IMPLICATIONS OF PROTEIN FLEXIBILITY FOR DRUG DISCOVERY

Simon J. Teague

Proteins are in constant motion between different conformational states with similar energies. This has often been ignored in drug design. However, protein flexibility is fundamental to understanding the ways in which drugs exert biological effects, their binding-site location, binding orientation, binding kinetics, metabolism and transport. Protein flexibility allows increased affinity to be achieved between a drug and its target. This is crucial, because the lipophilicity and number of polar interactions allowed for an oral drug is limited by absorption, distribution, metabolism and toxicology considerations.

Protein flexibility: diverse views

Induced fit, conformer selection and induction. The intrinsic mobility of proteins has often been ignored in drug design, despite recent accounts of its importance in the design of non-peptide peptidomimetics[1]. When it has been acknowledged, it is often conceptualized as 'induced fit'. The ligand (drug) is imagined to bind to the lowest-energy conformation of the protein, which is then distorted in order to accommodate the ligand.

In the biochemical literature, a different model dominates. Here, a protein is assumed to pre-exist in a number of energetically similar conformations. Conformational selection is defined as a process in which a ligand selectively binds to one of these conformers, thereby increasing the proportion of this conformer in the total protein population and producing the observed effect. Conformational induction is used to describe a process in which the ligand converts the protein into a conformation that it would not spontaneously adopt in its unligated state.

The constant cycling of a protein between a resting and activated conformation is also implicit in modern definitions of agonist and antagonist behaviour. An agonist shifts the equilibrium to the active conformation by preferentially binding to it. Neutral antagonists have equal affinity for both the resting and activated conformer, whereas inverse agonists bind preferentially to the resting conformation. Allosteric inhibitors are also common. For proteins with allosteric sites, the equilibrium between conformationally distinct states, with different substrate affinities, can be modulated by binding at these sites, which are remote from the active site[2].

From a thermodynamic standpoint, the induced-fit and conformational-selection models are equivalent. However, the concepts of conformational selection and induction are more useful when describing and exploiting many of the phenomena encountered during drug design. The effect of protein mobility on ligand binding can be illustrated using a simple model of an energy landscape (FIG. 1). It is clear that some protein–ligand complexes result from conformations of the protein that go well beyond those which might be anticipated for normal substrate processing. In these cases, ligand binding results in conformational induction, which produces a cascade of movements and sometimes even major conformational transitions. In many ways, these changes can be likened to a refolding of the protein by the ligand. These ligands are usually hydrophobic, and protein folding is often described using the 'hydrophobic zipper' model[3]. This involves the formation of transient, stable compact hydrophobic units (CHUs). Mutual recognition between these units eventually results in a slide towards a new or refolded protein structure. For instance, in the folding of chicken lysozyme, Trp62 plays a central role in a network of hydrophobic clusters that guide folding. Similarly, piperidine-based inhibitors of renin[4] (FIG. 2) interact with hydrophobic residues such as Trp75, which results in an enzyme conformation not previously observed for this enzyme. In another example, roughly half of

AstraZeneca R&D Charnwood, Bakewell Road, Loughborough, Leicester LE11 5RH, UK.
e-mail: simon.teague@astrazeneca.com
doi:10.1038/nrd1129

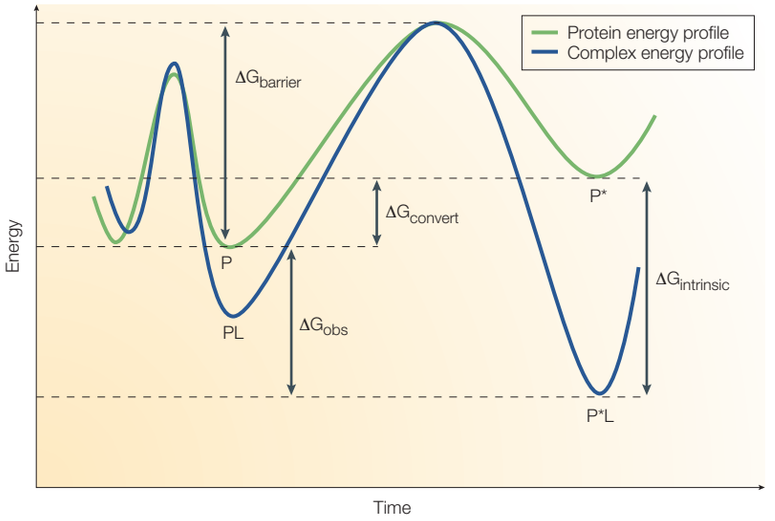


Figure 1 | **Protein mobility and ligand binding.** A protein is considered to exist in two conformations (P and P*) with an energy difference $\Delta G_{convert}$. The ligand (L) can bind the protein (P) to give a complex (PL), or bind to P* to give a complex (P*L). Although P* has a higher free energy, it might offer greater scope for interaction with L. For instance, P* might represent a conformer in which the binding site has opened and exposed hydrophobic patches. This is energetically unfavourable, but offers the potential for favourable interactions with the hydrophobic moiety of a suitable incoming L, thereby giving rise to a large, favourable interaction $\Delta G_{intrinsic}$. The resulting complex (P*L) has a lower energy than that of the complex PL. The observed affinity of L for the protein conformational ensemble is governed by ΔG_{obs}. Slow binding kinetics might well be observed, as P* is a higher-energy conformer than P and an energy barrier ($\Delta G_{barrier}$) must be surmounted before optimal binding to L can take place.

thymidylate synthetase is refolded by 1843U89 as lipophilic parts of this drug interact with residues such as Ile79 and Phe176 (FIG. 3)[5]. Conformational induction by hydrophobic ligands and protein folding through the formation of CHUs is strikingly similar.

Evidence for protein mobility. Protein mobility is required for function. In enzymes, the substrates must be absorbed, the transition state orientation enforced and the products desorbed. For membrane-bound receptors, the message of ligand binding must be transmitted from outside the cell to the inside. Channels must open and close. Proteins also often need to bind various ligands and substrates sequentially or perform different functions in different cellular environments. Motion can take place over picoseconds to hours and with amplitudes as large as 1.5 nm[6].

Protein dynamics have often been studied by NMR. For instance, it has been shown that it is possible to correlate the motion of key catalytic residues with the frequency of transition-state formation and decay[7]. X-ray crystallography of protein–ligand complexes routinely reveals ligands with 70–100% of their surface area buried. This can only be achieved as a consequence of protein flexibility. Although most X-ray structure determinations are time-averaged, the data are often collected at low temperature, and occasionally two protein molecules with quite different conformations are observed in the same crystallographic unit cell. For example, it has been shown that in the complex between the aspartic proteinease Plasmepsin II and an inhibitor Pepstatin A, two different conformations of Plasmepsin II are present within the unit cell. Each conformer binds the inhibitor in a different way[8]. Likewise, open and closed conformers of renin have been crystallized in the same unit cell[9].

Indirect evidence for protein mobility is also strong. G-protein-coupled receptors cycle spontaneously between inactive and active forms, as spontaneous complexes between the receptor and G-protein can be immunoprecipitated in the absence of the required agonist. Cells transfected with high levels of wild-type receptors show increased basal levels of cellular activity in the absence of agonist. Mutations that impart constitutive activity to G-proteins result in a higher affinity for agonists, but not for antagonists, when compared with the endogenous receptor[10].

Common types of mobility in proteins. Proteins can be considered as being composed of a number of CHUs connected by loops. The CHUs act as relatively rigid bodies that move with respect to each other[11]. Ligand-binding sites are typically composed of residues located on loops, or from two or more domains. Domain movements must trap ligands and so must be fast, implying a low energy barrier for transition between conformational states. Enzymes are selected to possess regions of instability, as too much stability adversely affects enzyme function[12]. Some common mechanisms for domain movement have been outlined[13] (FIG. 4).

Protein mobility concomitant with ligand binding

Multiple conformations of a few residues. As the amount of structural data available for ligand–protein complexes has increased, so have examples in which multiple ligands have been observed bound to different conformations of the same protein. Some examples involving drug-like molecules at therapeutically important receptors have been reviewed[14].

The structure of donepezil (Aricept; 3) (FIG. 5), a marketed anticholinergic drug, bound to *Torpedo californica* acetylcholinesterase (TcAchE) has been determined at 2.5 Å resolution[15]. It has a K_i = 3.35 nM for acetylcholinesterase. It binds in the active-site gorge through π–π STACKING and CATION–π INTERACTIONS. It is not a classical transition-state mimic, but straddles the catalytic triad without interacting directly with it. The mobility of Phe330 in TcAchE is indicated by the wide range of conformations that it adopts when it binds different inhibitors . The flexibility of Phe330, in comparison with the rigidity of the rest of the gorge, was highlighted. It is believed that Phe330 behaves as a swinging gate, which guides acetylcholine towards the active site, while simultaneously isolating the reaction centre from the rest of the gorge.

In order to function, residues involved in the catalytic cycle of an enzyme will often be mobile. This will sometimes allow for the binding of larger inhibitor molecules, which freeze the key enzymatic side chains in an open conformation. For example, His64 is crucial in the catalytic cycle of human carbonic anhydrase[16],

π–π STACKING
Weak non-covalent interactions between the faces of two aromatic moieties, one of which is electron rich and the other electron deficient.

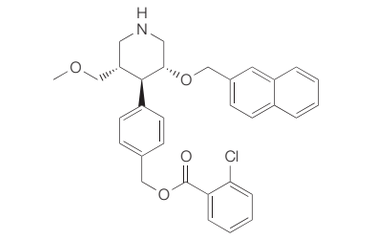
CATION–π INTERACTION
A surprisingly strong non-covalent force in which cations bind to the π face of aromatic rings. The interaction can be considered an electrostatic attraction between a positive charge and the quadrupole moment of the aromatic moiety, typically the side chains of phenylalanine, tyrosine and tryptophan residues.

in which it acts as a proton shuttle. It is observed in one conformation, accompanied by a water molecule, when inhibitors such as (4; K_i = 1.52 nM) (FIG. 6a) are bound. When the complex with compound (5) (FIG. 6a) was determined at 1.6 Å resolution, it revealed that the larger *N*-alkyl group results in an alternative and probably less favourable conformation of His64. However, the previously observed water molecule was also eliminated, resulting in an entropic gain, as the larger inhibitor still shows K_i = 0.6 nM. The position of His64 was already known to be delicately balanced. Its position is also altered either by mutations (Thr200Ser, His94Cys) or by a change of pH from 8.0 to 5.7 (REF. 17).

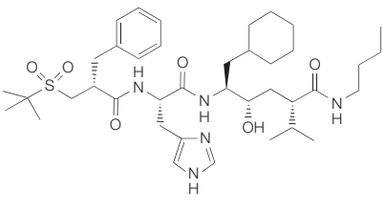
Movement of a large number of residues. The complex between *Drosophila melanogaster* acetylcholinesterase (DmAChE) and the tacrine analogue (6; K_i = 4 nM) (FIG. 7a), solved at a resolution of 2.7 Å, shows displacement of nine aromatic residues when compared with the native structure. The mobility of these residues is further highlighted by comparison between this complex and another with an iodo analogue (7; K_i = 1.09 nM) (FIG. 7b), in which a further round of changes is observed[18]. It is important to note that just like the protein residues, the acridinamine ring systems of these inhibitors do not exactly overlay. This accommodation of a new ligand by mutual mobility of both protein and ligand moieties is often ignored when medicinal chemists consider structure–activity relationships in the absence of such structural data.

The complex between renin and inhibitor (1; IC_{50} = 2 nM) (FIG. 2) has been determined at 2.97 Å resolution, and dramatic structural changes in renin were revealed. Similar conformations had never been observed during the previous decade of X-ray crystallographic study with peptidic ligands[4]. Comparison with a typical peptidomimetic TRANSITION-STATE ANALOGUE inhibitor (2) showed not only a movement of the loop between residue Thr72 and Ser81, but also further changes in the protein around the bulky lipophilic 4′-chlorobenzoyl substituent. The 4′-substituent and hydrophobic side chains of renin act to form a new compact hydrophobic unit or region within the protein. The side chain of Tyr75 is repositioned by a –120° χ^1 ANGLE rotation of the side chain of Trp39, which is flipped together with a number of small main-chain and side-chain conformational changes. The 4′-chlorobenzoyl group is essentially solvent inaccessible with ~130 $Å^2$ of hydrophobic surface buried. This probably contributes considerably to the overall stability of the complex and hence the high affinity of (1). The observed conformational induction can be likened to refolding of the protein. Alternatively, these changes can be rationalized as a complex series of conformational selections[1].

Enthalpy–entropy compensation. Backbone and side-chain flexibility can either increase or decrease as a result of ligand binding. Decreases are often associated with enthalpy–entropy compensation, freezing motion in order to participate effectively in binding interactions. In



1



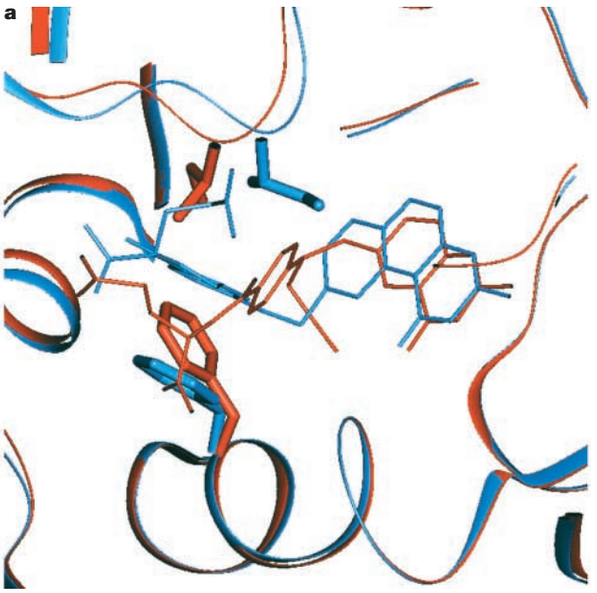
2

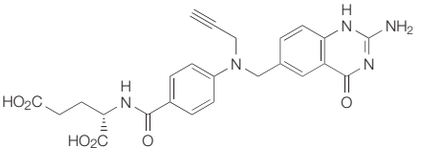
Figure 2 | **Novel conformation of renin observed on binding a non-peptidomimetic ligand.** The complex between renin (blue) and compound (1) in red overlaid with renin (yellow) bound to a transition-state analogue inhibitor (2). For clarity, (2) has been hidden and only those residues within a 10 Å sphere around the active site are displayed. The side chains for residues Trp39 and Tyr75 from each complex are displayed as thick lines in blue for the complex with (1) and yellow for the complex with (2). The differences in the positions are particularly striking, but many other changes in the conformation of the protein are also required to accommodate (1). The changes can be likened to a local refolding of the protein. (Protein Data Bank code 1RNE for the renin–(2) complex; the coordinates for the complex with compound (1) were kindly provided by Dr H.-P. Märki, F. Hoffmann-La Roche Ltd).

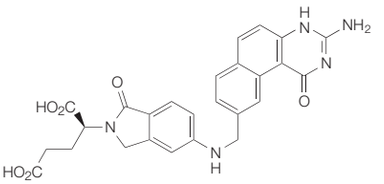
TRANSITION-STATE ANALOGUE
A compound designed to mimic the properties or geometry of the transition state of an enzymatic reaction, but which is stable to processing by the enzyme. Such compounds inhibit the enzyme by binding more tightly than the natural substrate(s).

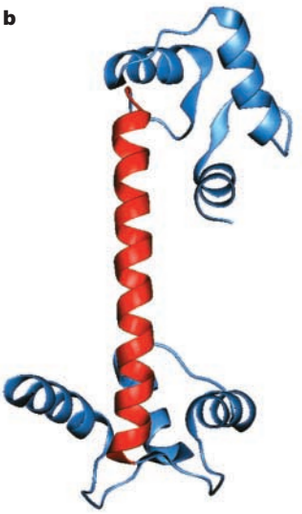
χ^1 ANGLE
χ^1 is the torsional angle that describes the conformation(s) of atoms which form the side-chain of a particular amino-acid residue relative to the peptide backbone.

other cases, increases in conformational entropy can contribute to stabilization of complexes. Small hydrophobic ligands can bind by inducing subtle changes in protein structure that result in an increase in entropy in regions of the binding protein[19]. In aminoacyl-transfer RNA (tRNA) synthetases, there is a great need to discriminate between similar amino-acid side chains. For instance, isoleucine, threonine and alanine must be selected for in the presence of valine, serine and glycine. This needs to be achieved while deriving only modest affinity, in order to facilitate product release. One way that these proteins reconcile these factors is to use both highly mobile and highly static regions. Exquisite levels of recognition are achieved from the highly specific, shape-matched, amino-acid binding pocket. The high binding energies that would be derived are offset by stabilization of previously flexible regions elsewhere in the protein. For prolyl-tRNA and histidyl-tRNA synthetase, high specificity for activation and aminoacylation of 3′-tRNA is seen through the use of conformational









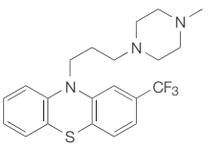


Figure 3 | **Conformational induction of proteins by ligands. a** | Overlay of the complexes between thymidylate synthetase and 1843U89 (blue) and CB3717 (red). The side chains of residues Ile79 and Phe176 are drawn in thick sticks and colour coded for their respective complexes. Their differing positions in the two complexes can be clearly seen. Many other conformational changes are also induced as the inhibitors bind to the enzyme (Protein Data Bank codes 1TSD and 2TSC). **b** | Unligated Ca^{2+}-calmodulin. Residues Asp64–Lys94, which comprise the central α-helix, are coloured red. The residues of the terminal lobes are in blue. **c** | Ca^{2+}-calmodulin complex with trifluoperazine. Two molecules of the ligand are bound and are coloured in yellow and green. The ligands induce a profound conformational change (Protein Data Bank codes 1A29 and 3CLN).

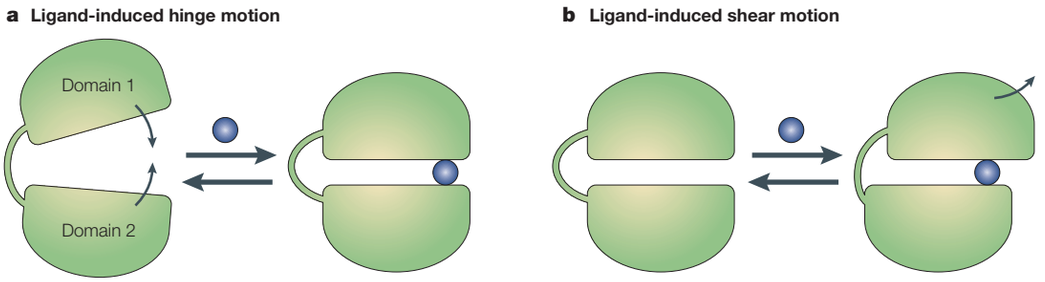


Figure 4 | **Common mechanisms for domain movement. a** | Hinge bending is one common protein motion. The ligand-induced motion is perpendicular to the plane of the domain interface. Hinge motion is typified by fairly large changes in main-chain angles for a few residues. Often these are located in a hinge region, which can also contain sites for allosteric modulation. **b** | A second common movement is shear motion. Here, the domains move parallel to the domain interface. This motion is typified by taking place amongst closely packed domains and involving the concatenation of many small motions. The movements in a number of therapeutically important proteins can be viewed as videos from a web-site devoted to their study (see Further information).

ordering of loops close to the active site[20]. For most complexes, the loss of a ligand increases protein mobility, priming the protein for its next encounter.

Strategies used at present in drug design usually generate hydrophobic and conformationally constrained ligands. The thermodynamic signature of these ligands is an entropically dominated binding affinity often accompanied by an unfavourable binding enthalpy. Conformationally constrained ligands cannot easily adapt to changes in the geometry of the binding site. As such, they are highly susceptible to drug resistance mutations or naturally occurring genetic polymorphisms. Ligands with a favourable binding enthalpy component to their affinity are less susceptible[21].

Ligand binding often reveals protein flexibility

TABLE 1 brings together a large number of pharmaceutically relevant targets that have been shown to be flexible. This extensive list of present and historical pharmaceutical targets provides the basis for questioning the relevance of the rigid receptor hypothesis, which is commonly used during *in silico* design.

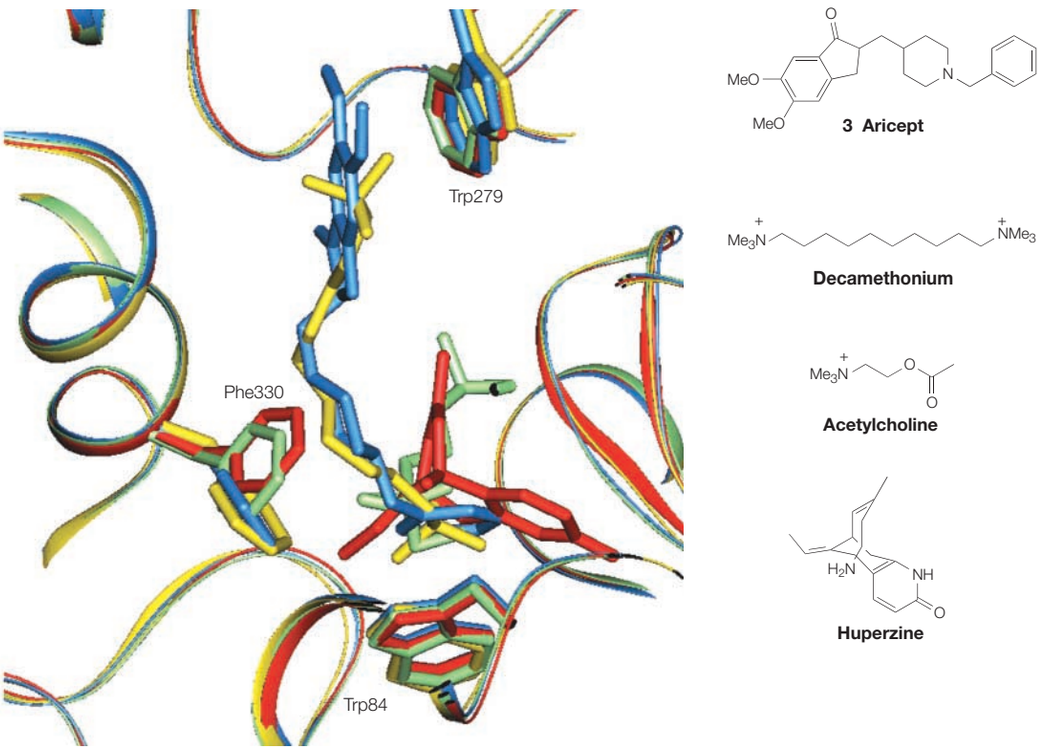


Figure 5 | **Multiple conformations of a single residue.** Overlay of native and three *Torpedo californica* acetylcholinesterase–ligand complexes using the protein C-α atoms (Protein Data Bank codes 2ACE, 1EVE, 1VOT and 1ACL). Key protein side chains are indicated by thick lines, as are the inhibitors. The colour codes are: donepezil (Aricept, blue), decamethonium (yellow), native (green), huperzine (red). The flexibility of Phe330, in comparison with the rigidity of the rest of the gorge, is highlighted.

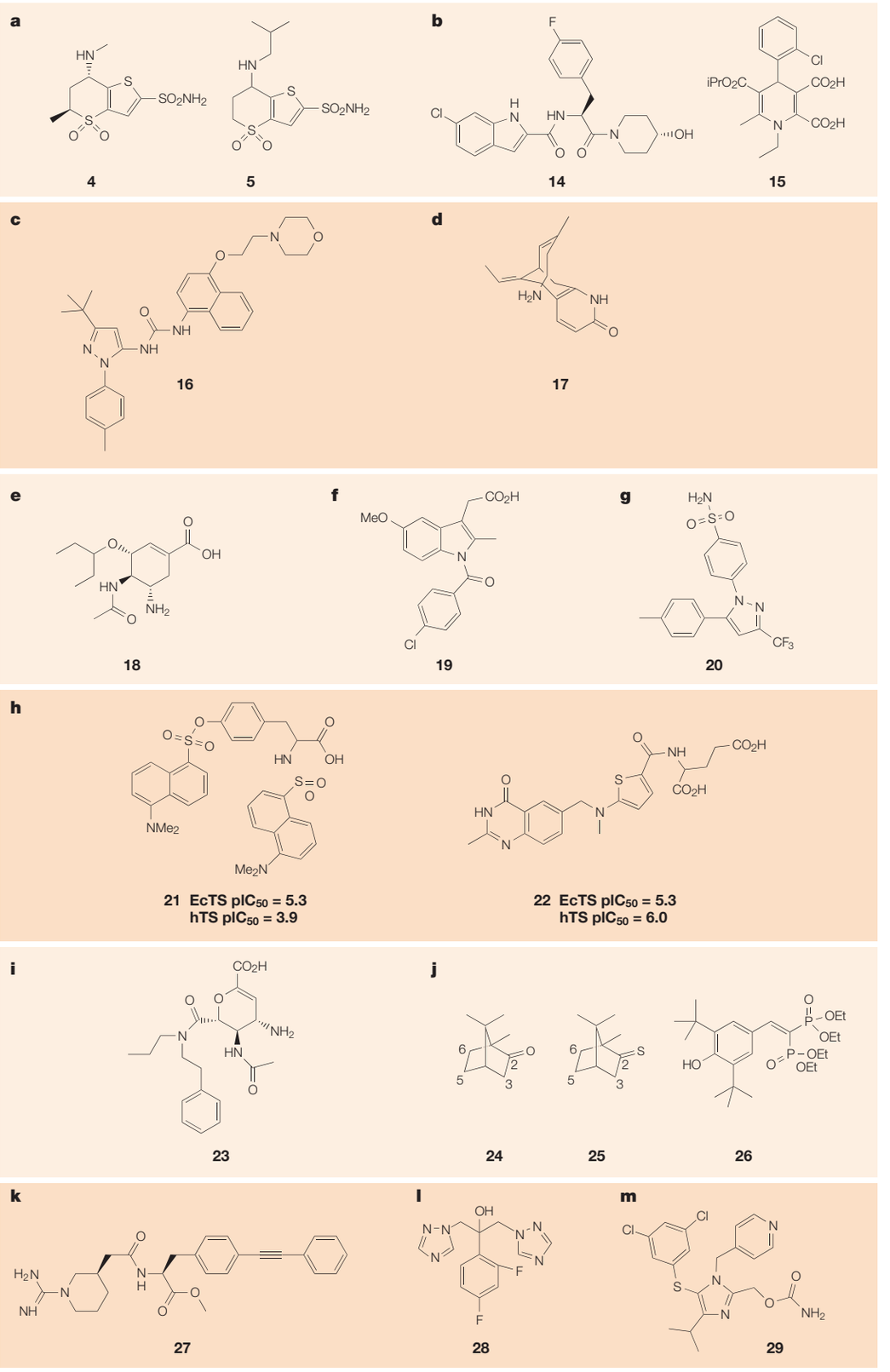


Figure 6 | **Chemical structures of various compounds highlighted. a** | Inhibitors of human carbonic anhydrase. **b** | Inhibitors of glycogen phosphorylase. **c** | BIRB 796. **d** | Huperzine. **e** | Oseltamivir. **f** | Indomethacin. **g** | Celecoxib. **h** | Inhibitors of thymidylate synthase. **i** | Inhibitor of the A form of influenza neuraminidase. **j** | Substrates of proteins involved in xenobiotic metabolism. **k** | Inhibitor of interleukin-2. **l** | Fluconazole. **m** | S-1153. EcTS, *Escherichia coli* thymidylate synthase; hTS, human thymidylate synthase.

A database of pairs of ligated and unligated protein structures from the Protein Data Bank (PDB) has been compiled, and the number and identity of binding-pocket residues that undergo side-chain conformational change have been determined[22]. A study restricted to drug-like molecules, which tend to possess regions of hydrophobicity to aid transport across membranes, would be useful. However, the errors commonly incorporated into structural models during their derivation from the observed electron density would need to be carefully taken into account.

The role of protein flexibility in future pharmaceutical targets is likely to be even greater than in the past. Presently underexploited target classes, such as ion channels, nuclear hormone receptors and transporters, have functions that are inextricably bound up with their structural flexibility. Similarly, allosteric modulation provides opportunities to more subtly affect biological processes and are likely to be increasingly sought after. Promiscuity in ligand binding is also intimately connected with protein flexibility, and the proteins involved in metabolism are likely to continue to provide a focus for investigation throughout the pharmaceutical industry.

Figure 7 | **Movement of a large number of residues: example 1. a** | The complex between *Drosophila melanogaster* acetylcholinesterase and compound (6) in yellow showing displacement of nine aromatic residues when compared with the native structure in blue (Protein Data Bank codes 1QO9 and 1DX4). **b** | Overlay of the complexes between *Drosophila melanogaster* acetylcholinesterase complexed with compounds (6) in red and (7) in yellow. Both the side chains of the protein and the positions of the inhibitors are altered (Protein Data Bank codes 1QON and 1DX4).

Consequences for ligand design

One site but many ligands and surprising orientations. With flexible receptors, binding-site shape and size is to some extent determined by the ligand. One consequence is that there are numerous instances in which structurally diverse ligands occupy almost exactly the same volume within a binding site. For instance, efavirenz (8), nevirapine (9), UC-781 (10) and Cl-TIBO (11) (FIG. 8), despite their structural diversity, bind to human immunodeficiency reverse transcriptase (HIV-RT) within much the same volume[23]. A diverse collection of putative ligands bind in only a few locations, rather than all over a protein because the CHUs exclude much of the available volume, leaving suitable sites composed of residues located between their interfaces and on loops.

A flexible protein offers several possible binding modes, even for structurally similar ligands. As a result, surprising orientations are actually quite common. One example is that of flufenamic acid (12) and diclofenamic acid (13), which bind to human transthyretin in opposing orientations (FIG. 9)[24]. Flufenamic acid binds with the carboxylate pointing outwards at the entrance of the hormone-binding cavity, but diclofenamic acid binds with the carboxylate pointing down into the interior of the receptor.

Allosteric inhibitors. Biological systems display feedback regulation, and respond to change in ways that are dependent on a number of different physiological signals. It is therefore unsurprising to find that activity at many binding sites can be regulated using other sites on the protein. Allosteric sites are commonly located on loops that act as hinges about which the CHUs move as they pass between conformational states.

The binding site of allosteric inhibitors can be remote from the active site. The site for HIV-RT inhibitors, such as nevirapine (9), is 10 Å from the active site[25], and is located on a loop connecting β-strands seven and eight. The catalytic aspartic acid residues,

Table 1 | **Some examples of mobile residues observed in protein–ligand complexes**

Protein	Ligand	Key residues	PDB code	References
Acetylcholinesterase	Huperzine (17)* Donepezil (Aricept) (3) Compounds (6) and (7)	Trp84, Phe330 Phe330 Nine aromatic residues	2ACE, 1VOT 1EVE 1QO9, 1DX4, 1QON	57 15 18
Aldose reductase	Zopolrestat Tolrestat	Loop including Phe122, Leu300 Leu300, Phe122	1ABN, 1MAR 1AH3	58 59
D-Amino-acid oxidase	Many	Loop 216–228, especially Tyr224	1AA8, 1DDO, 1DAO	60
Calmodulin	Trifluoperazine	Extensive rearrangements	1LIN	61
Carbonic anhydrase	Compounds (4) and (5)	His64	1CIN, 1CIM, 1CIL	16
CYP51	Fluconazole (28)	Several	1EA1	46
COX-1	Indomethacin (19)	Arg277, Gln358		62
COX-2	SC-558	Many residues	1CX2, 6COX	63
Cystathionine γ-synthase	CTPCO	Arg423, Tyr163	1QGN, 1I48	64
Elongation factor Tu aurodox	Aurodox	His85 and other extensive changes	1EXM, 1HA3	50
Oestrogen receptor	Tamoxifen Raloxifene	Many residues, especially helix 12 Leu540	3ERT 1A52, 1ERR	65 66
Glyceraldehyde-3-phosphate dehydrogenase	Adenosine analogues	Met39 and many N-terminal residues	1EVY, 1I32, 1I33	67
Glycogen phosphorylase	CP-526,423	Arg60, Val64, Lys191	1EXV, 1EM6	68
HIV-RT	Nevirapine (9), 1051U91, 9-Cl-TIBO (11), α-APA, HEPT	Tyr181, Tyr183, Tyr188	1VRT, 1TVR 1EP4	69
HIV protease	Cyclic ureas Hydroxyethyl amine dipeptide isostere Bis tert-amides	Asp29, Asp30, Gly48 Asp29, Asp30, Thr80, Val82	1HWR, 1HVR, 1QBS, 1DMP, 1QBR, 1QBT 1IDA, 1IDB	70 71 72
HMG-CoA reductase	Statins	Many residues, including Leu562, Val683, Leu853, Ala856 and Leu857	11HW8, 1HW9, 1HWI, 1HWJ, 1HWK, 1HWL	73
MMP (*Bacteriodes fragilis*)	SB225666	Trp49		74
MMP-3	Hydroxamates	Loop residues 222–231	456C, 830C, 966C	75
MMP-2	SC-74020	Phe148, Arg149	1HOV	76
MMP-8	Barbiturate	Loop Ala206–Asn218	1JJ9	77
eNOS	3-Bromo-7-nitroindazole	Glu363	2NSE, 1D0O	78
NAD(P)H:quinone oxidoreductase	ARH019	Phe106, Tyr128′, Phe232′	1D4A, 1H66, 1GG5, 1H69	79
Neuraminidase	Compound (23)	Glu276	1BJI	39
MAP p38 kinase	SB1-4 BIRB 796 (16)	Tyr35 Asp168, Phe169, Gly170	1A9U,1BL6, 1BMK, 1BL7 1KV2	80 31
P450 BM3	Substrates	Many residues	Not deposited	81
$P450_{cam}$	Antifungal imidazole Phenylimidazoles	Tyr96, Phe193 Helix 1 and Tyr96	1PHA, 1PHB 1PHD, 1PHE, 1PHF, 1PHG, 2CPP	82 83
Renin	Compounds (1) and (2)	Many residues, especially Trp39, Tyr75	1RNE	4
Secretory phospholipase A_2	LY315920	His64	1DB4, 1DB5,1DCY	84
Thrombin	LY178550 Benzo[b]thiophene Diphenylalanine dipeptides	Trp60, His57 Trp60D, insertion loop Trp60D, insertion loop	Native, 1D4P 1D3D, 1D4P Not deposited	85 86 87
Thymidylate synthase	Phenolphthalein analogues 1843U89	Arg23 Many residues, especially Ile79, Phe176	1TSL, 1TSM 1TSD, 2TSC	88 5

*Number refers to compound number in the text. COX, cyclooxygenase; CTPCO, 5-carboxymethylthio-3-(3′-chlorophenyl)-1,2,4-oxadiazol; CYP, cytochrome P450; eNOS, endothelial nitric oxide synthase; HIV, human immunodeficiency virus; HIV-RT, HIV reverse transcriptase; HMG, 3-hydroxy-3-methyl-glutaryl; MAP, mitogen-activated protein; MMP, matrix metalloproteinase; NAD(P)H, nicotinamide-adenine dinucleotide (phosphate) (reduced); PDB, Protein Data Bank.

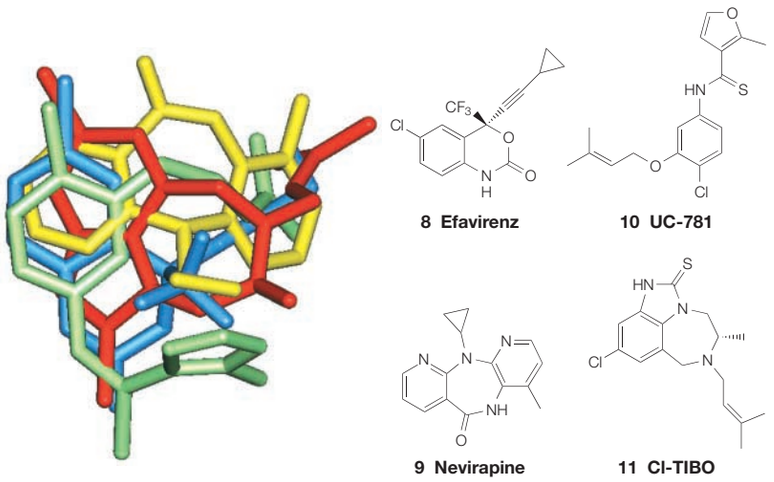


Figure 8 | **Structurally diverse ligands in a single binding site.** Overlay of four human immunodeficiency virus reverse transcriptase inhibitor complexes using the protein C-α atoms efavirenz (blue), nevirapine (yellow), UC-781 (green), Cl-TIBO (red) (Protein Data Bank codes 1FK9, 1VRT, 1JLG and 1TVR). These structurally diverse inhibitors occupy the same volume in the binding site.

which are part of the polymerase active site, are contained within strands four, seven and eight. These move as a single CHU up to 2 Å when compared with the native structure[26]. The allosteric binding pocket is described as being the palm of a hand for which the two domains form a thumb and fingers.

In another example, CP320626 (14; IC_{50} = 0.2 μM) (FIG. 6b) binds to glycogen phosphorylase b at a subunit interface in the centre of this dimeric enzyme[27]. Before the inhibitor binds, the site contains thirty molecules of water per subunit, enclosed in a cavity capped by α-helices at either end. As a result of binding, nine water molecules are displaced. The inhibitor is 83% buried with a total surface area of 536 $Å^2$ and becomes inaccessible to solvent. The binding site is 33 Å from the catalytic site and 15 Å from the allosteric site for AMP. The AMP site is also the binding site of the inhibitor Bay W1807 (15; IC_{50} = 2 nM) (FIG. 6b) (REF. 28). Likewise, allosteric inhibitors of fructose 1,6-bisphosphate[29] and haemoglobin[30] bind at the interfaces between subunits of multimeric proteins.

Binding kinetics. If a compound is binding to a conformer that is of higher energy and therefore only sampled occasionally, it would be expected that the rate of association would be decreased. This is represented by the barrier $\Delta G_{barrier}$ in FIG. 1. In the case of the p38 MAP kinase inhibitor BIRB 796 (16; K_d = 0.1 nM) (FIG. 6c), a loop positioned near the mouth of the ATP pocket, between the two domains of the enzyme, samples both 'in' and 'out' conformations[31]. The loop consists of Asp-Phe-Gly, and usually takes up a conformation in which the central Phe is located inside the hydrophobic binding pocket. For the out conformation, to which the inhibitor binds, this Phe is solvent exposed and has moved by ~10 Å. The rate of association (k_{on} M^{-1} s^{-1}) for BIRB 796 is 8.5 × 10^5, whereas for other inhibitors not requiring this conformational change to bind, k_{on} = 4.3 × 10^7. In a binding screen, BIRB 796 requires two hours pre-incubation before a maximal inhibitor potency is observed, but with inhibitors that bind to lower energy conformers, equilibrium is reached within 30 minutes. As a large hydrophobic surface area is buried, BIRB 796 is also slow to come off the enzyme, k_{off} = 8.3 × 10^{-6}. This value is low, with more typical inhibitors having k_{off} ~1. Slow kinetics can translate into extended dosing protocols and improve absorption, distribution, metabolism and excretion properties.

Slow inhibition kinetics have been observed for the marketed acetylcholinesterase inhibitor huperzine[32] (17) (FIG. 6d). This compound binds to the enzyme with Trp84 and Phe330 in different conformations to that observed in the native enzyme, and derives its affinity largely through hydrophobic collapse of the protein around the ligand and cation–π interactions. It shows k_{on} = 1 × 10^6 and k_{off} ~2 × 10^{-3}, which is in marked contrast to other inhibitors of similar potency which display rapid on/off rates. Oseltamivir carboxylate (18) (FIG. 6e) — a neuraminidase inhibitor with IC_{50} = 1 nM — also derives much of its affinity through hydrophobic fit, and shows slow, biphasic binding kinetics because of the time taken in reorientation of the Glu276 side chain[33,34].

Inhibitors of cyclooxygenase-1 (COX-1), such as indomethacin (19) (FIG. 6f), show a rapid and reversible initial binding to the enzyme followed by a slow first-order decay of activity of the synthase–inhibitor complex[35]. In the final complex, Arg277 and Gln358 have been shown to be in new positions relative to that in the native enzyme. Celecoxib (20) (FIG. 6g) displays an initial competitive interaction[36] with COX-2 (K_i =13 μM) followed by reversible but time-dependent inactivation k_{inact} = 0.08 s^{-1}. COX inhibitors show the full range of kinetic behaviours: competitive (ibuprofen), weak binding time-dependent (naproxen), tight binding time-dependent (indomethacin) and irreversible (aspirin). The apparent inhibition constants are strongly dependent on the assay conditions, making simple comparisons between the IC_{50} values for the different isoforms of questionable validity.

Selectivity between related receptors. Many therapeutic targets are single members of a family of highly related proteins, so selective inhibition is often required. This selectivity can be obtained as a consequence of protein movement. Highly conserved sites can have the same residues located in almost the same positions when single, low-energy conformations are compared. This is a consequence of the restraints imposed on the site in order to bind the same natural ligand or carry out the same chemistry. However, non-binding-site residues are usually less homologous, which allows different types or degrees of motion to take place. So individual conformers of two otherwise rather similar proteins can be quite different and these differences can be exploited to obtain selective inhibitors.

Protein flexibility has been highlighted as the key to obtaining bacterial-specific thymidylate synthase inhibitors[37]. The binding sites of this essential enzyme

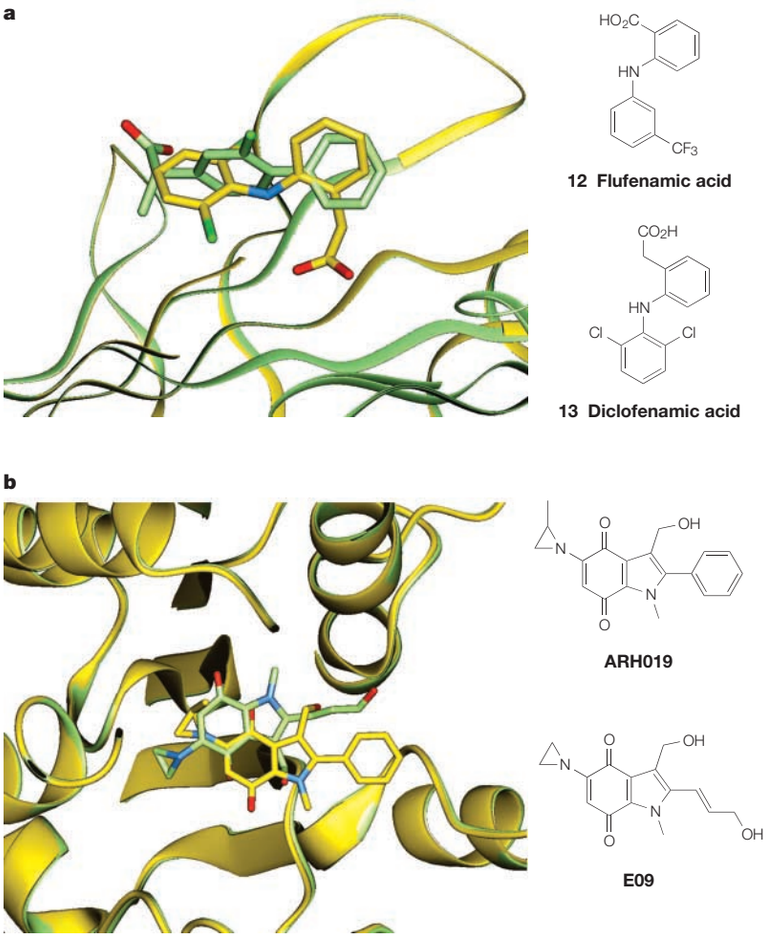


Figure 9 | **Surprising orientations of structurally similar ligands. a** | Overlay of the complex between transthyretin and flufenamic acid (green) with the complex of transthyretin and diclofenamic acid (yellow). The relative orientations are surprising: the carboxyl groups protrude either into solvent or directly down into the channel (Protein Data Bank codes 1DVT and 1DVX). **b** | Overlay of the complexes between nicotinamide-adenine dinucleotide (phosphate) (reduced):quinone oxidoreductase 1 with inhibitors AHR019 (yellow) and E09 (green). The relative orientation of the indole rings is surprising considering the structural similarities between the inhibitors (Protein Data Bank codes 1GG5 and 1H69).

are highly conserved across species. The complex between (21) (FIG. 6h) and *Escherichia coli* thymidylate synthase (EcTS) was determined at 2.08 Å resolution and compared with the complex of an unselective compound (22) (FIG. 6h) with the same enzyme. Compound (22) harnesses the protein flexibility inherent in the enzyme's catalytic mechanism, by binding to the active site in a more open form of the enzyme. The specificity is primarily the result of van der Waals contacts with specificity residues — Ile55, Trp83 and Val262 — which are not conserved between EcTS and the human orthologue.

Thymidylate synthase catalyses two separate reaction steps involving covalent bond formation. This requires reorientation of the substrate and cofactor at various stages in the enzymatic cycle. A movie derived from study of the enzyme during various parts of its reaction pathway is available (TABLE 2, which is linked online)[38]. Protein mobility can also result in resistance to inhibition of one isoform when it is desirable for both to be inhibited. For instance, the active sites of influenza A and B neuraminidases are composed of eleven conserved residues and compounds that inhibit both forms are required. However, compounds such as (23) (FIG. 6i) show potent inhibition of only the A form ($IC_{50} = 2$ nM) with little inhibition of the B form ($IC_{50} = 4$ μM)[39]. Binding of (23) is accompanied by movement of Glu276 in both the A and B forms, in which it forms a new salt bridge with Arg224. These side chains are also present in both forms and are located in almost exactly the same position in both the sialic acid and inhibitor-bound complexes. However, the B form seems to pay a much heavier energetic penalty for the movement when binding the inhibitor. Non-conserved residues further from the binding pocket, particularly at positions 240 and 405, are larger in the B form, which constrains the movement of Glu276.

Proteins involved in metabolism and transport. Proteins involved in metabolism are required to bind to a wide range of substrates, which is facilitated by protein mobility. These promiscuous proteins have often been observed binding diverse compounds, or a single molecule, in several different orientations. Multiple binding modes have been seen for cytochrome $P450_{cam}$[40]. Camphor (24) (FIG. 6j) is hydroxylated to a single product. Thiocamphor (25) (FIG. 6j) is observed to bind in two orientations, both of which are different to that observed for camphor. The major products of its metabolism are the 5- and 6-exo hydroxylated materials. However, in the complex with $P450_{cam}$, these atoms are farthest from the oxygen activated by the iron, indicating that these might be unproductive complexes and that other binding orientations are possible.

Another example of directed promiscuity is provided by the pregnane X receptor (PXR), a ligand-activated transcription factor that controls expression of the genes for the cytochrome P450 CYP3A enzymes[41]. The structure of the complex between PXR and the cholesterol-lowering drug SR12813 (26) (FIG. 6j) ($K_d = 41$ nM for this receptor) was determined at 2.75 Å. SR12813 was observed bound to PXR in three distinct orientations. The binding site of PXR accommodates a wide range of ligands using several mobile loops around the binding cavity, which is largely smooth and hydrophobic with a small number of polar residues spaced around it. Interestingly, the receptor is, at least by sequence comparisons, similar to known nuclear hormone receptors, which are themselves highly selective for their cognate hormone. This highlights the difficulties involved in predicting flexibility, selectivity and even function from simple sequence comparisons.

Similar conclusions have been drawn for non-specific transporters. For instance, the multi-drug binding protein QacR was determined complexed with six different cationic ligands and shows binding site flexibility[42]. As most docking and scoring algorithms are usually based on rigid receptor assumptions, it might be anticipated

Table 2 | **Movies of conformational changes in proteins**

Protein	Motion	URL	Comment
Website-based movies			
Dihydrofolate reductase[89]	Complex	http://chem-faculty.ucsd.edu/kraut/dhfr.html	This movie of the conformational changes that this enzyme undergoes during its catalytic cycle has been assembled from isomorphous crystal structures analogous to the five reaction intermediates and the transition state. Movements are illustrated in loop M20, rotation of the two halves of the enzyme and other conformational changes during cofactor and substrate capture followed by product release.
Thymidylate synthase[38]	Complex	http://www.msg.ucsf.edu/stroud/tsmovie.html	A movie of the conformational changes that occur during the enzymatic cycle constructed from eight X-ray crystallographic determinations at various points along the reaction pathway, together with detailed mechanistic and kinetic studies. The way in which conformational changes are coupled to function is examined in detail.
Acetylcholin-esterase	Complex	http://www.weizmann.ac.il/~joel/movies/millvx_full.mov	A movie of the conformational changes in the active site residues of acetylcholinesterase as it binds and processes its substrate.
Movies generated from a database of protein movements available on the web*			
Maltose	Domain motion: hinge mechanism	http://www.nature.com/nrd/journal/v2/n7/extref/nrd1129-s4.mpg	Maltose being engulfed by its binding protein. Ligand–protein X-ray complexes often show the ligand buried within the protein. This represents a 'flash-photograph' at the end of a series of events that together represent the total binding process.
HIV protease	Fragment motion: hinge mechanism	http://www.nature.com/nrd/journal/v2/n7/extref/nrd1129-s2.mpg	
Acetylcholin-esterase	Domain motion: hinge mechanism	http://www.nature.com/nrd/journal/v2/n7/extref/nrd1129-s1.mpg	
Thymidylate synthase	Fragment motion: shear mechanism	http://www.nature.com/nrd/journal/v2/n7/extref/nrd1129-s6.mpg	
P450BM-3	Domain motion: predominantly shear	http://www.nature.com/nrd/journal/v2/n7/extref/nrd1129-s5.mpg	
HIV-RT	Domain motion: partial refolding of tertiary structure	http://www.nature.com/nrd/journal/v2/n7/extref/nrd1129-s3.mpg	

*The construction and format of the database (see Further information, MolMovDB) are described in REFS 90, 91. Many existing movies and the software tools for constructing new movies are freely available. HIV, human immunodeficiency virus; HIV-RT, HIV-reverse transcriptase.

that any attempt at '*in silico*' metabolism predictions will meet with considerable difficulties without access to direct experimental information on relevant, accessible protein conformations.

Exploiting protein mobility in drug design

Library and inhibitor design. Usually, the opening of a binding pocket exposes non-polar residues to solvent and is likely to be energetically demanding. Some studies show that if the exposed non-polar surface area is large in the open conformation, the extent of opening of the domains is smaller[43]. This places a limit on the size of putative ligands during library design. As the size and shape of a binding site is to some extent dependent on the ligand, libraries can be viewed as offering a range of molecules to the receptor with lead-like physical properties[44]. The flexible binding site is then allowed to select a suitable lead, one possessing properties that allow scope for further elaboration into an orally dosed drug.

Presently, enzyme inhibitor design is frequently modelled using transition-state mimics. Successful drug design, however, usually requires deriving a large interaction (ΔG_{obs}; FIG. 1) from a modestly sized drug molecule. Molecules that derive their affinity by mimicking large natural substrates are often unsuccessful. Although they typically have high affinities, these are derived from a large number of low-affinity interactions. The result is often a compound with high molecular mass and lipophilicity, giving rise to poor absorption, distribution, metabolism and toxicity characteristics. However, protein flexibility makes it possible to exploit high-affinity interactions and so obtain a high affinity from a small drug. Some of the required affinity can be derived from changes in the protein rather than simply from interactions at the drug–protein interface. The fact that structurally diverse inhibitors bind at the same site also provides scope for competing organizations to obtain separate intellectual property claims.

Forcing new conformations. Deriving a high-affinity compound by forming a complex with an alternative receptor conformation is highly attractive. Predicting

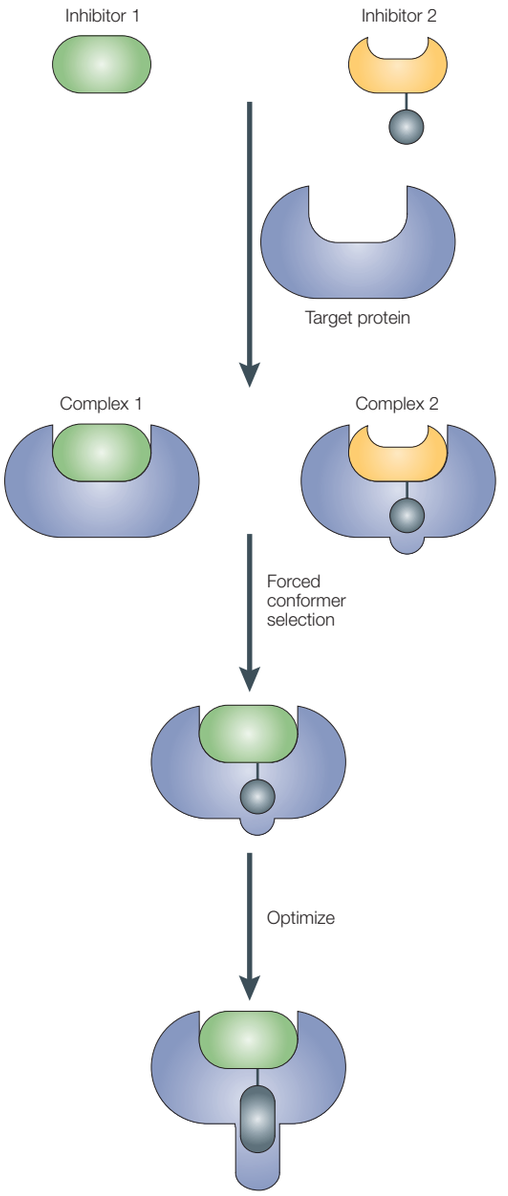


Figure 10 | **Forcing new conformations.** A target protein is crystallized with two inhibitors to generate complexes with micromolar affinities. One inhibitor (2) results from the interception of a minor conformer. The information that certain residues are flexible can then be exploited by further elaborating inhibitor 1. The process rather perversely involves deliberately, but intelligently, engineering steric clashes into complexes in which a ligand has previously been shown to fit snugly.

protein flexibility, however, is difficult. Sometimes sufficient structural information is available to aid in the discovery process (FIG. 10).

New side-chain conformations can be targeted for chosen residues. These can be prioritized on the basis of the study of complexes with diverse ligands, involvement in function, lipophilicity and position on loops or at the faces of CHUs. Main-chain B-FACTORS are not reliable predictors of flexibility. To a great extent, they result from small (~1 Å) excursions of the backbone on the pico- to microsecond timescale. Motions involved in binding and catalysis are typically on the micro- to millisecond timescale. Mobile sections of a protein often have high solvent accessibility and few scaffolding hydrogen bonds between the domains. Making and breaking large numbers of hydrogen bonds can provide a barrier to conformer interconversion.

Another method of searching for new conformations is a site-directed screening method based on tethering[45]. The method introduces cysteine residues proximate to the binding site and uses these to deliver putative ligands by reversible disulphide formation from a library pool of diverse disulphide monomers. X-ray crystallography of the complex between compound (27) (FIG. 6k) and interleukin-2 revealed positions on the protein that are mobile by comparing the complex with the APO FORM. Characteristically, the hydrophilic moiety of (27) bound to a relatively conformationally fixed region of the protein, and the hydrophobic moiety bound to a mobile one. Unfortunately, combinatorial libraries typically vary hydrophobic regions more than hydrophilic ones for reasons of synthetic expediency. So, libraries, each based on different amidine variations, but with more limited variation in the lipophilic parts of the library members, would be more likely to to be effective in lead discovery than a library synthesized from a single amidine template but with much more extensive variation in the lipophilic parts of the molecule. The authors were able to demonstrate the effectiveness of the technology by discovering alternative amidines that bound the hydrophilic region and fused aromatic groups that bound the mobile, hydrophobic one.

The degree of substrate specificity of a receptor is likely to be correlated with its flexibility. For example, proteolytic enzymes or relatively nonspecific endonucleases will possess a higher degree of flexibility, and, in general, enzymes with narrow substrate specificities will possess fewer and less diverse conformations. The degree of flexibility of a receptor is therefore intimately connected with its function. Likewise, residues involved in function are often mobile and tend to be observed in 'new' positions when inhibitors are bound.

Mutation resistance. Protein mobility is important to understanding mutation-derived resistance to drugs. Mutations selected to combat one drug will often confer resistance to several others of the same mechanistic class. These mutations can be located either at the binding site or a considerable distance away.

Mapping of mutations identified in azole-resistant *Candida albicans* cytochrome P450 14α-sterol demethylases (CYP51) indicates that resistance to fluconazole (28) (FIG. 6l) develops in regions involved in orchestrating the movement of CYP51 through different conformational stages, rather than simply involving residues directly contacting the inhibitor[46]. These mutations render the enzyme too stiff to accommodate the inhibitor at some point on its journey to the final binding site.

B-FACTORS
B-factors (also known as temperature factors) model the effect of static and dynamic disorder associated with an atom when it is observed by X-ray analysis. They are related to the atom's mean square displacement during observation.

APO FORM
Enzyme (receptor) in its free state without substrates or inhibitors bound.

Resistance mutations to inhibitors of acetylcholinesterase in insects have also been mapped. Two of the four most common resistance mutations, Gly265Ala and Ile161Val, are not within the active site but are located in the second shell of residues around it[18]. Small changes in hydrophobic packing affect the mobility of binding-site residues.

Flexibility in an inhibitor can sometimes offset mutational resistance. S-1153 (29) (FIG. 6m), an HIV-RT inhibitor, shows low susceptibility to resistance mutations. It is suggested that the relative flexibility of this inhibitor allows it to adapt more readily to mutated binding pockets[47]. The compound also uses several hydrogen-bond interactions to the amide backbone of the protein. These remain unaffected by side-chain mutations.

Likewise, resistance mutations to HIV protease have been mapped to regions that are spatially removed from the drug- or substrate-binding sites[48]. The authors conclude that a less rigid inhibitor, or one that targets the open conformation of the enzyme, would be desirable. In a remarkable study, a mutant drug-resistant HIV-1 protease was shown to have ten out of eleven mutations outside the binding pocket. These ten mutations alone stabilized the protease and inhibited it. The active-site mutation alone had only a slight inhibitory effect and increased the structural stability of the protease[49].

Usually, large inhibitors, which more completely fill the binding pocket, prove to be most susceptible to resistance mutations and genetic polymorphisms, as steric exclusion is more facile. Resistance mutations often cluster around the hydrophobic moieties of an inhibitor, indicating the importance of hydrophobic shape matching to inhibitor potency[50].

Summary

The evidence that proteins, which constitute therapeutic targets, are mobile is substantial. It is also inconvenient for medicinal chemists undertaking detailed structure–activity relationship studies of a set of structurally related ligands or attempting to model interactions by considering only the protein–ligand interface. The lipophilic groups, from which much of the interaction energy between a ligand and a protein is derived, differ from functional groups, which loom large in the psyche of synthetically trained chemists. Ligands possessing suitable physical properties for oral activity usually contain some lipophilic substituents, which also provides the opportunity for high-affinity interactions with alternative protein conformers. When planning synthetic targets using structural data, chemists should bear in mind that because of the inherent mobility of proteins, a good fit between a ligand crystallized in the binding pocket and the receptor is expected. However, this does not imply that there is no room for larger or additional substituents at certain points around the molecular scaffold. X-ray crystallographic data provide little information about affinity; therefore, observing an interaction between a particular part of a ligand and a receptor gives little clue as to its overall importance to the total affinity observed.

Binding sites are often characterized by regions of both very high and very low mobility[51]. Some common failures in ligand docking caused by protein flexibility have been discussed[52]. Ligand-design algorithms might improve their treatment of flexibility by requiring complementarity where the site is rigid and flexibility where it is not[53]. This could be more profitable than attempting to calculate precise binding geometries and interaction energies assuming a rigid binding site, as the effect on affinity of even small changes in protein geometry and water position is large[54]. Progress in improving computational methods to accommodate protein flexibility has recently been reviewed[55], as has the impact of the energy-landscape perspective on the complexity of structure prediction in molecular-recognition processes[56].

The process whereby a drug interacts with a target protein can be imagined as a collision between two flexible objects. The ligand collides with the protein surface and displaces loosely bound water molecules from its surface, but does not diffuse away immediately, as would be the case for a simple elastic collision. Instead, the ligand and protein roll on each other and sample surface area and conformational space, with suitable substrates being transported towards the active site with a matching process occurring between various conformations of the ligand and protein.

Proteins exist as an ensemble of conformers. These have similar energies, and proteins have evolved such that they flip between these conformations in a facile way. Modern definitions of agonism, antagonism and allosteric modulation incorporate the concept of multiple conformations. The binding of structurally diverse inhibitors at the same binding site, and the surprising binding orientations of similar ligands, derives from protein flexibility. Slow binding kinetics and exquisite selectivity between highly related receptors are also the result of protein flexibility. The promiscuous proteins involved in metabolism and transport require highly flexible binding sites. The inherent flexibility of proteins can be exploited in drug design and is likely to continue to play an important role in the understanding and design of future pharmaceutical molecules.

1. Bursavich, M. G. & Rich, D. H. Designing non-peptide peptidomimetics in the 21st century: inhibitors targeting conformational ensembles. *J. Med. Chem.* **45**, 541–558 (2002).
2. Marvin, J. S. & Hellinga, H. W. Manipulation of ligand binding affinity by exploitation of conformational coupling. *Nature Struct. Biol.* **8**, 795–798 (2001).
3. Baldwin, R. L. Making a network of hydophobic clusters. *Science* **295**, 1657–1658 (2002).
An excellent overview of the role of compact hydrophobic domains in protein folding.
4. Oefner, C. *et al.* Renin inhibition by substituted piperidines: a novel paradigm for the inhibition of monomeric aspartyl proteases. *Chem. Biol.* **6**, 127–131 (1999).
5. Weichsel, A. & Montford, W. R. Ligand-induced distortion of an active site in thymidylate synthase upon binding anticancer drug 1843U89. *Nature Struct. Biol.* **2**, 1095–1101 (1995).
6. Falke, J. J. A moving story. *Science* **295**, 1480–1481 (2002).
A summary of the relationship between mobility and function and its study by various methods.
7. Eisenmesser, E. Z., Bosco, D. A., Akke, M. & Kern, D. Enzyme dynamics during catalysis. *Science* **295**, 1520–1523 (2002).
8. Lee, A. Y., Gulnik, S. V. & Erickson, J. W. Conformational switching in an aspartic proteinase. *Nature Struct. Biol.* **5**, 866–871 (1998).

9. Rahuel, J., Priestle, J. P. & Grutter, M. G. The crystal structures of recombinant glycosylated human renin alone and in complex with a transition state analogue inhibitor. *J. Struct. Biol.* **107**, 227–236 (1991).
10. Kenakin, T. Agonist receptor efficacy 1: mechanisms of efficacy and receptor promiscuity. *Trends Pharmacol. Sci.* **16**, 188–192 (1995).
A useful and brief account of the importance of conformational ensembles in pharmacology.
11. Ma, B., Shatsky, M., Wolfson, H. J. & Nussinov, R. Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci.* **11**, 184–197 (2002).
A key paper giving an account of ligand-binding phenomena from the perspective of lessons learned in the study of protein folding.
12. Shoichet, B. K., Baase, W. A., Kuroki, R. & Matthews, B. A relationship between protein stability and protein function. *Proc. Natl Acad. Sci. USA* **92**, 452–456 (1995).
13. Gerstein, M., Lesk, A. M. & Chothia, C. Structural mechanisms for domain movements in proteins. *Biochemistry* **33**, 6739–6749 (1994).
14. Davis, A. M. & Teague, S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed. Engl.* **38**, 736–749 (1999).
15. Kryger, G., Silman, I. & Sussman, J. L. Structure of acetylcholinesterase complexed with E2020 (Aricept): implications for the design of new anti-Alzheimer drugs. *Structure* **7**, 297–307 (1999).
An outstanding account of the insights gained from the complex of a commercially important inhibitor with its target protein from the leading group in cholinesterase research.
16. Smith, G. M. *et al.* Positions of His-64 and a bound water in human carbonic anhydrase II upon binding three structurally related inhibitors. *Protein Sci.* **3**, 118–125 (1994).
17. Baldwin, J. J. *et al.* Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. *J. Med. Chem.* **32**, 2510–2513 (1989).
References 16 and 17 describe an early, successful and little-publicised account of the discovery of an important class of drugs using crystallography. The account provides many lessons including mutual fit of both ligand and drug, and the introduction of an amine to improve pharmaceutical properties.
18. Harel, M. *et al.* Three-dimensional structures of *Drosophila melanogaster* acetylcholinesterase and of its complexes with two potent inhibitors. *Protein Sci.* **9**, 1063–1072 (2000).
19. Zidek, L., Novotny, M. V. & Stone, M. J. Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nature Struct. Biol.* **6**, 1118–1121 (1999).
20. Yaremchuk, A., Tukalo, M., Grotli, M. & Cusack, S. A succession of substrate induced conformational changes ensures the amino acid specificity of *Thermus thermophilus* prolyl-tRNA synthetase: comparison with histidyl-tRNA synthetase. *J. Mol. Biol.* **309**, 989–1002 (2001).
21. Velazquez-Campoy, A., Luque, I. & Freire, E. The application of thermodynamic methods in drug design. *Thermochimica Acta* **380**, 217–227 (2001).
22. Najmanovich, R., Kuttner, J., Sobolev, V. & Edelman, M. Side-chain flexibility in proteins upon ligand binding. *Proteins* **39**, 261–268 (2000).
23. Ren, J. *et al.* High resolution structures of HIV-1RT from four RT–inhibitor complexes. *Nature Struct. Biol.* **2**, 293–302 (1995).
24. Klabunde, T. *et al*. Rational design of potent human transthyretin amyloid disease inhibitors. *Nature Struct. Biol.* **7**, 312–321 (2000).
25. Smerdon, S. J. *et al*. Structure of the binding site for nonnucleoside inhibitors of the reverse transcriptase of human immunodeficiency virus type 1. *Proc. Natl Acad. Sci. USA* **91**, 3911–3915 (1994).
26. Esnouf, R. *et al*. Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nature Struct. Biol.* **2**, 303–308 (1995).
A beautiful, visually appealing and detailed account of the structural basis of allosteric modulation in these inhibitors.
27. Oikonomakos, N. G., Skamnaki, V. T., Tsitsanou, K. E., Gavalas, N. G. & Johnson, L. N. A new allosteric site in glycogen phosphorylase b as a target for drug interactions. *Structure* **8**, 575–584 (2000).
28. Zographos, S. E. *et al*. The structure of glycogen phosphorylase b with an alkyl-dihydropyridine–dicarboxylic acid compound, a novel and potent inhibitor. *Structure* **5**, 1413–1425 (1997).
29. Wright, S. W. *et al.* Anilinoquinazoline inhibitors of fructose 1,6-bisphosphatase bind at a novel allosteric site: synthesis, *in vitro* characterization, and X-ray crystallography. *J. Med. Chem.* **45**, 3865–3877 (2002).
30. Abraham, D. J. *et al*. Allosteric modifiers of hemoglobin: 2-[4-[[(3,5-disubstituted anilino)carbonyl]methyl]phenoxy]-2-methylpropionic acid derivatives that lower the oxygen affinity of hemoglobin in red cell suspensions, in whole blood, and *in vivo* in rats. *Biochemistry* **31**, 9141–9149 (1992).
31. Pargellis, C. *et al.* Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nature Struct. Biol.* **9**, 268–272 (2002).
A key paper for a detailed account of the relationship between protein mobility and kinetics.
32. Ashani, Y., Peggins, J. O. III. & Doctor, B. P. Mechanism of inhibition of cholinesterases by huperzine A. *Biochem. Biophys. Res. Commun.* **184**, 719–726 (1992).
33. Graul, A., Leeson, P. A. & Castaner, J. Oseltamivir phosphate: anti-influenza, neuraminidase (sialidase) inhibitor. *Drugs Future* **24**, 1189–1202 (1999).
34. Wang, M. Z., Tai, C. Y. & Mendel, D. B. Mechanism by which mutations at His274 alter sensitivity of influenza A virus N1 neuraminidase to oseltamivir carboxylate and zanamivir. *Antimicrob. Agents Chemother.* **46**, 3809–3816 (2002).
35. Kulmacz, R. & Lands, W. E. M. Stoichiometry and kinetics of the interaction of prostaglandin H synthase with anti-inflammatory agents. *J. Biol. Chem.* **260**, 12572–12578 (1985).
36. Gierse, J. K., Koboldt, C. M., Walker, M. C. Seibert, K. & Isakson, P. C. Kinetic basis for selective inhibition of cyclo-oxygenases. *Biochem. J.* **339**, 607–614 (1999).
37. Fritz, T. A., Tondi, D., Finer-Moore, J. S., Costi, M. P. & Stroud, R. M. Predicting and harnessing protein flexibility in the design of species-specific inhibitors of thymidylate synthase 1,2. *Chem. Biol.* **8**, 981–995 (2001).
An insightful paper from one of the leading groups in rational design.
38. Stroud, R. M. & Finer-Moore, J. S. Conformational dynamics along an enzymatic reaction pathway: thymidylate synthase, 'the movie'. *Biochemistry* **42**, 239–247 (2003).
39. Taylor, N. R. *et al.* Dihydropyrancarboxamides related to zanamivir: a new series of inhibitors of influenza virus sialidases. 2. Crystallographic and molecular modeling study of complexes of 4-amino-4H-pyran-6-carboxamides and sialidase from influenza virus types A and B. *J. Med. Chem.* **41**, 798–807 (1998).
Probably the best account in the literature of the rationalization of selectivity between related receptors.
40. Raag, R. & Poulos, T. L. Crystal structures of cytochrome P-450CAM complexed with camphane, thiocamphor, and adamantane: factors controlling P-450 substrate hydroxylation. *Biochemistry* **30**, 2674–2684 (1991).
41. Watkins, R. E. *et al.* The human nuclear xenobiotic receptor PXR: Structural determinants of directed promiscuity. *Science* **292**, 2329–2333 (2001).
42. Schumacher, M. A. *et al*. Structural mechanisms of QacR induction and multidrug recognition. *Science* **294**, 2158–2163 (2001).
43. Sinha, N., Tsai, C.-J. & Nussinov, R. Building blocks, hinge bending motions and protein topology. *J. Biomol. Struct. Dyn.* **19**, 369–380 (2001).
An important account of the relationship between hydrophobic domains, mobility, function and topology.
44. Teague, S. J., Davis, A. M., Leeson, P. D. & Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed. Engl.* **38**, 3743–3748 (1999).
45. Arkin, M. R. *et al.* Binding of small molecules to an adaptive protein–protein interface. *Proc. Natl Acad. Sci. USA* **100**, 1603–1608 (2003).
An interesting proof of principle using a significant pharmaceutical target.
46. Podust, L. M., Poulos, T. L. & Waterman, M. R. Crystal structure of cytochrome P450 14α-sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors. *Proc. Natl Acad. Sci. USA* **98**, 3068–3073 (2001).
47. Ren, J. *et al*. Binding of the second generation non-nucleoside inhibitor S-1153 to HIV-1 reverse transcriptase involves extensive main chain hydrogen bonding. *J. Biol. Chem.* **275**, 14316–14320 (2000).
48. Rose, R. B., Craik, C. S. & Stroud, R. M. Domain flexibility in retroviral proteases: structural implications for drug resistant mutations. *Biochemistry* **37**, 2607–2621 (1998).
49. 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**, 631–638 (2003).
50. Vogeley, L., Palm, G J., Mesters, J R. & Hilgenfeld, R. Conformational change of elongation factor Tu (EF-Tu) induced by antibiotic binding. Crystal structure of the complex between EF–Tu, GDP and aurodox. *J. Biol. Chem.* **276**, 17149–17155 (2001).
51. Luque, I. & Freire, E. Structural stability of binding sites: consequences for binding affinity and allosteric effects. *Proteins* **4**, S63–S71 (2000).
An insightful paper presenting the link between protein stability and function for a number of therapeutically important proteins.
52. Verkhivker, G. M. *et al*. Deciphering common failures in molecular docking of ligand-protein complexes. *J. Comput. Aided Mol. Des.* **14**, 731–751 (2000).
53. Morton, A. & Matthews, B. W. Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: linkage of dynamics and structural plasticity. *Biochemistry* **34**, 8576–8588 (1995).
54. Murray, C. W., Baxter, C. A. & Frenkel, A. D. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. *J. Comput. Aided Mol. Des.* **13**, 547–562 (1999).
An honest and thorough account of problems with docking and scoring algorithms. Refreshing in an area which has often been oversold.
55. Carlson, H. A. Protein flexibility and drug design. How to hit a moving target. *Curr. Opin. Chem. Biol.* **6**, 447–452 (2002).
56. Verkhivker, G. M. *et al*. Complexity and simplicity of ligand-macromolecule interactions: the energy landscape perspective. *Curr. Opin. Struct. Biol.* **12**, 197–203 (2002).
57. Raves, M. L. *et al*. Structure of acetylcholinesterase complexed with the nootropic alkaloid, (–)-huperzine A. *Nature Struct. Biol.* **4**, 57–63 (1997).
58. Wilson, D. K., Tarle, I., Petrash, J. M. & Quiocho, F. A. Refined 1.8 Å structure of human aldose reductase complexed with the potent inhibitor zopolrestat. *Proc. Natl Acad. Sci. USA* **90**, 9847–9851 (1993).
59. Urzhumtsev, A. *et al*. A 'specificity' pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil. *Structure* **5**, 601–612 (1997).
60. Todone, F. *et al*. Active site plasticity in D-amino acid oxidase: a crystallographic analysis. *Biochemistry* **36**, 5853–5860 (1997).
61. Vandonselaar, M., Hickie, R. A., Quail, J. W. & Delbaere, L. T. J. Induction of a major conformational change in calcium–calmodulin (Ca^{2+}–CAM) by trifluoperazine (TFP). *Protein Eng.* **10**, S47 (1997).
62. Pouplana, R., Perez, C., Sanchez, J., Lozano, J. J. & Puig-Parellada, P. The structural and electronic factors that contribute affinity for the time-dependent inhibition of PGHS-1 by indomethacin, diclofenac and fenamates. *J. Comput. Aided Mol. Des.* **13**, 297–313 (1999).
63. Kurumbail, R. G. *et al*. Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature* **384**, 644–648 (1996).
64. Steegborn, C., Laber, B., Messerschmidt, A., Huber, R. & Clausen, T. Crystal structures of cystathionine γ-synthase inhibitor complexes rationalize the increased affinity of a novel inhibitor. *J. Mol. Biol.* **311**, 789–801 (2001).
65. Shiau, A K. *et al*. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95**, 927–937 (1998).
66. Brzozowski, A M. *et al*. Molecular basis of agonism and antagonism in the estrogen receptor. *Nature* **389**, 753–758 (1997).
67. Suresh, S. *et al*. Conformational changes in *Leishmania mexicana* glyceraldehyde-3-phosphate dehydrogenase induced by designed inhibitors. *J. Mol. Biol.* **309**, 423–435 (2001).
68. Rath, V. L. *et al*. Human liver glycogen phosphorylase inhibitors bind at a new allosteric site. *Chem. Biol.* **7**, 677–682 (2000).
69. Hopkins, A. L. *et al*. Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT series reveal conformational changes relevant to the design of potent non-nucleoside inhibitors. *J. Med. Chem.* **39**, 1589–1600 (1996).
70. Ala, P. J. *et al*. Molecular recognition of cyclic urea HIV-1 protease inhibitors. *J. Biol. Chem.* **273**, 12325–12331 (1998).
71. Tong, L. *et al*. Crystal structures of HIV-2 protease in complex with inhibitors containing the hydroxyethylamine dipeptide isostere. *Structure* **3**, 33–40 (1995).
72. Melnick, M. *et al.* Bis tertiary amide inhibitors of the HIV-1 protease generated via protein structure-based iterative design. *J. Med. Chem.* **39**, 2795–2811 (1996).
73. Istvan, E. S. & Deisenhofer, J. Structural mechanism for statin inhibition of HMG-CoA reductase. *Science* **292**, 1160–1164 (2001).
74. Scrofani, S. D. B. *et al*. NMR characterization of the metallo-β-lactamase from *Bacteroides fragilis* and its interaction with a tight-binding inhibitor: role of an active-site loop. *Biochemistry* **38**, 14507–14514 (1999).
75. Lovejoy, B. *et al*. Crystal structures of MMP-1 and -13 reveal the structural basis for selectivity of collagenase inhibitors. *Nature Struct. Biol.* **6**, 217–221 (1999).

76. Feng, Y. *et al*. Solution structure and backbone dynamics of the catalytic domain of matrix metalloproteinase-2 complexed with a hydroxamic acid inhibitor. *Biochim. Biophys. Acta* **1598**, 10–23 (2002).
77. Brandstetter, H. *et al*. The 1.8- Å. crystal structure of a matrix metalloproteinase 8–barbiturate inhibitor complex reveals a previously unobserved mechanism for collagenase substrate recognition. *J. Biol. Chem.* **276**, 17405–17412 (2001).
78. Raman, C. S. *et al*. Crystal structure of nitric oxide synthase bound to nitro indazole reveals a novel inactivation mechanism. *Biochemistry* **40**, 13448–13455 (2001).
79. Faig, M. *et al*. Structure-based development of anticancer drugs. Complexes of NAD(P)H:quinone oxidoreductase 1 with chemotherapeutic quinones. *Structure* **9**, 659–667 (2001).
80. Wang, Z. *et al*. Structural basis of inhibitor selectivity in MAP kinases. *Structure* **6**, 1117–1128 (1998).
A good account of both the flexibility of Tyr side chain as different inhibitors bind and the basis of the compound's selectivity for p38 against extracellular signal-regulated kinase.
81. Modi, S., Sutcliffe, M. J., Primrose, W. U., Lian, L.–Y. & Roberts, G. C. K. The catalytic mechanism of cytochrome P450 BM3 involves a 6 Å movement of the bound substrate on reduction. *Nature Struct. Biol.* **3**, 414–417 (1996).
82. Raag, R., Li, H., Jones, B. C. & Poulos, T. L. Inhibitor-induced conformational change in cytochrome P-450cam. *Biochemistry* **32**, 4571–4578 (1993).
83. Poulos, T. L. & Howard, A. J. Crystal structures of metyrapone and phenylimidazole inhibited complexes of cytochrome P-450cam. *Biochemistry* **26**, 8165–8174 (1987).
84. Schevitz, R. W. *et al*. Structure-based design of the first potent and selective inhibitor of human non-pancreatic secretory phospholipase A_2. *Nature Struct. Biol.* **2**, 458–465 (1995).
A nice account of structure–activity relationships by crystallography of complexes along the road to the discovery of potent agents.
85. Chirgadze, N. Y. *et al*. The crystal structure of human α-thrombin complexed with LY178550, a nonpeptidyl, active site-directed inhibitor. *Protein Sci.* **6**, 1412–1417 (1997).
86. Chirgadze, N. Y. *et al*. The crystal structures of human α-thrombin complexed with active site-directed diamino benzo-[b]thiophene derivatives: a binding mode for a structurally novel class of inhibitors. *Protein Sci.* **9**, 29–36 (2000).
An interesting account of the discovery of some surprising compounds. Significantly, reference 87 also describes compounds that interact with the insertion loop, which is known to control access of substrates to the active site.
87. Tucker, T J. *et al*. Design of highly potent noncovalent thrombin inhibitors that utilize a novel lipophilic binding pocket in the thrombin active site. *J. Med. Chem.* **40**, 830–832 (1997).
88. Stout, T. J. *et al*. Structure-based design of inhibitors specific for bacterial thymidylate synthase. *Biochemistry* **38**, 1607–1617 (1999).
89. Sawaya, M. R. & Kraut, J. Loop and sub-domain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence. *Biochemistry* **36**, 586–603 (1997).
90. Krebs, W. G. & Gerstein, M. The morph server: a standardized system for analyzing and visualizing macromolecular motions in a database framework. *Nucleic Acids Res.* **28**, 1665–1675 (2000).
91. Echols, N. D., Milburn, D. & Gerstein, M. MolMovDB: analysis and visualization of conformational change and structural flexibility. *Nucleic Acids Res*. **31**, 478–482 (2003).

Acknowledgements

I wish to acknowledge Dr A. M. Davis for invaluable discussions concerning this manuscript, and Dr S. St-Gallay for assistance with preparation of the figures.

Online links

FURTHER INFORMATION
Database of Macromolecular Movements with Associated Tools for Geometric Analysis:
http://molmovdb.mbb.yale.edu/molmovdb/
Protein Data Bank: http://www.rcsb.org/pdb/
Access to this interactive links box is free online.