4th Meeting of Structural Biologists

Academic and University Center, Nové Hrady, March 10 - 12, 2005

Conference Committee

Bohdan Schneider, Ivana Kutá Smatanová, Radomír Kužel, Dalibor Štys, Jindřich Hašek

Lectures - Thursday, March 10

L1

STRUCTURAL BIOLOGY AT YORK STRUCTURAL BIOLOGY LABORATORY; LABORATORY INFORMATION MANAGEMENT SYSTEMS FOR STRUCTURAL GENOMICS

Jan Dohnálek

Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovského nám. 2, 162 06 Praha 6

The Laboratory of Structural Biology of the University of York has played an important role in development of methods of modern 3D structural analysis of biological macromolecules by means of X-ray crystallography in the last 30 years. Research groups of the Laboratory specialized in different fields form a compact and friendly environment which influenced careers of many scientists also in the Czech and Slovak scientific communities. Current research topics range from industrial applications of knowledge gained from the 3D structural insight into protein structure and protein-ligand interactions through to structural genomics efforts and development of new computational methods in areas of structural analysis, bioinformatics, automation of structure determination process and also development of laboratory information management systems.

Structural genomics projects generate large amounts of experimental data, which are required to be safely recorded and accessible in an organized manner to provide maximum efficiency of information content.

Laboratory Information Management System stands for an ensamble of electronic databases, interfaces and utility programs that enables systematic recording (manual and automatic, i.e. machine produced) of crucial data during research projects at each experimental stage. Such a system should possess several fundamental properties:

1. Well defined but still flexible data model.
2. Ease of data input.
3. Flexible data accessibility.
4. Security.
5. Hierarchy of information type.
6. Compatibility with other LIMS systems.

Structural genomics centers throughout the world developed different approaches to this problem. LIMS systems developed for the purpose of structural genomics activities can be utilized also in research groups with less ambitious goals as for quantity. Several freely available systems for academics were released in the past few years Xtrack[1], CLIMS[2], MOLE[3], PIMS, eHTPX[4], HalX[5]. These provide variable levels of implementation of the above given criteria and of other capabilities. Ease of installation and use and applicability of several of these systems will be discussed.

Introduction of LIMS systems is supported by the Czech Ministry of Education, project no. 1K05008.

1. Harris M. and Jones, T. A. (2002) Xtrack - a web-based crystallographic notebook, *Acta Cryst.* **D58**, 1889-1891.
2. Fulton, K. F., Forster, R., Jodun, R. Ervine, S., Ly, W., Robilliard, L., Sonsini, J., Whelan, D., Whisstock, J. C. & Buckle, A. M. (2004) CLIMS: Crystallography Laboratory Information Management System. *Acta Cryst*. **D60**, pp 1691-3.
3. http://www.mole.ac.uk, MOLE - Mining Organising and Logging Experiments
4. Allan, R., Keegan, R., Meredith, D., Winn, M., Winter, G., Diprose, J., Mayo, C., Launer, L., Fillon, J., Young, P., (2004) e-HTPX HPC, Grid and Web-Portal Technologies in High Throughput Protein Crystallography. In: All Hands e-Science 2004, Nottingham, U.K. 2004.
5. http://halx.genomics.eu.org/. HalX. A free-source LIMS (Laboratory Information Management System) for small to large scale structural biology labs.

L2

HEXAMERIC STRUCTURE OF COLD ACTIVE BETA-GALACTOSIDASE FROM *ARTHROBACTER* SP. C2-2

T. Skálová[1], J. Dohnálek[1], E. Vondráčková[1], H. Petroková[1], P. Lipovová[2], V. Spiwok[2], H. Strnad[2], B. Králová[2] and J. Hašek[1]

[1]*Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovského nám. 2, 162 06 Praha 6, Czech Republic*

[2]*Institute of Chemical Technology, Technická 5, 166 28 Praha 6, Czech Republic*

Arthrobacter sp. C2-2, a soil bacteria found in Antarctica, belongs to psychrotrophic, i.e. cold tolerant, microorganisms. Two beta-galactosidases, isoenzymes C-2-2-1 and C-2-2-2, were isolated from this bacterium. In this contribution, we would like to present an X-ray structure of C-2-2-1 beta-galactosidase from *Arthrobacter* sp. C2-2, which was solved at 1.9 Å resolution.

The beta-galactosidase belongs to glycosyl hydrolase structural family 2. It is composed of 1023 amino acid residues and can be divided into five structural domains. The active site of the beta-galactosidase is localized within the TIM barrel domain. The beta-galactosidase cleaves beta-galactosides by separating of its terminal galactose group and it also catalyzes transglycosylation. Typically, it cleaves lactose into galactose and glucose.

The fold of this beta-galactosidase is similar to that of *Escherichia coli*, which belongs to the same glycosyl hydrolase family 2 and structure of which was studied in details. It is known that *Escherichia coli* beta-galactosidase forms tetramers and that its tetrameric state is necessary for activity.

On the contrary to this, we found that C-2-2-1 beta-galactosidase from *Arthrobacter* sp. C2-2 forms compact hexamers in crystal structure.

The research was supported by the Grant Agency of the Czech Republic (project 204/02/0843/A) and by the Academy of Sciences of the Czech Republic (projects AVOZ40500505 and B500 500 512).

L4

MODERN X-RAY IMAGING TECHNIQUES AND THEIR USE IN BIOLOGY

Radomír Kužel

Faculty of Mathematics and Physics, Charles University, 121 16 Praha 2, Ke Karlovu 5

In the last year contribution for the Meeting of Structural Biologists, recent non-traditional applications of some X-ray techniques - powder diffraction, X-ray reflectivity, grazing incidence, standing waves method in biology were shown [1]. Here the principles of modern X-ray imaging techniques - *X-ray microscopy, tomography* and *single-particle X-ray diffraction* are explained and examples are shown.

X-ray imaging

An excellent review of the use of hard X-rays for imaging was given by J. Härtwig [2] at last year X-top conference in Pruhonice near Prague.

Imaging means revealing of inhomogeneities and singularities in the sample and recording them as a 2-D information through variations in physical effects used for probing like reflection, transmission, absorption.

There are several possibilities of imaging: Imaging with a parallel and extended beam (the spatial resolution is mainly a function of the detector), imaging by using optical elements like lenses (the spatial resolution is mainly a function of the optical element), imaging with a micro beam - scanning (the spatial resolution is mainly a function of the spot size), Bragg-diffraction-imaging (more complex dependence of the spatial resolution).

The simplest techique is the imaging based on the *absorption effect*. The images can be detected by high-resolution films or CCD cameras and scintillators. An interesting application is usage of absorption edge of selected elements. The image is taken at two wavelengths - below and above the absorption edge and then the images are subtracted. In that way, specific contrast can be obtained. It has been used in angiography. However, the main stream seems to be *micro-tomography*, a quantitative description of a slice of matter within a bulky object, based on several radiographs collected at various angles. Images are reconstructed from many projections. Important features of SR are used - high spatial and time resolution and high flux.

When a coherent wavefield propagates through an object, phase differences arise between different parts of the wavefront. These are due to spatial variations in the refractive index of the object, which for x-ray wavelengths is given by

$$n = 1 - \quad + i$$

The first term ( ) determines the phase of the waves and causes their refraction. The second term - absorption ( ) affects the amplitude of the waves. For small wavelengths (<0.1 nm) the phase contrast dominates absorption contrast in the x-ray regime. An advantage of phase contrast is that even small spatial variations in the refractive indexcan be detected. Phase contrast x-ray imaging therefore has great potential for application in medicine and in biology, enabling one to differentiate between different kinds of soft tissue. However, for phase contrast, coherent radiation is necessary. A nice review of the coherent scattering and it use has recently been published in [3].

The structural detail in a phase contrast image depends on the distance between the object and the detector. Three regimes for imaging can be distinguished. In the contact or near-contact regime, the detector is placed directly behind the object. In this case only the absorption contrast can be seen. As one moves further away from the object, interferences build up, and one enters first the *Fresnel diffraction* regime, then the *Fraunhofer regime*. In the Fresnel diffraction regime, phase contrast can be exploited to greatest advantage.

Three-dimensional objects can be reconstructed by repeating the above imaging method for different orientations of the sample. The tomography in phase sensitive mode is called holotomography. The phases are retrieved from the images obtained at different distances in addition to different orientation.

Micro-focussing X-ray beams are obtained by means of different optical elements - capillaries, Fresnel lens, bent mirrors, refractive lens, waveguides. They can be used not only for pure imaging but also for chemical mapping of the objects in 3D (fluotomography). The techniques are based on the detection of photons coming from fluorescent emission, Compton and Rayleigh scattering [e.g. 4].

X-ray microscopy, X-ray tomography and their applications in biology

The technology of X-ray microscopes has improved considerably, in recent years [e. g. 5 - 8].

Although there are currently many powerful techniques for imaging biological cells, each with its own unique strengths and limitations, there remains a gap between the information that can be obtained with light microscopy and electron microscopy.

Soft X-ray microscopy combines features associated with both light and electron microscopy. It is fast and relatively easy to accomplish (like light microscopy), and it produces high-resolution, absorption-based images (like electron microscopy). As with light microscopy, one can examine whole, hydrated cells. In the energy range of the photons used (between the K shell absorption edges of carbon [284 eV, 4.4 nm] and oxygen [543 eV, 2.3 nm]), organic material absorbs approximately an order of magnitude more strongly than water, producing a quantifiable natural contrast and eliminating the need for contrast enhancement procedures to visualize cellular structures.

Using this approach, superb structural information can be obtained from whole, hydrated cells at better than 35-nm resolution. In addition, molecules can be localized using protocols that combine the ease of immunfluorescence labeling with the higher-resolution capabilities of X-ray imaging.

It has been shown that the biological specimens do not suffer mass loss or morphologicla changes at radiation doses up to about 10^{10} Gray [6]. This makes possible studies where multiple images of the same specimen are are needed. X-ray tomography is quite straightforward and rapidly generates 3-D, quantifiable information from whole cells. Because of their low numerical aperture optics, X-ray microscopes allow one to treat high resolution images as simple projections through a specimen.

The main advantage of X-ray microscopy is its ability to study single cells in their entirety, rather than be limited to thickness of 400 nm, as in the case of electron tomography. X-ray microscopes add a unique set of capabilities to biological research. They can be used for three-dimensional imaging of intact cells, and to obtain information on the chemical and trace element markup of unlabeled cells, which cannot be done using other methods.

In [9], the yeast, *Saccharomyces cerevisiae*, was examined using X-ray tomography and unique views of the internal structural organization of these cells at 60-nm resolution was demonstrated. Because proteins can be localized in the X-ray microscope using immunogold labeling protocols, tomography enables 3-D molecular localization. The time required to collect the data for each cell was 3 min, making it possible to examine numerous yeast and to collect statistically significant high-resolution data.

Phase tomography by using of X-ray interferometers has been suggested recently [10]. The image of the tissue of a rat kidney was obtained with 0.1 nm X-rays and X-ray interferometer with a 40 m analyzer.

Oversampling and single-particle diffraction

Protein crystallography using storage ring synchrotron light sources has made it possible to determine the structures of many proteins. However even the brightest storage ring x-ray sources require that the protein be prepared as a crystal – an orderly array of protein molecules, spaced and oriented identically. The millions of molecules in the crystal scatter x-rays in a distinctive pattern that can be used to determine the structure of the protein molecule. The necessary scattering data can be collected in only a few hours at a synchrotron. The difficult part of protein crystallography is producing a usable crystal – this takes 99% of the time and effort. Some of the proteins most important to life processes are difficult or impossible to crystallize. If it were possible to determine the structure of a protein without the need to form a crystal, progress in understanding proteins could be accelerated significantly.

Overcoming the crystallization difficulty requires the development of the new methodologies. One approach is to use NMR to image protein molecules in solvent. However, it is only applicable primarily to macromolecules in the lower molecular weight range. Another approach under rapid development is single molecule imaging using cryo electron microscopy (cryo-EM). The highest resolution currently achievable by this technique is ~7 Å for highly symmetrical viruses. The main limitations to achieving

better resolution by cryo-EM are radiation damage, specimen movement and low contrast. Due to the much weaker interaction between matter with X-rays than with electrons, the X-ray methods require much higher radiation dose to achieve the same resolution by X-rays than by electrons. Another is the difficulty of focussing X-rays. By using zone plates, the best focus currently achievable is ~30 nm for soft X-rays and ~100 nm for hard X-rays [11]. No techniques at present can provide three-dimensional imaging at nanometer resolution of the interior of particles in the micron size range.

An x-ray diffraction intensity is proportional to the square amplitude of the Fourier transform of an object image (an electron number density) in the kinematical theory. If the phase of the Fourier transform is also measured, the object image can be reconstructed by back Fourier transformation. However the phase is not directly obtained in x-ray diffraction measurement. Therefore the half of necessary information for image reconstruction is missing. The problem is known as the phase problem.

Single-particle diffraction is a methodology of extending crystallography to determine the 2D and 3D structures of nano crystals and noncrystalline samples by using coherent x-rays and electrons. In this approach, coherent diffraction patterns are recorded and then converted directly to high resolution images by using the oversampling phasing method.

The oversampling method is a solution to the phase problem [6, 12-20]. In the method a zero density region around the sample object is assumed. If the zero density region is larger than the sample region, more than half of total information is known in real space. Therefore it is in principle possible to use the zero density information to make up for the missing phase information. Thus, the method is based on the measurement of the diffraction intensity in between the Bragg peaks of a traditional crystal. Intensities must be measured at finer-than-Nyquist intervals in diffraction space. The Nyquist sampling interval (inversely related to specimen size) is as fine or finer than the Bragg interval (inversely related to the unit cell size).

The quality of image reconstruction of experimental diffraction patterns is a function of the oversampling ratio a parameter to characterize the oversampling degree. It was observed that the quality of reconstruction is strongly correlated with the oversampling ratio. When the oversampling ratio is around 5 or larger, the reconstructed images with high quality were obtained. When the oversampling ratio is less than 5, the images became noisy. When the oversampling ratio is very close to or less than 2, the images were extremely noisy and barely recognizable [22].

In the oversampling method the following iterative procedure is taken:

1. The magnitude of Fourier transform (i.e., square root of the measured diffraction intensities) is combined with the current best phase set. A random phaseset is used for the first iteration.

2. Applying the inverse fast Fourier transform, a new electron density function is obtained.

3. Constraints are enforced on the electron density function. By pushing the electron density outside the support and the negative electron density inside the support close to zero, and retaining the positive electron density inside the support, a new electron density is defined.

4. Applying the fast Fourier transform to the new electron density, a new set of phases is calculated. After setting the phase of central pixel to zero, this new phase set is used for the next iteration.

This can also be expressed as follows:

..... Image of n-th iteration Fourier transformation Set Fourier amplitudes to known (measured) values Back Fourier transformation Set the density in the zero density region and at negative density pixels gradually to zero Image of (n+1)-th iteration ..

By employment of an iterative algorithm, the phase information could be recovered from computer-generated oversampled diffraction patterns of small specimens that are perfect or imperfect crystals, or have a repeated motif without orientational regularity, or are an unrepeated motif, such as an amorphous glass, a single molecule or a single biological cell [15].

It is expected that extention of the current crystallographic methodology to crystals as small as a single unit cell will be possible, i.e. to non-crystals and individual biomolecules could be studied by using the beam of a free-electron-laser source of X-rays [23]. The resulting continuous diffraction patterns would be invertible into atomic-resolution images of the molecule. The method does not require any structural knowledge concerning the specimen and does not require data to atomic resolution (although it can use such data if present). After a few hundred to a few thousand iterations, the correct phase set and image are recovered.

However, the oversampling technique imposes a high radiation dose on the specimens compared with the situation in crystallography, in which it is usual for the pattern to be sampled at the (much less fine) Bragg spacing (the inverse of the size of the unit cell). An approach to overcoming the degrading effects of radiation damage is to record the diffraction pattern in a time shorter than the time of the damage process itself.

Theoretical simulations by Hajdu and collaborators [24] show that, within about 10 femtoseconds, biomolecules can withstand an X-ray intensity of ~3.8 x 10^6 photons/Å^2 with minimal structural changes. A 2D diffraction pattern can hence be obtained from a single exposure of a biomolecule before the radiation manifests itself.

To get 3D images, it is necessary to record a series of diffraction patterns by rotating the specimen perpendicular to the beam. Unfortunately, any biological sample is destroyed in a single shot. However, if the sample is reproducible, and that single-shot diffraction images can be collected from individual sample particles exposed to the beam one-by-one in unknown orientations. The images can be sorted into classes that correspond to a distinct view (orientation) of the sample [25]. Images within each class are then averaged; if the classification is correct, the signal adds constructively but the noise does not. It was shown that less then one photon per independent pixel can be enough for classification, even in the presence of a Poisson-type photon noise [25].

The method of single-particle diffraction is especially useful in two fields - structural study of nanoparticles and study of macromolecules. In the former field, several successfull experiments have been performed [26, 27] for example on gold nanoparticles. The method has been used for imaging of whole *Escherichia coli* bacteria [28] using coherent x-rays with a wavelength of 2 Å. A real space image at a resolution of 30 nm was directly reconstructed from the diffraction pattern. An R factor used for characterizing the quality of the reconstruction was in the range of 5%, which demonstrated the reliability of the reconstruction process. The distribution of proteins inside the bacteria labeled with manganese oxide has been identified and this distribution confirmed by fluorescence microscopy images.

Review of the single-particle diffraction method and discussion of its possibilities for biological applications can be found in [29]. The main hope is given to new XFELs (X-ray free lectron laser) The X rays produced by LCLS (Linac Coherent Light Source) will be fully transversely coherent. The pulse ($> 10^{12}$ photons) length, initially in the range of 100–200 fs, can be shortened to below100 fs, and with additional research and development it is expected to approach 1 fs.

Summary

The method of coherent X-ray diffractive imaging has been dramatically developed during last few years (in addition to the above papers, see for example also [30, 31]). It has a great potential especially in physics, materials science and biology. In life sciences such a technique is needed to determine the internal structure of assemblies of macromolecules, protein complexes, and virus particles at a resolution sufficient to recognize known proteins and determine their relationship to each other.

X-ray diffraction microscopy, a combination of coherent and bright X-rays with the oversampling phasing method, is a newly developed methodology that may overcome the need of crystallization for obtaining diffraction data. Due to the loss of the amplification from a large number of unit cells inside crystals, the major limitation of the application to structural biology seems to be radiation damage. It is hoped to overcome this problem by cryo technologies and femtosecond-pulsed X-ray lasers enabling to get a record from a biomolecule before it is destroyed.

1. R. Kužel, Some recent non-traditional applications of X-ray scattering techniques in biology, *Materials Structure*, **11** (2004) 19-24.
2. J. Härtwig, Imaging with hard synchrotron radiation, 7th Biennial Conference on High Resolution X-Ray Diffraction and Imaging, Book of Abstracts, September 7-10, 2004, Průhonice near Prague, T3, lecture.
3. F. van der Veen and F. Pfeiffer, Coherent x-ray scattering, *J. Phys.: Condens. Matter* **16** (2004) 5003–5030.
4. B. Golosio, A. Simionovici, A. Somogyi, L. Lemelle, M. Chukalina, A. Brunetti, Internal elemental microanalysis combining x-ray fluorescence, Compton and transmission tomography, J. Appl. Phys. **94** (2003) 145-156.
5. C. Jacobsen, J. Kirz, X-ray microscopy with synchrotron radiation, Nature Structural Biology, Synchrotron Supplement, (1998), 650-653.
6. J. Maser, A. Osanna, Y. Wang, C. Jacobse, J. Kirz, S. Spector, B. Winn & D. Tennant¸ X-ray microscope: I. Instrumentation, imaging and spectroscopy, *Journal of Microscopy*, **197** (2000), pp. 68-79.
7. J. Maser, A. Osanna, Y. Wang, C. Jacobse, J. Kirz, S. Spector, B. Winn & D. Tennant¸ X-ray microscope: II. Tomography, *Journal of Microscopy*, **197** (2000), pp. 80-93.
8. S. Lagomarsino and A. Cedola, X ray microscopy and nanodiffraction, in Encyclopedia for Nanoscience and Nanotechnology, Ed. L. Nalwa, 2004. http://www.ifn.cnr.it/IFN/Roma/L3Alessia/Xray.pdf
9. C. A. Larabell and M. A. Le Gros, X-ray Tomography Generates 3-D Reconstructions of the Yeast, Saccharomyces cerevisiae, at 60-nm Resolution, *Molecular Biology of the Cell*, **15**, (2004) 957–962.
10. Atsushi Momose, Phase-sensitive imaging and phase tomography using X-ray interferometers, *Optics Express*, **11**, (2003) No. 19, 2303.
11. W. Merner-Isle, T. Warwick, D. Attwood (eds), *X-Ray Microscopy*, American Institute of Physics. 2000.
12. D. Sayre, in *Imaging Processes and Coherence in Physics. Springer Lecture Notes in Physics,* (1980) ed. Schlenker, M. (Springer, Berlin), Vol. 112, pp. 229-235.
13. D. Sayre, *Note on 'superlarge' structures and their phase problem in Direct Methods of Solving Crystal Structures*, H. Schenk Eds. New York: Plenum, 1991 pp 353.
14. J. Miao, D. Sayre & H. N. Chapman, *J. Opt. Soc. Am.* A **15**, 1662 (1998).
15. J. Miao & D. Sayre, On possible extensions of X-ray crystallography through diffraction-pattern oversampling, *Acta Crystallogr.* **A56**, (2000) 596.-605.
16. J. Miao, J. Kirz and D. Sayre, The oversampling phasing method, *Acta Crystallographica.* **D56** (2000) 1312-1315.
17. J. Miao, P. Charalambous, L. Kirz, & D. Sayre, (1999) *Nature* **400,** 342–344.
18. J. Miao, D. Sayre, & H. N. Chapman, *J. Opt. Soc. Am. A* **15,** (1998) 1662–1669.
19. J. Miao, T. Ishikawa, B. Johnson, E. H. Anderson, B. Lai, & K. O. Hodgson, *Phys. Rev. Lett.* **89** (2002) 088303.
20. J. Miao, T. Ohsuna, O. Terasaki, K. O. Hodgson, & M. A. O'Keefe, *Phys. Rev. Lett.* **89** (2002) 155502.
21. U. Weierstall, Q. Chen, J. C. H. Spence, M. R. Howells, M. Isaacson, & R. P. Panepucci, *Ultramicroscopy* **90,** (2002) 171–195.
22. J. Miao, Tetsuya Ishikawa, E. H. Anderson, K. O. Hodgson, Phase retrieval of diffraction patterns from noncrystalline samples using the oversampling method, *Phys.Rev.* **B67** (2003) 174104.
23. C. Pellegrini, J. Stöhr: X-ray free-electron lasers—principles, properties and applications, *Nuclear Instruments and Methods in Physics Research A* **500** (2003) 33–40.
24. R. Neutze, R. Wouts, D. Spoel, E. Weckert, J. Hajdu, *Nature*, **406** (2000) 752-757.

25. G. Huldt, A. Szökeke and J. Hajdu, Diffraction imaging of single particles and biomolecules, *Journal of Structural Biology* **144** (2003) 219–227.

26. I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pferfer, & J. A. Pitney, *Phys. Rev. Lett.* **87** (2001) 195505.

27. Ian K. Robinson, F. Pfeiffer, I. A. Vartanyants, Yugang Sun and Younan Xia, Enhancement of coherent X-ray diffraction from nanocrystals by introduction of X-ray optics, *Optics Express*, **11**, (2003) No. 19, 2329.

28. J. Miao, K. O. Hodgson, Tetsuya Ishikawa, C. A. Larabell, M. A. LeGros and Yoshinori Nishino, Imaging whole Escherichia coli bacteria by using single-particle x-ray diffraction, PNAS - Proceedings of the National Academy of Sciences of the USA, **100** (2003), no. 1, 110–112.

29. J. Miao, H. N. Chapman, J. Kirz, D. Sayre, and K. O. Hodgson, TAKING X-RAY DIFFRACTION TO THE LIMIT: Macromolecular Structures from Femtosecond X-Ray Pulses and Diffraction Microscopy of Cells with Synchrotron Radiation, *Annu. Rev. Biophys. Biomol. Struct*. **33** (2004) 157–76.

30. S. Eisebitt, M. Lörgen, W. Eberhardt, J. Lüning, S. Andrews, and J. Stöhr, Scalable approach for lensless imaging at x-ray wavelengths, Appl. Phys. Let, **84**, (2004) no. 17, 3373-3375.

31. S. Marchesini, H. N. Chapman, S. P. Hau-Riege, R. A. London, A. Szoke, H. He, M. R. Howells, H. Padmore, R. Rosen, J. C. H. Spence, U. Weierstall, Coherent X-ray diffractive imaging: applications and limitations, *Optics Express*, **11**, (2003) No. 19, 2344.

Lectures - Friday, March 11, morning

L5

MOLECULAR MECHANISM OF 14-3-3 PROTEIN FUNCTION REGULATION

T. Obšil[1,2]

[1]Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, 128 43 Prague,

[2]Institute of Physiology, Academy of Sciences of the Czech Republic, 142 20 Prague

The 14-3-3 protein family represents one of the most important group among proteins recognizing phosphorylated targets [1, 2]. Many of 14-3-3 binding partners contain one of two consensus motifs, RSXpSXP, and RXY/FXpSXP, where pS denotes phosphoserine. 14-3-3 proteins function as molecular scaffolds modulating the conformation of their binding partners. As a result of this structural modulation they can: (i) affect the enzymatic properties of their partners [3]; (ii) interfere with the protein-protein interactions of their targets; or (iii) regulate the subcellular localization of their binding partners presumably by masking or obscuring a nearby targeting sequence, such as a nuclear localization sequence (NLS) or a nuclear export sequence (NES) [4]. The Cdc25 phosphatases, telomerase, histone deacetylase, or FoxO transcription factors have been suggested to be subject to such regulation.

All 14-3-3 proteins form very stable homo- and hetero-dimers with characteristic cup-like shape and a large, 40 Å-wide, deep central channel [2,3]. Each monomer consists of nine antiparallel -helices and an amphipathic groove where the phosphorylated segment of the binding partner is bound. The primary structure of 14-3-3 proteins is highly conserved with completely conserved residues forming either the dimer interface or the walls of the ligand binding groove. Maximal sequence diversity occurs within the flexible C-terminal stretch and it has been suggested that this region is involved in the regulation of binding properties of 14-3-3 proteins [2,5]. In addition, the C-terminal stretch of two vertebrate 14-3-3 isoforms (tau and zeta) contains a casein kinase I phosphorylation site at position 232 [6]. Phosphorylation of 14-3-3 proteins has been suggested as an important regulatory mechanism of individual isoforms [1].

We have used various methods of fluorescence spectroscopy to study structural changes of the C-terminal segment induced by ligand binding and phosphorylation at T232, a casein kinase I phosphorylation site located within this region. We have showed that the phosphopeptide binding changes the conformation and increases the flexibility of the 14-3-3 C-terminal stretch. Phosphorylation of 14-3-3 at T232 resulted in inhibition of ligand binding and significant change of the C-terminal stretch conformation [7]. In order to fully understand the role of the C-terminal stretch in the regulation of 14-3-3 protein binding properties we have also investigated the physical location of the C-terminal stretch and its changes upon the ligand binding. For this purpose we have used Förster resonance energy transfer (FRET) measurements and molecular dynamics simulation. These studies demonstrated that in the absence of the ligand the C-terminal stretch occupies the ligand binding groove of the 14-3-3 protein. The phosphopeptide binding displaces the C-terminal stretch from the ligand binding groove of the 14-3-3 protein molecule [8].

Our results are consistent with the hypothesis suggesting that the 14-3-3 C-terminal stretch functions as a suppressor of unspecific interactions between the 14-3-3 protein and inappropriate ligands. In the absence of the ligand the C-terminal stretch occupies ligand binding groove of the 14-3-3 molecule. Binding of the phosphopeptide with proper 14-3-3 binding motif displaces C-terminal stretch from the ligand binding groove. Phosphorylation of the T232, a casein kinase I phosphorylation site located within this region, likely mod-

ulates conformational changes of the C-terminal stretch and thus all its functions.

This work was supported by Grants 204/03/0714 and 309/02/1479 of the Grant Agency of the Czech Republic; by Grant B5011308 of the Grant Agency of the Czech Academy of Sciences; by Research Projects 1K03020, MSM 1131 00001, and 1132 00001 of the Ministry of Education, Youth and Sports of the Czech Republic, and by Research Project AVOZ 5011922.

1. H. Fu, R. R. Subramanian & S. C. Masters, *Annu. Rev. Pharmacol. Toxicol.*, 40 (2000) 617-647.
2. D. Liu, J. Bienkowska, C. Petosa, R. J. Collier, H. Fu & R. Liddington, *Nature*, 376 (1995) 191-194.
3. T. Obsil, R. Ghirlando, D. C. Klein, S. Ganguly & F. Dyda *Cell*, 105 (2001) 257-267.
4. A.J. Muslin & H. Xing, *Cell Signal.*, 12 (2000) 703-709.

A. B. Truong, S. C. Masters, H. Yang & H. Fu, *Proteins* 49 (2002) 321-325.

5. T. Dubois, C. Rommel, S. Howell, U. Steinhussen, Y. Soneji, N. Morrice, K. Moelling & A. Aitken, *J. Biol. Chem.* 272 (1997) 28882-28888.
6. V. Obsilova, P. Herman, J.Vecer, M. Sulc, J. Teisinger & T. Obsil, *J. Biol. Chem.,* 279 (2004) 4531-4540.
7. J. Silhan, V. Obsilova, J. Vecer, P. Herman, M. Sulc, J. Teisinger & T. Obsil, *J. Biol. Chem.,* 279 (2004) 49113-49119.

L6

SOLUTION STRUCTURE AND DYNAMICS OF PEPTIDES AND PROTEINS FROM RAMAN OPTICAL ACTIVITY

V. Baumruk,[1] J. Kapitán,[1,2] P. Bouř[2]

[1]Institute of Physics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 5, 121 16 Prague 2, Czech Republic, e-mail: baumruk@karlov.mff.cuni.cz

[2]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

In the postgenomic era, an increasing demand for structural information about biomolecules has appeared. This is partially met by the high-resolution nuclear magnetic resonance (NMR) spectroscopy and the X-ray crystallography. They are currently invaluable and will continue to be appreciated in the field of structural biology due to their ability to reveal structural details at the atomic resolution. However, there are limitations of their applicability. For example, many proteins are difficult to crystallize, and many others are too large or complicated to be currently approached by NMR. Although vibrational spectroscopic techniques do not yield the three-dimensional structures of proteins resolved at the atomic resolution directly, they provide useful alternative or supplementary information of these systems. Local arrangements of functionally important groups within macromolecules, or changes in these local structures that may relate to biological functions, are typical examples that can be conveniently probed by the vibrational spectroscopy.

A particularly informative method of vibrational spectroscopy is Raman optical activity (ROA) which refers to a small difference in the intensity of Raman scattering from chiral molecules in right- and left-circularly polarized incident light [1]. ROA spectrum of a chiral molecule can contain up to 3*N*-6 fundamental bands (*N* is the number of atoms), each associated with one of the normal modes of vibration. These contain information about conformation and absolute configuration of the particular part of the structure embraced by the normal mode.

The data obtained by Raman spectroscopy and ROA are complementary, as both Raman and ROA scattering are generated by related, yet subtly different, mechanisms. For proteins the ROA spectra tend to be dominated by the bands arising mainly from polypeptide backbone, detailing thus the secondary and tertiary structure. On the other hand, Raman spectra, which are insensitive to chirality, tend to be dominated by bands from sidechains that are particularly sensitive to local environment. However, the principal problem with analysis of both Raman and ROA spectra is in obtaining definitive assignments for all component bands in a spectrum, because of band overlaps and a lack of satisfactory models for some structural motifs. Information obtained by both the techniques contains qualitatively different information than that yielded by NMR spectroscopy, because the time scale of Raman scattering event is much shorter than that of the fastest conformational fluctuation in biomolecules. The vibrational spectrum represents a simple superposition (weighted sum) of the spectra of all the conformers present in the sample.

It has been shown that ROA is an excellent technique for studying the polypeptide and protein structures in aqueous solution [2]. Protein ROA spectra provide information on the secondary and tertiary structures of the polypeptide backbone, hydration, side-chain conformation, and structural elements present in denatured states. Backbone vibrations in polypeptides and proteins are usually associated with three main regions of the Raman spectrum: (i) the backbone skeletal stretch region (~870-1150 cm^{-1}); (ii) the extended amide III region (~1230-1340 cm^{-1}); and (iii) the amide I region (~1630-1700 cm^{-1}) [3]. The extended amide III region is particularly important for ROA studies because the coupling between N-H and C -H deformations is very sensitive to geometry and generates rich ROA band structure [4]. There are clear differences between ROA

spectra of -helix, -sheet and disordered conformations of model polypeptides that enable ROA to discriminate between ordered and disordered polypeptide sequences in folded proteins [5].

In addition to ROA bands characteristic for -helix and -sheet, sharp ROA bands have been observed in the extended amide III that appear to originate from loops and turns. It can be explained as a consequence of the fact that ROA reflects the local geometry of individual residues due to the local short-range nature of the vibrational coupling [5]. Because of its ability to detect loop and turn structure, ROA is able to monitor changes in the protein fold that are inaccessible to conventional techniques [6] and give interesting results even on the partially unfolded states associated with reduced proteins and molten globules [7]. Unfolded proteins which are often difficult to investigate by means of high-resolution techniques have proven an especially fruitful object of investigation for ROA spectroscopy as many of the proteins important for proteomics are expected to be relatively unstructured [2].

Analyses of the structural information content from ROA spectra of biological systems have ranged from totally empirical correlation to fully theoretical predictions. For proteins, PCA (principal component analysis) based pattern recognition approach to identify protein folds from ROA spectral band shapes is under development [8] while at least for short peptides, the full *ab initio* quantum mechanical ROA simulations are possible [9].

The aim of this work is not only to give an overview of fruitful application of ROA in peptide and protein science but also present very recent results from our laboratory.

Ministry of Education of the Czech Republic is gratefully acknowledged for support (No. MSM 0021620835).

1. L.D. Barron, *Molecular Light Scattering and Optical Activity*, Cambridge University Press, Cambridge 1982.
2. L.D. Barron, E.W. Blanch, L. Hecht, *Adv. Protein Chem.* **62** (2002) 51-90.
3. T. Miura, G.J. Thomas Jr., in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research*, pp. 261-315, Academic Press, New York 1995.
4. S.J. Ford, Z.Q. Wen, L. Hecht, L.D. Barron, *Biopolymers* **34** (1994) 303-313.
5. E.W. Blanch, L. Hecht, L.D. Barron, *Protein Sci.* **8** (1999) 1362-1367.
6. E.W. Blanch, L.A. Morozova-Roche, L. Hecht, W. Noppe, L.D. Barron, *Biopolymers* **57** (1999) 235-248.
7. L.D. Barron, L. Hecht, E.W. Blanch, A.F. Bell, *Prog. Biophys. Mol. Biol.* **73** (2000) 1-49.
8. E.W. Blanch, I.H. McColl, L. Hecht, K. Nielsen, L.D. Barron, *Vib. Spectrosc.* 35 (2004) 87-92.
9. P. Bouř, V. Sychrovský, P. Maloň, J. Hanzlíková, V. Baumruk, J. Pospíšek, M. Budešínský, *J. Phys. Chem. A* **106** (2002) 7321-7327.

L7

DROP COATING DEPOSITION RAMAN – A NEW METHOD IN RAMAN SPECTROSCOPY FOR BIOMOLECULAR SAMPLES OF LOW CONCENTRATIONS

V. Kopecký Jr., V. Baumruk, J. Kapitán, J. Štěpánek

Institute of Physics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 5, 121 16 Prague 2, Czech Republic, e-mail: kopecky@karlov.mff.cuni.cz

Raman spectroscopy can be a powerful probe of structure and function of biomolecules. However, its usage, particularly in biochemistry, is often limited by amount and concentration of samples. Recently reported new technique of non-enhanced Raman spectroscopy – a drop coating deposition Raman (DCDR) method [1] – based on a coffee ring effect [2] enables nondestructive measurements of solutions with concentration of biomolecules down to 1 μM (ca. 0,01 mg/mL). The method is based on a deposition of a small amount of sample solution (typically units of L) on a Teflon-coated stainless steel surface with almost no Raman signal. The hydrophobic surface enables drying of a sample. It has been demonstrated that even after drying its solution structure is preserved and Raman spectra taken from the dried deposit and (concentrated) solution are virtually identical [1].

Increasing number of papers demonstrates usefulness of DCDR method. It has been shown that DCDR can serve as an analytical tool in identification of insulin variants [3] or as a detection method of peptide tyrosine phosphorylation [4]. DCDR is useful also in chemometric analysis as was demonstrated by oligosaccharide identification and quantification in mixtures [5]. Possible coupling of DCDR with HPLC analysis and extension to proteomic analysis generally can bring new quality to scientific research.

Despite of significant increase of information about the practical usage of DCDR method, the basic questions remain unanswered: (i) how is it possible that the samples are almost identical in solution and in a drop deposit? (ii) are they really identical or little specific differences do exist? (iii) is it possible to estimate humidity of drop deposited samples? (iv) what samples remain intact after deposition and what samples are not suitable for DCDR method?

The aim of this work is to answer these questions as well as to give an overview of DCDR method and our first experience with it.

The authors thank M. Trchová (Institute of Macromolecular Chemistry, AS CR) for her kind assistance with Raman microscopy and K. Hofbauerová and K. Bezouška

(Department of Biochemistry, Charles University) for a generous gift of protein samples. Ministry of Education of the Czech Republic and the Grant Agency of the Czech Republic are gratefully acknowledged for support (No. MSM 0021620835, No. 202/05/0628, respectively).

1. D. Zhang, Y. Xie, M. F. Mrozek, C. Ortiz,V. J. Davisson, D. Ben-Amotz, *Anal. Chem.* 75 (2003) 5703–5709.
2. R. D. Deegan, O. Bakajin, T. F. Dupont, G. Huber, S. R. Nagel, T. A. Witten, *Nature* 389 (1997) 827–829.
3. C. Ortiz, D. Zhang, Y. Xie, V. J. Davisson, D. Ben-Amotz, *Anal. Biochem* 322 (2004) 245–252.
4. Y. Xie, D. Zhang, G. K. Jarori, V. J. Davisson, D. Ben-Amptz, *Anal. Biochem.* 332 (2004) 116–121.
5. M. F. Mrozek, D. Zhang, D. Ben-Amotz, Carbohydr. Res. 339 (2004) 141–145.

L8

STABILISATION OF PROTEIN STRUCTURES BY SOLVENT MOLECULES

Ľ. Urbániková and J. Ševčík

Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, 845 51 Bratislava, Slovak Republic, e-mail: lubica.urbanikova@savba.sk; jozef.sevcik@savba.sk

RNase Sa, Sa2 and Sa3, ribonucleases produced by various strains of *Streptomyces aureofaciens*, have proven to be excellent models for various types of studies including structure-function relationship, cytotoxicity, mechanism of enzymatic reaction, protein-protein and protein-nucleic acid recognition. They are small proteins consisting of 96-98 amino-acid residues with 56-69 % identity in primary structures. Properties of the enzymes have been thoroughly studied. More than ten years of structural studies of S.a. ribonucleases resulted in solving a number of 3D structures of free enzymes [1, 2, 3], mutants and complexes with several ligands at various resolutions including atomic. Structures of ten RNase Sa mutants have been used for the study of conformational stability of globular proteins orientated on better understanding of the contributio0n of individual hydrogen bonds to the protein stability [4, 5, 6]. Mutations were prepared with the aim to remove a limited number of hydrogen bonds at retaining the protein structure. It has been proved that the effect of a single amino-acid mutation on conformational stability of a protein highly depends on the location of the substitution and its environment in the structure.

High accuracy structures refined against high resolution data (2.0 - 1.0 Å) allowed inspection of subtle changes not only in protein itself but also in the solvent structure. There are evidences that protein-water hydrogen bonds contribute to the stability of the protein molecule. In the structures of all three S.a. ribonucleases and RNase Sa mutants, water molecules which are fully or mostly buried in the protein structures were identified. Their possible role in the stabilisation of protein structure will be discussed.

RNase Sa, Sa2, Sa3 and RNase Sa mutants were crystallised usually from phosphate buffer using ammonium sulphate as precipitant or additive, therefore it is not surprising that sulphate/phosphate anions were identified also in the structures. The structures of sulphate and phosphate anions are nearly identical; therefore the identity of the anions cannot be clarified. Their possible role in mediating crystal contacts and stabilization of protein molecules will be discussed. Structural observations will be extrapolated on behaviour of the enzymes in solution to support experimental findings on the stabilising role of phosphate buffer used in biochemical measurements and isolation of S.a. ribonucleases.

1. Sevcik J., Dauter Z., Lamzin V.S. & Wilson K.S., *Acta Cryst*. D**52** (1996) 327.
2. Sevcik J., Urbanikova L., Leland P.A. & Raines R.T., *J. Biol. Chem.* **277** (2002) 47325-47330.
3. Sevcik J., Dauter Z. & Wilson K. S. *Acta Cryst*. D**60** (2004) 1198-1204
4. Hebert E.J., Giletto A., Sevcik J., Urbanikova L., Wilson K.S., Dauter Z. & Pace C.N, *Biochemistry* **37** (1998) 16192.
5. Pace C.N., Horn G., Hebert E.J., Bechert J., Shaw K., Urbanikova L., Scholtz J.M. & Sevcik J., *J. Mol. Biol.* **312** (2001) 393.
6. Alston R.W., Urbanikova L., Sevcik J., Lasagna M., Reinhart G.D., Scholtz J.M. & Pace C.N., *Biophys. J.* **87**, (2004) 4036-47.

L9

E.COLI EXPRESSION, PURIFICATION AND BINDING STUDIES OF GALECTIN-4

Veronika Suchá[1], Gabriela Jeníková[2], Vladimíra Marková[2], Milan Fabry[1], Petr Malý[2], Jiří Brynda[1]

[1]Department of Recombinant Expression and Structural Biology, Institute of molecular genetics, Flemingovo nám. 2, 166 37 Praha 6

[2]Department of Molecular Glycobiology, Institute of molecular genetics, Vídeňská 1083, 142 20 Praha 4

Galectins belong to a family of carbohydrate-binding proteins that share conserved amino acid sequences and affinity for ß-galactoside sugars.

Galectin-4 is a monomer of about 36 kDa containing two carbohydrate-binding domains (CDR1 and CRD2, 40% identical) within a single polypeptide chain [1]. This protein is expressed only in the alimentary tract, from the tongue to the large intestine. Strong expression of galectin-4 can be induced in cancers from other tissues including breast and liver. This distinct induction can make a valuable diagnostic marker and target for the development of inhibitory carbohydrate-based drugs.

Galectins may also bind intracellular non-carbohydrate ligands and have intracellular regulatory roles in processes such as RNA splicing, apoptosis, and, as suggested most recently, the cell cycle [2].

The exact function of galectin-4 has not been exactly found out yet.

Full-length mouse galectin-4 and its respective carbohydrate recognition domains CRD1 and CRD2 were expressed in *E. coli* and consecutively purified. Lactose-binding affinity of CDR1 was determined by fluorescence assay based on fluorescence quantum yield of two tryptophan residues (Trp 62, Trp 75) located on opposite borders of the binding site.

Specificity of recombinant galectin-4 and its respective carbohydrate recognition domains CRD1 and CRD2 labeled with FITC was determined using a glycan array consisting of synthetic and natural carbohydrates attached to microtiter plates.
http://www.functionalglycomics.org/static/consortium/main.shtml

1. M.A. Gitt, C. Colnot, F. Poirier, K.J.Nani, S.H. Barondes, H. Leffler, *J. Biol.Chem*, **273** (1997), 2954-2960.
2. M.H. Huflejt, H. Leffler, Glycoconjugate J., **20** (2004), 247-255.

L10

INFLUENCE OF ADDITIVES ON CRYSTALLIZATION OF THE FLAVOPROTEIN WrbA

J. Wolfová[1], J. Carey[2] and I. Kutá Smatanová[3]

[1]Institute of Physical Biology, University of South Bohemia at Ceske Budejovice, Zamek 136, 373 33 Nove Hrady, Czech Republic,

[2]Chemistry Department, Priceton University, Washington Rd and William St, Princeton, NJ 08544-1009, USA

[3]Institute of Physical Biology, University of South Bohemia at Ceske Budejovice and Institute of Landscape Ecology, Academy of Science of the Czech Republic, Zamek 136, 373 33 Nove Hrady, Czech Republic

A novel flavodoxin-like protein, tryptophan (W)-repressor binding protein A, WrbA, was first identified as an *Escherichia coli* stationary-phase protein enhancing the binding interaction between DNA and the tryptophan repressor (TrpR), which regulates the biosynthesis of tryptophan in procaryotes. Later work [1] showed that WrbA doesn't specifically influence the binding of TrpR to DNA and thus left the physiological role of WrbA unclear. According to sequence analysis and homology modelling, WrbA was identified as the founding member of a new protein family, sharing the open, twisted / fold typical for flavodoxins [2]. The biochemical and biophysical studies of purified WrbA protein [1] revealed that the WrbA protein binds flavin mononucleotide (FMN) similarly to flavodoxins, but the affinity of WrbA for FMN is much lower. This may be associated with the structural differences in the flavin-binding pocket predicted by the computer analysis. Analytical ultracentrifugation in combination with the size-exclusion chromatography showed the multimeric character of WrbA protein in solution. WrbA is apparently the first characterized case in which multimerization is associated directly with the flavodoxin-like domain itself. In all other multimeric flavodoxins the flavodoxin-like domain is fused to a multimerization domain [3]. WrbA protein and its homologs thus present a unique family among the typical flavodoxin-like proteins. Structural analysis may aid in understanding these unique properties and may reveal the physiological role of WrbA in the living organ-

isms. This was a motivation for searching of diffraction-quality crystals.

Crystallization conditions for growing of WrbA apoprotein crystals were found using standard and advanced crystallization techniques [4]. As the WrbA crystals grew as twinned plates additives were used to improve the quality of crystals. Crystals suitable for X-ray diffraction measurements grown in capillaries as well as in sitting drops will be measured at synchrotron at cryotemperature and the diffraction data will be used for solving protein structure in the future.

Limited proteolysis [5] of WrbA apoprotein led to preliminary identification of folded substructures and flexible parts of protein structure. Further analysis of the fragments gained by the proteolytic digestion will serve as a competent accessory to X-ray structure.

This work is supported by grant of the Ministry of Education of the Czech Republic (project KONTAKT ME640 to I.K.S.), by grants MSM6007665808 and AVOZ60870520, and by NSF grant INT-03-09049 to J.C.

1. R. Grandori, P. Khalifah, J.A. Boice, R. Fairman, K. Giovanielli & J. Carey, *J. Biol. Chem.,* 273 (1998) 20960-20966.
2. R. Grandori & J. Carey, *Trends Biochem. Sci.,* 19 (1994) 72.
3. J. Ostrowski, M.J. Barber, D.C. Rueger, B.E. Miller, L.M. Siegel & N.M. Kredich, *J. Biol. Chem.,* 264 (1989) 15796-15808.
4. J. Lopez-Jaramillo, J.M. Garcia-Ruiz, J.A. Gavira & F. Otalora, *J. Appl. Cryst.,* 34 (2001) 365-370.
5. J. Carey, *Methods Enzymol.,* 328 (2000) 499-514.

L11

COMPUTER SIMULATIONS OF REAL MOLECULAR SYSTEMS

Jindřich Hašek

Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovského nám. 2, CZ-162 06 Praha 6, Czech Republic

Introduction

Quality of any computer study of real macromolecular systems heavily depends on availability of experimental information about the structure. Vast majority of the protein structures deposited in the "Protein structure database" (PDB) /1/ originates from X-ray structure analysis. It is gladdening that these data are widely used by many other experimental and theoretical methods, however one can find a lot of striking examples of incorrect use and misinterpretations even in the well established journals. This paper shows some basic facts that must be observed when using the protein structures deposited in the PDB.

Most of programs are able to use atomic coordinates from PDB files automatically, skipping all other information about physical and chemical background. Other programs easily add missing hydrogens and fill in the atomic types and thus everybody can virtually grow into expert and draw nicely looking pictures of molecules in few days. Thus, it happens from time to time that we are surprised when we read about a "successful" interpretation of IR spectra resulting from the "fact" extracted from PDB that the liganded HIV protease is dimer and the unliganded monomer - the authors did not noticed that the ordinary PDB file shows the symmetrically independent part of molecule only and that the related monomer of unliganded protein (of course missing in the original *.pdb file) forms with the second part a compact unit.

To avoid many other similar "obstacles", one should observe and work with other parameters hidden in the data deposited in the database. Here is a short overview of some important actions without which the user of PDB can potentially easily destroy an afford spend in his future work.

Molecular Model for Energy Calculation

Diffraction experiment provides as its primary result the map of electron density showing a time averaged view of all conformations and all intermolecular interactions with solvent and possibly with other molecules contained in the solvent drop surrounding the protein crystal at the time of measurement. The map shows all atoms which have some defined positions in the structure. Thus, we can see solvent molecules bound to protein in the first hydration shell, but the freely movable bulk water form only the uniform background without any possibility to talk about water positions. However, to start with energy calculations, one needs a single conformation of some low energy state of the molecular system.

Choice of the representative structure(s)

PDB often contains more structure determinations of the same molecule with different ligands, in different molecular environments or in different pH and buffers. Choice of the system which is near the intended subject of study is surely worth of some afford.

Simultaneously, one requires a clear and unambiguous view of structure. Thus, one of the important issues is accuracy and reliability of the selected structure determination. One should evaluate the following:

- Check the overall quality of the structure determination, i.e. the resolution, R and R_{free} factors. These figures have only informative character. Do not believe much structures with $R > 0.25$ and $R_{free} > R + 0.1$. Notice specially, that the resolution does not mean accuracy. For example Res = 4 Å means "unreliable", Res = 3 Å means "doubtful", Res =2 Å reli-

able with accuracy 0.1 Å, Res =1 Å means accuracy 0.02 Å and you can start to interpret position of hydrogen atoms.
- Check for chemical composition of buffer, pH, etc. for all candidates selected from PDB.
- Check always the alignment of your sequence with both SEQREC and ATOM records.
- Check for disorder or any missing parts of molecule. In critical cases decide about the chain truncation or modelling of the fragments missing in experimental PDB file.
- Check the quality of the structure determination for each residuum in selected structures.

Temperature of measurement

In spite of the fact that many diffraction experiments are done under very low temperatures derived from liquid nitrogen, the reported structure always corresponds to the solvated protein structure under ambient temperatures. It is because the sample should be flash frozen so quick that solvent water remains amorphous in all cases, otherwise no diffraction from protein could be measured. Thus, the protein structure measured under low temperatures always corresponds to the fold at the ambient temperature, i.e. to the situation just before the flash freezing.

One of many advantages of flash freezing of the protein sample is that the local vibrations of atoms are greatly restricted, some side chains with multiple conformations are fixed in lower energy states and thus one receives much sharper view of the low energy states of the whole molecular system. This is the reason, why the use of "low temperature" data is generally preferred.

Choice of a single conformation from many

Conformational flexibility of macromolecule is in diffraction experiment represented by two factors. One is atomic displacement factor describing the motion of atom by ellipsoidal probability density of finding the atom. When the motion of atom is so large that the ellipsoidal description cannot describe the motion well, then the concept of multiple conformation is used. The disordered residuum or other part of molecule is build into the electron density several times with restriction that the sum of occupation factors is one for each atom. Thus, contrary to NMR results were one is offered by a series of representative models, the crystallographers leave this step on the user of their results. At first site it may sound horrible that let say eighteen residues with two possible conformations theoretically lead to 262 144 structure models. However, it is relatively very easy to go through individual conformations and to decide which is the best for your purpose.
- Check all residues with multiple conformations in the main chain or side chains.
- Choose the suitable conformational model
- Check the interesting areas of the chosen structure at the map of electron density.

Idealization of the model

X-ray crystallography provides the structure of macromolecule in its natural environment with all interactions with solvent and other neighbour molecules. In calculations, one always works with some idealized system, so that the problem solved has prevailing affect on calculated energies and the changes in the distant parts of the system have possibly a negligible effect on the calculation.
- Identify the areas at the surface where the intermolecular contacts take place and analyse the types of interactions. Decide how it suits to the purpose of your work. Decide whether to incorporate it to your model, to truncate it, to fix positions, or to minimize the relevant residues.

Water molecules

Solvent is essential for protein conformation and therefore it needs a special care. Protein crystal is a regular arrangement of loosely connected protein molecules with solvent filling the gaps between them. Solvent form usually about 50 % of the crystal mass. The diffraction methods can identify well localized positions for water (ions or other solvent molecules), usually at the surface or inside of the protein. In the well made structure analysis it represents usually up to 10 % of solvent molecules. Thus, the experiment gives only the most important water positions. The bulk solvent (the solvent molecules which have no fixed position to macromolecule) is represented by continuum. However, crystallographers usually do not concentrate on problem of hydration and some of them simple give no information about waters in their resulting PDB files. Therefore, to work with reliable hydration, one have to run programs for searching the empty cavities and crevices inside of protein and to fill them by reasonably oriented water molecules.

Keep in mind that the "crystallographic waters" means the average well localised sites for water molecules with residence times running from tenths ps at the protein surface to tenths ns for water positions deep inside of protein. Thus, it depends on purpose whether to tether the crystallographic waters at their average positions or to leave them free in dynamics.

Hydrogens

Neutron diffraction which provides reliable determination of hydrogen atoms is not widely used method. The X-ray crystallography has problems to localize hydrogen atoms in proteins and therefore their positions are not given in the PDB file as a rule. If you find hydrogens in the PDB file, ask some crystallographer to assess the meaning and reliability.

Add hydrogens automatically, however, the overall assessment of all possible variants namely inside of protein is always necessary. Keep in mind that the ionization inside of protein need not correspond to the solution pH and that all that is a dynamic process.

Equilibration of the starting structure model

The structure deposited in the PDB file represents the time averaged structure model and the atomic positions are subject of some experimental errors. Therefore, the preliminary equilibration of the structure model using the energy minimization of your choice is always necessary. The procedure is simple:

- Assign parameters of all atom types (atom types). If it is done by program automatically, check all (atom by atom) carefully.
- Tether all non-hydrogen atoms by soft forces and minimize hydrogen positions freely.
- Optimize the starting model by your favorite minimization method.

Final check of the starting model

The primary result of diffraction experiment is a map of electron density. Therefore, the final structure model should correspond well the map "pipes". Of course deviations are expected at all places of truncation, rebuilding or any other intervention during the model system formation.

- Check how the resulting structure model fits the experimental map of average electron density.
- The difference should be interpreted as caused by interventions made before during the model building
- The results may differ slightly according to the method of energy calculation and the tethering potential.

Molecular Dynamics

High complexity of molecular simulation methods and applications does not allow general comments. Thus we can show only example of some misinterpretations following from static perception of the X-ray structure model and wrong representation of hydration.

Hydration is of basic importance for any protein structure and thus any afford consisting in unconstrained molecular dynamics without well represented water around the protein might be approved only in the case that one wants to see how easy is to dry and destroy his protein.

Some other problems of molecular dynamics found in literature follow from the fact that the macromolecules pass through their conformational space to slow (microseconds) so that in majority cases one cannot receive a plausible statistical sample by calculations of reasonable length. Here is a place for exploitation of logical extrapolations of simplified case studies of molecular dynamics, showing for example the ways how water molecules circulate at the protein surface, how water molecules penetrate in between the protein loops, how fast and in what form the ions penetrate inside of protein, what is average time necessary to deliver water molecule in the specific place inside of protein important for explanation of delay in spreading the solution pH into the protein bulk in the course of enzymatic reactions, etc.

Another area for extrapolation and case studies is in an evaluation of conformation changes and motions of large molecular fragments. Here the molecular dynamics is successfully combined with suitably chosen additional external potentials forcing the molecular fragments to pass through the required paths.

Conclusion

Finally it is important to stress that the basic experimental result of diffraction experiment is the electron density in the average unit cell averaged over the time of measurement. This map contains full information about the conformation variability of the protein in the drop of solvent under temperature before flash freezing including the protein interactions with water and other components of the buffer used for experiment, and it reflects also the intermolecular interactions with neighbour molecules. The unbiased use of experiment requires more than copying the atom coordinates from PDB file. The presented coordinates are only some interpretation of the map by the scientist who made the final refinement of the structure model into the electron density.

Acknowledgement

The study was supported by GA AV CR KJB4050312 and MSMT 1K05008.

APPENDIX A

Some servers that can help in assessment of molecular models.

Search through the PDB:

Protein Structure database searching tools:
http://www.ebi.ac.uk/msd-srv/msdlite/apps/query
Protein Structure database searching tools:
http://www.rcsb.org/pdb/

Design of ligands:

ProDrug server:
http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg_beta

Search for experimentally observed interactions of various functional groups in the database of organic and organometalic structures (CCDC):
Main page of the Czech regional center of CCDC is **http://www-xray.fzu.cz/csd/csd.html**
For registration to use the database of organic and organometalic structures follow the instructions at **http://www-xray.fzu.cz/csd/regform.html**

Ligand binding sites:

http://sumo-pbil.ibcp.fr/cgi-bin/sumo-database
http://relibase.ebi.ac.uk/reli-cgi/rll?/reli-cgi/general_layout.pl+home

1. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, *Nucleic Acids Research,* 2000, 28, 235-23.

Lectures - Friday, March 11, afternoon

L12

THE USE OF STEREOSCOPICAL VISUALISATION IN STRUCTURAL BIOLOGY

M. Hušák

Department of Solid State Chemistry, Institute of Chemical Technology Prague, Technická 5, CZ 166 28 Prague

It is not always very easy to understand the structure of complex biological molecules from 2D drawings only. A lot of operations as e.g. interpretation of electron density maps or manual ligand docking require precise orientation in the graphic data. Stereoscopic 3D visualization based on different image delivery for each eye is an useful tool in such situations. In addition the 3D stereoscopic presentation techniques could be very attractive for public presentations as well as for teaching purpose.

Following table summarize several parameters of the most important available 3D stereoscopic visualization methods:

Stereoscopy is supported in several chemical structure-visualization codes. Next table gives an overview of the most interesting one supporting stereoscopy through OpenGL standard:

Technique	**Basic principle**	**Number of observers**	**Quality**	**Cost**
Anaglyph	Different colors for each eye view	No restriction	low	cheap
Active based on CRT monitor	Fast switching between 2 images	1-3	high	cheap
Passive stereoscopic projection	Different light polarization for each view (IMAX)	No restriction	high	expensive
Active projection based on DLP projector	Fast switching between 2 images in DLP chip	No restriction	very high	middle
Autostereoscopic monitor	Multiple viewing zones with different image	1-3	middle	middle

Code name	**Characteristic**	**Source**
DS Viewer Pro	Commercial software for visualization of any type of molecular structure	http://www.accelrys.com/
VMD	Open source software for macromolecular visualization and visualization of MD calculations results	http://www.ks.uiuc.edu/Research/vmd/
O	Electron density maps interpretation	http://www.bioxray.dk/~mok/o-files.html
MolMol	Open source software for macromolecular visualization	http://hugin.ethz.ch/wuthrich/software/molmol/index.html

L13

TRITON: GRAPHIC SOFTWARE FOR MODELLING PROTEIN MUTANTS AND CALCULATION REACTION PATHWAYS

Martin Prokop, Jiří Damborský, Jaroslav Koča

National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

One of the objectives of protein engineering is to propose and construct modified enzymes with improved catalytic activity for substrate of interest. The rational engineering of an enzyme requires to know which amino acid residues of the protein are involved in the catalysis and how to modify them to achieve an increased activity.

The program TRITON is a graphical tool for modelling protein mutants and assessment of their activities [1,2]. Protein mutants are modelled based on the wild type structure by homology modelling using the external program MODELLER. Enzymatic reactions taking place in the mutants active site are modelled using the semi-empirical quantum mechanic program MOPAC. Activities of the mutants can be estimated by evaluation the changes in energies of the system and partial atomic charges of the active site residues during the reaction. The program TRITON offers graphical tools for preparation of input data files, for calculation and for the analysis of generated output data. Implementation ensures overall integrity of consecutive steps of the modelling of mutants and calculation of reaction pahways. The program and its methodology were proven by several studies performed on haloalkan dehalogenase enzymes [3-5]. Calculated results showed qualitative agreement with experimental data.

The program TRITON can run under operating system IRIX, Linux and NetBSD. The software is available at **http://ncbr.chemi.muni.cz/triton/triton.html**.

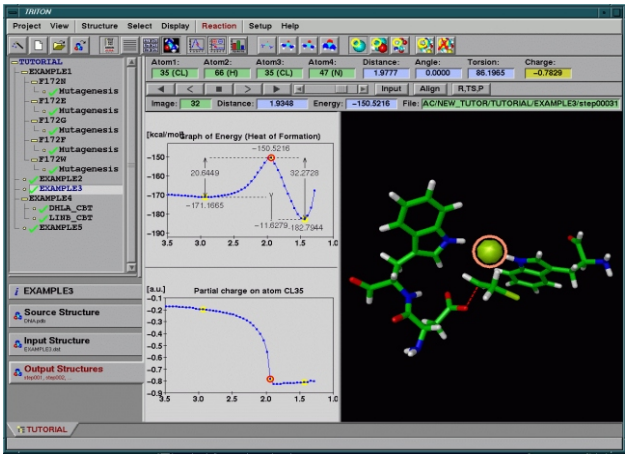

Figure 1. The main window of the program TRITON. The tree list of the projects and folders for access to the data of specific subprojects are situated on the left. The plot of the reaction coordinate versus the potential energy is situated in the middle-top window and the plot of the reaction coordinate versus partial atomic charge on a selected atom in the middle-bottom window. A 3D model of the structure is in the right window.

1. M. Prokop, J. Damborský, J. Koča, *Bioinformatics,* **16** (2000) 845-846.
2. J. Damborský, M. Prokop, J. Koča, *Trends Biochem. Sci.,* **26** (2001) 71-73.
3. J. Damborský, M. Boháč, M. Prokop, J. Koča, *Prot. Engng.,* **11** (1998) 901-907.
4. M. Boháč, Y. Nagata, Z. Prokop, M. Prokop, M. Monincová, M. Tsuda, J. Koča, J. Damborský, *Biochemistry,* **41** (2002) 14272-14280.
5. A. Oakley, M. Klvaňa, M. Otyepka, Y. Nagata, M.C.J. Wilce, J. Damborský, *Biochemistry,* **43** (2004) 870-878.

L14

STRUCTURAL STUDIES OF CARBONIC ANHYDRASE IX

Pavel Mader, Renata Štouračová, Jiří Brynda, Milan Fábry, Magdalena Hořejší, Vlastimil Král, Juraj Sedláček

Department of Recombinant Expression and Structural Biology, Institute of Molecular Genetics, Flemingovo nám. 2, CZ - 166 37 Praha 6, mader@img.cas.cz

Carbonic anhydrase IX (CA IX) belongs to a family of zinc metalloenzymes that catalyze the reversible hydration of carbon dioxide: $CO_2 + H_2O$ $H^+ + HCO_3^-$. These enzymes play a role in pH regulation, CO_2 and HCO_3^- transport, bone resorption, production of biological fluids, ureagenesis, gluconeogenesis, and lipogenesis [1]. CA IX is a unique member of the CA protein family. In contrast to the other isozymes, it has been implicated to play a role in regulation of cell proliferation, adhesion, and malignant cell invasion [2, 3]. This integral membrane protein was shown to function as a cell adhesion molecule with the binding site localized in the extracellular proteoglycan-like (PG) domain; the same site is recognized by the monoclonal antibody M75 thus abrogating cell adhesion in *in vitro* test [2]. Interestingly, CA IX is overexpressed in human epithelial tumours derived from tissues that normally do not express this isozyme, including carcinomas of cervix, lung, kidney, and breast [4]. In contrast, tumours originating from CA IX-positive tissues, such as stomach, tend to have lowered expression of CA IX [4]. CA IX is also overexpressed in von Hippel-Lindau (VHL)-defective tumors and under hypoxic conditions [5]. There are several reasons to consider CA IX as a suitable target molecule for cancer therapy: (i) it is expressed ectopically in various commonly occurring carcinomas, which are rather resistant to conventional therapy; (ii) the antigen is exposed on the cell surface; (iii) normal expression of CA IX is restricted to the luminal epithelia of the alimentary tract, with limited accessibility to immune cells, antibodies and many drugs [6].

The aim of our research is to help elucidate the molecular basis of CA IX involvement in cell proliferation by solving the structure of recombinant forms of extracellular domain of CA IX. Further, we plan to solve structure of CA IX complexes with specific sulfonamide inhibitors, and recombinant scFv fragment derived from monoclonal antibody M75, and finally to identify the putative CA IX binding partner.

To obtain recombinant extracellular CA IX domain, various expression systems were tested, however only Drosophila Schneider 2 Cells yielded sufficient amount (60 $mg.l^{-1}$) of native protein in serum free medium. The protein was purified in 3 chromatographic steps: affinity chromatography on p-aminobenzene sulfonamide agarose, anion exchange chromatography, and gel filtration. The first crystals of CA IX have been obtained using vapour diffusion hanging drop method [Figure 1].

Structure of M75 Fab fragment in complex with its epitope peptide [Figure 2] derived from PG domain of CA IX, which will also be presented, should be helpful in rational drug design of compounds suitable for use in human oncology.

1. W.S. Sly, P.Y.Hu, *Annu. Rev. Biochem.*, 64 (1995) 375–401.
2. J. Závada, Z. Závadová, J. Pastorek, Z. Biesova, J. Ježek, J. Velek, *Br. J. Cancer* 82 (2000) 1808-13.
3. M. Hilvo, M. Rafajová, S. Pastoreková, J. Pastorek, S. Parkkila, *J. Histochem. Cytochem.* 52(10) (2004) 1313-22.
4. M. Leppilampi, J. Saarnio, T. J. Karttunen, J. Kivelä, S. Pastoreková, J. Pastorek, A. Waheed, et al, *World. J. Gastroenterol.* 9 (2003) 1398–1403.
5. C. C. Wykoff, N. J. Beasley, P. H. Watson, K. J. Turner, J. Pastorek, A. Sibtain, G. D. Wilson, et al, *Cancer Res.* 60 (2000) 7075–7083.
6. S. Pastoreková, J. Závada, *Cancer Therapy* 2 (2004) 245-262.

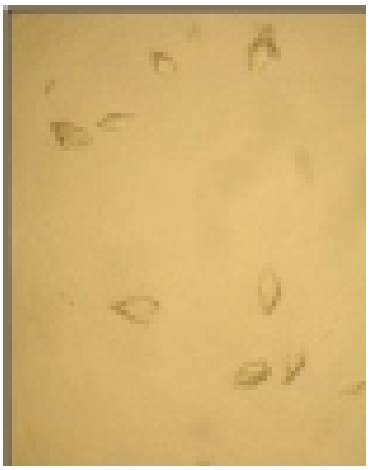

Figure 1: Crystals of CA IX obtained at: 0.2 M Magnesium acetate, 0.1 M Sodium Cacodylate pH 6.5, 20 % PEG 8000; protein concentration: 7 $mg.ml^{-1}$

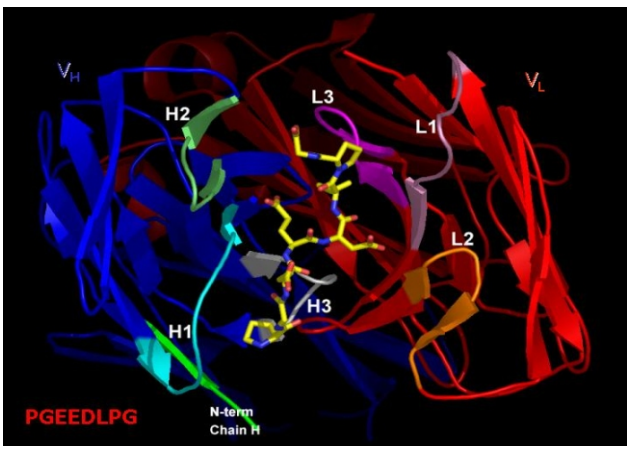

Figure 2:
Fab M75 in complex with the epitope peptide PGEEDLPGEEDL

L15

ARGININE-REPRESSOR PROTEIN FROM E. COLI: A COMPUTATIONAL STUDY

L. Lüley,[1] T. Stockner[2] and R. Ettrich[1]

[1]Laboratory of High Performance Computing, Institute of Physical Biology USB and Institute of Landscape Ecology AS CR, University of South Bohemia, Zámek 136, 373 33 Nové Hrady, Czech Republic, e-mail: luley@greentech.cz

[2]Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 4N1, Canada.

Arginine biosynthesis in *Escherichia coli* is regulated by the negative feedback mechanism at the hand of the hexameric arginine repressor protein (ArgR) with the corepressor L-arginine.

ArgR differs from most other bacterial repressors in functioning as a hexamer. L-Arginine binds to ArgR in the C-terminal domain of the repressor. Binding to operator DNA occurs in the N-terminal domain. The molecular structures of both domains have recently been solved [1, 2], but due to the size of the ArgR and flexible linker region between C and N-terminal domains, the whole structure is still challenge for experimentalists.

Allosteric activation of the hexameric ArgR for specific operator DNA binding appears to involve alteration in its quaternary structure. Current models for activation include subunit assembly and/or domain rearrangements in response to binding of the coeffector L-arginine. It is supposed [3] that subunit assembly plays no role in activation, although communication among subunits of the ArgR hexamer is required for specific DNA binding. It is expected [3] that DNA is also an allosteric effector of ArgR.

We have built several ArgR homology models for *Escherichia coli* which were further refined by molecular dynamics. Currently the effect of binding of L-arginine on the quaternary structure of the repressor protein is examined using long molecular dynamics runs. Our results indicate a potential binding site for DNA on the N-domain of he activated repressor.

This research was supported by FRVS (3303/2005), by the Ministry of Education, Youth and Sports of the Czech Republic (MSM6007665808) and by the Academy of Sciences of the Czech Republic (Institutional research concept AVOZ60870520)

1. Van Duyne, G. D., Ghosh, S., Maas, W. K. & Sigler, P. B. (1996). Structure of the oligomerization and L-arginine binding domain of the arginine repressor of Escherichia coli. *J. Mol. Biol.* **256**, 377-391.
2. Grandori R., Lavoie T.A., Pflumm, M., Tian, G., Niersbach, H., Maas, W.K., Fairman, R. and Carey, J. (1995) ''The DNA-binding domain of the hexameric arginine repressor.'' *J. Mol. Biol.* **254**, 150-162.
3. Szwajkajzer, Danuta, Dai, Lizhong, Fukayama, June W., Abramczyk, Bozena, Fairman, Robert, and Carey, Jannette. 2001. Quantitative analysis of DNA binding by the E. coli arginine repressor. *J. Mol. Biol.* 312:949.

L16

SELECTED INTERACTIONS OF CISPLATIN IN CELLULAR ENVIRONMENT. ACTIVATION OF CISPLATIN AND ITS COORDINATION TO THE DNA BASES AND AMINO ACIDS

Jaroslav V. Burda

Department of Chemical Physics and Optics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Prague 2, Czech Republic.

Solvation effects were studied for cis/transplatin, as well as for some other charged complexes [1]. These investigations were performed using very accurate approach and it was found that $Pt(NH_3)_2(OH)_2$ together with $[Pt(NH_3)Cl(OH)_2]^-$ complex represent the most stable species on the hydration energy surface. A similar hydration energy surface was also calculated for corresponding square planar Pd(II) complexes [2]. Obtained energies (heats of solvation) clearly show that the different metal reactivity (Pt(II) versus Pd(II)) to DNA bases is not caused by different thermodynamical properties of these metals but due to kinetical reaction factors [3]. An important role is also played by the solvent environment which was explored using COSMO model [4].

Further, the interactions of the square planar Pt(II) complexes with DNA bases, base pairs and aminoacids were examined. The influence of the platinum complexes on tautomer forms of guanine and adenine was established [5]. It was found that the adenine imino-tautomer is more stabilized under the influence of charged platinum cations. However, this preference was eliminated for neutral $PtCl_2(NH_3)$ adduct, giving a similar energy difference between imino- and amino-forms of adenine like in the cas of non-metalated adenine tautomers. On the contrary, guanine

keto-form is further stabilized over enol-form under platination.

A base pair enhancement under platination was also explored [6]. Platination causes some geometry distortion in the H-bond arrangement. This effect was also observed for other metals [7-9]. It was shown that no (pairwise) base pair enhancement was observed for AT pair. Nevertheless, platinated guanine base exhibits more firmly connection with cytosine.

Model study on Pt-bridges of the purine DNA bases shows weaker adenine coordination in accord with the other studies in this branch. In dependence on base (guanine or adenine), symmetrical or non-symmetrical arrangement was observed [10, 11].

Recently interactions of cisplatin with amino acides cysteine and methionine were performed, clearly pointing on stronger Pt-cysteine binding in comparison with methionine-containing complex. This correspond to the experimentally known fact that irreversible adducts of cisplatin with cysteine are formed in living body on the contrary to creation of reversible Pt-methionine complexes [12]. Also, a model for general amino acids interaction was studied where glycine molecule was employed for this purpose [13].

1. Burda, J. V.; Zeizinger, M.; Sponer, J.; Leszczynski, J. *J. Chem. Phys.* 2000, *113*, 2224.
2. Zeizinger, M.; Burda, J. V.; Šponer, J.; Kapsa, V.; Leszczynski, J. *J. Phys. Chem. A,* 2001, *105*, 8086.
3. Burda, J. V.; Zeizinger, M.; Leszczynski, J. *J. Chem. Phys.* 2004, *120*, 1253.
4. Burda, J. V.; Zeizinger, M. *J. Comput. Chem.* 2004, submitted.
5. Burda, J. V.; Šponer, J.; Leszczynski, J. *J. Biol. Inorg. Chem.* 2000, *5*, 178.
6. Burda, J. V.; Šponer, J.; Leszczynski, J. *Phys. Chem. Chem. Phys.* 2001, *3*, 4404.
7. Burda, J. V.; Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* 1997, *101*, 9670.
8. Šponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* 1998, *102*, 5951.
9. Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* 1999, *103*, 2528.
10. Burda, J. V.; Leszczynski, J. *Inorg. Chem.* 2003, *42.*, 7162.
11. Zeizinger, M.; Burda, J. V.; Leszczynski, J. *PCCP* 2004, *6*, 000.
12. Zimmermann, T.; Zeizinger, M.; Burda, J. V., *J. Phys. Chem. B* 2005, submitted.
13. Burda, J. V.; Orozco, M.; Luque, F. J., *J. Chem. Phys.* 2005, submitted.

L17

HIERARCHICAL VERSUS NON-HIERARCHICAL PROTEIN FOLDING

Peter Flecker

Johannes Gutenberg Universität Mainz/FRG

The soybean Bowman Birk inhibitor (sBBI) is an attractive model protein for structural and mutational studies directed towards protein structure, folding and functionality. The parent protein of this study was a Lys16Arg and Met27Ile double replacement (termed as rBBI) of sBBI expressed in *E. coli*. A binary arrangement a trypsin- and a chymotrypsin-reactive subdomain in this protein was very useful for a detection of structural/functional irregularities transmitted from one subdomain into another by means of comparative titration and activity determination with trypsin and chymotrypsin. This can be used for a clear cut distinction between an independent (hierarchical) versus a coupled (non-hierarchical) folding of subdomains. In the present work we have studied the significance of a short -turn forming segment near the exit of the trypsin binding loop of BBI. The mutants may be categorised into four classes according to the extent of the observed irregularities. The parent rBBI and one of the mutations distinguished themselves from the other variants by a spontaneous formation of two fully active subdomains upon refolding in solution. The other variants belonging to classes 2-4 were characterised by significant irregularities after refolding in solution. External assistance by means of Trypsin-Sepharose as a template with complementary structure was required with these mutants in order to obtain a fully active state of the variants. The class 2 mutations carrying hydrophobic residues of medium or large size gave only local irregularities in the trypsin-inhibitory region of the molecule and indistinguishable amounts and activities against chymotrypsin after refolding in solution and on the template. The amounts and the activity of the chymotrypsin-inhibitory subdomain were decreased or even completely abolished with the class 3 variants in accord with more global irregularities. The class 4 variants distinguished themselves from the class 2-3 variants by an unusual 2:1 stoichiometry with trypsin and chymotrypsin after template-directed folding. The mutants displaying local or more global irregularities could be distinguished by means of molecular modelling studies in accord with the experimental findings. The models seemed to be in line electrostatic interactions across the interdomain boundary and controlling a coupled folding of subdomains in accord with non-hierarchical models of protein folding.

L16

COMPUTATIONAL ANALYSIS OF ENZYME ACTIVE SITE ROUTES ANATOMY

Petřek Martin[1], Otyepka Michal[2], Banáš Pavel[2] and Damborský Jiří[1]

[1]*National Centre for Biomolecular Research, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic*

[2]*Department of Physical Chemistry, Palacký University Olomouc, tř. Svobody 26, 771 46 Olomouc, Czech Republic*

Enzymes are catalysts of biological systems that determine the patterns of chemical transformations. The enzymatic reaction is complex, consisting of substrate binding, one or more chemical reactions and product release. It is necessary to consider all these steps in study of enzymatic activity. The rate-limiting step can be substrate binding or product release in some instances. The molecular dynamics simulation is effective tool to study flexibility of biomolecules. The goal of this work was to develop a simple, quick and flexible technique to monitor radius of a tunnel leading to the active site during molecular dynamic simulation. The presented algorithm utilizes alpha shape theory in combination with grid search algorithm [1].

The haloalkane dehalogenases (EC 3.8.1.5) are microbial enzymes cleaving a carbon-halogen bond in a broad range of halogenated compounds [2]. These enzymes were used for the case study. The X-ray structures of three haloalkane dehalogenases are known: LinB from *Sphingomonas paucimobilis* UT26 [3], DhlA from *Xanthobacter autotrophicus* GJ10 [4] and DhaA from *Rhodococcus* sp. NCIMB13064 [5]. The active site is buried between the main domain (the hydrolase fold domain) and the cap domain. The several possible access routes were denoted as main tunnel, upper tunnel and slot. The enzymes LinB, DhaA and DhlA differ in number of routes as well as in radius of each tunnel leading to the active site. LinB has the most open active site accessible through all three tunnels, DhaA has accessible upper tunnel and slot and DhlA has only one access via main tunnel [6]. Number and size of entrance tunnels seems to correlatewith the rate limiting step. The rate-limiting step of DhlA and DhaA is a product release, while the chemical reaction is the rate-limiting step of LinB [7]. The amino acid Leu177 of LinB localized in a gorge of upper tunnel was mutated to study the influence of upper tunnel radius upon enzyme activity. The substitution in analogous amino acid of DhaA, Cys176, was obtained by direct evolution method. Cys176Tyr mutation increased enzyme activity with 1,2,3-trichloropropane 3-times [8]. The developed program was used to determine the preferred route to the active site (Fig 1.) in the molecular dynamic simulations of wild type enzymes DhaA, DhlA, LinB and Cys176Tyr mutant of DhaA in complexes with 1,2,3-trichloropropane. The free enzymes differ in radii of access route gorge and also in mean path cost (Table 1). DhlA has the narrowest and expensive active site route while DhaA and LinB tunnels have higher radii and are cheap. Cys176Tyr mutation (DhaA) causes change in preferred (low cost) route form upper tunnel (WT-DhaA) to slot (Cys176Tyr-DhaA).

Table 1 Mean radii (Å) of active site access routes and mean route costs. The normalized cost function is defined as $C_n(P) \quad \frac{1}{N} \quad \frac{1}{r_{\max}(x)}$, where $r_{\max}(x)$ is the maximal radius of a hypothetical ball that can be inserted into the node x just touching the protein surface. The small constant is here only for technical purposes to get rid of a singularity of the function in

enzyme	radius	cost
DhlA	0.3 0.1	2.1 2.0
LinB	0.7 0.3	1.6 1.2
DhaA	0.9 0.2	0.5 0.1

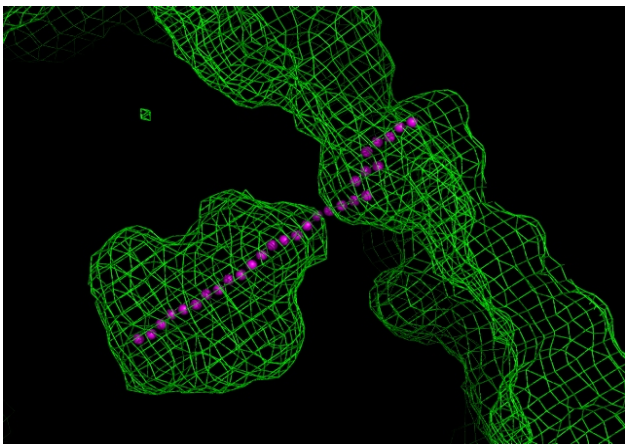

Figure 1. Active site of LinB enzyme represented by solvent accessible surface (wire model) and the easiest way from the active site (magenta points), calculated at 60 x 60 x 60 grid

1. Dijkstra, E.W., A note on two problems in connection with graphs. Numer. Math., 1959. 1: p. 83-89.
2. Damborsky, J., et al., Structure-specificity relationships for haloalkane dehalogenases. Environmental Toxicology and Chemistry, 2001. 20(12): p. 2681-2689.
3. Nagata, Y., et al., Purification and characterization of haloalkane dehalogenase of a new substrate class from a g-hexachlorocyclohexane-degrading bacterium, Sphingomonas paucimobilis UT26. Applied and Environmental Microbiology, 1997. 63: p. 3707-3710.
4. Keuning, S., D.B. Janssen, and B. Witholt, Purification and characterization of hydrolytic haloalkane dehalogenase from Xanthobacter autotrophicus GJ10. Journal of Bacteriology, 1985. 163: p. 635-639.
5. Kulakova, A.N., M.J. Larkin, and L.A. Kulakov, The plasmid-located haloalkane dehalogenase gene from

Rhodococcus rhodochrous NCIMB 13064. Microbiology, 1997. 143: p. 109-115.

6. Otyepka, M. and J. Damborsky, Functionally relevant motions of haloalkane dehalogenases occur in the specificity-modulating cap domains. Protein Science, 2002. 11(5): p. 1206-1217.
7. Chaloupkova, R., et al., Modification of activity and specificity of haloalkane dehalogenase from Sphingomonas paucimobilis UT26 by engineering of its entrance tunnel. Journal of Biological Chemistry, 2003. 278(52): p. 52622-52628.
8. Bosma, T., et al., Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. Applied and Environmental Microbiology, 2002. 68(7): p. 3582-3587.

Lectures - Saturday, March 12, morning

L17

WHERE THE ASYMMETRY IN GENE EXPRESSION STARTS

Imrich Barák

Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, 845 51 Bratislava 45, Slovakia; phone: ++421 2 5930 7418, fax: ++421 2 5930 7416, e-mail: imrich.barak@savba.sk

A general problem in developmental biology is the process by which cells of one type give rise to two dissimilar daughter cells. Although, these two daughter cells are genetically identical and contain identical chromosomes they differ from each other morphologically, physiologically and also they can have a different fate. *Bacillus subtilis* sporulation represents an excellent model to study one of the simplest differentiation process in great detail. Many interesting questions have been raised in the course of studies of this relatively simple cell differentiation process. Firstly, how does the cell decide when to sporulate and what are the signals? Secondly, what is the mechanism by which morphological asymmetry arises and how is the sporulation septum formed? Thirdly, how is the hierarchical regulatory cascade of sigma factors switched on and how are changes in gene expression coordinated with changes in cell morphology? Several decades of study of this process have provided insight into cell cycle regulation and development. Important advances were achieved in our understanding of asymmetric gene expression during spore formation. The work in our laboratory is emphasized on development stages that lead to asymmetric septum formation and especially to activation of the first compartment specific sigma factor - ^{F}.

Work in author's laboratory is supported by grant 2/1004/22 from the Slovak Academy of Sciences and Wellcome Trust Project and Collaborative Research Initiative Grants 066732/Z/01/Z.

L19

RNA CONFORMATIONAL CLASSES

David Micallef[1], Laura Murray[2], Helen M. Berman[1], Jane Richardson[2], Zdeněk Morávek[3], Bohdan Schneider[4]

[1]Rutgers, The State University of New Jersey, Department of Chemistry and Chemical Biology, NJ-08854, USA. [2]Duke University, Durham NC, USA. [3]Faculty of Mathematics and Physics, Charles University, Ke Karlovu 2, Prague, Czech Republic. [4]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, CZ-16602 Prague, Czech Republic, bohdan@rcsb.rutgers.edu

RNA conformations have been analyzed by different knowledge-based approaches. Duarte *et al*. [1] reduced nucleotide six dimensional (6-D) conformational space to two pseudo torsions around pseudo bonds P_i-C4'$_i$ and C4'$_i$-P_{i+1}. This approach, while appealing by its simplicity, neglects correlations between nucleotide torsion angles and therefore fails to recognize NA conformational families. Hershkovitz *et al*. [2] have analyzed nucleotide torsional space. An automated pattern-recognition approach concentrated on binning of torsions with the highest variability, namely , , , and . The analysis of correlations *within* a nucleotide precludes analysis of correlations of torsions *between* nucleotides, especially the key correlation between the torsions at the phosphodiester link.

Murray *et al*. [3] have studied a ribose-to-ribose unit called "suite", not a nucleotide. Ingenious analysis of two 3-D torsional spaces of "heminucleotides", — — and — — , have lead to identification of 47 distinct conformations of suites. Schneider, Morávek and Berman [4] have studied relationships between nucleotide torsion angles with the main emphasis on the torsions at the phosphodiester link, and , and considering the other main descriptors of RNA conformations, namely , de-

scribing ribose pucker, , and , the torsion describing base orientation around the glycosidic bond.

The approaches of Murray *et al*. [3] and Schneider, Morávek and Berman [4] are to some extent complementary and their combination brings a more detailed knowledge about the RNA conformational behavior. The talk presents across-the-database analysis of RNA structures combining these two approaches. The original data containing all dinucleotides (di-nt) from the NDB public archive were filtered by crystallographic criteria, resolution and temperature factor, and by checking their stereo chemical quality, mainly atom-atom close contacts as described by Murray *et al*. [3]; the resulting data matrix contains about four thousand di-nt fragments. The fragments were divided into two groups, the majority of "A-RNA like" with ~ ~300°, and the "non-A-like" rest. In each group, seventeen 3-D scatter grams ("maps") were analyzed by the Fourier averaging technique developed previously [5]. The information content of the maps was estimated by Shannon's entropy, $S \sim - \sum P_i \ln P_i$, where P_i is a fraction of di-nt fragments which can be assigned to the i-th peak. The maps with the highest information content (the largest S) were then used to cluster the data points. Conformational clusters were independently determined using protocols described in both [3] and [4] and their geometries compared and validated. The final RNA conformational families were then determined as consensus geometries determined by [3] and [4].

BS is grateful to support by grant LN00A032 from the Ministry of Education of the Czech Republic.

1. Duarte, C.M., Wadley, L.M. and Pyle, A.M. RNA structure comparison, motif search and discovery using a reduced representation of RNA conformational space. *Nucl.Acids Res.*, 31 (2003) 4755-4761.
2. Hershkovitz, E., Tannenbaum, E., Howerton, S.B., Sheth, A., Tannenbaum, A. and Williams, L.D. Automated identification of RNA conformational motifs: Theory and application to the HM LSU 23S rRNA. *Nucl.Acids Res.*, 31 (2003) 6249-6257.
3. Murray, L.J.W., Arendall III, W.B., Richardson, D.C. and Richardson, J.S. RNA backbone is rotameric. *Proc.Natl.Acad.Sci USA*, 100 (2003) 13904-13909.
4. Schneider, B., Morávek, Z. and Berman, H.M. RNA conformational classes. *Nucl. Acids Res.* 32 (2004) 1666-1677.
5. Schneider, B., Cohen, D.M., Schleifer, L., Srinivasan, R., Olson, W.K. and Berman, H.M. A systematic method to study the spatial distribution of water molecules around nucleic acid bases. *Biophys.J.*, 65 (1993) 2291-2303.

L20

NEW NMR METHODS FOR CONFORMATIONAL STUDIES OF BIOMOLECULES

R. Fiala and L. Žídek

National Centre for Biomolecular Research, Masaryk University, Kotlářska 2, CZ-611 37 Brno, Czech Republic

In the last three decades, high resolution NMR has taken its place beside X-ray crystallography as a means to obtain high-resolution structures of biomolecules. Traditionally, the structure determination by NMR relies on the measurement of interatomic distances that are accessible by the nuclear Overhauser effect (NOE) and the determination of torsional angles that can be derived from scalar coupling constants. The analysis of NMR spectra consist of two main steps – identification of the signals in the spectra corresponding to individual atoms in the molecule (assignment), and extracting the constraints that can be used for the calculation of the structure.

Originally, the NMR studies of biomolecules were based almost exclusively on the resonances of ^{1}H [1], since the nucleus provides the best sensitivity due to the high gyromagnetic ratio and a high (99.98%) natural abundance. However, the relatively low dispersion of the proton NMR spectra led to severe signal overlap even with molecules of moderate complexity. This problem was overcome by the introduction of multidimensional spectroscopic techniques and by the use of samples isotopically enriched by ^{13}C and ^{15}N. These advances made the assignment procedure more reliable by following the through-bond interactions rather than conformationally dependent interatomic distances. The structural constraints, however, are still based mainly on the ^{1}H -^{1}H NOE and J-couplings. While their values can characterize the local structure with very good accuracy, they provide virtually no information on the relative orientations of distant regions of a molecule (NOEs are typically observable only for the distances shorter than 6 Å). The current research in NMR methods therefore concentrates on obtaining additional structural constraints, preferably those providing an absolute configuration with respect to a fixed common frame of reference.

Recently, high resolution NMR started using interactions that were previously the domain of solid-state NMR and in liquids manifested themselves only in the relaxation properties, namely residual dipolar couplings (RDC) [2] and chemical shift anisotropy (CSA) [3]. In order to obtain structurally dependent RDCs and CSA in the high-resolution liquid NMR spectra, it is usually necessary to introduce certain degree of orientation into the solution of the biomolecule. While the high magnetic field itself induces certain amount of orientation of the molecules due to their inherent anisotropies in magnetic susceptibility, an alignment medium has to be added to the sample in most cases to make the measurement of RDCs a chemical shift changes due to CSA practically feasible. A great variety of alignment media has been proposed so far with lipid bicelles, filamentous bacteriophage particles and polyacrylamide gels

being the most commonly used. In the partially oriented samples, the effects of dipolar couplings and CSA are typically scaled down by about three orders of magnitude, allowing thus their measurement with sufficient accuracy while retaining the simplicity and relaxation properties of liquid-phase spectra.

In practice, the RDCs are measured as changes of scalar couplings. Their values are a function of orientation of the internuclear vector with respect to the magnetic field, which represents here the reference direction common to all parts of the molecule. The constraints obtained from RDCs can be complemented by the changes of chemical shift that arise from CSA. The changes in the chemical shifts between and isotropic and oriented phases provide information on the orientation of chemical shielding tensors relative to the molecule's alignment frame. Angular information can also be obtained from the cross-correlated relaxation interference [4]. In this case, no partial alignment of the sample is necessary. Angles between two interatomic vectors can be determined from the measured effects of the cross-correlated relaxation directly, without the need of an empirical calibration curve. The interference between the dipolar and CSA relaxation mechanisms is also used in TROSY (transverse relaxation optimized spectroscopy) techniques [5] that, while not providing any additional structural constraints by themselves, greatly expanded the range of molecules amenable to studies by the methods of NMR spectroscopy towards higher molecular weights.

1. K. Wüthrich: NMR of Proteins and Nucleic Acids, New York 1986, Wiley.
2. J. H. Prestegard, C. M. Bougault, and A. I. Kishore, *Chem. Rev.* 104 (2004) 3519-3540.
3. G. Cornilescu and A. Bax, *J. Am. Chem. Soc.* 122 (2000) 10143-10154.
4. B. Reif, M. Hennig, and C. Griesinger, *Science* 276 (1997) 1230-1233.
5. K. Pervushin, R. Riek, G. Wider, and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* 94 (1997) 12366-12371.

L21

EFFECT OF MOLECULAR VIBRATIONAL MOTION ON NMR PARAMETERS IN ZWITERANION LALANINE-LALANINE

V. Sychrovský, P. Bouř, V. Špirko, J. Šebestík, M. Buděšínský

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Praha

Calculation of NMR shifts as well as spin-spin coupling constants [1] was performed for the zwiter-anionic form of dipeptide LAla-LAla including the effect of water solvent. Experiment was done for LAla-LAla molecule labelled by ^{13}C, ^{15}N isotopes, what enables the complete NMR spectra assignment. The NMR parameters were calculated for different molecular geometries of a di-peptide in order to systematically sample the conformational space corresponding to the mutual rotation of both amino acids along the peptide bond. The analysis of surfaces (i.e. NMR properties calculated on the grid of smoothly varying geometry parameters) of calculated NMR parameters revealed their distinct partitioning according to their character. The single-dimensional character of NMR parameter is attributed to the autonomous behaviour within single amino acid unit, while for the two-dimensional NMR parameters the molecular motion of individual amino acids are coupled. The vibration correction for each NMR parameter was calculated as an average value of the parameter over the anharmonic vibrational molecular motion. Prediction of molecular geometry from the calculated values of NMR parameters is discussed.

1. Sychrovsky V., Grafenstein J., Cremer D.; J.Chem. Phys. 113(9), 3530-3547, 2000.