratio of the relevant Harker section. In fact, the position of the inc ion could not be determined from the non-filtered atterson function (not shown). The correct peaks for determination of the coordinates of the inc are now seen at about 2σ above the noise as the only peaks not on any a is. Fig. 5( ) shows the corresponding atterson calculated from the inc position.

The position of the single inc ion in the asymmetric unit was now used for calculation of phase probability distributions in *SHA* . The resulting electron-density map is now of relatively poor uality and does not contain enough information for tracing of an initial model. The overall map correlation between the *SHA* map and the final refined $2 - _c$ map is 22 . However, when subjected to 70 cycles of solvent ipping in *S L* , the map improves dramatically and the resulting electron density is of a uality close to the final refined map (Fig. 6). The overall map correlation between the solvent- attened map and the final refined map is about 80 . In the solvent- attened map, all parts of the protein main and side chains are visible and in some places multiple conformations can be seen.

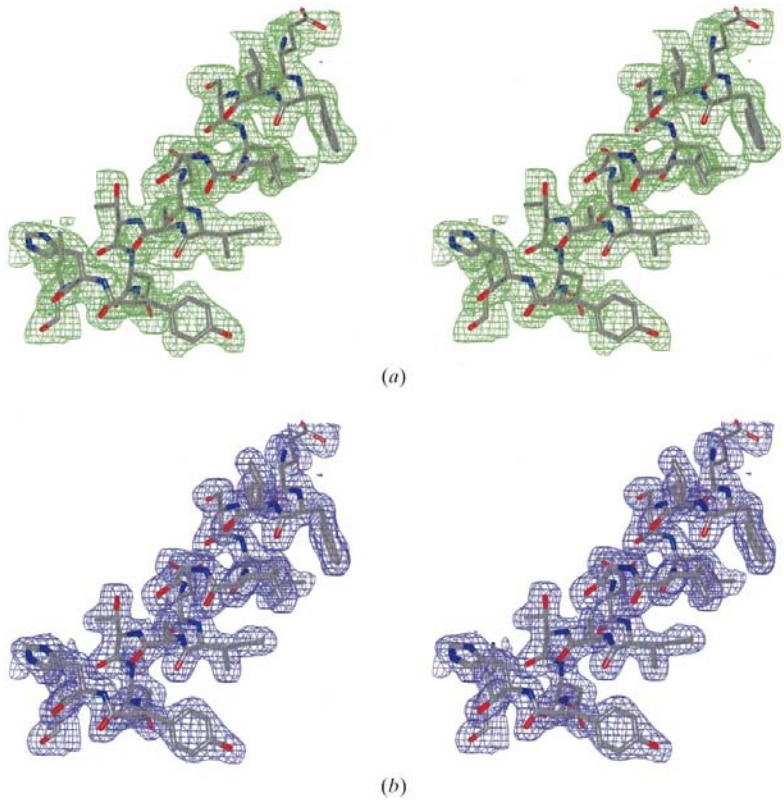

Figure 6
Two stereographs showing the electron density resulting from phasing with a single inc ion. (*a*) The electron density after solvent attening in *S L* . Contours are at 0.5σ. ( ) The refined $2 - _c$ electron density from *SHELX*. Contours are at 1.5σ.

.6. Effect of completeness and redundancy

The data set used here is .0 complete and has a high redundancy of 7.0. In order to investigate the effect of these two parameters on the phasing, a series of truncated data sets were produced by gradually removing images at the end of the φ-range. In contrast to, for e ample, applying an upper resolution limit, this does not produce artificially truncated data since it merely corresponds to terminating the data collection at an earlier stage. As can be seen in Fig. 7, in such a case the redundancy will decrease almost linearly when images are removed, whereas the completeness will remain at almost 100 until a certain point where it starts to

drop dramatically. hasing in *SHA* was carried out with 13 different truncated data sets ranging from the above 100 completeness and a redundancy of 7.0 down to 70 completeness and a redundancy of 1.5 (Fig. 7). As judged by the uality of the resulting electron-density maps, the outcome of the phasing does not seem to depend much on redundancy. At 8.7 completeness and a redundancy of 3.8, the correlation between the solvent- attened *SHA* map and the final refined map is still 72 . However, when the (anomalous) completeness starts to drop, the map correlation coefficients also deteriorate rapidly. In fact, at 3.3 completeness and a redundancy of 2. , solvent attening is no longer able to bootstrap the initial phases. In other words, the final correlation reaches its ma imum of 80 e actly at the point where the completeness reaches 100 . Thus, for reliable and consistent phasing using weak single-wavelength data, complete but not necessarily highly redundant data seems to be needed, provided it is measured accurately.

. Discussion

In the e periment with AD phasing using the strong anomalous scatterer holmium, we compared the AD and 2 − $_c$ maps for our assessment of the e tent of model bias. In a further step, however, we took advantage of the accuracy of the independent phases to calculate an unusual type of difference map with Fourier coefficients (− $_c$, φ), where represents the measured structure factors, $_c$ the structure factors calculated from the final model and φ the phases output from the solvent- attening step. This map shows a vast amount of detail close to the structure as well as long stretches of density in the solvent regions. ndoubtedly, this type of map presents valuable information with respect to the lack of ability of the refined model to describe the measured data. e did not, however, further assess the significance of this additional information, but this type of map could be valuable as an investigation tool in future studies.

bserving chemically reasonable density features in the AD map or in the AD-phased difference map is not sufficient by itself to justify an update of the model. After modelling of the additional conformers found from inspection of the AD map, the traditional cross-validation indicator ($_{free}$) should theoretically decrease as a result of bias reduction. In these e periments, we have used the same 3 free set of re ections that was used in the original refinement against the 1.05 A data to judge whether the side-chain modifications were significant. However, after a few more cycles of refinement in *SHELX* 7, we could not observe a significant change in $_{free}$ regardless of the resolution range. This is consistent with the idea that a global uality indicator based on a subset of the available measurements is insensitive to small alterations in the model. After all, the newly introduced conformations constitute few atoms all with occupancies well below unity and the total additional scattering power of these atoms is thus very small. In addition, it must be borne in mind that the same number of additional parameters are introduced to the refinement regardless of the occupancy of the atoms involved.

In another approach to assess the significance of the improved modelling, we used the same refinement program to estimate the standard deviations of the refined occupancies of the additional side-chain conformations. Inversion of the full least-s uares matri for all but the smallest proteins is a computationally challenging procedure. However, by fi ing the positional parameters of all atoms we were able to invert the full Hessian matri consisting of atomic displacement parameters and occupancies for the entire protein (200 residues). For the lysine residue mentioned above (LysA48), the occupancies of the two conformations refine to 0.771 (old conformation) and 0.22 (new conformation) with an estimated standard deviation (e.s.d.) of 0.01 8. The reason there is only one e.s.d. per two conformations is that modelling of a double conformation only introduces one additional parameter, namely the occupancy of the first conformation α. The occupancy of the second conformation is then constrained to $1 - \alpha$. The overall results show that all additional conformations are indeed significant at the 3σ level.

The apparent contradiction between the $_{free}$ and the e.s.d.s can be e plained by the lack of sensitivity of the free factor as a significance test in this case. Most of the information concerning the alternate conformations resides in the high-resolution data and a 3 sample may not be enough for proper cross-validation of these fine details. This lack of sensitivity could also arise from an inherent inability of the atomic parameter model to represent accurately high-resolution features in the electron density. hile an atomic model provides a convenient way to e change information about a molecule (or store it in a database), it may only be one of several possible interpretations of the e perimental data. Thus, to represent certain features of a structural model, an

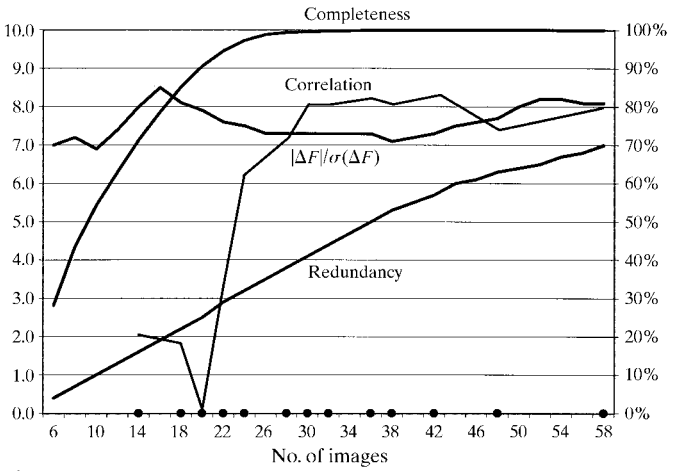

Figure
The breakdown of data completeness, redundancy and $|\Delta|/\sigma|\Delta|$ with the number of images included in the processing. 'Correlation' is the map correlation between the solvent- attened AD map output from *S L* at any given stage and the final refined 2 − $_c$ map from *SHELX*. Completeness, map correlation and $|\Delta|/\sigma|\Delta|$ are measured in percentage on the right *y* a is, whereas the redundancy is measured in absolute values on the left *y* a is. The solid circles on the a is indicate the truncated data sets used for phasing.

electron-density map calculated from good phase estimates can in some cases provide more information than the refined atomic model. However, when the phases are relatively weak, describing the phases by discrete phase values (as re uired to perform a Fourier transform) is not a satisfactory appro - imation. In such cases, the resulting phase probability distribution is broader and possibly bimodal, which makes it impossible to appro imate by a single phase value.

hen the phase information is too weak to provide an accurate initial electron-density map, the e perimental information is best represented in reciprocal space as structure-factor amplitudes with phase probability distributions. To use this information with the least possible bias, it is necessary to refine the atomic model against a ma imum-likelihood target in reciprocal space. This approach has recently been implemented in several programs (*X L* , Brunger *et al.*, 1 87 *E AC*, Murshudov *et al.*, 1 7 *S E* , Bricogne, 1 3, 1 7 Bricogne & Irwin, 1 6) however, none of these programs yet support both refinement of anisotropic factors and partially occupied conformations.

The uality of the phase estimates which are based on the single-wavelength anomalous signal from inc is remarkable for data at this resolution (2.0 A). Compared with a structure solution by molecular replacement alone, this approach produces an electron-density map free from bias of any search model, as only the position of the n atom is used for phasing.

ur results do not only prove the usefulness of density-modification procedures in phasing, but also the power of *SHA* to create a precise initial envelope of the molecule. In other words, the map correlation coefficient calculated directly after the *SHA* run is not a good indicator of the uality of the calculated phases. The real discriminator is the density-modification procedure, which is capable of dramatically improving the map uality in the case of a useful initial envelope.

. Conclusions

The present study first of all underlines the fact that anomalous scattering, when present, provides e tremely valuable phase information to high resolution. In addition, AD phasing has an important and usually overlooked advantage over I phasing. In AD, the knowledge of the heavy-atom substructure can be used not only to model the anomalous diffraction but also to predict the normal diffraction of these atoms as a known fragment of the whole molecule. Taking this information into account (which is automatically performed in *SHA*) significantly reduces the degree of bimodality in the phase probability distribution as described by Hendrickson & Teeter (1 81).

In practice, any detectable anomalous signal should be e ploited even if the data was not deliberately collected to optimi e it. In general, it is likely that AD phasing can be carried out as far as the resolution limit for crystals containing anomalous scatterers. uch scatterers are either naturally present in the macromolecule or can be introduced by a number of well known techni ues (e-Met, heavy-atom derivatives, bromouridine *etc*). Even atoms have been shown to be useful sources of anomalous signal in some cases (Hendrickson & Teeter, 1 81 Dauter *et al.*, 1). hases derived from such e periments will not only help to reduce bias in crystallographic refinement, but also in structure solution by molecular replacement as well as ligand-binding studies using Fourier difference maps.

The results presented in this paper also demonstrate the importance of proper data treatment in macromolecular crystallography. ignificant anomalous or isomorphous differences may be partly disguised by badly measured intensities and can thus impede the direct determination of heavy-atom positions. Additionally, we have shown that near-complete (anomalous) data is crucial for the estimation of initial phases by the AD method. In the case of molecular replacement, a small anomalous signal can help circumvent model bias. ur results also show that with modern phasing techni ues based on fully edged ma imum likelihood (such as *SHA* in combination with *S L*), less information is re uired to produce an interpretable map than has been the case previously. Here, we have shown how an ordinary 2.0 A native data set, which was not even collected to ma i-mi e the anomalous contribution of the inc, can be used to solve the structure of human psoriasin from scratch. Thus, for relatively small metal-binding proteins, a full-blown multiple-wavelength (MAD) or multiple-derivative (MI) e periment will in many cases be unnecessary.

The authors wish to thank igrun Nols e for collecting the original diffraction data on the calcium-bound psoriasin crystals and Dr aymond Brown for help and discussions with respect to the filtering of anomalous differences. Funding was provided by a special programme for biotechnology under the Danish Natural cience esearch Council and a Hallas-M ller fellowship from the Novo Nordisk Foundation to MK.

References

Abrahams, J. . & Leslie, A. G. . (1 6). *Acta Cryst.* D**5** , 30–42.
Bijvoet, J. M. (1 4). *at re L* , , 888–8 1.
Biou, V., hu, F. & amakrishnan, V. (1 5). *E* . (16), 4056–4064.
Bricogne, G. (1 3). *Acta Cryst.* D , 37–60.
Bricogne, G. (1 7). *et s E y l.* , 14–1 .
Bricogne, G. & Irwin, J. J. (1 6). *r cee s t e CC St y ee e . acr lec lar e e e t*, edited by E. Dodson, M. Moore, A. alph & . Bailey, pp. 85– 2. arrington Daresbury Laboratory.
Brodersen, D. E., Et erodt, M., Madsen, ., Th gersen, H. C., Celis, J. E., Nyborg, J. & Kjeldgaard, M. (1 8). *Str ct re*, **6**, 477–48 .
Brodersen, D. E., Nyborg, J. & Kjeldgaard, M. (1). *c e stry*, (6), 16 5–1704.
Brunger, A. T., Kuriyan, J. & Karplus, M. (1 87). *Sc e ce*, **5**, 458–460.
Burling, F. T., eis, . I., Flaherty, K. M. & Brunger, A. T. (1 6). *Sc e ce*, , 72–77.
Chen, L., ose, J. ., Breslow, E., ang, D., Chang, .- ., Furey, . F. Jr, a , M. & ang, B.-C. (1 1). *r c. atl Aca . Sc . SA*, , 4240–4244.
Collaborative Computational roject, Number 4 (1 4). *Acta Cryst.* D**5** , 760–763.

Coster, D., Knol, K. . & rins, J. A. (1 33). . *ys.* **6** , 345–36 .
Dauter, ., Dauter, M., La Fortelle, E. de, Bricogne, G. & heldrick, G. M. (1). . *l. l.* , 83– 2.
Dauter, ., Lam in, V. . & ilson, K. . (1 7). *C rr.* . *Str ct. l.* , 681–688.
French, G. . & ilson, K. . (1 78). *Acta Cryst.* A , 517–525.
Furey, . F., obbins, A. H., Clancy, D. ., inge, B. C. & tout, C. D. (1 86). *Sc e ce*, (473), 704–710.
Green, D. ., Ingram, V. M. & erut , M. F. (1 54). *r c.* . *S c. L Ser. A*, **5**, 287–307.
Haas, D. J. (1 68). *Acta Cryst.* B , 604.
Harvey, I., Hao, ., Duke, E. M. H., Ingledew, . J. & Hasnain, . . (1 8). *Acta Cryst.* D**5** , 62 –635.
Helliwell, J. . (1 2). *acr lec lar Crystall ra y t Sy c r tr a at* . Cambridge niversity ress.
Hendrickson, . A. (1 1). *Sc e ce*, **5** , 51–58.
Hendrickson, . A., mith, J. L. & heriff, . (1 85). *et s E y l.* **5**, 41–55.
Hendrickson, . A. & Teeter, M. M. (1 81). *at re L* , , 107–113.
Hodel, A., Kim, . & Brunger, A. T. (1 2). *Acta Cryst.* A , 851–858.
Honl, H. (1 33). . *ys.* , 1–16.
Jones, T. A. & Kjeldgaard, M. (1 7). *et s E y l.* , 173–207.
Kleywegt, G. A. & Jones, T. A. (1 6). *Acta Cryst.* D**5** , 826–828.
La Fortelle, E. de & Bricogne, G. (1 7). *et s E y l.* **6**, 472–4 4.
Laue, M. von (1 16). *A* . *ys.* **5** , 433–446.
Leslie, A. (1 3). *r cee s t e CC St y ee e* . *ata C llect a r cess* , edited by L. awyer, N. Isaacs & . Bailey, pp. 44–51. arrington Daresbury Laboratory.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1 7). *Acta Cryst.* D**5** , 240–255.
Nava a, J. (1 0). *Acta Cryst.* A **6**, 61 –620.
Nava a, J. (1 4). *Acta Cryst.* A**5** , 157–163.
Nols e, ., Thirup, ., Et erodt, M., Th gersen, H. C. & Nyborg, J. (1 7). *Acta Cryst.* D**5** , 11 –121.
twinowski, . (1 3). *r cee s t e CC St y ee e* . *ata C llect a r cess* , edited by L. awyer, N. Isaacs & . Bailey, pp. 56–62. arrington Daresbury Laboratory.
twinowski, . & Minor, . (1 7). *et s E y l.* **6**, 307–326.
ugrath, J. . (1 7). *et s E y l.* **6**, 286–306.
ead, . J. (1 7). *Str ct re*, , 11–14.
odgers, D. . (1 7). *et s E y l.* **6**, 183–203.
ossmann, M. G. & Blow, D. M. (1 62). *Acta Cryst.*, **5**, 24–31.
oyer, . E. Jr, Hendrickson, . A. & Chiancone, E. (1 8). . *l. C e* . **6** (35), 21052–21061.
heldrick, G. M. (1 3). *SHELX* 3 *A r ra r Crystal Str ct re e e e t.* niversity of Gottingen, Germany.
heldrick, G. M. (1 7). *rect et s r S l acr lec lar Str ct res*, edited by . Fortier, pp. 11 –130. Dordrecht Kluwer Academic ublishers.
heldrick, G. M. & chneider, T. . (1 7). *et s E y l.* **6**, 31 –343.
mith, J. L., Hendrickson, . A. & Addison, A. . (1 83). *at re L* , , 86–88.
mith, J. L., Hendrickson, . A., Hon atko, . B. & heriff, . (1 6). *c e stry*, **5**, 5018–5027.
allace, B. A. (1 86). *ys.* . (1), 2 5–306.
allace, B. A., Hendrickson, . A. & avikumar, K. (1 0). *Acta Cryst.* B **6**, 440–446.
ang, B.-C. (1 85). *et s E y l.* **5**, 0–112.