The impact of genomics on drug design

Lee J. Beeley and D. Malcolm Duckworth

Within the past few years, genomics has made a dramatic impact on the pharmaceutical industry, and it is viewed as an effective way to accelerate the discovery of pioneer drugs. In this review we describe how the application of genomics is generating new protein targets for drug discovery. Efficient linking of these new targets with structural studies, high-throughput screening and combinatorial/automated synthesis will provide significant benefits for both the industry and patients.

Genomic information is being generated at a rate that is testing our ability to make efficient use of the data. Proprietary and public databases are expanding rapidly as the genome and cDNA projects get into full swing. However, these data are meaningless unless steps are taken to transform the raw information into knowledge and then innovation. For the pharmaceutical industry, successful use of the data will yield a plethora of novel protein therapeutics, diagnostic tools and new molecular targets. *En route* to these commercial end products will be exciting discoveries that will provide a greater understanding of cellular signalling and the mechanistic basis of disease states. Many steps are required to process the genomic data into knowledge that, if used innovatively, will give a competitive edge both in intellectual property rights and improved cycle times to those involved in designing new drugs The main beneficiaries from this genomic revolution will be patients, with the design of drugs that more effectively target specific diseases. Significant progress has been made through conventional molecular biology, but linking mass genome and cDNA sequencing with positional cloning, high-throughput screening (HTS) and combinatorial/automated syntheses promises even greater rewards (Figure 1).

Alliances

A measure of how important genomics is becoming to the pharmaceutical industry is obtained by counting the number of strategic alliances that have been made public since 1993. According to information sources within SmithKline Beecham (SB), there have been at least 180 recorded company–company and company–academic institution alliances up to the end of 1995, covering a gamut of activities, from large-scale gene sequencing and analysis, through high-throughput screening (HTS), to combinatorial chemistry. All the major pharmaceutical companies are involved. Some initially chose to invest in exclusivity to databases of partial cDNA sequences from the human genome, as SB did with Human Genome Sciences (HGS). Others have chosen to licence Incyte Pharmaceuticals' genomic database products in non-exclusivity deals, while Merck (USA) have decided to fund a large-scale sequencing initiative, putting information into the public domain (Table 1). Many have chosen to focus on specific disease areas, by searching for genetic origins of the diseases (Table 2), or by entering collaborations that seek novel genes involved in specific cell-regulatory mechanisms, such as the Ras–mitogen-activated protein kinase pathway. Whatever the alliance, the successful companies will be those that can identify novel, disease-relevant genes and develop new products from them. Effective use of bioinformatics is the key to making successful use of the

Lee J. Beeley, SmithKline Beecham Pharmaceuticals, Department of Medicinal Chemistry, Great Burgh, Epsom, Surrey, UK KT18 5XQ.
D. Malcolm Duckworth*, SmithKline Beecham Pharmaceuticals, Department of Medicinal Chemistry, New Frontiers Science Park North, Third Avenue, Harlow, Essex, UK CM19 5AW. *tel: +44 1279 622017, fax: +44 1279 622348, e-mail: malcolm_duckworth@sbphrd.com

  1359-6446/96/$15.00. PII:S1359-6446(96)10041-6

genomic data. In the context of this review, bioinformatics is the process by which sequence data is transformed into useful biological information.

Table 1. Human cDNA large-scale sequencing alliances

Genome partner	Pharmaceutical company
Human Genome Sciences (HGS)	SmithKline Beecham (1993)
	Takeda (1995)(sub-licencee/SB)
	Synthelabo (1996) (sub-licencee/SB)
	E Merck (1996) (sub-licencee/SB)
	Schering-Plough (1996)
Incyte Pharmaceuticals	Pfizer (1994)
	Pharmacia & Upjohn (1994)
	Novo Nordisk (1995)
	Hoechst Marion Roussel (1995)
	Abbott (1995)
	Johnson & Johnson (1996)
	Roche (1996)
	Zeneca (1996)
	BASF (1996)
	Schering AG (1996)
	Vysis (1996)
Washington University (St Louis)	Merck (USA) (1994)

Table 2. Alliances targeted to specific disease areas

Genome partner	Pharmaceutical company
Millennium Pharmaceuticals	Eli Lilly (atherosclerosis, oncology)
	Roche (obesity, diabetes)
	Astra (inflammatory respiratory disorders)
	AHP (CNS diseases)
Myriad Genetics	Bayer (obesity, osteoporosis, asthma)
	Novartis (cardiovascular diseases)
Sequana Therapeutics	GlaxoWellcome (diabetes)
	Boehringer Ingelheim (asthma)
	Corange (osteoporosis)
Darwin Molecular Corporation	Rhône-Poulenc Rorer (cancer)

Bioinformatics

The genomics-to-drug-target process breaks down into several steps. First, efficient access to the data in a useable form is vital. Next, relevant data must be filtered out using sophisticated manual and automatic search tools. Then the genomic data should be (hyper)linked to other information sources and tools, such as protein secondary structure and fold prediction programs to transform the information into knowledge. Finally, to prioritize genes as targets for drug design, it will be necessary to correlate the genes with their biological function, thus involving a link to experimental work. However, no matter how efficient this process is, transformation of the knowledge into commercially successful innovations will require imaginative scientists, who have a combined knowledge of disease mechanisms and protein structure as well as a training in 'end user' bioinformatics. Obtaining new molecular targets will not be difficult, but identifying and exploiting the most appropriate targets in a timely manner will stretch the capabilities of most pharmaceutical companies.

Many molecular targets are being discovered by specific gene-hunting procedures such as homology, expression and positional cloning, but eventually the majority are likely to arise from nucleotide and protein sequence database searching. Where is the key information stored? There are both public and proprietary databases containing the raw genomic data. Of the public databases, the US-based GenBank and the European Molecular Biology Laboratory nucleotide databases are the workhorses of those doing sequence analysis. A range of protein databases that contain the deduced amino acid sequences are also available, and these include PIR[1], SWISS-PROT[2] and OWL[3], a non-redundant composite database. These public databases can be accessed via the Internet or commercially available CD-ROMs, and they can be searched using a variety of search tools that may run using mainframe or personal computers. Until recently, the nucleotide databases contained mainly complete DNA and mRNA sequences for individual genes, but their contents are now being swelled by cosmid sequences and expressed sequence tags (ESTs)[4,5] . The cosmids and EST sequences emerge from parallel strategies. Cosmids, the product of genome-sequencing projects, are large regions of complete DNA that do not discriminate between promoter, coding or 'junk' DNA sequences except by putative annotations. Because cosmid sequences are derived from DNA, they provide no clues to the tissue distribution of a particular protein. ESTs are being generated in many laboratories as a means of unravelling the coding regions of DNA. They are derived from cDNA libraries, and yield the partial sequences of mRNA in a particular tissue or cell line. Overlapping ESTs form contigs (assemblies), which potentially code for the expressed proteins. A significant proportion of the ESTs are proprietary, and the companies at the forefront of these initiatives, such

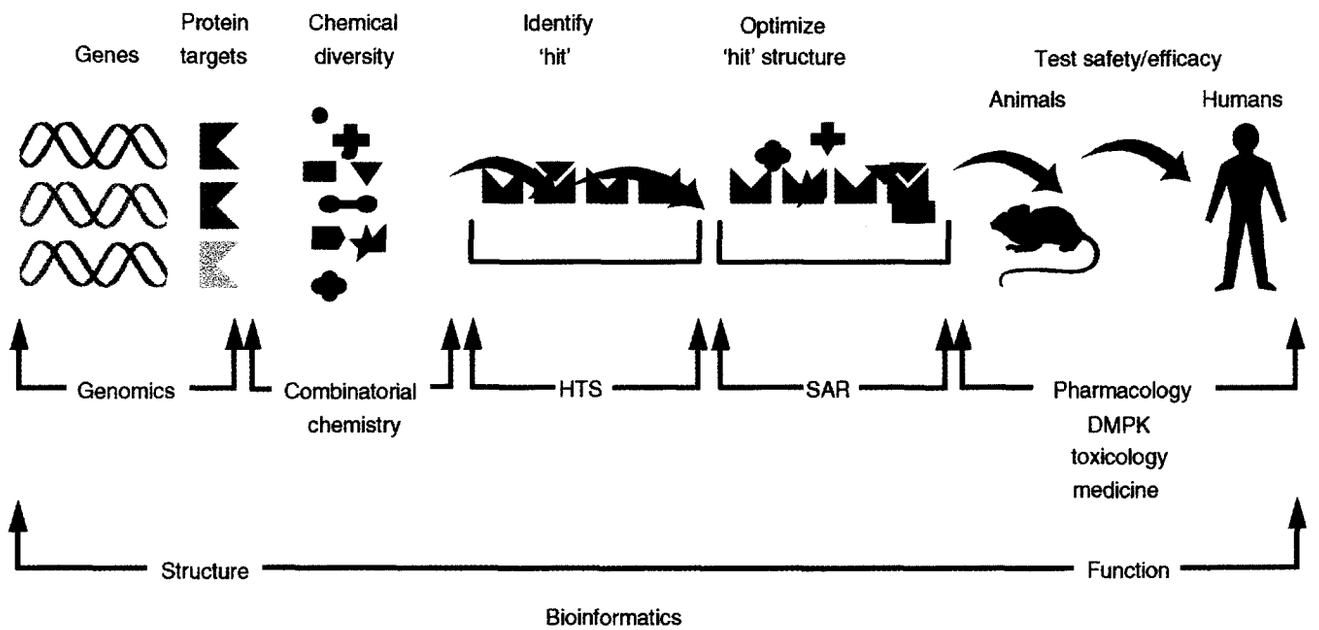


Figure 1. *Key steps in the drug discovery process. HTS stands for high-throughput screening, SAR for structure–activity relationships and DMPK for drug metabolism and pharmacokinetics.*

as SB and Incyte Pharmaceuticals, are developing their own relational databases to integrate the public-domain and proprietary sequences. The EST- and genome-sequencing strategies complement one another and, although most of the EST sequencing will be completed first, the Human Genome Project will provide very valuable additional information. A vast amount of data will also be available from the genomic sequences of many other species, for example mouse, nematode and various bacteria, before the completion date of 2003 predicted[6] for the Human Genome Project.

Whatever the source of a new gene, once a putative coding region has been identified the deduced protein sequence is used to search a database such as OWL using analytical tools such as BLAST[7] or FASTA[8]. Similarity of a protein sequence to a protein of known function or the presence of a particular motif in the sequence allows it to be classified (Figure 2). However, many genes (up to 50%) cannot be classified by sequence homology. For unknown genes that are thought to be of particular interest due to their chromosomal or tissue localization, or their association with a particular disease, a clue to their function can sometimes be gained from secondary or tertiary structure predictions of the deduced protein sequence. Many proteins with no significant sequence similarity share both a common structural motif (or fold) and common function. For example, cytokines with a four-helix-bundle structure frequently have very low homology even within a subfamily[9], but they interact with a family of receptors that signal through homologous proteins[10]. Clearly, as more gene sequences become available, new, large families of proteins will emerge that cannot be classified on the basis of sequence homology or structural predictions. In this instance, the structural determination of one key family member by nuclear magnetic resonance (NMR) or X-ray crystallography could result in the assignment of function to the whole family. Dealing with unknown genes is an important task, particularly for those companies with access to proprietary databases, as we will probably soon reach the situation where there are partial nucleotide sequence data on most genes.

Genomics to drug targets

Conventional molecular biology has yielded many new targets through expression and homology cloning, and these methods are still very productive. Here we give a few examples that indicate the potential of the rapidly expanding EST databases and of positional cloning of specific disease-related genes for identifying new targets. The identification of genes encoding a protease, Cathepsin K, and a lipoprotein-associated phospholipase, Lp-PLA_2, illustrate how the novel genes can be found from ESTs and how they can be linked to disease processes. The power of positional cloning is exemplified

by the *ob* gene, which encodes leptin. There are other examples, such as the hereditary breast and ovarian cancer gene *BRCA*2, which have been documented elsewhere[11,12].

Cathepsin K

In humans the skeleton is maintained by a balance between osteoblasts (bone-building cells) and osteoclasts (bone-absorbing cells). If osteoclast activity outstrips osteoblast activity, the skeleton becomes porous and weak; osteoporosis can then develop. Research scientists at SB/HGS identified a novel protease out of 5,000 randomly selected ESTs from a human osteoclast cDNA library, and classified it as a member of the cathepsin family based on similarity to published sequences. The full-length clone was expressed in *Escherichia coli* and resulted in a protease of 329 amino acids, 56% identical to the pre-proenzyme form of human cathepsin S. This new protein is now known as cathepsin K (Refs 13,14). It is highly abundant in the osteoclastoma cDNA library, representing approximately 5% of all clones sequenced, and was absent or extremely rare in other libraries. Experiments to confirm its biological role are still under way[15]; however, it is tempting to speculate that inhibition of this protease will selectively interfere with bone resorption and thus inhibitors may have a role in osteoporosis treatment.

Lp-PLA$_2$

In many instances, scientists are working with purified, isolated enzymes that are only available in small amounts, thus limiting the opportunity of obtaining a 3D structure to assist a rational-drug-design approach. The large number of unidentified sequences in EST cDNA databases offer a potential solution to this problem. They can provide a convenient route to a full-length clone, and consequently to the protein of interest, as was the case for SB with Lp-PLA$_2$ (Ref. 16). This phospholipase is thought to play an important pathophysiological role in atherosclerosis by generating lysophosphatidylcholine and oxidized fatty acids during the oxidative modification of low-density lipoprotein. Peptide fragments generated by digesting the purified protein with cyanogen bromide or trypsin provided short amino acid sequences, but none had high-scoring matches to any sequences in the public databases. However, there were previously unidentified clones within the SB/HGS EST database from which enough sequence information was forthcoming to search for the full-length sequence. A cDNA library from a human T-cell lymphoma library contained the correct clone in this instance. The clone coded for a protein of 441 amino acids (GenBank accession number U24577); it could be expressed in a baculovirus system giving approximately 7–8 mg/l of Lp-PLA$_2$, and it had activity indistinguishable from that of native enzyme in a tritiated platelet-aggregation-factor turnover assay. Clearly, there is an opportunity now to search the databases to look for novel, related enzymes.

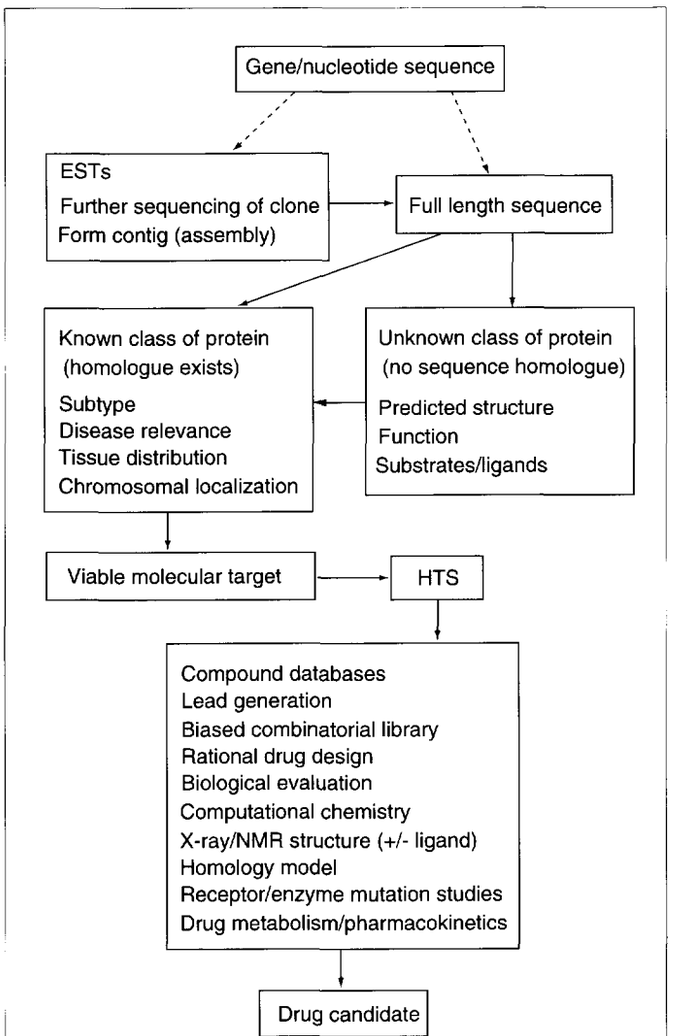


Figure 2. *From nucleotide sequence to drug design.*

Leptin receptor

At least five different rodent models of diabetes and obesity have been developed. One model, the *ob/ob* (obese) mouse, was discovered in the Jackson Laboratory (Bar Harbor, ME, USA), and recent analysis of the genetic defect responsible for this phenotype has initiated an unprecedented level of competition to capture the intellectual property rights and scientific accolade. The story of the *ob* gene, which encodes

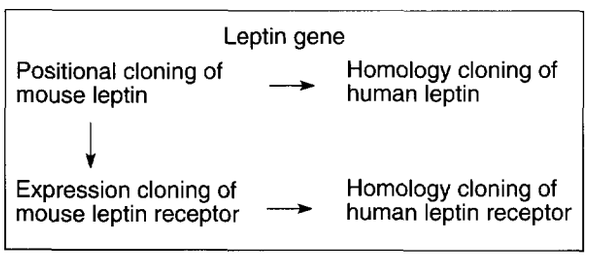


Figure 3. *The discovery of leptin and its receptor.*

a protein called leptin, is an example of the combined power of positional, homology and expression cloning and the use of mouse genomics to define human genes (Figure 3).

It has been proposed for some time that the *ob/ob* mouse lacks an unknown circulating hormone that regulates appetite and metabolism. Positional cloning of the *ob* gene[17], which encodes a 167 amino acid reading frame with a putative signal sequence, confirmed the existence of this hormone. The mutation present in *ob/ob* mice introduces a stop codon at amino acid 105 leading to a nonfunctional protein. Leptin displays little sequence homology to other proteins. However, a combination of its low level of sequence similarity to prolactin, logic and the examination of exon–intron structure yielded the tentative prediction that leptin is a cytokine in the leukaemia inhibitory factor (LIF) subfamily of 'gp130 users'. Therefore, leptin is likely to signal through a JAK–STAT (Janus kinase–signal transducer and activator of transcription) pathway. This hypothesis was supported by structural predictions suggesting a four-helix-bundle protein[18,19]. Overexpression and purification of alkaline phosphatase derivatives of leptin allowed the expression cloning of its receptor[20]. In line with predictions, the receptor appears to be a class 1 cytokine receptor of the 'gp130 users' subtype, and it has protein motifs characteristic of binding JAKs. Human leptin is being evaluated as a potential protein therapeutic for treating obesity and, with the cloning[20] of the gene for the human leptin receptor a screen should be available to search for leptin mimetics.

Screening and lead generation

Once a novel gene of functional relevance has been identified, it is essential to have other key technologies available that will lead to the identification of possible drug candidates. These technologies include HTS, protein expression, structural studies and automated/combinatorial chemistry. HTS and combinatorial chemistry have made an enormous impact on medicinal chemistry in the past few years; both have changed laboratory working practices, causing the chemists to rethink not only lead discovery but also lead optimization strategy. The need to produce much larger numbers of compounds much more rapidly has driven the change to increased automation, with automated nonpeptide chemical synthesis now a reality. Many companies are investing in automation. Some have developed their own 'in-house' systems to perform solution-phase and solid-phase chemistry, while others have chosen to invest in purpose-designed machines through external alliances. For example, two consortia have been formed; one by the US company Argonaut Technologies, who have linked up with Amgen, Berlex, Bristol-Myers Squibb, Merck, Pharmacopeia, Parke-Davis and SB to develop a machine called Nautilus[21], and the second by The Technology Partnership in the UK, who have linked up with BASF, Ciba-Geigy, Pfizer and SB to develop a machine called Myriad[22]. Prototype machines capable of carrying out many reactions in parallel will soon be available, leading to automated processes for synthesizing compounds up to 1000 times faster than was previously possible. Parke-Davis have also formed a spin-off company, Diversomer Technologies Inc., which aims to accelerate compound synthesis for lead generation and structure–activity relationships[23].

The most ambitious and exciting endeavour with regard to automation is the link-up of SB with the David Sarnoff Research Centre, a leader in micro-engineering, microfluidics and computer control, to form the company Orchid Biocomputer[24]. The agreement is to collaborate on and commercialize a business-card-sized chip capable of performing the parallel synthesis of 10,000 individual compounds, and to link it to a chip for automated screening against protein targets. Fifty of these chips completing four cycles per year would provide two million individual compounds per year! There is a reasonable chance therefore, that a lead identified using this technology will also satisfy the criteria for a drug development candidate, thus truly accelerating the drug discovery process.

Structural studies and rational drug design

Drug design benefits enormously from structural information on the protein of interest. Structural data are vital for rational drug design and for creating biased combinatorial libraries. What of the possibility of obtaining 3D structures for the new proteins emerging from genomic efforts? For non-membrane-bound proteins, such as proteases, there is

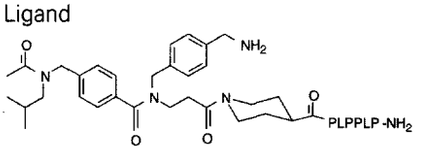
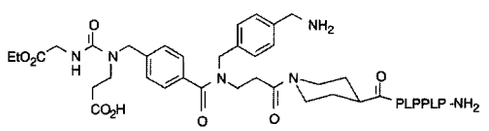
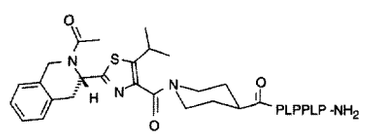
a good chance that a 3D structure will be forthcoming, provided that adequate expression of the protein can be achieved. Success rates have definitely improved as a result of technology and methodology advances in both NMR and protein X-ray crystallography. However, to accelerate drug discovery it is imperative that the structures become available to medicinal chemists early enough to have a significant impact on the rational design process.

When no 3D structure is forthcoming, homology modelling can provide a reasonable basis on which to try to predict protein structure[25]. A limitation to this is the enormous difference between the number of protein sequences and the number of solved structures. Currently there are 142,737 sequences in OWL compared with 4,162 full-release atomic-coordinate entries in PDB[26]. This difference is further exacerbated if one takes into account the redundancy caused by multiplicity of structures with the same fold.

For many membrane-bound proteins, the prospects of obtaining X-ray data are low owing to problems in preparing 3D crystals, hence true homology modelling is not possible. This is unfortunate because this category includes some of the most important targets for the pharmaceutical industry: the G-protein-coupled receptors (GPCRs) through which the monoamine neurotransmitters and the neuropeptides function. This lack of accurate structural data is particularly frustrating because the discovery of novel GPCRs has been an especially fruitful aspect of the genomics research, with over 40 novel examples identified as a result of the SB–HGS alliance. There is limited structural information for the largest GPCR subfamily, the 'rhodopsin-like' receptors[27,28], which includes the receptors for 5-hydroxytryptamine, dopamine, noradrenaline, neurokinins and endothelins, and this information has been used to generate computer models[29,30]. Because the structural data are of low resolution (cryoelectron microscopy) it is likely that the models will be far from perfect. For now, the best we can do is slowly refine them on the basis of the results of the numerous site-directed mutagenesis (SDM) studies that have been reported[31] and by identifying selective ligands. The second largest family of GPCRs is the secretin family, for which there are no reported computer models. These receptors all have large peptide ligands, for which structural data are also limited. Selective nonpeptide ligands have recently been discovered for a member of this family, the corticotropin-releasing-factor receptor CRF_1 (Ref. 32) which may herald the publication of some different models and additional SDM work.

The genomic initiatives are likely to identify many novel neuropeptides. An important link between sequence, structure and function will be made by coupling these neuropeptides to the orphan GPCRs. A method that addresses this is the focus of high-throughput functional screening methodology currently being developed by Cadus Pharmaceuticals. It utilizes the pheromone-response pathway of the yeast *Saccharomyces cerevisiae*[33]. This yeast can be engineered to express both a receptor and libraries of secreted random peptides, to select for peptides that bind to the receptor. An agonist response can be measured by changes in yeast growth or a colorometric response.

Box 1. Ligands for the Src SH3-binding domain

Ligand	K_d (μM)[a]
	3.4
	6.6
	11

[a]Dissociation constant measured by fluorescence perturbation assays.

The process in action

Combinatorial chemistry and lead generation have been reviewed elsewhere[34]. Exploiting molecular diversity to explore multiparameter space is an important aspect of combinatorial library design, but where does it end? It makes sense to take on board the advantages of preparing libraries of compounds to exploit leads, but find ways of constraining the limitless diversity. Given structural knowledge of the active site of a protein target, it should be possible to use all of its biostructural features to limit diversity and guide library design. Work on thrombin inhibitors suggests that this approach is feasible[35], and it is highly probable that further examples remain confidential to companies. The potential of this approach is illustrated by an example wherein peptide–nonpeptide hybrid molecules for the Src

homology 3 (SH3) domain of the protein tyrosine kinase Src were discovered[36]. Although the protein target in this instance has not emerged from genomic efforts, there will be related targets in the future to which these concepts can be applied. This example brings together many of the technologies and concepts that have been highlighted in the review: protein sequence alignments and structural studies to identify binding sites, all linked to combinatorial chemistry and HTS to identify specific compounds.

The novel ligands were designed as a result of the following events. A peptide structural motif was identified by aligning of many SH3-binding domain sequences; this formed the starting point for biased peptide libraries. As a result, peptides with affinities at least comparable to those of peptides derived from natural ligands were identified. Structural investigations (NMR) of SH3 domain–peptide complexes revealed that these domains bind peptide ligands in either of two orientations (classes I and II) involving three pockets. Using this structural knowledge of the ligand-binding domain, a resin-linked peptide sequence PLPPLP (P = Proline, L = Leucine), biased towards ligands favouring the class I orientation, was designed to fill two of the pockets known to bind LP dipeptides. By attachment to the N-terminal proline through the pyrrolidine nitrogen, molecular fragments could be oriented into the third pocket lined by the nSrc and RT loops, which are common to all SH3 domains and are the primary determinants of ligand specificity.

In order to identify ligands for the SH3 domain from the protein tyrosine kinase Src, a sensitive binding assay with colorimetric detection was developed[36]. From a library of approximately 1.1 million discrete compounds, several compounds were identified that had affinities within one order of magnitude of the highest-affinity SH3 ligands known (Box 1).

Examples of potential drug candidates emanating from a combination of HTS and combinatorial chemistry should soon start to appear. Selectide identified low-molecular-weight (550 Da) factor Xa inhibitors by this approach that were described as potential pre-clinical candidates for use as anticoagulants, but no structures were given[37].

Conclusion

Database searching and positional cloning will yield an abundance of potential molecular targets. The ability to rapidly select the targets that are both relevant to a particular disease and viable for small-molecule interactions will determine the survival of many pharmaceutical companies. The key to sifting out the relevant targets is the linking of genomics to function. Then, efficient coupling of genomics to drug design via other platform technologies, such as HTS and combinatorial chemistry, will significantly decrease the cycle time for drug discovery. With these technologies in place, the long-term prospects for patients are good, with the possibility of even more selective treatments that overcome the underlying causes of a particular disease.

REFERENCES

1 George, D.G. *et al.* (1996) *Nucleic Acids Research* 24, 17–20
2 Bairoch, A. and Apweilet, R. (1996) *Nucleic Acids Research* 24, 21–25
3 Bleasby, A.J., Akrigg, D. and Attwood, T.K. (1994) *Nucleic Acids Research* 22, 3574–3577
4 Adams, M.D. *et al.* (1991) *Science* 252, 1651–1656
5 Ashton, M.J., Jaye, M. and Mason, J.S. (1996) *Drug Discovery Today* 1, 11–15
6 *BioWorld Today* (1996) 7 (No. 72), 1 and 3
7 Altschul, S.F. *et al.* (1990) *J. Mol. Biol.* 215, 403–410
8 Lipman, D.J. and Pearson, W.R. (1985) *Science* 227, 1435–1441
9 Boulay, J-L. and Paul, W.E. (1993) *Curr. Biol.* 3, 573–581
10 Ihle, J.N. *et al.* (1994) *Trends Biochem. Sci.* 19, 222–227
11 Wooster, R. *et al.* (1995) *Nature* 378, 789–792
12 Tavtigian, S.V. *et al.* (1996) *Nat. Genet.* 12, 333–337
13 Human Genome Sciences (1996) US Patent 5,501,969
14 Drake, F.H. *et al.* (1996) *J. Biol. Chem.* 271, 12511–12516
15 Bossard, M.J. *et al.* (1996) *J. Biol. Chem.* 271, 12517–12524
16 Lawrence, G.M.P. *et al.* (1996) in *Genomes, Molecular Biology and Drug Discovery* (Browne, M.J. and Thurlby, P.L., eds), p. 183, Academic Press
17 Zhang, Y. *et al.* (1994) *Nature* 372, 425–432
18 Beeley, L.J., Smith, R.A.G. and Arch, J.R.S. (1996) 3rd International Symposium on Obesity and NIDDM, July 16–18, University of Buckingham, UK
19 Madej, T., Boguski, M.S. and Bryant, S.H. (1995) *FEBS Lett.* 373, 13–18
20 Tartaglia, L.A. *et al.* (1995) *Cell* 83, 1263–1271
21 Martin, J.F. (1996) CHI Conference on Solid Phase Synthesis: Developing Small Molecule Libraries, February 1–2, Coronado, CA, USA
22 *Research Fortnight* (1996) May 15, 5
23 DeWitt, S.H. and Czarnik, A.W. (1996) *Acc. Chem. Res.* 29, 114–122
24 *Scrip* (1995) No. 2034, 11
25 Bates, P.A., Jackson, R.M. and Sternberg, M.J.E. (1996) in *Genomes, Molecular Biology and Drug Discovery* (Browne, M.J. and Thurlby, P.L., eds), pp. 73–86, Academic Press
26 *Protein Data Bank Quarterly Newsletter* (1996) No. 75, 1
27 Henderson, R. *et al.* (1990) *J. Mol. Biol.* 213, 899–929
28 Schertler, G.F., Villa, C. and Henderson, R. (1993) *Nature* 362, 770–772
29 Hibert, M.F. *et al.* (1991) *Mol. Pharmacol.* 40, 8–15
30 Strader, C.D. *et al.* (1995) *FASEB J.* 9, 745–754, and references cited therein
31 Van Rhee, A.M. and Jacobson, K.A. (1996) *Drug Dev. Res.* 37, 1–38
32 De Souza, E.B. *et al.* (1995) *Ann. Rep. Med. Chem.* 30, 21–30
33 Kay, B.K. and Paul, J.I. (1995) *Mol. Diversity* 1, 139–140
34 Ashton, M.J., Jaye, M.C. and Mason, J.S. (1996) *Drug Discovery Today* 1, 71–78
35 Weber, L. *et al.* (1995) *Angew. Chem., Int. Ed. Engl.* 34, 2280–2282, and references cited therein
36 Combs, A.P. *et al.* (1996) *J. Am. Chem. Soc.* 118, 287–288, and references cited therein
37 Seligman, B. (1995) IBC Conference on Synthetic Chemical Libraries in Drug Discovery, 24–25 April, London, UK