Biochemical Pharmacology 78 (2009) 217–223

Contents lists available at ScienceDirect

Biochemical Pharmacology

journal homepage: www.elsevier.com/locate/biochempharm

Commentary

The chemist as astronaut: Searching for biologically useful space in the chemical universe

David J. Triggle *

School of Pharmacy and Pharmaceutical Sciences, State University of New York, 126 Cooke, SUNY at Buffalo, Buffalo, NY 14260-1608, United States

ARTICLE INFO

Article history:
Received 30 January 2009
Accepted 19 February 2009

Keywords:
Chemical space
Pharmacological space
Fragment assembly
Privileged structures
Natural products
Scaffolds

ABSTRACT

Chemical space whether defined by small molecules or large proteins is larger than can be usefully explored. One of the challenges of drug discovery is thus the definition of the overlap between chemical space, biologically useful space and pharmacological space and how this may be employed in the discovery of new small molecule drugs. Despite the decrease in drug discovery productivity in recent years there is no shortage of targets for small molecule intervention, including stroke, pain, neurodegenerative diseases, inflammation and bacterial and viral infections.

Only an extremely small fraction of available chemical space has thus far been explored and it is likely that prior synthetic constraints and bias to existing frameworks and scaffolds have contributed to this. Several approaches are being employed to explore more fruitful paths to discovery. These include recognition that existing therapeutic entities already occupy validated pharmacological space and thus are good leads, the use of molecular fragments that permits a broader exploration of chemical space, and the role of templates that permit fragments to combine to generate active species. Finally, a new focus on natural product-like scaffolds from both synthetic methodologies and the genetic reengineering of biosynthetic pathways is likely to prove valuable.

However the exploration of chemical space will alone not solve the current deficit in drug discovery productivity. It is necessary to recognize that cellular environments are not the dilute homogeneous solutions of many screening systems and that a more integrated systems approach will serve to maximize any success of chemical space exploration.

1. Definitions

Chemical Space:

"The total descriptor space that encompasses all of the small carbon-based molecules that could, in principle, be created".

Biological Space:

"Those parts of chemical space in which biologically active compounds reside". [11]

Pharmacological Space:

"Pharmacological space attempts to chart the limits of chemical space, targets space, disease space, in order to reduce and systematize the search for new drugs in these spaces". (Hopkins AL. Pharmacological space. In: Wermuth CG, editor. The practice of medicinal chemistry. 3rd ed. London/New York: Academic Press; 2008 [Chapter 25].)

2. Where have all the molecules gone?[1]

From the perspective of drug discovery there is a shortage of new molecules being introduced into the therapeutic arena [1,2]. With only17 new molecular entities introduced in 2007 success has decreased and the cost per introduction continues to increase in a non-sustainable manner.[2] In reality, molecules have not, of course, gone anywhere. They, like their component atoms, were with us at the beginning and will be with us at the end – presumably in some eternally and infinitely cold entropy-dead universe. And, in any event, our increasingly confident synthetic chemical methodologies virtually guarantee that if an organic

* Tel.: +1 716 645 7315; fax: +1 716 645 3688.
E-mail address: triggle@buffalo.edu.

[1] With apologies to Pete Seeger composer of "***Where have all the flowers gone***?".

[2] "*Anyone who believes exponential growth can go on forever in a finite world is either a madman or an economist*". Kenneth Boulding, Economist, 1971. Quoted in Ehrlich P, Ehrlich AH. The population explosion; 1990.

0006-2952/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bcp.2009.02.015

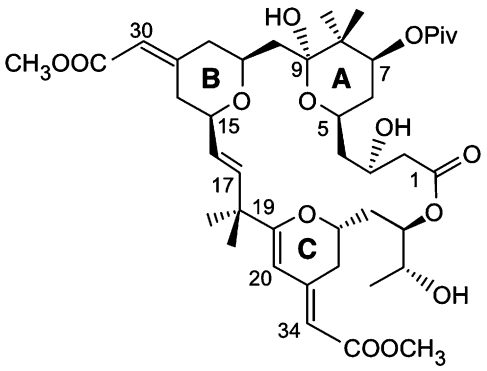

Fig. 1. The structure of bryostatin 16 (for further details of this synthesis see Trost and Dong [4]).

structure can be drawn then it can be synthesized [3]. An outstanding recent example is the recently described 16-step synthesis of bryostatin 16 (Fig. 1) by Trost and Dong [4]. Thus, with apologies to Shakespeare, the fault in therapeutic discovery lies not within our molecules, but within ourselves.

Indeed, the opportunities for new small molecule therapeutic intervention are large indeed. Three examples, will suffice, namely the complete failure of drug intervention in stroke despite over 1000 experimental treatments having been reported [5], pain where significant opportunities for improved therapy still exist decades of effort and the existence of a $1 trillion market [6], and the role that small molecules play as gene regulators including the recently described role of valproic acid in greatly enhancing the induction of pluripotent stem cells by defined factors [7].

The failure of innovation in new drug discovery has, of course, multiple origins including the tackling of more complex diseases, regulatory hurdles, failure of new technologies to be as productive as originally believed, industry business emphasis on life-style diseases, blockbuster entities and "*me-too*" drugs, an excessively reductionist approach to the understanding of disease, and failure to explore the full diversity of the chemical universe for biological ends [8,9].

3. The chemical universe: how big is big?

Our universe consists of some 10^{11} galaxies each containing some 10^{11} stars for a total population of $\sim 10^{22}$ stars [10]. It is unlikely that more than an entirely trivial fraction of this universe will ever be explored by humans, and any detailed exploration will likely be confined to the planets of our own solar system. We are on our own. Whether carbon-based life exists outside of planet Earth is not known, although the presence of water on Mars, the detection of organic molecules under conditions existing during celestial collisions and star formation, and the presence of water and geologic activity on Enceladus (a moon of Jupiter) provides further support for the view that given the right conditions and enough time the arrival of life is inevitable. However, this may not be life as we know it (or even like it!).

Within this universe chemical space is large. It is estimated that the virtual space of small molecules (M. Wt. < 500) that could in principle be created is some 10^{60} [11], far outside the capability of even the most ambitious combinatorial chemist. Similarly, the number of possible proteins of 300 residue length using the 20 amino acid code is $>10^{390}$, a number far exceeding the number of atoms, estimated at 10^{80}, in the universe. Thus, even though the human proteome count may be as large as 10^6, and derived from as few as 20,000 genes, it is clear that Nature is both parsimonious and conservative in her explorations of the chemical universe [12]. Even the total number of protein sequences in terrestrial species, estimated as $\sim 5 \times 10^{10}$, is negligible compare to the theoretically possible total number [13]. The challenge for both medicinal chemist and pharmacologist is to determine the overlap of chemical and biological space, to map biologically useful chemical space and to define those areas where molecules of therapeutic utility may be found recognizing that such utility depends upon both pharmacodynamic and pharmacokinetic properties [14].

Chemical space may, in fact, be enormously larger than that already estimated for our universe. The presence of life as we define it is remarkably dependent on the precise values of a number of physical constants – gravity, electromagnetism, strong force, weak force, number of dimensions and the cosmological constant – absent this "fine tuning" our universe would not exist (or would not be observed) in its current form [15]. According to this view, "*. . .I find that the universe in some sense must have known that we were coming*" [16]. An extreme interpretation would therefore hold that biologically useful chemical space is already predefined and we need only to seek those boundaries to complete our search for this knowledge. However, an alternative explanation held by a number of prominent cosmologists, notably string theorists, argues that the universe appears the way it is because we observe it and that the "fine tuning" and apparent order that we observe is simply coincidental and represents only one permutation in as many as 10^{500} multiple universes [17]. These causally distinct and unobservable (to us) universes presumably define alternative chemical space and this may well be non-carbon and non-water based [18]. Perhaps even now on some fraction of these universes medicinal chemists and pharmacologists are trying to define biologically useful chemical space in silicon-based chemistry in a hydrocarbon milieu! For the purpose of this commentary I will focus, however, only on our universe and the exploration of its chemical space.

4. How much chemical space has been explored?

Only a small fraction has so far been explored and as the preceding considerations indicate the best that we can hope to achieve is but to sample a small fraction of this theoretically available space. Furthermore, what has been explored in the area of small molecule organic materials likely represents a biased analysis of chemical space. The CAS Registry contains over 33 million inorganic and small organic compounds and framework analysis of a subset of some 24 million that represents small molecules containing only C, H, B, Si, N, P, As, O, Se, Te and the halogens revealed a wide range of sizes, but with a highly biased distribution such that a mere 143 framework types describe 50% of the molecules and, furthermore, this distribution follows a power–law relationship [19]. Lipkus and co-workers at the Chemical Abstracts Service (Columbus, Ohio) argue that this relationship derives from a "rich-get-richer" scenario whereby the availability of synthetic access to a particular framework ensures its increased use.[3] Thus organic chemistry space has tended to focus on a quite limited number of structural motifs. It is scarcely surprising, therefore, that a similarly limited distribution of structural motifs is to be found in the chemical space occupied by known drugs. In two pioneering analyses conducted in 1996 and 1999 Bemis and Murcko analyzed the molecular frameworks and side-chains of some 5000 compounds then existing in the Comprehensive Medicinal Chemistry Database [20,21]. Consistent with the

[3] We may also describe this as Robert Merton did as "The Matthew Effect" derived from Matthew 25:29 (King James version): "For unto everyone that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath" (Merton RK. The Matthew effect in science. Science 1968;159:56–63.). It is at least interesting that the Darwinian rules of the so-called "free market" appear to apply to the chemist's choice of what to synthesize.

analysis by Lipkus they found 32 frameworks were sufficient to describe the shapes of 50% of the drugs in the database and that of the 15,000 occurrences of side-chains in these drugs approximately 75% were from 20% of the side-chains. These studies suggest that the search for biologically useful chemical space may have been limited by synthetic chemical constraints, including cost, synthetic ease and bias to known structural motifs.

5. Just how big is biologically useful chemical space?

The boundaries of biological space are determined by the genetic code that maps 64 3-letter codons to 20 amino acids and one stop signal (two other stop codons are not employed for translational termination and can thus be used for *in vitro* modification to code for "unnatural" amino acids (see below)). In principle, a multiplicity of genetic codes could exist, but the existing code is highly optimized to achieve a minimal impact of translational errors whereby similar codons code for chemically related amino acids and because the simplest amino acids have more assigned codons reflecting that these amino acids are more often employed in protein assembly [22]. Although 1:1 correspondence between the triplet codon has been assumed to be essential for fidelity of translation it has recently been described that in a ciliate of the genus Euplotes a single codon generates both cysteine and selenocysteine [23]. If this phenomenon occurs elsewhere then the genetic code encompasses more chemical space than previously believed.

The genetic code is, however, also optimized for the ability to carry information additional to that necessary for protein coding. Such parallel information includes codes for the binding of regulatory and structural proteins, for splicing signals and for RNA secondary structure [24]. Nor is such parallel information likely confined only to the protein coding sequences of DNA. It is clear, for example, that transposable elements, which constitute a significant fraction of eukaryotic DNA, are important in genome evolution and also contribute to genetic instability and to genetically driven diseases [25].

Accordingly, a relatively limited area of chemical space can, by coding for multiple levels of information, be amplified into a far larger area of biologically useful space. That this code is apparently universal on earth does not preclude the existence of alternative code constructs elsewhere that may employ similar building blocks but became "frozen" at a different stage of evolution [26] and would thus define a different region(s) of biologically useful chemical space. It is at least theoretically possible that alternative life forms, or at least markers in the geologic record of their prior existence may yet be found on earth itself [27].

How much of this biologically useful space is amenable to actual drug development remains, however, in question. Estimates of the size of the druggable genome indicates totals of between 2000 and 3000 or approximately 10% of the actual gene count [28]. In an effort to map globally pharmacological space Paolini et al. [29] identified 836 human genes for which small molecule compounds with binding affinity of $<10\ \mu M$ have been identified. They further identified 727 human targets with at least one similarly binding compound and identified 158 human proteins as the primary targets for clinically approved small molecule drugs. This, of course, is significantly less than the "embarrassment of targets" once predicted at the beginning of the genome project. Additional limitations are imposed from considerations of protein structure and classification. Consistent with the relatively small number of extant protein sequences there also appears to be a limited number of shapes into which proteins fold [13]. There exists perhaps as many as 10,000 total folds, but the distribution amongst sequences is far from uniform with the great majority of protein families belonging to approximately 1000 folds. This distribution has the properties of a power law common to a wide variety of biological and social situations and also by the distribution of the molecular frameworks of small molecule organic compounds [19]. These observations, together with the demonstration that significant ligand binding site similarity exists within superfolds of functionally unrelated proteins [30] are also consistent with Nature's parsimony in the exploitation of available chemical space for biological ends. This view has been emphasized most recently by Marth who has argued that there exists a mere 68 molecules – 8 nucleosides, 32 saccharides, 8 lipids and 20 amino acids – that generate the four macromolecular components of all cells – nucleic acids, proteins, glycans and lipids [31].

Biologically encoded chemical space can, however, be expanded by reengineering DNA to code for unnatural amino acids that are subsequently translated into proteins that may have novel biological properties [32]. In excess of 60 additional amino acids have been incorporated into proteins employing, amongst other methods, the amber stop codon TAG and an orthogonal tRNA-synthetase pair that can incorporate the new amino acid. Thus, in studies with ion channels Henry Lester and co-workers have studied the modification of specific residues of interest and also the effect of incorporating fluorescent residues, tethered ligands and "caged" residues in nAChR and the 5-HT_3AR ligand-gated channels [33]. Similar studies with the K_V1.4 channel have demonstrated that the size of the introduced residues in the inactivation peptide is critical in N-type inactivation [34]. Of related interest is the ability to resurrect ancient genes by inferring the sequence of an ancient protein by phylogenetic tracing and then through DNA synthesis express that protein [35]. An example of this technique was the characterization of the ancestral EF-Tu elongation factor that proved to have an optimum GDP binding temperature of 65 °C presumably reflecting that the ancestor of bacteria was adapted to this temperature [36].

Expansion of biologically encoded chemical space is also achievable through the de novo synthesis of genomes. The recent synthesis of the genome (582,970 base-pairs) of *Mycoplasma genitalium* [37] indicates the opportunities for the laboratory synthesis of entirely novel genomes that may encode novel chemical structures of biological interest and that would complement current work on re-engineered bacteria that synthesize molecules such as the anti-malarial artemesinin [38].

Despite boundary limitations imposed by the existing genetic code there is an obvious wealth of biologically useful chemical space. Alternative splicing and multiple regulatory mechanisms including a diversity of recently discovered RNA signaling mechanisms and epigenetic modification, ensure that a single DNA sequence is expressable in multiple forms in both time- and quantity-dependent forms. This is familiarly expressed in antibody diversity and combinatorial protein chemistry, but is seen elsewhere also. Thus many microorganisms use clonal phenotypic variation of surface antigens as an immune-evasion strategy. This ensures their infective success [39]. *Giardia lamblia*, an intestinal parasite with disturbingly unpleasant effects expresses only one variant-specific surface protein (VSP) at a time, selected from a repertoire of approximately 200 genes. Spontaneous switching at approximately every tenth generation of the surface protein occurs through an iRNA-dependent silencing of all but one of the RNA copies from the VSP genes [40].

Such diversity of expression is also revealed in the remarkable collection of neuropeptides secreted by mollusks of the *Conus* genus [41,42]. Here, some 700 species of these predatory snails each secretes between 100 and 200 different peptides targeted against a constellation of known ion channels and neurotransmitter receptors. There likely exists in excess of some 100,000 potent neuropeptides secreted by this one genus. This diversity is made possible by a set of rapidly diversifying genes (nine members

are so far known) that encode these peptides and whose diversification reflects the distinct ecological niches that are inhabited by the different *Conus* species. The majority of these peptides contains several disulfide bridges and thus represent relatively rigid structures where the different sequences of amino acids define the individual biological specificity. They may be thought of as "privileged structures".

The term "privileged structure" was indeed first applied to the benzodiazepine nucleus by Evans et al. in their search for CCK-A antagonists derived from the natural product asperlicin [43]. Subsequently, this nucleus has served as the scaffold for a variety of discrete receptor systems (Fig. 2). It has, however, long been recognized that certain substructures are common molecular components of active drugs [44]. Notably, Ariens noted the frequent presence of 1,1-diphenylmethyl and arylpiperidine units in many drug classes [45] and this concept was expanded by Wiley and Rich [46] who described some 16 templates including the ubiquitous 4-aryl-1,4-dihydropyridine nucleus ([47]; Fig. 3). The existence of such privileged structures reflects that Nature is extremely proficient in "*variation on a theme*" to generate molecular diversity and biological differentiation from a limited toolbox of molecular units.

6. Exploring chemical space for new therapeutic entities

> "We shall not cease from exploration, and the end of our exploring will be to arrive where we started and know the place for the first time". (Eliot TS. Little Giddings. The Four Quartets; 1942.)

In the search for new therapeutic entities there are but three principal paths of exploration – random walks, building on existing active structures, and using macromolecules as templates for molecular assembly.

6.1. Random walks

> "Two roads diverged in a yellow wood, and sorry I could not travel both,. . .then took the other just as fair",

wrote Robert Frost in 1915, thus, and unknowingly, anticipating the challenges in exploring molecular diversity by a random walk. Given the dimensions of small molecule space the probability of successful search by a random walk is remote, hence the failure of the first implementations of combinatorial chemistry that failed to include considerations of drug "quality" [48]. Increasingly, attention has been to focussed libraries based on productive molecular scaffolds (existing drugs and privileged structures), and mimicking the combinatorial chemistry of Nature through diversity-oriented methods [49].

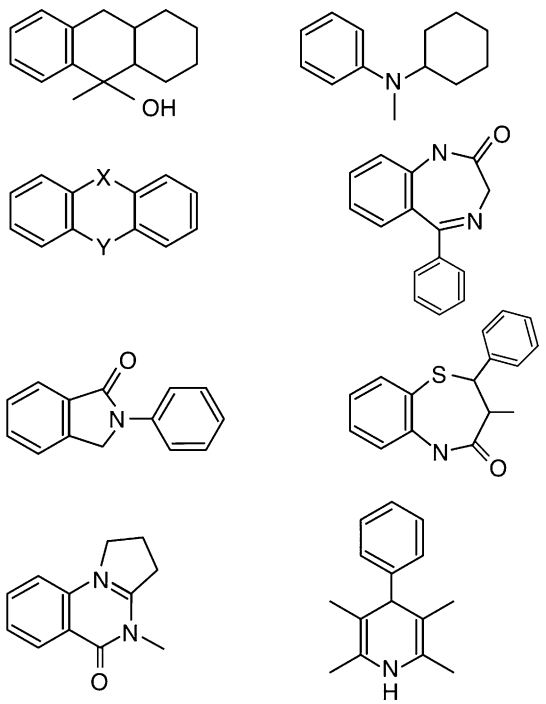

Fig. 3. Some scaffolds that are a common component of pharmacologically active molecules (for further examples of scaffolds see Wiley and Rich [46]).

6.2. Building on existing structures

> "The most fruitful basis for the discovery of a new drug is to start with an old drug" (Sir James Black, Nobel Laureate; 1994).

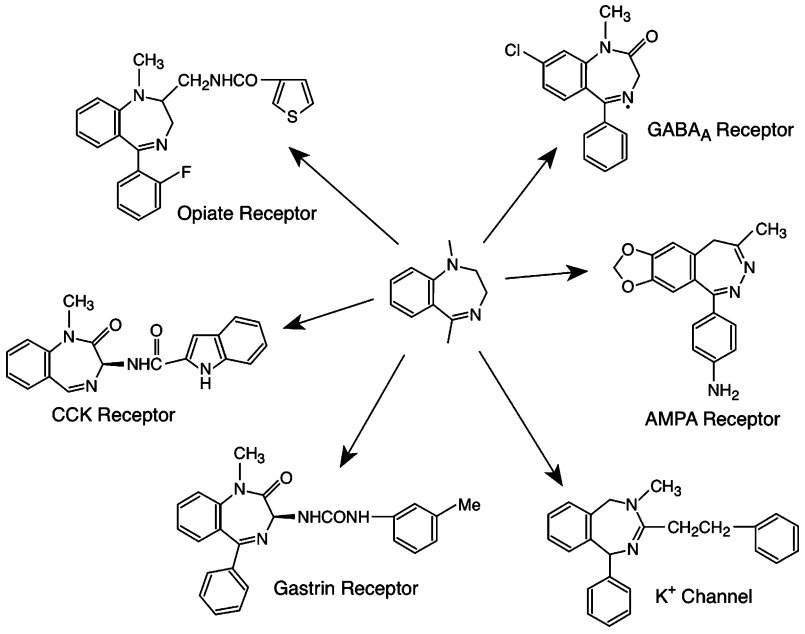

Fig. 2. The benzodiazepine scaffold as a privileged structure.

It may well be argued that all of our successful approaches to drug discovery are variations on this theme. Is this then simply an example of exploring under the lamppost because that is where the light is or rather that this light – Nature's light – is already illuminating the biologically useful chemical space? Clearly, there exist many examples of drug classes that represent variations on an original theme: familiar examples would include the β-blocker family, tricyclic antidepressants, the statins, the 1,4-dihydropyridine calcium channel blockers, etc. [50]. Sometimes, and pejoratively, referred to as "me too" drugs and indeed often produced in response to competitive and economic forces this approach frequently produces molecules with differing pharmacodynamic and pharmacokinetic properties. Additionally, since these molecules are already occupying validated pharmacological space they can be employed as leads for new therapeutic indications. The "selective optimization of side activities" (SOSA) approach screens a library of existing drugs (with known human safety, toxicity and bioavailability profiles) and screens this against new targets. Subsequent structural optimization then converts a side-effect of an existing drug into a principal activity of a potential new drug [51]. Hong and Sullivan [52] have advocated building a comprehensive library of the ~10,000 known clinically active entities together with their known major metabolites and establishing this as a public resource for screening, the results of which would be available in a public database. Partial validation of this approach was derived by examining the phenotypic side-effect similarities of some 746 marketed drugs to derive a network of drug–drug relations, a significant fraction of which were generated by distinct molecular entities from different therapeutic indications [53,54].

6.3. Natural product-like chemical space

> "Molecules forged in the crucible of evolution". (Triggle DJ. The shape of medicines to come. Med Chem Res 2004;13:315.)

Historically, natural products and their derivatives have been a major source of therapeutic agents. Almost 50% of the NCE's introduced in the 1980s and 1990s were related directly or indirectly to natural product structures [55]. There is now a resurgence of interest derived from genetic engineering of biosynthetic pathways that generate natural product scaffolds [56], and from improved synthetic methods that can more readily generate the complex ring structures of natural products [57]. Traditionally, such complex natural product structures have been approached on a "one molecule at a time" basis. More recent synthetic methods strive for diversity-oriented synthesis to permit from simple starting blocks the generation of a large collection of diverse and complex entities. A recent example of such synthetic methodology is provided in a scheme, based on a "build-couple-methodology", which generates 80 different molecular skeletons in modular fashion in a small number of steps [58]. To facilitate further exploration of natural product-based chemical space Waldmann and co-workers have provided as structural classification of natural products (SCONPs) based on the underlying scaffolds present in natural products [59]. This hierarchical scaffold organization can provide guidance for the selection of underlying molecular frameworks for the development of natural product-based chemical libraries.

6.4. Fragment assembly: one small step at a time

> "Great things are done by a series of small things brought together" (Vincent van Gogh, 1853–1890)

Fragment-based drug discovery rests on the basic concept that molecular complementarity is more easily and efficiently explored with small molecular fragments of size up to 12 heavy (non-hydrogen) atoms since the number of such fragments is $\sim 10^7$, far smaller than the $>10^{60}$ of larger drug-like molecules. Thus, a larger proportion of "fragment-like" space can be probed compared to "drug-like" space in conventional combinatorial chemistry plus high throughput screening [60]. However, since the fragments typically have low affinity, frequently in the millimolar range, the assays must employ sensitive biophysical techniques such as NMR and X-ray crystallography. Examples of recent applications of this technique are detailed by Congreve et al. [61].

Not infrequently drug molecules contain within the structure a substructure that is also active as another drug. Examination of a total of some 1386 marketed drugs revealed that approximately 15% of them are contained within other drugs and that some 30% of the total contain other drugs as substructure fragments [62]. Thus, a subset of drugs was identified that can be used as building blocks in fragment-based drug discovery.

Template-guided synthesis may be viewed as an extension of fragment assembly whereby a biological macromolecule is employed as template on which molecular fragments assemble and then through covalent bond formation – "click chemistry" – link together to form the active molecule [63]. An excellent and early example of the success of this approach was the synthesis of an AChE inhibitor with femtomolar affinity from smaller tacrine and phenanthridinium motifs ([64]; Fig. 4).

7. But what have we ignored?

Our search for biological activity in the chemical universe has focused increasingly in the past two decades on *in vitro* high-throughput technologies in a reductionist approach to drug discovery. In our adherence to "bigger is better" and "faster is best" policies for discovery technology we have ignored critical issues and this has undoubtedly contributed to the current lag in the development of new therapeutic entities.

As an early systems biologist Ezekiel "*connected dem dry bones*", thus recognizing that in biological systems everything is connected to everything else. Such connectivity should not be surprising since it was also recognized at the human whole body level by Travers and Milgram in the "six degrees of separation" experiment [65]. The reductionist approach of many contemporary drug discovery programs has been discussed by Peterson [66] and Williams [67]. Williams has argued, persuasively, that: "*unfettered approach to*

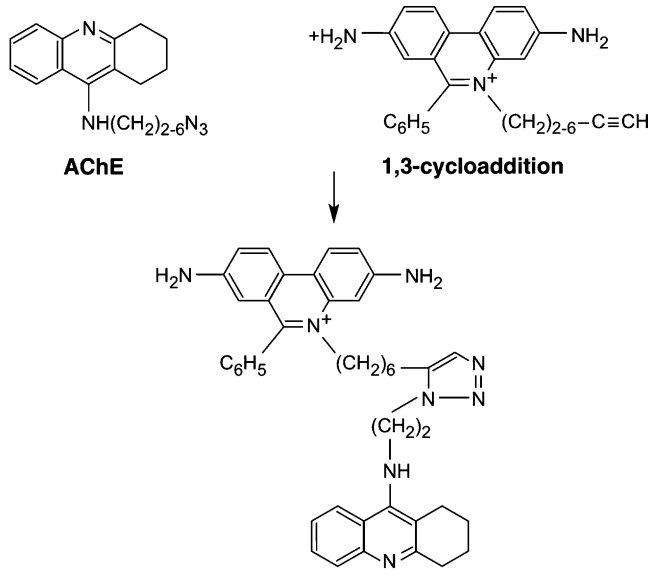

Fig. 4. Click chemistry formation of a potent AChE inhibitor (for details of this approach see Lewis et al. [64]).

such technologies, where exclusion rather than integration has been the hallmark, has markedly reduced the intellectual component of the biomedical research endeavor, with perceived technological "quick fixes" displacing the integrative, hierarchical approach of pharmacology that, with medicinal chemistry, represents the core of the drug discovery process".

In related manner we have also ignored that in cellular systems chemical processes take place in concentrated solutions in small and crowded environments [68]. A dominant characteristic of cell interiors is they contain a high concentration of macromolecules. Even extra-cellular environments can represent areas where molecule concentrations can be remarkably high as with neurotransmitters in synaptic clefts. Thus, acetylcholine concentrations in skeletal muscle synaptic concentrations are estimated to be as high as 500 μM [69]. Typically, however, macromolecule interactions are typically studied in dilute solutions that fail to accommodate the kinetic and thermodynamic effects imposed on reactions by molecular crowding. Additionally, many drug interactions take place on membrane surfaces with high concentrations of receptor sites rather than in dilute homogenous solutions. Under these conditions the kinetics of drug interaction may be significantly altered through such processes as non-specific attachment to cell surface followed by two-dimensional diffusion to the receptor site and by repeated association and disassociation from receptor sites that may be present in localized high concentrations. This phenomenon of "reduction of dimensionality" was initially explored by Adam and Delbruck in a treatment of insect olfaction [70] and a more general analysis has been provided by Axelrod and Wang [71]. Under such conditions of cellular crowding and facilitated interaction at surface sites many ligand–macromolecule interactions are not of high affinity, but rather achieve their selectivity where necessary through cellular constraints, including localization by release and removal constraints and by the presence of efficient inactivating mechanisms such as enzymes and transporters for neurotransmitters. Thus, high affinity of a ligand is not an obligatory prerequisite for selectivity or efficiency of interaction provided that anatomical or other constraints confine the ligand to its site of action. Such selectivity could in principle be achieved for a parenterally administered drug by an appropriate delivery mechanism that confines its locus of action.

Finally, the search for therapeutic entities has been driven in large part on the premise that high selectivity and specificity of action is necessary to ensure interaction at a single target only. However, there is increasing evidence that target promiscuity is key to the clinical efficacy of a number of important drugs, including mood disorder, anticonvulsant and anti-arrhythmic drugs [72,73]. The design of promiscuous as opposed to monogamous entities represents however a challenge to medicinal chemistry where it will be necessary to obtain a balanced pharmacodynamic and pharmacokinetic activity at each designated target site. Currently available drugs with multiple actions have been arrived serendipitously not rationally [74].

References

[1] Business Insights. Drug approval trends at the FDA and EMEA. www.global-businessinsights.com. Accessed December 10, 2008.
[2] Owens J. Big pharma slims down to increase productivity. Nat Rev Drug Discov 2007;6:173–4.
[3] Hudlicky Reed JW. The way of synthesis. New York: Wiley-VCH; 2007.
[4] Trost BM, Dong G. Total synthesis of bryostatin 16 using atom-economical and chemoselective approaches. Nature 2008;456:485–8.
[5] O'Collins VE, MacLeod MR, Donnan GA, Horky LL, van der Worp BH, Howells DW. 1,026 experimental treatments in acute stroke. Ann Neurol 2006;59:467–77.
[6] Max MB, Stewart WF. The molecular epidemiology of pain: a new discipline for drug discovery. Nat Rev Drug Discov 2008;7:647–58.
[7] Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, et al. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 2008;26:795–7.
[8] Cuatrecasas P. Drug discovery in jeopardy. J Clin Invest 2006;116:2837–42.
[9] Williams M. Systems and integrative biology as alternative guises for pharmacology: prime time for an iPharm concept. Biochem Pharmacol 2005;70:1707–16.
[10] Baldi P. Chemoinformatics, drug design, and systems biology. Genome Inf 2005;16:281–5.
[11] Dobson CM. Chemical space and biology. Nature 2004;432:824–8.
[12] Simon GM, Cravatt BF. Challenges for the "chemical-systems" biologist. Nat Chem Biol 2008;4:639–42.
[13] Koonin EV, Wolf YI, Karev GP. The structure of the protein universe and genome evolution. Nature 2002;420:218–23.
[14] Lipinski C, Hopkins A. Navigating chemical space for biology and medicine. Nature 2004;432:855–61.
[15] Rees M. Just six numbers. The deep forces that shape the universe. New York, NY: Basic Books; 2002.
[16] Dyson F. Disturbing the universe. New York, NY: Harper Row; 1979.
[17] Susskind L. The cosmic landscape. String theory and the illusion of intelligent design. Boston, New York: Little, Brown & Company; 2006.
[18] The Limits of Organic Life in Planetary Systems. Washington, DC: National Academies Press; 2007. www.nap.edu.
[19] Lipkus AH, Yuan Q, Lucas KA, Funk SA, Bartelt III WF, Schenck RJ, et al. Structural diversity of organic chemistry: a scaffold analysis of the CAS Registry. J Org Chem 2008;73:4443–51.
[20] Bemis GW, Murcko MA. The properties of known drugs. 1. Molecular frameworks. J Med Chem 1996;39:2887–93.
[21] Bemis GW, Murcko MA. The properties of known drugs. 2. Side chains. J Med Chem 1999;42:5095–9.
[22] Freedland SJ, Hurst LD. The genetic code is one in a million. J Mol Evol 1998;47:238–48.
[23] Turanov AA, Lobanov AV, Fomenko DE, Morrison HG, Sogin ML, Klobutcher LA, et al. Genetic code supports targeted insertion of two amino acids by one codon. Science 2009;323:259–61.
[24] Itzkovitz S, Alon U. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. Genome Res 2007;17:405–12.
[25] Kazazian Jr HH. Mobile elements: drivers of genome evolution. Science 2004;303:1626–32.
[26] Crick FHC. The origin of the genetic code. J Mol Biol 1968;38:367–79.
[27] Davies P. Are aliens among us? Sci Am 2007;62–9.
[28] Russ Ap, Lampel S. The druggable genome: an update. Drug Discov Today 2005;10:1607–10.
[29] Paolini GV, Shapland RHB, van Hoorn WP, Mason JS, Hopkins AL. Nat Biotechnol 2006;24:805–15.
[30] Russell RB, Sasieni PD, Sternberg MJE. Supersites within superfolds. Binding site similarity in the absence of homology. J Mol Biol 1998;282:903–18.
[31] Marth JD. A unified vision of the building blocks of life. Nat Cell Biol 2008;10:1015–6.
[32] Xie J, Schultz PG. Adding amino acids to the genetic repertoire. Curr Opin Chem Biol 2005;9:548–54.
[33] Beene DL, Dougherty DA, Lester HA. Unnatural amino acid mutagenesis in mapping ion channel function. Curr Opin Neurobiol 2003;13:264–70.
[34] Wang W, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee K-F, et al. Genetically encoding unnatural amino acids foe cellular and neuronal studies. Nat Neuorsci 2007;10:1063–72.
[35] Thornton JW. Resurrecting ancient genes: experimental analysis of extinct molecules. Nat Rev Genet 2004;5:366–75.
[36] Gaucher EA, Thomson JM, Burgan MF, Benner SA. Inferring the palaoeenvironment of ancient bacteria on the basis of resurrected proteins. Nature 2003;425:285–8.
[37] Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaeveri J, et al. Complete chemical synthesis, assembly and cloning of a *Mycoplasma genitalium* genome. Science 2008;319:1215–9.
[38] Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndunga JM, et al. Production of the antimalarial drug precursor artemesinic acid in engineered yeast. Nature 2006;440:940–3.
[39] Deitsch KW, Moxon ER, Wellems TE. Shared themes of antigenic variation and virulence in bacterial, protozoal and fungal infections. Microbiol Mol Biol Rev 1997;61:281–93.
[40] Prucca CG, Slavin I, Quiroga R, Elias EV, Rivero FD, Saura A, et al. Antigenic variation in *Giardia lambia* is regulated by RNA interference. Nature 2008;456:750–4.
[41] Terlau H, Olivera BM. *Conus* venoms: a rich source of novel ion channel-targeted peptides. Physiol Rev 2004;84:41–68.
[42] Olivera BM. *Conus* peptides: biodiversity-based discovery and exogenomics. J Biol Chem 2006;281:31173–7.
[43] Evans BE, Rittle KE, Bock MG, DiPardo RM, Whitter WL, Veber DF, et al. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. J Med Chem 1988;31:2235–46.
[44] Patchett AA, Nargund RP. Privileged structures—an update. Ann Rep Med Chem 2000;36:289–98.
[45] Ariens EJ. Stereochemistry in the analysis of drug action. Med Res Rev 1987;7:367–87.
[46] Wiley RA, Rich DH. Peptidomimetics derived from natural products. Med Res Rev 1993;13:327–84.

[47] Triggle DJ. The 1,4-dihydropyridine nucleus: a pharmacophoric template. Part 1. Actions at ion channels. Mini Rev Med Chem 2003;3:215–23.
[48] Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 2007;6:881–90.
[49] Verdine, GL. The combinatorial chemistry of nature. 1996;384(Suppl.):11–3.
[50] Wermuth CG. Similarity in drugs: reflections on analogue design. Drug Discov Today 2006;11:348–54.
[51] Wermuth CG. Selective optimization of side activities: another way for drug discovery. J Med Chem 2004;47:1303–14.
[52] Hong CR, Sullivan Jr DJ. New uses for old drugs. Nature 2007;448:645–6.
[53] Crukhorn S. New targets for old drugs. Nat Rev Drug Discov 2008;7:251.
[54] Campillos M, Kuhn M, Gavin A-C, Jensen LJ, Bork P. Drug target identification using side-effect similarity. Science 2008;321:263–6.
[55] Newman DJ, Cragg GM. Natural products as sources of new drugs over the past 25 years. J Nat Prod 2007;70:461–77.
[56] Clardy J, Walsh C. Lessons from natural molecules. Nature 2004;432:829–37.
[57] Schreiber SL. Target-oriented and diversity-oriented organic synthesis in drug discovery. Science 2000;287:1964–9.
[58] Morton D, Leach S, Cordier C, Warriner S, Nelson A. Synthesis of natural-product-like molecules with over eighty distinct scaffolds. Angew Chem Int Ed 2009;48:104–9.
[59] Koch MA, Schuffenhauer, Scheck M, Wetzel S, Casaulta M, Odermatt A, et al. Charting biologically relevant chemical space: a structural classification of natural products (SCONP). Proc Nat Acad Sci USA 2005;102:17272–7.
[60] Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat Rev Drug Discov 2007.
[61] Congreve M, Chessari G, Tisi D, Woodhead AJU. Recent developments in fragment-based drug discovery. J Med Chem 2008;51:3661–80.
[62] Siegel MG, Vieth M. Drugs in other drugs: a new look at drugs as fragments. Drug Discov Today 2006.
[63] Kolb HC, Sharpless KB. The growing impact of click chemistry on drug discovery. 2003;8:1128–37.
[64] Lewis WG, Green LG, Grynszpan F, Radic Z, Carlier PR, Taylor P, et al. Click chemistry *in situ*: acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. Angew Chem Int Ed 2002;41:1053–7.
[65] Travers J, Milgram S. An experimental study of the small world problem. Sociometry 1969;32:425–43.
[66] Peterson RT. Chemical biology and the limits of reductionism. Nat Chem Biol 2008;4:635–8.
[67] Williams M. Systems and integrative biology as alternative guises for pharmacology: prime time for an iPharm concept? Biochem Pharmacol 2005;70: 1707–16.
[68] Ellis RJ. Macromolecular crowding: obvious but unappreciated. Trends Biochem Sci 2001;26:597–604.
[69] Land B, Salpeter EE, Salpter M. Kinetic parameters for acetylcholine interaction in intact neuromuscular junctions. Proc Nat Acad Sci USA 1981;78:7200–4.
[70] Adam G, Delbrick M. Reduction of dimensionality in biological diffusion processes. In: Rich A, Davidson N, editors. Structural chemistry and molecular biology. San Francisco: W.H. Freeman; 1968. p. 198–215.
[71] Axelrod D, Wang MD. Reduction-of-dimensionality kinetics at reaction-limited surface receptors. Biophys J 1994;66:588–600.
[72] Roth BL, Sheffler DJ, Kroeze WK. Magic shotguns versus magic bullets; selectively non-selective drugs for mood disorders and schizophrenia. Nat Rev Drug Discov 2004;3:353–9.
[73] Bianchi MT, Pthmanathan J, Cash SJ. From ion channels to complex networks: magic bullet versus magic shotgun approach to anticonvulsant pharmacotherapy. Med Hypotheses 2008.
[74] Hopkins AL, Mason JS, Overington JP. Can we rationally design promiscuous drugs? Curr Opin Struct Biol 2006;16:127–36.