

The impact of genomics on medicinal chemistry

Brian W. Metcalf

SmithKline Beecham Pharmaceuticals
709 Swedeland Road
King of Prussia, Pennsylvania 19406-0939

Sequencing of cDNA Libraries

High throughput sequencing as applied to cDNA libraries has rapidly led to a buildup of expressed sequence tags (EST's) in the public and corporate data bases. By definition, this approach gives the partial sequences of expressed genes from the various cells and tissues used for the cDNA library preparation, but not that of the intervening "non-coding" DNA (ref. 1). While these EST's are only partial sequences of 300 nucleotides or so, bioinformatic approaches are able to assemble such EST's that comprise the same gene by the identification of overlapping fragments. Assemblies of EST's (contigs) are then created giving an estimate of the number of expressed genes (as distinct from gene fragments) in the library under consideration. Taken to their logical conclusion such approaches will eventually identify the complete repertoire of expressed human genes and thereby an extremely large number of potential human drug targets. Such estimates range from 80,000 to 100,000. Analogously, application of such high throughput sequencing techniques to microbial genomes will identify all possible targets for intervention in infection control.

The challenge facing the pharmaceutical industry is to prioritize which of the myriad of possible targets it is productive to attack. Such decisions will be made on the basis of homology reasoning to known genes, as well as using information gleaned from new techniques that characterize the localization and temporal expression of the identified genes. For example, genes that are preferentially expressed in diseased tissue, compared to healthy tissue, are likely to be the new targets of the pharmaceutical industry (ref. 2). The industry is therefore in the throes of a revolution as the discovery process now begins with a gene sequence rather than with the traditional disease focus. The magnitude of the opportunity is illustrated by the number of EST's available for analysis in SmithKline Beecham's data bases. These comprise sequences generated by its collaborators at Human Genome Sciences (HGS) and augmented by those in the public data bases.

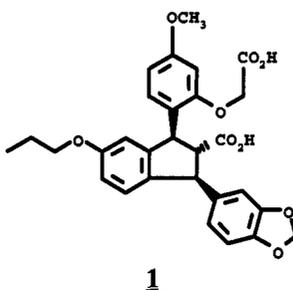
Over 1 million EST's are in the SB/HGS data base and this is increasing at the rate of 20,000 per month

Seven Transmembrane Receptors (7TMR's)

Seven transmembrane receptors contain sequences that span the cell membrane seven times so that the N terminal portion is extracellular and the C terminal portion is intracellular. Such receptors are G protein coupled and the coupling occurs via the intracellular loops. The superfamily of 7TMR's, also known as G-protein coupled receptors, has historically proven to be a rich source of drug targets. Some examples of clinically and commercially successful drugs which are ligands for this superfamily are shown below.

<u>Drug</u>	<u>7TMR</u>
propranolol	β_1, β_2
cimetidine/ranitidine	H ₂
terfenidine	H ₁
clonidine	α_2
prazosin	α_1
losartan	AT ₁
bupirone	5-HT _{1A}

Furthermore, newly-discovered members of the superfamily continue to provide molecular targets for recent drug discovery programs and a number of selective antagonists are under development. In a recent example from SB, a research effort directed at the endothelin A and B receptors (ET_A and ET_B) has led to several drug candidates under development. The SB program required the synthesis of over 325 compounds prepared in an iterative synthesis/testing paradigm to achieve the first development candidate. In total, 1000 compounds were made over several years in the search for and optimization of different structural types. The structure of one of these, 1 (SB 209670), is shown below (ref. 3).



It therefore is evident that the identification of novel 7TMR's and their ligands and of orphan 7TMR's (those members for which the endogenous ligand is unknown) may provide a likely starting point for new drug discovery programs. Using motifs characteristic of members of the 7TMR superfamily to search the SB/HGS data base, followed by cloning the full length genes, scientists at SB have identified over 50 novel full length clones, the majority of which have been stably expressed in mammalian cells for functional studies. (D. Bergsma, et al, unpublished).

Over 50 full length orphan 7TMR's are under study at SB

The Chemist's Response to Genomics - Combinatorial Chemistry

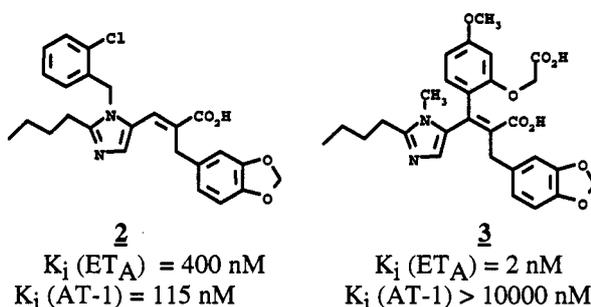
As illustrated by the magnitude of the opportunity represented by the 7TMR superfamily, the multiplicity of targets emanating from rapid high throughput sequencing of cDNA libraries cannot be addressed by "one target at a time" iterative medicinal chemistry paradigms. First, the time window of opportunity is narrow owing to competitive forces and many competitors now have access to similar information. Second, the magnitude of the problem is such that opportunities will be wasted if one is restricted to iterative methods. Third, rapid identification of antagonists for members of, say the 7TMR family, could offer tools to allow dissection of function of many of these receptors. Thus, we enter an era of "reverse pharmacology" where, in some cases, finding the antagonist first would allow the biological significance of the target to be ascribed.

Combinatorial chemistry can take on many forms. "Split and mix" procedures on solid phase lead to each solid phase particle (bead) containing only one molecular entity. Large libraries are possible, but active beads require structural deconvolution. Alternatively, the concurrent syntheses of many compounds using array technology where the individual components are spatially localized obviates deconvolution of structure. Library sizes, however, are limited by current technology. No one approach appears

superior, and SB has invested in both. The "split and mix" approach was used to great effect by Lam, et al (ref. 4) in the generation of a pentapeptide library using nineteen naturally-encoded amino acids to afford 19^5 peptides (2,476,099) on beads where each bead carries only one peptide. The library was screened in a biopanning procedure against a monoclonal antibody, as a model receptor, and high affinity peptides were isolated. The challenge facing medicinal chemists is whether the concepts illustrated in the Lam study can be extrapolated to the combinatorial preparation of "drug-like" organic molecules.

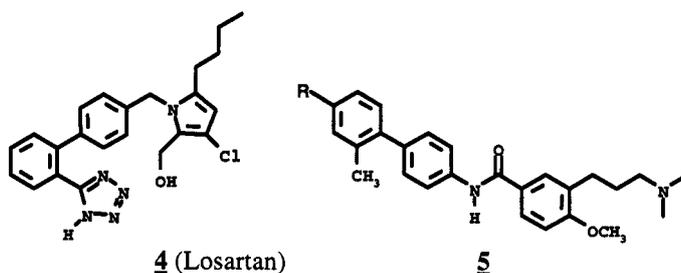
Permissive Structures for 7TMR ligands

It has been known for some time that permissivity exists amongst ligands for the 7TMR superfamily. For example, in the biogenic amine arena, dopamine has affinity for a number of 5-HT receptors, and noradrenaline for some dopamine receptors. This promiscuity suggests that when confronted with a new member of the 7TMR superfamily one should first screen a collection of ligands previously identified as having affinity for other members of the family. Accordingly, cross screening of compound collections from other 7TMR efforts revealed that the imidazole **2**, a relatively potent antagonist of the angiotensin-1 receptor had affinity for the ET_A receptor. Further modification of **2** led to the potent and selective ET_A antagonist **3** (ref. 5).



Permissivity amongst ligands therefore applies not only to receptors that bind the biogenic amines, but also to the 7TMR's that recognize acidic ligands. Extending the concept, faced with fifty members of the family at once as drug targets, one could consider selecting a motif common to some of these ligands and constructing a combinatorial library around that structure.

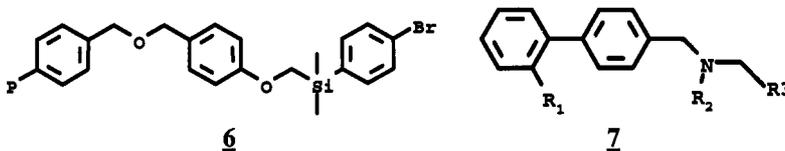
We wished to apply the "split and mix" approach to a template representative of a permissive ligand for 7TMR's. Such a resource could then be applied to the multiplicity of novel 7TMR's available from genomic approaches. The template we chose was the biphenyl structure present in the angiotensin antagonist Losartan, **4** (ref. 6) and in a number of ligands for 5-HT_{1D} receptors as represented in **5** (ref. 7).



Solid Phase Synthesis and Library Construction

A novel arylsilane polymer **6** has been prepared starting from a polystyrene resin so that the aryl bromide substituent allows elaboration. Library components can be released from the polymer by protodesilylation in a "traceless" fashion (ref. 8). A library of the general structure **7** has thus been prepared. The following organic reactions, normally performed in solution, were adapted to the solid

phase: lithium-halogen exchange, palladium-catalyzed coupling with arylboronic acids, Schiff base formation, reduction of the Schiff base, Grignard addition to the Schiff base, acylation of the resulting amine and protodesilylation. Points of diversity are noted and the library can be rapidly expanded. A number of other libraries of organic molecules have been reported in the literature (ref. 9, 10, 11).



Analytical Methods

Advances in the solid state synthesis field have been facilitated by the advent of the mass spectrometric techniques, electrospray-, matrix-assisted laser desorption- and time-of-flight mass spectrometry (ref. 12) and by magic angle spinning nmr (ref. 13). Using these techniques we are able to elucidate the structure of molecules from a single bead.

Limitations of Combinatorial Chemistry and the Future of Medicinal Chemistry

Unlike solution phase organic chemistry where each reaction product in a linear sequence can be purified prior to the next step, solid phase chemistry allows no purification until final release of product from the bead. It does however offer the advantage of allowing reactions to be driven to completion by using an excess of reagents that can be simply washed away from the product which remains bead-bound. Consider, however, a solid phase linear sequence with each step providing a product of 90% purity. The result after six steps would be a compound of maximum purity of 53% attached to the bead. Perhaps it is not coincidental that the longest linear solid phase syntheses thus far reported is that of Goff and Zuckerman (ref. 9) who prepared a library of isoquinolinones in five linear steps.

Can molecules of sufficient molecular complexity to satisfy potency, selectivity and pharmacokinetic criteria for drug candidates be prepared directly by combinatorial methods? This author's opinion is that some percentage, perhaps 10 - 20% of drug candidates will be identified by combinatorial means and will not require further elaboration. In the majority of cases, however, combinatorial chemistry will provide a lead which will require optimization by iterative methods. Paradoxically, the success of combinatorial chemistry will create a greater need for resources for iterative chemistry.

REFERENCES

1. M.D. Adams, J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merrill, A. Wu, B. Olde, R.F. Moreno, A.R. Kerlavage, W.R. McCombie and J. C. Venter, *Science*, **252**, 1651 (1991).
2. M.R. Fannon, *Trends in Biotechnology*, **14/8**, 294 (1996).
3. J.D. Elliott, M.A. Lago, R.D. Cousins, A. Gao, J.D. Leber, K.F. Erhard, P. Nambi, N.A. Elshourbagy, C. Kumar, J.A. Lee, J.W. Bean, C.W. Debrosse, D.S. Eggleston, D.P. Brooks, G. Feuerstein, R.R. Ruffolo, J. Weinstock, J.G. Gleason, C.E. Peishoff and E. Ohlstein, *J. Med. Chem.* **37**, 1153 (1994).
4. K.S. Lam, S.E. Salmon, E.M. Hersh, V.J. Hruby, W.M. Kazmierski and R.J. Knapp, *Nature*, **354**, 82 (1991).
5. J. D. Elliott, D. L. Bryan, P Nambi and E. H. Ohlstein, *Proceedings of the 14th American Peptide Symposium*, P. T. P. Kaumaya and R. S. Hodges Eds., p. 673.
6. A.T. Chiu, D.E. McCall, W.A. Price, P.C. Wong, D.J. Carini, J.V. Duncia, R.R. Wexler, S.E. Yoo, A.L. Johnson and P.B.M.W.M. Timmermans, *J. Pharm. Exp. Ther.* **252**, 711 (1990).
7. J.W. Clitherow, D.I.C. Scopes, M. Skingle, C.C. Jordan, W. Feniuk, I.B. Campbell, M.C. Carter, E.W. Collington, H.E. Connor, G.A. Higgins, D. Beattie, H.A. Kelly, W.L. Mitchell, A.W. Oxford, A.H. Wadsworth and M.B. Tyers, *J. Med. Chem.*, **37** (15), 2253-2257, (1994).

8. B. Chenera, J.A. Finkelstein, and D. F. Veber, *J. Amer. Chem. Soc.* **117**, 11999 (1995).
9. D.A. Goff and R.N. Zuckerman, *J. Org. Chem.* **60**, 5748 (1995).
10. C.G. Boojamra, K.M. Burow and J.A. Eliman, *J. Org. Chem.* **60**, 5742 (1995).
11. J.J. Burnbaum, M.H.J. Ohlmeyer, J.C. Reader, I. Henderson, L.W. Dillard, G. Li, T.L. Randle, N.H. Sigal, D. Chelsky and J.J. Baldwin, *Proc. Nat. Acad. Sci.* **92**, 6027 (1995).
12. C.L. Brummel, J.C. Vickerman, S.A. Carr, M.E. Hemling, G.D. Roberts, W. Johnson, J. Weinstock, D. Gaitanopoulos, S.J. Benkovic and N. Winograd, *Analytical Chemistry*, **68**, 237 (1996).
13. R.S. Garigipati, B. Adams, J.L. Adams and S. K. Sarkar, *J. Org. Chem.*, **61**, 2911 (1996).