kallikrein and papain, all but papain being implicated in the progression of cancer. Among the many active compounds detected following iterative deconvolution of the library, one component (Xaa = Trp, Yaa = Trp; 5μM) was found to be an inhibitor of the serine protease plasmin.

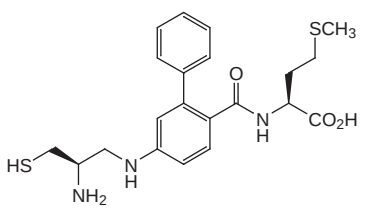
The data demonstrated that for plasmin, the S2′ subsite preferentially binds hydrophobic and, especially, aromatic amino acids. By contrast, the S2′ subsites of cathepsin B and papain do not appear to have strong preferences for any particular amino acid.

2 Abato, P. *et al.* (1999) Combinatorial library of serine and cysteine protease inhibitors that interact with both the S and S′ binding sites. *J. Med. Chem.* 42, 4001–4009

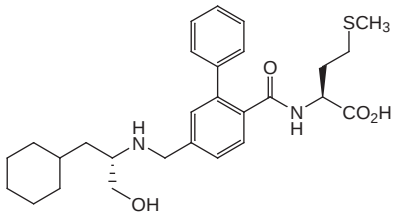
Protein farnesyltransferase inhibitors

The GTPase, RAS p21, is a molecular switch that fulfils a key function in cellular message transduction. Many human tumours, including lung, colon and pancreatic cancers, appear to contain a RAS mutant that is permanently in the 'on' state. As prenylation, catalyzed by protein farnesyltransferase (FTase), is an essential RAS post-translational modification, there has been much interest in the discovery of inhibitors of this enzyme as anticancer agents.

Using known active FTase inhibitor structures, combinatorial library chemistry has been employed in the discovery of novel and more potent inhibitors[3]. Using the known inhibitor, FTI276 (**iii**) as a starting point, a library of benzylic amines has been prepared via the reductive amination of a resin-bound aldehydic intermediate, followed by trifluoroacetic acid (TFA)-catalyzed cleavage from the resin. The preferred compound from this series (**iv**) is an inhibitor of FTase at nanomolar concentrations. Orthodox solution-phase medicinal chemistry has been employed to generate potent analogues that attenuate tumour growth in a nude mouse xenograft model of human pancreatic cancer.



iii



iv

3 Henry, K.J. *et al.* (1999) Discovery of a series of cyclohexylethylamine-containing protein farnesyltransferase inhibitors exhibiting potent cellular activity. *J. Med. Chem.* 42, 4844–4852

Nick Terrett
Discovery Chemistry
Pfizer Central Research
Sandwich, Kent, UK
fax: +44 1304 655419
e-mail: nick_terrett@sandwich.pfizer.com

A genome analysis production line

Rational approaches to speed up the drug discovery process using bioinformatic analyses of genomic DNA are intriguing, as they inherently deliver all the targets with all their associated information. Hence, an assay developer is able to select bioinformatically proposed targets that might not have become obvious from classical experiments.

Automated strategy of Directed Minimal Sequencing

As the time and cost pressures for genomic sequence analysis rapidly increase, Lion Biosciences AG (Heidelberg, Germany) has automated crucial process steps in their Directed Minimal Sequencing strategy. This approach sequences small genomic DNA fragments, which are pre-ordered by high-throughput hybridization. A 'minimal tiling path' is selected from a clone map to generate a sequence scan of the genome or a genomic region at homogeneous minimum coverage. On demand, a patent-quality sequence can be obtained, with bioinformatic analyses predicting potentially interesting loci, and the overall sequencing being reduced by >80%. Sequencing can be further reduced if, for example, a cDNA of interest is already known and can be hybridized to the clone map. Hence, the clones representing the exons and the regions in-between can be selected for directed genomic gene sequencing.

Advantages

By comparison with the classical shotgun sequencing, described by Fleischmann R.D. and coworkers[1], the Directed Minimal Sequencing strategy is suggested to have three major technical advantages.

Firstly, accurate (99.5%, mostly double-stranded) contiguous genomic sequences are produced online, independent of the genome size. Hence, the genomic DNA structure can be analyzed and gene-finding programs activated immediately after the first sequences are obtained.

Secondly, integration with automated genome assembly and functional whole-genome annotation (bioSCOUT™) enables the maximum utilization of the minimal sequencing results with little human intervention: assembly is reduced to a 'click' of the mouse, as clones have already been assembled in the mapping process. Optional finishing is performed on clones selected from the present maps instead of using PCR or additional backbone libraries. Open reading frames are automatically predicted and annotated.

 1359-6446/00/$ – see front matter PII: S1359-6446(99)01463-4

Thirdly, minimally ordered clones can be subjected to expression profiling experiments before, or instead of, sequence analysis. Clones that show differential expression levels can be individually selected for sequence analysis. This focuses the work on the functionally active regions and can reduce the overall sequencing to the minimum required for a defined biological question. If new questions arise, expression-profiling experiments can be designed to detect new genomic regions of interest, which can also be directly sequenced. The effort to access the complete genomic structure of all differentially expressed genes of a given set of profiling experiments is therefore reduced to one or two sequencing machine runs, together with the bioinformatic analyses.

The need for automation

To achieve a finished minimal sequence output in the megabase-per-week range using a relatively small team, automation of repetitive work has been essential, not only to guarantee the required quality, but also to speed up sequential processes. In collecting the library, bacterial clones are automatically transferred into microtitre plates using picking robots (Q-bot, Genetix, UK). The same robot replicates libraries and generates replica filters for physical mapping. One replica filter can contain up to 1.2 Mb of shotgun clones, which can be used repeatedly. A statistically significant proportion of all clones is qualitatively checked for insert presence and length homogeneity.

All information is then stored in the Laboratory Information Management System, a distributed system for documentation and steering of production processes and their assigned resources. It supports the qualitative and quantitative acquisition and reporting of the main production data. Every day, up to 1000 hybridizations are performed and evaluated using diverse self-made devices and software tools. Image analysis is time-critical and, thus, highly automated. For >70% of the hybridization results, human intervention is minimized to a few seconds per experiment.

Sequencing and expression profiling

For sequencing or expression profiling, the error-free extraction or re-arraying of the minimal tiling path is performed automatically using a Genesis 100/8 robot (Tecan, Crailsheim, Germany) integrated with a robotic workstation (CRS, Burlington, Canada). It can inoculate up to eight 384-well PCR plates in a working day from up to 200 384-well source plates. Sequencing template preparation and cycle sequencing is automated using various pipetting robots. To load more sequencing reactions onto one sequencing machine in less time, this process has been automated to receive up to 70 kb of DNA sequence per machine run. The disadvantages of pouring gels and the long time required for electrophoresis compared with capillary sequencing are compensated by a reduced input into template purification, a reduced total cost and a higher reliability.

The described strategy can be used for gigabase genome projects or, for direct microbial expression profiling, for comparative genomics. Expression profiling experiments on 'minimal tiling path' arrays can be designed to extract exons from genetically identified regions of the human genome to lead them to direct sequence analysis.

Reference

1 Fleischmann R.D. *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus*-influenza RD. *Science* 269, 496–512

Patrik Scholler
LION bioscience AG
Im Neuenheimer Feld 515-517
D-69120 Heidelberg, Germany
tel: +49 6221 40 38 00
fax: +49 6221 40 38 301
e-mail: contact@lion-ag.de

Collaborations…

Lexicon Genetics (The Woodlands, TX, USA) has announced that it will be collaborating with **The Scripps Research Institute** (TSRI; La Jolla, CA, USA) to generate knockout mice using its proprietary homologous recombination technology for a gene identified by TSRI researchers. The collaboration will then involve the determination of the function of the gene and its potential role in human disease. Randall Riggs, Vice President of Business Development at Lexicon said, 'We will be using biological and physiological screens from our Seek–Target–Validation (S–T–V™) program, in collaboration with the efforts of TSRI researchers, to determine the targeted gene's function and disease tolerance'.