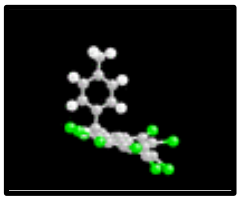

SESAME WORKSHOP/SCHOOL ON

BIOINFORMATICS & STRUCTURAL MODELING

3-8 SEPTEMBER 2001

Organised By
Z. Sayers, Sabanci University, Turkey, I. Sagi, Weizmann Institute, Israel, M. Vlassi, N.C.S.R. "Democritos", Greece

Sponsored By

UNESCO

Sabancı Üniversitesi

TUBITAK

Israeli Academy of Sciences and Humanities

Table of Contents

*See list of participants for institutional affiliations of speakers

Report on the SESAME Workshop/School on Bioinformatics and Structural Modelling

Istanbul, Turkey, September 3- 8, 2001

I. *Introduction*

During the Sesame (**S**ynchrotron-light **E**xperimental **S**cience and **A**pplications in the **M**iddle-**E**ast) Structural Molecular Biology (SMB) meeting in Cyprus in December 2000 (see the SESAME web site www.sesame.org.jo for a report) ideas for future projects were explored, keeping in mind the importance of developing a local user base for SESAME and establishing strong collaborations among scientists from the region. The idea for a bioinformatics workshop was conceived not only because bioinformatics is a fast developing field which can be used to complement structural studies carried out on synchrotrons, but also because bioinformatics tools are on the World Wide Web and hence accessible to people even in parts of the world where other facilities may not be present. People located in different countries can share data and analysis skills, establish collaborations and develop projects together in the electronic medium at a relatively low cost.

The 'Bioinformatics and Structural Modelling' workshop/school was hosted at Sabanci University (SU) near Istanbul and Drs. Zehra Sayers (SU, Turkey), Irit Sagi (Weizmann Institute, Israel) and Metaxia Vlassi (NCSR "Demokritos", Greece) were co-organisers. These three scientists met for the first time at the Cyprus SMB workshop. With help from Dr. H. Winick, who is one of the driving forces behind the SESAME project, it was possible to get sponsorship from UNESCO. TUBITAK (Turkish National Scientific and Technical Research Council), Sabanci University and the Israeli Academy of Sciences and Humanities also provided funds which made it possible to cover local expenses and also to provide travel allowances for all participants.

II. *Executive summary*

Bioinformatics is a fast developing where computational tools are applied to problems and analyses in a wide range of fields from molecular biology and medicine to structural biology. The workshop/school was at post-graduate level, and the twenty-four students from different countries (including Armenia, Bosnia and Herzegovina, Egypt, Greece, Israel, Jordan, Morocco, Poland, Turkey and the USA) came from diverse disciplines such as medicine, computer science, physics and biology.
The lecturers included C. Baysal (Sabanci University, Turkey), I. Berezovsky (Weizmann Institute, Israel), E. Eliopoulos (University of Athens, Greece), A. Gürsoy (Bilkent University, Turkey), D. Lancet (Weizmann Institute, Israel), S. Mobashery (Wayne State University, USA), R. Najmanovich (Weizmann Institute, Israel), P. Rizkallah (Daresbury Laboratory, UK) I. Sagi (Weizmann Institute, Israel), Z. Sayers (Sabanci University, Turkey), U. Sezerman (Sabanci University, Turkey), M. Vlassi (NCSR "Demokritos",

Greece), S. Wakatsuki (Photon Facotry, Japan), H. Winick (Stanford University,USA), A. Yonath (Weizmann Institute, Israel), and P. Zielenkiewicz (Polish Academy of Sciences, Poland).

The workshop/school was designed to have lectures in the morning and practical sessions in the afternoon. Lectures covered introduction to experimental techniques for structure determination based on synchrotron radiation, e.g. X-ray crystallography (M. Vlassi, P. Rizkallah), small angle scattering (Z. Sayers) and EXAFS (I. Sagi). Complementarity of results obtained with different techniques (Z. Sayers) as well as emergence of novel methods for dynamic studies (I. Sagi) was emphasised. Analysis of the crystallographicaly determined structures of bacterial ribosome and ribosome-antibiotics complexes (A. Yonath) highlighted possibilities for drug design. Structure based modelling and molecular dynamics simulations of biological systems at scales from those of catalytic sites of enzymes to large structures such as surface components of bacteria (S. Mobashary) provided examples for how 3D modelling tools may be used for bridging experimental data and developing alternative strategies for rational drug design. A survey of a target oriented structural genomics project using synchrotron X-ray protein Crystallography was also given (S. Wakatzuki). These lectures focused on results obtained using various synchrotron radiation sources around the world e.g. APS, Photon Factory, ESRF, DESY, Daresbury, SSRl and Brookhaven.
Talks on tools for utilisation of data from the Human Genome project (D. Lancet) and new strategies for structural motif recognition (I. Berezovsky) introduced different aspects of analysis of DNA and protein sequence data bases. Secondary structure predictions (E. Eliopoulos), molecular dynamics simulations (C. Baysal), homology modelling U. Sezerman), threading (P. Zielenkiewcz) and docking methods (R. Najmanovich) were presented both in lectures as well as used as topics for practical exercises.
Practical sessions were carried out in a computer lab at SU equipped with 15 PCs and the necessary peripherals (e.g. printers, multimedia facilities etc.). Each PC was used by a workgroup of two students, that allowed everybody to gain direct experience.

Practical sessions were computer based exercises in the form of tutorials prepared as web pages that were accessed through the local website developed for the purposes of the workshop. The topics of each day's tutorial were related to the morning lectures and covered use of web-based resources related to macromolecular structure: search and retrieval (e.g. ENTREZ, SRS) of data from various databases (such as PFAM, PROSITE, MIME, SwissProt and the PDB structural database), sequence similarity searches (e.g.BLAST), sequence alignments (CLUSTALW, t-COFFEE), secondary structure predictions (PHD, GOR-IV and Joint prediction: by Eliopoulos), sequence-prototype based prediction of the closed loops (developed by I. Berezovsky), homology modelling (SwissModel and 3D-PSSM), threading (developed by P. Zielenkiewicz) and ligand-protein contact analysis (developed by R. Najmanovich). Students were also introduced to 3D visualisation programs (e.g. rasmol and Swiss-pdb viewer) which were used to display, compare and analyse known 3D structures and models. One protein with unknown 3D structure belonging to a well-studied protein family served as a test case for all practical sessions.

Availability and rapid accumulation of DNA sequence data necessitate heavy use of electronic means for storage, transmission and analyses. Effective use of sequence-based information reveals previously unknown relationships - ranging from the identification of unknown genes to the prediction of protein structures. Parallel to improvements in experimental methodologies for faster structural data collection and analysis, computational tools are now being developed aimed at predicting structures and elucidation of structure-function relationships as well as for drug design. Today Web-based facilities such Entrez at NCBI or Biology Workbench at SDSC, worldwide accessible databases e.g. Swiss Prot, PDB as well as web-based tools (BLAST, CLUSTAL W), and servers for secondary structure predictions (e.g. PHD, JOINT, GOR-IV) and 3D modelling (e.g. SwissModel, 3d-pssm) are at the disposal of scientists from many countries. The ease of access to data and facilities provide opportunities for competitive research regardless of the location of the scientist.

The workshop/school confirmed that bioinformatics and computational biology are emerging fields which could be developed in SESAME countries as a prelude to the synchrotron activities. This would help to build a user base with background in structural analysis and encourage establishing collaborations in the region.

It is also important to note that the workshop/school provided a medium for future collaborations among scientists and new ideas for directions to follow. Since then Drs. Z. Sayers (Turkey) and E. Charvalos (Greece) have prepared a joint research proposal to be submitted to The Turkish Scientific and Technical Research Council and The Ministry of Development in Greece. Also a joint proposal from Drs. M. Vlassi (Greece) and A. Soukri (Morocco) is being submitted to NATO.
As a possible future direction a proposal, for a computational facility that will be based in Jordan at the SESAME site, to be taken to representatives of respective countries and to the SESAME Council, is also being prepared.

The great success of the workshop has opened a window of new opportunities to expose young scientists to cutting edge scientific projects and to train them in these fields. The nice and relaxed scientific atmosphere has promoted scientific interactions and new hope for the future of the Middle East region.

III. *Abstracts*

New chapters in protein structure

Dr. Igor Berezovsky, The Weizmann Institute of Science, Structural Biology Department, Rehovot, 76100 Israel

In a globular protein the polypeptide chain returns to itself many times, making numerous chain-to-chain contacts. The stability of these contacts is maintained primarily by van der Waals interactions. The loops closed by the van der Waals interactions provide a principally novel view of protein globule organization: the loop-n-lock structure. This opens a new perspective in understanding protein folding as well: the consecutive looping of the polypeptide chain and the locking of the loop ends by tight van der Waals interactions. Loop-n-lock elements are found to have a preferred size of about 30 amino acid residues and these elements are connected consecutively into the linear arrays. Such linear arrangement has obvious implications for protein evolution and cotranslational folding. Search for families of sequences of the size
25-30 revealed several prototypes, each member of such family is found to be folded into the loop-n-lock structure's element.

Selected references:

1. Berezovsky, I.N., Grosberg, A.Y. & Trifonov, E.N. 2000. Closed loops of nearly standard size: common basic element of protein structure. FEBS Letters 466, 283-286.
2. Berezovsky, I.N. & Trifonov, E.N. 2001. Evolutionary aspects of protein structure and folding. Mol.Biol. 35, 233-239.
3. Berezovsky, I.N. & Trifonov, E.N. 2001. Van der Waals locks: loop-n-lock structure of globular proteins. J. Mol. Biol. 307, 1419-1426.
4. Berezovsky, I.N. & Trifonov, E.N. 2001. Loop fold nature of globular proteins. Protein Engineering 14, 403-407.
5. Trifonov, E.N., Kirzhner, A., Kirzhner, V.M. & Berezovsky, I.N. 2001. Distinct stages of protein evolution as suggested by protein sequence analysis. J. Mol. Evol. 53, in press.
6. Berezovsky, I.N., Kirzhner, A., Kirzhner, V.M. & Trifonov, E.N. 2001. Protein folding: looping from the hydrophobic nuclei. PROTEINS: Structure, Function, and Genetics, in press.
7. Berezovsky, I.N. 2001. Protein structure: chapters which have been missed. Recent Research Developments in Protein Folding, Stability and Design. (Eds. M.M. Gromiha and S. Selvaraj) Vol.1, Transworld ResearchNetwork, Trivandrum, India, in press.

Pathway Analysis Tool for Integration and Knowledge Acquisition

Rengul Atalay and Atilla Ozturk, Departments of Molecular Biology and Genetics and Computer Engineering and Information Science, Bilkent University, Bilkent, Ankara, Turkey

As the focus of the scientific community shifts from genomics to proteomics, data produced about cellular processes accumulates rapidly. It is essential to develop tools for representing, storing and visualizing this data. We define a regular semantic for describing cellular processes as a graph of actors and events. This model is designed to facilitate knowledge integration; accumulation and analysis. We present an integrated environment, named PATIKA, based on this model. PATIKA is composed of a server-side database and client-side editors to provide an integrated, multi-user environment for manipulating and visualizing cellular processes. PATIKA provides fundamental tools for modeling and integrating cellular processes with a user-friendly interface. Furthermore, it features powerful querying options, functional computation support and automated pathway layout. PATIKA is currently in development and first release is scheduled for Spring 2002.

Protein Secondary Structure Prediction and Modeling

Elias Eliopoulos, Genetics Laboratory, Dept. of Agricultural Biotechnology, Iera Odos 75, Athens 118 55, Greece.

Proteins owe their function to their unique tertiary structure. This is determined mainly by the primary structure (aminoacid sequence) and can be considered as an aggregate of secondary structure elements, alpha helices, beta strands and turns. If secondary structure could be accurately predicted from the primary sequence the problem of determining the tertiary structure is reduced to a packing problem i.e. to correctly pack secondary structure elements. Approaches to predict secondary structure from the aminoacid sequence fall into categories. Empirical and statistical techniques first introduced in the 1970 are based on first order correlations. Alternative structure prediction techniques include neural networks, homology modeling and lattice models. Although the accuracy of predictions remains not fully satisfactory, protein class prediction and motif determination can readily improve the results. Many software packages accessible through the web can provide secondary structure prediction and modeling of unknown protein sequences.

Harvesting the human genome: a world-wide endeavour

Doron Lancet, The Crown Human Genome Center, The Weizmann Institute of Science, Rehovot 76100 Israel

As the Human Genome Project progresses and the First Draft is nearly finished, a most urgent need is data integration and utilization. At our Genome Center, serving as a national laboratory for Israel, we have devised a model for a role for smaller projects worldwide, in which we combine gene discovery, integrated database development and a focused effort in the field of DNA microarrays. All three activities revolve around the development of a strong capacity in computational genomics and bioinformatics.

For gene discovery, our main effort is through collaborations with the Israeli medical community. We strive to utilize the unique attributions of populations in Israel, constituting many genetically-defined ethnic groups. Six gene discovery projects are underway, some already culminated in the identification of a gene underlying a monegenic disease. We depart from a linkage map, and utilize the power of sequence data mining and integration to identify gene candidate. Large scale DNA sequencing is used to reveal mutations within all identified exons.

In the Field of data integration, we have developed three software tools. The first is GeneCards, a novel functional genomics compendium combining automated data mining and context-related navigation support (Trends in Genetics 13: (1997). Bioinformatics 14:656-664 (1998); URL: bioinfo.weizmann.ac.il/cards). It automatically identifies new HUGO approved gene symbols, extracts relevant information from multiple public databases, and creates a Card for each gene. The second toll is Unified DataBase (UDB), in which novel concepts og genome-wide map and sequence integration are implemented (Genome Digest 4(3): 15 (1997), URL: bioinfo.weizmann.ac.il/udb). It merges method-specific genome maps with genomic sequence information (Sequence Based Repositioning). The third tool is GESTALT, (Bioinformatics 2000 May; 16(5): 482-483) a GEnomic Sequence Total Analysis and Lookup Tool (http://bioinfo.weizmann.ac.il/GESTALT). It constitutes a workbench for automatic integration and visualization of large-scale genomic sequence analyses.

For DNA arrays, we have acquired and Affymetrix GeneChip system, which has already been successfully used with human, mouse and yeast expression arrays, as well as with

P53 mutation array In parallel, we have established the complementary array spotting and scanning technology, to meet additional species and gene group requirements. The data are integrated with our in-house developed GeneCards.

Applications of Molecular Modeling and Dynamics Simulations to Systems of Importance to Biological Sciences

S. Mobashery, Department of Chemistry, Wayne State University, Detroit, MI 48202

Computational capabilities and development of sophisticated softwares in the past 5-10 years have made giant strides. It is now possible to use computational analyses for biological systems that are intractable to study by other techniques. These are often extremely large systems whose structural aspects are difficult to investigate by other means (NMR, X-ray, force microscopy) or they are dynamic systems that process events at fleeting timescales. This presentation will concentrate on a few such applications and will demonstrate how dynamics simulations will provide valuable information in understanding of important biological processes.

Drugs and their protein targets: Side chain flexibility

Rafael J.Najmanovich, Plant Sciences Department, Weizmann Institute of Science, Israel

Detailed knowledge of ligand binding at the molecular level is important in order to understand most cellular processes. The exponential growth of the Protein Data Bank (PDB) creates the opportunity for bioinformatics to play a crucial role in the study of ligand binding both in terms of mining the information present in the PDB as well as in creating docking prediction algorithms.

Analysis of side chain flexibility is an important step in order to include side chain flexibility on docking predictions. For that purpose, we built a database of Holo and Apo protein pairs present in the PDB to study the effect of ligand binding on side chain flexibility. Our study reveals that in 85% of studied binding sites, less than 3 side chains undergo conformational changes. Furthermore, there are clear differences in the propensities of different amino acid side chains to undergo structural changes upon ligand binding. A similar flexibility scale was obtained by analyzing side chain flexibility in the vicinity of point mutations. This study suggests that the side chain flexibility scale is general rather than a consequence of point mutations or ligand binding.

Our group developed web tools that can aid the study of interactions between drugs and protein targets. Four such tools are LPC, CSU, MutaProt and LigProt. LPC and CSU characterize the atomic contacts between ligands and proteins (LPC) and between structural units (CSU) (http://bioinfo.weizmann.ac.il:8500/oca-bin/lpccsu). The databases used to study side chain flexibility are also available through the WWW, allowing for searching, analysis of atomic contacts and visualization of pairs of superimposed structures present in the PDB that differ in either the presence of ligands (LigProt) (http://www.weizmann.ac.il/ligprot) or point mutations (MutaProt) (http://bioinfo.weizmann.ac.il/mutaprot). These tools are useful not only in bioinformatics studies but also in experimental molecular biology, for example, to guide amino acid substitutions.

We are presently modifying our chemical-compatibility and surface-complementarity based docking algorithm (LigIn). We are using a genetic algorithm for the search and

optimization of ligand position and conformation (ligand flexibility). Side chain flexibility is modeled by a machine learning approach in conjunction with the use of side chain rotamer libraries.

Similar/Different Structure/Function from Similar/Different Sequences

Pierre Rizkallah, Daresbury Laboratory, Warrington, WA4 4AD, U.K.

Structural molecular biology has provided theoreticians with a wealth of information to predict unknown structures by homology with known ones. This has proved successful in routine usage, but certain limitations apply:

- When sequence identity drops below 50%, it becomes more difficult to predict the structure. Below 20% sequence identity, structural similarity becomes random.

- Structural similarity has not been a very good predictor of function. Often, point mutations have dramatic consequences on the Biochemistry involved. An extreme form of functional modification is the use of disabled enzymes as structural proteins for eye lenses (crystallins).

In this presentation, two case studies will be reviewed:

- Monocot mannose binding lectins, from flower bulbs of amaryllidaceae (snowdrop, amaryllis, and daffodil), lilleaceae (bluebell) and alliaceae (garlic), are defence proteins of non-immune origin. Their ability to bind mannose makes them able to cross-link between gut cells of herbivores and foraging insects, deterring them from feeding. The lectins have a conserved structural fold, a triangular prism of beta sheets, with a binding site on each face. The sequence is also highly conserved, with identity over 50% and similarity even higher. Yet their ability to inhibit *in vitro* HIV infection of human T-cells varies significantly. Daffodil lectin (NPL) is 10 times more effective than bluebell lectin (SCA), while garlic lectin (ASA) is ineffective. Detailed structural comparisons reveal the underlying reason for this functional divergence to be the quaternary structure. A monomer exchanges the last strand of the beta barrel with another molecule, making a tight canonical dimer. The dimers coalesce into tetramers, in crystals of NPL and SCA, but not those of ASA. The coalescence is achieved by a strong hydrophobic surface patch interaction in NPL, which persists in solution. The equivalent surface patch in SCA is mutated to have polar residues, and the coalescence occurs via a salt bridge. These residues are mutated to aromatic side chains in ASA, which are bulky and keep the dimers apart. Tetramer formation

extends the overall surface area that can be mustered to cross link mannose containing cells or viruses such as HIV. The varying 'blanket coverage' explains the differential ability of the monocot lectins in inhibiting HIV infection.

- Lobster and other marine crustacea use crustacyanin for camouflage to avoid predation. The absorption properties of the blue chromphore, astaxanthin, are minutely modified by binding to the crustacyanin complex. This is the protein product of a multigene operator, which produces at least 5 different subunits. These associate in pairs to bind astaxanthin, then associate in higher complexes to produce a range of environments around the chromophore, which can be modified in order to produce the colour change necessary for camouflage. The structure determination of two of these subunits, A1 and A3, shows a conserved fold, namely the lipocalin motif of proteins that bind non-polar substrates. A1 and A3 have only 25% sequence identity. Intriguingly, the A1 structure could not be determined with molecular replacement using other lipocalin models that have reasonable identity (RBP, MUP and others). It was eventually determined at 1.4Å resolution using a combination of softer X-rays and Xe derivative. Yet the A1 model could be easily used to solve the A3 structure in the A1/A3 complex at 3.2Å resolution, despite the lower identity in comparison with the other lipocalins. The A1 and A3 structures were more similar to each other than to any other lipocalin structure despite the relatively low identity, hence the ability to use the A1 model in molecular replacement. This is an example of a specific function requiring a specific motif, which can be achieved by divergent sequences.

Introduction to Protein Structure

Metaxia Vlassi, Institute of Biology, NCSR "Demokritos", Athens-Greece

The proteins belong to one of the most important classes of biological molecules and have a key role in almost all biological functions in living systems. The enzymes that control chemical processes in all organisms, hemoglobin that is responsible for the transfer of oxygen in the blood, insulin, the hormones, the antibodies, are only few examples of this family of ubiquitous molecules. Their importance is reflected to their name "proteins" which derived form the Greek word "πρότειος" meaning the first. All protein molecules are polymers built up from 20 types of aminoacids linked end-to-tail forming a polypeptide chain that is folded in three-dimensional space (tertiary structure) in a certain manner, characteristic for every protein. All functional properties of every protein depend on its three-dimensional structure: In some cases tiny atomic movements of the order of magnitude of 0,1Å (1Å=10^{-8}cm) can change dramatically the protein properties. Thus knowledge of the three-dimensional structure of proteins is a prerequisite for the proper understanding of their biological function.

Unfortunately, in spite of all efforts over the last 30 years and of recent significant technological advances, there is no direct method to deduce the 3D structure from the aminoacid sequence. This, known as the protein-folding problem, remains a problem because the 20 different aminoacids can be combined into an enormous number of different proteins with a big number of possible conformations for each one. In addition there is a vast number of ways in which similar structural domains can be generated by different aminoacid sequences.

Instead the tertiary structure of proteins can be determined experimentally by two different methods:

1) X-ray crystallography: The x-rays, as an electromagnetic radiation, interact with the electrons of the protein molecules arranged in a crystal, producing a characteristic diffraction pattern, which depends on the 3D structure of the molecule that produces it. X-ray crystallography is capable to determine protein structures at even the level of atomic resolution (1Å). This method requires the presence of crystals, the formation of which, in the case of proteins, consists the so-called bottleneck of crystallography and 2)

Nuclear magnetic resonance (NMR): NMR makes use of the magnetic spin properties of atomic nuclei within the molecule in order to obtain distance constraints between them. NMR does not require protein crystals and can be used on protein molecules in concentrated solutions. It is, however, restricted in its use to relatively small protein molecules.

Over the past 30 years the structures of around 13.000 proteins have been determined by one of the above methods and this number is increasing rapidly every year. Although the structural database (PDB) is relatively small compared to the number of available sequences (more than 500.000), it enabled us to discern patterns and motifs, to identify common structural themes, to relate structure and function and to see fundamental relationships between different proteins. In addition knowledge of these structures allowed us to predict the structure and function of unknown gene products based on sequence similarity to proteins of known structure.

Target Oriented Structural Genomics Using Synchrotron X-ray Protein Crystallography

Soichi Wakatsuki, Photon Factory, KEK, Japan

The recent years have seen a number of structural genomics (SG) projects started, for instance the NIH funded consortia in the US and the Protein Factory in Germany. In Japan, there are a major SG effort by RIKEN using both NMR and synchrotron X-ray techniques and several others, notably those on membrane proteins and rice genomics. During the last 12 months, the Photon Factory (PF) and Japanese universities have been proposing a network of structural biology consortia to participate in the national effort of determining protein structures with emphasis in medical, pharmaceutical and other industrial applications. The primary goal of the RIKEN project is to participate in the worldwide effort of determining all the representative structures in the structure space. In contrast, the proposal of the university consortia aims at target-oriented structural genomics with a strong emphasis on structure-function relationships. Each consortium will consist of X-ray protein crystallography and/or NMR groups tightly coupled with those specialized in medical, pharmaceutical and biological sciences that share the same biological interests in their pursuit of structure-function relationships.

The PF is a second-generation synchrotron radiation (SR) facility in Japan where a new structural biology group was created 15 months ago. Together with four universities and three institutes, the PF group is proposing a SG project on protein transport and post-translational modification of proteins. The project includes systematic structural analyses of proteins and complexes involved in the transport and oligosaccharide modification processes, as well as development of various high-throughput techniques using SR. The final goal is to develop a technology to produce medically/biologically active human glycoproteins using lower organisms such as yeast. The talk will start with an introduction to the SR protein crystallography methods followed by high-throughput techniques and some initial results of structural analysis from the SG project even though it hasn't officially started yet.

Overview and the Status of the SESAME Project

Herman Winick, Stanford Synchrotron Radiation Laboratory, Stanford Linear Accelerator Center, Stanford University

The SESAME Project has made significant progress since it was first suggested in September, 1997. This includes

1. UNESCO as an umbrella organization
2. Completion of a 113 page design report
3. Presentations at several scientific meetings
4. 11 countries have joined the project; more expected
5. Controlled/documented dismantling of BESSY I and storage in Berlin
6. Selection of a site offered by the government of Jordan in Alaan
7. Offer by the government of Jordan to provide funds to construct a building
8. Hiring of a Technical Director
9. 3 scientific workshops in 2000 (reports on the SESAME web site; www.sesame.org.jo)
10. Scientific collaborations among scientists from SESAME member countries
11. Workshop on accelerator science and technology
12. Training of 20 engineers and scientists in accelerator science and technology
13. Scientists from SESAME member countries working in SR labs in the US and Europe

However, the funds needed to realize the project (upgrading of BESSY I, construction of beam lines and user-support facilities, etc.) are not yet identified. Tension and violence in the region are hampering efforts to secure these funds and continue to make progress in developing the project.

It is important for scientists from the region to increase their involvement and provide more of the leadership in developing the project, defining the scientific program, beam lines, user-support facilities, and securing the necessary funds.

IV. *APPENDICES*

Appendix I List of Participants

Abdelghaffar, Houria
University of Mohammedia , Faculty of Science and T echnics, Mohammedia City , Morocco
E-mail: ahouria@caramail.com
Atalay, Rengul
Bilkent University, MBG Dept., 06533 Bilkent Ankara TURKEY
E-mail: rengul@bilkent.edu.tr
Babur,Özgün
Bilkent University, Dept of MBG, Bilkent06533 Ankara
E-mail: ozgun@bilkent.edu.tr
Bakkal, Suphan
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: suphanbak@su.sabanci.edu.tr
Bakr, Mohamed
Ain Shams Uni, X-RAY Unit, Physics Department, Faculty of Science, Cairo,Egypt
E-mail: bakr_moh_mah@hotmail.com
Baysal, Canan
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: canan@sabanciuniv.edu.tr
Berezovsky, Igor
The Weizmann Institute of Science Structural Biology Department Rehovot, 76100 Israel
E-mail: igor.berezovsky@weizmann.ac.il
Bilecen, Kivanc
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: kbilecen@su.sabanci.edu.tr
Charvalos, Ekatherina
School of Health and Caring Professions, TEI of Athens, Paraskevi, Athens Greece.
E-mail: gasson@aegean.gr
Dyczkowski, JERZY
Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Bioinformatics Unit
Ul. Pawinskiego 5A 02-106 Warszawa, Poland
E-mail: jerzyd@ibb.waw.pl
Eliopoulos, Elias
Genetics Laboratory, Dept. of Agricultural Biotechnology, Iera Odos 75, Athens 118 55, GREECE.
E-mail: eliop@aua.gr
Gursoy, Atilla
Bilkent University ,Dept. of Computer Eng. and Information Science 06533 Bilkent Ankara TURKEY
E-mail: agursoy@cs.bilkent.edu.tr
Ibrahim, Sameh
Ain Shams Uni, Physics Department, Faculty of Science, Cairo11566, Egypt
E-mail: mr_samh@hotmail.com
Kosakovski, Jaroslaw
Polish Academy of Sciences
E-mail: jarkos@ibbrain.ibb.waw.pl
Lancet, Doron
The Weizmann Institute of Science Molecular Genetics Department Rehovot, 76100 Israel
E-mail: doron.lancet@weizmann.ac.il
Malkawi, Hanan
Yarmouk University, Dept. of Biological Sciences. P. O. Box. 4933, Irbid-Jordan

E-mail: hanan@yu.edu.jo
Mobashery, Shahriar
Institute for Drug Design Department of Chemistry Wayne State University Detroit, Michigan 48202
E-mail: som@chem.wayne.edu

Najmanovich, Rafael Josef
Weizmann Institute of Science Plant Sciences Department REHOVOT 76100 - ISRAEL
E-mail: rafael.najmanovich@weizmann.ac.il
Ozbey, Suheyla
Hacettepe Univ., Dept. of Physics, Beytepe, Ankara
E-mail: sozbey@hacettepe.edu.tr
Ozturk, Umit
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: umit@su.sabanci.edu.tr
Pojskic, Naris
Institute for Genetic Engineering and Biotechnology Kemal begova 10, 71000 Sarajevo,
Bosnia and Herzegovina
E-mail: naris.pojskic@gmx.net , naris@ingeb.b
Rizkallah, Pierre
CCLRC Daresbury Laboratory, Keckwick Lane, Daresbury, Warrington, WA4 4AD
E-mail: p.j.rizkallah@dl.ac.uk
Rosenblum, Gabriel
Weizmann Institute of Sciences, Dept. of Structural Biology, 76100 Rehovot, Israel
E-mail: gabrielr@weizmann.ac.il
Sagi, Irit
Weizmann Institute of Sciences Department of structural Biology Israel
E-mail: irit.sagi@weizmann.ac.il
Sağdıçoğlu, Adviye Gülçin
Middle East Technical University,Biochemistry Department, Graduate School of Natural and Applied Sciences, Ankara,Turkey
E-mail: gulcin@metu.edu.tr
Sahin, Mert
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: merts@su.sabanci.edu.tr
Sayers, Zehra
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: zehra@sabanciuniv.edu
Sezerman, Ugur
Sabanci University Faculty of Engineering and Natural Sciences Orhanli, 81474 Tuzla-Istanbul, Turkey
E-mail: ugur@sabanciuniv.edu.tr
Shahkhatuni, Aleksan
Molecular Structure Research Center/National Academy of Sciences, Azatutian str., 26, Yerevan Armenia, 375014
E-mail: alexsh@msrc.am
Shatnawi, Moneeb
Yarmuk University, Physics Department, Irbid, Jordan
E-mail: m-shatnawi@usa.net
Siedlecki, Pawel
Polish Academy of Sciences, Institute of Biochemistry and Biophysics, Bioinformatics Unit
Ul. Pawinskiego 5A 02-106 Warszawa, Poland
E-mail: henryk@ibb.waw.pl
Solomon, Ariel
Weizmann Institute of Sciences, Dept. of Structural Biology, P.O.Box: 126, Rehovot, Israel
E-mail: ariel.solomon@weizmann.ac.il
Soukri, Abdelaziz

Université Hassan II, Department of Biology, Faculté des Sciences I Ain-chock, BP 5366 Maarif-Casablanca, Morocco
E-mail: a_soukri@hotmail.com
Vlassi, Metaxia
Institute of Biology N.C.S.R. "DEMOKRITOS" 15310 Ag. Paraskevi, Athens**,** Greece
E-mail: **meta@mail.demokritos.gr**
Wakatsuki, Soichi
Institute of Materials Structure Science KEK 1-1 Oho, TsukubaIbaraki, 305-0801 JAPAN
E-mail: soichi.wakatsuki@kek.jp
Winick, Herman
Stanford University, SSRL/SLAC MS 69; 2575 Sand Hill Road, Menlo Park CA 94025-7015, USA
Email; winick@slac.stanford.edu
Yonath, Ada
The Weizmann Institute of Science Structural Biology Department Rehovot 76100, Israel
E-mail: ada.yonath@weizmann.ac.il
Yousef, Mohammad
Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA
E-mail: yousef@sb.fsu.edu
Zielenkiewicz, Piotr
Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Bioinformatics Unit
Ul. Pawinskiego 5A 02-106 Warszawa, Poland
E-mail: Piotr@ibbrain.ibb.waw.pl

Appendix II

WORKSHOP/SCHOOL ON BIOINFORMATICS & STRUCTURAL MODELLING

September 3-7, 2001 Sabancı University, Istanbul Turkey

Monday, September 3, 2001
Chairman; I. Sagi

09:00-09:15	Welcome and Introduction	Z. Sayers
09:15-10:00	Overview and status of the SESAME project	H. Winick
10:00-10:30	Coffee break	
10:30-11:15	SR methods for 3D structure determination	Z. Sayers
11:15-12:00	Introduction to protein structure	M. Vlassi
12:00-13:00	Lunch Break	

13:00-17:30 Exercises
Introduction to different databases and specific protein families

Tuesday, September 4, 2001
Chairman; M. Vlassi

09:15-10:00	Similar/different function from similar/different structure	P. Rizkallah
10:00-10:30	Coffee break	
10:30-11:15	3D structure of bacterial ribosome	A. Yonath
11:15-12:00	3D structure and structure/function relationships	C. Baysal
12:00-13:00	Lunch Break	

13:00-17-30 Exercises
Introduction to display programs for visualization of 3D structure of proteins

Wednesday, September 5, 2001
Chairman; H. Winick

09:15-10:00	Secondary structure prediction software	E. Eliopoulos
10:00-10:30	Coffee break	
10:30-11:15	New Chapters in protein structure	I. Berezovsky
11:15-12:00	Threading	P. Zielenkiewicz
12:00-13:00	Lunch Break	
13:00-13:45	Homology modelling	U. Sezerman

13:45-17:30 Exercises
Secondary structure prediction and related databases and software

Thursday, September 6, 2001
Chairman; P. Rizkallah

09:15-10:00	Harvesting the human genome: a world-wide endeavor	D. Lancet
10:00-10:30	Coffee break	
10:30-11:15	Target oriented structural genomics using synchrotron X-ray protein crystallography	S. Wakatzuki
11:15-12:00	Structure dynamic studies of metalloenzymes using synchroton radiation	I. Sagi
12:00-13:00	Lunch Break	
13:00-17:30	Exercises	

Homology searches and alignment strategies

Friday, September 7, 2001
Chairman; Z. Sayers

09:15-10:00	Applications for Molecular Modeling and Dynamics Simulations to Systems of Importance to Biological Sciences	S. Mobashery
10:00-10:30	Coffee break	
10:30-11:15	Drugs and their protein targets: Side chain flexibility	R. Najmanovich
11:15-12:00	Pathway analysis tool for integration and knowledge acquisition	R. Atalay
12:00-13:00	Lunch Break	
13:00-17:30	Exercises and general discussion	

3D modelling

Appendix III

A Draft Proposal for a SESAME Computational Center

Developments in recombinant DNA technology and availability and rapid accumulation of DNA sequence data revolutionised approaches used in modern biology. Today, analyses which lead to model building for biological systems, necessitate electronic means for storage, transmission and predictive calculations. Effective use of available information reveals previously unknown relationships - ranging from the identification of unknown genes to the prediction of protein structures for functional analyses. Parallel to improvements in experimental methodologies for faster structural data collection and analysis, computational tools are now being developed aimed at predicting structures and elucidation of structure-function relationships as well as for drug design. Web-based facilities such Entrez at NCBI or Biology Workbench at SDSC, worldwide accessible databases e.g. Swiss Prot, PDB as well as web-based tools (BLAST, CLUSTAL W), and servers for secondary structure predictions (e.g. PHD, JOINT, GOR-IV) and 3D modelling (e.g. SwissModel, 3d-pssm) are at the disposal of scientists from many countries. The ease of access to data and facilities provide opportunities for competitive research regardless of the location of the scientist.
During the Sabanci University workshop where methods for predicting 3D structure of biological molecules, retrieving structural information from sequences of proteins and accessing facilities that are available through the world wide web for these applications were discussed, it became clear that a simple computational facility which could serve as a starting point for SMB studies of SESAME would be very useful. Such a facility would help to build a user base with background in structural analysis and encourage establishing collaborations in the region. It could also be used in computations for the synchrotron.

Requirements for the facility:
Deatiled requirements will be presented elsewhere.
Initially a facility which could be housed at the the exisiting buildings at the site at Aalan is planned.
The equipment will include a PC with advanced graphics cards, a printer.
A scientist with a background either in computer science or bioinformatics is recommended.

GROUP PHOTO