Towards a pharmacological genetics

In the past, pharmacological approaches have been instrumental in identifying proteins involved in biological processes. Combinatorial chemistry has the potential to make this type of approach even more powerful.

Chemistry & Biology September 1994, **1**:3–6

Introduction

To understand a complex biological system we must dissect it. Three general approaches, used alone or in combination, have been effective historically. These are biochemical fractionation and reconstitution, genetic analysis, and pharmacology. A modern graduate curriculum in molecular biology tends to emphasize the first two approaches. Biochemical fractionation is generally useful for purifying components of a system and elucidating their properties, whereas genetics is uniquely powerful at opening up for molecular analysis systems that are difficult to dissect with biochemical approaches. In contrast, pharmacological approaches, such as the use of small molecule inhibitors or ligands to dissect a complex pathway are usually given less emphasis, despite their historical importance. For example, in enthusing over a clever genetic screen for mutants in mitosis it is easy to forget the importance of the drug colchicine in discovering tubulin [1], arguably the most important protein in that process. In fact, in many complex systems, particularly those relevant for human health, the first molecular insights came from studying a drug or ligand and using the small molecule to access an important protein.

The modern emphasis in pharmacological research is on rational drug design rather than using drugs as tools. The objective is thus first to obtain the target protein using genetics and biochemistry, then to design a small molecule that binds to it with therapeutic effects. Here I will argue that one of the consequences of recent developments in drug design technology, specifically in the area of molecular diversity, has been to open up powerful new avenues to an approach that is essentially old-time pharmacology, that is, using drugs to identify unknown proteins from complex systems. Furthermore, I will argue that with a little help from the chemists this technology could be as useful in small laboratories as it is expected to be in large companies.

A new pharmacology

Molecular diversity, which is also referred to as combinatorial synthesis, is one of the most exciting and rapidly growing areas in drug design [2–4]. The goal of such work is to synthesize complex and diverse mixtures of compounds from which lead drugs are identified by selection or screening. Many different chemistries have been proposed and developed for use in combinatorial syntheses. They have in common the use of relatively simple, but diverse, building blocks (monomers) that are coupled together to give linear or cyclic polymers (Fig. 1a).

Combinatorial synthesis is a radical departure from the traditional goal of chemists to synthesize pure, characterized compounds, and its inspiration has come partly from the overlap of chemistry and genetics. The first combinatorial chemicals were biopolymers, polynucleotides and polypeptides, made either chemically or in living cells. These biopolymers tend to have relatively low diversity in their building blocks, four nucleotide bases for polynucleotides, for example, or twenty amino acids for polypeptides, but they achieve huge diversity by incorporating large numbers of polymeric units. For systems such as phage display [5] an approach that is essentially genetic selection can be invoked, and for most biopolymers well-developed, matrix-bound synthetic chemical approaches are available. Recently, several combinatorial chemistries for more drug-like molecules have been developed. Here the goal is to generate diverse libraries of low molecular weight molecules, and this entails using more diverse monomer sets in shorter polymers that are resistant to enzymatic degradation. It is too early to predict which chemistries will work best for making drugs, but promising candidates include *N*-substituted polyglycines (peptoids; Fig. 1b) and tetra-substituted benzodiazepines (Fig. 1c).

Combinatorial chemistry has opened up a whole new area of pharmacology that depends on molecular diversity. It is now possible to produce in a single microtiter dish more different small molecules than exist in the Aldrich catalogue; such mixtures are termed compound libraries. The catch is that the few interesting compounds are mixed together with many irrelevant ones, and the challenge is to develop methods for selecting, screening or indexing the libraries to allow for the identification of single molecules from the mixtures. These technologies are being developed in parallel with the chemistries [6,7]. With synthesis and selection schemes in hand, we are now in a position to add vast numbers of new compounds to a biological system in order to find a novel inhibitor (or activator). The ability to generate a diverse library of compounds and then select for an inhibitor of a particular biochemical event is a pharmacological analogue of the typical genetic experiment, in which mutants are randomly generated and those that inhibit a particular step in a complex process are selected.

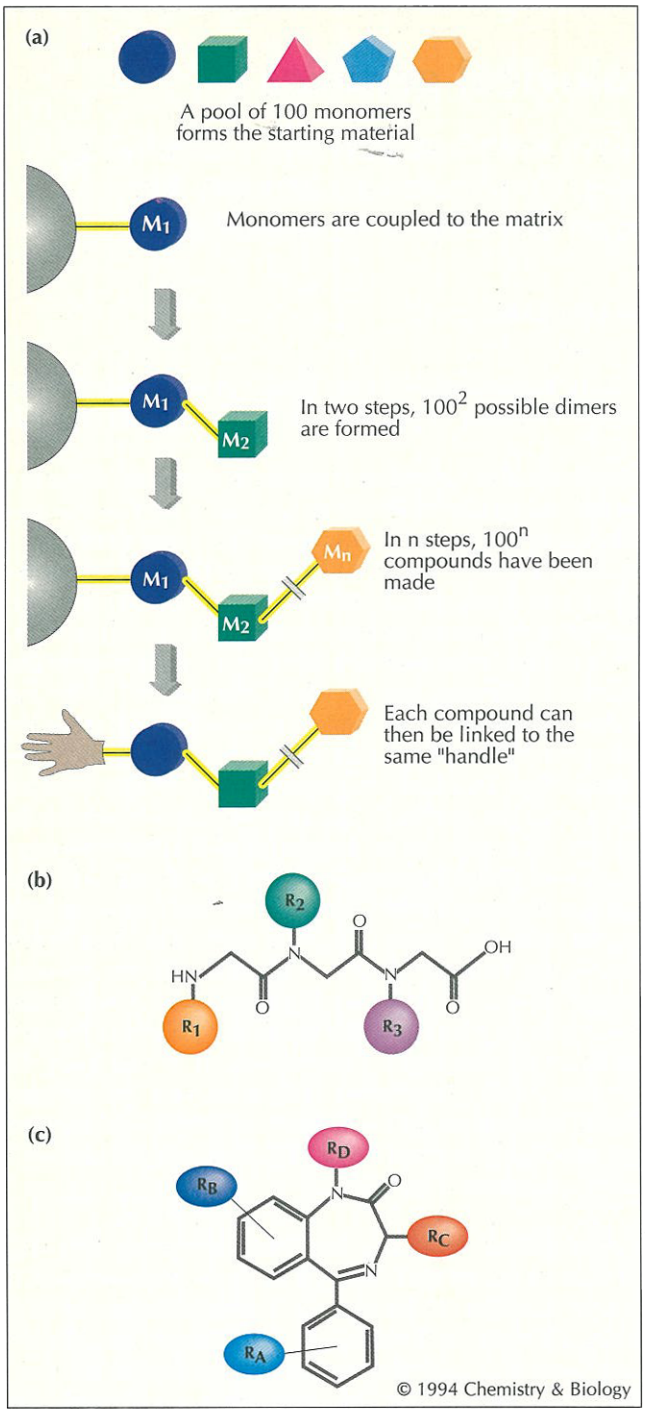

Fig. 1. Design of combinatorial molecules. **(a)** Making a generic combinatorial library. The monomer units, M_1–M_n provide the variable part of the molecule responsible for protein targeting. The number of different drugs in the library is given by the *n*th power of the number of different monomer units used in synthesis. Both linear and cyclic backbones are used in many specific chemistries. Once the combinatorial library has been formed, each of the compounds can be transferred onto a 'molecular handle'. The handle is constant, and is designed to facilitate synthesis and target protein identification, for example, by incorporating a cross-linking group or a tag suitable for affinity chromatography. **(b)** Oligo *N*-substituted glycines, or peptoids [11]. This is a backbone that has many of the synthetic and biomimetic advantages of polypeptides but which is protease resistant. **(c)** The combinatorial benzodiazepine backbone [12]. This backbone is present in several important conventional drugs.

Systems and selections

What kinds of biological systems will be most suited to dissection by combinatorial pharmacology? They will typically be complex, and refractory to dissection by traditional biochemistry. This includes any regulatory or morphological process exhibited by a whole living cell, a permeabilized cell or a complex extract derived from a cell. Three conditions will be required for successful screening: the system must allow ready access to the novel compounds; the compounds must be stable over the assay period; and the assay must be robust, and thus not readily perturbed by non-specific effects such as low concentrations of organic solvent. The access requirement strongly favours permeabilized cells and extracts, as designing appropriate small molecules that can cross cell membranes effectively is non-trivial. The type of system I have in mind includes such processes as mitosis and cell-cycle regulation in *Xenopus* egg extracts [8], vesicle trafficking in permeabilized cells [9] or any of a number of complex signal transduction pathways reconstituted *in vitro*. These are the prejudices of a basic cell biologist, and the generality of combinatorial pharmacology is limited mainly by our imaginations. It is unclear how serious a problem the impermeability of the cell membrane presents for screening; a compound library engineered to cross membranes would allow readier access to medically relevant complex systems in living cells.

I will illustrate the possibilities of the combinatorial pharmacology with a specific system from my laboratory, mitotic chromosome assembly in a cytoplasmic extract derived from *Xenopus* eggs [10]. In this system the DNA in added sperm nuclei is converted into condensed mitotic chromosomes through a series of morphological steps that correspond to specific molecular events, each of which presumably requires multiple unidentified protein components. A small molecule inhibitor of chromosome assembly would be an interesting lead compound for cancer therapy, but the primary goal of our research is to dissect the system, identify key proteins, and determine their function.

Aliquots from a compound library would be added to aliquots of the assay mixture, and the result screened for inhibition of chromosome assembly and accumulation of specific intermediates. The compounds in the library could be free in solution or bound on a bead; free compounds would presumably inhibit chromosome assembly by directly binding to a target protein, whereas bound compounds would inhibit the process by sequestering the protein onto the bead. After identifying an initial hit from a library aliquot, the specific inhibitor would be isolated. For soluble compounds the most straightforward approach would be to resynthesize the compound pool in multiple aliquots that each contain a subset of the original mixture. The library would thus continue to be subdivided until the active pool contained only the compound of interest. For bead-bound compounds the best inhibitor could be identified by direct analysis or by more complex indexing strategies.

From drug to protein

With a specific inhibitor (or activator) in hand, the next step would be to identify the target protein. The most useful strategies would be affinity chromatography and/or cross-linking to a radioactive ligand. The beauty of the combinatorial approach is that the drugs can be pre-designed to incorporate chemical groups appropriate for these techniques. I will refer to the part of the molecule designed to facilitate protein identification as the handle (Fig. 1a). The handle could be present during the initial screen, or added later to a convenient attachment point. It could be added as one of the monomer units, or it could replace the matrix attachment site used during synthesis. Designing effective handles that combine ease of synthesis with lack of interference with the targeting reaction will require some chemical creativity.

Returning to chromosome assembly, the specific inhibitor that we fished out from the compound library would then be synthesized with a handle, attached to a column matrix, and used to bind the protein target (Fig. 2). Eluting the affinity column with a soluble version of the inhibitor would increase the specificity of this technique, and when isolating a native protein it may increase the elution efficiency to use a modified form of the inhibitor on the column that binds with lower affinity, then elute with the original higher affinity form. The ease of inhibitor modification that is inherent in the combinatorial chemistry should greatly facilitate such approaches. Affinity chromatography may fail if the compound targets a specific protein–protein (or protein–DNA) complex, and in this case cross-linking would be useful. The handle would then be made to incorporate a radioactive tracer, and a reactive group would be attached to the inhibitor and used to label the protein of interest (Fig. 2). Powerful denaturing fractionation techniques, such as electrophoresis and reverse phase chromatography, would then be used to isolate the protein of interest for sequencing. The two techniques could be combined by developing a handle that contains both a cross-linking group and a tag suitable for affinity chromatography such as biotin.

What we need from the chemists

To realize the possibility of pharmacological genetics the three things we most need from chemists are advice, molecules, and access. The requirements for the targeting part of the drug molecules (Fig. 1a) correspond well with the current goals of combinatorial pharmaceutical development programs. We do not yet know what chemistries will be most effective for targeting specific proteins, and trying to inhibit complex biochemical systems may prove to be a useful way to identify effective chemistries. The main difference between laboratory and company uses of combinatorial chemistries will be the less stringent laboratory requirements for clinically important properties, such as biostability and pharmacokinetics.

The molecular requirements for the handle are more specific. As we want to optimize the molecules for affinity chromatography from the outset, it may be best to perform the screen with diverse molecules to which a uniform hydrophilic spacer is already attached. This chain

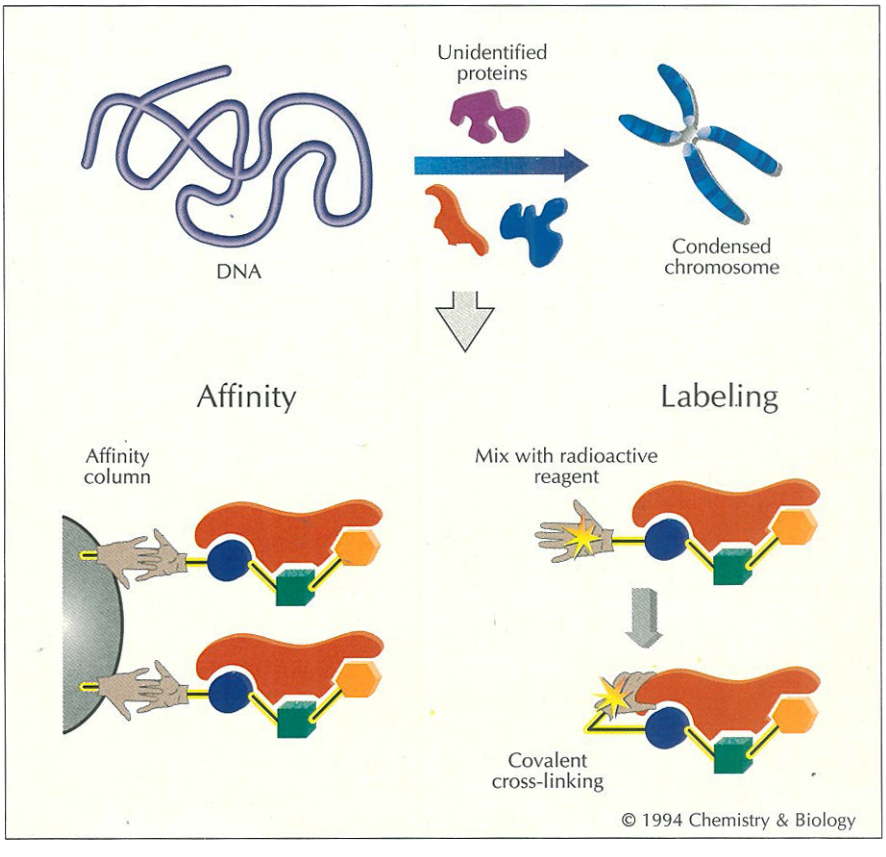

Fig. 2. Protein identification. Once a specific drug has been selected from the library, two strategies can be used to identify its target protein. On the left the drug is attached to a column matrix by its handle and used for affinity chromatography. The same matrix could be used for both synthesis and protein isolation for bead-bound drug libraries. On the right the handle is a radioactive, reactive (or photoreactive) group that can be used to covalently label the target protein. The protein is then isolated by conventional methods, using the attached drug as an affinity tag if necessary.

could also aid in water solubility; for this purpose, a short polyethylene glycol or polyamide springs to mind, or perhaps a biotin or similar ligand. From the chemical perspective it would probably be easier to introduce a uniform attachment point on the diverse molecules, such as a carboxylic acid or amine group, that could be converted to an ester or amide during the screening phase, and then be replaced by a spacer arm for the protein isolation phase. One risk of this approach would be that the attachment point would become part of the targeting site, and thus targeting would be impaired once the attachment group was added. In addition, it would be necessary to avoid using the chemical group that is used for attachment in the variable part of the molecule, and this would somewhat limit the potential diversity of the library. Clearly some chemical ingenuity will be required, along with optimization in real screens.

The other requirements are advice and access to molecules. To maximize the impact on protein discovery we need to get the combinatorial chemistries into the laboratories of academic chemists and biologists. As a collaborative effort, a combinatorial pharmacology screen could be made as straightforward as a genetic screen, with intellectual rewards for both the chemist and the biologist.

The realm of the possible

A combinatorial pharmacology screen sounds good on paper, but would it actually work? The answer hinges on two issues. What kinds of proteins will be drug targets, and can we sample sufficient diversity within the confines of a reasonable synthesis and screening workload? At this stage in the technology the pharmacology cannot sample as much diversity as real genetics, at least for simple organisms like yeast where millions of organisms can be screened on a few dishes. Any protein is a target for mutation, but is any protein a potential drug target? In principle yes, but currently we can only assert this with confidence for receptors and enzymes, where the active site presents a logical small-molecule target. Even if we could only target enzymes, combinatorial pharmacology would still be an enormously powerful tool. For example, most complex processes require enzymes that use nucleotide triphosphates in some way, such as kinases, GTPases or mechanochemical enzymes. Identifying just these enzymes would be a powerful start to dissecting any process, and one challenge for chemists would be to develop a combinatorial chemistry that specifically targets this class of molecule. In the end we will not learn the limits of combinatorial pharmacology until we try it, and by using complex systems rather than single proteins as targets we maximize our chance of encouraging early hits. In fact, starting with complex systems may be the best way of addressing the question of what types of protein we can effectively target for drug design.

References

1. Shelanski, M.L. & Taylor, E.W. (1967). Isolation of a protein subunit from microtubules. *J. Cell Biol.* **34**, 549–554.
2. Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P. & Gordon, E.M. (1994). Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem.* **37**, 1233–1251.
3. Gordon, E.M., Barrett, R.W., Dower, W.J., Fodor, S.P. & Gallop, M.A. (1994). Applications of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies, and future directions. *J. Med. Chem.* **37**, 1385–1401.
4. Kenan, D.J., Tsai, D.E. & Keene, J.D. (1994). Exploring molecular diversity with combinatorial shape libraries. *Trends Biochem. Sci.* **19**, 57–64.
5. Dower, W.J. (1992). Phage power. *Curr. Biol.* **2**, 291–253.
6. Kobayashi, R., Wigler, M. & Still, W.C. (1993). Complex synthetic chemical libraries indexed with molecular tags. *Proc. Natl. Acad. Sci. USA* **90**, 10922–10926.
7. Brenner, S., & Lerner, R.A. (1992). Encoded combinatorial chemistry. *Proc. Natl. Acad. Sci. USA* **90**, 5381–5383.
8. Murray, A.W. (1991). Cell cycle extracts. *Methods Cell Biol.* **36**, 581–605.
9. Pryer, N., Wuesthube, L.J. & Schekman, R. (1992). Vesicle mediated protein sorting. *Annu. Rev. Biochem.* **61**, 471–516.
10. Hirano, T. & Mitchison, T.J. (1991). Cell cycle control of higher-order chromatin assembly around naked DNA *in vitro*. *J. Cell Biol.* **115**, 1479–1489.
11. Simon, R.J., *et al.*, & Bartlett, P.A. (1992). Peptoids: a modular approach to drug discovery. *Proc. Natl. Acad. Sci. USA* **89**, 9367–9371.
12. Bunin, B.A., Plunkett, M.J. & Ellman, J.A. (1994). The combinatorial synthesis and chemical and biological evaluation of a 1,4-benzodiazepine library. *Proc. Natl. Acad. Sci. USA*, **91**, 4708–4712.

Tim J. Mitchison, Department of Pharmacology, UCSF, San Francisco, CA 94143, USA.