Multi-target lead discovery for networked systems

Curtis T Keith & Grant R Zimmermann
CombinatoRx, Boston, MA, USA

With the advent of systems biology, the pharmaceutical industry needs to look again at the 'one-drug, one-target' paradigm.

The 'one-drug, one-target' principle is a core assumption of our current drug discovery paradigm and has been a mainstay of the pharmaceutical industry. This approach has yielded many notable successes, even if the underlying correspondence of biology (the targets) and chemistry (the drugs) is frequently more complicated than one-to-one in reality.

However, there is now widespread recognition that the prevailing approaches have not provided sufficient increases in the number of new chemical entities reaching the market to support increasing R&D expenditures. This has prompted some researchers to make long-range projections for the productivity of the industry, often basing their analyses on the wealth of genomic and other data amassed in recent years. Questions such as how many good disease-modifying druggable targets remain and how many truly differentiated (first-in-class) NCE drugs will the future hold, are common. A favored approach has been to estimate the intersection set of 'disease-modifying targets' and those proteins that are 'druggable' by current definitions, i.e., with existing 'rule-of-five' chemistries. At the same time, others dispute this line of reasoning altogether, suggesting that the disease-modifying gene set is limited only by our current biological understanding, and that the size of the druggable set is bound only by the current limitations of synthetic chemistry.

We propose a fundamental reconsideration of the one-drug, one-target premise. An alternative systems biology perspective recognizes that a drug target may comprise a set of two or more interacting proteins that are connected in a disease network. This perspective broadens the target definition to cases where a complementary action elsewhere in the system is required. This effectively expands the disease-modifying gene set and concomitantly the overlap with the druggable genome. In the limit, the intersection set encompasses the entire druggable genome. This view provides a framework for envisioning many new multi-target mechanisms that cannot be achieved by single-point, stand-alone interventions. Formulating a path to discovering such drugs presents challenges and has important implications for lead discovery and product development.

Expanding target space

The human genome contains approximately 30,000 genes. These give rise to 150,000 or more proteins via splice variations, and even more forms after post-translational modification. Traditionally, only a small fraction of these are predicted to be directly related to human disease. For example, approximately 10% of genes (~3000) have been described as disease-modifying based on knockout studies in mice, and this number has often been used as a ballpark figure for the number of disease-modifying targets. The druggable genome has been described as a set of ~3000 genes whose predicted proteins can be expected to bind rule-of-five drugs. This number is extrapolated from known drug-binding proteins and includes all protein family members with the same fold. The extent of overlap between the disease-modifying and druggable sets of genes is a topic of debate, but is estimated to be as small as 600-1500 based on the yeast genome and the number of antifungal drug targets (Figure 1).

Figure 1. Expanding disease-modifying target space. The partial overlap (green) between the druggable genome (blue) and the set of disease-modifying genes (yellow) present in the human genome (adapted from Hopkins and Groom). By applying compounds in combination, the set of disease-relevant genes may grow to encompass nearly the entire druggable genome (bottom).
Current drugs are generally thought to act at ~300-400 of these targets, and this relatively small intersection set has prompted some to suggest stark implications for the pharmaceutical industry in the long term.

Others analyse suggest that the number of genes that actually contribute to clinical pathophysiology, directly or indirectly, is significantly greater than 3000. Some estimates put the number of disease-modifying genes as high as 10,000 by recognizing that genes which are ‘linked’ through a pathway to a disease-modifying node in the network may also be suitable drug targets. Moreover skeptics point out that as our understanding of biology improves, proteins currently discounted as non-disease-modifying (such as house-keeping genes), may be found to contribute to disease.

From a systems perspective, the key biological entities relevant to human disease (thereby candidate disease-modifying targets for drugs) are actually coordinated sets of proteins with functional connections or pathway relationships. Embracing this view recognizes that many proteins are not disease-modifying in isolation, but rather, acquire therapeutic relevance by virtue of their functional relationship to another gene (that on its own may or may not be disease-modifying). Interpreting this principle in the broadest sense, it may be that most genes, including the entire druggable genome, are in fact disease-modifying in the context of a complementary action, eg, treatment with another agent or the presence of a genetic mutation(s) (Figure 1).

**Accessing novel mechanisms**

In many cases, action at multiple targets in a disease network will result in a new mechanism that is not represented by the targets on their own. This is particularly true when the interaction is ‘synthetic’ in the genetic sense, or synergistic in the compound sense (also termed coalism). If this is true, how large is the novel mechanism space that might be accessed by multi-target action? How many of these might be disease-modifying?

Figure 2. Many novel multi-target mechanisms for lead discovery. A small fraction of genes will interact with functional consequence when tested in combination (eg, 1%), and a small fraction of those interactions will be disease-modifying (eg, 1-10%). As a result, there may be as many as 10,000 to 100,000 novel multi-target mechanisms to be exploited for therapeutic benefit.

Genetic double-deletion experiments in yeast suggest that ~1% of non-essential genes induce a synthetic lethal or sick phenotype in the background of another non-essential knockout. Extrapolation to the set of ~32,000 human genes defines a space of 500 million pairwise gene combinations. An estimate of a 1% synthetic interaction rate implies that for a particular phenotype, there could be as many as 5 million gene combinations where the effect of the combination on phenotype is significantly different from the sum of the components. Different assays and disease backgrounds are likely to produce different sets of interacting genes, as well as different rates. Even assuming just a small fraction of functionally interacting pairs for a particular disease model are ultimately useful, the set of new network-based mechanisms could be ~10,000-100,000 (Figure 2). This represents a potential wealth of unexplored targets - and moreover, only considers gene pairs, whereas higher order combinations may also generate desirable systemic effects.

**Existing combination drugs**

Despite best efforts to validate targets at the beginning of the discovery process, development of single-target, magic-bullet drugs continues to be challenged by the ‘systems’ nature of biology. Clinical drop-off because of insufficient efficacy or failure to meet desired endpoints is common, even when predicted pharmacodynamic actions are achieved. This may be attributable in part to the reality that disease systems are dynamic and complicated, having partially redundant pathways that can be difficult to account for in model systems or recognized during target validation. Clinicians have historically overcome single-target insufficiency by administering multiple drugs as a combination therapy resulting in greater benefit for patients. A survey of currently marketed combination drugs provides compelling examples of multi-target therapeutics achieving efficacy superior to single-agent interventions in multiple therapeutic areas. Three examples from infectious disease are discussed below, other combinations used in areas including oncology, inflammation, asthma, and heart disease are not discussed.

Antiviral drug development for HIV has embraced combination therapy almost from the beginning. In modern HAART therapy, two reverse transcriptase inhibitors are given in a triple compound cocktail with a third agent, usually a protease inhibitor. By simultaneously inhibiting two critical viral systems, the cocktail achieves a greater level of efficacy. Perhaps more importantly, by slowing viral replication, emergence of resistant strains is reduced.

Augmentin, (developed by GSK) combines amoxicillin and clavulanic acid, a competitive inhibitor of β-lactamase that is the primary resistance mechanism against penicillin. Augmentin is a rare example of a combination in which one of the components is not used clinically as a single agent. Indeed, clavulanic acid has no therapeutic benefit outside the combination with the active component amoxicillin. In the case of a penicillin-resistant infection, the amoxicillin is only weakly active because it is rapidly degraded by the resistance mechanism of the bacteria. In this clinical context, Augmentin behaves as a ‘synthetic-lethal’ (coalism) combination where neither of the components have any detectable activity in isolation.
Bactrim is the combination of trimethoprim and sulfamethoxazole, both effective inhibitors of the folic acid pathway in bacteria required for nucleoside biosynthesis (Figure 3 top). Sulfamethoxazole is a competitive inhibitor of dihydropteroate synthase (DHPS). Trimethoprim inhibits dihydrofolate reductase (DHFR) and is used as a single-agent to treat urinary tract infections. By targeting two points in the folate synthesis pathway simultaneously, the combination of these drugs (Bactrim) is able to achieve greater efficacy than either in isolation and is recommended as the first-line treatment of respiratory tract infections (primarily *Streptococcus pneumoniae*) by the World Health Organization. The synergistic activity of the combination of these drugs can also be detected in vitro using the dose-ratio matrix screening methodology in methicillin-resistant *Staphylococcus aureus* (Figure 3).

**Multi-target lead discovery**

The success of existing combination drugs indicates that multi-target therapeutics can be clinically effective. Additionally, the networked nature of biology suggests a vast wealth of potential drugs acting via coordinated targets remains to be discovered. Rational design or prediction using in silico network models may one day identify effective multi-target interventions. Success with such an approach will lie far in the future and we propose alternative empiric methods for the present.

Conceptually, empiric multi-target discovery is the genetic (or chemical genetic) approach applied in combination space. This requires complex assays as phenotypic models of disease and probe agents that perturb function of individual components of the system (perturbagens).

An advantage of using combinations as genetic (or chemical genetic) tools for discovery is that they can reveal disease-modifying targets and mechanisms that would be missed by traditional single-agent approaches.

**Cell-based phenotypic assays**

An ideal phenotypic assay system encompasses all of the relevant aspects of disease biology. Reasonable approximations can be made using phenotypic cellular models, which capture some, but not all biological complexity. Cells integrate the multiplicity of effects induced by compound treatment (as single agents or in combination), and play a particularly important role in the discovery of drugs with multi-target mechanisms. Phenotypic assays do not require detailed a priori knowledge of target function or connectivity, yet still test targets in a comparatively ‘native’ environment (relative to cell-free, target-based screens). Moreover, by focusing on the end-state of the system, these assays facilitate discovery of solutions that yield the desirable outcome.

**Probes for modulating activity**

Probe agents that conditionally perturb genes or protein function are systematically tested in combination to look for disease-modifying effect - ideally, restoration of a ‘healthy’ state in the phenotypic assay. A variety of agents can be applied in cell-based assays as perturbagens, including drugs or other small molecules, biomolecules such as antibodies, cytokines, growth factors, etc. Ideally these are capable of either agonism or antagonism, and can independently modulate the activity of individual functions of multi-domain proteins. Chemical genetics has the advantage that modulating function via small molecules is not inherently limited to loss-of-function effects, as is the case using siRNA knock-down. On the other hand, target coverage by siRNA libraries is currently more comprehensive, although Schreiber and others are endeavoring to improve chemical genetic libraries for the future.

**A pragmatic approach**

Truly comprehensive versions of chemical genetic screening for multi-target mechanisms must wait until larger sets of well-defined functional probes have been created. However, a variation on the combination chemical genetic approach
that is currently feasible utilizes a different compound set – the existing pharmacopeia. These compounds are less preferable for purely chemical genetic studies because of reduced target confidence, limited target diversity, and in some cases, marginal specificity.

However, they possess other valuable characteristics instead, such as well-characterized pharmacology and ADMET, ability to manufacture, formulate, etc. These qualities allow them to be advanced to proof-of-concept clinical trials in combination much more rapidly than would be possible for NCEs. This also provides the opportunity to streamline clinical validation of newly identified multi-component targets.

Numerous challenges needed to be overcome to create a drug discovery platform for systematically evaluating combinations of compounds in complex cellular models of disease. This form of discovery requires cell-based screening at high-throughput to efficiently search the large numbers of combinations that arise in factorial space. It has also necessitated the creation of new LIMS and informatics capabilities to allow combination analytics and automated synergy determination.

Dose-response surfaces measured for each pair of compounds are compared to theoretical additive-interaction surfaces (computed using standard reference models) to determine if an interaction is synergistic (Figure 4).

In the near term, this approach can be applied not only for repurposing approved drugs, but should also be used for valuable compounds that modulate shelved targets that had been abandoned due to insufficient efficacy. Many of these assets can be resuscitated - in the first case, to achieve greater effect in their original indications, but alternatively they can be repurposed into new indications in the context of a combination.

**System combinations**

Combination drugs are the standard-of-care in many therapeutic areas, including cancer, diabetes and infectious disease. They will become even more prevalent in the future as multigenic diseases become increasingly important target diseases for our industry. To date, most combination drugs have arisen at the clinical stage owing to efforts to combine for superior patient treatment. The application of multi-target principles to earlier stages of discovery will be more efficient and allow the discovery of fundamentally new types of drugs, including combinations of compounds that are entirely inactive when used on their own.

The reductionist view of biology that prevailed in much of the twentieth century is giving way to a ‘systems’ view of biology, and this is having an impact on our understanding of disease and how it should be treated. For some disease systems, action at a single node may in fact be sufficient for desired therapeutic outcome. More often than not, optimal therapeutic intervention will require coordinated action at multiple points. Achieving this by utilizing mixtures of targeted small molecules seems the most straightforward extension of previous successes with combination drugs. However, a compelling alternative is focused multi-selectivity or polyvalency within a single molecule. These approaches share a common principle at their origin, but the technical challenges and risks associated with the two paths are very different. Only the future will reveal whether both approaches become widely used or whether one will dominate.

Finally, there is another pragmatic reason to embrace the concept of multi-
targeted action that should be considered. Current pharmaceuticals are estimated to act through as few as 400 molecular targets, and nearly half of these fall into just a handful of categories (GPCRs, kinases, proteases, NHRs, and PDEs). These are enzymes or ligand receptors with active sites that can be successfully targeted by low-molecular weight compounds using current chemistries. We should maximize these chemically tractable targets, using them in combination and thereby broadening their disease applications wherever possible.

Acknowledgements

We would like to acknowledge Jeb Ledell, Joseph Lehar and Ken Mullen for assistance in the preparation of figures.