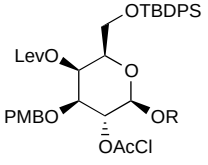
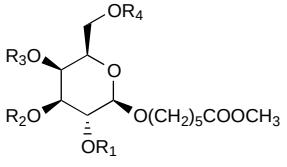
Oligosaccharide libraries

This column has frequently featured approaches to oligosaccharide libraries, mainly because these products are important pharmacological agents, but also because the synthesis of oligosaccharides poses significant problems. A recent paper from The Scripps Research Institute describes a useful orthogonal protection strategy that extends the accessibility of oligosaccharide libraries made in solution [Wong C-H. *et al. J. Am. Chem. Soc.* (1998) 120, 7137–7138].

The key advance described in this paper is an effective orthogonal protection–deprotection strategy that allows



3



4

the sequential removal of four different protecting groups from a core building block (**3**). Each of the four hydroxyls is protected by a different group: chloroacetyl (ClAc), *p*-methoxybenzyl (PMB), levulinyl (Lev), and *t*-butyldiphenylsilyl (TBDPS) selectively removed by sodium bicarbonate, trifluoroacetic acid, hydrazine and hydrogen fluoride-pyridine, respectively. After each deprotection the hydroxyl revealed can be coupled with a choice of seven glycosyl donors to generate 56 disaccharides, 1176 trisaccharides and, ultimately, 38,416 pentasaccharides.

This strategy has been used to prepare 45 protected oligosaccharides (**4**) in multimilligram amounts. Work is in progress to increase the number of compounds synthesized and screen for compounds that bind to lectins and antibodies.

Nick Terrett
Discovery Chemistry
Pfizer Central Research
Sandwich, Kent, UK
fax: +44 1304 618422
e-mail: nick_terrett@sandwich.pfizer.com

Emerging molecular targets

Caspase-8: an initiator of Fas-mediated apoptosis

Apoptosis – programmed cell death – is responsible for most cell deaths observed during embryogenesis, metamorphosis and normal tissue turnover. During apoptosis, protein-degrading enzymes known as caspases are activated and cleave critical cellular substrates leading to the ultimate death of the cell. The caspases are produced as inactive precursors that must themselves be activated by other caspases, so there is a cascade of cleavage and activation to induce apoptosis. In the immune system, the Fas molecule is a key inducer of the apoptotic process.

A recent paper by a group from Harvard Medical School describes an investigation into a human T-cell line variant found to be resistant to Fas-induced apoptosis [Juo, P. *et al. Curr. Biol.* (1998) 8, 1001–1008]. Examination of the cells for all the molecules known to be involved in Fas-induced apoptosis showed that the variant lacked the expression of caspase-8 but that all the other proteins were found to be present at normal levels. The group also showed that reintroduction of caspase-8 into these cells restored their sensitivity to Fas-mediated apoptosis. This implicates caspase-8 as the key enzyme responsible for the initiation of apoptosis in Fas-triggered cell death.

As malfunctions in the Fas system are involved in a variety of diseases, caspase-8 inhibitors or activators may have therapeutic potential for the correction of Fas-pathway defects.

Lymphoid sulphotransferase

During the inflammatory process, lymphocytes migrate from the bloodstream into the tissues by adhering to the endothelial cells lining specialized blood vessels. The process of adhesion is mediated through the binding of L-selectin on the surface of the lymphocytes to glycoproteins on the endothelial cell surface. Previous studies have shown that L-selectin binding requires sulphation of the endothelial surface glycoproteins but, until now, the identity of the sulphotransferase enzyme responsible for this process has been unknown.

Bertozzi's group has recently reported the identification of an enzyme specific to lymphoid tissue that can sulphate synthetic glycoprotein analogues, and suggest that this enzyme may be responsible for sulphation of glycoproteins recognized by L-selectin [Bowman, K.G. *et al. Chem. Biol.* (1998) 5, 447–460]. In addition to providing an insight into the mechanism of lymphocyte homing, the identification of this enzyme may lead to useful therapeutic agents for inflammatory control.

Andrew Lloyd

Bioinformatics

Integrated approach to bioinformatics

The tremendous pace of biomedical science is leading to an 'information explosion', an increased dependency on computers and an absolute need for effective and efficient use of bioinformatics. Bioinformatics is essential for the management and analysis of biological information and serves to eliminate *ad hoc* planning, enable informed decision-making, provide relevant information and mobilize all available biological information to take the discovery process forward and produce commercially viable products and patents. Consequently, up-to-date and rapid access to the right information (computer databases) and the right analytical tools (computer programs)

  1359-6446/98/$19.00 PII: S1359-6446(98)01254-9

is critical to a company's success and survival in today's tough and competitive world.

The challenge – especially for a small biotechnology or pharmaceutical company – is how to integrate and harness the power of bioinformatics. The challenge is not easy, because molecular bioinformatics is experiencing an exponential increase in sequence information – database content is doubling about every 14 months [Benton, D. *Trends Biotechnol.* (1996) 14, 261–272], the total number of databases is increasing and there are a plethora of programs (both public domain and commercial) available to analyse the data. The 'ideal' bioinformatics system would allow a company easy and effective integrated access to all molecular databases and bioinformatic analytical tools. Moreover, it would all be done at low cost and with absolute confidentiality.

Unfortunately, an ideal bioinformatics system does not exist and there is still a great deal of diversity in the design, structure and implementation of both databases and analytical programs. A lot of work is being carried out in the public domain to attempt to integrate and relate databases (for example, Entrez at the NCBI, http://www.ncbi.nlm.nih.gov/Search/Entrz/index.html and SRS at EMBL, http://www.embl-heidelberg.de/srs/srsc) and even

Table 1. Advantages and disadvantages of various approaches to bioinformatics systems (databases and analytical programs)[a]

	Advantages	**Disadvantages**
Databases		
Develop in-house	Build proprietary information and data particular to the company's needs; confidentiality maintained	Not all information requirements met; may still need access to the main public databases; time and cost involved developing database
Use public domain databases on line	Up-to-date information	Access time determined by Internet traffic; confidentiality compromised[c]
Download the required public databases[b]	All public domain databases available in-house and analyses on them confidential	Time and cost required for computer infrastructure and labour; would need about 90Gb of hard disk space to store some of the main sequence and secondary databases locally
Analytical programs		
Design in-house	Build proprietary tools particular to the company's needs; confidentiality maintained	May duplicate resources and costs when something already exists elsewhere to do the same job; time and cost involved in developing programs
'Off-the-shelf' analysis programs	Confidentiality maintained; quick and easy access to tools	Programs can quickly become out of date; stuck with the vendors 'solution' and support; can be expensive
Download programs	Flexibility maintained – can use many other programs; easy access to appropriate tools; confidentiality maintained	Program execution time determined by computer infrastructure; different programs require different operating systems; support may not be available
On-line programs	Flexibility maintained – can use many other programs; easy access to appropriate tools	Program execution time determined by internet traffic and host computer infrastructure; confidentiality compromised[c]
Collaboration	Confidentiality maintained – but depends upon how close the collaboration and the existence of contracts; more minds working on the problem	Problems with partners (e.g. political, legal, social, technical and personal); issues with proprietary information
Fee-for-service	Range of services available; limits start-up capital; flexibility maintained; confidentiality maintained – contracts usually tight	Not all tools or databases available; cost can be high for some services

[a]For more information see: Littlejohn, T.G. (1997) in *Humans to Proteins. Advances in Computational Life Sciences* (Vol. 2) (Michalewicz, M.T., ed.), pp. 155–166, CSIRO Publishing; and Littlejohn, T.G. *et al.* (1996) *Aust. Biotechno.* 6, 211–217.

[b]Could also subscribe to a CD-ROM version of the database, but these are published less frequently.

[c]Any submission through the Internet (eg. a sequence) may invalidate patent application based on that submission.

analytical tools (see Biology Workbench at http://biology.ncsa.uiuc.edu/). However, there are still issues that prevent effective and efficient use of available information [Karp, P. *Trends Biotechnol.* (1996) 14, 273–279]. Also, an analytical program that can do everything does not exist, so different programs are required for different jobs. But what are the options available to a company?

Bioinformatic options

A company can access public domain databases (eg. sequence databases such as GenBank) at the point of origin, use databases on-line, download the main public domain databases to their own system, or develop their own database. Similarly, the company can purchase off-the-shelf analytical software, use programs on-line, form a collaboration with some other company or institution with the necessary analytical expertise, design and develop their own analytical system in-house, use public domain analytical services, or use a fee-based analytical service. These approaches are not mutually exclusive – a company can use one or a combination of them – and each has it advantages and disadvantages (Table 1).

ANGIS and EnCompass

One option that a small company may find attractive out of all of the alternatives offered above is a fee-based service. This type of service limits a company's start-up capital requirements, maintains security and confidentiality, and still provides flexibility, individuality and independence. An example of this type of service is ANGIS (Australian National Genomic Information Service). The service is through a user-friendly, interactive, Web-based interface (WebANGIS at http://www.angis. org.au/ WebANGIS) that offers integrated access to numerous databases and analytical tools. It also provides facilities for easy storage and manipulation of sequence information. The ANGIS service is only available to subscribers within Australia; however, EnCompass Bioinformatics (a recent commercial spin-off from ANGIS) has been developed to provide the same services to the rest of the world (http://www.enbio.com.au).

Databases presently offered by ANGIS (and EnCompass) include: GenBank, OMIM, AceDB, Swiss-Prot, and Ribosomal Database. Analytical programs include:

- Database searching (eg. BLAST, FASTA, and Blitz)
- Sequence alignment (eg. Bestfit, SeqH, lalign, and Dotplot)
- Multiple sequence alignment (eg. Pileup, clustalW, and Profalign)
- Pattern recognition (eg. FindPatterns, Profile analysis, erepeat, and quicktandem)
- Protein structure and function prediction (PHD)
- Gene discovery (eg. GRAIL, and Proscan)
- Molecular phylogeny (eg. clustree, ednaml, ednadist)
- Restriction mapping (eg. mapplot, mapsort)
- File management facility

ANGIS also provides services in education, training [Bottomley, S. *Drug Discovery Today* (1998) 3, 426–428] and consultancy in bioinformatics. ANGIS doesn't claim to be the 'holy grail' of bioinformatics systems but it does provide a range of databases and tools in a useful integrated environment. ANGIS charges a yearly subscription of about AUS$1200 per group within Australia.

Steve Bottomley and*
Tim Littlejohn‡
**School of Biomedical Sciences*
Curtin University of Technology
Perth, Western Australia
e-mail: ibottoml@info.curtin.edu.au
‡Head of ANGIS
University of Sydney.
New South Wales
e-mail: tim@angis.org.au

Erratum

In the September issue of *Drug Discovery Today*, the Book review by David E. Szymkowski concerning *Clinical Trials of Genetic Therapy with Antisense DNA and DNA Vectors* gave the wrong publisher in the book details. We apologize for this error. The correct book details are as follows:

Clinical Trials of Genetic Therapy with Antisense DNA and DNA Vectors edited by Eric Wickstrom, Marcel Dekker, 1998. $185.00 (xvii + 427 pages, hardback) ISBN 0-8247-0085-6

Readers may also be interested to learn that since the publication of this book, the Isis 2922 oligonucleotide for treating retinitis induced by cytomegalovirus (CMV) has earned final approval by the FDA. It is the first antisense drug to reach the market – 31 years after publication of the first suggestion to use a synthetic oligomer for finding a naturally occurring sequence.