basic molecular mechanism of HIF regulation has provided exciting therapeutic opportunities. Whether these findings can now lead to the development of useful therapeutic interventions will depend largely on collaborative investigations between basic laboratory and clinical researchers.

References

1 Semenza, G.L. (2000) HIF-1 and human disease: one highly involved factor. *Genes Dev.* 14, 1983–1991
2 Lando, D. *et al.* (2003) Oxygen-dependent regulation of hypoxia-inducible factors by prolyl and asparaginyl hydroxylation. *Eur. J. Biochem.* 270, 781–790
3 Hewitson, K.S. and Schofield C.J. (2004) The HIF pathway as a therapeutic target. *Drug Discov. Today* 9, 704–711
4 Berg, T. *et al.* (2002) Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts. *Proc. Natl. Acad. Sci. USA* 99, 3830–3835

David Lando
dl280@cam.ac.uk

Opening the bottlenecks

Glenn E. Dale, Morphochem AG, Schwarzwaldallee 215, CH-4058, Basel, Switzerland; e-mail: glenn.dale@morphochem.ch

The Protein Expression conference held 28–29 April 2004 in London, UK, brought together around 100 delegates from academia and the pharmaceutical industry to address the major issues in protein production. Much of the conference was devoted to the expression of soluble proteins for crystallography, with sessions describing high-throughput techniques and alternative expression systems and analysis. The importance of structural biology research has been highlighted in the past few years, not only as an integral part of drug discovery programs in the pharmaceutical industry, but also through the efforts in academia and the many structural genomics programs that have been established. Statistics arising from the various structure-genomics programs reveal variable success rates from the cloned protein to the structure determination [1]. The expression of a protein has a direct impact on the successful application of crystallography. If a protein can be expressed in a soluble form and purified as a monodisperse species then there is a 70% chance of obtaining diffraction quality crystals, whereas proteins that are aggregated or polydisperse in solution have <10% chance of crystallizing [2].

High throughput and automation

Jean-Dominique Guitton (Aventis; http://www.aventis.com) began the conference by describing the road towards the industrialization of protein structure determination. He pointed out that there were 28 different structural genomics initiatives, however the success rate from cloned gene to solved structure was, on average, less than 5%. He described the main bottlenecks in the process as protein expression and the crystallization, whereas the purification, data collection, data processing and structure refinement processes have generally been industrialized. He presented a robotic platform for vector construction (with a throughput of 90 constructs per month) and a robotic solution to performing high-throughput expression trials in *Escherichia coli*. He stated that, using a parallel approach of multiple constructs and expression conditions, they were able to reduce the time for structure determination from >14 months on average to five months.

Konrad Büssow (Protein Structure Fabrik in Berlin; PSF, http://proteinstructurefabrik.de) presented a cost effective approach to high-throughput protein production for structural analysis. PSF developed a method for parallel expression and purification of recombinant proteins with a hexahistidine tag (His-tag) or glutathione S-transferase (GST)-tag from bacterial expression systems. Proteins are expressed in 96-well microplates and are purified by a fully automated procedure on a pipetting robot. Up to 90 µg of purified protein can be obtained from 1 ml microplate cultures. Grant Cameron (NextGen Sciences; http://nextgensciences.com) described the Expressionfactory™, an instrument that automates the cloning, expression and purification within a single platform. When combined with the proprietary software the Expressionfactory™ enables the parallel exploration of different constructs, host cells and growth conditions to optimally produce the desired protein.

Angelo Gunasekera (Abbott; http://www.abbott.com) described an approach where the gene of interest was cloned into *Bsa*I site of a vector (IISC vector system) and how this approach was used for expression of essential *Streptococcus pneumoniae* genes. They tried several *E. coli* host cells with each

1359-6446/04/$ – see front matter PII: S1359-6446(04)03203-9

construct to test for expression and found large variations among the hosts depending on vector insert. Harry Yim (Invitrogen; http://www.invitrogen.com) presented the Gateway system and the latest advances. An interesting item was the addition of a T7 based *in vitro* protein synthesis system using *E. coli* extracts. Sabine Wizemann (Roche Applied Science; http://roche-applied-science.com) presented the Rapid Translation System for cell-free expression using *E. coli* extracts and wheat-germ extracts. The addition of various cofactors and chaperones to the extracts helped improve the solubility of various proteins. Both *in vitro* expression systems offer an alternative to cellular expression.

Towards soluble proteins

The next bottleneck to be addressed was the production of soluble protein. This was challenged in several presentations ranging from alternative fusion proteins to the high-throughput refolding of proteins. Fusion proteins are the favourite and often considered the universal remedy for insoluble protein expression. *E. coli* expression is still the host of choice for the structural genomics initiatives and possibly crystallography in general, due in part to its amenability to high throughput as well as the lack of posttranslational modifications, which often disturb the crystallization process. Many researchers included a panel of fusion proteins in the initial construct selection process. Tauseef Butt (Lifesensors; http://www.lifesensors.com) presented successful fusions with small ubiquitin-like modifier (SUMO). Lukas Leder (Novartis; http://www.novartis.com) described an elegant assay for determining protein expression using S-tag coupled to a variety of fusion partners. The S-tag when fused to the protein of interest is reconstituted with purified ribonuclease S, resulting in an active protein. The addition of an Rnase S FRET substrate to the lysed cells enabled the detection of expressed protein. The method enabled a rapid screen of suitable fusion partners.

Aire Geerlof (European Molecular Biology Laboratory; http://www.embl-hamburg.com) described a successful approach with the co-expression of molecular chaperones with a target protein. He presented several constructs where the various chaperone partners were co-expressed. Using molecular chaperones he tested 56 proteins with an 80% success rate. Glenn Dale (Morphochem; http://www.morphochem.de) applied an alternative approach. Mutant proteins were constructed to improve the solubility and the crystallization properties of the target protein. The method proved successful for a number of difficult proteins and various strategies were described including point mutations, insertions and deletions as well as fusion partners.

Protein refolding is often seen as tedious and too unpredictable to be considered by many as a method of choice. However, in recent years many methods have been established that have improved the outlook. There is now a refolding kit available, the FoldIt™ matrix screen, which can be purchased from Hampton Research (http://www.hamptonresearch.com) for those who would like to make an initial attempt. Laurent Vuillard (Avidis; http://www.avidis.fr) presented a high-throughput approach to protein refolding in which they screen as many as 40 conditions. Bernd Gerhartz (Novartis; http://www.novartis.com) presented a protein family approach and described the refolding of the Family A1 pepsin-like aspartyl proteases, which includes eight members. Following the initial protocol established for pepsin A they could expand and develop reproducible refolding protocols for the other members (except Napsin A).

Membrane proteins, more specifically G-protein coupled receptors (GPCRs), are often heralded as the holy grail of crystallography. Kenneth Lundstrom (bioXtal; http:// www.bioxtal.com) presented the strategy of MePNet towards GPCRs. MePNet is a research network dedicated to GPCR structural genomic technology improvement. The MePNet consortium is coordinated by BioXtal (http://www.mepnet.org) and is financially supported by industrial participants, including major pharmaceutical companies and Biotech startups. He presented the results of the work on 101 GPCR targets using *E. coli*, *Pichia pastoris*, and Semliki Forest Virus. Expression is analyzed by western blot and positive expression has been seen with >90% of the targets in one or more of the systems. To date, ten targets have been solubilized and five have been refolded and purified.

Concluding remarks

The expression of proteins for HTS and the manufacturing of therapeutic proteins have been emphasized in previous conferences but with the advent of structural genomics, protein production for biostructural research has changed the emphasis. There is clearly a need to improve the success rate of the structural genomics initiatives and this will begin with the production of a protein amenable to crystallization. This challenge was addressed in the many talks ranging from high-throughput expression and purification of a range of constructs to the single amino acid changes to induce the crystallization process. The various methods and approaches exemplified at this conference leave researchers with the opinion that we will soon see the bottlenecks of protein expression and crystallization disappearing and the promise of the various structural genomices initiatives being realized.

References

1 Service, R.F. (2002) Tapping DNA for structures produces a trickle. *Science* 298, 948–950
2 Zulauf, M. and D'Arcy A. (1992) Light scattering of proteins as a criterion for crystallization. *J. Cryst. Growth* 122, 102–106