EUROPEAN **S**YNCHROTRON **R**ADIATION **F**ACILITY

INSTALLATION EUROPEENNE DE RAYONNEMENT SYNCHROTRON

ESRF

Experiment Report Form

The double page inside this form is to be filled in by all users or groups of users who have had access to beam time for measurements at the ESRF.

Once completed, the report should be submitted electronically to the User Office using the **Electronic Report Submission Application:**

http://193.49.43.2:8080/smis/servlet/UserUtils?start

Reports supporting requests for additional beam time

Reports can now be submitted independently of new proposals – it is necessary simply to indicate the number of the report(s) supporting a new proposal on the proposal form.

The Review Committees reserve the right to reject new proposals from groups who have not reported on the use of beam time allocated previously.

Reports on experiments relating to long term projects

Proposers awarded beam time for a long term project are required to submit an interim report at the end of each year, irrespective of the number of shifts of beam time they have used.

Published papers

All users must give proper credit to ESRF staff members and proper mention to ESRF facilities which were essential for the results described in any ensuing publication. Further, they are obliged to send to the Joint ESRF/ ILL library the complete reference and the abstract of all papers appearing in print, and resulting from the use of the ESRF.

Should you wish to make more general comments on the experiment, please note them on the User Evaluation Form, and send both the Report and the Evaluation Form to the User Office.

Deadlines for submission of Experimental Reports

- 1st March for experiments carried out up until June of the previous year;
- 1st September for experiments carried out up until January of the same year.

Instructions for preparing your Report

- fill in a separate form for each project or series of measurements.
- type your report, in English.
- include the reference number of the proposal to which the report refers.
- make sure that the text, tables and figures fit into the space available.
- if your work is published or is in press, you may prefer to paste in the abstract, and add full reference details. If the abstract is in a language other than English, please include an English translation.

ESRF	**Experiment title:** EMBL Grenoble Outstation BAG (Protein-RNA Group of Stephen Cusack)	**Experiment number**: LS 1814/LS 1945
Beamline: various	**Date of experiment**: from: September 2000 to: August 2001	**Date of report**: 31/8/01
Shifts:	**Local contact(s)**: various	*Received at ESRF:*

Names and affiliations of applicants (* indicates experimentalists):

Dr. Stephen Cusack, EMBL Grenoble Outstation

Dr. Catherine Mazza, Dr. Carmen Berthet, Dr. Michael Härtlein

Dr. Anya Yaremchuk, Dr. Mikhail Tukalo, Dr. Wim Burmeister, Mr. Paul Backe,

Dr. Klemens Wild.

Report:

(1) Human nuclear cap-binding complex.

The hetero-dimeric nuclear cap-binding complex (CBC) binds to 5' capped polymerase II transcripts. It enhances the efficiency of several mRNA maturation steps and is essential for U snRNA nuclear export in multicellular eukaryotes. The complex comprises a large subunit CBP80 and small subunit CPB20. The crystal structure of human CBC was determined by Se-Met MAD using data to 3Å resolution measured on FIP (BM30) from which SHELXD found 57/57 selenium sites. (Selenium incorporation was done CBP80 expressed in insect cells). Phases were transferred to native data measured to 2Å resolution on EH2 and phase extension performed using 3-fold NCS. The final map could largely be built into automatically by WarpNtrace. The structure shows that the large subunit, CBP80, comprises three domains each containing consecutive helical hairpins and resembling the so-called MIF4G domain found in several other proteins involved in RNA metabolism. The small subunit, CPB20, has an RNP fold and associates with the second and third domains of CBP80. Site-directed mutagenesis revealed four residues of CBP20 which are critical for cap binding. A model for cap binding is proposed based on these results and the known mode of binding of RNA to RNP domains.

Publication. **Crystal structure of the human nuclear cap-binding complex**. Mazza, C., Ohno, M., Segref, A., Mattaj, I. and Cusack, S. (2001). *Molecular Cell*, **8**, 383-396.

(2) **Human signal recognition particle**.

The signal recognition particle (SRP) is a universally conserved ribonucleoprotein complex that mediates the co-translational targeting of secretory and membrane proteins to cellular membranes. The *Alu*-domain of the mammalian signal recognition particle (SRP), consisting of the proteins SRP9/14 bound to the 5' and 3' terminal sequences of SRP RNA, is responsible for retarding ribosomal elongation of signal sequence containing proteins prior to engagement with the translocation machinery in the endoplasmic reticulum. A crystal structure at 3.3Å resolution of human SRP9/14 bound to the conserved 5' domain of *Alu*-domain RNA

shows that a stable complex could be formed even before the rest of the SRP RNA is transcribed. From a second structure, with an 88 nucleotide RNA, a model for the full *Alu*-domain consistent with extensive biochemical data is deduced. A novel feature is the idea that there could be a reversible switch in the folding of the *Alu*-domain which might be of functional importance in the mechanism of elongation arrest. The Alu-domain structure is likely to be conserved in other small cytoplasmic RNPs and retroposition intermediates that contain SRP9/14 bound to RNAs transcribed from *Alu* or *Alu*-derived elements in genomic DNA. **Publication**: Weichenreider, O., Wild, K., Strub, K. and Cusack, S. (2000). **Structure and assembly of the *Alu* domain of the mammalian signal recognition particle**. *Nature*, 408, 167-173.

A crucial early step in SRP assembly in archaea and eukarya is binding of protein SRP19 to specific sites on SRP RNA; this is a pre-requisite to SRP54 binding. We have determined at 1.8Å resolution the crystal structure of human SRP19 in complex with its primary binding site on helix 6 of SRP RNA which consists of a stem-loop structure closed by an unusual GGAG tetraloop. The structure solution was done by MAD on bromine uridine-substituted RNA using data from ID29. Protein-RNA interactions are mediated by the specific recognition of a widened major groove and the tetraloop without any direct protein-base contacts, and include a complex network of highly ordered water molecules. A model of the assembly of the SRP core comprising SRP19, SRP54 and SRP RNA based on crystallographic and biochemical data is proposed. **Publication**: Wild, K., Sinning, I. and Cusack, S. **Crystal structure of an early protein-RNA assembly complex of the human signal recognition particle**. *Science*. In press. (Collaboration with Irmi Sinning group, Univ. Heidelberg).

(3) ***T. thermophilus*** **tyrosyl-tRNA synthetase** complexed with tRNAtyr(GUA).

Bacterial tyrosyl-tRNA synthetases (TyrRS), unlike those of archae and eukaryotes, possess a flexibly linked C-terminal domain of about 80 residues which has hitherto been disordered in crystal structures of the enzyme. Prokaryotic $tRNA^{tyr}$ has a long variable arm which distinguishes it from all tRNAs except $tRNA^{ser}$ and $tRNA^{leu}$. We have determined the structure of *Thermus thermophilus* TyrRS at 2.0Å resolution in a new $P2_12_12_1$ crystal form in which the C-terminal domain is ordered and find that the fold is very similar to the C-terminal domain of ribosomal protein S4. Several crystal forms of the complex of *T. thermophilus* TyrRS and $tRNA^{tyr}$ have been characterised, the best being a trigonal form diffracting to 2.8Å resolution. In this structure the C-terminal domain is observed to bind to the long anti-codon stem and long variable arm of the tRNA. The orientation of the long variable is of $tRNA^{tyr}$ is distinct to that of $tRNA^{ser}$ (which helps the respective synthetases distinguish between the two) and this can now be understood in terms of the respective core structures of the tRNAs. Specific recognition of Gua-34 and Uri-35 anti-codon bases is observed. Furthermore the $tRNA^{tyr}$ binds across the two subunits of the dimeric enzyme and remarkably the mode of recognition of the class I TyrRS for its cognate tRNA resembles that of a class II synthetase in being from the major groove side of the acceptor stem.

Publication: Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. Yaremchuk, A., Kriklivyi, I., Tukalo, M. and Cusack, S. (submitted).

(4) Structure of the fibre head of adenovirus strain Ad3.

Adenoviruses of serotype Ad3 (subgenus B) use a still unknown host cell receptor for viral attachment whereas viruses from all other known subgenera use the Coxsackie and Adenovirus receptor (CAR). The receptor binding domain (head) of Ad3 fibre protein has been expressed in *Echerichia coli* inclusion bodies. After denaturation and renaturation using a rapid dilution method, crystals of trimeric head were obtained. The 1.6 Å resolution X-ray structure shows a strict conservation of the beta-sheet scaffold of the protein very similar to the head structures of the CAR-binding serotypes Ad2, Ad5 and Ad12. The conformation of the loops is different, with the exception of the AB loop which forms the centre of the interface in the Ad12-CAR complex structure. The structure explains why a mutation in Ad5 of one residue in the AB loop to glutamic acid, as in Ad3, abrogates binding to CAR. It is possible that the Ad3 receptor binding site is nevertheless similarly situated to the CAR binding site. The structure has been solved by molecular replacement using data collected at beam lines ID14-1.

Publication: Structure of the adenovirus fibre receptor domain of Ad3, a non CAR-binding serotype Durmort, C., Mitraki, A., Stehlin, C., Schoehn, G., Drouet, E., Cusack, C. & Burmeister, W. P. (2001). *Virology* **285**, 302-3122.

(5) Leucyl-tRNA synthetase.

Leucyl-tRNA synthetase (LeuRS) is a large monomeric class I synthetase closely related with IleRS and ValRS. Each contains a homologous insertion domain of ~ 200 residues, which is thought to permit them to hydrolyse ('edit') cognate tRNA that has been mischarged with a chemically similar but non-cognate amino acid. We have found that *Thermus thermophilus* leucyl-tRNA synthetase (LeuRSTT) is capable of editing homocysteine, norleucine and norvaline, the latter activity occurs in tRNA-dependent manner. We have previously determined the crystal structure of LeuRSTT in complex with leucine and a leucyl-adenylate analogue at 2Å resolution giving the first view of this enzyme and its architectural similarities and differences from IleRS as well as a detailed picture of the leucine binding site and editing domain. Very recently we have determined structures of LeuRSTT at 2Å resolution, with AMP bound in the editing site (data from ID14-EH2 data) and with a norvalyl-adenylate analogue bound in both the synthetic and editing sites (ID14-EH1 data). These structures show for the first time how a non-cognate amino acyl-adenylate is bound in the editing site and suggest a mechanism for pre-transfer editing althpugh leave open the question of how the adenylate translocates between the two active sites, 33Å apart. The structure is also very consistent with rectn biochemical results which implicate Thr-252 in the editing site in permitting the non-cognate norvaline to bind but excluding the cognate leucine.

Manuscript in preparation.

(6) **Human hnRNP K**.

HnRNP K is a multifunctional, multi-domain RNA/DNA-binding and nucleocytoplasmic shuttling protein. Specific functions have been found as a cytoplasmic translational repressor (by binding to CU-rich RNA sequences in the 3' UTR of certain mRNAs) and as a nuclear transcription factor e.g. for c-myc (by binding to CT-rich DNA sequences). Nuclear acid binding ability is modulated by tyrosine phosphorylation by c-SRC kinase. The protein has three K-homology (KH) domains which bind nucleic acid (usually RNA). Our aim is to make complexes of domains of hnRNP K with interacting partners (nucleic acid or SH3 domains). We have solved the structure of the third KH domain (domain of 86 residues) at 0.9Å resolution, measured on ID14-EH2 by molecular replacement using a previously published NMR structure. Refinement is in progress.