Chapter 2 Odd Solutions

1. There is of course no 'right' answer to this. In most cases I would expect CATH and SCOP to agree. The functional annotations should also agree. The link to prodom is not very stable and may be causing problems.

3. The protein has seven exons, corresponding to the sequences
MAPSRKFFVGGNWKMNGRKQSLGELIGTLNAAKVPADT,
EVVCAPPTAYIDFARQKLDPKIAVAAQNCYKVTNGAFTGEI,
SPGMIKDCGATWVVLGHSERRHVFGESDE,
LIGQKVAHALAEGLGVIACIGEKLDEREAGITEKVVFEQTKVIA,
DNVKDWSKVVLAYEPVWAIGTGKTATPQQ,
AQEVHEKLRGWLKSNVSDAVAQSTRIIYG, and
GSVTGATCKELASQPDVDGFLVGGASLKPEFVDIINAKQ.

By looking at the locations of these exons in the structure (for example PDB file 1wyi), it can be seen that there is a reasonable correlation between exon boundary and secondary structure. None of the exon boundaries comes in the middle of regular secondary structure. Many of the boundaries occur one or two residues into a regular secondary structure element: thus the exon 1/2 boundary is just within the second strand; exon 3/4 is just within helix 4, and so on. Triosephosphate isomerase is a β-barrel and consists of alternating strand–helix (or sometimes more than one helix). One might therefore think that the exons would each code for a strand and a helix, but they don't; however, almost all code for one strand plus some helix. Some of the exon boundaries come at the 'top' of the barrel and some at the 'bottom.' Thus, overall, the hypothesis of exon boundaries matching domain boundaries is rather well supported by this example.

5. (a) Why are eukaryotic modular proteins often extracellular? There are several possible answers to this. Each is probably partly true. (i) Eukaryotic cells occurred relatively late in evolution, and the need for extracellular proteins occurred even later, especially in multicellular organisms. By this stage in evolution, the major folds and functions had stabilized, implying that the easiest way for evolution to come up with new functions was to take existing functional modules and string them together in new ways. (ii) The extracellular environment is harsher than the intracellular environment, and only certain folds are suitable for use as extracellular domains. These are the modules. (iii) Several the new functions required in eukaryotic extracellular proteins are related to formation of the extracellular matrix, attachment of cells to surfaces, and so on, which require long stringy proteins. Modular proteins are ideally suited to such functions.
(b) Why are eukaryotic extracellular proteins stabilized by disulfides? A trivial answer is that intracellular proteins cannot be stabilized by disulfides, because they would be reduced inside the cell; therefore the only possible location for disulfide-linked proteins is outside the cell. (This is not quite true: there are some disulfide-linked proteins inside cells, but very few.) A less trivial answer, as stated

in (ii) above, is that the extracellular environment is relatively harsh and variable, and therefore proteins need to work harder to remain folded and functional. Disulfides help in this regard.

7. The purpose of this question is to make you think, not necessarily to produce a single 'correct' answer. What I had in mind was that the power tools that need plugging in are analogous to membrane proteins that only function because they are embedded in a membrane that directly provides the energy for function. Suitable examples might be F_oF_1-ATPase, in which the protein functions only if there is a proton gradient; if this source of power is switched off (for example, by the addition of an uncoupler that runs down the gradient), the enzyme no longer functions. You could extend this analogy to many other examples such as symporters and antiporters.

My analogy with battery-powered tools is anything that uses a 'power pack' to make it go. The obvious analogy is enzymes that are coupled to ATP hydrolysis, including of course motor proteins. One could again extend the analogy to anything that uses 'high-energy' substrates, including the ribosome, and DNA and RNA polymerases.

N1. Five times as much: it is as simple as the ratio of K_d values.

N3. (a) $\Delta G = -RT\ln K_d = -34.2\,\text{kJ mol}^{-1}$.
(b) The correction is that described near the start of section 2.2.11. If we write

$$\Delta G = \Sigma\Delta G_i + \Delta G_{t+r} + \Delta G_{conf}$$

and then assume that in this model system there is no loss of conformational free energy on binding, then

$$\Delta G = \Sigma\Delta G_i + \Delta G_{t+r}$$

or

$$\Sigma\Delta G_i = \Delta G - \Delta G_{t+r}$$

The loss in free energy due to translation and rotation on bringing two molecules together is roughly $50\,\text{kJ mol}^{-1}$, and thus the intrinsic free energy is $50\,\text{kJ mol}^{-1}$ stronger than that in (a), or $-84.2\,\text{kJ mol}^{-1}$.
(c) The additional domain will also have an intrinsic affinity of $-84.2\,\text{kJ mol}^{-1}$. Five rotatable bonds frozen contributes $5 \times 5\,\text{kJ mol}^{-1}$, so the overall affinity is $-84.2 - 84.2 + 50 + 25\,\text{kJ mol}^{-1} = -93.4\,\text{kJ mol}^{-1}$. Translated into a dissociation constant this gives $K_d = 4 \times 10^{-17}\,\text{M}$. No contest! If what you want is strong binding, then adding an extra pair of interacting domains is the way to go.

N5. (a) The peptide is bound one-quarter of the time. This is indeed significant. It gives a much higher bound fraction than in the intermolecular case, and there is enough bound peptide for the peptide to be almost entirely located next to the binding site, and therefore blocking access for other peptides. (b) The effective concentration is $K_{intermolecular}/K_{intramolecular} = 830/3\,\mu\text{M}$ or $280\,\mu\text{M}$. This is actually a very weak value (larger is better!), but still enough to give useful binding.