From left to right, front row: P. Bösecke, J. Gorini, O. Diat, B. Rasmussen, J. Lescar. Back row: J.M. Bois, D. Pognant, F. Felisaz, F. Lapeyre, J. Grimes, E. Mitchell.

ID2: A Brilliant Source for Probing Large Biomolecular Assemblies

J. Lescar[1], E. Mitchell[1], J. Gorini[1], O. Diat[1], P. Bösecke[3], J. Grimes[2], J.-M. Bois[2], F. Felisaz[2], L. Claustre[2], D. Pognant[2], F. Lapeyre[2] and B. Rasmussen[2]

1 ESRF, Experiments Division
2 EMBL, Grenoble Outstation
3 Max Planck Institut, Hamburg

The cell is an amazingly complex system of interacting macromolecules. One of the major challenges in the structural molecular biology of today is to understand how complex integrated systems function, in a dynamic way, within the living cell. Instead of studying individual biomolecules as separate pieces, there is a need to address structural problems which involve assemblies composed of many biomacromolecules (generally proteins in association with lipids, sugars, or nucleic acids) and with total molecular weights superior to 100 kDa and up to tens of MDa. Surprisingly, such enormous systems can be well ordered and can give rise to near atomic resolution structures. Conversely, on a smaller scale, increasing numbers of ab initio drug design programs require the collection of very high resolution data for single proteins which are important pharmaceutical targets.

Large bio-assemblies tend in general to give only tiny crystals with volumes less than 10^5 mm^3, even after years of crystallisation trials. Such crystals diffract very weakly and adequate data cannot be collected using conventional X-ray sources. This has hitherto prevented structure determination of important biological systems like membrane proteins.

Furthermore large macromolecular assemblies yield crystals with large unit cells (some virus crystals have unit-cell parameters up to 1500 Å). These sample peculiarities require both a high photon flux through a small cross-section and also an essentially parallel beam. With such characteristics, the many closely-spaced spots from a large unit-cell crystal can be resolved and a good signal-to-noise for weak reflections can be obtained.

The High Brilliance beamline ID2, with a highly collimated undulator source, ideally fulfils the requirements for crystallography of biological complexes. ID2 provides one of the most brilliant X-ray source for Macromolecular Crystallography in the world and was one of the first ESRF beamlines to become operational in September 1994. The beamline is devoted to macromolecular crystallography experiments for half of the usable time, the other half being devoted to small-angle scattering (including on biological samples such as muscle fibres) on the second end-station. ID2 is equipped with two undulators (46 mm and 26 mm) which produce a small (100 x 100 μm) parallel beam with a flux of about 6 x 10^{12} ph/s/100 mA through the sample.

In addition, two cryocooling devices (Oxford Instruments Cryostream and a coaxial FTS system) allow regulation of the sample temperature between 100 K and 293 K. Data collection at low temperatures, preferably at 100 K if the crystal can be frozen, is essential to avoid radiation damage and obtain more accurate measurements. A five-circle Huber diffractometer with a long 2θ arm allows positioning of the detector up to 1m away from the crystal. This allows diffraction spots from crystals with very large unit cells (up to 1500 Å for crystals containing whole viral particles) or large mosaicity (which was the case for the membrane protein Photosystem I) to be resolved. In addition, the 2θ arm can also be swung (± 20 degrees) allowing high resolution data to be collected even at long crystal-to-detector distances.

Table 1

Viruses	
• *Small RNA bacteriophages*	*Liljas et al. (1996) Structure 4 543-554*
• *Structure of Blue Tongue Virus*	*Stuart et al. (1997) In preparation*
Drug design	
• *HIV-I Reverse Transcriptase*	*Stuart et al. (1995) Structure 3 915-926*
• *Isopenicillin Synthase*	*Hajdu et al. (1995) Nature 375 700-704*
• *HIV Protease*	*Lescar et al. (1997) J. Mol. Biol. In the press (collaboration with Institut Pasteur Paris)*
Membrane proteins	
• *Structure of photosystem I*	*Krauss et al. (1997) Nature Struct. Biol. 3 965-973*
• *High resolution studies of the light-harvesting complex from photosynthetic bacteria*	*Freer et al. (1997) In preparation*
• *Structure of cytochrome c oxidase at 2.8 Å resolution*	*Iwata et al. (1995) Nature 376, 660-669*
Nucleic acid-protein interactions	
• *Crystallographic studies on ribosomes*	*Yonath et al. (1996) Structure 4 513-518*
• *Aminoacyl-tRNA synthethases and their tRNA complexes*	*Cusack et al. (1996) EMBO journal 15 6351-6334*

A number of structural biology projects, including complete virus particles, targets for drug-design, membrane proteins and nucleic acid-protein complexes have made use of the beamline (some of them are listed in Table 1). For many systems ID2 has provided the experimenters with data to higher resolution than previously achieved, allowing a more detailed understanding of fundamental biological phenomena, for example viral replication or plant photosynthesis. Of the many demanding projects that were successfully carried out over the past two and a half years of operation, we have selected four examples to illustrate how a high brilliance X-ray source can be utilised to probe and understand large biological assemblies.

VIRUSES; BLUE TONGUE VIRUS

Understanding viruses that infect cattle and how to combat them

Blue Tongue Virus (BTV), an orbivirus, is a member of the Reoviridae family and infects ruminants and domestic cattle, causing diseases of great economic importance. The Blue Tongue Virus project has been described in some detail in the ESRF Highlights 1995/1996. We would like to summarise the overall status of this project. The virus particle is very large in structural terms (800 Å in diameter and with a molecular weight of 60 MDa) and is composed of 4 capsid proteins (with differing copy numbers ranging from 120 to 780) and 10 unique double-strand ribonucleic acids, each associated with its transcription complex (made up of 3 or 4 proteins). Core particles, with the two outer capsid proteins removed, have been crystallised by D. Suart's group at Oxford and diffraction data have been collected at ID2 on several occasions. These data to 3.5 Å resolution could not have been collected at any other beamline. The Blue Tongue Virus project is probably the largest biological assembly giving well-diffracting and interpretable diffraction data and a near atomic resolution structure (3.5 Å).

HIV1-REVERSE-TRANSCRIPTASE (RT)

Improved drug design to prevent AIDS?

The genome of HIV (the causative agent of AIDS) encodes three enzymes, each with an essential function in the infectious cycle of the virus. One of these enzymes, the reverse transcriptase or RT, is responsible for the conversion (transcription) of the viral genome into DNA molecules which are then incorporated into the host-cell genome. As a crucial element in the infectious cycle, this enzyme has been the primary target for antiviral therapy, and analogues of the building blocks of DNA such as AZT or DDI have been used as tentative AIDS therapeutic agents because they bind to RT and prevent its function. Several mutations however have been observed within the RT molecule, which enable it to escape the binding of these drugs. Detailed structural information of mutated forms of the RT molecule are needed in order to understand the way these mutations allow the RT to escape binding of existing drugs and to help design others. Since RT is a large enzyme containing several subdomains linked by flexible segments, crystals have been extremely difficult to produce and usually diffract poorly.

The technical characteristics of ID2 have helped the Oxford group to collect higher resolution data on these projects and produce a clearer understanding of this enzyme.

MEMBRANE PROTEINS; PHOTOSYSTEM I

Understanding plant photosynthesis

Photosynthesis is the process whereby solar light energy is converted into a chemical form readily available to living systems. Two groups of plant photosystems (numbered PSI and PSII according to the nature of their terminal electron acceptor: either an iron sulphur cluster or a quinone molecule) have been extensively studied through biochemical, spectroscopic and structural methods. The large chlorophyl antenna of PSI capture and channel excitation energy to the primary electron donors (chlorophyl molecules) located near the light-sensitive side of the membrane, which further transfer the electron to Fe_4S_4 clusters. On the basis of data collected on ID2 to 3.8 Å resolution, which have been recently extended to 3.5 Å resolution, the geometry of the cofactors, which are important structural elements within the electron transfer chain, could be imaged

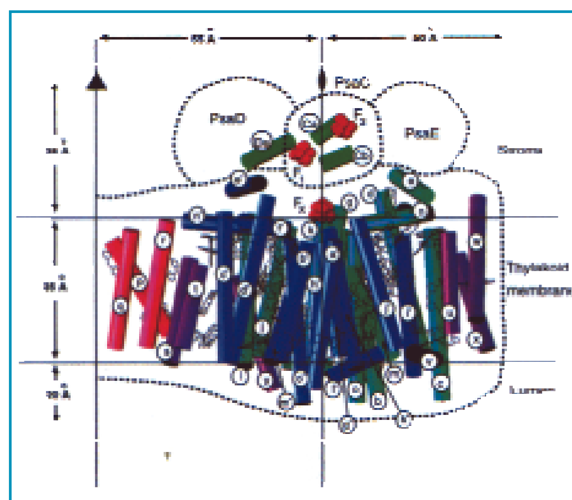

Fig. 1: (from Krauss et al., Nature Struct. Biology, 3, 965) Overall view of the photosynthetic system I molecule - Courtesy of N. Krauss, Institut für Kristallographie Berlin, Germany) α-helices are represented as tubes and labelled a*-y. Transmenbrane and surface α-helices of the peripheral subunits are purple, those involved in trimerisation red, those of the extrinsic stromal subunits are green. The porphyrin headgroups of chlorophyll* a *are indicated as wire models, atoms of iron-sulphur clusters as red spheres.*

much more precisely (Figure 1). The fine collimation of the beam provided by ID2 proved to be essential for resolving broad spots due to a large mosaicity of the crystals. This structure determination has provided a basis for understanding photosynthetic reaction centres involving iron sulphur clusters. Furthermore, this result has an interesting implication for understanding the evolution of photosynthetic systems within plants: despite a low level of primary sequence homology (the sequence of amino-acids of the two PS bear no obvious resemblance), structural analogy between PSI and PSII indicates that these two classes of photosynthetic systems may derive from a common ancestral molecule.

NUCLEIC ACID-PROTEIN COMPLEXES

Understanding the basic machinery of life: protein biosynthesis

Proteins are the «workhorses» of the living cell: they are used as enzymes to catalyse metabolic reactions, as antibodies to protect higher organisms against pathogens, as structural motors and as signal transducers between the outside of the cell and the nucleus which stores the genetic information.

The way genetic information is specifically translated into proteins depends in part upon specific interactions between nucleic acids and protein molecules. In order to elongate the growing polypeptide chain with the correct amino acid, one important step is the correct attachment of cognate amino acids onto their cognate tRNA molecules, which is catalysed by specific enzymes called tRNA synthetases. Much work has been done at the EMBL in Grenoble by the group headed by S. Cusack to understand the structural basis of these interactions. ID2 has enabled high-resolution structure determination of several aminoacyl-tRNA synthetases and their complexes with cognate tRNA (Table 1). Protein biosynthesis is carried out by numerous enzymes, tRNAs and other protein factors which act in concert with a macromolecular complex called the ribosome, where the synthesis of the growing polypeptide actually takes place. This «factory» consists of many proteins (up to 73) and several RNA molecules which ensure a coordinated translation of the genetic code. Several crystals from different ribosomal systems have been grown and are being characterised by the group of Dr Yonath at Hamburg and at the ESRF on ID2. Together, through obtaining better images of these macromolecular assemblies, this work opens the way to a better understanding of the mechanism and the interactions crucial for the protein biosynthesis machinery. ■

Reference

N. Krauss, W.-D. Schubert , O. Klukas, P. Fromme, H.T. Witt, W.. Saenger(1996) Photosystem I at 4 Å resolution represents the first structural model of a joint photosynthetic reaction centre and core antenna system. Nature Struct. Biol. 3, 965-973.