Cytotechnology **38:** 3–10, 2002.

Chemical genomics: massively parallel technologies for rapid lead identification and target validation

Uli Hacksell*, Norman Nash, Ethan S. Burstein, Fabrice Piu, Glenn Croston & Mark R. Brann
ACADIA Pharmaceuticals Inc., 3911 Sorrento Valley Boulevard, San Diego, CA 92121, USA
(Author for correspondence; E-mail: uhacksell@acadia-pharm.com; Fax: +1 (858) 558-2872)*

Accepted 31 March 2002

Key words: 5-HT_{2A} inverse agonist, chemical genomics, m1 agonist, GPCR, psychosis, R-SAT, uHTS

Abstract

Chemical genomics is a new research paradigm with important applications in drug discovery. It links genomic targets with small-molecule chemistries thereby allowing for efficient target validation and lead compound identification. ACADIA's chemical-genomics platform consists of a large and diverse small-molecule library (800,000), a reference drug library (2,000), druggable genomic targets (>300) and a cell-based functional assay technology (R-SAT^{TM}; Receptor Selection and Amplification Technology) that allows for ultra-high throughput screening (>500,000 data points/week) as well as high throughput pharmacology and profiling over a wide range of targets. Two examples are presented that illustrate the success of our chemical-genomics approach: (i) The validation of inverse agonism at serotonin 5-HT_{2A} receptors as an antipsychotic mechanism and the subsequent discovery of potent and selectively acting 5-HT_{2A} inverse agonists, currently in preclinical development, and (ii) the discovery of the first ectopically binding subtype-selective muscarinic m1 agonist.

Abbreviations: AD, Alzheimer's disease; GPCR, G protein coupled receptor; R-SAT, Receptor selection and amplification technology; TIP, Target-interaction profile; uHTS, ultra high throughput screening.

Introduction

The unprecedented progress in biological sciences during the last two decades, most recently demonstrated by the deciphering of the human genome (Lander, 2001; Venter, 2001), has provided scientists with an enormous knowledge base as well as sophisticated research tools and technologies. Unfortunately, the scientific progress has not yet translated into a more efficient/productive drug discovery and development process. In fact, during the past two decades a trend has emerged in which fewer new chemical entities (NCEs) have entered the market (Fig. 1) despite the fact that more resources have been allocated to pharmaceutical research and development (R&D) (cf Drews, 1996). In part, these opposing trends may be explained by the increasing demands on NCEs posed by regulatory authorities and the delay caused by the slow drug R&D process itself (it takes 10–15 years to proceed from the research stage to a marketed drug). More surprisingly, it is apparent that the adoption of new technologies in biotechnology and pharmaceutical firms has led to little increase in the productivity and quality of the discovery process. The drug discovery process may even have been slowed down because the 'modern/rational drug discovery paradigm' starts with a costly and time consuming target validation process using sophisticated genomic tools. These tools include linkage studies that attempt to couple genetic traits to a disease, transgenic animals in which particular genes are knocked out or knocked in, and array technologies providing extremely complex gene expression patterns caused by various interventions. It now appears that some drug discovery efforts may involve an even greater degree of early complexity by the addition of proteomics information (Pandey, 2000) to help in target validation. Beyond new technologies, new research processes are needed to increase drug discovery productivity.

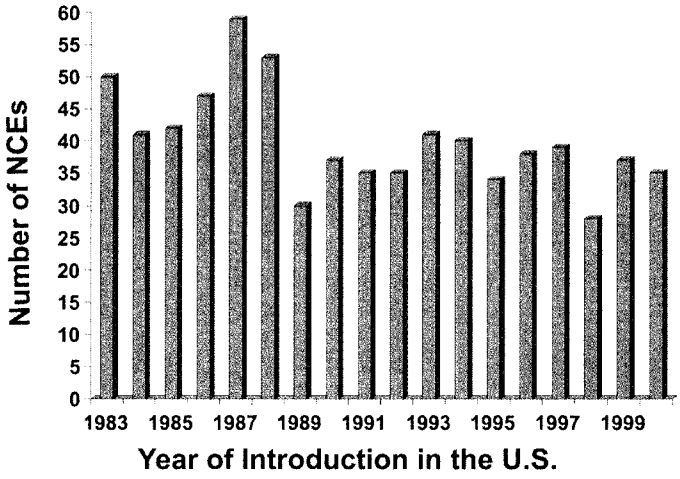

Figure 1. New chemical entities (NCEs) approved as drugs by the FDA from 1983 – 2000. Despite the unprecedented increase in R&D spending and a number of scientific breakthroughs in the past two decades, the number of approved new chemical entities (NCEs) has not been increasing.

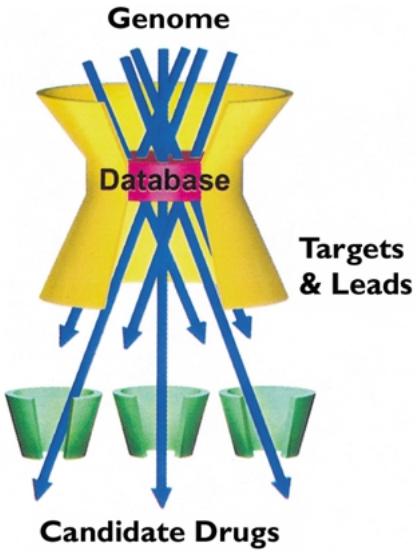

Figure 2. Chemical-genomics strategy. A massively parallel, chemistry driven drug discovery process in which ultra-high throughput functional screening, high throughput pharmacology and profiling over a wide range of targets provide a foundation for prioritization, target validation and lead compound identification. Sophisticated modern genomic technologies are applied after lead identification and target selection to confirm the initial target validation. Application of a chemical-genomics approach as opposed to a sequential functional genomics process for target validation and drug discovery shortens the time from target identification to candidate drug by about 50%.

Chemical-genomics

A new drug discovery paradigm, chemical genomics (or chemogenomics), can be used to speed up drug discovery while still efficiently exploiting the information offered by modern genomic technologies (cf. Willson, 2001; Caron, 2001; Landro, 2000 and Lenz, 2000). The chemical-genomics approach involves the identification of small-molecule tool compounds and their use to establish the function of genomic targets and the involvement of targets in pathophysiology, thereby validating targets as worthy of drug discovery efforts. In order for a chemical-genomics effort to succeed one needs access to a technology platform that allows for massively parallel screening of large and diverse small-compound libraries over all relevant genomic drug targets (Fig. 2). Such an effort will link complementary areas in genomic and chemical space, that is, identify small-molecule ligands that interact with the various targets. The next step is to assay selectivity and thus identify the compounds that interact selectively with only one of the targets. Such target-specific small molecules can then be used as research tools to probe and establish the function of the individual targets. Scientists have been using tool compounds to elucidate biochemical and pharmacological mechanisms for more than a century (Drews, 2000 and Triggle, 2000). Although many tool compounds were believed to be acting specifically, we now recognize that they often interacted with a variety of targets that were known or unknown at that time. The difference is that we now

have access to all the targets of interest and, therefore, can truly assess the target specificity of the tool compounds.

A productive chemical-genomics platform

ACADIA has built a uniquely broad and productive chemical-genomics platform (Fig. 3). It consists of two chemical libraries, the diverse compound library (800,000 compounds) and the evidence-based compound library (2000 drugs), and a large number (>300) of genomic targets. These targets belong to gene families known to be druggable, GPCRs (G protein coupled receptors) and nuclear receptors, as well as many other gene families like intracellular kinases whose potential as drug targets is less exploited. The link between these chemistries and these genomic targets is provided by our functional assay technology, Receptor Selection and Amplification Technology (R-SAT™), (Messier, 1995) which is essentially universally applicable and provides us with uHTS (ultra-high throughput screening), high throughput pharmacology and profiling capabilities (Croston, 2002).

Our assay technology, R-SAT (Fig. 4), is at the heart of our chemical-genomics platform. It harnesses the molecular steps that drive oncogenic cell growth to provide a convenient and homogenous read-out for drug and gene screening assays. Our scientists have engineered cells so that a vast array of genetically-defined drug targets can be linked to genes that control oncogenic cellular growth. It is important to note, however, that this engineering process does not alter the drug target but only co-opts the downstream signaling pathway of a drug target. Drug interactions with targets in R-SAT assays are measured as an increase or decrease in cell growth (regardless of the native signaling pathway). All assays are performed on targets in their native, non-mutated state, so that drug-target interactions occur in an environment similar to the normal cellular milieu.

In the simplest version of R-SAT, a gene representing a desired drug target is transiently transfected into a large batch of cultured cells. These cells are exposed to individually plated, small-molecule compounds. Agonists for the drug target induce partial oncogenic cellular transformation. This release from contact inhibition results in explosive cell growth in cells expressing the target while all other cells remain quiescent. Cells expressing the target are, thus, selected and amplified when exposed to an agonist. In many cases, drug targets exhibit constitutive activity in this system and oncogenic cell growth is accelerated in the absence of a ligand (merely upon expression of the target alone in these cells). In this manner orphan receptors can be readily assayed. Under these conditions, inverse agonists can be identified by their ability to suppress constitutive activity. Alternately, neutral antagonists can be identified in R-SAT by their ability to block the actions of agonists and inverse agonists. Thus, R-SAT returns the full spectrum of potential pharmacologic activities using one asssy system.

Multiple drug targets can be screened simultaneously using R-SAT. As depicted in Figure 4, cells can be transiently transfected with sets of genes for desired drug targets. An agonist interaction with any one of the targets in the mix will induce oncogenic growth in those cells containing a drug-specific target while the other cells that do not contain this target will not grow. Simple deconvolution using individual target assays determines which of the targets in the multiplex is driving the agonist interaction. In this way large numbers of drug targets can be screened against large numbers of compounds adding to the ultra-high throughput of this system.

In practice, target genes are co-transfected with reporter genes such as beta-galactosidase. In the quiescent cells the reporter gene product diminshes over time, while the number of reporter gene products increases within the oncogenically stimulated cells. Thus, drug/target interactions can be detected in simple colorimetric assays or other efficient detection systems. This application does not rely on transcriptional regulation of marker genes and thus represents a distinct and novel technology relative to the various proprietary transcription-based reporter gene assays.

Our experience in the area of chemical genomics is very positive. We have screened our diversity library against over 100 genomic targets leading to the identification of selective chemotypes for 62 targets. These targets include a wide variety of GPCRs and nuclear receptors. Our efforts have already generated 16 projects and programs resulting from solid target validations and identified selective lead chemistries. To illustrate this success two case stories are reviewed below. The first describes a successful attempt to find target specific chemistries in an area where others have failed. The specific chemistries discovered are currently being used to validate the function of the muscarinic m1 receptor and its involvement in various disease states. The second example describes how we applied our chemical-genomics platform to

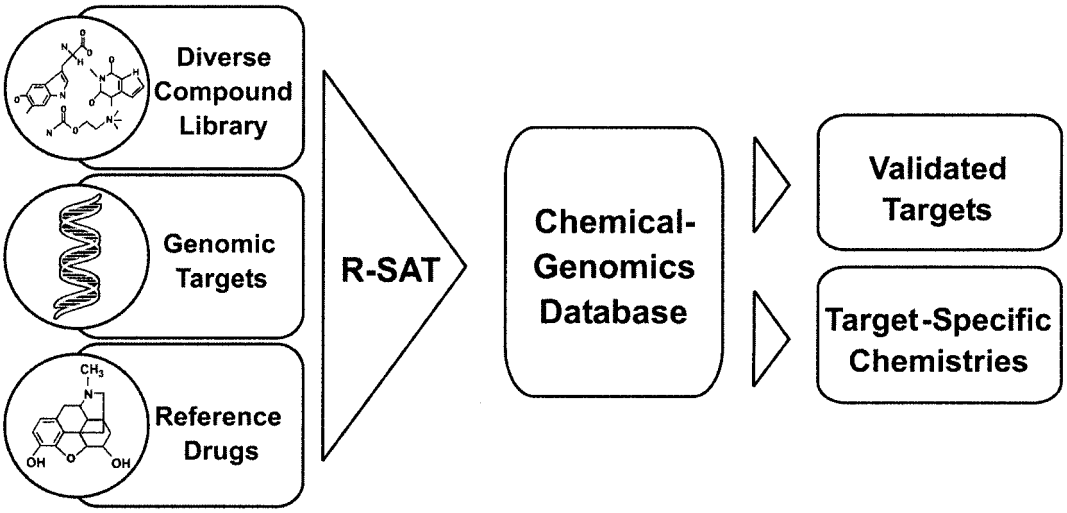

Figure 3. A chemical-genomics platform that allows for rapid target validation and lead identification. Its key-components are briefly described below. Diverse compound library: Approximately 800,000 small-molecule individual compounds collected from about 100 different sources for uHTS. The library is being modified and expanded based on chemical diversity considerations and empirical screening data. Evidence-based library: A comprehensive collection (>2000) of compounds with known clinical CNS effects and a wide range of clinical uses. Genomic targets: a large number of genomic targets have been collected using strong internal genome mining, molecular cloning and signal transduction expertise. Currently, more than 300 of these are on line as assays. The majority of these targets are GPCRs and nuclear receptors. R-SAT (Receptor Selection and Amplification Technology): This technology provides rapid lead identification by integrating functional HTS, pharmacology and profiling. It can be applied to a wide range of genomic targets and can be multiplexed such that many targets can be screened simultaneously. See also Figure 6.

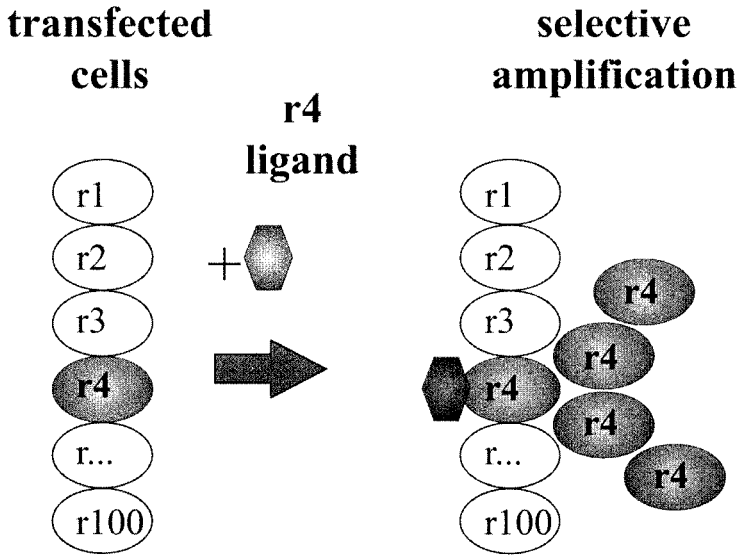

Figure 4. R-SAT is based on the principle of genetic selection: cells which express the receptor for an agonist ligand will be selected and amplified relative to cells that do not express a receptor for the ligand. r1, r2 . . . are different transiently transfected receptors. r4 is a receptor that is stimulated by the added ligand.

obtain evidence-based target validation of a novel antipsychotic mechanism of action and how we quickly moved on to discover and develop a drug candidate in this area.

Ectopically binding muscarinic m1 agonists

Based on its predominant distribution in the cerebral cortex and hippocampus, areas involved in higher order cognitive functions and in Alzheimer's disease (AD) pathology, the m1 muscarinic receptor subtype appears to be the preferred target for anti-dementia drugs. Indeed, many companies over the past 15 years have attempted to discover and develop m1-selective agonists. Despite initial claims to the contrary, no truly m1 receptor selective agonist has emerged from these efforts (Wood, 1999; Eglen, 1999).

We have identified the first truly m1-selective muscarinic agonists. Compounds from this chemical series interact with a unique ectopic site involving the non-conserved regions of the receptor in the putative 1st and 7th transmembrane domains. These compounds have no agonist activity at any other muscarinic receptor or at any of a large number of receptors that have been examined. We also have demonstrated that compounds from this structural class suppress spontaneous and anticholinergic-induced locomotor activity in mice but do not induce cholinergically-mediated, side-effects (e.g., salivation) in these animals unlike other muscarinic agonists. These compounds currently are being optimized and will be considered for future development as anti-dementia agents aimed at treating the behavioral impairments associated with AD. As part of our chemical-genomics effort we also use our m1 specific agonists as pharmacological tool compounds in relevant disease models to fully explore the therapeutic potential of drugs with this unique profile.

Our lead compound in the muscarinic program, AC-42 (Fig. 5), was identified by functional (R-SAT) screening of our diversity library (Spalding, 2002). The reason why we successfully identified an m1-specific agonist while big pharmaceutical firms repeatedly failed is most likely related to differences in screening protocols. Since our assay was functional we could detect agonist activity regardless of the site of interaction between the receptor and the small-molecule agonist. In contrast, the pharmaceutical firms used radioligand binding assays and screened for affinity, not function. Hence, they were unable to find ectopically binding agonists because the radioligands used interacted with conserved regions in the muscarinic receptors. As stated above, our m1-specific agonists interact with an ectopic site that is not conserved in the muscarinic family and is physically distinct from the binding site utilized by acetylcholine and other classical muscarinic agonists.

Evidence-based target validation

Because the currently used drugs were discovered before research tools had been developed to enable rigorous assessment of target specificity, most of the marketed drugs have non-selective pharmacological profiles. This means that by establishing what targets each of these drugs interact with and by comparing these data with their clinical effects one might be able to correlate target interactions with clinical effects. Since this appeared to be an attractive alternative approach to traditional target validation and as we knew that we had created a technology platform that allowed us to profile a large number of drugs over a large number of genomic targets, we started to assemble a library of clinically used drugs with a focus on those with CNS activities (the evidence-based library). Today, this library consists of about 2000 drugs.

High throughput profiling of the evidence-based library (Figure 6) provided us with target-interaction profiles (TIPs) of a large number of CNS drugs. Most of these drugs have well-established clinical profiles and are relatively non-specific in terms of their mode of action. An internal database that provides access to the TIPs now allows us to correlate specific target interactions with clinical effects (therapeutic or side-effects). When studying the TIPs of the antipsychotic drug classes it became apparent that in addition to their antagonistic interaction with dopamine D_2 receptors, almost all of these drugs are fully efficacious inverse agonists at the serotonin 5-HT_{2A} receptors. Even more interesting was that the 5-HT_{2A} inverse agonist potency was particularly pronounced among the modern, atypical antipsychotic agents (Weiner, 2001). Having obtained this initial validation of the 5-HT_{2A} inverse agonist mechanism we embarked on an antipsychotic drug discovery project. uHTS using assays configured to detect the natural constitutive activity of the 5-HT_{2A} receptor provided a wealth of chemistries with inverse efficacy at 5-HT_{2A}. This allowed us to pick a chemical starting point for a discovery program on the basis of potency, efficacy, specificity, drug-likeness, synthetic (combinatorial) feasibility and patentability. One of

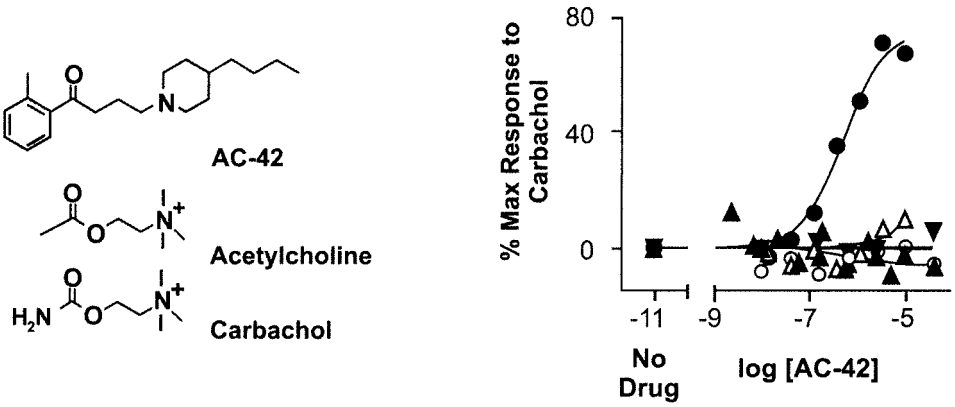

Figure 5. Chemical structure of the m1 selective agonist AC-42 the non-selective muscarinic agonist, carbachol (CCh) and the endogenous muscarinic ligand acetylcholine (ACh) and agonist activity of AC-42 on the m1 (●), m2 (○), m3 (▲), m4 (△) and m5 (■) muscarinic receptor subtypes. Assays were carried out using R-SAT. Points represent the mean of duplicate determinations. Data are typical of over 10 experiments. Typical pEC50 values for carbachol in this assay are (mean +/- S.D.): m1: 6.1 +/- 0.2; m2: 6.4 +/- 0.4; m3: 6.4 +/- 0.2; m4: 6.5 +/- 0.3; m5: 6.3 +/- 0.5 (From Spalding, 2002).

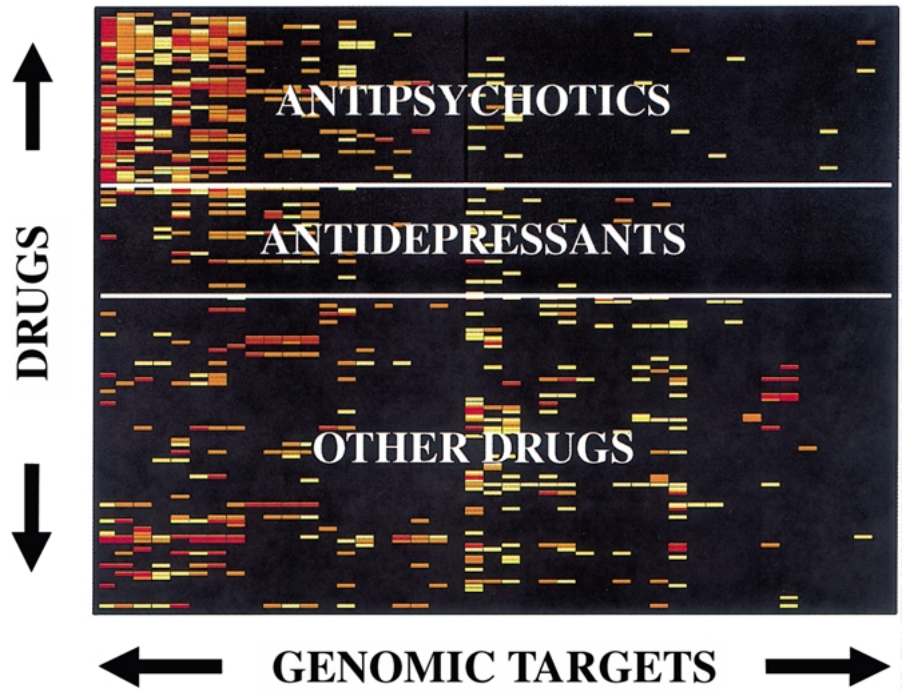

Figure 6. Heat map of Target Interaction Profiles (TIPs) resulting from high throughput profiling of 250 CNS-active drugs against monoaminergic GPC receptors (data reflect potencies as agonists and antagonists or inverse agonists at these receptors). The targets studied include 10 serotonin, 8 adrenergic, 5 muscarinic, 4 dopaminergic and 3 histaminergic receptors. Full dose response curves were generated to determine the drug potencies. Color coding: red = potency < 10nM, orange = potency between 10 and 100 nM, yellow = potency between 100 nM and 1000 nM, black indicates potency less than 1000 nM or no activity.

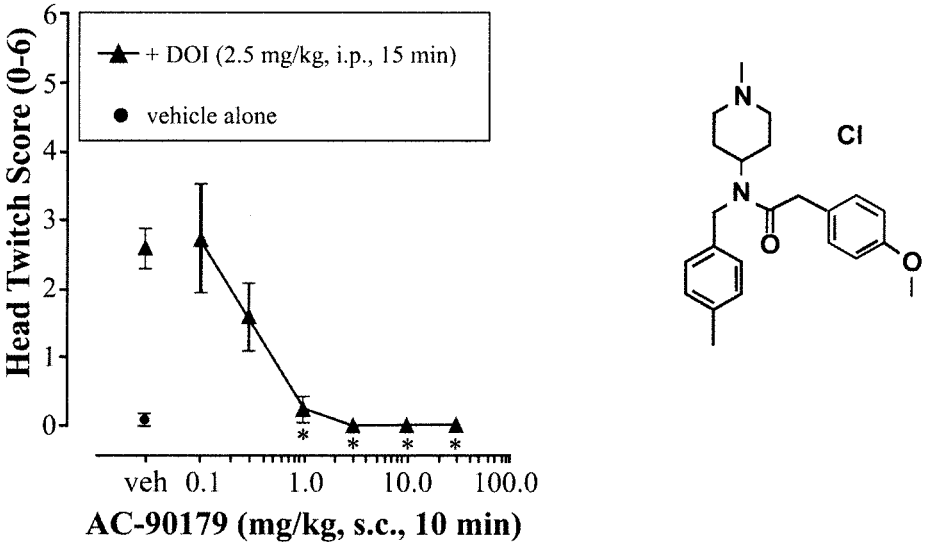

Figure 7. Antipsychotic-like behavioral effect of AC-90179, head twitch score as a function of AC-90179 dose, and the chemical structure of AC-90179. Filled triangles, AC-90179 or vehicle in combination with 2.5 mg/kg DOI; filled circle, vehicle dosed alone. Each point represents the mean (n = 8-24), and vertical lines represent standard error of the mean. Asterisks indicate statistical significance ($p < 0.05$) compared with the respective vehicle control. (From Weiner, 2001).

the confirmed hits, AC-90179, fulfilled most of these criteria and was selected as a starting point for optimization efforts. In fact, the potency and specificity of AC-90179 allowed us to use it for pharmacological validation of the 5-HT_{2A} hypothesis in a pivotal psychosis model (Fig 7). Optimization of AC-90179 provided ACP-103, a development candidate which is currently in IND-track development.

Conclusions

Chemical genomics is a practical and powerful approach to drug discovery that allows one to focus resources on druggable targets where chemistries are available that enable both target validation and uniquely efficient drug discovery efforts. However, a truly successful chemical-genomics effort requires access to chemistries (large and optimally diverse libraries) as well as assay technologies that provide the opportunity for functional uHTS. It is apparent that affinity screening is suboptimal as it is unable to identify ectopically binding chemistries and does not allow for screening of orphan receptors. Having the requisite technologies available in house, we have demonstrated that a small organization which applies a chemical-genomics approach can be very productive. Since 1997, our efforts have resulted in 16 drug discovery projects/programs, three of which are already in development.

References

Caron PR, Mullican MD, Mashal RD, Wilson KP, Su MS & Murcko MA (2001) Chemogenomics: a gene family approach to parallel drug discover. Drug Discovery World 2: 57–62.

Croston GE (2002) Functional and Cell-Based uHTS in Chemical Genomic Drug Discovery. Trends Biotech 20:110–115

Debouck C & Metcalf B (2000) The impact of genomics on drug discovery. Ann Rev Pharmacol Toxicol 40: 193–207.

Drews J & Ryser S (1996) Innovation deficit in the pharmaceutical industry. Drug Info J 30: 97–107.

Drews J (2000) Drug discovery: A historical perspective. Science 287: 1960–1964.

Eglen, ME, Choppin A, Dillon MP & Hegde S (1999) Muscarinic receptor ligands and their therapeutic potential. Curr Op Chem Biol 3: 426–432.

Harris S & Foord SM (2000) Transgenic gene knock-outs: functional genomics and therapeutic target selection. Pharmacogenomics 1: 433–443.

Lander ES et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.

Landro JA, Taylor ICA, Stirtan WG, Osterman DG, Kristie J, Hunnicutt EJ, Rae PMM & Sweetnam PM (2000) HTS of the new millennium. The role of pharmacology and flexibility. J Pharmacol Toxicol Method 44:273–289.

Lenz GR, Nash HM & Jindal S (2000) Chemical ligands, genomics and drug discovery. Drug Discovery Today 5: 145–156.

Messier TL, Dorman CM, Brauner-Osborne H, Eubanks D & Brann MR (1995) High throughput assays of cloned adrener-

gic, muscarinic, neurokinin, and neurotrophin receptors in living mammalian cells. Pharmacol Toxicol 76: 308–311.

Pandey A & Mann M (2000) Proteomics to study genes and genomes. Nature 405: 837–846.

Spalding TA, Trotter C, Skjærbæck N, Messier TL, Currier EA, Burstein ES, Li D, Hacksell U & Brann MR (2002) Discovery of an Ectopic Activation Site on the m1 Muscarinic Receptor. Mol Pharmacol 61: 1297–1302.

Szymkowski DE (2001) Too many targets, not enough target validation. Drug Discovery Today 6: 397.

Triggle DJ (2000) Pharmacological receptors: a century of discovery – and more. Pharm Acta Helv 74: 79–84.

Venter JC et al. (2001) The Sequence of the human genome. Science 5507: 1304–1351.

Weiner DM, Burstein ES, Nash N, Croston GE, Currier EA, Vanover KE, Harvey SC, Donohue E, Hansen HC, Andersson CA, Spalding TA, Gibson DFC, Krebs-Thomson K, Powell SB, Geyer MA, Hacksell U & Brann M (2001) 5-Hydroxytryptamine2a Receptor Agonists as Antipsychotics. J Pharmacol Exp Ther 299: 268–276.

Willson TM, Jones SA, Moore JT & Kliewer SA (2001) Chemical-genomics: Functional analysis of orphan nuclear receptors in the regulation of bile acid metabolism. Med Res Rev 21: 513–522.

Wood MD, Murkitt KL, Ho M, Watson JM, Brown F, Huner AJ & Middlemiss DN (1999) Functional comparison of muscarinic partial agonists at muscarinic receptor subtypes hM1, hM2, hM3, hM4 and hM5 using microphysiometry. Br J Pharmacol 126: 1620–1624.