

A game of concentration	523
Under the subproteome	528
Russian scientists lead in pharmacophore space	528
Box 1: Making chemical genomics compute	528

## High-throughput screening goes to school

High-throughput screening, traditionally the domain of big pharmaceutical companies, is now creeping into academic laboratories in the guise of chemical genomics. Though the technique can be demanding and expensive, it is already yielding impressive results, as Alan Dove finds out.

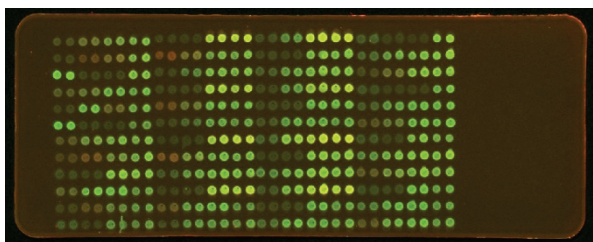
For years, scientists in academic labs and small biotechnology companies have suffered from throughput envy. Peering over the wall at their colleagues in big pharmaceutical firms, they saw massive banks of multiwell plates, vast libraries of novel chemicals and armies of robots processing thousands of experiments daily. The concept of high-throughput screening was appealing, but the budgets and staffs of these operations remained far out of reach for most basic scientists.

Now, declining equipment prices, the opening of a major national screening center and new high-throughput core facilities at some universities are feeding the development of a new field, variously called chemical genomics, chemical genetics or chemical proteomics. Practitioners disagree on the terminology—most just apply their favorite term to their own work—but the common theme is the use of high-throughput assays using small molecules to glean fundamental biological insights rather than direct pharmaceutical leads.

### Papers over pills

In industry, high-throughput chemical screens typically entail purifying a specific protein target, then testing the company's compound library to identify hits that change that protein's activity. The protein targets themselves are chosen on the basis of biological activity, perceived 'druggability', and the company's existing intellectual property covering particular drugs and targets.

Although basic scientists may envy the pharmaceutical companies' gear, they have approached high-throughput screening with a different set of priorities. "In academia, you have the luxury to get rewarded



Cellzome uses reverse phase microarrays to study subproteomes. (Image courtesy of Cellzome.)

for your work even if it doesn't bring you to solid [intellectual property]," says Andrei Gudkov, Scientific Director of the Small Molecule Screening Core at the Cleveland Clinic in Cleveland, Ohio, USA.

Chemical 'omics, which occurs primarily in places like Gudkov's academic core facility, begins instead with a general biological phenomenon, such as cancer cell division or DNA replication. Researchers develop a suitable high-throughput assay for the process, often using whole cells or cell fractions rather than purified proteins, then screen libraries of compounds to see which ones affect it. These hits then serve as laboratory tools to probe the biology.

The approach is less direct than testing a single protein target, but Gudkov argues that it is less biased. "Sometimes your favorite molecule was picked from a plethora of others ... not based on objective reasons but simply reflecting the history of the subject," he says. Testing the effects of novel chemicals on whole cells removes that bias, potentially revealing entirely new targets.

Indeed, Gudkov now has the data to back that claim: in 2005, he and his colleagues used chemical genomics to uncover a new tumor suppression mechanism used by the much-studied protein p53 (ref. 1). That work, along with other successes in the field, has spurred more academic research-

ers to dip their toes into the high-throughput pool.

Although several major research centers now have screening core facilities, they do not generally offer turnkey solutions; one does not simply drop off an assay at the core facility and pick up useful hits a few months later. "It never works like this for a screening facility," says Gudkov.

Instead, researchers should expect to spend some time collaborating with the screening core to optimize the assay. Once the test is suitable for high-throughput techniques, the postdoc or graduate student leading the project will generally spend a few months working directly in the screening core, first with a small pilot set of compounds and then with a more comprehensive library.

Newcomers to chemical genomics are not the only ones facing a steep learning curve. With the field still in its infancy, granting agencies have not yet figured out how to assess proposals for academic high-throughput screening. Chemical genomics study sections at the US National Institutes of Health (NIH) include members familiar with the technique, but the standards are still a moving target.

At one recent study section meeting he attended, Gudkov says "the most shocking thing for me ... was the complete lack of solid criteria for selection or evaluation of qualified proposals." As the field becomes more established, experts expect that clearer standards will emerge.

### A game of concentration

The NIH, for its part, is keen to see chemical genomics take off. As part of the recent Roadmap initiative, the Institutes have launched the NIH Chemical Genomics



Mass spectrometry goes high-throughput. (Image courtesy of Cellzome.)

Center (NCGC), which is both a national-level core facility and a collection of ‘centers of excellence’ scattered around the country.

The main NCGC facility is the brain-child of two experienced high-throughput screeners who learned the field at Merck: Christopher Austin, NCGC Director, and James Inglese, who runs the NCGC

Biomolecular Screening and Profiling Division. “Coming to the NIH, I sat around thinking ... ‘what would I do differently if I were starting from scratch?’” says Inglese. Among other things, he decided that the new facility should pay more attention to pharmacology than pharmaceutical screeners do.

In a typical pharmaceutical high-throughput screen, researchers test millions of compounds in an assay, all at the same concentration. When screeners find hits this way, they must test the compounds’ dose-response characteristics and toxicities separately, and the data are difficult to compare between screens. “There was a compromise along the way, and the decision was basically let’s just forget about pharmacology ... just see what’s active,” says Inglese.

At the NCGC, Inglese and his colleagues developed an alternative approach, which they call quantitative high-throughput screening<sup>2</sup>. Testing each compound at seven or more concentrations, they obtain dose-response data and hits simultaneously. That septuples the size of the job, of course, but because the team was starting from scratch, they were able to use the latest miniaturization techniques, including 1,536-well plates and high-precision liquid handlers.

Driving the screening process is a robot built by GNF Systems and Kalypsys, and now marketed separately by both companies. Besides handling smaller volumes

and denser plates than previous systems, the GNF-Kalypsys robot doubles as a compound-storage system. “We basically prepare the library and load it onto the robot once every six months,” says Inglese.

The system is also more reliable than older high-throughput screening tools because of a combination of the heavy-duty robot and an assay design that places different compound concentrations on separate plates, which makes it easier to keep the titrations accurate. “We’ve gotten wind of this from other people who’ve tried [quantitative high-throughput screening], and people who’ve had robotic systems that are less robust generate noisy data,” says NCGC director Christopher Austin.

As an NIH facility, the NCGC is available to researchers worldwide. “We’re always looking for good assays,” says Austin, adding that when it comes to targets, “anything which is encoded by any genome in any species is fair game.” Researchers can submit proposals to the center, which reviews them to ensure that the assays are suitable for high-throughput screening. Before send-

ing a proposal, though, Austin recommends making a phone call. “Unless people have a lot of experience developing assays, which most academics don’t, ... we really much

prefer that they just contact us,” he says.

Once researchers have found a few interesting hits, they might want to turn to a related NIH effort, called PubChem. Run by



A high-precision robot manipulates 1,536-well plates at the NIH Chemical Genomics Center. (Image courtesy James Inglese, NCGC.)

the National Library of Medicine, PubChem is a rapidly expanding database of small-molecule structures, with information on their bioactivities. Links to GenBank, Medline and similar resources allow users to browse through all of the information available on specific compounds. At this writing, PubChem contained data on at least 10 million compounds, and the database is growing daily.

"What one can do then from PubChem is to go seamlessly from a gene to a protein to a small molecule that binds to that protein, and see the crystal structure and the protein bound to it, and you can go to papers on what that small molecule and that protein and that gene do," says Austin. Combining free databases like PubChem with new screening software (see **Box 1**) might eventually provide academic scientists with high-throughput screening capabilities that rival those in industry.

### Under the subproteome

Industry, of course, is not sitting still. While major pharmaceutical companies are still heavily invested in screening specific protein targets, some are starting to explore chemical genomics approaches as well. Meanwhile, a few small biotechnology companies hope to use chemical genomics to find targets the big players missed.

"We're looking at drug-proteome interactions," says Gitta Neubauer, vice president of research operations at Cellzome, who describes the company's approach as "chemical proteomics". As one of the early adopters of the strategy, Cellzome has established collaborations with pharmaceutical giants Johnson and Johnson, and Ortho-McNeil, but the company is also developing its own drug leads.

Rather than screen drugs against individual proteins, Cellzome screens 'subproteomes', collections of proteins purified by their affinity to particular compounds. "We

have a set of small molecules that defines a tractable part of the proteome, and these small molecules can be immobilized on a matrix," says Gerard Drewes, the company's vice president for discovery research. Drewes adds that this part of the technique is "quite similar to affinity chromatography."

Having selected an appropriate subproteome—for example, all kinases and their associated protein complexes—the company's researchers then search for compounds that will elute the kinases or their partners from the matrix. The scientists can identify the eluted proteins, which are the test compound's targets, by mass spectrometry. "We test our components against many different kinds of things in one assay," says Neubauer.

To run these complex assays, the researchers use a combination of off-the-shelf robotic systems from Perkin-Elmer, and custom-built or modified hardware developed in-house.

Cellzome's general approach has also proven useful for basic research. Using a similar affinity purification and mass spectrometry technique, the company recently performed a genome-wide screen for protein complexes in *Saccharomyces cerevisiae*, isolating a total of 491 putative complexes, 257 of which were newly identified<sup>3</sup>. Cellzome is now collaborating with academic researchers on other 'interactome' projects, but since completing the yeast work, Drewes says the company's main focus has shifted to subproteomes involved in specific human diseases.

## BOX 1 MAKING CHEMICAL GENOMICS COMPUTE

As chemical genomics starts to take off, researchers in the field are already finding themselves swamped with data and discovering that there is little software available to analyze it. Applications built to handle traditional high-throughput screening seldom contemplate proteome-scale analyses, and existing proteomics software is not suited for high-throughput experiments.

"When we first started, we greatly underestimated processing time," says NCGC's Inglese. His experiments, which are essentially traditional high-throughput screens done at multiple compound concentrations, typically take about 30 h to run, but analyzing the data initially took his team nearly six months. "That's because there really wasn't any software developed or available that was capable of doing those kinds of calculations," he says.

To address that, bioinformatics researchers at the NIH are now developing computer applications custom-tailored to chemical genomics. The new programs are still infested with bugs, but the team hopes to release them to the scientific community once they are refined.

Other researchers are working on a different bioinformatics problem in chemical genomics: reading out and analyzing the complex assays that many of the new projects use. At the end of 2005, for example, the Broad Institute (Cambridge, Massachusetts, USA) released its CellProfiler application, which automates phenotype analysis in cultured cells. The program has since undergone several revisions, with updates coming out every few months. CellProfiler has already been used to analyze data for more than a half-dozen published papers.

The progress has been impressive, but some of the most useful informatics systems are still available only to industry. "They've been doing the research for some time, and they have very often tools that the academics don't have," says Reg Richardson, European Marketing Manager for ChemBridge Corporation.

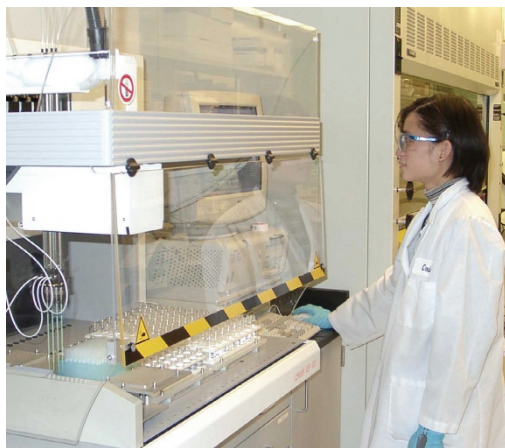
In particular, industry scientists have proprietary 'virtual screening' programs that can test millions of compounds without dipping a single pipet tip. Richardson says pharmaceutical companies commonly do such screens on ChemBridge's entire catalog before placing an order. "They will download the database, perform virtual screening on all half-million compounds and select a relatively small number of hits from that virtual screening," he says.

### Russian scientists lead in pharmacophore space

Regardless of the assay, anyone hoping to do high-throughput screening needs to pick an appropriate library of compounds to test. For big pharmaceutical companies, a team of staff chemists caters to that need, but academic researchers lack that luxury. Fortunately, several types of compound libraries are now available for purchase, covering a wide range of pharmacological complexity and price.

Standard chemical suppliers offer libraries that encompass nearly all known drugs in clinical use—about 4,000 unique compounds. Although quite small by high-throughput standards, that type of library has some distinct advantages. "It can be [screened] very fast, and people are usually very happy if something already approved for use in humans can affect





ChemBridge provides a variety of chemical libraries. (Image courtesy of ChemBridge.)

their target,” says the Cleveland Clinic’s Gudkov.

If this quick, cheap approach yields no useful hits, researchers can turn to larger libraries available from specialty suppliers. In the past decade, these suppliers have dipped into a trove of compounds produced by chemistry students at Russian universities. The students must synthesize new chemical structures as part of their advanced training, and since the collapse of the Soviet Union, selling these compounds to high-throughput library suppliers has kept many of the country’s scientific centers afloat.

“Because there is no well-established Russian pharmaceutical industry, those compounds would do nothing if companies like Chembridge and others didn’t buy them,” says Reg Richardson, European marketing manager for ChemBridge. With one of the largest collections of these compounds—now approaching 500,000—ChemBridge is a common source for the libraries used in chemical genomics.

Although the known-drug library is quite small, the complete ChemBridge collection is enormous, and most scientists prefer to screen a mid-size collection instead. To address that need, ChemBridge offers a subset of the library, called DIVERset. “DIVERset, which is a subset of about 50,000 compounds, covers something like 60–65% of the pharmacophore space,” says Richardson.

The pharmacophore space is a theoretical construct used widely in high-throughput screening. Chemists attach descriptors to

particular chemical moieties, such as hydrogen bond acceptor, hydrogen bond donor and aromatic group, then classify compounds based on the number and types of descriptors in them. Two compounds with radically different structures might have very similar pharmacophore descriptors, so screeners would only need to include one of the two in a library to cover that portion of the pharmacophore space.

ChemBridge also offers subsets of DIVERset, structured so that a group of 10,000 compounds covers nearly the same pharmacophore space as the 50,000-compound set, though at lower resolution. “People can buy the whole 50,000, and we’d be pleased if they did, but few can afford to buy that, so they tend to buy 5,000–10,000,” says Richardson. These smaller subsets, with enough material for two dozen assays, cost about \$10,000. Investigators who prefer to pick their own subset of the overall library can also browse ChemBridge’s entire database, which is available for free downloading.

Whichever approach they choose, chemical genomics researchers are clearly learning that it is not the size of the assay that matters—but how you probe it.

1. Gurova, K.V. *et al. Proc. Natl. Acad. Sci. USA* **102**, 17448–17453 (2005).
2. Inglese, J. *et al. Proc. Natl. Acad. Sci. USA* **103**, 11473–11478 (2006).
3. Gavin, A.C. *et al. Nature* **440**, 631–636 (2006).

Alan Dove is a science writer based in the New York City area (alan.dove@gmail.com).

## SUPPLIERS GUIDE: COMPANIES OFFERING CHEMICAL BIOLOGY REAGENTS AND EQUIPMENT

Company	Web address
Abbott Molecular	<a href="http://www.abbottmolecular.com">http://www.abbottmolecular.com</a>
Abgene	<a href="http://www.abgene.com">http://www.abgene.com</a>
Agilent	<a href="http://www.agilent.com">http://www.agilent.com</a>
Anaspec	<a href="http://www.anaspec.com">http://www.anaspec.com</a>
Apogee Instruments	<a href="http://www.ccd.com">http://www.ccd.com</a>
Applied Biosystems	<a href="http://www.appliedbiosystems.com">http://www.appliedbiosystems.com</a>
Applied Imaging	<a href="http://www.aicorp.com">http://www.aicorp.com</a>
Applied Precision	<a href="http://www.appliedprecision.com">http://www.appliedprecision.com</a>
Applied Scientific Instrumentation	<a href="http://www.asiimaging.com">http://www.asiimaging.com</a>
Andor Technology	<a href="http://www.andor.com">http://www.andor.com</a>
Aurora Biomed	<a href="http://www.aurorabiomed.com">http://www.aurorabiomed.com</a>
B-Bridge	<a href="http://www.b-bridge.com">http://www.b-bridge.com</a>
BD Biosciences	<a href="http://www.bdbiosciences.com">http://www.bdbiosciences.com</a>
Beckman Coulter	<a href="http://www.beckman.com">http://www.beckman.com</a>
Biaffin GmbH	<a href="http://www.biaffin.com">http://www.biaffin.com</a>
BioImagene	<a href="http://www.bioimagene.com">http://www.bioimagene.com</a>
Biometra	<a href="http://www.biometra.de">http://www.biometra.de</a>
Bio-Rad	<a href="http://www.bio-rad.com">http://www.bio-rad.com</a>
BioTec	<a href="http://www.biotec.co.jp">http://www.biotec.co.jp</a>
BioTek	<a href="http://www.biotek.com">http://www.biotek.com</a>
Biotrend Chemikalien GmbH	<a href="http://www.biotrend.com">http://www.biotrend.com</a>
Biovision AG	<a href="http://www.peptidomics.com">http://www.peptidomics.com</a>
Bruker Daltonics	<a href="http://www.bdal.com">http://www.bdal.com</a>
Caliper Life Sciences	<a href="http://www.caliperls.com">http://www.caliperls.com</a>
Cambrex	<a href="http://www.cambrex.com">http://www.cambrex.com</a>
CambridgeSoft software	<a href="http://www.cambridgesoft.com">http://www.cambridgesoft.com</a>
Cell Profiler software	<a href="http://www.cellprofiler.org">http://www.cellprofiler.org</a>
Cell Signaling Technology	<a href="http://www.cellsignal.com">http://www.cellsignal.com</a>
Cellomics	<a href="http://www.cellomics.com">http://www.cellomics.com</a>
CGI Pharmaceuticals	<a href="http://www.cellulargenomics.com">http://www.cellulargenomics.com</a>
ChemBridge	<a href="http://www.chembridge.com">http://www.chembridge.com</a>
Chemicon	<a href="http://www.chemicon.com">http://www.chemicon.com</a>
Clemex Technologies	<a href="http://www.clemex.com">http://www.clemex.com</a>
Corbett Life Science	<a href="http://www.corbettlifescience.com">http://www.corbettlifescience.com</a>
CyBio	<a href="http://www.cybio-ag.com">http://www.cybio-ag.com</a>
Data Translation	<a href="http://www.datx.com">http://www.datx.com</a>
Deerac Fluidics	<a href="http://www.deerac.com">http://www.deerac.com</a>
DiscoverX	<a href="http://www.discoverx.com">http://www.discoverx.com</a>
DVC Company	<a href="http://www.dvcco.com">http://www.dvcco.com</a>
Elsevier MDL	<a href="http://www.mdli.com">http://www.mdli.com</a>
EMD Biosciences	<a href="http://www.emdbiosciences.com">http://www.emdbiosciences.com</a>
Eppendorf	<a href="http://www.eppendorf.com">http://www.eppendorf.com</a>
Exalpha Biologicals	<a href="http://www.exalpha.com">http://www.exalpha.com</a>
GE Healthcare	<a href="http://www4.gelifesciences.com">http://www4.gelifesciences.com</a>
Genetix	<a href="http://www.genetix.com">http://www.genetix.com</a>
Genomic Solutions	<a href="http://www.genomicsolutions.com">http://www.genomicsolutions.com</a>
Gentra	<a href="http://www.gentra.com">http://www.gentra.com</a>
Gilson	<a href="http://www.gilson.com">http://www.gilson.com</a>
Globozymes	<a href="http://www.globozymes.com">http://www.globozymes.com</a>
GNF Systems	<a href="http://gnfsystems.com">http://gnfsystems.com</a>
Hamilton Robotics	<a href="http://www.hamiltoncomp.com">http://www.hamiltoncomp.com</a>
Hudson Control Group	<a href="http://www.hudsoncontrol.com">http://www.hudsoncontrol.com</a>

Company	Web address
Imaging Research	<a href="http://www.imagingresearch.com">http://www.imagingresearch.com</a>
Imagenex	<a href="http://www.imgenex.com">http://www.imgenex.com</a>
Improvision	<a href="http://www.improvision.com">http://www.improvision.com</a>
Intelligent Imaging Innovations	<a href="http://www.intelligent-imaging.com">http://www.intelligent-imaging.com</a>
Intrexon	<a href="http://www.intrexon.com">http://www.intrexon.com</a>
Invitrogen	<a href="http://www.invitrogen.com">http://www.invitrogen.com</a>
Jerini AG	<a href="http://www.jerini.de">http://www.jerini.de</a>
Kalypsys	<a href="http://www.kalypsys.com">http://www.kalypsys.com</a>
KBiosystems	<a href="http://www.kbiosystems.com">http://www.kbiosystems.com</a>
Kinasource	<a href="http://www.kinasource.co.uk">http://www.kinasource.co.uk</a>
Kinexus Bioinformatics	<a href="http://www.kinexus.ca">http://www.kinexus.ca</a>
Lab Services BV	<a href="http://www.lab-services.nl">http://www.lab-services.nl</a>
Labcyte	<a href="http://www.labcyte.com">http://www.labcyte.com</a>
LaVision Biotech	<a href="http://www.lavisionbiotech.de">http://www.lavisionbiotech.de</a>
Leica Microsystems	<a href="http://www.leica-microsystems.com">http://www.leica-microsystems.com</a>
MBL International	<a href="http://www.mblintl.com">http://www.mblintl.com</a>
Media Cybernetics	<a href="http://www.mediacy.com">http://www.mediacy.com</a>
Micro Luminetics	<a href="http://www.cryocam.com">http://www.cryocam.com</a>
Microtec Niton	<a href="http://niton.com/en">http://niton.com/en</a>
Millipore	<a href="http://www.millipore.com">http://www.millipore.com</a>
Mirero	<a href="http://www.gaia-zone.com">http://www.gaia-zone.com</a>
Molecular Devices	<a href="http://www.moleculardevices.com">http://www.moleculardevices.com</a>
New England Biolabs	<a href="http://www.neb.com">http://www.neb.com</a>
NIH Chemical Genomics Center	<a href="http://www.ncgc.nih.gov">http://www.ncgc.nih.gov</a>
Nikon	<a href="http://www.nikonusa.com">http://www.nikonusa.com</a>
Nventa	<a href="http://www.nventacorp.com">http://www.nventacorp.com</a>
Olympus	<a href="http://www.olympusamerica.com">http://www.olympusamerica.com</a>
Optronics	<a href="http://www.optronics.com">http://www.optronics.com</a>
Parallab	<a href="http://www.parallabs.com">http://www.parallabs.com</a>
Pepscan Systems	<a href="http://www.pepscan.nl">http://www.pepscan.nl</a>
Perkin-Elmer	<a href="http://www.perkinelmer.com">http://www.perkinelmer.com</a>
Pierce Biotechnology	<a href="http://www.piercenet.com">http://www.piercenet.com</a>
Phosphosolutions	<a href="http://www.phosphosolutions.com">http://www.phosphosolutions.com</a>
Promega	<a href="http://www.promega.com">http://www.promega.com</a>
Proteodyne	<a href="http://www.proteodyne.com">http://www.proteodyne.com</a>
PubChem	<a href="http://pubchem.ncbi.nlm.nih.gov">http://pubchem.ncbi.nlm.nih.gov</a>
Qiagen	<a href="http://www1.qiagen.com">http://www1.qiagen.com</a>
QImaging	<a href="http://www.qimaging.com">http://www.qimaging.com</a>
Reaction Biology Corp.	<a href="http://www.reactionbiology.com">http://www.reactionbiology.com</a>
Rigaku Automation	<a href="http://www.rigaku.com">http://www.rigaku.com</a>
Roche Applied Science	<a href="http://www.roche-applied-science.com">http://www.roche-applied-science.com</a>
Rockland Immunochemicals	<a href="http://www.rockland-inc.com">http://www.rockland-inc.com</a>
Roper Scientific	<a href="http://www.roperscientific.com">http://www.roperscientific.com</a>
RTS Life Science	<a href="http://www.rtlsifescience.com">http://www.rtlsifescience.com</a>
Scanalytics	<a href="http://www.scanalytics.com">http://www.scanalytics.com</a>
Sias	<a href="http://www.sias.ch">http://www.sias.ch</a>
Sigma-Aldrich	<a href="http://www.sigmaaldrich.com">http://www.sigmaaldrich.com</a>
Stratagene	<a href="http://www.stratagene.com">http://www.stratagene.com</a>
Synoptics	<a href="http://www.synoptics.co.uk">http://www.synoptics.co.uk</a>
Systat	<a href="http://www.systat.com">http://www.systat.com</a>
Tecan Group	<a href="http://www.tecan.com">http://www.tecan.com</a>
Thermo Scientific	<a href="http://www.thermo.com">http://www.thermo.com</a>

## SUPPLIERS GUIDE: COMPANIES OFFERING CHEMICAL BIOLOGY REAGENTS AND EQUIPMENT (CONTINUED)

Company	Web address
Titertek	<a href="http://www.titertek.com">http://www.titertek.com</a>
Torcon Instruments	<a href="http://www.torconinstruments.com">http://www.torconinstruments.com</a>
TTP Labtech	<a href="http://www.ttplabtech.com">http://www.ttplabtech.com</a>
Upstate	<a href="http://www.upstate.com">http://www.upstate.com</a>
USBiological	<a href="http://www.usbio.net">http://www.usbio.net</a>
VayTek	<a href="http://www.vaytek.com">http://www.vaytek.com</a>

Company	Web address
Wavemetrics	<a href="http://www.wavemetrics.com">http://www.wavemetrics.com</a>
Velocity11	<a href="http://www.velocity11.com">http://www.velocity11.com</a>
Xiril	<a href="http://www.xiril.com">http://www.xiril.com</a>
Zeiss	<a href="http://www.zeiss.com">http://www.zeiss.com</a>
Zinsser Analytic	<a href="http://www.zinsser-analytic.com">http://www.zinsser-analytic.com</a>