EMBL Hamburg Structural Biology Projects
Progress report (1/9/2008)

Enzymes:

AspAT: aspartate aminotransferase from *P. falciparium*, data collected on ID23-2 to 2.9A, solved by molecular replacement (R/Rfree 26/29). Data also collected on ID29 and id14-1.

Diffraction data of the tHisF triple mutant from *Thermotoga maritima* crystals in complex with Cu were collected at ID19 and ID23-2 to a maximum resolution of 2.6Å. A solution was found using molecular replacement but due to intrinsic problems of the crystals it has not been possible to refine the model. New crystallization conditions giving other crystal forms are still to be screened.

The structure of SAG1039, a novel protein from *Streptococcus Agalactiae*, was solved using beam-time at ID19. The structure reveals a homo-dimeric four helical bundle, with similarities to the expected virulent-factor-complex.

A glutathione S-transferase from mosquito flight muscle has been solved in three different crystal forms containing 2, 4 and 8 molecules per asymmetric unit. In some structure glutathione is bound but unfortunately in none of them is there any evidence for the insecticide (temefos) that we were trying to establish the binding site for. The refinement of the three structures is essentially complete.

For interlukin 4 inducing protein (IPSE) data have been collected to 2.5 Angstrom and the structure has been solved by molecular replacement. Refinement is almost complete.

Although data sets were collected, crystals of the kinase domain of the ethylene receptor did not diffract well enough to allow structure solution.

The plant *Thalictrum flavum* produces an enzyme belonging to a novel class of S-adenosylmethionine-dependent N-methyltransferases specific for benzylisoquinoline alkaloids. After expression in E. coli and purification, the protein crystallized in space group P21. A native and a derivative dataset were collected at ID23 at wavelengths of 1.0 A and 1.7 A respectively. The structure of the enzyme was determined using SIRAS method using Xenon derivative to resolution 2.0 A.

Data for two mutants of Strictosidine synthase (G208L and G208F) have been collected. The structures have been refined.

Muscle proteins:

Titin A165-A164 tandem Ig domains. A SAD dataset was collected from Se-Met crystal on ID-29 with a resolution up to 2.8 A, the structure is under refinement with Rfactor around 30%

Proteins from Mycobacterium tuberculosis:

Rv3676 (Mtb-cAMP Receptor Protein complex with DNA motifs). Two native data sets (2.9-3.0 Angstrom) collected from BAG trips the molecular replacement was succesfull but there is no clear density for DNA. Refinement is progressing with presently Rfree/Rwork =0.38/0.36).

The structure of MSMEG_4756, the Mycobacterium smegmatis Acyl-Carrier protein Synthase, was solved to 1.8Å using beam-time at ID23.2 and ID19. The crystals of the Acyl-Carrier protein Synthase was obtained in the attempt to crystallize over-expressed

protein in M. smegmatis from M. tuberculosis.
Current goal is solve the structure of the homologous Rv2523c from M. tuberculosis, being a important drug target.
CFP10/ESAT6. The structure solution has shed light on the evolution of proteins belonging to the WXG100 protein family and provided information about the newly discovered type VII and VIIb secretion systems.
Native and many potential derivative data sets have been collected for the N-terminal domain of Rv3220c. Phase information has so far not been forthcoming.
Data were collected on crystals of a mutant of LeuD (an enzyme in the leucine biosynthetic pathway). The structure was solved using the SAD method with selenomethionine containing protein. Refinement is almost complete.

Transcription factors:
Oct-4/Sox2/UTF1 complex. Protein-Protein:DNA complex that is regulating pluripotency. 1 data set collected to 3.8 resolution, no solution. Crystallization optimization in progress. Oct4:PORE complex. Protein:DNA complex that is regulating pluripotency. 4 data sets collected to 3.2-2.6 A resolution, 1 structure has been fully refined so far.

Peroxisomal proteins:
Data were collected on a complex of the ubiquitin conjugating enzyme Pex4p and its binding partner, the peroxisomal membrane protein Pex22p. The crystals diffracted to around 3Å and i collected a complete data set. So far, molecular replacement has been unsuccessful and a number of halide/Hg soaked xtals were tried but these did not diffract. SeMet proterin is being produced

Viral proteins:
The Structure of ADRP domain from feline infectious peritonitis virus was solved to 3.1A resolution. Data was collected on ID-29 beamline.
Native X-ray data has been collected for the RNA-dependent RNA polymerase fromthe rabbit hemorrhagic disease virus to a 6.5 A resolution on the ID23-2
beamline. Optimisation of crystals is ongoing.

Lipid binding proteins:
The structure of the fatty acid and retinol binding protein, FAR7, from C. elegans has been solved and refined to 1.8Å resolution using data collected on ID23-2 and ID29. Data for another of these proteins, FAR6, has also been collected to a resolution of 3.0Å , although here crystal optimization is still in progress