of producing exquisite machines of nanometer size that can operate at room temperature in aqueous environments. Our current understanding of these processes and the products made by them is not yet at the point where we are able to fully design or customize any but the simplest proteins; however, our knowledge is rapidly improving and integration of these biological molecules in crucial functional roles in NEMS devices has already begun. Advances in technology are quickening the pace of discoveries emerging from academic and industrial laboratories, which in turn lead to more technical innovation. This feedback loop will result in future devices and techniques that are presently unimaginable: autonomously operating mobile robots *in vivo* that sense and respond to their environments and dispense drugs that are mechanical in nature.

References

1 Fritz, J. *et al.* (2000) Translating biomolecular recognition into nanomechanics. *Science* 288, 316–318
2 Wu, G. *et al.* (2001) Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nat. Biotechnol.* 19, 856–860
3 Soong, R.K. *et al.* (2000) Powering an inorganic nanodevice with a biomolecular motor. *Science* 290, 1555–1558
4 Yasuda, R. *et al.* (1998) F_1-ATPase is a highly efficient molecular motor that rotates with discrete 120 degree steps. *Cell* 93, 1117–1124
5 Yasuda, R. *et al.* (2001) Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. *Nature* 410, 898–904
6 Kacher, C.M. *et al.* (2000) Imaging microtubules and kinesin decorated microtubules using tapping mode atomic force microscopy in fluids. *Eur. Biophys. J.* 28, 611–620
7 Hess, H. *et al.* (2001) Light-controlled molecular shuttles made from motor proteins carrying cargo on engineered surfaces. *Nanoletters* 1, 235–239
8 Limberis, L. and Stewart, R.J. (2000) Toward kinesin-powered microdevices. *Nanotechnology* 11, 47–51
9 Benson, D.E. *et al.* (2001) Design of bioelectronic interfaces by exploiting hinge-bending motions in proteins. *Science* 293, 1641–1644
10 Harder, P. *et al.* (1998) Molecular conformation in oligo(ethylene glycol)-terminated self-assembled monolayers on gold and silver surfaces determines their ability to resist protein adsorption. *J. Phys. Chem. B* 102, 426–436
11 Nakagaki, T. *et al.* (2000) Maze-solving by an amoeboid organism. *Nature* 407, 470
12 Tsunoda, S.P. *et al.* (2001) Large conformational changes of the ε subunit in the bacterial F_1F_0 ATP synthase provide a ratchet action to regulate this rotary motor enzyme. *Proc. Natl. Acad. Sci. U. S. A.* 98, 6560–6564
13 Karcher, R.L. *et al.* (2001) Cell cycle regulation of myosin-V by calcium/calmodulin-dependent protein kinase II. *Science* 293, 1317–1320

The *Discussion Forum* provides a medium for airing your views on any issues related to the pharmaceutical industry and obtaining feedback and discussion on these views from others in the field. You can discuss issues that get you hot under the collar, practical problems at the bench, recently published literature, or just something bizarre or humorous that you wish to share. Publication of letters in this section is subject to editorial discretion and company-promotional letters will be rejected immediately. Furthermore, the views provided are those of the authors and are not intended to represent the views of the companies they work for. Moreover, these views do not reflect those of Elsevier, *Drug Discovery Today* or its editorial team. Please submit all letters to Rebecca Lawrence, News & Features Editor, *Drug Discovery Today*, e-mail: Rebecca.Lawrence@drugdiscoverytoday.com

The design of combinatorial libraries ▼

Combinatorial chemistry has an enormous impact on classical organic chemistry. Solution-phase syntheses are supported by polymeric reagents and scavengers. Automation of parallel syntheses and new purification methods, such as solid phase extraction (SPE) and automated column chromatography, are widely accepted by organic chemists. But is there a corresponding impact on success in drug discovery? When combinatorial chemistry started, huge libraries of mixtures of (impure) compounds dominated synthetic strategies. Screening results of such libraries were frustrating – hits could not be confirmed or biological activities 'disappeared' after deconvolution. Roger Lahana hit the nail on the head by formulating: 'When trying to find a needle in a haystack, the best strategy might not be to increase the size of the haystack' [1].

Therefore, better methods had to be developed to discover the needle. As long as we depend on HTS with all its inherent problems, rational design of targeted or focused libraries is the better alternative to the synthesis of huge, random libraries [2–4]. Lipinski's rule-of-five for the identification of analogs, which most probably lack sufficient oral bioavailability, was a first step in this direction. A further step forward resulted from the training of neural nets with drugs and chemicals, performed in parallel at BASF (Ludwigshafen, Germany) and Vertex (Cambridge, MA, USA), and later also at several other places [3,4]. The trained nets differentiate libraries with a higher percentage of biologically active compounds from those that include 'mere' chemicals. Genetic algorithms aid in the efficient selection of drug-like, chemically diverse libraries that can be produced from cheap building blocks [3,4].

1359-6446/02/$ – see front matter PII: S1359-6446(02)02274-2

Already established tools, such as similarity searching, clustering, nonlinear mapping (including Kohonen maps) and other virtual screening tools, aid in the selection of the most promising sublibraries. If the 3D structure of the biological target is known, docking methods are applied in later stages, for example, GOLD, FlexX [3,4] and, more recently, FlexE [5], which simultaneously considers the flexibility of the ligand and the binding site.

However, filtering and virtual screening of huge libraries is a waste of time if drug-like libraries can be generated from scratch. The program Comb^{i}Gen, developed at the University of Innsbruck (Austria), combines and modifies several hundreds of structure-optimized, drug-specific ('privileged') fragments to produce, *de novo*, structurally diverse, virtual combinatorial libraries with a high percentage of drug-like molecules [6]; ~1,000,000 structures can be generated per CPU day at a standard personal computer.

Another strategy of the future will be the combinatorial design of ligands within their binding site, following a concept that was first formulated by Böhm *et al.* [7] and which inherently applies the algorithms of the program FlexX. Instead of docking thousands of ligands separately, a recent extension of the Flex software suite, called FlexX-C [8], constructs a whole library within the binding site of the protein, step-by-step. Although such approaches are highly appealing, all current docking methods suffer from the insufficient precision of the scoring functions. Further research in this direction is required to proceed with combinatorial ligand design.

References

1 Lahana, R. (1999) How many leads from HTS? *Drug Discov. Today* 4, 447–448

2 Rose, S. (2002) Statistical design and application to combinatorial chemistry. *Drug Discov. Today* 7, 133–138

3 Böhm, H.-J. and Schneider, G. (eds) (2000) *Virtual Screening for Bioactive Molecules. Methods and Principles in Medicinal Chemistry.* (Vol. 10), Wiley–VCH

4 Schneider, G. and Böhm, H-J. (2002) Virtual screening and fast automated docking methods. *Drug Discov. Today* 7, 64–70

5 Claußen, H. *et al.* (2001) Efficient molecular docking considering protein structure variations. *J. Mol. Biol.* 308, 377–395 (http://www.biosolveit.de)

6 Wolber, G. and Langer, T. (2001) Comb^{i}Gen, a novel software package for the rapid generation of virtual combinatorial libraries. In *Rational Approaches to Drug Design, Proceedings of the 13th European Symposium on Quantitative Structure–Activity Relationships,* (Höltje, H-D. and Sippl, W., eds), pp. 390–399, Prous Science (http://pharmazie.uibk.ac.at/combigen.html)

7 Böhm, H-J. *et al.* (1999) Combinatorial docking and combinatorial chemistry: design of non-peptide thrombin inhibitors. *J. Comput.-Aided Mol. Design* 13, 51–56

8 Rarey, M. and Lengauer, T. (2000) A recursive algorithm for efficient combinatorial library docking. *Persp. Drug Disc. Des.* 20, 63–81

Hugo Kubinyi
University of Heidelberg, Germany
and BASF AG
Ludwigshafen, Germany

Functional proteomics: separating the substance from the hype ▼

The completion of the draft sequence of the human genome by the public consortium and Celera Genomics (Norwalk, CT, USA) has stirred controversy regarding the lower than expected number of genes in humans. Although the debate will continue for the next several years regarding the 'true' number of genes, the consensus is that the next bottleneck is the identification of the genes' function(s). Proteins, the main executors of biological function, have been estimated to be magnitudes greater in both number and complexity than genes. From the pharmaceutical industry's point of view, proteins represent the majority of molecules targeted by the drugs in the market today. An immediate requirement in the pharmaceutical industry is to prioritize which protein targets are suitable bases for developing novel and effective therapies.

The new opportunity created by the plethora of gene sequences with unknown function and the urgent need to understand protein function has led to the frequent misuse of the term 'functional proteomics'. Increasing interest from the financial markets has compounded this misuse. As the buzz moves from genomics to proteomics, it is important to separate the substance from the hype. Originally, proteomics focussed on the identification of the protein complement of the genome [1] but has since been extended to include everything associated with proteins on a global scale [2]. It is thus necessary to subdivide proteomics to fully grasp this burgeoning field, and the effort by Blackstock and Weir [3] to do this represented the first attempt to clarify the jargon. This communication extends the subdivision of proteomics and concurs that dividing proteomics and understanding each segment will facilitate conquering the complexity of proteins, which will lead to a better understanding of their biological function(s). Extending from the Blackstock and Weir nomenclature, proteomics can be subdivided into the following categories: (1) Expression proteomics; (2) Cell-map or interaction proteomics; (3) Functional proteomics; and (4) Structural proteomics.

Expression proteomics

Expression proteomics involves the creation of quantitative maps of expressed proteins from cell or tissue extracts (by separation methods such as 2D gels and MS), similar to cDNA arrays for transcription profiling. An alternative terminology used is 'differential proteomics', which refers to the comparison of differences between biological samples. Expression of functional proteins and the advent of protein chips represent a burgeoning area of proteomics research and fall under this category. This comparative profiling of proteins, noting their

 1359-6446/02/$ – see front matter PII: S1359-6446(02)02275-4