ESRF	**Experiment title:** Structure determination of UDP-galactomutase	**Experiment number:** LS-1477
Beamline: ID14 4	**Date of experiment:** from: 26/08/99 to: 27/08/99	**Date of report:** 11/01/2000
Shifts:	**Local contact(s):** Gordon Leonard	*Received at ESRF:* 10 JUIN 2000

Names and affiliations of applicants (* indicates experimentalists):

Dr. David Sanders *

Dr. Marie-France Giraud *

Dr. Stephen McMahon *

Dr. James Naismith

Report:

UDP-Galactopyranose mutase (mutase)

Se-Met mutase crystals were brought to Grenoble in sitting drops and were mounted and frozen on location. We first attempted to collect an anomalous scattering curve using one of our crystals, however, we were unable to accurately locate the peak energy value and so used a value determined from a previous experiment. As past experiment have shown that our crystals are highly mosaic, we collected data on small crystals (0.2x0.2x0.01). These crystals are diamond shaped plates. Data was collected in 1° frames, for 30 sec. exposures (symmetry P21).

A native data set and two complete Se-Met data sets (3 wave-lengths each) were collected, however, neither of the sets were collected from a single crystal. The data was processed using Denzo and Scalepack, and is summarized below :

Set	Resolution	R_{merge}	Completeness	Mosaicity	Redundancy
Peak 1	2.8 Å	8.3 %	95.8 %	1.72	1.7
Inflection 1	2.8 Å	7.9 %	96.4 %	1.75	1.7
Remote 1	3.2 Å	15.0 %	98.4 %	1.93	1.85
Peak 2	3.2 Å	7.6 %	77.8 %	1.37	1.2
Inflection 2	3.2 Å	8.8 %	76.9 %	1.26	1.35
Remote 2	3.2 Å	8.1 %	88.4 %	0.81	1.5
Native	2.7 Å	9.0 %	99.4 %	0.73	2.9

Data set 1 proved to be the most useful and will be the one referred to here, although data set 2 was also used, it was fairly non-isomorphous between data sets (Scaling gave an R_{factor} of 27 % between Remote 2 and both Inflection 2 and Peak 2), and didn't improve any of the procedures.

SOLVE was used to find the Se-Met sites. To 3.5 Å, the program found 9 Se-Met sites, with a FoM of 51.8 and a Z-score of 35.0. To 2.8 Å, the FoM and Z-score were 33.9 and 36.5 respectively, with the same 9 sites found.

NCSFIND was then used to calculate all of the symmetry mates for the 9 sites, and a match and 8 of the sites were found to be symmetry related (4 & 4), with an obvious two-fold axis. The centre of the two-fold was calculated and a 50 Å spherical mask was generated using NCSMASK. At this point MLPHARE was also used to attempt to improve the sites found by SOLVE. The native data was added in with MLPHARE, as was the second Se-Met data set, and the results after DM were as follows :

SOLVE alone (to 3.2 Å) Correlation (start-finish)	0.166 - 0.723
MLPHARE (to 2.7 Å)	0.038 - 0.696
(to 3.2 Å)	0.045 - 0.666

The maps were then visually inspected, and the SOLVE alone maps were deemed to be the best. MLPHARE and the Se-Met data was then used as a SIROAS experiment. This didn't prove to be very effective, as attempts to refine the anomalous signal demonstrated to us that the anomalous data was very poorly measured or absent.

The additional phase estimates (though poor) from the SIROAS were combined with the original SOLVE data through DM (Correlation 0.167 - 0.736) and the maps generated from this were used to improve the mask by placing a 'pseudo-protein' into the density. This

process was repeated a number of times and the resolution was extended to 2.7 Å with the native data set. The best mask, when used in DM, gave a correlation increase from 0.208 to 0.843.

This result was combined with a previously collected, non-isomorphous Se-Met data set (3.5 Å) and an additional native data set (2.9 Å) (fairly isomorphous) using DM_Multi. The native data set (2.7 Å) had a Correlation increase of 0.935 from 0.155 and an R_{factor} decrease of 0.635 to 0.381.

The results from DM_Multi were then used to generate a final set of maps which were used for building the protein model. The model was built using O. First a bones model was calculated by MAPMAN, and this was used as a guide for a preliminary trace trhough connected density. This was improved by mapping secondary structure elements onto the trace. This gave 28 secondary structure elements that were then used in DEJAVU to find similar structures. The best match found only matched 12 secondary structure elements, dispersed over the whole protein rather than a localized region. This made it very difficult to determine the directionality and linkage of the secondary structure elements.

A best-guess model was taken through different types of refinement (CNS) with no improvement of the maps (best $R_{free} = 0.49$). At this point, we believe that more data is required in order to solve the phase problem enough to solve the structure.