Molecular engineering and design

The ability to impart any desired property to biological macromolecules, or to manipulate their activity in a predictable manner using extrinsic factors (such as small molecules), is the goal of the biomolecular engineer and drug designer. Such powers, while still far from fully realized, are starting to have an impact on a vast range of disciplines—from medicine to microelectronics. Although the goals of the new engineers and designers are ambitious (as they should be) the tools and the knowledge base they have at their disposal are, as yet, primitive. Two articles in this issue of *Nature Structural Biology* address the concerns of the engineers and designers. The first, from Johan Jansonius and colleagues, probes the molecular basis of enzyme–substrate specificity[1] and the second discusses the achievements of, and prognosis for, fighting AIDS through drugs designed against the HIV protease[2].

A pair of related enzymes, aspartate aminotransferase (AspAT) and tyrosine aminotransferase (TyrAT), have presented a fascinating challenge for those interested in understanding the basis of enzyme–substrate specificity. The two enzymes, as their names suggest, serve to shuttle an amino group between specific amino acids and the corresponding oxo-acids: for AspAT these are the dicarboxylic amino acids aspartate and glutamate; and for TyrAT they are the dicarboxylic amino acids Asp and Glu as well as the aromatic amino acids tyrosine, phenylalanine and tryptophan. How does TyrAT cope with both acidic and aromatic substrates? The eTyrAT protein has so far resisted all attempts to determine its structure, so Jansonius and colleagues have taken an indirect approach to solving the puzzle: they have chosen to analyse a hexamutant of eAspAT which has a substrate specificity more typical of eTyrAT. They have solved the structure of the mutant 'eTyrAT-like' eAspAT in isolation and in the presence of inhibitors that mimic the Asp/Glu, and the aromatic amino-acid substrates.

Reassuringly, the structure of the mutant eAspAT in the presence of an inhibitor (maleate, a dicarboxylic amino-acid analogue) is essentially the same as that of the wild-type enzyme–substrate complex. The two carboxylate groups form strong hydrogen bonds/ion pairs with the side chains of Asp 386 and Arg 292* (the asterisk indicating that the Arg 292 is from the second subunit in the homo-dimeric enzyme). The single carboxylate group of the aromatic inhibitors makes a similar contact with Asp 386 and the phenyl ring is found in the same binding pocket as the second carboxylic-acid group of the maleate inhibitor. As a result, the Arg 292* side chain is forced out of the binding pocket and into the surrounding solvent. It is thought that, in the wild type enzyme, excess substrate-binding energy is used to drive the conversion of the dimeric enzyme from the open (ligand-binding) conformation to the to the closed (catalytic) conformation. The structure of the mutant complex without ligand is already in the closed conformation and the energy of substrate binding can be used to expel the Arg 292* side chain from the binding pocket. Thus, for AspAT and TyrAT, substrate specificity is fine-tuned by residues of the 'second shell', that is, those that are remote from the active site and do not necessarily make direct contact with the substrate. The study provides yet another indication of the subtly and complexity of protein-based recognition processes.

Looked at from a theoretical perspective, almost all protein engineering can be defined as 'tinkering'. That is, the underlying structure of the protein is not significantly altered, and the new specificity, domain or quaternary structure so generated relies heavily on

the pre-existing, pre-evolved polypeptide chain. In this way questions concerning productive folding and stability are, for the most part, neatly side-stepped.

The *de novo* design of enzymes has been achieved only at the level of catalytic peptides. An example is provided by the design and construction of a 14-residue circular peptide with the ability of catalysing the decarboxylation of oxaloacetate[3]. The 'minizyme' achieves a modest rate enhancement (10^3–10^4 over the same reaction catalysed by simple amines; 20-fold over a random coil analogue of the same peptide), which, if one takes the former estimate of the rate enhancement, is in the range achieved by catalytic antibodies[4].

An understanding of the mechanisms underlying macromolecular function need not necessarily be limited to the study of biological systems. The relatively new discipline of supramolecular chemistry (born in the 1960s) is closely aligned with the experiments aimed at generating stripped-down peptide catalysts, specific receptors and ligands except that here the building blocks are for the most part small organic molecules, with the occasional metal ion thrown into the mix. One of the major themes in the field is to mimic, effectively, biological catalysis. Although the nature of the small molecules is often very different from those found in living cells, the underlying principals of catalysis will be the same, so insights gleaned here inform the research of the biological enzymologists.

Of course, if one can invoke some means of selection to generate the desired end product then the researcher is no longer limited by his or her (lack of) knowledge of the system under study—the only requirement is a robust strategy to select the desired end product and a sufficiently large pool of variation to guarantee a thorough (or at least reasonable) sampling of the 'variation' universe. But in using such methodologies one must sacrifice the intellectual satisfaction, not to mention the practical benefit, of knowing exactly how catalysis, binding and so on are occurring. This is not to say that the end products of the *in vitro* selection will not be informative—novel solutions to biological problems will probably be of great value in understanding the parameters that constrain a particular biological process.

Our vulnerability to newly emerged pathogens (AIDS, Ebola virus), as well as a raft of all-too-familiar diseases, has lent a particular urgency to the quest for ways to combat these well-known and not so well-known killers. The ability to design small molecule drugs at will against disease-related macromolecules hold out much hope for the treatment of many of these conditions.

John Erickson, in a *News & Views* article in this issue, outlines some of the problems faced by those trying to design drugs directed against the HIV protease. Perhaps one of the most striking lessons that has been learnt from HIV is the ability of the viral proteins to acquire mutations that allow them to evade the inhibitory effects of the drug. This is true even where, for HIV protease, the active site residues are highly conserved and also contribute to the dimer interface so that, one would have imagined (perhaps naively) that resistance mutations would have been difficult to acquire without seriously compromising the function of the enzyme and therefore the viability of the virus. This is not so: the various drugs both on the market and in clinical trials are notable for the speed with which resistant strains of the virus appear.

The message so far is obvious, and for those who hoped for an early and effective treatment for AIDS (and other diseases), depressing: we do not know nearly enough about the parameters that constrain the basic architecture of proteins (which provide most disease target molecules) to design effective drug treatments at will. But such a failure of comprehension should be a spur to those unravelling the fundamental nature of biomolecules. Let's hope the science funding agencies recognize this, too.

1. Malashkevich, V.N., Onuffer, J.J., Kirsch, J.F. & Jansonius, J.J. *Nature struct. Biol.* 2, 548–553 (1995).
2. Erickson, J. *Nature struct. Biol.* 2, 523–529 (1995).
3. Johnson, K., Illumine, R.K., Widmer, H. & Benner, S.A. *Nature* 365, 530–532 (1993).
4. DeGrado, W.F. *Nature* 365, 530–532 (1993).