

# Preface

Joel L. Sussman\* and Israel Silman†

The concept of structural genomics (SG) arose in the mid-to-late 1990s in both the USA and Japan, triggered by the success achieved in applying high-throughput (HTP) sequencing methods to whole genomes. It was envisaged that application of a similar HTP approach to obtaining the three-dimensional structures of a substantial fraction of the entire set of proteins of a given organism, the “proteome,” would be an efficient way of filling in the gaps in observed “fold-space.” The decision to adopt such an approach resulted in the investment of large sums of money, i.e. hundreds of millions of dollars, in large-scale structural genomics projects in both countries. Thus, in Japan, the Protein Research Group was established at the RIKEN Genomic Sciences Center in 1998, and in the USA, the Protein Structure Initiative (PSI), funded by the NIH/NIGMS, commenced at nine major centers in 2000. These projects were characterized by concentration of resources in a small number of large centers, by development of novel, automated technologies which would allow a HTP pipeline approach to structure determination, and a focus on novel folds as a major target criteria.

Europe was slower in implementing HTP approaches to structural biology. Various national efforts such as the Protein Structure Factory

---

Departments of \*Structural Biology and †Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel

(PSF) in Berlin, the Oxford Protein Production Facility (OPPF), and the Genopoles in France, led the way. But it was only towards the end of 2002 that the first Europe-wide project began, an EC-funded Integrated Project entitled Structural Proteomics in Europe (SPINE). While benefiting from the technological achievements of the US and Japanese programs, and itself also concerned with developing cutting edge technologies as a means of achieving its objectives, SPINE from the outset focused these technologies on biomedically relevant targets. Indeed, as implied by its name, it aimed to establish a pan-European biomedically oriented structural **proteomics** program, placing significant emphasis on **functional aspects** of the target proteins studied. SPINE was followed by a series of EC-funded programs of various scopes and funding, some placing an emphasis on technological development, and others on attacking various classes of targets.

A similar emphasis on the use of the emerging HTP technologies to solve structures of biomedical relevance was adopted by the Structural Genomics Consortium (SGC), established in 2003 with the support of Canadian and British sponsors from both the public and private sectors, with laboratories in Oxford, Toronto and, subsequently, in Stockholm.

From the outset, the scale of funding of the PSI met with considerable criticism, especially in the USA, which has been going through a period during which funding for research by individual PIs has been hard to come by. Many critics have argued that the \$270M spent on funding the Pilot Phase of the PSI, over a period of five years, could have been more effectively utilized to fund hypothesis-driven research directed towards targets of fundamental or applied interest. Nevertheless, the achievements of the various SG/SP consortia, viewed in aggregate, and achieved in less than a decade, are impressive. In aggregate, the US PSI centers (September 2000 to June 2005) have determined over 1100 structures (<http://www.nigms.nih.gov/Initiatives/PSI/Background/PilotFacts.htm>). During PSI-2, which is still ongoing, ~1200 structures have already been solved, of which the vast majority share less than 30% sequence identity with any structure already deposited in the PDB, and many represent novel folds. As a result of the efforts of all the consortia (US, Japanese and

European), 5968 protein structures had been deposited in the Protein Data Bank as of 11-Dec-2007 ([http://www.rcsb.org/pdb/static.do?p=general\\_information/pdb\\_statistics](http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics)). Although some of these structures may be redundant, or even appear uninteresting at first sight, many are of the highest technical quality, of fundamental and/or medical importance and, taken overall, provide a valuable database. Moreover, it has been reported that, in 2005, structures arising out of structural genomics and structural proteomics efforts accounted for 44% of the total number of novel structures reported. Although many of the novel structures solved by the SG/SP centers were, on the surface, low-hanging fruit, which filled gaps either in a given proteome or in fold space without yielding novel functional information, other targets have been, as already mentioned, of great fundamental and/or medical importance. Furthermore, filling in fold space provides a robust body of templates for homology modeling, which can rapidly take advantage of these templates as computational techniques become more sophisticated, and computing power increases. Thus, in our view, whatever policy decisions are taken with respect to funding of large-scale SG/SP projects, what has been achieved so far will have a lasting impact on biological and biomedical research. Indeed, the EC, through a Specific Support Action (SSA), established the Forum for European Structural Proteomics (**FESP**, see <http://www.ec-fesp.org>) to assess the current status and make recommendations for future European infrastructure requirements in the SG/SP area. We feel, therefore, that it is timely to publish a book in which these achievements are presented with a look to their potential impact on biological and biomedical research in general.

In this volume, we have tried to bring together experts capable of addressing all aspects of the SG/SP effort, from target selection, through the various techniques for expressing and purifying proteins and protein complexes and the methodologies for solving their structures, to their impact on drug design and on coping with emerging diseases. In view of the ongoing debate on SG/SP funding, we have also included a special chapter dealing with policy, which includes sections written by several scientists and officials who have been closely associated with the decision-making processes.