

Experiment Report Form

	Experiment title:	Experiment number:
Beamline: ID14, ID29 BM30, BM14A	Date of experiment: from: March 2001 to: March 2003	Date of report: Mars 2003
Shifts: 24	Local contact(s): Joanne McCarthy, Andy Thompson, Gordon Leonard, Bill Shepard, Ingar Leiros, Frank Borel, Steffi Arzt, Lilian Jacquemet, Jean-Luc Ferrer, Philippe Carpentier	<i>Received at ESRF:</i>
Names and affiliations of applicants (* indicates experimentalists): Abergel Chantal Claverie Jean-Michel		

Report:

The review period correspond to 01.03.2001 – 01.03.2003, BAG and CRG experiments (**LS-1926 and MX-56, 30-01-569, 30-01-501, 30-01-599, 14-S-606**)

Structural Genomics of Orfan Genes:

The complete nucleotide sequences of more than 60 microbial and four eukaryote genomes are already available in the public domain and many more genomic projects are in progress throughout the world. Despite this accumulation of data, newly sequenced microbial genomes continue to reveal up to 50% of functionally uncharacterized "anonymous" genes. A majority of these anonymous genes encode proteins with

homologues in several organisms, but a significant fraction remains that exhibit no clear similarity to any other protein sequence in the databases. This set of unique - apparently species specific - sequences are referred to as "ORFans". The biochemical and structural analysis of ORFan gene products is both of evolutionary and functional interest.

The allocated beam time allowed the resolution of the following crystal structures:

E. coli IVY orfan gene product, a new type of type C lysozyme inhibitor (MAD) PDB1GPQ

Structure of the *Pseudomonas Aeruginosa* IVY protein (MR): Discovery of a new family of bacterial type C lysozyme inhibitor. PDB1HKE

References :

Abergel C., Monchois V., Lembo F. and Claverie J.-M Discovery of a new family of proteins. (in preparation).

Monchois V., Abergel C., Sturgis J., Jeudy S. and Claverie J.-M. (2001) *Escherichia coli* YkfE "ORFan" gene encodes a potent inhibitor of C-type lysozyme. *J. Biol. Chem.* 276 : (21) 18437-18441

Abergel C., Monchois V., Chenivesse S., Jeudy S. and Claverie J.-M. (2000) Crystallization and preliminary crystallographic study of b0220, an "ORFan" protein of unknown function from *Escherichia coli*. *Acta Cryst. D56* : 1694-1695

Structural genomics of *E. coli* conserved genes of unknown function in search of new anti-bacterial targets:

With more than 100 antibacterial drugs at our disposal in the 1980's, the problem of bacterial infection was considered solved. Today, however, most hospital infections are insensitive to several classes of antibacterial drugs, and deadly strains of *Staphylococcus aureus* resistant to vancomycin - the last resort antibiotic- have recently begin to appear. Other life-threatening microbes, such as *Enterococcus faecalis* and *Mycobacterium tuberculosis* are already able to resist every available antibiotic. There is thus an urgent, and continuous need for new, preferably large-spectrum, antibacterial molecules, ideally targeting new biochemical pathways. Here we report on the progress of our structural genomics program aiming at the discovery of new antibacterial gene targets among evolutionary conserved genes of uncharacterized function. A series of bioinformatic and comparative genomics analyses were used to identify a set of 221 candidate genes common to Gram-positive and Gram-negative bacteria. These genes are now submitted to a systematic 3-D structure determination protocol including cloning, protein expression and purification, crystallization, X-ray diffraction, structure interpretation, and function prediction. Bioinformatics is used to optimize most stages of this production process. Out of 110 genes processed in our laboratory - and 17 months into the project - 108 have been successfully cloned, 93 have exhibited detectable expression, 75 have led to the production of soluble protein, 42 have been purified, 12 have led to usable crystals, and 7 structures have been determined.

The allocated beam time allowed the resolution of the following crystal structures:

E. coli yqhE (MR) PDB1MZR

E. coli ydhF (MAD) (Refinement)

E. coli yliB (MAD) (Refinement)

E. coli ydiB (MAD) (1NPD Northeast Structural Genomics Research Consortium (Nesg) Target Er24)

E. coli yhbO (MR) (Refinement)

E. coli yecD (MR-MODELLER) PDB1J2R

E. coli yggV (MAD) (PAD1K7K Structural Genomics Consortium)

References:

Chantal Abergel , Bruno Coutard, Deborah Byrne, Sabine Chenivesse, Jean-Baptiste Claude, Céline Deregaucourt, Thierry Fricaux, Celine Giancesini-Boutreux, Sandra Jeudy, Régine Lebrun, Caroline Maza, Cédric Notredame, Olivier Poirot, Karsten Suhre, Majorie Varagnol and Jean-Michel Claverie. Structural genomics of highly conserved microbial genes of unknown function in search of new antibacterial targets. JSFG (submitted).

Other Projects:

Structure determination of the *E. coli* periplasmic PAL protein (MAD) PDB1OAP

Structure determination of the Acidithiobacillus Ferrooxidans Cytochrome C4 (MAD) PDB1H1O

References:

Abergel C., Walburger A., Chenivesse S. and Lazdunski C. (2001) Crystallization and preliminary crystallographic study of the peptidoglycan-associed lipoprotein from Escherichia coli. *Acta Cryst. D57* : 317-319

Chantal Abergel, Anne Walburger, Emmanuelle Bouveret and Jean-Michel Claverie. First Insigh in the conserved periplasmic domain of the *Escherichia coli* Pal protein (In preparation).

Chantal Abergel, Wolfgang Nitschke, Guillaume Malarte, Mireille Bruschi, Jean-Michel Claverie and Marie-Thérèse Giudici-Orticoni. The structure of *Acidithiobacillus ferrooxidans* c4-cytochrome: a model for complex-induced electron transfer tuning. *Structure Fold Des.* (In press)

Summary of project status during review period:

Protein Name ^a	Data set ^b	Beam-line	Date	Protein size	Unit cell dimensions (Å, °)	Space Group	Crystal size (mm ³)	Arom. Scatt.(s)	d _{max} (Å)	F _{sym} (%)	Structure Status ^b	Publication Status ^c	Comments
yvP/HEWL	native	ID14-EH2	Sept 2001	129/ 129	52.3,60.76,78.35	P21	0.4x0.4x0.3	none	1.8	5.7	Completed	In Preparation	
aphE	unusable	ID14-EH2	Sept 2001	335	57.2,69.8,86.4,76.1,71.8 ,69.4	P1	0.15x0.15x0.1	none	4	NA	Poor Data	Not Applicable	
PAL	native	ID14-EH2	Sept 2001	107	88.6,88.6,88	I4122	0.4x0.4x0.4	none	1.93	5.2	Completed	In Preparation	
yqkE	native	ID14-EH2	Sept 2001	275	139.2,145.7,79.5	C2221	0.5x0.5x0.5	none	2.16	7.5	Completed	In Preparation	
PAL	MAD	ID29	Nov 2001	107	89.2,89.2,89	I4122	0.5x0.5x0.6	Se	2.3	4.6	Completed	In Preparation	
TOLR	MAD	ID29	Nov 2001	100	46.4,46.4,188.8	P43212	0.4x0.4x0.2	Se	3.9	25	More Phasing Needed	Not Applicable	
yqhE-NADP	unusable	ID14-EH2	Dec 2001	335	139.2,145.7,79.5	C2221	0.2x0.2x0.2	none	4	NA	Poor Data	Not Applicable	
ydhF	MAD	BM30	Feb 2002	298	87.7,87.7,66.2	P63	0.3x0.2x0.2	Se	2.6	9.1	Under Refinement	In Preparation	
ybgL	unusable	BM30	Feb 2002	244	105.9,129.45,8	C2221	0.15x0.2x0.1	Se	3.5		Poor Data	Not Applicable	
ydiB	MAD	ID14-EH4	Apr 2002	288	158.5,156.5,39.9	P64	0.3x0.3x0.6	Se	2.8	10	More Phasing Needed	Not Applicable	
yhbB	MAD	ID14-EH4	Apr 2002	512	152.4,82.6,93.3	P21212	0.2x0.2x0.1	Se	3.0	10	More Phasing Needed	Not Applicable	
yggV-dGDP	ligand	ID14-EH4	Apr 2002	197	79.3,79.3,78.9	P43212	0.4x0.4x0.4	none	2.3	5.4	More Phasing Needed	Not Applicable	
yecD	native	ID14-EH1	June 2002	188	108.9,140.50,4	P21212	0.2x0.4x0.4	none	1.3	5.8	Completed	In Preparation	
aphE	unusable	ID29	Sep 2002	335	57.7,86.76,72.69	P1	0.2x0.1x0.1	Se	3.8		Poor Data	Not Applicable	
ydhF-NADP	ligand	ID14-EH2	Dec 2002	298	174.5,174.5,98.90,90.12 0	R3	0.1x0.1x0.1	None	2.9	10	Under Refinement	Not Applicable	
ycgM	unusable	ID14-EH2	Dec 2002	219	77.8,77.8,88.3	P4	0.1x0.3x0.3	None	2.9		Poor Data	Not Applicable	