

**Searching for Molecular Solutions – Additional Material****CHAPTER 9**

These Files contain additional material relevant to **Chapter 9** of *Searching for Molecular Solutions*. The page numbers of the book pertaining to each section are shown in the Table below, the corresponding page number for this file, and the title of each relevant section.

**Contents:**

~ Book Reference Page Number	Page Number in this File	Section	
		No.	Title
319	3	<a href="#">A14</a>	Rational Design Examples See <a href="#">Subsections</a> below
338	19	<a href="#">A15</a>	The Interactome and Biological Parsimony See <a href="#">Subsections</a> below

<b>Subsection titles for Section A14</b>	<b>Page Number in this File</b>
Rationally Facing a Potent Viral Challenge	<a href="#"><u>3</u></a>
Elusive Tumor Targets and Rational Bullets	<a href="#"><u>11</u></a>
<b>Subsection titles for Section A15</b>	<b>Page Number in this File</b>
The Incredibly Intricate Interactome: Polyomic Problems and Promise	<a href="#"><u>19</u></a>
Biological thrift and drug targeting	<a href="#"><u>22</u></a>
The Dark Side of Thrift – Drug Cross-Reactivity	<a href="#"><u>27</u></a>
The Bright Side of Thrift – Target Diversity, Drug Repositioning, and More	<a href="#"><u>32</u></a>

## Section A14: *Rational Design Examples*

Relevant to the extended treatment of Rational Design approaches (from p. 319 of *Searching for Molecular Solutions*) as specific case examples.

### *Rationally Facing a Potent Viral Challenge*

Pathogens can present both opportunities and special challenges as drug targets, depending on the specific threat involved. For a long time, it might have seemed that the 'bigger they are, the harder they fall' principle held true in this area, since eukaryotic parasites and bacteria appeared to be much more druggable than viruses. But size *per se* is not the real issue, but rather the ability to identify non-host targets which are crucial to the pathogen life-cycle. It is the relative simplicity of viruses, combined with their co-opting of host processes, which renders them less tractable as drug targets. But as a general statement, it is very evident that they are not at all undruggable entities, though the details of this of course vary on a case-by-case basis. The story of the development of effective drugs against targets encoded by the human immunodeficiency virus (HIV) has many connections with rational molecular design, and is of interest from a number of different but related viewpoints. It is also one of the best exemplars of the accelerating pace of innovation in drug discovery. A gap of only 15 years exists between the medical recognition of AIDS itself in 1981 and the first clinical application of highly active anti-retroviral therapy (HAART), which though not a cure can convert AIDS from a death sentence into a chronic disease. In this intervening time (agonizingly long as the world watched the AIDS death toll, but the blink of the eye in historical terms) the viral life cycle had to be

defined, specific viral targets identified, and multiple avenues explored for anti-HIV drug discovery.

Although the HIV life-cycle offers a number of potential targets for therapy, an important early observation was that (in common with other retroviruses) HIV structural and functional genes are expressed as fused continuous *polyproteins*<sup>♥</sup>, which require cleavage at multiple points by a virally-encoded protease<sup>1</sup>. Genetic abrogation of functional HIV protease in viruses cultivated in cell culture proved that it was essential for viral propagation, and thereby established its significance as a drug target<sup>3</sup>. This small protease (99 amino acid residues) forms its active site as a dimer, and cleaves a variety of substrates at different rates<sup>4</sup>. While empirically-obtained data has provided some leads for further development, the structure of the HIV protease has been pivotal in drug design. Also, consideration of the nature of the peptide substrates for the protease has been very important.

Molecular design based on an enzyme substrate is a conceptual descendant of strategies screening analogs of known target protein ligands as potential antagonists, as we have seen with the development of cimetidine, as described in Chapter 9 of *Searching for Molecular Solutions*. In general, development of enzyme inhibitors from first principles can be placed on a much more rational footing by using the predicted structures of transition-state intermediates as a starting point for synthesis of analogs. (We might recall from Chapter 7 that the same knowledge can be applied in a different mode, towards the generation of catalytic antibodies). But the story of the development of the first clinical HIV

---

<sup>♥</sup>This is an interesting example of the efficient compacting of virally-encoded information, as also noted with SV40 Large T antigen (see the file SMS–CitedNotes-Ch9/Section 30; from the same ftp site). The protease itself is part of the Gag-Pol polyprotein, and cleaves itself out<sup>1</sup>. From the HIV genome, three polyproteins are expressed (Gag, Gag-Pol, and Env), and several other accessory proteins are expressed early in infection from separately spliced viral mRNAs<sup>2</sup>.

protease inhibitor is firmly to the right of the rational design spectrum <sup>▼</sup> (Fig. 9.1), since it incorporates not only the structure of the target protein (the protease enzyme), but also the nature of the enzymatic cleavage of the substrate.

To design an effective substrate-mimic (mimetic) which binds and inhibits the action of an enzyme, knowledge of the normal substrate processing is obviously important. Where the substrate is a peptide sequence cleaved through the enzyme's catalytic activity, a designed *peptidomimetic* should be non-hydrolyzable and thereby prevent catalytic turnover. To do this, all or part of the substrate molecule needs to be altered such that the relevant moiety is replaced with another chemical group of similar size and shape (an *isosteric* replacement). Here knowledge of the structure of the binding site, as well as the substrate itself, is clearly of great utility as a guide, and the developmental history of inhibitors of the HIV protease has consequently had major input from virtual screening approaches <sup>5</sup> and QSAR <sup>6</sup>. Such strategies led to the development of the first effective anti-HIV protease inhibitor in clinical use, saquinavir <sup>4</sup> (Fig. 9A14.1, below).

---

<sup>▼</sup>Since such design makes use of available knowledge of the substrate, it is not *de novo* in the sense of designing a ligand for a protein when provided only with the target structure. (Further discussed in Chapter 9 of *Searching for Molecular Solutions*).

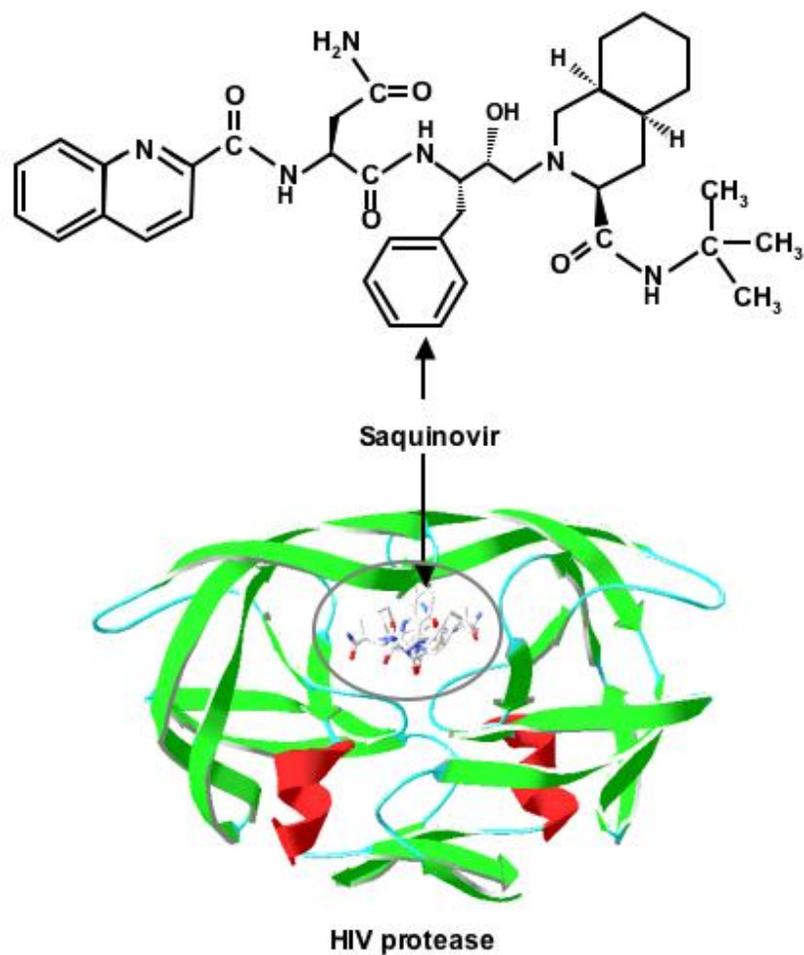


Fig. 9A14.1

Structural formula of the HIV protease inhibitor Saquinovir (Fortovase; Invirase) and in complex with the symmetrical HIV protease dimer<sup>7</sup> (saquinovir encompassed by gray oval). Green segments  $\beta$ -strands; red,  $\alpha$ -helices. Source: [Protein Data Bank](#).<sup>8</sup>; [1HXB](#). Image generated with [Swiss-pdb viewer](#)<sup>9</sup>.

As we have seen in Chapter 8, there is more to drug development than just target binding. In any circumstances, a compound with excellent affinity and selectivity for a target might still fail as a practical drug if any of its performances for the absorption, distribution, metabolism, excretion and toxicity (ADMET) criteria are judged as unacceptable. The issue of drug bioavailability is especially important in chronic diseases, within which HIV infection is certainly classifiable owing to the lack of any effective means for permanent viral eradication<sup>10</sup>. Many promising anti-HIV protease drugs arising through ingenious design efforts have unfortunately fallen by the wayside owing to poor bioavailability. Saquinavir (Fig. 9A14.1) itself has very sub-optimal bioavailability, but this can be enhanced by improved formulations for its oral administration<sup>4,11</sup>. Another strategy is to improve bioavailability with other drugs given at the same time. For example, the HIV protease inhibitor ritonavir itself has reduced efficacy *in vivo* due to serum binding, but usefully retards drug metabolism by inhibiting a specific cytochrome P450 isoform. In so doing, it improves the bioavailability of other potent protease inhibitors when co-administered with them<sup>4,12</sup>.

The biggest challenge with anti-HIV drugs is coping with a constantly moving target. Owing to their error-prone replication, retroviral populations in general are characterized by high mutability and evolvability, and HIV is certainly no exception in this regard. The resulting variability is such that a large replicating set of HIV is really a population of variants constantly driven by high reproductive error rate and selection (removing deleterious forms), which can be modeled as a 'genotype cloud' or *quasispecies*<sup>13</sup>. HIV provides us with some spectacular examples of evolution in action, but this is a major headache for pharmacologists and clinicians, not to mention a threat to the large global infected population. Mutations in the HIV protease which diminish or abolish the efficacy of protease inhibitors were observed soon after their introduction, forcing continued re-design efforts. While the problem is aggravated by the high mutation rate of HIV, it can be seen as a generalizable consequence of 'promiscuous' protein activities.

The binding of a drug by any enzyme can be viewed as an additional promiscuous (albeit non-natural) function<sup>14</sup>. But precisely because the drug-binding is obviously not central to the enzyme's catalytic function, the promiscuous activity can often be modified without compromising enzymatic activity. In general, mutations conferring drug resistance often occur in protein loops normally involved with substrate-binding, rather than by altering the catalytic mechanism or the protein framework<sup>14</sup>, a much greater 'design' challenge for evolution. In the specific case of HIV protease, a variety of mutations (proximal to the active site) have been observed where changes in the binding affinities for specific drugs are indeed observed<sup>15</sup>. Other mutations have nonetheless been found outside this region, and while the fundamental scaffold of the protein is preserved, these 'distal' mutations can change the geometry of the active site to the detriment of drug binding, but to the benefit of the enzyme's resistance<sup>15,16</sup>. To acquire resistance against some drugs, the protease appears to need multiple mutations, which comes at a cost of decreased fitness (loss of protease efficiency)<sup>15,17</sup>. This tends to be overcome by secondary compensating mutations as prolonged drug selection drives viral evolution<sup>17,18</sup>. A further complication in this regard is that mutations in the *substrate* sites for the protease can themselves act as resistance mutations<sup>19</sup>. In other words, when even wild-type HIV protease is targeted at specific variant substrates, at least some inhibitors (potent in normal circumstances) fail to effectively block the enzyme activity.

Later-generation drugs against the HIV protease have been rationally designed to counter multiple drug-resistance mutations. Darunavir, approved for clinical use in 2006, is one such example<sup>20</sup>. We have focused on the protease in this sketch, but it should not be thought that other rational targets have not been investigated, and these are found at multiple stages of the viral life-cycle<sup>6</sup>. Recent approvals have been granted for novel drugs antagonizing HIV host cell entry, reverse transcription, and integration into host cell DNA<sup>21-23</sup>. Targeting HIV accessory protein interactions is also a possibility<sup>24</sup>. Acquisition of resistance to

any virally-encoded target is logically to be expected in the same manner as for the protease, but possession of a multiplicity of anti-viral weapons is clearly very useful for trying to short-circuit the continual arms race between the evolving virus and drug designers. Ideally, a drug should be designed to interact strongly with regions of a target which are known to be conserved, and have flexibility in its modes of interaction with additional target residues which are subject to variation, in order to present the greatest possible functional target range<sup>25,26</sup>. 'Adaptive inhibitors' of this type have been mooted for the HIV protease<sup>27</sup>, but the ability of mutations to change the protease active site geometry renders the generation of a universal inhibitor a true challenge for design.

HIV, and the acquired collapse of immunity which it engenders, present formidable hurdles to overcome before a complete cure can be contemplated. Apart from its targeting of the CD4 helper T cells of the immune system itself, by crossing the blood-brain barrier HIV can infect the brain and nervous system, with potentially devastating consequences<sup>28</sup>. The ability of HIV to reverse-transcribe and integrate into host genomes is particularly problematic for complete viral eradication, as viral rebound after cessation of anti-retroviral therapy is attributed to activation of HIV proviruses within latently infected cells<sup>29</sup>. This complexity demands a multi-pronged overall assault against the threat, beyond small molecule drugs where appropriate. For example, anti-protease aptamers have been isolated<sup>30</sup>, although these have not yet moved into the clinic. A number of genetic approaches also have been actively pursued, including ribozymes and RNAi<sup>31</sup>, and zinc finger nucleases or other 'genomic editing' techniques may be very useful for engineering viral resistance in patient lymphocytes<sup>32</sup>. Antibody therapies for blocking viral uptake or killing of infected cells are also an option<sup>32,33</sup>.

An effective prophylactic vaccine, though elusive to date, would block the spread of the infection but not cure individuals already burdened with the virus. Rational

---

▼ This was referred to in the file SMS–CitedNotes-Ch4/Section 8B; from the same ftp site.

strategies with small molecules have gone far towards preventing viral propagation, but the above-mentioned genomic persistence of proviral DNA might seem virtually insoluble. With small molecules alone, indeed it might be, but deployment of artificial macromolecular tools has the potential to remove the offending foreign viral DNA. Promising work *in vitro* has been performed with the Cre recombinase from the bacteriophage P1 <sup>▼</sup>. Directed evolutionary approaches have been used to change the specificity of Cre such that it excises HIV proviral segments and ‘cures’ infected cells in culture <sup>35</sup>. Obviously, many challenges remain in terms of rendering this treatment clinically effective, but the process has become possible in principle.

Before leaving this topic, we could consider the Cre recombinase as a useful molecule bestowed upon us by a bacterial virus. Somewhat ironically, considering its terrible impact, HIV too has made a significant contribution to biotechnology in the form of *lentiviral* vectors <sup>\*</sup>, which have numerous applications <sup>36</sup>. The efficiencies and economies of the HIV life-cycle can also be studied from the point of view of systems biology, briefly considered in Chapter 9.

We thus learn from our enemies, and while AIDS is new, the next enemy we will think about in a druggable sense is almost certainly as old as multicellular life....

---

<sup>▼</sup> The application of this recombinase and its specific target ‘lox’ sites has become important in genomic engineering following targeted homologous recombination, especially for the generation of mice with specifically modified genetic alleles <sup>34</sup>.

<sup>\*</sup> Lentiviruses (which include HIV) are a subset of retroviruses which can infect non-dividing cells and integrate into their genomes, a property which makes them valuable when permanently inactivated as viruses and converted into vectors for delivering payloads for expression.

### *Elusive Tumor Targets and Rational Bullets*

Many animal tumors and some tumors of humans are known to arise as a direct or indirect consequence of infection with specific viruses, and HIV-mediated immunosuppression promotes the formation of many cancers, especially those associated with other viral agents<sup>37</sup>. But broadly speaking, any invasive life-threatening tumor can be viewed as a 'parasite' in the sense that it replicates aggressively at the expense of its host. And at least for specific tumors of dogs and Tasmanian devils<sup>38-40</sup> the label of parasite is literally apt, since the tumor itself acts as a transmissible agent between members of the host species. Irrespective of this distinction, tumors arise through somatic mutational changes, which enable Darwinian selection of variants with favorable replicative or host-evasive properties<sup>41</sup>. At the same time, cancer-related somatic changes can afford a foothold for drug targeting, which is otherwise confronted with a 'parasite' which is virtually the same as Self.

In Chapter 9 of *Searching for Molecular Solutions*, we looked at some essentially empirical high-throughput processes for obtaining anti-cancer drugs, and now is an appropriate juncture to look at the same target area with a rational eye. So for this to succeed, either a tumor-related 'non-self' target is required, or a host target whose inhibition will not prove to be excessively toxic. The latter has been the traditional domain of cytotoxic drugs, which trade on the high replication rates of cancer cells. The hope in such circumstances is that the malignant transformed cells can be killed quicker than the host would be through the same treatment, and not surprisingly side-effects are often severe<sup>▼</sup>. A 'magic bullet' therefore requires a special handle on a tumor cell which can distinguish it from its host, often a veritable needle in a haystack.

---

▼ An example of a promising host pathway for anti-tumor drug development (*Sonic hedgehog*) and its potential for side-effects is given in [Section A15](#) (Biological Parsimony) below.

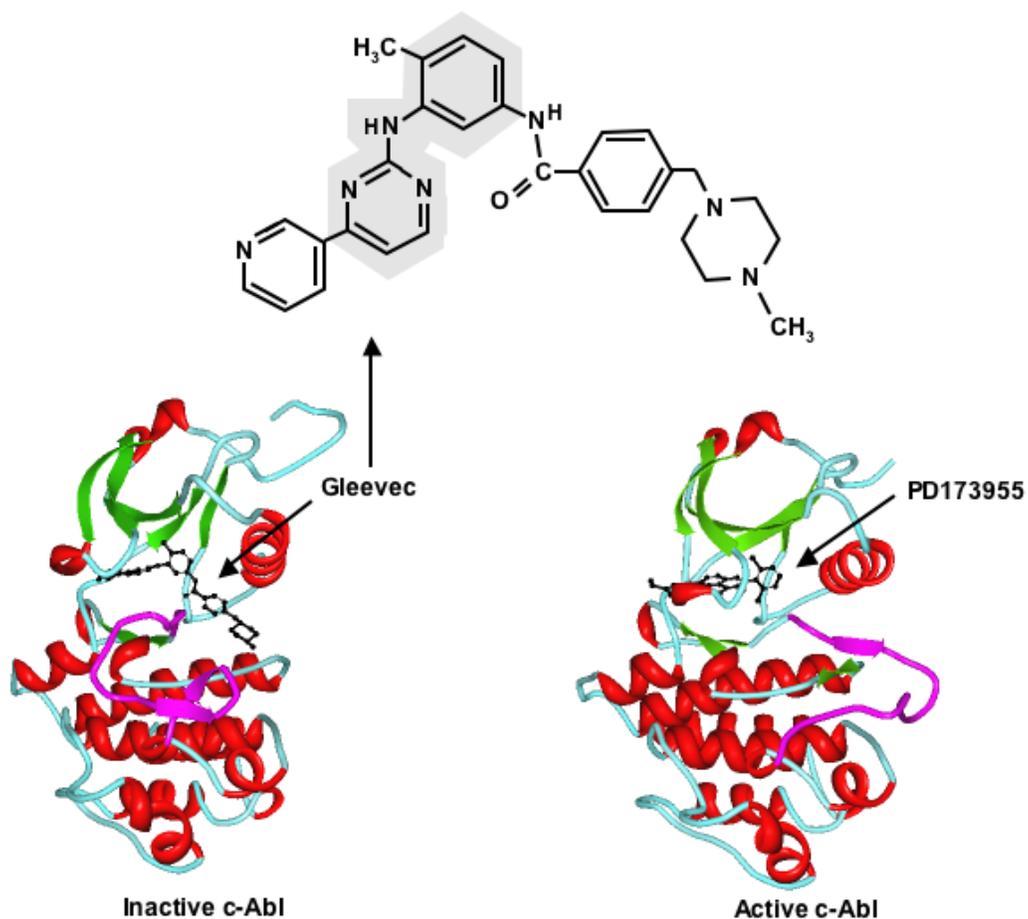
As briefly noted in Chapter 3, one source of somatic change associated with tumorigenesis is chromosomal translocation, where different chromosomal segments fuse aberrantly and either activate local gene expression or cause the expression of an abnormal fusion protein. The example we will look at further in this context is found in chronic myelogenous leukemia, and this translocation is also of historic importance as the 'Philadelphia chromosome'. This karyotypic abnormality results from a reciprocal translocation between specific sites on chromosomes 9 and 22, resulting in a 'minichromosome' product given the 'Philadelphia' title by virtue of its place of discovery in 1960 <sup>42</sup>.

This was the first such consistent abnormality described <sup>▼</sup>, and is present in the vast majority of cases of chronic myelogenous leukemia. The specific translocation which characterizes this type of leukemia results in the expression of a protein from the fused reading frames of the *bcr* and *c-abl* genes <sup>\*</sup>. This BCR-ABL gene product retains tyrosine kinase activity in common with the normal *c-abl* gene product, but in an unregulated (constitutive) fashion and at higher levels <sup>45</sup>. (The N-terminal BCR region of the fusion uncouples the ABL segment from its normal control mechanisms, but the latter is still responsible for the kinase activity of the whole aberrant protein). Subsequent studies showed that BCR-ABL was leukemogenic when expressed in mice, and its kinase activity correlated with its transformation ability. Such findings collectively validated the choice of BCR-ABL as a tumor target <sup>45,46</sup>. Most importantly, this target was absent in normal cells, and thus an ideal drug exclusively active on this fusion protein should in turn only act against the greater tumor target itself.

---

<sup>▼</sup>To put this into perspective, accurate counting of the human chromosome number itself was not reported until 1956 <sup>43</sup>, three years after the published structure of DNA.

<sup>\*</sup>The *bcr* gene ('breakpoint cluster region') was discovered and defined through the Philadelphia chromosome translocation itself, but the *c-abl* gene was discovered as the normal cellular homolog of a gene (*v-abl*) from the Abelson murine leukemia virus <sup>44</sup>.



**Fig. 9A14.2**

Top; Structure of gleevec (Imatinib), alternatively spelt as glivec. Gray shaded area superimposed on structural formula indicates 2-phenylaminopyrimidine core of original lead. The structure of gleevec bound to c-abl kinase is also shown as indicated, in contrast to the compound PD173955<sup>47</sup>. (Green segments indicate  $\beta$ -strands; red segments  $\alpha$ -helices, bound drugs black). While gleevec only binds when the kinase activation loop (purple segment) is in the inactive conformation, PD173955 can bind in the active conformation as shown. Source: [Protein Data Bank](#)<sup>8</sup>; [1IEP](#) (abl-gleevec) and [1M52](#) (abl-PD173955). Images generated with Protein Workshop<sup>48</sup>.

The targeting of BCR-ABL links us to the historical acquisition of kinase inhibitors in general. Empirical screens of natural products identified compounds with tyrosine kinase inhibitory activity, including quercetin<sup>49</sup> and genistein<sup>50</sup> (also well-known for their effects as phytoestrogens). One such compound, erbstatin<sup>51</sup>, served as a scaffold starting point for the synthesis and screening of functionally useful tyrosine kinase inhibitors<sup>52</sup>. But random chemical library screens were also instituted in attempts to gain kinase lead compounds, and one promising direction was indicated by the core structure of 2-phenylaminopyrimidine<sup>53</sup> (Fig. 9A14.2-A), from which strong inhibitors against the serine-threonine protein kinase C<sup>▼</sup> were obtained<sup>55</sup>. The same structural core proved also useful against specific receptor tyrosine kinases (surface receptors whose signaling is based on tyrosine kinase activity)<sup>56</sup>. Structure-function relationship studies revealed specific sites on the 2-phenylaminopyrimidine core which abolished inhibition towards protein kinase C and conferred inhibitory activity towards tyrosine kinases, an obviously highly valuable finding for obtaining desirable kinase selectivity<sup>46,53</sup>. After specifically optimizing the molecular design based on the ABL kinase target, gleevec was derived<sup>57</sup> (Fig. 9A14.2-A). Although this compound is not completely specific for BCR-ABL (affecting certain other tyrosine kinases), it proved to be very valuable as a therapeutic against chronic myelogenous leukemia bearing the Philadelphia chromosome and the *bcr-abl* translocation<sup>45,46</sup>.

Protein kinases in general use ATP as a cofactor for transferring a phosphate group onto the correct protein substrate(s), and kinases accordingly possess ATP binding sites. While a structural mimic of ATP might serve as a competitive drug inhibitor of kinases, it would from first principles need to be of high affinity to counter the high intracellular levels of the ubiquitous ATP molecule<sup>45,58</sup>. Although initial kinetic and modeling studies indicated that gleevec bound to the active site of the ABL kinase segment of the BCR-ABL fusion in a competitive fashion<sup>46,58</sup>,

---

▼ 'Protein kinase C' is actually an extensive multigene family of related serine-threonine kinases

structural analyses of bound ABL-drug complexes revealed a distinct interaction mode. Gleevec binds and stabilizes an inactive form of the kinase, where the 'activation loop' <sup>▼</sup> is a conformation incompatible with enzymatic activity, but certain other kinase inhibitors can also interact with the active state of the ABL enzyme <sup>47</sup> (Fig. 9A14.2-B). The inactive conformation provides an accessible pocket for gleevec binding adjacent to the ATP binding site, although part of the gleevec molecule extends into the adenine-binding region of the ATP site itself. Gleevec also indirectly competes for the ATP site by its stabilization of inactive form, which has an intrinsically lower affinity for ATP <sup>59</sup>).

The gleevec story illustrates how a steadily increasing knowledge base translates into increasing refinement of design and its shift towards a rational ideal. Experimentally-acquired structure-activity relationship data feeds into all subsequent design efforts in a positive feed-back loop. The identification of gleevec as an effective BCR-ABL kinase inhibitor bootstrapped upon the early empirical data which pointed to 2-phenylaminopyrimidine as a good lead scaffold. This phase of the drug development corresponded to a semirational ligand-based screen, where the initial lead drug is analogous to a known natural ligand which is used as the springboard for further synthesis. Systematic application of structure-function information allowed the move in chemical space from the initial phenylaminopyrimidine-based lead to gleevec through screening of many compounds (on the order of hundreds <sup>53</sup>), but a relatively small number compared to chemical libraries used for primary screening.

Further along the developmental line, the gleevec binding mode to the ABL kinase domain was not predicted in advance of structural information. Yet this knowledge allows further refinement of rational design to deliberately create

---

<sup>▼</sup> The activation loop is involved with the normal control of ABL kinase activity, mediated by its phosphorylation, and by other means <sup>45</sup>. In the active conformation, the loop is in an extended (or 'open' state) which correctly positions the catalytically active residues and forms a platform structure for substrate binding <sup>47</sup>.

inhibitors which bind to the inactive forms of kinases<sup>59</sup>. The latter 'Type II inhibitors' (as opposed to the Type I class which interact directly with the ATP-binding site of the active form) may offer the best opportunities for gaining kinase selectivity<sup>▼</sup>. Structure-based dissection of localized weak bonding interactions between Type II inhibitors and their targets offers the ongoing prospect for steadily increasingly rationality of kinase inhibitor design for selectivity and potency<sup>59</sup>.

The Darwinian aspect of tumor survival is rarely more starkly demonstrated than in their acquisition of resistance to drugs, and this has been the experience with gleevec. Any drug treatment provides a powerful selective pressure towards the amplification of variant tumor clones with improved survival. Since the introduction of gleevec into the clinic, a significant fraction of patients have been observed to become refractory to this otherwise-effective drug treatment, which can (as with other cancers) occur through a variety of mechanisms. These can include mutations which affect access of the drug (such as cellular efflux mechanisms), overexpression of BCR-ABL, or the acquisition by leukemic cells of alternate pathways for proliferation<sup>60</sup>. Nevertheless, most phenotypic resistance results from mutations within the kinase domain of BCR-ABL itself, and in consequence a second generation of drugs has been developed with the aim of countering this trend. One result of structural studies of ABL-gleevec complexes was the realization that gleevec binding, though selective, had a low tolerance for sequence alterations in its contacts with ABL. Structure-based modeling of the effects of such ABL mutations led to the design of new and more robust inhibitors approved for clinical use (such as nilotinib and dasatinib<sup>61</sup>). Unlike gleevec, dasatinib can bind the active form of the ABL kinase<sup>62</sup> (in a similar manner to the PD173955 complex in Fig. 9A14.2B). Yet there exists a

---

▼The pocket in the inactive kinase state which Type II inhibitors such as gleevec bind has greater diversity between different kinases than the ATP-binding site itself, thereby affording more opportunities for selectivity. Unfortunately, the Type II binding mode is apparently not available for some kinases<sup>59</sup>.

specific mutation of ABL (threonine-315 replaced with an isoleucine) against which even these second-generation drugs are ineffective<sup>62,63</sup>. This 'gatekeeper' position can be altered to isoleucine without affecting ATP binding, but such a change sterically interferes with drug binding to the hydrophobic pocket adjacent to the ATP site<sup>60</sup>. Inhibitors which could act outside the ATP site (that is, in a non-competitive manner with ATP) would potentially side-step this problem. For any kinase, non-competitive inhibitors can act through an allosteric mechanism (binding to a different site and causing a conformational change), or through binding to the substrate site rather than the ATP cofactor site<sup>58</sup>. A substrate-competitive BCR-ABL inhibitor has shown promise in overcoming the threonine-315 gatekeeper mutational resistance escape route for gleevec resistance. In this case, the lead compound was isolated from a focused library screen of derivatives of non-ATP-based kinase inhibitors<sup>64</sup>.

Another approach to drug resistance with general implications for protein ligand design comes from a feature of many natural protein folds involving the exposure of residues to aqueous solvent. Most internal hydrogen bonding associated with the protein backbone is protected from water, but in some cases this does not apply, and such 'underwrapped' hydrogen bonds are destabilized by competing water-based interactions<sup>65</sup>. Conversely, if water is excluded such hydrogen bonds are substantially stabilized, and this can be mediated through a protein-protein or protein-ligand binding event. Protein regions with such 'packaging defects' ('dehydrons', as stabilized by dehydration<sup>65,66</sup>) may be exploited naturally for the stabilization of specific protein interactions, but also serve as potential 'sticky' sites for artificial non-competitive drug targeting. This principle is rationally applicable to kinases<sup>66</sup>, and has been used for the redesign of gleevec for improved selectivity towards c-Kit, a receptor tyrosine kinase towards which gleevec is also active<sup>67</sup>. We should note also that the dehydron identification strategy has been used to rationalize drug activities against our previous example of the HIV protease, and for finding new 'drug epitopes' on other HIV targets<sup>68</sup>.

There is much more that could be said on kinase targets of other tumors and corresponding drug design, beyond our current scope. But the kinome of protein phosphorylation in general remains a huge drug target<sup>69</sup>, especially if we note that as well as kinases (transferring phosphates to substrates) we must include phosphatases which remove phosphate groups<sup>▼</sup>.

---

<sup>▼</sup> With respect to phosphatases, an example is given below ([Section A15](#)) in terms of the inhibition of the important regulatory phosphatase calcineurin by cyclosporin A and FKBP, in the form of ternary complexes with immunophilins.

## Section A15: *The Interactome and Biological Parsimony*

Relevant to p. 338 of *Searching for Molecular Solutions*, where the unfolding of genomic complexity is schematically depicted. Note that this also extends the genomic / chemogenomic presentation for Chapter 8 Cited Notes (file SMS–CitedNotes-Ch8/Section 22; from the same ftp site).

### *The Incredibly Intricate Interactome: Polyomic Problems and Promise*

One of the recurring themes of *Searching for Molecular Solutions* is the relevance of the Generation of Diversity in the field of molecular discovery, and great diversification occurs during the ‘horizontal’ somatic expression of genomes, as well as through the ‘vertical’ transmission of genomes through reproduction. It is clear that the genomes of multicellular organisms undergo expression during somatic growth in a highly complex manner, and some of the central processes involved in this densely-packed unfolding of genomic complexity are listed in Table 9.N1-A (Molecular Diversifiers ♡). All this vast intricacy arises from long linear strings of a polymeric molecule with an alphabet of four members, but this occurs in a highly entangled manner with respect to its own encoded products (Fig. 9.7 of *Searching for Molecular Solutions*). All the products of the various ‘omics participate in complex communications as part of the interactome, but it is more accurate to speak of interactome in the plural with respect to multicellular organisms, or even unicellular organisms which exist in distinct differentiated states \*. This is so because of the diversity which the unfolding of genomic information confers upon different cell lineages. Each of

---

♡ Within file SMS–CitedNotes-Ch9/Section 30; from the same ftp site.

\* An example of this the establishment of spore formation (sporulation) in some (Gram-positive) bacteria, which involves extensive cell-state specific changes in gene expression<sup>70,71</sup>.

these will clearly share common subsystems, but distinct differentiated cell types will have distinct global expression patterns and associated global interactomes.

We speak of the 'human genome', but human and almost all other eukaryotic cells are in reality an interactome between two separate genomes, nuclear chromosomes and mitochondria. Although the human mitochondrial genome is tiny compared to the 'main' genome (approximately 16.5 vs. 3.3 million kilobases for the mitochondrial and human haploid genomes, respectively), mitochondria are present in about 1000 copies per normal human cells <sup>72</sup>, and these organelles are fundamentally important for cellular energy transactions through the synthesis of adenosine triphosphate (ATP). As noted in Chapter 3 of *Searching for Molecular Solutions*, it is generally accepted that mitochondria derive from formerly free-living prokaryotic 'endosymbionts' whose genomes have undergone extensive simplification over evolutionary time. But this ancient association with prokaryotes at the cellular level is not the end of the story concerning interactions between multicellular organisms and bacteria. Humans and other mammals carry extensive populations of bacteria within their digestive tracts which are very significant for higher-level functioning, including the production of vitamins and other cofactors, and immune system conditioning <sup>73-76</sup>. This has led to the consideration of mammals as 'superorganisms' composed of their genomically-encoded 'parts' and their accompanying prokaryotic symbiotes <sup>77</sup> ♡. As such, the composite genome at the superorganismal level has been termed the 'hologenome' <sup>79</sup>, which is not appropriate when applied as a unitary entity at the level of evolutionary selection \*. Yet it is unquestionable that there are mutual interactions between the proteomes and metabolomes of humans (and other multicellular animals) and their prokaryotic symbionts, which might be

---

♡The superorganism concept has also been applied to social insects such as ants and bees <sup>78</sup>.

\* Considering superorganisms (and their associated 'hologenomes') as units of evolutionary selection is contentious, as is the related area of group selection, which has had a turbulent history in modern evolutionary theory <sup>80-83</sup>. Certainly symbiotic organisms co-evolve, but not as a single selectable unit.

considered as a higher-level interactome of sorts. But since the gut is technically still an external environment, relaxation of the definition of interactome to this level would then conjure up 'interactomes' of animals and the plants upon which they depend, and sweep in all other ecological relationships. This becomes a matter of definition, but for the present purposes let us restrict 'interactome' to the inter-relationships between the 'omes within cells (as in Fig. 9.7 of *Searching for Molecular Solutions*).

Diversification of interactomes equates with the long-recognized phenomenon of cell differentiation, and changes in cell-specific interactomes results from diversification in the functional deployment of each of the 'omic progeny encoded by the genome<sup>♥</sup>. The specific temporal and spatial unfolding of this 'polyomic GOD' constitutes the 'phenogenetic logic'<sup>84</sup> by which complex multicellular systems develop. In this context we should note that the power of immunological GOD for the generation of specific binding molecules (a major theme of Chapter 3 of *Searching for Molecular Solutions*) is but a subset of a higher-level polymic GOD. But always it must be stressed that despite parallels which can be drawn between somatic diversification mechanisms and evolutionary processes, somatic and germline diversity are utterly distinct from the viewpoint of natural evolution<sup>84</sup>. As also noted in Chapter 3, natural genetic diversification and selection is demonstrably capable of giving rise a somatic system (the adaptive immune system) which mimic its own attributes.

The diversification of an organism's phenogenetic logic by mutation is evolutionarily significant. But we have also noted other facets of this 'logic' previously as well, and these include combinatorics and modularity. The latter are an inherent aspect of the parsimonious nature of evolutionary design, and the relevance of this to drug discovery prompts us to move a little further down this track.

---

<sup>♥</sup> See also the file SMS–CitedNotes–Ch8/Section 22 (At Home with 'Omics); from the same ftp site.

### *Biological thrift and drug targeting*

Consider the proposition, 'Evolutionary selective pressures favor the acquisition of efficiency in the architecture of complex natural biomolecular systems'. The accuracy of this is highly dependent on one's notion of 'efficiency' in this context. One indeed treads on very dangerous ground to invoke absolute efficiency as an inevitable consequence of evolution<sup>85</sup>. This theme we have come across previously, in the form of natural molecular systems which are believed to be 'locked in' and unable to traverse deep fitness valleys to alternative (and more efficient) states<sup>♥</sup> Yet where efficiency is a correlate of biological fitness, it should increase in natural systems subject to their freedom to move in the applicable terrain of fitness landscapes. A successful organism may then show a relative increase in system efficiency over a reproductively inferior competitor, even if it is blocked from attaining a theoretical efficiency maximum. Modularity and combinatorics are strategic pathways by which such improvements are attained, and in essence deliver the message that it is advantageous to parsimoniously build multiple things out of a limited set of parts than to start from scratch each time. The biological deployment of a relatively limited set of materials from an evolutionary molecular 'toolkit' we might term 'biological thrift'.

One aspect of this kind of thrift applies at the level of small biomolecules. In Fig. 9A15.1A two sets of biomolecules are shown where there are easily discerned similarities within each. This is no coincidence, since they are inter-related through metabolic pathways converting amino acid precursors into other biological mediators. Common precursors thus give rise to small molecule mediators with very diverse functions. On the one hand, this might seem like a trite, mundane observation for which data extends back over a century<sup>\*</sup>.

---

<sup>♥</sup> A putative example of this with the enzyme Rubisco is considered in the file SMS–Extras-Ch5/Section A6; from the same ftp site.

<sup>\*</sup>Adrenaline, for example, was described at the turn of the 20<sup>th</sup> century, and synthesized in 1904 ([Link to further information](#)).

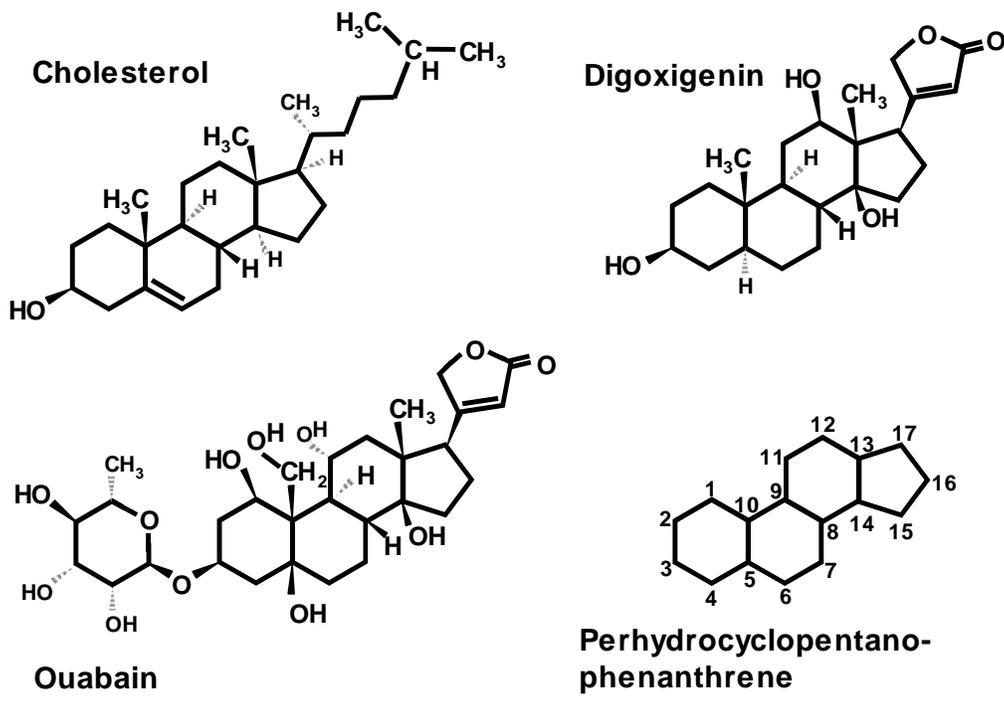
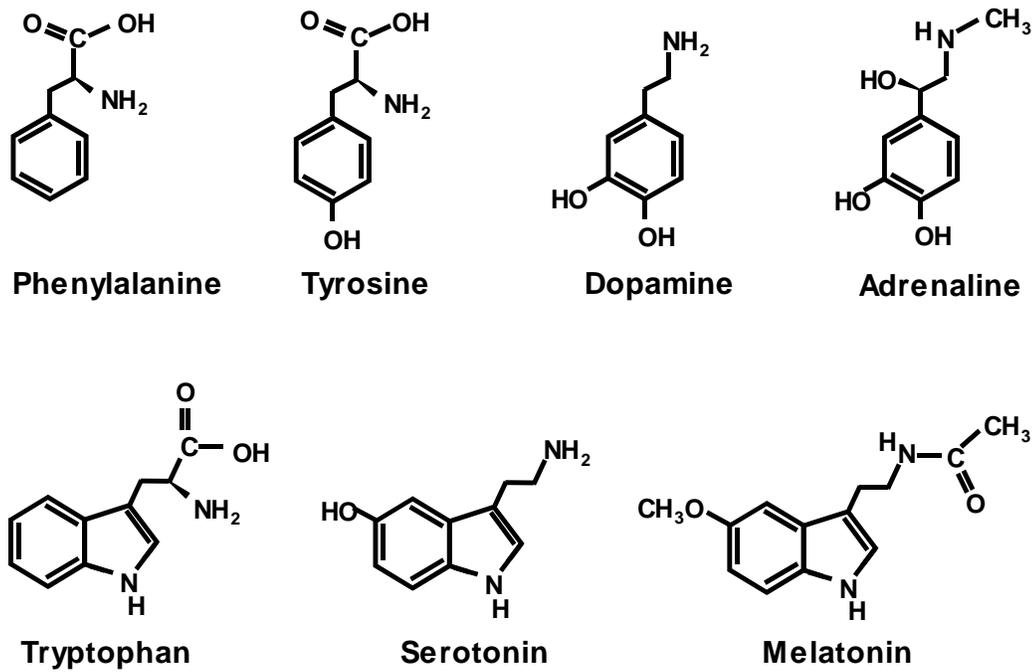


Fig. 9A15.1

**Fig. 9A15.1.** Sets of natural small molecules of diverse functions and related structures. **A**, Mediators derived from amino acids: Dopamine (neurotransmitter) and adrenaline (epinephrine; hormone / neurotransmitter) are derived from phenylalanine and tyrosine; serotonin (neurotransmitter / hormone) and melatonin (neurohormone) are derived from tryptophan. (Another example is the pleiotropic mediator histamine, derived from the amino acid histidine (noted in Fig. 9.2 of *Searching for Molecular Solutions*) **B**, Some natural small molecules with the steroid perhydrocyclopentanophenanthrene skeleton. Cholesterol is structural membrane component, digoxigenin and ouabain are cardioactive compounds from plant sources. (Digoxigenin as an aglycone (sugarless) derivative of the cardiac glycoside digoxin, both from the foxglove plant [*Digitalis lanata*]).

---

But at the same time, one can easily imagine biosystems where in each case mediators performing the same range of functions as in Fig. 9A15.1 were quite distinct and unrelated molecules. Of course, the great complexity of biosystems is such that it is not trivial to dissect chance from necessity in the factors driving the evolution of mediators and their receptors. For example, both serotonin and melatonin are active in the central nervous system, but with very distinct functions. The amino acid tryptophan is a precursor of serotonin<sup>86</sup>, and serotonin in turn is the precursor of melatonin<sup>87</sup>. Was it inevitable that melatonin, its biosynthetic enzymes, and its receptors should have evolved as mediators of mammalian circadian rhythms, or could some other small molecule have stepped into this role? Is melatonin the most economical solution to this requirement, or is it merely another example of the chance evolutionary fixation of one option over other alternatives? Moreover, the question does not end here, as many single small molecule mediators have multiple receptors and diverse roles themselves. In this regard, it is accurate to refer to adrenaline and serotonin as both hormones and neurotransmitters<sup>▼</sup>.

---

▼ Adrenaline (epinephrine) has nine receptors of the GPCR class<sup>88</sup>, and serotonin has thirteen GPCR receptors and one receptor of a totally different type (a ligand-gated ion channel)<sup>89,90</sup>.

Certainly the enzymes which shape small molecules have deep evolutionary roots. For example, plants use steroid-like hormones in common with animals. an important class of plant hormones possess the steroid ring system or a very close analog of it. These 'brassinosteroids' are synthesized by enzymes which have been shown to be homologous between plants and animals, pointing to a common (albeit distant) evolutionary origin between plant and animal steroid hormones<sup>91,92</sup>. Certain plant products with the characteristic steroid ring system are useful drugs (Fig. 9A15.1B).

Although in itself this does not explain existing patterns of small biomolecular function, it is a logical proposition that it is evolutionarily easier to evolve enzymes making variant derivatives of a small molecular framework than to evolve the catalytic underpinnings for an entirely different framework. An observation consistent with this is that derivatives of biological precursor molecules are not arbitrarily diversified, but are subject to constraints imposed through the acquisition of altered synthetic enzyme specificities. Certain molecular sites accordingly tend to be favored for modification over others, and one example of this is the diversification of molecules with the steroidal ring system, where the 17-position is a prominent site for varied substitutions (Fig. 9A15.1B).

Many of the mechanisms of protein evolution in general (especially gene duplication and divergence, noted in Chapter 2 of *Searching for Molecular Solutions*) and enzyme evolution in particular (including catalytic promiscuity, Chapters 2 and 5) are likely to constrain the biological use of small molecules towards 'thriftness'. This parsimony is an inevitable consequence of efficiencies forced by natural selection (in the guarded sense noted above), with the synthetic machinery for pre-existing biological mediators as the raw material for diversification. Evolution of entirely new genes may be selectively advantageous only when no other 'parsimonious' evolutionary pathways are available.

Dobzhansky's famous aphorism, 'Nothing in biology makes sense except in the light of evolution'<sup>93</sup> is never more compelling than in this issue of biological thrift.

As noted earlier, a thrifty approach to building a complex system cannot rely on constructing all subcomponents from scratch, but must use a limited 'parts list' combined as reassemblable modules. Though highlighted by results from whole-genome sequencing projects<sup>94</sup>, the modularity principle has been underscored through direct physical analyses of protein interactomes<sup>95</sup>. Many examples of modularity at the protein level can be invoked, including basic processes such as transcription (relevant factors used in different combinations in different contexts) and signal transduction (signal pathway members varying in different cellular differentiation backgrounds). Another case in point is recombination, and here we could recall the process involved in immunoglobulin and T cell receptor somatic gene rearrangements (Chapter 3 of *Searching for Molecular Solutions*), where specific recombinases<sup>▼</sup> co-operate with 'generalist' factors to produce the final recombined structures. This immunological example can be taken still further to consider modularity at higher systems levels, as when one compares similarities between the immune and nervous systems<sup>\*</sup>.

A modular arrangement of the components of an interactome inherently means that modularity and connectivity (which we visited in Chapter 2) are themselves connected concepts. As we have seen, not all gene products are equally involved in cellular networks, and some key elements lie at interconnection 'nodes'. One consequence of this is that some heritable diseases of monogenic origin can have devastating effects in multiple systems, with complex phenotypes. Crippling a gene product with high connectivity is likely to have wider ramifications than a less well-connected counterpart. An accurate analysis of gene product networks and their the connectivity should have predictive power

---

▼ These RAG recombinases are considered in more detail in the file SMS–Extras–Ch3/Section A1; from the same ftp site.

\* See the file SMS–Extras–Ch3/Section A4; from the same ftp site.

for target identification, a rational design-related topic considered in Chapter 9 of *Searching for Molecular Solutions*. But for the present purposes, this aspect of modularity brings us back to the relevance of biological thrift for drug activity.

### *The Dark Side of Thrift – Drug Cross-Reactivity*

The implications of evolution on drug targeting are significant for all members of any species. For the development of safe and effective drugs, the parsimonious nature of evolutionary processes means that an identified target, or a target pathway represented by a specific protein, is never an isolated entity. And this inevitably raises the problem of unwanted side-effects in otherwise useful drug molecules. An 'off-target' effect can arise from clear-cut similarities in drug binding sites of protein with evolutionary family relationships (once relevant phylogenetic, genomic and structural information is available), or by what at least superficially appears to be a random chance cross-match between a drug and a site on an unrelated protein.

Artificial xenobiotics may give the best examples of chance multi-target effects, as with the example of the well-known insecticide p, p'-dichlorodiphenyl-trichloroethane (DDT). This compound was identified as an insecticide through empirical screening, and subsequently shown to be an insect neurotoxin by binding to the voltage-gated sodium channel in insect neurons<sup>▼</sup>. An important side-effect of many chlorinated hydrocarbon insecticides (including DDT) was revealed to be estrogenic activity, or mimicking of estrogens<sup>97-100</sup>. DDT, for example, can directly bind and activate the estrogen receptor (albeit much more weakly than normal hormone<sup>101,102</sup>), and DDT can even support the growth of estrogen-dependent tumor cells in culture<sup>103</sup>. Now, the 'selection' for artificial insecticides was empirical searching for insect-killing abilities, and certainly not

---

▼ It appears that only a three amino acid residue difference in the human vs, insect sodium channel is the determinant of the differential toxicity of DDT to insects<sup>96</sup>.

as 'xenoestrogens'. It follows in turn that the interaction of such artificial compounds with estrogen receptors is an undirected chance event. (Even so, the probability for such events may be increased through the relatively limited number of protein folds in biological protein sequence space, as noted in Chapter 2 of *Searching for Molecular Solutions*).

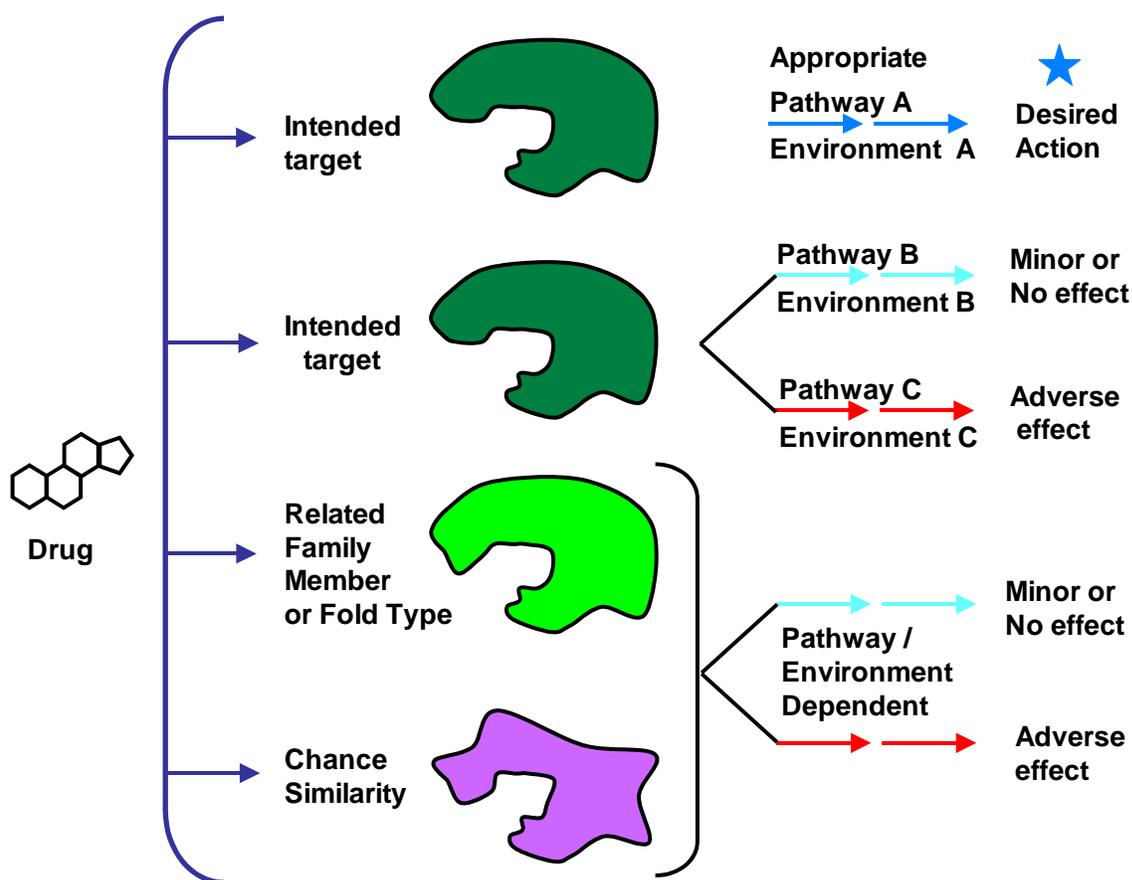


Fig. 9A15.2

Types of unwanted drug effects. The intended target and the desired drug-induced phenotype define the desired drug effect. But in a different intracellular context (variant differentiation or activation state; 'pathway' B, C) or extracellular environment,

modulation of the same target may produce different effects, sometimes adversely. A protein in the same family as the target, or one with a related fold may bind and become modulated by the drug. Rarely, a unrelated fold may reproduce a drug-binding site by chance. In either of these cases, the outcome may range from minor to serious, subject to cell context ('pathway') and environment-specific factors.

---

Yet the pharmacological effect of such interactions is also determined by the specific cellular context in which the 'off-target' binding occurs. Again owing to evolutionary thrift, biological protein interaction networks (and interactomes in general) will have different cross-connections in specific cellular environments resulting from differentiation in a multicellular organism. And this can apply even to the intended target itself, if the drug binds the 'right' biomolecule but in the 'wrong' time and place, with potentially serious consequences. These scenarios are depicted schematically in Fig. 9A15.2 above.

In order for the cellular components of complex organisms to differentiate and grow in a coordinated manner, they must receive, and in turn transmit, a constant flow of information encoded in molecular signaling pathways. Many drug targets are themselves embedded within such complex signaling networks, whose precise switching on and off is often a key factor determining correct developmental patterns. The modular and 'thrifty' nature of these networks means that signaling components during an early phase of development may be re-deployed during later stages of the life-cycle of a multicellular organism. They may also potentially corrupt, erupt and contribute to pathological states during any stage of growth. And therein lies a drug targeting problem which falls into the 'right target / wrong setting' conundrum introduced above. A case in point is the

*Sonic Hedgehog* signaling pathway <sup>♥</sup>, of prime importance in mammalian development, but also clearly associated with certain tumors, especially basal cell skin cancers and a specific predominantly pediatric brain tumor (medulloblastoma) <sup>105,109</sup>. Small molecule inhibitors of the Hedgehog pathway (obtained both from natural sources or via high-throughput screening of chemical libraries) have been very promising in cellular and animal models of medulloblastoma <sup>109-111</sup>. Unfortunately, although adult mice appeared unharmed by one of these artificial inhibitors, immature mice showed bone malformations upon exposure to it <sup>112</sup>. If such problems were also manifested in humans, clearly treatment options with such compounds for childhood medullablastoma become limited, although not necessarily abandoned <sup>\*</sup>.

In the area of developmental disruption (or teratogenesis), a much more famous example of drug adverse effects is the history of thalidomide. More accurately, we should say 'infamous', since the story is well-known: the use of thalidomide as a sedative in pregnant women led to thousands of birth defects in the late 1950s and early 1960s. Here the situation is more complex, since thalidomide is

---

<sup>♥</sup>The Hedgehog pathway was discovered as developmentally crucial in the fruitfly *Drosophila*, and so named owing to mutations in the Hedgehog gene producing a curled-up, 'prickly' appearance in larvae <sup>104,105</sup>. In vertebrates three Hedgehog homologs exist, termed Indian, Desert, and Sonic, which bind the same receptor (the Patched gene product) but with distinct activities <sup>106</sup>. The Sonic variety may sound auditory-related, but in fact was named after the videogame / cartoon character, a bit of humor which has not met with universal approbation <sup>107</sup>. Certain whimsical gene designations patterned after well-known trade names have attracted threats of legal action <sup>108</sup>, but *Sonic Hedgehog* appears to be off the hook. But there are unlikely to be many more Sardonic Hedgehogs in future.

<sup>\*</sup> Compartmentalizing drugs into the desired target site and away from areas associated with bone development might be possible. Even where the same targets are bound at the same sites by distinct inhibitors, the drugs may differ in their pharmacological distribution *in vivo* <sup>112</sup>. From the point of view of the developmental effects of inhibiting Hedgehog signaling, it is interesting to note that a natural Hedgehog pathway inhibitor from a plant source, cyclopamine, is a teratogen in sheep <sup>113</sup>, and engineered mutations of *Sonic Hedgehog* in mice or corresponding natural mutations in humans have similar phenotypes to cyclopamine effects <sup>114,115</sup>.

likely to interact with multiple protein targets, not all of which are involved with signaling pathways in early growth and development<sup>116</sup>. Indeed, some of the 'other activities' of thalidomide are beneficial, as we will see shortly in the following [subsection](#). One such potential area of therapeutic benefit is in the thalidomide-mediated inhibition of TNF- $\alpha$ <sup>116,117</sup>, and thalidomide may join the growing list of alternative means for attaining this pharmacological end, some of which were considered in Chapter 7 of *Searching for Molecular Solutions*. It is a little ironic in this regard that some of the anti-TNF- $\alpha$  monoclonal antibody products currently used for treatment of rheumatoid arthritis have been themselves accused of increasing susceptibilities to infectious diseases and malignancies<sup>118</sup>, although the extent of such risks is controversial<sup>119,120</sup>. But very few drugs of any sort can truly be magic bullets, and there is in fact a gray area between 'minimal side-effects' and 'adverse effects' (Fig. 9A15.2). Accordingly, in the real world it is often a matter of deciding on which side of the line a risk / benefit analysis falls.

Given the difficulties and expenses arising from drug withdrawals due to unforeseen adverse effects<sup>121</sup>, it is not surprising that concerted efforts have been directed at developing assays to heading off such problems at an early stage. These initiatives have been referred to as 'fail-fast' strategies<sup>122</sup>. (It is greatly preferable for an ultimately doomed pharmaceutical venture to cease before massive investment in large-scale clinical trials is undertaken). Testing in animal models (even with primates) is not always a good predictor of responses in humans, as seen in 2006 with the well-reported disastrous preliminary trial of an anti-CD28 superagonist antibody<sup>123,124</sup>. Effective *in vitro* monitoring of drug cross-reactivities will inevitably have to address the roles of therapeutic compounds in perturbing complex cellular signaling networks, in order to reveal 'hidden phenotypes' produced by unexpected drug activities. Intracellular detection of changes in protein-protein interactions logically offer a coherent

answer to this problem, and the protein complementation assay <sup>♥</sup> has emerged as a powerful approach towards this end. Use of this assay with sizable numbers of interactive proteins allows high-throughput assays for drug effects on cellular signaling networks <sup>122,125</sup>. A significant observation from such studies has been that structurally unrelated drugs which cluster in their effects on protein-protein interaction networks also tend to share phenotypic properties <sup>125</sup>, a feature of predictive value for the assessment of novel drugs as they emerge from the developmental pipeline.

But 'hidden' information gleaned from protein complementation assays (or by any other means) need not necessarily always prove to be bad news. When one door closes on a drug targeting opportunity, another door may (sometimes) open. The reasons for this are also intimately connected with the parsimony of biosystems wrought by evolution.

*The Bright Side of Thrift – Target Diversity, Drug Repositioning, and More*

Sometimes binding of a drug by a family of targets (rather than absolute monospecificity) is not detrimental, and may offer positive advantages. If the targets are produced by foreign organisms and cross-reactive through shared evolutionary antecedents, this is not so surprising. A broad-spectrum antibiotic active against a wide range of pathogenic bacteria is clearly more generally applicable than a drug which is very limited in its antibacterial selectivity <sup>\*</sup>. But drugs against human targets can also be useful with a broader basis of specificity, provided the target range is relevant to the disease process in

---

<sup>♥</sup> This assay is described in more detail in the file –CitedNotes-Ch4/Section 9; from the same ftp site.

<sup>\*</sup> A caveat should be noted here. Very broad-spectrum antibiotics may be more disruptive of normal gut bacteria (the non-human part of the so-called 'superorganism' mentioned earlier), and this can sometimes cause significant problems, as with severe diarrhea associated with antibiotic-associated overgrowth of the anaerobic bacterium *Clostridium difficile* <sup>126</sup>.

question, and adverse effects are not excessive. Consider an example in the field of oncology. The inhibitor sorafenib (an approved drug for renal cancer) was developed through optimization of a lead compound identified from a high-throughput screening for inhibitors of the cellular C-Raf kinase, an important signal transduction mediator <sup>127</sup>. Subsequently, it was found that sorafenib had inhibitory activity towards other kinases, and in particular certain receptor tyrosine kinases involved with transmitting angiogenic signals (promoting blood vessel growth which favors solid tumor proliferation) <sup>127</sup>. This cross-specificity was certainly no accident, since mammalian kinases share common evolutionary origins, although they have diversified into multiple subclasses (the kinome, as noted above) <sup>128</sup>. The salient feature which enables their genomic enumeration is the conserved eukaryotic protein kinase catalytic domain, and kinases with substantial sequence identity are most likely to be inhibited by the same groups of compounds <sup>129</sup>. As far as sorafenib is concerned, its multi-kinase targeting appears to be therapeutically useful <sup>127</sup>, especially where cancer (as a life-threatening disease) can entail acceptance of increased risk with commensurate clinical benefit. This 'polypharmacology' exhibited by sorafenib is by no means unique among therapeutic drugs <sup>130</sup>, and the efficacy of some drugs is correlated with their relatively 'promiscuous' targeting <sup>131</sup>. The global analysis of pharmacological structure-activity data is expected to assist future rational predictions of such drug multi-targeting <sup>132</sup>.

If a drug modulates more than one biochemical pathway, perhaps failure in one arena need not necessarily rule out the potential utility of the drug in general. Or even if a drug is already successful in one application, perhaps it can also be beneficial in a different role. These statements constitute the essential basis of *drug repositioning*, which seeks to re-employ old drugs in fruitful new roles. The high costs of bringing entirely new drugs to market (referred in Chapter 8 of *Searching for Molecular Solutions*) renders repositioning (or 'repurposing') a drug a very cost-effective option if it is possible. And the fact that is indeed often feasible is rarely based on chance, but rather another facet of biological thrift,

either at the level of multi-target evolutionary similarities or through the biological deployment of the same target in multiple pathways (as in Fig. 9A15.2). Drug repositioning has accounted for almost half of 'new' therapeutics launched in recent times<sup>133</sup>. Among the many examples which could be cited in the repositioning field in general, one of the more dramatic is the above-mentioned case of thalidomide, which has progressed from a demonized teratogen to being perceived as a useful anti-inflammatory and immunomodulatory drug.

Thalidomide appears to interact with multiple targets with varied biological effects<sup>116,117,134</sup> and is currently licensed by the FDA for treatment of a complication of leprosy<sup>135</sup>. Its anti-TNF- $\alpha$  / anti-inflammatory activity has led to testing of numerous chemical analogs for improved safety<sup>134,136</sup>.

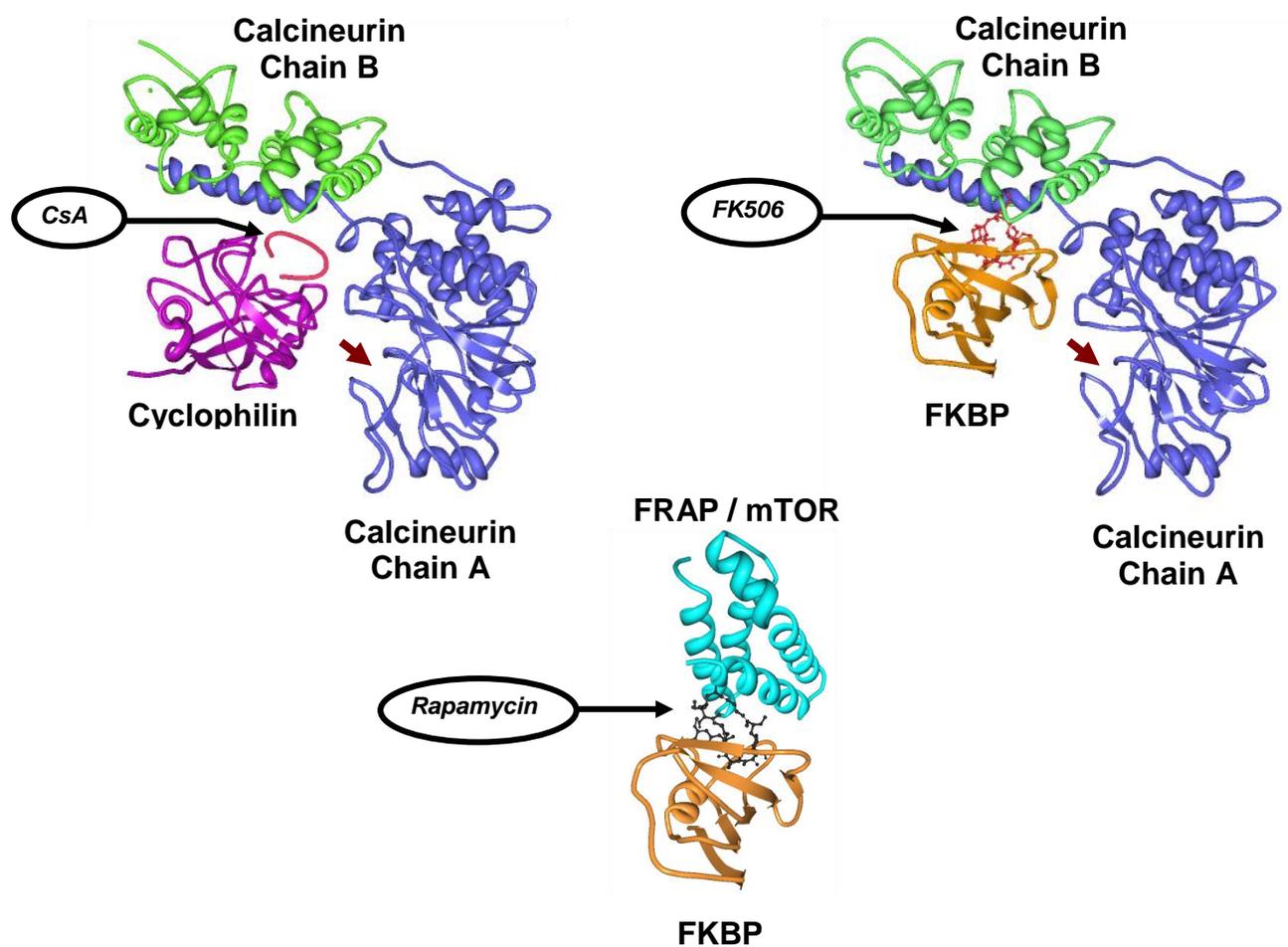
Many other examples can be proffered. Cimetidine (referred to in Chapter 9 as a paradigm for serendipitous discovery) has been found to have certain anti-tumor activity<sup>137</sup>. The ubiquitous analgesic acetaminophen has been attributed with cardioprotective effects, at least under some conditions<sup>138,139</sup>. Sildenafil (viagra) is of interest, since its original 'repositioning' from an anti-anginal drug to an effective therapy for erectile dysfunction was based on serendipitous observations. And the uses of this drug now include the treatment of pulmonary hypertension, and possibly other conditions<sup>140</sup>. The original saga of viagra should not imply that repositioning need remain a haphazard and random affair, though, as it can be approached and studied in a systematic manner. Data from high-throughput pathway analyses (as with the protein complementation assay referred to above) and collation of existing structure-activity information can allow useful predictions to be made. Computational virtual screening can be used in an 'inverse docking' mode (noted in Chapter 9), useful for identification of new drug-target interactions.

As with the development of thalidomide analogs, drug repositioning may require re-optimization of a drug in a new direction based on existing knowledge. The side-activities initially observed in a drug's conventional therapeutic mode can

serve as the springboard for redirection of activity such that the off-target effect becomes the main game, and the original activity is minimized or abrogated. Since by definition one begins this process with a drug molecule, it has been proposed that this 'Selective Optimization of Side Activities' strategy is more efficient than conventional high-throughput screening, and is broadly applicable in drug discovery <sup>141</sup>.

Finally, perhaps the most important positive aspect of biological thrift, and one that is obviously evolutionarily-related, confronts us in the form of virtually the entire panoply of natural product drugs which serve us. Much of the time, the rationale is clear from the existence of targets with common evolutionary underpinnings between the source organism and humans, but sometimes the link takes a little more investigation. Consider a specific set of examples, in the form of two distinct natural product immunosuppressants: cyclosporin A and FK506 (from fungal and bacterial sources, respectively), which have had a major impact on organ transplantation through their effects on T cell activation. These compounds bind distinct cellular proteins, cyclophilin and FK506-binding protein (FKBP), but they share a common target in the form of the protein phosphatase calcineurin, a calcium-responsive regulator of multiple signaling pathways <sup>142,143</sup>. The cyclophilin-cyclosporin A and FK506-FKBP complexes form ternary complexes with calcineurin (Fig. 9A15.3 below) and inhibit its role in T cell activation via transcription factors of the NFAT family <sup>143</sup>. A third immunomodulatory natural product, rapamycin, binds the same FKBP as the FK506 molecule, but forms a ternary complex with the unrelated protein FRAP / mTOR with distinct signal pathway roles <sup>142</sup>.

---



**Fig. 9A15.3**

Comparisons between ternary structures of cyclosporin A (CsA) – cyclophilin – calcineurin<sup>144</sup>, FK506-FKBP-calcineurin<sup>145</sup>, and rapamycin-FKBP-FRAP/mTOR<sup>146</sup>. Sources: [Protein Data Bank](#).<sup>8</sup> [1MF8](#), [1TCO](#), and [1FAP](#) respectively. Calcineurin itself consists of a regulatory domain B (green) and a catalytic domain A (blue; catalytic cleft shown with diagonal arrows). Images generated with Protein Workshop<sup>48</sup>.

A very large body of work has gone into the study of these natural products and related molecules, and the remarkable three-way interaction which they exhibit in order to manifest their biological activities has acted as the inspiration for the 'surface borrowing' strategy for enhancement of drug recognition, as further discussed below. But for our immediate purposes, let's focus on the relevance of such natural immunosuppressants to evolutionary thrift. Since fungi and certainly prokaryotes are hardly exemplars of adaptive immune systems, why should any of their metabolic products affect mammalian immune functioning in such a highly specific manner? The answer is reasonably clear. Both the 'immunophilins' (cyclophilin and FKBP) and calcineurin are highly conserved, with homologs clearly identifiable across the domain of eukaryotes <sup>147-149</sup>. Although the yeast homologs of cyclophilin and FKBP are not essential for yeast viability <sup>148</sup>, calcineurin confers survival benefits under some conditions <sup>150</sup>, suggesting that organisms able to block the calcineurin pathway in competitors would gain a selective advantage <sup>151,152</sup>.

Similar points can be made for the rapamycin target FRAP / mTOR. In other words, the existence of the natural immunosuppressants is attributable to cogent evolutionary factors, and their effect on mammalian immune systems is consistent with the same principles. This agency is squarely due to the evolutionary parsimony which has conserved key signaling pathways and constituents over hundreds of megayears, with selection diversifying their roles into multiple areas. In this regard, we should also note that cyclosporin A and FK506 served as both leads and drugs, meaning that their functional properties were not substantially improved through a wide range of artificial derivatizations <sup>142,144</sup>. This observation makes sense when interpreted in the light of a molecular ligand-target interaction 'tuned' by evolution (especially given its tripartite nature

♥ [Fig. 9A15.3], which would certainly would not be expected from a chance binding event).

There are some additional features of the natural immunosuppressants relevant to this section. From a therapeutic point of view, the great specificity of cyclosporin A and FK506 for calcineurin is a major bonus, since agents which cross-reacted significantly with other cellular protein phosphatases would have a high probability of unacceptable toxicity. The reason for this specificity lies in the fact that the active site of calcineurin is not involved with complex formation (Fig. 9A15.3). An inhibitor binding directly to the evolutionarily-conserved active site would be much more likely to show cross-activity on other phosphatases <sup>153</sup>, along the lines of the ‘family member’ interactions depicted in Fig. 9A15.2. And yet this high specificity in itself cannot avoid the fact that calcineurin is used in a variety of physiological circumstances, including neural tissues. In fact, as its name would suggest, calcineurin was first isolated from neural sources <sup>154</sup>, and is expressed in much higher levels in the nervous system than T cells, despite its importance for regulation of the latter. But it is this expression differential which renders the natural immunosuppressants clinically useful for organ transplantation, in that immunoinhibitory effects can be obtained in doses below the threshold of unacceptable side-effects in other systems \*. Yet the same widespread roles and expression levels of calcineurin mean that the natural immunosuppressants cannot be ‘repositioned’ for the treatment of autoimmune diseases. Much general information regarding cell signaling pathways has been

---

♥ This might beg the question as to why such a complex arrangement evolved in the first place, rather than molecules directly binding and inhibiting calcineurin. The ‘surface borrowing’ action of increasing the available contact surface for a small molecule may be one factor, along with the general utility of the immunophilin homologs as presenting proteins. The latter are peptidyl prolyl isomerases, and while (ironically) this is unrelated to their immunosuppression, they may be ‘designed’ to interact widely with a broad range of protein substrates, and therefore relatively easy to co-opt as presenters for foreign small molecules <sup>153</sup>.

\* These drugs still require close monitoring for adverse effects, though, and prolonged immunosuppression through their use is associated with increased levels of certain cancers <sup>155</sup>.

obtained from the saga of the natural immunosuppressive drugs, and this too is potentially useful in a therapeutic context, for exploiting the parsimonious sharing and overlap of such pathways. When a pathway known to be modulated by a specific drug turns up in another context, the old drug may suddenly gain a new application. For example, in the genetic disease tuberous sclerosis complex, one disease phenotype in neural tissue is associated with hyperactive FRAP/mTOR signaling. This proved to be reversible in mouse models by rapamycin treatment

<sup>156</sup>

An interesting strategy to improve the affinity of drug binding at protein surfaces has been directly inspired by the natural precedents of immunosuppressive ternary protein-drug complexes. In Chapter 8 of *Searching for Molecular Solutions* it was noted that drugging of protein-protein interaction surfaces has historically been a difficult challenge, although not without recent successes in specific cases. Here the generalizable 'take home message' from rapamycin and functionally analogous compounds (as in Fig. 9A15.3) is enhancement of a small molecule interaction by 'borrowing' protein-protein contacts to increase the overall binding affinity <sup>157</sup>. This surface borrowing effect is depicted in Fig. 9A15.4 below.

---

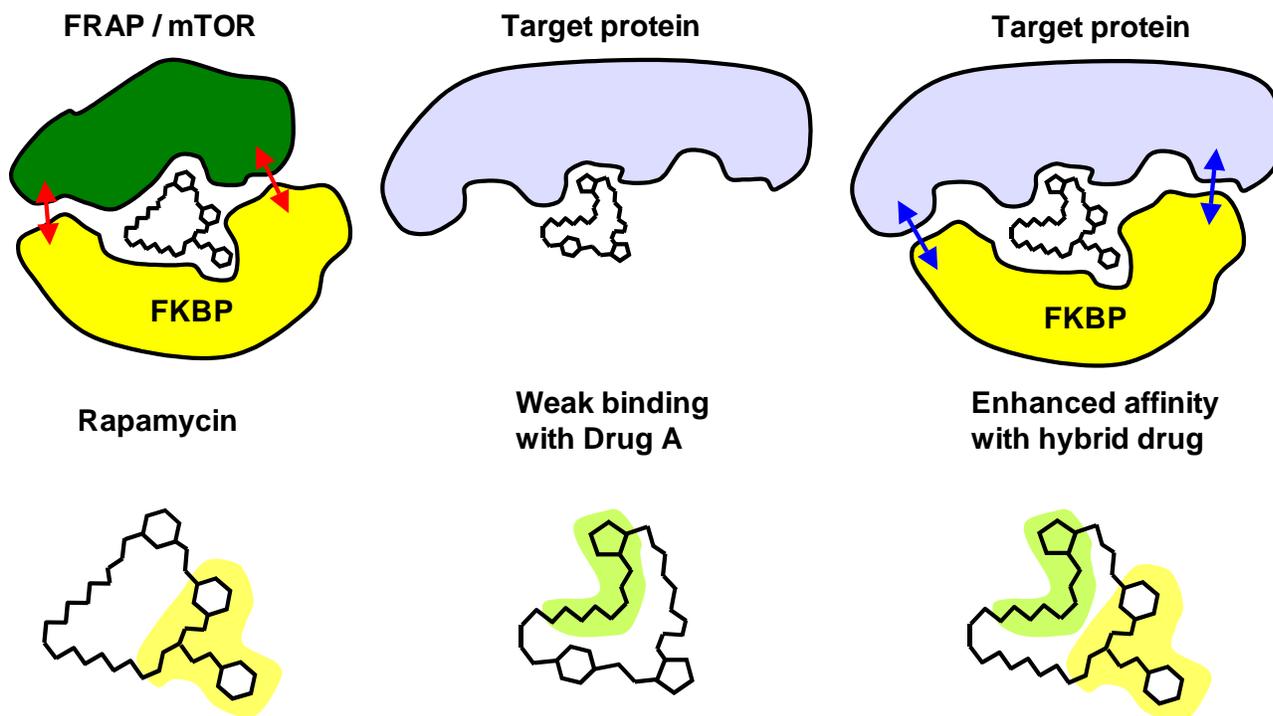


Fig. 9A15.4

Surface borrowing principle, using rapamycin as an example. If a drug moiety weakly binding a target protein of interest is joined with the FKBP-binding moiety of rapamycin, the hybrid can form a ternary complex with the target. If energetically favorable protein-protein contacts take place, the overall complex is stabilized relative to the binding of the target by the original drug alone.

The bifunctional nature of compounds such as rapamycin which interface with two separate binding pockets on the presenting and target proteins allows hybrid molecules to be constructed, such that the moiety binding the presenting protein is preserved and another known to bind a separate target is inserted. The presenting protein can then either enhance or diminish the interaction between

the target protein and the original drug <sup>153,157</sup>. Where the interaction between the target and presenter proteins is unfavorable, formation of the complex is blocked, and this can serve as a mechanism for generating cell-specific small molecule activity. (In such circumstances, co-expression of a presenter protein in a cell of interest can block a bifunctional molecule from accessing its cellular target, and ablate its normal phenotypic effects <sup>158</sup>). The presenting protein can also be engineered for modulation of ternary complex binding, as has been done for an aptamer-based adaptation of the surface borrowing principle <sup>159</sup>. In this study, the goal achieved was a specific cellular orthogonal system with a unique presenter protein (modified FKBP), a bifunctional small binding molecule, and an aptamer only recognizing the small molecule in the presence of the presenter protein <sup>159</sup>.

To conclude this section, although hits and leads obtained from completely artificial chemical libraries inform us that chance molecular interactions can lead to useful drugs, evolution indeed has already done the 'hard yards' for us. In fact, it is hard to find any examples of natural products in clinical use whose modes of action can be ascribed to 'pure chance', once the relevant molecular mechanisms have been unraveled. And very often, the explanation for why a fungal or bacterial product should modulate human physiology is attributable to the parsimony of evolutionary pathways and their components, preserved in their essential forms over vast gulfs of time.

## References:

1. Dunn, B. M., Goodenow, M. M., Gustchina, A. & Wlodawer, A. Retroviral proteases. *Genome Biol* **3 (reviews)**, 3006.1-3006.7 (2002).
2. Frankel, A. D. & Young, J. A. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* **67**, 1-25 (1998).
3. Kohl, N. E. et al. Active human immunodeficiency virus protease is required for viral infectivity. *Proc Natl Acad Sci U S A* **85**, 4686-90 (1988).
4. Huff, J. R. & Kahn, J. Discovery and clinical development of HIV-1 protease inhibitors. *Adv Protein Chem* **56**, 213-51 (2001).
5. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* **3**, 935-49 (2004).
6. Debnath, A. K. Application of 3D-QSAR techniques in anti-HIV-1 drug design--an overview. *Curr Pharm Des* **11**, 3091-110 (2005).
7. Krohn, A., Redshaw, S., Ritchie, J. C., Graves, B. J. & Hatada, M. H. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. *J Med Chem* **34**, 3340-2 (1991).
8. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat Struct Biol* **10**, 980 (2003).
9. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-23 (1997).
10. Kempf, D. J. et al. ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc Natl Acad Sci U S A* **92**, 2484-8 (1995).
11. Buchanan, C. M. et al. Pharmacokinetics of saquinavir after intravenous and oral dosing of saquinavir: hydroxybutenyl-beta-cyclodextrin formulations. *Biomacromolecules* **9**, 305-13 (2008).
12. Molina, J. M. & Hill, A. Darunavir (TMC114): a new HIV-1 protease inhibitor. *Expert Opin Pharmacother* **8**, 1951-64 (2007).
13. Bull, J. J., Meyers, L. A. & Lachmann, M. Quasispecies made simple. *PLoS Comput Biol* **1**, e61 (2005).
14. Aharoni, A. et al. The 'evolvability' of promiscuous protein functions. *Nat Genet* **37**, 73-6 (2005).
15. de Mendoza, C. & Soriano, V. Resistance to HIV protease inhibitors: mechanisms and clinical consequences. *Curr Drug Metab* **5**, 321-8 (2004).

16. Muzammil, S., Ross, P. & Freire, E. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. *Biochemistry* **42**, 631-8 (2003).
17. Menendez-Arias, L., Martinez, M., Quinones-Mateu, M. & Martinez-Picado, J. Fitness variations and their impact on the evolution of antiretroviral drug resistance. *Curr Drug Targets Infect Disord* **3**, 355-371 (2003).
18. Nijhuis, M. et al. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. *Aids* **13**, 2349-59 (1999).
19. Nijhuis, M. et al. A novel substrate-based HIV-1 protease inhibitor drug resistance mechanism. *PLoS Med* **4**, e36 (2007).
20. McCoy, C. Darunavir: a nonpeptidic antiretroviral protease inhibitor. *Clin Ther* **29**, 1559-76 (2007).
21. Fadel, H. & Temesgen, Z. Maraviroc. *Drugs Today (Barc)* **43**, 749-58 (2007).
22. Moyle, G. et al. Potential for new antiretrovirals to address unmet needs in the management of HIV-1 infection. *AIDS Patient Care STDS* **22**, 459-71 (2008).
23. Pace, P. & Rowley, M. Integrase inhibitors for the treatment of HIV infection. *Curr Opin Drug Discov Devel* **11**, 471-9 (2008).
24. Betzi, S. et al. Protein protein interaction inhibition (2P2I) combining high throughput and virtual screening: Application to the HIV-1 Nef protein. *Proc Natl Acad Sci U S A* **104**, 19256-61 (2007).
25. Velazquez-Campoy, A. & Freire, E. Incorporating target heterogeneity in drug design. *J Cell Biochem Suppl Suppl* **37**, 82-8 (2001).
26. Nezami, A. et al. High-affinity inhibition of a family of Plasmodium falciparum proteases by a designed adaptive inhibitor. *Biochemistry* **42**, 8459-64 (2003).
27. Ohtaka, H. & Freire, E. Adaptive inhibitors of the HIV-1 protease. *Prog Biophys Mol Biol* **88**, 193-208 (2005).
28. Banks, W. A., Ercal, N. & Price, T. O. The blood-brain barrier in neuroAIDS. *Curr HIV Res* **4**, 259-66 (2006).
29. Joos, B. et al. HIV rebounds from latently infected cells, rather than from continuing low-level replication. *Proc Natl Acad Sci U S A* **105**, 16725-30 (2008).
30. Joshi, P. J., Fisher, T. S. & Prasad, V. R. Anti-HIV inhibitors based on nucleic acids: emergence of aptamers as potent antivirals. *Curr Drug Targets Infect Disord* **3**, 383-400 (2003).
31. Rossi, J. J., June, C. H. & Kohn, D. B. Genetic therapies against HIV. *Nat Biotechnol* **25**, 1444-54 (2007).
32. Choudhry, V. et al. Antibody-based inhibitors of HIV infection. *Expert Opin Biol Ther* **6**, 523-31 (2006).

33. Huber, M., Olson, W. C. & Trkola, A. Antibodies for HIV treatment and prevention: window of opportunity? *Curr Top Microbiol Immunol* **317**, 39-66 (2008).
34. Mallo, M. Controlled gene activation and inactivation in the mouse. *Front Biosci* **11**, 313-27 (2006).
35. Sarkar, I., Hauber, I., Hauber, J. & Buchholz, F. HIV-1 proviral DNA excision using an evolved recombinase. *Science* **316**, 1912-5 (2007).
36. Wiznerowicz, M. & Trono, D. Harnessing HIV for therapy, basic research and biotechnology. *Trends Biotechnol* **23**, 42-7 (2005).
37. Angeletti, P. C., Zhang, L. & Wood, C. The viral etiology of AIDS-associated malignancies. *Adv Pharmacol* **56**, 509-57 (2008).
38. Murgia, C., Pritchard, J. K., Kim, S. Y., Fassati, A. & Weiss, R. A. Clonal origin and evolution of a transmissible cancer. *Cell* **126**, 477-87 (2006).
39. Pearse, A. M. & Swift, K. Allograft theory: transmission of devil facial-tumour disease. *Nature* **439**, 549 (2006).
40. Siddle, H. V. et al. Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. *Proc Natl Acad Sci U S A* **104**, 16221-6 (2007).
41. Merlo, L. M., Pepper, J. W., Reid, B. J. & Maley, C. C. Cancer as an evolutionary and ecological process. *Nat Rev Cancer* **6**, 924-35 (2006).
42. Nowell, P. C. Discovery of the Philadelphia chromosome: a personal perspective. *J Clin Invest* **117**, 2033-5 (2007).
43. Gartler, S. M. The chromosome number in humans: a brief history. *Nat Rev Genet* **7**, 655-60 (2006).
44. Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. *Cell* **22**, 777-85 (1980).
45. Hunter, T. Treatment for chronic myelogenous leukemia: the long road to imatinib. *J Clin Invest* **117**, 2036-43 (2007).
46. Capdeville, R., Buchdunger, E., Zimmermann, J. & Matter, A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat Rev Drug Discov* **1**, 493-502 (2002).
47. Nagar, B. et al. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res* **62**, 4236-43 (2002).
48. Moreland, J. L., Gramada, A., Buzko, O. V., Zhang, Q. & Bourne, P. E. The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinformatics* **6**, 21 (2005).

49. Graziani, Y., Erikson, E. & Erikson, R. L. The effect of quercetin on the phosphorylation activity of the Rous sarcoma virus transforming gene product in vitro and in vivo. *Eur J Biochem* **135**, 583-9 (1983).
50. Akiyama, T. et al. Genistein, a specific inhibitor of tyrosine-specific protein kinases. *J Biol Chem* **262**, 5592-5 (1987).
51. Umezawa, K. Isolation and biological activities of signal transduction inhibitors from microorganisms and plants. *Adv Enzyme Regul* **35**, 43-53 (1995).
52. Yaish, P., Gazit, A., Gilon, C. & Levitzki, A. Blocking of EGF-dependent cell proliferation by EGF receptor kinase inhibitors. *Science* **242**, 933-5 (1988).
53. Traxler, P. et al. Tyrosine kinase inhibitors: from rational design to clinical trials. *Med Res Rev* **21**, 499-512 (2001).
54. Steinberg, S. F. Distinctive activation mechanisms and functions for protein kinase Cdelta. *Biochem J* **384**, 449-59 (2004).
55. Zimmermann, J. et al. Phenylamino-pyrimidine (PAP) derivatives: a new class of potent and selective inhibitors of protein kinase C (PKC). *Arch Pharm (Weinheim)* **329**, 371-6 (1996).
56. Buchdunger, E. et al. Selective inhibition of the platelet-derived growth factor signal transduction pathway by a protein-tyrosine kinase inhibitor of the 2-phenylaminopyrimidine class. *Proc Natl Acad Sci U S A* **92**, 2558-62 (1995).
57. Buchdunger, E. et al. Inhibition of the Abl protein-tyrosine kinase in vitro and in vivo by a 2-phenylaminopyrimidine derivative. *Cancer Res* **56**, 100-4 (1996).
58. Bogoyevitch, M. A. & Fairlie, D. P. A new paradigm for protein kinase inhibition: blocking phosphorylation without directly targeting ATP binding. *Drug Discov Today* **12**, 622-33 (2007).
59. Liu, Y. & Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat Chem Biol* **2**, 358-64 (2006).
60. Daub, H., Specht, K. & Ullrich, A. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat Rev Drug Discov* **3**, 1001-10 (2004).
61. O'Hare, T., Eide, C. A. & Deininger, M. W. New Bcr-Abl inhibitors in chronic myeloid leukemia: keeping resistance in check. *Expert Opin Investig Drugs* **17**, 865-78 (2008).
62. Tokarski, J. S. et al. The structure of Dasatinib (BMS-354825) bound to activated ABL kinase domain elucidates its inhibitory activity against imatinib-resistant ABL mutants. *Cancer Res* **66**, 5790-7 (2006).
63. Tanaka, R. & Kimura, S. Abl tyrosine kinase inhibitors for overriding Bcr-Abl/T315I: from the second to third generation. *Expert Rev Anticancer Ther* **8**, 1387-98 (2008).
64. Gumireddy, K. et al. A non-ATP-competitive inhibitor of BCR-ABL overrides imatinib resistance. *Proc Natl Acad Sci U S A* **102**, 1992-7 (2005).

65. Fernandez, A. & Scheraga, H. A. Insufficiently dehydrated hydrogen bonds as determinants of protein interactions. *Proc Natl Acad Sci U S A* **100**, 113-8 (2003).
66. Crespo, A. & Fernandez, A. Kinase packing defects as drug targets. *Drug Discov Today* **12**, 917-23 (2007).
67. Fernandez, A. et al. An anticancer C-Kit kinase inhibitor is reengineered to make it more active and less cardiotoxic. *J Clin Invest* **117**, 4044-54 (2007).
68. Fernandez, A., Rogale, K., Scott, R. & Scheraga, H. A. Inhibitor design by wrapping packing defects in HIV-1 proteins. *Proc Natl Acad Sci U S A* **101**, 11640-5 (2004).
69. Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* **1**, 309-15 (2002).
70. Errington, J. & Illing, N. Establishment of cell-specific transcription during sporulation in *Bacillus subtilis*. *Mol Microbiol* **6**, 689-95 (1992).
71. Duncan, L., Alper, S. & Losick, R. Establishment of cell type specific gene transcription during sporulation in *Bacillus subtilis*. *Curr Opin Genet Dev* **4**, 630-6 (1994).
72. Legros, F., Malka, F., Frachon, P., Lombes, A. & Rojo, M. Organization and dynamics of human mitochondrial DNA. *J Cell Sci* **117**, 2653-62 (2004).
73. Hooper, L. V., Midtvedt, T. & Gordon, J. I. How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr* **22**, 283-307 (2002).
74. Xu, J. et al. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* **299**, 2074-6 (2003).
75. Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. Host-bacterial mutualism in the human intestine. *Science* **307**, 1915-20 (2005).
76. Tappenden, K. A. & Deutsch, A. S. The physiological relevance of the intestinal microbiota--contributions to human health. *J Am Coll Nutr* **26**, 679S-83S (2007).
77. Goodacre, R. Metabolomics of a superorganism. *J Nutr* **137**, 259S-266S (2007).
78. Queller, D. C. & Strassmann, J. E. The many selves of social insects. *Science* **296**, 311-3 (2002).
79. Zilber-Rosenberg, I. & Rosenberg, E. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev* **32**, 723-35 (2008).
80. Dawkins, R. *The Selfish Gene* (Oxford University Press, Oxford, 1976).
81. Wilson, D. S. & Sober, E. Reviving the superorganism. *J Theor Biol* **136**, 337-56 (1989).
82. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560-3 (2006).
83. Wilson, D. S. & Wilson, E. O. Rethinking the theoretical foundation of sociobiology. *Q Rev Biol* **82**, 327-48 (2007).
84. Weiss, K. M. The phenogenetic logic of life. *Nat Rev Genet* **6**, 36-45 (2005).
85. Crick, F. *What Mad Pursuit. A Personal View of Scientific Discovery* (Basic Books, 1988).

86. Jonnakuty, C. & Gragnoli, C. What do we know about serotonin? *J Cell Physiol* **217**, 301-6 (2008).
87. Klein, D. C. Arylalkylamine N-acetyltransferase: "the Timezyme". *J Biol Chem* **282**, 4233-7 (2007).
88. Philipp, M. & Hein, L. Adrenergic receptor knockout mice: distinct functions of 9 receptor subtypes. *Pharmacol Ther* **101**, 65-74 (2004).
89. Hoyer, D., Hannon, J. P. & Martin, G. R. Molecular, pharmacological and functional diversity of 5-HT receptors. *Pharmacol Biochem Behav* **71**, 533-54 (2002).
90. Millan, M. J., Marin, P., Bockaert, J. & Mannoury la Cour, C. Signaling at G-protein-coupled serotonin receptors: recent advances and future research directions. *Trends Pharmacol Sci* **29**, 454-64 (2008).
91. Li, J., Nagpal, P., Vitart, V., McMorris, T. C. & Chory, J. A role for brassinosteroids in light-dependent development of Arabidopsis. *Science* **272**, 398-401 (1996).
92. Russell, D. W. Green light for steroid hormones. *Science* **272**, 370-1 (1996).
93. Ayala, F. J. "Nothing in biology makes sense except in the light of evolution": Theodosius Dobzhansky: 1900-1975. *J Hered* **68**, 3-10 (1977).
94. Sharom, J. R., Bellows, D. S. & Tyers, M. From large networks to small molecules. *Curr Opin Chem Biol* **8**, 81-90 (2004).
95. Gavin, A. C. et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631-6 (2006).
96. O'Reilly, A. O. et al. Modelling insecticide-binding sites in the voltage-gated sodium channel. *Biochem J* **396**, 255-63 (2006).
97. Robison, A. K., Schmidt, W. A. & Stancel, G. M. Estrogenic activity of DDT: estrogen-receptor profiles and the responses of individual uterine cell types following o,p'-DDT administration. *J Toxicol Environ Health* **16**, 493-508 (1985).
98. Steinmetz, R. et al. Novel estrogenic action of the pesticide residue beta-hexachlorocyclohexane in human breast cancer cells. *Cancer Res* **56**, 5403-9 (1996).
99. Rosselli, M., Reinhart, K., Imthurn, B., Keller, P. J. & Dubey, R. K. Cellular and biochemical mechanisms by which environmental oestrogens influence reproductive function. *Hum Reprod Update* **6**, 332-50 (2000).
100. Tapiero, H., Ba, G. N. & Tew, K. D. Estrogens and environmental estrogens. *Biomed Pharmacother* **56**, 36-44 (2002).
101. Kuiper, G. G. et al. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology* **139**, 4252-63 (1998).
102. Klotz, D. M. et al. Identification of environmental chemicals with estrogenic activity using a combination of in vitro assays. *Environ Health Perspect* **104**, 1084-9 (1996).

103. Robison, A. K., Sirbasku, D. A. & Stancel, G. M. DDT supports the growth of an estrogen-responsive tumor. *Toxicol Lett* **27**, 109-13 (1985).
104. Mohler, J. Requirements for hedgehog, a segmental polarity gene, in patterning larval and adult cuticle of *Drosophila*. *Genetics* **120**, 1061-72 (1988).
105. Lupi, O. Correlations between the Sonic Hedgehog pathway and basal cell carcinoma. *Int J Dermatol* **46**, 1113-7 (2007).
106. Pathi, S. et al. Comparative biological responses to human Sonic, Indian, and Desert hedgehog. *Mech Dev* **106**, 107-17 (2001).
107. Maclean, K. Humour of gene names lost in translation to patients. *Nature* **439**, 266 (2006).
108. Simonite, T. Pokemon blocks gene name. *Nature* **438**, 897 (2005).
109. Romer, J. T. et al. Suppression of the Shh pathway using a small molecule inhibitor eliminates medulloblastoma in *Ptc1(+/-)p53(-/-)* mice. *Cancer Cell* **6**, 229-40 (2004).
110. Berman, D. M. et al. Medulloblastoma growth inhibition by hedgehog pathway blockade. *Science* **297**, 1559-61 (2002).
111. Sanchez, P. & Ruiz i Altaba, A. In vivo inhibition of endogenous brain tumors through systemic interference of Hedgehog signaling in mice. *Mech Dev* **122**, 223-30 (2005).
112. Kimura, H., Ng, J. M. & Curran, T. Transient inhibition of the Hedgehog pathway in young mice causes permanent defects in bone structure. *Cancer Cell* **13**, 249-60 (2008).
113. King, R. W. Roughing up Smoothened: chemical modulators of hedgehog signaling. *J Biol* **1**, 8 (2002).
114. Chiang, C. et al. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383**, 407-13 (1996).
115. Roessler, E. et al. Mutations in the human Sonic Hedgehog gene cause holoprosencephaly. *Nat Genet* **14**, 357-60 (1996).
116. Meierhofer, C. & Wiedermann, C. J. New insights into the pharmacological and toxicological effects of thalidomide. *Curr Opin Drug Discov Devel* **6**, 92-9 (2003).
117. Melchert, M. & List, A. The thalidomide saga. *Int J Biochem Cell Biol* **39**, 1489-99 (2007).
118. Bongartz, T. et al. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *Jama* **295**, 2275-85 (2006).
119. Schiff, M. H. et al. Safety analyses of adalimumab (HUMIRA) in global clinical trials and US postmarketing surveillance of patients with rheumatoid arthritis. *Ann Rheum Dis* **65**, 889-94 (2006).
120. Askling, J. & Dixon, W. The safety of anti-tumour necrosis factor therapy in rheumatoid arthritis. *Curr Opin Rheumatol* **20**, 138-44 (2008).

121. Smith, D. A. & Schmid, E. F. Drug withdrawals and the lessons within. *Curr Opin Drug Discov Devel* **9**, 38-46 (2006).
122. Michnick, S. W., Ear, P. H., Manderson, E. N., Remy, I. & Stefan, E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat Rev Drug Discov* **6**, 569-82 (2007).
123. Dowsing, T. & Kendall, M. J. The Northwick Park tragedy--protecting healthy volunteers in future first-in-man trials. *J Clin Pharm Ther* **32**, 203-7 (2007).
124. Schraven, B. & Kalinke, U. CD28 superagonists: what makes the difference in humans? *Immunity* **28**, 591-5 (2008).
125. MacDonald, M. L. et al. Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nat Chem Biol* **2**, 329-37 (2006).
126. Bartlett, J. G. Narrative review: the new epidemic of Clostridium difficile-associated enteric disease. *Ann Intern Med* **145**, 758-64 (2006).
127. Wilhelm, S. et al. Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nat Rev Drug Discov* **5**, 835-44 (2006).
128. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-34 (2002).
129. Vieth, M., Sutherland, J. J., Robertson, D. H. & Campbell, R. M. Kinomics: characterizing the therapeutically validated kinase space. *Drug Discov Today* **10**, 839-46 (2005).
130. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat Rev Drug Discov* **5**, 993-6 (2006).
131. Imming, P., Sinning, C. & Meyer, A. Drugs, their targets and the nature and number of drug targets. *Nat Rev Drug Discov* **5**, 821-34 (2006).
132. Paolini, G. V., Shapland, R. H., van Hoorn, W. P., Mason, J. S. & Hopkins, A. L. Global mapping of pharmacological space. *Nat Biotechnol* **24**, 805-15 (2006).
133. Graul, A. I. et al. The year's new drugs and biologics--2007. *Drug News Perspect* **21**, 7-35 (2008).
134. Teo, S. K., Stirling, D. I. & Zeldis, J. B. Thalidomide as a novel therapeutic agent: new uses for an old product. *Drug Discov Today* **10**, 107-14 (2005).
135. Teo, S. K. et al. Clinical pharmacokinetics of thalidomide. *Clin Pharmacokinet* **43**, 311-27 (2004).
136. Tweedie, D., Sambamurti, K. & Greig, N. H. TNF-alpha inhibition as a treatment strategy for neurodegenerative disorders: new drug candidates and targets. *Curr Alzheimer Res* **4**, 378-85 (2007).
137. Lefranc, F., Yeaton, P., Brotchi, J. & Kiss, R. Cimetidine, an unexpected anti-tumor agent, and its potential for the treatment of glioblastoma (review). *Int J Oncol* **28**, 1021-30 (2006).

138. Spiler, N. M., Rork, T. H. & Merrill, G. F. An old drug with a new purpose: cardiovascular actions of acetaminophen (paracetamol). *Curr Drug Targets Cardiovasc Haematol Disord* **5**, 419-29 (2005).
139. Jaques-Robinson, K. M., Golfetti, R., Baliga, S. S., Hadzimichalis, N. M. & Merrill, G. F. Acetaminophen is cardioprotective against H<sub>2</sub>O<sub>2</sub>-induced injury in vivo. *Exp Biol Med (Maywood)* (2008).
140. Ghofrani, H. A., Osterloh, I. H. & Grimminger, F. Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. *Nat Rev Drug Discov* **5**, 689-702 (2006).
141. Wermuth, C. G. Selective optimization of side activities: the SOSA approach. *Drug Discov Today* **11**, 160-4 (2006).
142. Mann, J. Natural products as immunosuppressive agents. *Nat Prod Rep* **18**, 417-30 (2001).
143. Lee, M. & Park, J. Regulation of NFAT activation: a potential therapeutic target for immunosuppression. *Mol Cells* **22**, 1-7 (2006).
144. Jin, L. & Harrison, S. C. Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. *Proc Natl Acad Sci U S A* **99**, 13522-6 (2002).
145. Griffith, J. P. et al. X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. *Cell* **82**, 507-22 (1995).
146. Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239-42 (1996).
147. Cyert, M. S., Kunisawa, R., Kaim, D. & Thorner, J. Yeast has homologs (CNA1 and CNA2 gene products) of mammalian calcineurin, a calmodulin-regulated phosphoprotein phosphatase. *Proc Natl Acad Sci U S A* **88**, 7376-80 (1991).
148. Dolinski, K., Muir, S., Cardenas, M. & Heitman, J. All cyclophilins and FK506 binding proteins are, individually and collectively, dispensable for viability in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **94**, 13093-8 (1997).
149. Fox, D. S. & Heitman, J. Good fungi gone bad: the corruption of calcineurin. *Bioessays* **24**, 894-903 (2002).
150. Breuder, T., Hemenway, C. S., Movva, N. R., Cardenas, M. E. & Heitman, J. Calcineurin is essential in cyclosporin A- and FK506-sensitive yeast strains. *Proc Natl Acad Sci U S A* **91**, 5372-6 (1994).
151. Crabtree, G. R. & Schreiber, S. L. Three-part inventions: intracellular signaling and induced proximity. *Trends Biochem Sci* **21**, 418-22 (1996).

152. Cardenas, M. E., Sanfridson, A., Cutler, N. S. & Heitman, J. Signal-transduction cascades as targets for therapeutic intervention by natural products. *Trends Biotechnol* **16**, 427-33 (1998).
153. Vogel, K. W., Briesewitz, R., Wandless, T. J. & Crabtree, G. R. Calcineurin inhibitors and the generalization of the presenting protein strategy. *Adv Protein Chem* **56**, 253-91 (2001).
154. Klee, C. B., Crouch, T. H. & Krinks, M. H. Calcineurin: a calcium- and calmodulin-binding protein of the nervous system. *Proc Natl Acad Sci U S A* **76**, 6270-3 (1979).
155. Buell, J. F., Gross, T. G. & Woodle, E. S. Malignancy after transplantation. *Transplantation* **80**, S254-64 (2005).
156. Ehninger, D. et al. Reversal of learning deficits in a Tsc2<sup>+/-</sup> mouse model of tuberous sclerosis. *Nat Med* **14**, 843-8 (2008).
157. Briesewitz, R., Ray, G. T., Wandless, T. J. & Crabtree, G. R. Affinity modulation of small-molecule ligands by borrowing endogenous protein surfaces. *Proc Natl Acad Sci U S A* **96**, 1953-8 (1999).
158. Sellmyer, M. A., Stankunas, K., Briesewitz, R., Crabtree, G. R. & Wandless, T. J. Engineering small molecule specificity in nearly identical cellular environments. *Bioorg Med Chem Lett* **17**, 2703-5 (2007).
159. Plummer, K. A., Carothers, J. M., Yoshimura, M., Szostak, J. W. & Verdine, G. L. In vitro selection of RNA aptamers against a composite small molecule-protein surface. *Nucleic Acids Res* **33**, 5602-10 (2005).