High-throughput screening for drug discovery

James R. Broach & Jeremy Thorner

Recent progress in elucidating the mechanisms underlying human disease has dramatically increased the number of protein targets available for potential drug treatment. Concurrently, new approaches have increased the number of compounds that can be tested for activity against these targets. Together, these trends have stimulated the adoption of high-throughput screening as a primary tool for early-stage drug discovery.

HIGH-THROUGHPUT screening (HTS) is the process by which large numbers of compounds can be tested, in an automated fashion, for activity as inhibitors (antagonists) or activators (agonists) of a particular biological target, such as a cell-surface receptor or a metabolic enzyme. The primary goal is to identify high-quality 'hits' or 'leads' (compounds that affect the target in the desired manner) that are active at a fairly low concentration and that have a new structure. The lower the concentration at which the compound acts, the more likely that it will exhibit specificity and, as a corollary, the less likely that it will have undesired side effects. If different chemotypes can be identified using the same screen then medicinal chemists will have a broader range of options for modifying the lead. The greater the number and diversity of compounds that are run through a screen, the more successful it is likely to be, a fact that further propels rapid developments in HTS.

Goals and limitations of HTS

Although a number of currently marketed drugs, such as cyclosporin A and mevastatin, have emerged directly from high-throughput screens, HTS generally does not actually identify a drug. Rather, the primary role of HTS is to detect lead compounds and supply directions for their optimization. This limitation exists because many properties critical to the development of a successful drug cannot be assessed by HTS; therefore, the final compound that eventually becomes a drug is unlikely to have been the molecule present in the initial library. HTS cannot evaluate: (1) bioavailability (a drug must be absorbed efficiently after oral ingestion and accumulate in the target tissue in reasonable quantities); (2) pharmacokinetics (a drug must persist in the body for a reasonable time period); (3) toxicity (a drug should have a minimum of nonspecific side effects); and (4) absolute specificity (a drug should act on the desired target in human physiology and nothing else). Thus, medicinal chemistry and pharmacological study are required to convert a compound that emerges from HTS into a useful drug. Finally, any screening process is subject to both false negatives (compounds active against the target that fail to score in the assay) and false positives (compounds that are not active against the target but score as hits in the assay). False negatives do not pose a significant problem, as long as a reasonable hit rate is achieved with a particular assay. Pursuing false positives, on the other hand, is a drain on time and resources; thus, suitable controls in the primary screen and stringent secondary assays are mandatory to confirm the authenticity of an initial hit.

A well designed HTS can provide information in addition to the potency of a compound. Information on specificity can be obtained by running concomitantly a counter-screen with a related target, for example HIV protease versus a cellular aspartyl protease, or the serotonin 2A receptor versus the serotonin 2C receptor. Compounds that exhibit activity only against the primary target probably possess the necessary selectivity and are less likely to be generally toxic. Information about specificity is also cumulative. When arrays of compounds are run

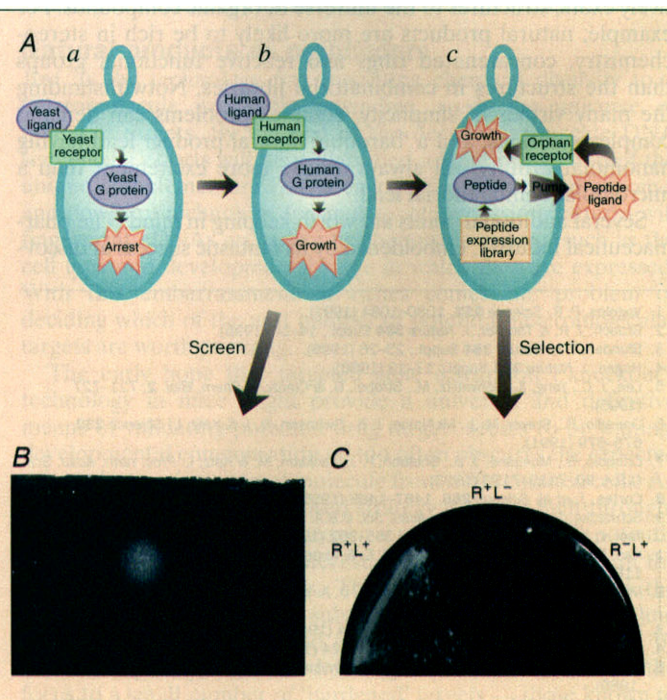

FIG. 1 Yeast-based screen and selection for identification of agonists for a human orphan G-protein-coupled receptor. *A*, Diagram of a yeast (*Saccharomyces cerevisiae*) strain. *A*, Its normal G-protein-coupled receptor-initiated signalling pathway results in cell-cycle arrest in response to ligand[22]. *b*, Substitution of a human GPCR (orphan receptor) and human Gα protein for the cognate yeast components and modification of the pathway output as described (ref. 9 and C. Klein, manuscript in preparation) yield a strain that can only grow if the receptor is stimulated. Thus, a receptor agonist can be detected by growth of the cells in the presence of the appropriate molecule from a library of compounds. *c*, Alternatively, an autocrine assay can be established by introducing into the test strain a random peptide expression library as described (ref. 23 and C. Klein, manuscript in preparation); a cell expressing a peptide that does not activate the receptor cannot grow, whereas a cell expressing a peptide that stimulates the receptor yields a colony. *B*, Individual compounds from a chemical library were spotted onto a lawn of a yeast strain, as in *Ab*, carrying a human orphan G-protein-coupled receptor. Agonists for the receptor are detected by growth of the strain on the plate around the site of compound application. A control plate (not shown) with a strain carrying a different receptor does not grow in response to any of the compounds. *C*, Peptide agonist for the orphan receptor isolated by selection for growth following transformation of the receptor-bearing strain with a random peptide expression library (ref. 23 and C. J. Klein *et al.*, in preparation), as in *Ac*. Strains that produce both the receptor and the peptide ligand (R^+ L^+) are capable of growth, whereas those that lack either the ligand (R^+ L^-) or the receptor (R^- L^+) fail to grow, as indicated.

repeatedly against different targets, past performance of a compound can provide indications about its specificity in the current screen. Also, the spectrum of compounds scoring positive (and to some extent those that do not) help pinpoint which structural features of the molecules are responsible for their efficacy. Such preliminary structure–activity relationships help the further optimization of the lead. Finally, in an HTS using live cells, some indication about the cytotoxicity of a hit can be obtained. These considerations highlight the need for computer-based methods for storage and retrieval of the voluminous data generated by HTS.

Implementation of HTS

An HTS requires four elements: (1) suitably arrayed compound libraries; (2) an assay method configured for automation; (3) a robotics workstation; and (4) a computerized system for handling the data. The 96-well microtitre plate is the standard format for automated assays, although arrays of compounds on chips[1] or on beads[2] are also used and assays can be performed on agar plates or other solid support (Fig. 1). Synthesis of combinatorial libraries can be accomplished in microtitre plates, thereby providing addresses for particular compounds generated by a given subset or series of reactions and thus identifying the compound. Moreover, concentrates of fermentation broths or natural product extracts, or pre-existing collections of chemicals (such as the repositories possessed by large pharmaceutical firms), can be dispensed in 96-well plates, either singly (simplex arrays) or as defined mixtures of 10–20 compounds per well (multiplex arrays). The latter approach permits a faster rate of screening, but requires deconvolution of the mixture to identify the true active component. Nonetheless, the more diverse (in source and structure) the compound collections used for HTS, the higher the likelihood of success.

Robotics systems for HTS range from simple automated dilution devices to complex workstations in which multiple functions are performed by one or more mechanical arms. Full automation — from sample dispensing to data collection — allows for round-the-clock operation, thereby increasing the screening rate. Given the variety of chemical libraries available, the large number of compounds in each, and the need to compare results from different screens, data collection and management are critical to automated HTS. Data bases of structures, assays performed, screening results and so on, must be relational (interlinked) so that necessary information can be extracted by a query from any perspective. One should be able to search for all compounds active at a certain threshold level in a particular screen, or for the behaviour of all compounds of similar structure in different screens.

Although any assay performed on the benchtop can, in theory, be applied in HTS, conversion to an automated format imposes certain constraints that affect the design of the assay in practice. Procedures that are routine at the bench (for example, centrifugation to remove debris or to collect beads, rinsing wells for enzyme-linked immunosorbent assays or radioimmunoassays, or fractionation techniques for separating substrate from product) are often extremely difficult to automate. Also, the more steps required for an assay, the more difficult to automate the HTS. The ideal assay is one that can be performed in a single well with no other manipulation other than simple injection of the sample to be tested.

Many types of *in vitro* assays can be readily converted to HTS. Even binding activity, such as ligand–receptor interaction (for example, platelet-derived growth factor (PDGF) docking to the extracellular domain of PDGF receptor) or protein–protein interaction (for example, src SH2 domain association with a tyrosine-phosphorylated site) can be assessed using proximity-dependent transfer methods. Such assays can be performed in homogeneous mixtures requiring no additions or fractionations because the output of such assays derives from the signal enhancement generated by bringing a source and a distance-dependent amplifier close together. A number of light-based readouts can be produced in this way. For example, the β-particles of a low-energy radionuclide attached to a ligand will stimulate the fluorescent emission of a scintillant in a bead to which the ligand's receptor is attached[3]. As another example, the rare earth lanthanide Eu^{2+}, when irradiated by light, can transfer its excitation energy in a nonradiative process to the fluorescent protein, allophycocyanin (APC), if the two are close (<90 Å), which can occur when a Eu^{2+}-derivitized ligand binds to an APC-linked receptor[4].

Cell-based assays are an increasingly attractive alternative to *in vitro* biochemical assays for HTS. Such *in vivo* assays require the ability to examine a specific cellular process and a means to measure its output. For instance, agonist activation of a cell-surface receptor or a ligand-gated ion channel can be followed by monitoring a coupled cellular response[5]. Although the immediate downstream event (for example, increased phosphatidylinositol-4,5-bisphosphate turnover, transient elevation in cytosolic Ca^{2+}, or phosphorylation of target proteins) may be difficult to evaluate in an automated format, the subsequent transcriptional changes can be more easily monitored. For example, binding of isoproterenol to β-adrenergic receptor causes a transient rise in cyclic AMP level, activating protein kinase A (PKA), which translocates to the nucleus and phosphorylates a transcription factor (CREB) that recognizes cAMP response elements (CREs); CREB activation can be detected and quantified by measuring the expression level of a reporter gene whose transcription is driven by an enhancer containing CREs[6]. Enhancer elements that couple gene expression to distinct signal transduction pathways are now known[7] and reporter genes that

TABLE 1 Reporter genes useful for cell-based HTS

Reporter genes	Advantages	Disadvantages
β-Galactosidase (bacterial)	Well characterized; stable; inexpensive substrates; little interference from test compounds; simple readouts (readily automated)	Endogenous activity (mammalian cells); tetrameric (nonlinear response at low concentration)
Luciferase (firefly)	Dimeric; high specific activity; no endogenous activity (low background)	Requires addition of cofactor (luciferin) and presence of O_2 and ATP
Alkaline phosphatase (human placental)	Secreted protein (avoids need for membrane-permeable substrates); inexpensive colorimetric and highly sensitive luminescent assays available[19]	Endogenous activity in some cell types; optimal at pH 9.8
β-Lactamase (bacterial)	Monomeric; highly sensitive, membrane-permeant, fluorogenic substrates available*; no endogenous activity	Membrane-permeant fluorescent substrates not yet commercially available
Green fluorescent protein (jellyfish)[20,21]	Monomeric; no substrate needed (no manipulations required for assay); no endogenous activity; brighter mutants and colour variants available	Detection is relatively insensitive because of lack of enzymatic amplification

*G. Zlokarnik & R. Y. Tsien, personal communication.

generate products that can be adapted to the HTS format are available (Table 1).

Cell-based assays have notable advantages over *in vitro* assays. First, the starting material (the cell) self-replicates, avoiding the investment involved in preparing a purified target, in chemically modifying the target to suit the screen and so on. Second, the targets and readouts are examined in a biological context that hopefully mimics the normal physiological situation. Third, cell-based assays can provide insights about bioavailability (a compound must enter the cell to affect an intracellular target) and cytotoxicity (whether a compound compromises cell growth). However, mammalian cells are expensive to culture and difficult to propagate in the automated systems used for HTS.

An alternative is to recapitulate the desired human physiological process in a microorganism, such as yeast. For instance, signalling through human G-protein-coupled receptors (GPCRs) has been reconstituted in yeast (refs 8, 9 and C. Klein *et al.*, manuscript in preparation) to yield a facile growth response or a reporter gene readout (Fig. 1). Similarly, mammalian ion channels can be reconstituted in yeast to yield a readily assayed growth response[10]. In addition, protein–protein interactions, including RAS–RAF association[11] and peptide hormone receptor binding[12], have been faithfully reproduced using the yeast two-hybrid system. A variation on this theme, dubbed the tribrid system[13], permits scrutiny of tyrosine phosphorylation-dependent protein–protein interactions in yeast, and other methods detect activation of receptor-tyrosine kinases[14] or intracellular tyrosine kinases[15] directly in yeast. Finally, many mammalian transcription factors operate in yeast, including glucocorticoid receptor[16], the RAR and RXR families of receptors[17], and JAK-activated STATs (B. H. Cochran, personal communication).

The ease and low cost of growing yeast, their ready genetic manipulation, and their resistance to solvents make yeast an attractive option for cell-based HTS. Although yeasts are surrounded by a cell wall and their plasma membrane composition differs somewhat from that of their mammalian counterpart, any compound that can penetrate a yeast cell should have no problem entering a human cell. On the other hand, despite their drawbacks, screening in mammalian cells does have some advantages, such as a context for the clinically relevant target that presumably resembles more closely the milieu found in human tissues, greater amenability to fluorescent readouts of cell responses (such as Ca^{2+} transients, membrane potential changes, and reporter genes), and better drug permeability. Thus, there are trade-offs in the choice of a system for a cell-based screen.

Future directions for HTS

In the next few years, quantitative changes in both genetic and instrumentation engineering, including advances in nanotechnology, will undoubtedly extend the purview, and accelerate the pace of, HTS. For instance, new reporter genes and substrates that give sensitive fluorescent readouts without cell disruption should allow rapid tuning of cell lines and miniaturization of transcriptional assays (G. Zlokarnik and R. Y. Tsien, personal communication), resulting in faster screening with much less material. It seems likely that another fundamental feature of cell-based assays — the possibility of devising conditions wherein the desired compound provides a growth advantage — will cause a qualitative shift in how HTS is conducted. Application of such positive selections should allow examination of hundreds-of-thousands to millions of compounds per day, rather than the tens-of-thousands now accessible using the best screening formats. Phage display, in which panning with a target permits isolation of those rare phage in a population that carry an engineered peptide sequence in the coat protein that mediates interaction with the target, was one of the first uses of selection in drug discovery[18]. More recently, expression of random peptide libraries in a yeast strain designed to respond to activation of an incorporated GPCR by growth (Fig. 1) has allowed direct selection of new peptide ligands for orphan GPCRs (C. Klein *et al.*, manuscript in preparation). The next stage in the further development of this kind of strategy could be to introduce an enzyme or an entire pathway of therapeutic interest into a test strain designed to respond to potential drug leads by growth.

Finally, advances in combinatorial chemistry interface well with HTS approaches based on either a growth readout assay or a highly sensitive reporter system. For example, hundreds-of-thousands of tagged beads[2], each bearing a different compound, could be sprinkled onto a lawn of yeast or mammalian cells designed to grow or change fluorescence in response to an appropriate agonist (or antagonist) for the target of interest. Beads eliciting a growth (or light) response could be recovered and the identity of the hits determined by appropriate decoding.

The goal of HTS is to provide useful drug leads. The ability to accomplish this objective has been enhanced by the advent of methods by which large numbers of assays can be performed on highly specific and biologically relevant targets using rich and highly diverse libraries. Conversely, the ability to survey vast arrays of chemicals and the application of selection procedures in HTS raises the possibility that these methods can be used to identify systematically the biological function of new genes discovered by the human genome project.

This new tack may soon reverse the standard approach to drug discovery. Rather than relying on thorough characterization of a target before embarking on a drug discovery programme, HTS should allow exploration of a significantly larger number of targets and provide the means to identify the reagents necessary to determine whether any given target is relevant to a disease state of interest. Thus, further developments in HTS will ensure that detection of initial leads is not the rate-limiting step in drug discovery and will provide new avenues for the treatment and prevention of disease. □

James R. Broach is in the Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA and at Cadus Pharmaceuticals Inc., 777 Old Saw Mill Road, Tarrytown, New York 10591-6705, USA. Jeremy Thorner is in the Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720-3202, USA.

1. Fodor, S. *et al. Nature* **364**, 555–556 (1993).
2. Ohlmeyer, M. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 10922–10926 (1993).
3. Udenfriend, S., Gerber, L. D., Brink, L. & Spector, S. *Analyt. Biochem.* **161**, 494–500 (1985).
4. Mathis, G. *Clin. Chem.* **39**, 1953–1959 (1993).
5. Levitzki, A. *Curr. Opin. Cell Biol.* **8**, 239–244 (1996).
6. Kee, B. I., Arias, J. & Montminy, M. R. *J. biol. Chem.* **271**, 2373–2375 (1996).
7. Treisman, R. *Curr. Opin. Cell Biol.* **8**, 205–215 (1996).
8. King, K., Dohlman, H. G., Thorner, J., Caron, M. G. & Lefkowitz, R. J. *Science* **250**, 121–123 (1990).
9. Price, L. A., Kajkowski, E. M., Hadcock, J. R., Ozenberger, B. A. & Pausch, M. H. *Molec. cell. Biol.* **15**, 6188–6195 (1995).
10. Tang, W. *et al. Molec. Biol. Cell* **6**, 1231–1240 (1995).
11. Van Aelst, L., Barr, M., Marcus, S., Polverino, A. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6213–6217 (1993).
12. Ozenberger, B. A. & Young, K. H. *Molec. Endocrin.* **9**, 1321–1329 (1995).
13. Osborne, M. A., Dalton, S. & Kochan, J. P. *Bio/technology* **13**, 1474–1478 (1995).
14. Kinoshita, N., Minshull, J. & Kirschner, M. W. *Cell* **83**, 621–630 (1995).
15. Superti-Furga, G., Johnsson, K. & Courtneidge, S. A. *Nature Biotechnol.* **14**, 600–605 (1996).
16. Kralli, A., Bohen, S. P. & Yamamoto, K. R. *Proc. natn. Acad. Sci. U.S.A.* **92**, 4701–4705 (1995).
17. Durand, B. *et al. EMBO J.* **13**, 5370–5382 (1994).
18. O'Neil, K. T. & Hoess, R. H. *Curr. Opin. struct. Biol.* **5**, 443–449 (1995).
19. Berger, J., Hauber, J., Hauber, R., Geiger, R. & Cullen, B. R. *Gene* **66**, 1–10 (1988).
20. Cubitt, A. B. *et al. Trends biochem. Sci.* **20**, 448–455 (1995).
21. Cheng, L. J. Fu, Tsukamoto, A. & Hawley, R. G. *Nature Biotechnol.* **14**, 606–609 (1996).
22. Bardwell, L. S., Cook, J. G., Inouye, C. J. & Thorner. J. *Devl Biol.* **166**, 363–379 (1994).
23. Manfredi, J. *et al. Molec. cell. Biol.* (in the press).

ACKNOWLEDGEMENTS. We thank P. Fernandes (Bristol-Myers Squibb), R. Tsien (HHMI and UC San Diego) and W. Wiesler (Cadus Pharmaceuticals), for their valuable insights into high-throughput screening, assay development and other aspects of the drug discovery process, L. Enquist, A. N. Glazer, J. Rine and C. Klein for their thoughtful comments on the manuscript and D. Webb, C. Klein and W. Wiesler for providing materials for Fig. 1.