# RECOGNITION AND SELECTIVITY OF BINDING: MOLECULAR CORRELATES

Professor T.R.C.Boyde, University of Hong Kong

for a meeting of the
Hong Kong Society for Biochemistry and Molecular Biology,

at the
Chinese University of Hong Kong,

February 22nd 2008

The paper will be prefaced with a few remarks on the origin and early history of the Hong Kong Biochemical Association, including where, when and how the idea of such a society came to life, and the role played by Professor Takashi Murachi, of Kyoto.

---

In its origins, 'recognition' carries an implication of consciousness; the sense used here is extended by analogy to what can be observed of lower organisms or molecules when they bind to or react with each other. Our specific concerns are how such properties are provided for in the molecular structures and how we might improve upon nature by synthetic means. This is most obviously done in the field of drug discovery, and the obvious choice of molecular correlate, for a pharmaceutical company, engaged in such an exercise, is perfection of fit. We need to examine in detail what that means and how the principles can be applied to larger molecules.

The beginning of my interest may be dated to attending a Harden Conference on the subject, in Britain, in 1974, so just a year after arriving in Hong Kong, and it certainly fired me up to come back and try to persuade my colleagues that this should be the departmental theme. Joe Tam owed his appointment in the department to that initiative. I wish it had been wholly successful, we would certainly be much further on; but my colleagues then, like nearly all biologists now, were alarmed by anything that looked like physical chemistry and found it impossible to abandon studies of metabolism. And yet the importance of precise intermolecular recognition was already obvious from the Watson-Crick discovery, if nothing else, and metabolism too depends upon the first example of biological recognition that was at all well studied at an atomic scale – namely enzyme specificity – so we begin there today, with a semantic analysis.

The old, old textbooks used to prate about enzyme specificity being 'absolute'. It never was true. The modern treatment is both quite different and highly instructive for consideration of reactions in which there is no catalysis but only simple binding. Enzyme specificity is nowadays expressed as the ratio of rates for two substrates, and though one of these may be the natural substrate or the best known substrate, nevertheless, no matter how arbitrary the choice, there must always be a comparator.

Turning to antibodies, the word specificity is used ambiguously. Greenspan [1] suggests that there is really a family of overlapping senses. Let us be content with two,

- qualitative, the identity of a target to which an antibody will bind
- quantitative, an antibody binds better to one target than another,

and for the second of these it is proposed to use instead the words 'selectivity' or 'discriminatory power'. Just that choice of words, 'better than another' shows that, as with enzymes, there can be no meaning unless a comparator is present as well as the principal target. We may say that an antibody or other ligand has a definite degree of selectivity for target U over target V. Notice the necessary but often arbitrary selection of which of two participating molecules shall be called the ligand and which the target, and avoidance of the word 'receptor' (because it may cause confusion except when dealing with living cells).

But what should be the units or means of quantitative expression? That depends upon the property of interest or capable of being measured and it seems there is a choice only between affinity, avidity and potency, the latter incorporating downstream biological effects. For the duration of this paper, the choice falls on affinity, for simplicity and because it allows for a fairly simple algebraic theory, not because the others are invalid. In fact the presentation here may be regarded as only a demonstration of what is possible, acknowledging that the theory will need to be worked over again for avidity and potency. The verbal forms, e.g. 'selectivity', may be found useful in all cases, but the symbols and verbal equivalents below will perhaps be best restricted to the affinity case, i.e., equilibrium in solution.

A Discrimination Constant, D, is defined: 'The ratio of the affinity constants of a ligand for two targets'; or, for improved clarity in restricted circumstances, "The ratio of the affinity constants of a ligand for a defined target and a defined cross-reactant."

$$^{X}D_{UV} = K_{XU}/K_{XV}$$

D does not directly represent the degree of discrimination actually achieved in an experiment (just as an equilibrium constant does not directly indicate the extent of reaction in a particular experiment). We may usefully define discrimination, ***d***, as the ratio of the concentrations of the bound forms of ligand (bound to the two targets). For example where ligand X reacts with targets U and V to yield complexes XU and XV

$$\boldsymbol{d} = C_{XU}/C_{XV} = D.C_U/C_V$$

The dimensionless quantities, D and ***d*** are readily interconverted if the concentrations (at equilibrium) of the free target species are known, or even merely the ratio. This way of illustrating the relationship is best for understanding principles, and reflects quite closely one common and valuable practical situation when there is a large excess of both target species over a high-affinity ligand, a situation that also often allows a relatively easy route to measurement.

A numerical example: Antibody X might have selectivity for U over V of $10^6$. If antibody Y has selectivity of $10^7$ for the same pair of targets, then clearly it is in principle a better discriminant, and probably this 10-fold difference can be realized and made useful in practice whatever the absolute values of the affinities concerned.

These concepts are independent of molecular structure. We do not need to know about structures to discuss whether they are valid or not.. We are left to discuss why selectivity matters more than affinity considered alone, and how selectivity can be improved.

**Selectivity versus Affinity**

Suppose a ligand binds with such high affinity that it never comes off:
1] we don't know whether it is binding to a specific epitope or non-specifically
2] we cannot compare its binding with other ligands of comparable or greater affinity
3] we have no way of doing competition experiments
4] if it has bound at first non-specifically, there is no way that it can dissociate and then re-associate in a specific manner, or no way that we can know this has happened

Thus except for strictly practical, operationally-defined purposes (where we should be speaking in terms of avidity) high affinity is by itself not useful.

But why bother with new symbols and a new terminology, when the affinity constants contain all the required information? Perhaps this is just a proposal to shift things around a little.
1] Higher affinity does not necessarily yield better discrimination; looking at affinity constants alone may conceal this fact.
2] There is often confusion about affinity for monovalent and polyvalent targets, or the single-site or 'intrinsic' affinity is confused with affinity proper. The new terminology requires explicit identification of the participants in the reaction being discussed and will thus help to minimize errors arising from this cause.
3] See about polyvalency below. The new terminology makes discussion easier.

**Molecular correlates for high affinity and selectivity**

The classic is chelation of Ca or Mg such as by EDTA, and it is a little surprising that this old, old example has not put an end to development of the theory; anyway it will serve as a starting point. Very high affinity is found and increases as successively more of the orbitals of the 'central atom' are involved in 'coordination bonds' with electron-donating groups of the chelating species. Nuclear and electronic charge, overlap and merging of the orbitals, geometry of the complex, especially because the orbitals are strongly directional: all affect the outcome. Despite high affinity, the reaction is readily reversible and is pH dependent because of effects on the ionization state and charge of potential electron-donating groups, whilst in the unbound state.

In the complex itself, up to nine factors may be seen at work, depending on how you count: **Number, stength** and **nature** of bonds formed, through the **matching** of complementary features which may allow, *inter alia* the formation of **directional bonds. Total contact area.** Molecular **geometry** yielding a **neat** fit without **strain**.

**Applications in drug discovery**.

Search is by computer modelling, or chemical search, or biologically-based search to seek out the potentially 'best' binder which is also druggable; meaning soluble, small molecular size, readily synthesized; and then proves safe in practice. Goodness of binding depends on overall molecular shape plus the number of bonds that can be formed between the candidate drug and the target site, and their quality; therefore partly upon the area of contact, partly on geometrically correct matching of complementary groups to yield ionic bonds, coordination bonds, hydrogen bonds with or without bridging water molecules, covalent bonds, hydrophobic bonds. Deliberately or otherwise, binding may induce conformational change in the target molecule and the effect of that on affinity may be in either direction. More important in considering conformational change are downstream biological effects, whether agonist or inhibitory, so that a lower affinity candidate may yield greater potency.

Looking at binding alone, the key is to provide the greatest possible number of bonding opportunities.

**Applications to specific binding proteins, including antibodies**.

We might say that this actually means all proteins; since all of them act in nature by way of specific binding to other proteins and/or non-protein molecules, and we look first at the binding of an individual binding site, say on an Fab fragment, to its target epitope. It must be emphasized, however, that neither in theory nor in practical application is this treatment confined to antibodies.

All of the above considerations affect affinity and selectivity of binding; **and in addition** we need to consider conformational change in the ligand molecule as well as the target, at least to a greater degree than for ordinary drug candidates, recognizing that change in conformation may either enhance affinity by allowing more bonding opportunities, or lessen affinity if too much of the potentially available bonding energy is taken up to overcome mechanical resistance: the net effect may be either way. Having said that, our simplifying assumption in the rest of this paper is that such single-site reactions may be treated as unitary, that the **intrinsic** affinity of each may be regarded as a constant, and hence that any other influences upon the total binding energy may be treated as due to things taking place outside the area of that unitary interaction. It does **NOT** follow that the overall binding affinity of a whole molecule such as a real-life Fab fragment is the same as the intrinsic affinity of the binding site: that would be to suggest that the rest of the molecule has no influence.

**More than one binding site: Consequences for Affinity**

If two sites are able to bind simultaneously to the same target:

1] There is enhanced area of contact and more bonding opportunities which should yield enhanced binding energy, perhaps the sum of the intrinsic affinities.

2] Offset by losses due to molecular strain or any other interference with bonding. If bad enough, this might prevent binding of the second site altogether.

3] Offset by losses due to diminished freedom of movement; that is, loss of entropy.

4] Affected in either direction by events in other parts of the molecule additional to the above. That is, there may be a direct or indirect contribution to binding energy from parts

of the molecule outside the conventionally-recognized binding site, either plus or minus, and in the bivalent binding reaction it may be more or less than in the monovalent – we cannot predict from existing theory. But even in the monovalent case this contribution may exist**:** we do not know its extent**:** the measured affinity of a monovalent Fab fragment is not the same thing as the intrinsic affinity of the binding site.

The above shows that the increase in affinity for a bivalent reaction is neither automatically, nor necessarily (nor even likely to be) obtainable by adding together the energies for the monovalent reactions. Nevertheless, there usually is an increase [2,3] provided that the ground rules are fulfilled; how big an increase is something to be considered later.

**A digression on affinity enhancement and adding together binding energies**.

We shall have to return to this, especially about entropy, but it may be worth clearing the air right now on some points.

First, the Gibbs standard free energy of a reaction is related to the equilibrium constant, $\Delta G^o = -RT\ln K$, so energy and equilibrium constant can always be calculated one from the other, though few people take the trouble to understand what is meant by a standard state. Adding free energies corresponds to multiplying the equilibrium constants.

To repeat, the increase in affinity for a bivalent reaction is neither automatically, nor necessarily (nor even likely to be) obtainable by adding together the energies for the monovalent reactions The situation was clearly laid out by W.P.Jencks (of Brandeis) 25 years ago [4] and our ability to calculate basic thermodynamic parameters is scarcely better today, so that his analysis stands, though to read the literature you would not believe that any such work had been published. Some believe that affinity enhancement *should* correspond to multiplication of K values, others 'prove' that affinity is not improved by bivalency, on the basis of mind-numbingly stupid experiments or grave misunderstanding of the theory, other still reinvent elaborate theories based on kinetics and a 'mechano'-type model of molecular structure. The best employ a statistical mechanics type of approach though of only approximate nature.

Bill Jencks died a year ago, and this paper is dedicated to his memory. Papers on various aspects were in early draft at that time and I thought of getting his advice and help over publication, but too late. However, even he didn't seem to know that a true thermodynamic equilibrium constant is dimensionless: it has no units, neither of moles per litre nor litres per mole. This makes little practical difference, it's true, provided we all stick to the same conventions about concentrations and activities, but it does explain some little fads of the physical chemists, and why we can take the logarithm of K.

**More than one binding site: Consequences for Selectivity**.

If a bivalent antibody has higher affinity for its bi- or polyvalent target, then it also has greater selectivity for that target, than the corresponding Fab fragment. There is really nothing startling about this conclusion; in fact I believe that we all recognize its truth in our daily work. Some examples, however, may help to show the power of 'selectivity' to illuminate experimental and natural phenomena.

1] A bivalent antibody should select a polyvalent target even in the presence of soluble forms of the monovalent epitope. This has obvious advantages for dealing with a foreign cellular or viral target, or rogue cells from within, where in both cases soluble confounding molecules can and do occur.
2] Here is a plausible explanation for the evolution of bi- and higher valencies in antibodies. This idea has been around for many years: 'selectivity' only makes it easier to appreciate the underlying physics.
3] Roughly speaking, the earliest antibody response to infection is decavalent IgM, wherein the affinity of individual Fab units is low. But the potential affinity and selectivity for appropriate targets can nevertheless be exceedingly high (and of course there are downstream effects also). More generally, polyvalency in antibodies allows for high affinity and selectivity without high affinity of the individual binding site, provided always that the target is polyvalent and both or all binding sites can react.

**Heteropolyvalency**. **Always artificial? Specificity in polyvalent ligands**.

From the beginning, forty years ago (by Nisonoff [5], who I already knew about because he had published on transaminase kinetics), the objective has in nearly every case been to link together **two different molecules**, for experimental or therapeutic purposes. Lacking the insight provided by 'selectivity', the workers concerned have seen only dimly, if at all, the opportunity or the advantages presented by a reagent that might react simultaneously with **two different epitopes on the same target**, whether on the same macromolecule or merely present on the same surface.

A theoretical analysis must run in parallel with that for hompolyvalency. (Don't underestimate the importance or the difficulty of that insight. It must be difficult and important because it has taken me almost twenty years to get there.) It also leads to a new insight into the meaning of specificity, since a heteropolyvalent antibody shows specificity and selectivity for its corresponding heteropolyvalent target, that is to say a **new specificity not present in nature**. It follows that the natural homopolyvalent antibody must also have a specificity profile different from that of the individual Fab units, and indeed we have just seen that this is so, it exhibits a specificity for the homopolyvalent target that the Fab cannot possess, in addition to the underlying specificity for the monovalent target.

**Spacing and linking between binding sites**.

The examples of drug discovery and chelation point strongly to an ideal situation wherein the spacing and orientation of binding sites or key bonds exactly match the corresponding features on the target. But when we are assembling an artificial polyvalent antibody intended to react with a polyvalent target, we don't have that refinement of control and instead either make use of the tools to hand and get the assembly together any old how or provide for sufficient flexibility in the linker that the whole molecule can adopt the necessary conformation. Herein lies a major point of controversy. There is a strong body of opinion, led by no less than George Whitesides [6], that a flexible linker must lose so much entropy upon being tied down in the complex as to obviate the advantage of polyvalency, and I am obliged again to digress for a while to examine this. Possibly Whitesides is right to think that a long aliphatic chain will have the disadvantages cited, losing about 0.5 kcal/mol per freely rotating bond in the original

chain that is immobilised in the complex. Detailed computation and some experimental results combine to suggest that the effect is nothing like as bad for linkers made of either aminoacid or nucleotide chains, the latter of particular relevance to this talk. Let me present just one model computation [7].

The problem set is, what effect does length of an oligonucleotide spacer have upon the binding energy attributable to a second binding site, once the first has already bound, and this can be modeled by a simple oligonucleotide chain having complementary sequences at the ends, separated by a non-hybridising spacer sequence e.g ACCCCC$(A)_n$GGGGGA. Then the thermodynamic parameters may be calculated as a function of n, with results shown in this table, which are startling and remarkably similar in principle to what has been calculated for a polyaminoacid spacer by Zhou [8].

Conditions: NaCl 0.150mol/l, MgCl2 0.002 mol/l, 37degCelsius
Free energy and enthalpy kcal/mol; entropy (cal/mol/K); Tm deg Celsius.

| n | ΔG | ΔH | ΔS | Tm |
|---|---|---|---|---|
| 1 | -3.5 | -32.4 | -93.2 | 74.4 |
| 2 | -4.5 | -39.4 | -112.6 | 76.9 |
| 3 | -5.2 | -40.4 | -113.5 | 82.7 |
| 4 | -5.3 | -38.2 | -106.1 | 86.8 |
| 5 | -6.4 | -43.2 | -118.7 | 90.8 |
| 6 | -5.7 | -43.2 | -120.9 | 84.0 |
| 7 | -5.4 | -43.2 | -121.9 | 81.2 |
| 8 | -5.4 | -43.2 | -121.9 | 81.2 |
| 9 | -5.4 | -43.2 | -121.9 | 81.2 |
| 10 | -5.4 | -43.2 | -121.9 | 81.2 |
| 15 | -5.0 | -43.2 | -123.2 | 77.5 |
| 20 | -4.6 | -43.2 | -124.5 | 73.9 |
| 25 | -4.0 | -43.2 | -126.4 | 68.5 |
| 30 | -3.5 | -43.2 | -128.0 | 64.2 |

The pattern may be described thus: ΔH increases to reach a plateau from n = 5 onwards. ΔS at first decreases rapidly (unfavorable to binding) but then becomes invariant from n = 7 to n = 10 and decreases only slowly thereafter. In consequence, ΔG and Tm exhibit a sharp peak at n = 5 but decline slowly after n = 7, such that the predicted loss of affinity is only about ten-fold from that point to n = 25. Thus an optimum of length is found, much as predicted by Zhou for certain of his peptide linkers, and a long tail as also predicted by Zhou [8].

Oligonucleotide linkers may offer other practical advantages:
1] Adducts with an oligonucleotide tail are easily made,
2] may be of fairly low molecular weight, e.g. about 20kDa,
3] readily self-assemble forming a link that is as stable as necessary,
4] having also a flexible segment that is as long as necessary.
5] Hybrids can be made with >2 Fab's,

6] plus additional effector or marker components,
7] in a few moments (given stock of the necessary adducts),
8] tailored for the individual case,
9] even within the body, on site, *in vivo*,
10] thus avoiding problems of tissue penetration,
11] and may eventually provide for intra-cellular access.

**Conclusions**.

Provided that all binding sites are able to react with epitopes on a single target without either undue strain or floppiness:-
1] Polyvalent binding leads generally to enhanced affinity compared with the monovalent reaction, and also 2] enhanced selectivity, and also 3] a distinctive specificity, namely for the polyvalent target.
4] Artificial heteropolyvalent ligands exhibit novel specificities not found in nature but equivalent natural phenomema exist [9].
5] Some flexibility in ligand and/or target may be essential for binding to occur between macromolecules [10].
6] Linkers (between binding sites of the ligand) composed of polyaminoacids or oligonucleotides may escape the expected general degrading (entropic) effect of flexibility and excessive length.
7] Linkers composed of oligonucleotides offer many other advantages.
8] Hetereopolyvalent ligands based on natural proteins, including antibodies, and with oligonucleotide linkers may prove very useful in laboratory, industry and clinic.

References (see also www.trcboyde.net)

1. Greenspan NS (2001) *Curr Top Microbiol Immunol* 260: 65-85
2. Greenbury CL, Moore DH, Nunn LAC (1965) *Immunol* 8: 420-431
Hornick CF, Karush F (1972) *Immunochemistry* 9: 325-340
3. Crothers DM, Metzger H (1972) *Immunochemistry* 9: 341-357.
4. Jencks WP (1981) *Proc Natl Acad Sci USA* 78: 4046-4050.
5. Nisonoff A, Rivers MM (1961) *Arch Biochem Biophys* 93: 463-466
Hong R, Nisonoff A (1965) *J Biol Chem* 210: 3883-end
6. Mammen M, Choi S-K, Whitesides GM (1998) *Angew Chem Int Ed* 37: 2754-2794.
7. Zuker M (2003) *Nucleic Acids Res* 31: 3406-3415
8. Zhou H-X, (2003) *J Mol Biol* 329: 1-8
9..Maraganore JM, Bourdon P, Jablonski J, Ramachandran KL, Fenton JW, II (1990) *Biochemistry* 29: 7095-7101.
10. Zhang Y, Xi Z Hegde RS, Shakked Z, Crothers DM (2004) *Proc Natl Acad Sci USA* 101: 8337-41.