Summary pre-project for the Bio-crystallography beamlines at SOLEIL.

Prepared by G. A. Bentley and A. Lewit-Bentley in collaboration with the Working Group for Biocrystallography at SOLEIL

19 October 2001

Contents

1. Executive summary

1.1 The needs of the French biocrystallography community

In this document, we analyse the developing needs of the French biocrystallography community for synchrotron beam time, and put forward a proposal for the construction of beamlines at SOLEIL that should meet these demands. This proposal has been made after close consultation with French biocrystallographers through a Working Group of 17 members drawn from representative laboratories throughout France (see annex 1). The Working Group has, in addition, solicited advice of experts from European synchrotron facilities during their formulation of the beamline and instrument specifications described here. The request for biocrystallography beamlines at SOLEIL takes into account the present use of synchrotron sources by the French community, and the planned evolution of their scientific programmes over the next 10 years. Indeed, three national structural genomics projects, accepted for support by the French Ministry of Research and Technology, will place heavy demands on future beam-time usage. There is, in addition, a growing awareness of the economic impact that synchrotron technology can have on industry (e.g., food products, pharmacology, vaccines). It is clear that practices within the biocrystallography community are evolving rapidly because of the new opportunities offered by access to powerful third-generation sources such as the ESRF, Grenoble. These new possibilities allow, for example, the analysis of very small crystals or large macromolecular assemblies, which has hitherto been impossible. But it has become equally clear that present demands have saturated currently available sources, and that the construction of new beamlines dedicated to biocrystallography is urgent. Following an inquiry conducted in collaboration with the French biocrystallography community, the Working Group has concluded that **a minimum of five experimental stations should be constructed at SOLEIL over the next ten year period** to respond to this expanding need.

1.2 Requirements: ring access, beamlines, stations and environment

Our proposal requests that a coherent and adequate portion of the storage ring be reserved for the biocrystallography beamlines. This corresponds to a contiguous segment of the ring, comprising two medium-length straight sections (7 m), one short straight section (3 m) and a bending magnet. Securing of a contiguous segment is crucial to realising this project since the small size of the ring limits the area available in the experimental zone. By grouping the biocrystallography stations within a common perimeter, an economic use of the limited space can be achieved by sharing facilities, such as sample preparation and a computing room, between them. This configuration would accommodate the five stations that we consider essential for meeting the users' future needs.

The specifications for photon energy and flux in biocrystallography require that the ring should operate at an energy of 2.75 GeV. The characteristics of the beamlines described below assume that this key parameter will be adopted. The nature of the beamlines and experimental stations, and the order of their construction, are summarised as follows:

(1) Beamline 1 (1 station): the first station should be constructed on a 7 m straight section and should provide high flux in the 5 – 17.5 keV energy range, allowing MAD experiments to be performed

with a wide selection of anomalous scatterers. We propose that two U26 undulators, aligned in tandem, be placed in the 7 m section, giving optimum intensity over the desired energy range. We propose, further, that this beamline be arranged to accommodate two distinct distances from the source, one closer with high demagnification (8:1) to function as a micro-focus station (focal spot size of 15 μm), and the other more distant with lower demagnification (3:1) to serve for MAD and high resolution experiments (focal spot size of 50 - 70 μm). The design of this multi-purpose beamline should foresee a rapid, simple and reproducible interchange of the instrument between the two positions for efficient exploitation of the two modes of operation. A similar arrangement has been constructed at the SLS, Switzerland.

(2) Beamline 2 (1 station): the commissioning of the first station should be closely followed by a second instrument positioned on the 3 m section, with a high flux optimised for energy at 12 - 14 keV for MAD experiments with elements such as Se or Br, as well as for standard high resolution data collection. One option would be to use a U20 undulator for this beamline. The details of the beamline 2 design should be finalised after construction of beamline 1 has begun in order to profit from experience gained from the latter. Construction of the second beamline, however, should not be delayed by more than 12 months after work on the first has commenced.

(3) Beamline 3 (2 stations): the next phase of construction should take place on the second 7 m section. We propose that two canted undulators (e.g., U26) be installed here to produce two beams with a small but sufficient angular separation (2 mrad) to allow two independent stations to operate. Since building these stations will start 4 – 6 years after the first phase of construction has begun, the final choice of their characteristics should be finalised at this date. Nonetheless, they would probably be optimised to operate with high flux at about 13 KeV. The use of canted undulators is a developing technology, but we anticipate that it should be sufficiently advanced in 4 – 6 years time to be implemented without need for further development.

(4) Beamline 4 (1 station): the fifth station should operate on the bending magnet section, and would be designed for MAD experiments over a wide energy range to benefit from the properties of its beam.

All stations should be designed to maximise through-put in the diffraction experiments by introducing automation at all levels, from beam intensity optimisation and wavelength selection, through to sample manipulation and diffraction quality assessment, and finally to data collection strategy and on-line data treatment. **The design of the beamlines must take into account all requirements for their full automation. Computing resources will have to meet the demands for high data through-put and automation, and an appropriate security filter should be installed to permit control of the experiment at a distance** (e.g., home laboratory).

To accommodate changing European Union norms, **the entire experimental environment should conform to at least P2 health security specifications**. Furthermore, considering that these norms are becoming stricter with regard to pathogens, a P3 security laboratory should be equipped within close proximity to the beamlines. It may equally be necessary that one of the beamlines constructed during phase II should conform to P3 containment norms.

1.3 Personnel

SOLEIL must provide a sound career environment for its personnel to ensure that top scientists are attracted to participate in both the construction and operation phases of the biocrystallography beamlines and stations. Strong support for in-house biology projects is therefore a prerequisite to maintain a high level of scientific dynamism around the operation of these instruments. Macromolecular crystallography does not centre around the beamlines only; it also relies an autonomous scientific infrastructure of molecular biology, biochemistry and functional studies, and the personnel requirements should take these into account as well.

2. The Scientific Case

2.1 Preamble

A decisive advance in the life sciences has recently taken place with the advent of ”large scale biology”. Automation of sequencing techniques has been instrumental for decoding of the complete human genome, as well as that of many other organisms. The possibility to identify and systematically clone each and every gene of an organism, together with the determination of its state of transcription (transcriptome) and translation (proteome) in different tissues at any given time, now permits a broad study of biological function at a level that would have been inconceivable ten years ago. Knowing the presence of, and the interactions between, each one of these gene products at a given moment is essential for understanding the complex network of biological processes inside the living cell. In particular, comprehension of these phenomena at the molecular level requires detailed knowledge of the atomic structure of each component as well as that of the molecular associations in which it participates. Such knowledge allows, among other things, to envisage the development of molecules for therapeutic intervention in malfunctioning cells. The medical and pharmaceutical industries of tomorrow will depend increasingly on the availability of atomic resolution data of macromolecules participating at different stages of the cell cycle.

The construction of SOLEIL will endow France with a resource of first rank for the advancement of fundamental research in biology and its applications to the domains of medicine, pharmacology and food industry. It is thus indispensable that the number and performance of the beamlines destined for biology meet the level required to conduct biology on a large scale in the years to come.

2.2 The demand for macromolecular structural data

The importance of structural data for the community of biologists can be gauged by the quality and impact of publications relating to protein structures determined using synchrotron radiation. We have confirmed this from an inquiry conducted with the French biocrystallography community: a total of 313 publications produced over the past three years (1998 – 2000) by 25 French laboratories had made direct use of synchrotron radiation. Moreover, their mean impact factor was 7.7, and for 17 of these publications, the impact factor was over 25 (Cell, Nature, Science).

The French biocrystallography community has contributed to a wide range of biological themes through their use of synchrotron radiation. These areas cover cell signalling (Cherfils, *et al*., 1998; Maignan *et al*., 1995; Welch *et al*., 1998), regulation of nuclear receptors (Renaud et al., 1995), membrane fusion (Réty *et al*., 1999), cytoskeletal dynamics (Gigant *et al*., 2000), protein synthesis (Sankaranarayanan *et al*., 1999; Dock-Bregeon *et al*., 2000; Weichenrieder *et al*., 2000), virology (Bizebard *et al*., 1995, Lescar *et al*., 2001), immunology (Bentley *et al*., 1995, Housset *et al*., 1997), photosynthesis (Royant *et al*., 2000), prions (Bousset *et al*., 2001) and metabolism (van Tilbeurgh *et al*., 1993), to cite just a few domains where major contributions have been made. Nonetheless, the tasks that remain to be accomplished are immense. An analysis of the yeast genome, for example, shows that only one third of potentially encoded proteins have a function directly identifiable from their sequence. Today, the palette of structures that can be explored is no longer limited by the size of the protein or macromolecular assembly: the recent success of the analysis of the ribosomal subunits

(Ban *et al*., 2000; Schluenzen *et al*., 2000; Wimberly *et al*., 2000) or RNA polymerase II (Cramer *et al*., 2000; Cramer *et al*., 2001) shows that once crystals of a protein have been obtained that diffract beyond 3.5 Å resolution, its structure can be determined even if it requires pushing crystallographic methodologies to their current limits.

The importance of structural information on proteins comes not only from the quest for a more complete knowledge of life processes, but also from the need, for example, to develop more effective medicaments. Structure determination has become a necessary step in this development: once a lead compound has been identified from the screening of chemical product libraries, it is optimised by cycles of chemical synthesis alternated with structure determination of complexes of targeted proteins with the improved drug. Among the most recent examples of this kind of application are the development if HIV protease inhibitors and anti-inflammatory non-steroid inhibitors of prostaglandin synthase.

One significant factor contributing to the growing impact of structural analysis in pharmaceutical and medical research is the accelerated rhythm of structure determination. Today, the determination of the structure of the protein of therapeutic interest complexed, for example, to a new drug can take less than 2 weeks from initial crystallisation trials to the analysis of the structural results. This stage of the process thus falls in phase with the other steps in drug optimisation, notably that of chemical synthesis. Moreover, it is no longer uncommon that the determination of a completely new structure be achieved within a few weeks, corresponding to dead-lines typically applied in industrial research projects.

The following factors must be taken into account in the growing demand for protein structure determination:

- The standardisation of crystallographic methodologies has placed structure determination more directly into the hands of the non-specialist, without these procedures necessarily becoming his principal activity. Indeed, biologists are integrating structural information more and more into their research and establishing an expanding network of collaborations with structural biologists. This is creating a major increase in the demand for structure determination, which must be resolved by the availability of beamlines in sufficient number, in combination with automations and the standardisation of crystallographic methodologies.
- The efficacy of recombinant technology for protein production in quantities sufficient for structural studies has improved considerably. These improvements now allow a systematic structural survey of all gene products from a wide variety of organisms, both prokaryotes and eukaryotes. As noted above, recent advances in crystallographic methodologies are pushing the limit on protein size for structure determination to a higher and higher molecular weight. The number of proteins that can, in practice, be studied by X-ray crystallography will correspondingly increase.
- The growing demand for structural results by industrial research in medicine, pharmacy and food products must be added to these factors. Such activities in these sectors will certainly increase and spread. Many laboratories in the French community have collaborations with industry.

- The launching of systematic structural analyses of the ensemble of proteins expressed by given organism, or of proteins having the same function in a series of different organisms, is currently under way in several concerted national and international programmes. In France, 11 laboratories are implicated in three structural genomics projects. This type of systematic effort will eventually extend to other, more technically difficult, classes of protein, such as membrane proteins, as experience in structural genomics technology accumulates.

Three national structural genomics projects have been financed by the French Ministry of Research and Technology since last year, one centred on a eukaryotic genome (human nuclear receptors), coordinated by D. Moras, IBGMC, Strasbourg, one on a prokaryotic genome (*Mycobacterium tuberculosis*), coordinated by P. Alzari, Institut Pasteur, Paris, and one on a lower eukaryotic genome (*Saccharomyces cerevisiae*) coordinated by J. Janin and R. Fourme, Orsay. The three projects are now under way and will expand following integration within an European initiative on structural genomics, accepted by the EU in 2001. Within the next few years, each of these projects will be producing tens of structures to be analysed using synchrotron radiation (as an example, 3 proteins have yielded good quality crystals within the past 3 months for the yeast project alone).

2.3 Beamline needs of the French biocrystallography community

Today, the beamlines available to French macromolecular crystallographers on third generation synchrotron sources are saturated, in particular those with the capacity to measure diffraction data over a range of different wavelengths that permit MAD experiments. The inquiry conducted by the Working Group revealed that current annual beam-time usage by French biocrystallographers is equivalent to 240 days on an ID14-like station at the ESRF; this corresponds to a full-time use of two ID14 stations. (A total of 120 days per year is available to the user community on each ID14 station.) **The present demand for beam time is created by 30 laboratories, including two from the private sector, with more than 450 personnel working directly on the different steps of macromolecular structure determination.** It is certain that the number of both scientists and laboratories will increase for the reasons elaborated above.

From the Working Group's inquiry, a clear consensus emerged for **a minimum of two biocrystallography stations to be in operation when the first beamlines of SOLEIL become available to users**. These should include one station operating over a wide wavelength range for MAD and SAD experiments using a broad selection of anomalous scatterers, and a second station with a high flux optimised in the energy range of 12 – 14 keV for both specialised MAD (Se, Br) and high resolution diffraction experiments. A strong emphasis was placed on the need to accommodate small crystals (10 – 50 μm), placing stringent requirements on the flux, beam properties and diffractometer design. A desire for automation at all levels was expressed, from flux and wavelength optimisation, sample manipulation, through to data collection and treatment. The importance of adequate support facilities, such as a small laboratory and a cold-room, at or immediately adjacent to the experimental area, was underlined. Users also requested that technical personnel dedicated specifically to the biocrystallography should be engaged. Over a ten-year term, the number of stations

available to the biocrystallography community should increase to five. These additional stations will need to cater for high through-put structure determination as structural genomics projects become fully operational and automation procedures upstream from the synchrotron phase take effect. Finally, there was strong insistence that the ring operate at 2.75 GeV, as the loss of brilliance by an order of magnitude at 2.5 GeV would compromise all wishes and expectations of a third generation synchrotron source.

2.4 References

Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905-20.

Bentley, G.A., Boulot, Karjallainen, K. and Mariuzza, R.A. (1995) Crystal structure of the β chain of a T cell antigen receptor. Science, 267, 1984-1987.

Bousset L, Belrhali H, Janin J, Melki R, Morera S. (2001) Structure of the Globular Region of the Prion Protein Ure2 from the Yeast *Saccharomyces cerevisiae*. *Structure*, **10**, 39-46.

Bizebard, T., Gigant, B., Rigolet, P., Rasmussen, B., Diat, O., Bosecke, P., Wharton, S.A., Skehel, J.J. and Knossow, M. (1995) Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. *Nature*, **376**, 92-4.

Cherfils, J., Menetrey, J., Mathieu, M., Le Bras, G., Robineau, S., Beraud-Dufour, S., Antonny, B. and Chardin, P. (1998) Structure of the Sec7 domain of the Arf exchange factor ARNO. *Nature*, **392**, 101-5.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R. and Kornberg, R.D. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science*, **288**, 640-9.

Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001) Structural basis of transcription: rna polymerase ii at 2.8 angstrom resolution. *Science*, **292**, 1863-76.

Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C.S., Ehresmann, C. and Moras, D. (2000) Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell*, **103**, 877-84.

Gigant, B., Curmi, P.A., Martin-Barbey, C., Charbaut, E., Lachkar, S., Lebeau, L., Siavoshian, S., Sobel, A. and Knossow, M. (2000) The 4 A X-ray structure of a tubulin:stathmin-like domain complex. *Cell*, **102**, 809-16.

Housset, D., Mazza, G., Grégoire, C., Piras, C., Malissen, B. and Fontecilla-Camps, J.C. (1997) The three-dimensional structure of a T-cell antigen receptor VαVβ heterodimer reveals a novel arrangement of the Vβ domain. EMBO J., 16, 4205-4216.

Lescar, J., Roussel, A., Wien, M.W., Navaza, J., Fuller, S.D., Wengler, G. and Rey, F.A. (2001) The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell*, **105**, 137-48.

Maignan, S., Guilloteau, J.P., Fromage, N., Arnoux, B., Becquart, J. and Ducruix, A. (1995) Crystal structure of the mammalian Grb2 adaptor. *Science*, **268**, 291-3.

Renaud, J.P., Rochel, N., Ruff, M. Vivat, V., Chambon, P., Gronemeyer, H. and Moras, D. (1995) Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature*, **378**, 681-689.

Réty S., Sopkova J. Renouard M., Osterloh D., Gerke V., Tabaries, S., Russo-Marie F. and Lewit-Bentley, A. (1999) Crystal structure of a complex of p11 with the annexin II N-terminal peptide. *Nature Struct. Biol*. **6**, 89-95.

Royant, A., Edman, K., Ursby, T., Pebay-Peyroula, E., Landau, E.M. and Neutze, R. (2000) Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin. *Nature,* **406**, 645-648.

Sankaranarayanan, R., Dock-Bregeon, A.C., Romby, P., Caillet, J., Springer, M., Rees, B., Ehresmann, C., Ehresmann, B. and Moras, D. (1999) The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site. *Cell*, **97**, 371-81.

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*, **102**, 615-23.

van Tilbeurgh, H., Egloff, M.P., Martinez, C., Rugani, N., Verger, R. and Cambillau, C. (1993) Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography. *Nature*, **362**, 814-20.

Walsh, M.A., Dementieva, I., Evans, G., Sanishvili, R. and Joachimiak, A. (1999) Taking MAD to the extreme: ultrafast protein structure determination. *Acta Crystallogr D Biol Crystallogr*, **55**, 1168-73.

Welch, M., Chinardet, N., Mourey, L., Birck, C. and Samama, J.P. (1998) Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. *Nat. Struct. Biol.* **5**, 25-29.

Weichenrierder, 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.

Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327-39.

3. Technical case

3.1 Overview

The estimated needs of the French biocrystallography community correspond to 5 stations, which should be built progressively over the commissioning phases I and II of SOLEIL. Given this fact, we propose to group all biocrystallography stations within one single zone of the experimental hall, within a contiguous ring segment comprising two 7 m straight sections, one 3 m straight section and one bending magnet exit (i.e. 4 exits out of 44, or 1/11th of the storage ring), to allow a rational approach to the design and exploitation of the experimental area. By re-grouping several ancillary facilities in spaces common to all 5 beamlines, we expect to make significant savings in space and cost of the entire operation.

We request that the biocrystallography beamlines lie close to the annex that houses the biochemistry/biology support laboratory.

Figure 1 (opposite page) illustrates the proposed layout of the biocrystallography beamlines, as well as the common areas.

The 5 beamlines should be built progressively, with each successive beamline design benefiting from the development and experience gained from the preceding installation. The Working Committee has agreed upon the following order:

1) 7 m straight section (1 station): a polyvalent, high performance MAD/SAD station, covering a wide energy range (5 – 17.5 keV) catering for a wide range of anomalous scatterers will be the first priority.
2) 3 m straight section (1 station): a station offering high flux with a limited wavelength range for specialised MAD and high resolution measurements (12 – 14 KeV around the Se absorption edge). The design of this beamline should be finalised no later than 12 months after work on the first beamline has begun in order to ensure that two biocrystallography stations will be operational when the synchrotron is first open to users.
3) 7 m straight section (2 stations): in the next stage, two stations should be built on the remaining medium straight section by the use of canted undulators. The precise characteristics should be defined later, after consultation of the community.
4) bending magnet source (1 station): this beamline should be equipped last and will be designed to exploit a large energy range.

Note: The construction of a microfocus beamline could be an option for phase II construction.

Figure 1: Layout of the biocrystallography beamlines. The ensemble of two 7 m straight sections (numbered 4 and 7), the 3 m straight section (5) and bending magnet (6), form a contiguous segment of 1/11th of the synchrotron ring. Construction should begin on a 7 m sections (4) to give a MAD station placed at 40 m from the source with a large energy range, with the possibility to reproducibly interchange at the 30 m position in order to function as a microfocus diffractometer. The second beamline should be constructed on 3 m section (5), and should be optimised to function in the 12-14 keV energy range. For the second 7 m straight section (7), we propose that two canted undulators be installed to allow the operation of two independent stations. The final section to be exploited should be the bending magnet (6).

3.2 Technical objectives for the first MAD beamline

In the technical specifications that follow we shall discuss the first beamline only.

The general requirements include robustness, as well as high reliability, automation and user-friendly operation of the instrument.

The following general objectives should be met:

- high flux on the sample (10^{12}-10^{13} photons/s)
- smooth and rapid tunability between 5 and 17.5 keV
- resolution in energy of $\leq 10^{-4}$
- focal spot of 50 - 70 μm in both directions for most samples
- possibility of focussing down to smaller samples
- high performance detector – fast (< sec cycle), low background, high resolution
- cryogenic sample environment
- full automation of the instrument and of data collection
- standard and automated data processing software
- appropriate ancillary equipment and facilities

3.3 Technical solutions

The technical choices that follow have been selected as viable possibilities on the basis of present day know-how. The final choice should be made by the project leader and his team when they have been found. We wish to point out that all the beamline elements we propose here are at the forefront of present technical developments, extremely promising but as yet not thoroughly tested.

We wish to stress that the major development effort on the SOLEIL beamlines should be directed at full and comprehensive automation, in order to achieve a complete and fully operational expert system, including remote control (from outside the facility). This system must, furthermore, be very robust, extremely reliable, very fast and totally user-friendly.

3.3.1 *Source*

The first instrument should be installed on a straight section of 7 m. Of the undulators available to date, the most adequate seems to be a combination of two U26-type undulators in tandem, giving an adequate flux over the desired energy range for the 2.75 GeV operation mode.

In order to obtain the desired beam size the sample, we require a source shape that is as symmetric as possible. Assuming a focal point with 3:1 demagnification, the FWHM size of the source should be 180-200 μm in the horizontal direction. In the vertical direction the FWHM source size should not be less than 50 μm.

The beamline optical components and beam must be kept extremely stable in order to work with focal spot sizes below 100 μm, and an overall beam positional stability better than 5 μm will be required in the vertical plane at the sample position. In the horizontal plane, a stability better than 10 μm would be desirable. It should also be noted that good reproducibility of the beam position between fills, particularly in the vertical direction, is necessary for MAD beamlines, where the energy is selected to a

high precision, and should be maintained during the experiment. A similar level of reproducibility of beam position and angle is therefore required between refills.
Good diagnostic information on the position and angle of the electron beam in the storage ring and of the X-ray beam in the beamline will be required. This information will allow the beamline operators to compensate for any systematic long-term drift in the overall beam position and keep the beamline well aligned and calibrated at any time. The entire beamline has to be under automatic control both for wavelength tuning and for alignment/beam optimisation. It is furthermore essential that the beamline operator should be able to control the undulator gaps.

3.3.2 *Optics*

Several options are possible:

- Si[111] double crystal monochromator - sagittal focussing second crystal (3:1 demagnification position) – vertical focussing mirror
- Si[111] channel cut monochromator - toroidal mirror at a 3:1 demagnification position
- Si[111] channel cut monochromator - vertical focussing mirror - horizontal focussing mirror (3:1 demagnification).

We propose the solution consisting of a vertically focussing mirror and a sagittally focussing double crystal monochromator. The system developed for the SLS beamline 6S appears perfectly suited to our needs:

- a Rh-coated mirror for vertical focussing,
- liquid N_2-cooled fixed-exit Si(111) monochromator with sagittal focussing,
- both mounted on dynamical bending devices developed by the SLS and ESRF.

Secondary slits can be used to remove scattered radiation.

As mentioned above, in order to use the most favourable 3:1 demagnification ratio, we require more adequate source dimensions. If these requirements are satisfied, we can also adopt the SLS – Beamline 6S option of an 8:1 demagnification position for the sample. The dynamic monochromator benders allow focussing of the monochromator elements in the sagital and vertical directions.
The entire system of mirror/monochromator/slits must be under computer control and integrated into the automatic beamline control.

3.3.3 *Diffractometer*

The device we propose here is the EMBL – ESRF micro-diffractometer, which combines the following characteristics:

- Crystals from 5 to 200 μm size can be remotely mounted, centred on the rotation axis and aligned with the beam.
- The phi-axis diffractometer combines:

- Sphere of confusion radius at sample position: 1 µm
- Precision: ±1 mdeg @ 10 deg/s
- Maximum rotation speed: 180 deg/s
- Beam definition aperture from 10 to 200 µm

- The video-microscope is used to view the sample and the beam display scintillator. Because the camera looks in the direction of the beam, the sample can be aligned with the beam without parallax error. It combines:
  - Sub-micron resolution, diffraction limited objective lens
  - High resolution 3CCD colour camera
  - Motorised zoom (X1 to X10; field 1.6 x 2.1 mm down to 0,16 x 0,21 mm)
  - Condenser lighting with polarizer
  - Motorized analyser
- Beam visualisation device is incorporated.

In order to satisfy the need to measure very small samples (around 15 µm) already at the first beamline, we should adopt a solution similar to that of beamline 6S at the SLS: the microdiffractometer can be mounted on a stable optical bench to allow displacement between two well defined positions:

1) microfocus position close to the monochromator (demagnification 8:1) which will allow focussing down to 15 µm on the sample
2) standard position (demagnification 3:1) for measurements with a focal spot of 60 µm, and low divergence.

This design can be achieved with the monochromator at 30 m, first sample position at 33.75 m and the second at 40 m.

The diffractometer will have to be integrated within the automated instrument control software. Automatic and rapid alignment of the diffractometer at the two positions must be ensured.

3.3.4 *Ancillary equipment for the diffractometer*

- Cryo-cooling of the sample. The Oxford Cryo-Stream is compatible with the microdiffractometer and very reliable in continuous use. It has to be connected to a continuous source of dry liquid nitrogen.
- A fluorescence detector mounted at 90° for XANES scans for MAD experiments.
- Automatic cryo-sample changer: an automatic crystal-mounting device is currently under development at the EMBL, Grenoble. It has been designed to be compatible with the micro-diffractometer and should be integrated within the automated instrument control software.
- General beamline elements, such as vacuum pumps and monitors, X-ray windows, slits, filters and monitors.

3.3.5 *Detector*

The present choice is the MarResearch Flat-panel Detector with

- a continuous active area of 420 x 350 mm^2
- pixel size of 140 μm^2
- point-spread function of less than 1 pixel
- very low noise
- fast read-out system (~ 1 s)

3.3.6 *Automation*

None of the operations carried out during a biological macromolecule diffraction data collection are beyond the possibilities of an automat and a good expert system. Indeed, full automation has become an absolute necessity with the development of structural genomics. The ultimate goal of building a third generation beamline is therefore:

- to provide an automatically controlled X-ray beam, under complete control by the user, without constant involvement of beamline staff,
- to implement automation of crystal mounting, data collection and data treatment in order to replace what constitutes nowadays largely repetitive tasks. All operations should allow remote monitoring by Telnet or Internet (installation of webcameras). This will require particular attention to network security.

All these requirements involve already existing know-how.

Beamline alignment automation

Complete automation of wavelength setting and beamline alignment requires the following:

- all operations required for wavelength setting and beamline alignment are achieved automatically by software, without the need of any human operation. This includes the adjustment of the undulator gap at the appropriate value, and the adjustment of the optical elements of the beamline, namely the double crystal monochromator and the focusing mirror. The user has simply to select the wavelength that he wishes to work with, and to start the automatic adjustment procedure.
- all automatic operations are ensured by the programme itself. This includes the possibility of automatic restart after any kind of functional breakdown.
- all alignment operations are fast and run as parallel processes.

As far as speed is concerned, the electronics and software should be chosen carefully in order to avoid severe slowing down of the system which might occur at runtime, because of too many software operations and network overload. Intensity maximisation for beam alignment should be carried out by electronic regulators rather than by rocking curve recording followed by numerical data treatment of the curve.

The beamline optics elements will have to be designed specifically, having the following requirements in mind :

- all adjustment actions will be motorised and encoded,
- all movements of beamline optics and monitors must be fast,
- a number of beam monitors (counters, beam position monitors or visualisation screens) should be installed, at appropriate positions never to loose track of the X ray beam,
- beam monitors with large dynamics will have to be selected, combined with monitored intensity attenuators.

Data collection automation

The goal of data-collection automation consists in programming an automatic run of multiple data collection series, i.e. series of scans on a series of crystals.

This sequence requires:

- automatic mounting of crystals on the diffractometer, from a cryogenic storage tank containing a series of frozen crystals
- automatic crystal centering on the centre of the diffractometer (this requires automatic adjustment of good crystal viewing conditions: illumination, magnification, contrast, followed by image analysis)
- short preliminary data collection and data treatment for crystal quality assessment
- optimal data collection strategy determination
- and finally, data reduction and structure resolution.

Expert system

The ultimate goal for automation at a macromolecular crystallography beam-line is the automation of the entire process, including wavelength selection, beam-line optimisation at the selected wavelength, crystal mounting, selection, data collection and data reduction. The expert system will take the necessary decisions at each step of this process.

3.3.7 *Specifications for computing*

Diffraction experiments on biological macromolecules at synchrotron sources generate large volumes of data that must be collected, treated and archived in a relatively short time. The requirements can therefore be expressed in terms of data storage space, rate of data transfer, CPU power and data backup capacity. At present, users collect their diffraction data themselves at the beamline. Since the high flux of X-rays permits relatively short collection times, the rotation of experiments on a given station is rapid and the through-put of users is important. This imposes a supplementary constraint on the choice of computing. It is important that the systems chosen be simple to operate, the ideal situation being where users encounter an environment equivalent to that of their home laboratory. **The choice should therefore correspond to standards used by the biocrystallography community.**

Transfer and safeguard of data

Detectors used for data collection at synchrotron sources must respond to a number of exacting specifications, some of which are critical for the optimal choice in computing. These parameters include the dimensions of the active surface of the detector, the pixel size and read-out time. For example, the solid-state detector recently developed by MarResearch produces an image of 16 Mb, with a readout time of 1-2 seconds. Detector images are initially stored on the disk of the workstation controlling the experiment, and then transferred to an independent server for storage and backup. For optimal functioning, all transfers, from detector to workstation to server, must be made as rapidly as possible. Here, the essential parameters are the speed and the availability of the network. For the first two stations proposed for construction, the internal network should be able to handle a peak data flow of 30 Mb/sec. The data storage capacity will need to accommodate 2-4 Gb per data set, and 100 Gb per day for each station. Thus, 30 SCSI disks of capacity 100 Gb would permit an autonomy of 15 days data storage when the diffraction stations are under optimal operation.

Data treatment

Given the rapid evolution in computing, in particular that of computing performance with respect to cost, the CPU power should not in itself pose any severe problems. By contrast, the data flow to the computer performing the data treatment will require careful attention. Here again, the organisation of the network around the storage server used in common by all biocrystallography stations would seem to be an adequate solution. The choice of computers should take into account the current programmes commonly used for data treatment since certain among them are not exportable to all machines.

Proposal

In figure 2, we propose a configuration that should satisfy the above requirements. At the level of each diffraction instrument, four workstations are necessary:

- A workstation to control the experiment, connected to the machine network
- A workstation to control the collection of data
- A workstation dedicated to data treatment
- A network access that would enable the user to plug in a portable computer

A computer room, situated in close proximity to the experimental stations and connected to a high-speed network, could house the storage server ('hot plug' disks) as well as workstations for calculations and graphics to enable rapid evaluation of structural results. The Internal network must possess a security system independent of that on the general network in order to allow the control of 'FEDEX-type' experiments from the outside (e.g. from the users home laboratory) aided by automation and web-cameras installed in the experimental hutch. It should allow for recovery of encrypted data at the experimenter's laboratory.

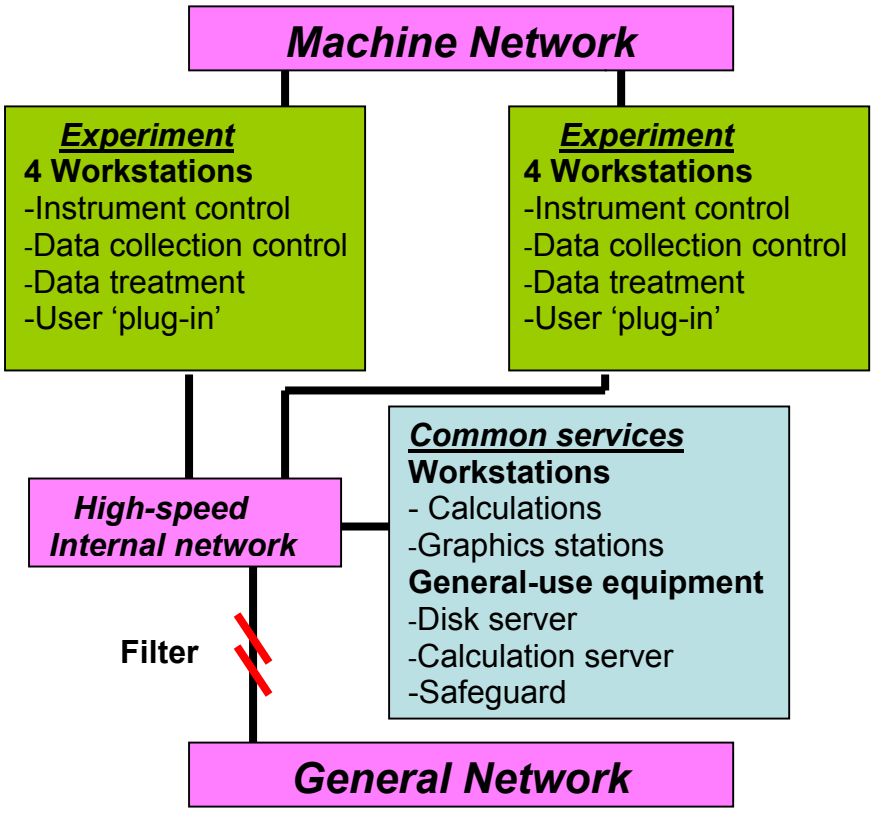

Figure 2: Proposed computing configuration for the first two experimental stations.

3.3.8 *General layout of the experimental area*

Each experimental station will have its own optics, experiment and control hutch, which must be air-conditioned. The experiment hutches must be designed according to P2 containment specifications (see annex 2) and should be equipped with an automatic supply of dry liquid nitrogen and the appropriate safety installation.

In the case of the straight section equipped with canted undulators, the optics hutch will have to accommodate two sets of optics.

All the biocrystallography stations will share a common sample preparation area (P2 containment), which will be air-conditioned and will contain work-benches, water (standard and distilled) and means for disposing of chemicals and biological samples. This area must include a cold room and a zone for the storage of frozen samples. This space should have a supply of liquid N_2.

They will also share a room with computer facilities for backing-up data and optional extra data treatment. This room can also serve as a meeting room. A small common workshop should be created adjacent to the sample preparation room. These rooms should be built at the same time as the first experimental station.

Estimated dimensions of the rooms:

Optics hutch: depending on the final choice for optics

Experiment hutch: 4 m x 4 m

Experiment control hutch: 4 m x 5 m

Sample preparation room: 30 m^2 with appropriate work benches.

Cold-room (adjacent to the sample preparation area): 5 m^2

Storage space: 4 m^2

Computer room: 30 m^2

Fluids requested:

- dry N_2 (gas) source (experiment hutch)
- liquid N_2 source in the optics and experiment hutches, frozen crystal storage and at 2 points within the area of the biocrystallography beamlines
- air-conditioning for all enclosed spaces (with temperature adjustment possible)
- compressed air in the experiment hutch and sample preparation space
- cooled water in the optics and experiment hutches
- de-ionised water in the optics and experiment hutches and sample preparation space

Summary of P2 safety requirements (see annex 2)

Definition: biological agents of group 2 can provoke human diseases for which there exists preventive and curative treatment.

Requirements: The space should be enclosed and equipped with doors and at least one window. It must be possible to seal the space if fumigation is required. If air-conditioned, air should be extracted.

All working surfaces must be easy to clean and resistant to disinfectants, acids, alkali and solvents. The entire laboratory must be easy to clean.
The laboratory must be equipped with a sink and taps that can be opened without the use of hands. Eye and hand disinfectant means must be provided. Storage for temporary clothing (lab-coats) must be provided.

3.3.9 *Personnel*

The personnel will play a critical role in both the construction and operation phases of SOLEIL, although the requirements in terms of number and area of competence will differ in certain respects at each stage. The project of construction should be under the responsibility of a scientist, and it would be desirable that the same person continue this role as instrument responsible during operation phase. This responsibility would include the overall conception of the project during the construction phase, overseeing the integration of the different elements – mechanics, electronics, computing – into the design. The role of second responsible would best be taken by a scientist or engineer with a speciality in computing for the development of programmes for control of the experiment. To this end, it would be important that this person have sufficient scientific background in X-ray diffraction measurements on biological macromolecules in order to implement an optimal experimental design into the software. As such this person would part take in scientific running of the instrument at the operation phase while at the same time continuing to implement developments into the software. An electronics or computing engineer would be required to implement the necessary interface for machine control. Finally, two technicians would be required, one to carry out the assembling of the mechanical components of the instrument, the other to introduce and develop automation in the experimental zone. It is important that the construction phase takes place within an adequate infrastructure providing technical support from design and drafting, electronics/computing, mechanics and optics. During the operation phase, the need for intensive technical support will be reduced to a level where personnel could be share between two biocrystallography stations. To meet the requirements of users, however, additional scientific staff will need to be engaged to act as local contacts.

It is strongly recommended that careful measures be taken to ensure that high quality scientists are attrached by the scientific environment of SOLEIL for the operation phase in order to match and justify the heavy investment put into the construction of the biocrystallography stations. This can only be achieved by providing an adequate scientific environment, either on-site and/or in very close proximity to SOLEIL, that will provide an autonomous scientific infrastructure to support in-house projects in structural biology.

Personnel requirements for construction phase

1 Scientist	Responsible
1 Engineer electronics/computing	System and real-time programming
1 Technician	Electronics and automation
1 Scientist / Engineer	Computing (applications software)
1 Technician (general)	Assembly, mechanics, general work

Personnel requirements for operation phase

Responsible	1 scientist/station
Electronics	1 engineer/2 stations
Computing	1 engine er/2 stations
Maintenance	1 technician/2 stations
Local contacts	3 scientists/engineers

ANNEX 1 – The Working Group for biocrystallography beamlines at SOLEIL and acknowledgements

Coordinators

G.A.Bentley (scientific)	IP, Paris	(scientific case, personnel)
R. Kahn (technical)	IBS, Grenoble	(technical case)

Specialists

J.L. Ferrer	IBS, Grenoble	(technical case)
W. Shepard	EMBL, Grenoble	(technical case)
A. Thompson	EMBL, Grenoble	(technical case)

Users

J. Cherfils	LEBS, Gif sur Yvette	(scientific case)
A. Houdusse	IC, Paris	(personnel)
M.H. Le Du	CEA, Saclay	(scientific case, experiment layout)
A. Lewit-Bentley	LURE, Orsay	(contact LURE/SOLEIL, experiment layout, technical case)
M. Knossow	LEBS, Gif sur Yvette	(scientific case)
P. Gouet	IBCP, Lyon	(computing)
A. Mitschler	IGBMS, Strasbourg	
L. Mourey	IPBS, Toulouse	(experiment layout, minutes of meetings)
E. Schmidt	Polytech., Palisseau	(experiment layout)
F. Rey	LGV, Gif sur Yvette	(scientific case)
A. Roussel	AFMB, Marseille	(computing)
H. vanTilbeurgh	Univ. Marseille	

ACKNOWLEDGEMENTS

External experts

H. Belrahli	ESRF, France
C. Schulze	SLS, Switzerland
E. Mitchell	ESRF, France
C. Nave	Daresbury, U.K.
M. Roth	IBS, France

Local experts

J.M. Dubuisson	LURE
B. Gagey	SOLEIL
M.P. Level	SOLEIL
O. Marcouillé	LURE
J.L. Marlats	LURE
A. Nadji	LURE

ANNEX 2 – P2 containment specifications

Definitions and norms for P2 containment are given in the following pages.