CROATICA CHEMICA ACTA CCACAA **72** (2–3) 673–684 (1999)

ISSN-0011-1643
CCA-2605

Original Scientific Paper

Crystallographic MAD Phasing Strategies Explored Using ELETTRA Sincrotrone Mn K-Edge Data to 2.1 Å and Use of CHESS Establishes the Diffraction Resolution Limit as 0.92 Å for the Protein Mn, Ca Concanavalin A*

Nicola S. Hunter,[a] *Thomas J. Gleichmann,*[a,§] *Madeleine Helliwell,*[a] *Luisa Barba,*[b] *Edoardo Busetto,*[c] *Andrea Lausi,*[c] *Helen J. Price,*[a] *and John R. Helliwell*[a,**]

[a] *Section of Structural Chemistry, Department of Chemistry, University of Manchester, M13 9PL, UK*

[b] *Istituto di Strutturista Chimica, CNR, Stada Statale 14, Km 163.5, Basovizza, 34102 Trieste, Italy*

[c] *Sincrotrone Trieste, Stada Statale 14, Km 163.5, Basovizza, 34102 Trieste, Italy*

Received October 12, 1998; revised April 26, 1999; accepted April 30, 1999

Multiwavelength anomalous dispersion (MAD) data have been collected from a single crystal of the protein concanavalin A so as to evaluate different combinations of wavelengths for crystallographic structure determination. Data were recorded to 2.1 Å resolution on a flash frozen crystal at three wavelengths about the Mn K-edge (1.8951 Å, 1.8940 Å, 1.800 Å) using synchrotron radiation at ELETTRA's Sincrotrone Trieste 'XRD' beamline. This is one of the longest wavelength K-edge MAD studies undertaken to date.

---

* Dedicated to Professor Boris Kamenar on the occasion of his 70th birthday.

§ Current address: Anorganisch-Chemisches Institut der Universität Münster, Wilhelm Klemm Str. 8, D-48149 Münster, Germany.

** Author to whom correspondence should be addressed. (E-mail: John.Helliwell@man.ac.uk)

Anomalous and dispersive Patterson maps are seen to be of high quality and indicate a high occupancy for the manganese binding site. This is confirmed also in the MAD phase determination and electron density maps. Finally 0.92 Å data recorded at CHESS indicates the prospects available for combined phasing strategies based on MAD to medium/high resolution along with ultra high resolution data.

Key words: Mn, Ca concanavalin A, protein crystallography, MAD phasing, ultra-high resolution, ELETTRA sincrotrone, CHESS synchrotron, Mn K-edge, lectin protein, direct methods in protein crystallography.

INTRODUCTION

Multiple wavelength anomalous dispersion phasing is the approach that side-steps the 'hit and miss' nature of multiple isomorphous replacement in protein crystal structure determination.[1] A general 'MAD' phasing approach, has been suggested involving seleno-methionine.[2] Concanavalin A is a metalloprotein where Mn and Ca are the natural metals. It is a legume lectin found in the jack bean (*Canavalia ensiformis*). The precise role of concanavalin A in the plant is not yet fully understood but it is thought to be involved in protein saccharide cross-linking and possibly affords the plant an anti-fungal plant defence strategy. The protein exists as a tetramer arranged as a dimer of dimers at neutral pH, with the monomer constituting the crystallographic asymmetric unit in the crystal form presented here. Each monomer consists of 237 amino acids and has a molecular weight of 25 kDa. There are two metal binding sites denoted S1 and S2. The S1 site can bind divalent transition metal ions such as Ni, Cd, Co, Zn and Mn. The S2 site binds Ca but may also accept Cd ions. This site can only be occupied once the S1 site is filled. In this study Mn K-edge MAD data were collected on the »native« protein *i.e.* as extracted from the jack bean for which the precise occupancy of the Mn ion in the S1 site is of interest.[3]

Concanavalin A crystals can also be prepared from protein in which the metals are removed by dialysis in acid and the transition metal binding site then repopulated by a single metal. These have been found to diffract at room temperature to a resolution which depends on the metal but is always higher than that observed for the native protein crystals.[3] Already it has been possible to undertake a 0.94 Å resolution study (comprising ≈ 117000 reflections) of cryo-cooled native concanavalin A using a »1k« CCD electronic area detector at CHESS.[4] Using the specially prepared 'pure' Mn, Ca concanavalin,[3] data have been measured to 0.92 Å at the Cornell CHESS synchrotron comprising ≈ 139 000 reflections.

EXPERIMENTAL DATA AND RESULTS

MAD Experimental Data Collection and Processing

MAD Data from ELETTRA

Crystals of native concanavalin A were obtained by dialysis, the methodology for which has already been established.[3] The ELETTRA synchrotron radiation source at Trieste was first used to measure the X-ray absorption spectrum at the Mn K-edge (see Figure 1) from a number of crystals packed into a thin walled glass capillary. MAD data were then collected from a single cryo-cooled crystal measuring $\approx 0.3 \times 0.2 \times 0.1$ mm. The cryo-protectant used was 50% methyl pentanediol.

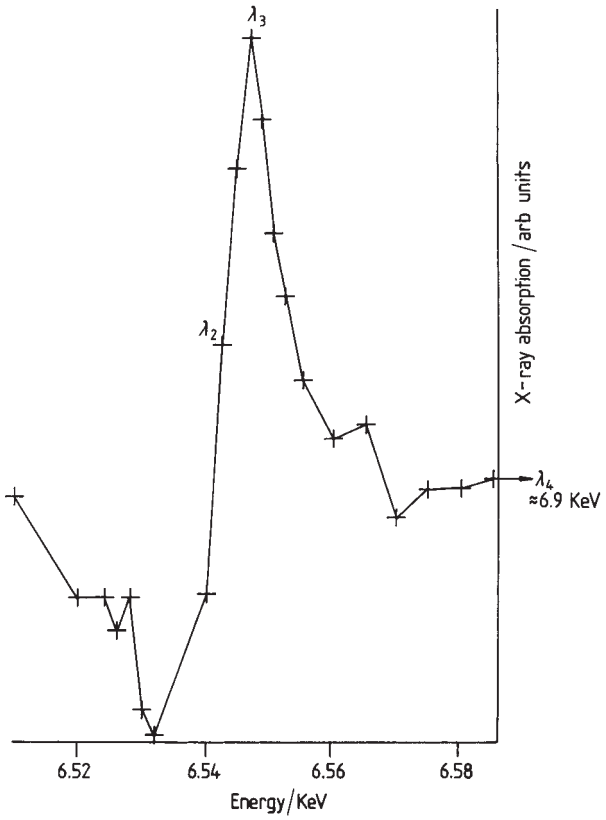


Figure 1. X-ray absorption scan across the Mn K-edge from native concanavalin A crystals measured at the ELETTRA Sincrotrone 'XRD' station. The wavelengths used for data collection are marked (for the details of the nomenclature see the main text).

Data were collected to 2.1 Å on a MAR Research image plate at the inflection point (λ_2 = 1.8951 Å), the top of the white line (λ_3 = 1.8940 Å) and at a remote wavelength away from the edge (λ_4 = 1.800 Å). The λ_2 and λ_3 data sets each consisted of 89.6° of data collected in 1.4° oscillations. The λ_4 data set consisted of 96.0° of data collected in 1.5° oscillations. A pre-edge reference wavelength (λ_1) was not collected here. Such an approach eases beamtime pressures and aids data collection efficiency for MAD experiments.[5] The three wavelength data sets were recorded in a total time of ≈ 24 hours on the 'XRD' ELETTRA multipole wiggler beamline. Each data set was processed and merged using Denzo and Scalepack.[6] Rotavata and Agrovata[7] were also used to obtain merging statistics. The cell dimensions of the crystal used for MAD data collection at ELETTRA refined to a = 89.01 Å, b = 85.75 Å, c = 61.41 Å, $\alpha = \beta = \gamma = 90°$. The space group of the crystal is *I*222. There were approx. 13000 independent reflections for each of the MAD wavelength data sets with a completeness of ≈ 94% although this was higher at λ_4. Full details are given in Table I. The data sets were put onto a common scale using CAD and SCALEIT,[7] treating the λ_2 set as the 'native'.

TABLE I

Statistics showing the quality of the MAD data as given from the output of Agrovata[7] (to 2.1 Å resolution)

	λ_2 (1.8951 Å)	λ_3 (1.8940 Å)	λ_4 (1.800 Å)
R_{merge} / %	8.4	8.4	8.7
R_{anom} / %	5.6	6.1	5.4
Total number of reflections	43407	47502	49970
Number of independent reflections	13064	13163	13972
Completeness / %	93.5	93.9	98.1
Multiplicity	3.3	3.6	3.6

Ultra-high Resolution Data from CHESS

For establishing the »ultimate resolution limit« of the Mn, Ca concanavalin A crystals, data were collected from a frozen crystal at CHESS on a multipole wiggler beamline with a 2k CCD detector.[8] Two crystals were used together here for realising high completeness of the data. One of the crystals also comprised a quick pass to obtain reflection intensities which would otherwise saturate the CCD detector. The crystal to CCD detector distance was 45 mm. Exposures were recorded with the detector displaced

vertically so that the diffraction pattern centre was respectively at the bottom edge, middle position, of the entire area. Either 0.3° (359 images) or 0.5° (193 images) crystal rotations with the exposure times per image being 9 or 15 seconds were used. In addition to these 'slow pass' exposures, 'quick pass' exposures were recorded with the diffraction pattern centre being in the centre of the detector active area, and 1° or 2° rotations (80 images) with exposure times of 5 seconds. A total of around 490,000 measurements yielded 138735 independent reflections (F^+ and F^- merged), see Table II, recorded in a total time of ≈ 24 hours at CHESS on the multipole wiggler 'A1' beamline.

TABLE II

The final merging data statistics[a] for the Mn, Ca concanavalin A »two crystal« CCD data measured at CHESS at 'ultra–high resolution' [b]

Resolution Limits Å		Completeness	R_{merge}	Unique reflections
50.0	2.27	99.4	2.9	11454
2.27	1.80	99.5	2.8	11160
1.80	1.57	99.7	3.3	11140
1.57	1.43	99.0	3.9	11005
1.43	1.33	99.2	4.7	10999
1.33	1.25	99.3	5.4	10990
1.25	1.19	98.7	6.6	10893
1.19	1.13	95.6	7.4	10578
1.13	1.09	91.1	8.9	10050
1.09	1.05	85.2	11.9	9388
1.05	1.02	77.9	16.6	8581
1.02	0.99	66.8	22.2	7532
0.99	0.96	56.2	29.0	6190
0.96	0.94	47.8	36.5	5246
0.94	0.92	33.6	44.4	3709
All data		83.4	3.5	138735

[a] From Scalepack.[6]

[b] In protein (macromolecular) crystallography the following terms are often used:
– low resolution ≈ 6 Å diffraction, molecular envelope and helices visible;
– medium resolution ≈ 3 Å diffraction, polypeptide chain is traced;
– high resolution ≈ 2 Å diffraction, restrained protein model refinement is possible;
– atomic resolution ≈ 1.2 Å diffraction, individual atoms can be seen (resolved) in electron density map;
– ultra-high resolution ≈ 0.95 Å diffraction, 'small molecule' refinement (X-ray data to parameter ratio ≥ 5) can be undertaken.

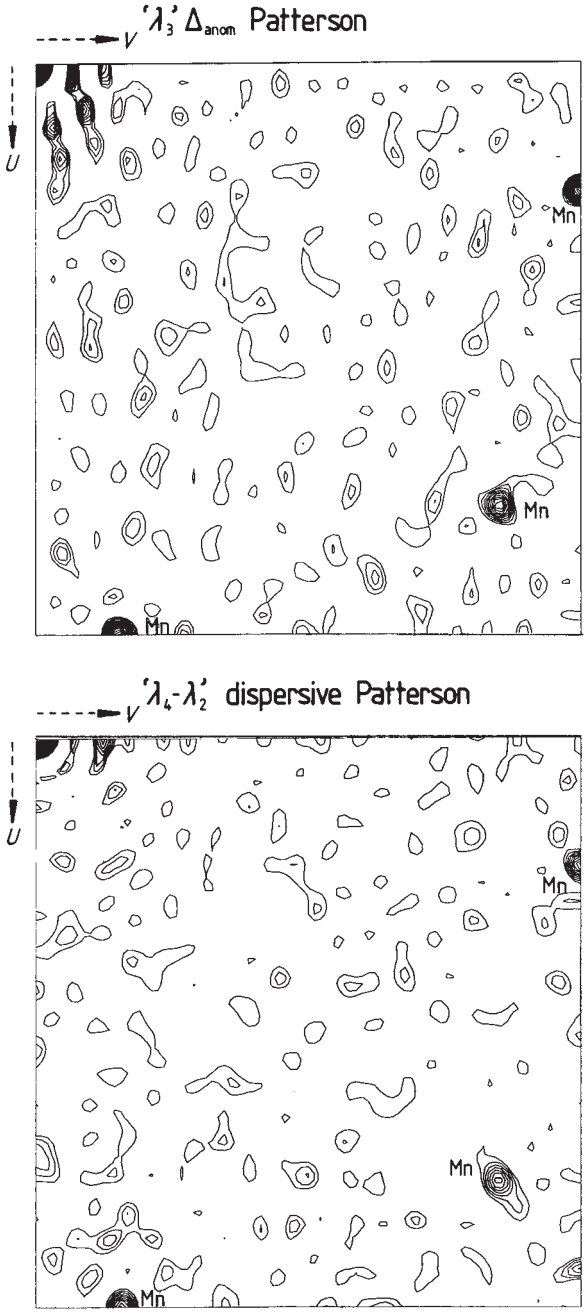


Figure 2. Difference Patterson Harker sections ($w = 0$ as example):
(a) λ_3 anomalous differences (»white line« maximum),
(b) $\lambda_4 - \lambda_2$ dispersive differences.

Anomalous and Dispersive Patterson Maps for the ELETTRA MAD Data

The clearest MAD anomalous Patterson maps were obtained from the λ_3 data set (Figure 2a) and the best dispersive Pattersons were obtained from the λ_4–λ_2 data sets (Figure 2b). This agrees with what would be expected since λ_3 gives the greatest f'' signal and λ_4–λ_2 gives the greatest $\Delta f'$ value (Table III).

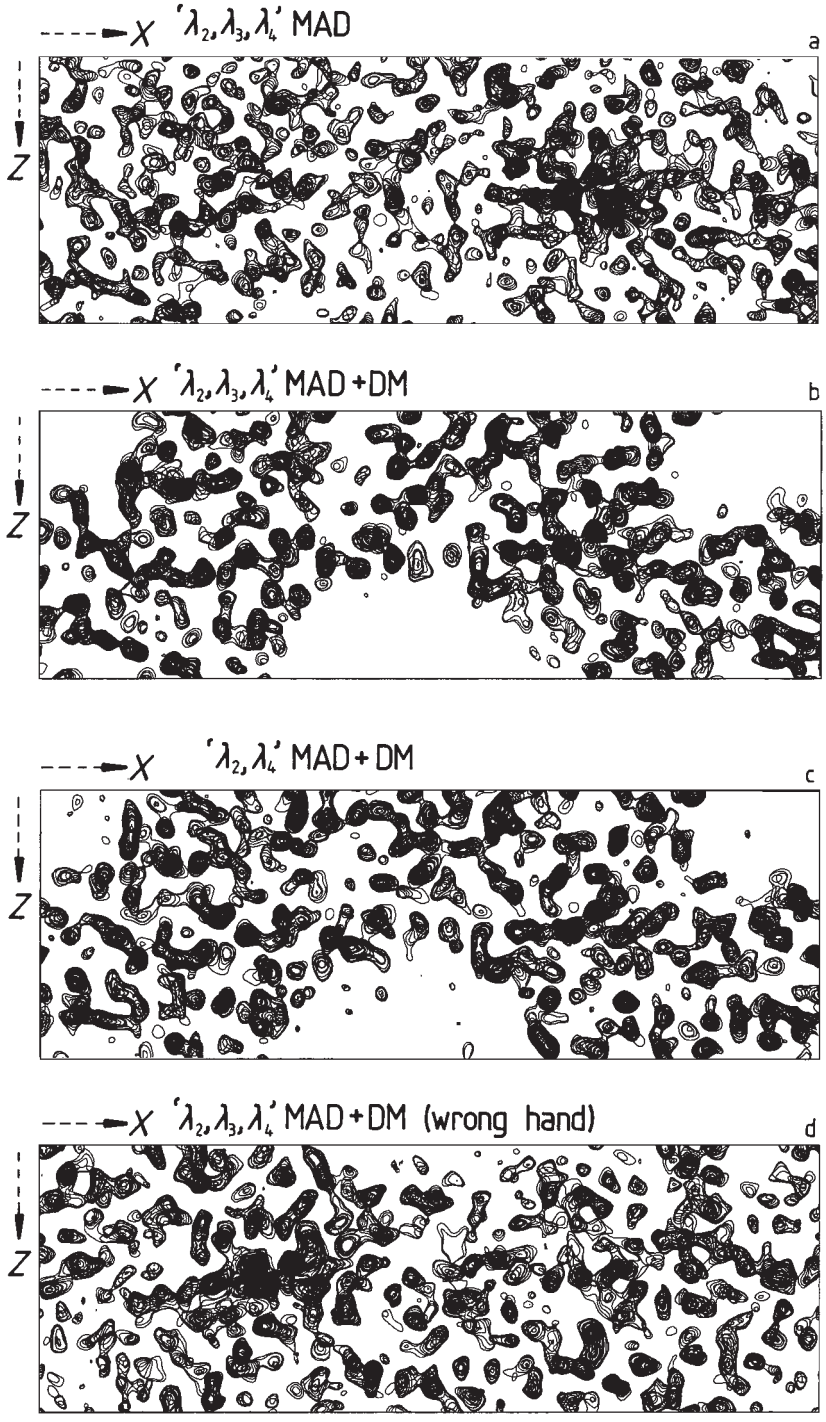
TABLE III

Values of f'' and $\Delta f'$ for Mn in Concanavalin A[a]

Data sets	f'' / e^-	$\Delta f'$ / e^-
λ_2	5.435	–
λ_3	9.688	–
λ_4	3.589	–
λ_4–λ_2	–	8.704

MAD Phasing and Electron Density Map Calculations, Including Phase Improvement Procedure

The phases from the MAD 3-wavelength data sets, based on the Mn atom coordinates alone, were used to calculate an anomalous difference Fourier map (based on λ_4 anomalous differences) and yielded the Ca atom position (the f'' for Ca at these wavelengths is present at 1.7 e^-). Both the Mn and Ca coordinates were then used in the CCP4 phasing program[7] MLPHARE to calculate MAD phases. The correct hand was determined by calculating MAD phases on both hands for the three wavelength combination. These phases were then input into DM[7,9] using the solvent flattening and histogram matching procedure assuming a 50% solvent content. Fourier maps calculated using FFT,[7] clearly showed the positive hand to be correct (compare Figure 3b and 3d). The MLPHARE and DM phase calculations were repeated for each two wavelength combination and corresponding electron density maps were produced. The figures of merit for the various wavelength combinations, before and after DM, are shown in Table IV. The electron density for the 2-wavelength (λ_2 with λ_4) case electron density (after DM) is shown in Figure 3c. For comparison Figure 3a shows the MAD 3-wavelength map without DM and Figures 3e and 3f the 'perfect' *i.e.* calculated electron density maps at 2.1 Å and 0.92 Å resolu-


X
'λ2, λ3, λ4' MAD
a
Z
X
'λ2, λ3, λ4' MAD+DM
b
Z
X
'λ2, λ4' MAD+DM
c
Z
X
'λ2, λ3, λ4' MAD+DM (wrong hand)
d
Z

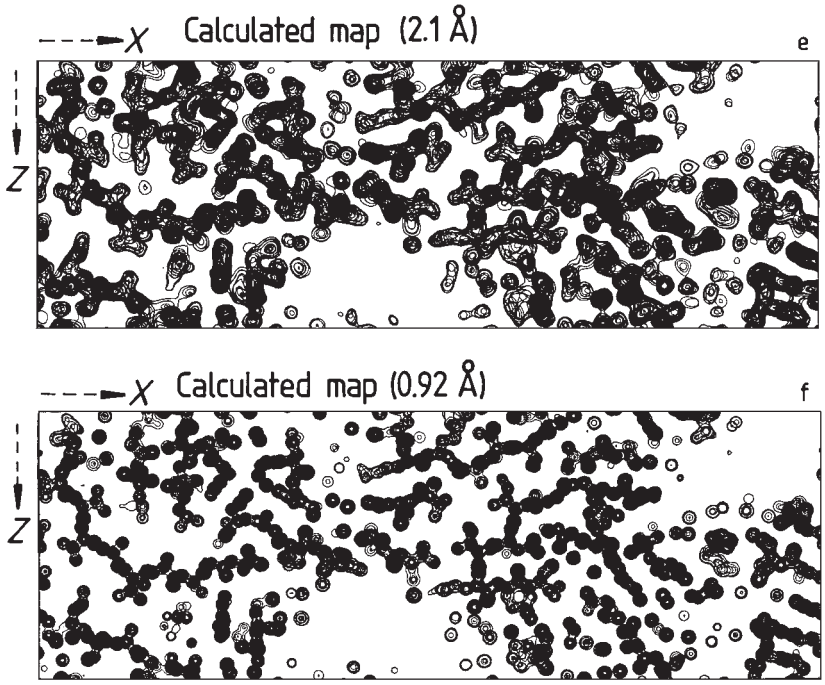


Figure 3. Electron density map slab for:
(a) MAD '3λ' phases (correct hand) to 2.1 Å (without 'DM'),
(b) MAD '3λ' + 'DM' phases (correct hand) to 2.1 Å,
(c) MAD '2λ' (*i.e.* λ_2 and λ_4) + 'DM' phases (correct hand) to 2.1 Å,
(d) MAD '3λ' + 'DM' phases (incorrect hand) to 2.1 Å,
(e) calculated phases from all atoms to 2.1 Å,
(f) calculated phases from all atoms to 0.92 Å.

tions. The calculated (*i.e.* »ideal«) electron density map, Figure 3e, shows the molecular boundary and course of the polypeptide chain in this example

TABLE IV

MAD phasing figures of merit (FOM)

Case	Description	Mean FOM-a	Mean FOM-c	Overall FOM	FOM (+DM)
1	$\lambda_2, \lambda_3, \lambda_4$	0.3539	0.2098	0.3385	0.576
2	λ_2, λ_4	0.1480	0.0988	0.1427	0.204
3	λ_2, λ_3	0.2319	0.0019	0.1914	0.357
4	λ_3, λ_4	0.2016	0.0813	0.1821	0.430

a = acentric reflections.
c = centric reflections.

slab of the unit cell. These are clear also in the MAD+DM map (Figure 3b) and to a good degree in the MAD only map (Figure 3a). Good connectivity of electron density and especially the protein molecular envelope, are evident in the 2-wavelength map MAD+DM (Figure 3c). The 0.92 Å resolution calculated map (Figure 3f) shows the atom by atom structure of this example portion of the protein and represents the ideal case to which experimental techniques in synchrotron radiation protein crystallography should aspire or regard as a target in future.

DISCUSSION AND CONCLUDING REMARKS

The MAD data from the ELETTRA Sincrotrone (Table I) and the ultra-high resolution data from CHESS (Table II) are of high quality. The issue of the occupancy of the S1 binding site in native concanavalin A (in the crystals at least!) has been settled *via* the Mn MAD K-edge experiment as being of high, rather than low Mn occupancy, as evidenced by the anomalous and dispersive Patterson and MAD phased electron density maps. The number of wavelengths needed for a well phased MAD electron density map for a protein of this size and for one Mn ion is an interesting question. Comparison of the 3- and 2-wavelength MAD+DM phased maps (Figures 3b and 3c) settles in favour of three rather than two wavelengths (*i.e.* Figure 3b is better; see also the phasing statistics in Table IV, but where λ_3, λ_4 however looks promising). Whilst two wavelengths, with F^+ and F^- at one and with F^+ or F^- at the second wavelength is sufficient in principle to solve uniquely for the phase angle,[1] and excellent experience with 2 wavelengths is seen with a MAD brominated oligonucleotide study,[10] three wavelengths (thus exciting the largest possible f'' and $\Delta f'$ are shown here to be needed, as is DM for MAD phase improvement. Nevertheless since improved electron density treatment *via* phase improvement procedures is an active area of research this should be kept as an open question for experimental and theoretical phasing methods in the future. Likewise the potentiality of combined MAD and ultra-high resolution data, along with DM style approaches and direct methods, is seen here with data being collectable to 0.92 Å (on CHESS), as well as three wavelengths to 2.1 Å (on ELETTRA). In the future then an SR beamline with exceptional brilliance, like an ESRF undulator, will allow these capabilities to be realised within one and the same SR beamline. Exciting prospects lie ahead for yet more automatic phasing and protein crystal structure solution, possibly opening up genome level numbers of protein crystal structures.[11] Specifically, MAD phasing, would yield the molecule positioning in the unit cell and course of the polypeptide chain. Then the ultra-high resolution data would be phased *via* partial structure

phasing followed by automatic atomic position refinement in an iterative manner. Finally, it has recently been shown[13] that atomic resolution (single wavelength) for a protein (14 kDa lysozyme), without a metal bound, can yield a direct structure solution. Since the majority of naturally occurring proteins are not metalloproteins (*i.e.* heaviest atom is sulphur) this is also an important step forward. Tests using the 0.92 Å data for the Mn, Ca concanavalin A seeking a direct methods structure solution, have commenced.[14]

Acknowledgements. – The ELETTRA Sincrotrone in Trieste is thanked for access to their MAD 'XRD' beamline for this experiment, funded under an ELETTRA proposal. The work is within the auspices of a collaboration in Structural Chemistry between Professor J. R. Helliwell and Professor V. Kaučič of the National Institute of Chemistry, Slovenia on metal aluminophosphates and metalloproteins; within this The Royal Society is thanked for travel support. CHESS, Cornell University, is thanked for beamtime provision for the ultra-high resolution data collection. The BBSRC is thanked for a grant (J. R. Helliwell) for funding of computer workstations and molecular graphics. T. J. Gleichmann was funded under an EU Host Institute award to J. R. Helliwell. N. S. Hunter was supported by a PhD research studentship from the BBSRC with J. R. Helliwell. Natasha Zabukovec-Logar is thanked for assistance with data collection at ELETTRA as is Dr. Ashley Deacon at CHESS to both of whom we are very grateful. J. R. Helliwell very warmly acknowledges Dr. A. J. Kalb (Gilboa) and Dr. J. Yariv (now retired) of The Weizmann Institute, Israel for many discussions and collaborations on concanavalin A and its complexes with monosaccharides; H. J. Price was funded under a British Council UK/Israel collaborative award to A. J. Kalb (Gilboa) and J. R. Helliwell.

REFERENCES

1. J. R. Helliwell, *Rep. Prog. Phys.* **47** (1984) 1403–1497.
2. W. Hendrickson, *Science* **254** (1991) 51–58.
3. A. J. Kalb (Gilboa), J. Yariv, J. R. Helliwell, and M. Z. Papiz, *J. Cryst. Growth* **88** (1988) 537–540.
4. A. Deacon, T. Gleichman, A. J. Kalb (Gilboa), H. Price, J. Raftery, G. Bradbrook, J. Yariv, and J. R. Helliwell., *J. Chem. Soc., Faraday Trans.* **24** (1997) 4305–4312.
5. J. R. Helliwell, *Proc. of Daresbury Study Weekend*, 26–28 January 1979, DL/SCI/ /R13, pp. 1–6.
6. Z. Otwinowski and W. Minor, DENZO and SCALEPACK Programs.
7. CCP4 package available from Daresbury Laboratory, see *Acta Crystallogr., Sect. D* **50** (1994) 760–763.
8. M. W. Tate, E. F. Eikenberry, S. L. Barna, M. E. Wall, J. L. Lowrance, and S. M. Gruner*, J. Appl. Crystallogr.* **28**, 196.
9. K. Cowtan and P. Main, *Acta Crystallogr., Sect. D* **54** (1998) 487–493 and references therein.
10. M. R. Peterson, S. J. Harrop, S. M. McSweeney, G. A. Leonard, A. W. Thompson, W. N. Hunter, and J. R. Helliwell, *J. Synchrotron Rad*. **3** (1996) 24-34.
11. N. E. Chayen and J. R. Helliwell, *Physics World,* May 1998.
12. S. Sasaki, KEK Report 88-14, Photon Factory, Japan (1989).

13. A. M. Deacon, C. M. Weeks, R. Miller and S. E. Ealick, *Proc. Natl. Acad. Sci. USA* **95** (1998) 9284-9289.
14. H. J. Price (1999) Ph.D. Thesis, University of Manchester, in preparation.

SAŽETAK

Kristalografski postupak određivanja faza tehnikom MAD ispitan korištenjem podataka prikupljenih sinkrotronom ELETTRA pri apsorpcijskom pragu Mn K uz rezoluciju od 2,1 Å. Za protein Mn, Ca konkanavalin A upotreba CHESS sinkrotrona ustanovljuje difrakcijsko ograničenje razlučivanja pri 0,92 Å

Nicola S. Hunter, Thomas J. Gleichmann, Madeleine Helliwell, Luisa Barba, Edoardo Busetto, Andrea Lausi, Helen J. Price i John R. Helliwell

Jedinični kristal proteina konkanavalina A korišten je za prikupljanje difrakcijskih podataka postupkom anomalne disperzije rentgenskog zračenja različitih valnih duljina (multiwavelength anomalous dispersion, MAD) kako bi se procijenila kombinacija različitih valnih duljina za kristalografsko određivanje strukture. Podaci su prikupljani uz rezoluciju od 2,1 Å s trenutačno smrznutog kristala primjenom sinkrotronskog zračenja (ELETTRA's Sincrotrone Trieste 'XRD' beamline) triju valnih duljina u blizini apsorpcijskog praga Mn K (1,8951 Å, 1,8940 Å, 1,800 Å). Do danas nije obavljen niti jedan MAD difrakcijski pokus s većom valnom duljinom koja bi bila bliže apsorpcijskom pragu K. Anomalne i disperzijske Pattersonove mape visoke su kakvoće i pokazuju visoku zaposjednutost veznog mjesta mangana. Ta je činjenica također potvrđena određivanjem faza i mapa elektronske gustoće primjenom postupka MAD. Konačno, podaci prikupljeni uz rezoluciju od 0,92 Å CHESS sinkrotronom ukazuju na mogućnosti udruživanja istraživanja temeljenih na eksperimentalnim podacima dobivenih tehnikom MAD uz osrednje i visoko i onih dobivenih uz krajnje visoko razlučivanje.