

Chapter 1
Principles of Supramolecular Chemistry and Molecular Recognition

1-1 Molecular Interactions and Chemical Bonds - The Nature of the Hydrogen Bonds

1-1-1 The kinds of chemical bonds

Three types of chemical bonds are generally recognized: covalent bonds, ionic bonds, and metallic bonds. There is some arbitrariness in this classification system and many chemical bonds are intermediate in type. Hydrogen bonds can better discussed and described within the context of chemical bonds generally.

The simplest atom, hydrogen (one proton and one extranuclear electron) has a 1s orbital, the wave function of lowest energy. In this orbital the electron cloud is spherically symmetrical around the central nucleus. Ordinary hydrogen gas consisted of diatomic molecules H-H bond is the simplest case of a covalent bond. The overlap of the 1S orbitals of two H atoms form a stable covalent bond. The covalent bond is sometimes referred to as an electron-pair bond. All electrons have the property of the spin, with two possible spins allowed. If the electron spins of the two electrons associated with two hydrogen atoms are the same, there is repulsion and no bond formation, whereas opposite spins lead to bond formation with the liberation of about 100 kcal/mol of H₂ molecules. Since the two H atoms in a H₂ molecule are identical, the electron pair is shared equally and the electron cloud is symmetrical. When a covalent bond is formed between unlike atoms the electron charge could may be asymmetric as seen in hydrogen chloride, in which the Cl molecule is more electronegative than hydrogen. It will concentrate the electron cloud toward itself, creating a partial negative charge on itself and leaving the H atom with a partial positive charge. Covalent bonds with displacement of electron charge could toward the more electronegative atoms or groups are called polar bonds and molecules or dipoles.

When electronegativity differences are sufficiently great, electrons may be transferred completely from one atom to another, creating negatively and positively charged ions. These ions form ionic bonds. The most common example of this type of bonding is found sodium chloride.

The best picture of bonding in metals indicates that the individual atoms have a positive charge and the electrons are delocalized throughout the entire metal. No bond can be considered to exist solely between any two atoms in a metal crystal. The delocalized electrons are responsible for the high electrical and thermal conductivity of metals.

1-1-2 The kinds of intermolecular forces (interactions)

Six types of intermolecular forces are generally recognized: repulsion due to the exclusion principle, orientation or electrostatic forces, polarization forces, dispersion interactions, charge transfer interactions, and hydrogen bonds¹. All intermolecular interactions normally involve energies of from less than 1 to some 15 kcal/mol and are, therefore, weak by comparison with ordinary chemical bond.

(1) *repulsion forces*

repulsion force is a short range interactions which occurs when two atoms or molecules approach each other. The repulsion between closed electron shells is due to the Pauli exclusion principle, in the sense that too many electrons occupy the same space. There is room for only one pair of electrons

in a bonding orbital. Generally, an antibonding electron exerts a stronger repulsion effect than the attractive force of a bonding electron. The rapidly rising repulsion energy is represented by a simple function:

$$E(R) \text{ (repulsion)} = \frac{b}{R^n} \quad \text{or} \quad \beta \exp\left(-\frac{R}{\rho}\right)$$

where b , n , β , and ρ are empirical factors and R is the distance between the two interacting atoms or groups. In general, the exponential form is preferred on the basis of theory and experimental results.

(2) Orientation forces

These forces are due to the mutual Coulombic attraction or repulsion of the net charges or electric moments carried by two interacting atoms or molecules. The interaction between two ions of net charge q and q' is given by the relation

$$E(R) \text{ (ion-ion)} = \frac{qq'}{\epsilon R}$$

Where ϵ is an effective value for the dielectric constant of the medium surrounding the two charges. When $\epsilon=15$ and $R=5\text{\AA}$, $E=4\text{kcal/mol}$ for two unit charges of opposite sign. Higher-order Coulombic interactions are ion-dipole and dipole-dipole interactions.

(3) Polarization forces

These forces arise from the polarization of one atom or molecule by an approaching atom or molecule. The polarization consisted of the redistribution of the charges of the first molecule or atom with concomitant effects on its permanent electric moment, due to the effect of the charges or permanent electric moments of the approaching atom or molecule. In the case of a charge q interacting with a spherical group or isotropic polarizability α , the polarization energy is

$$E(R) \text{ (ion-induced dipole)} = - \frac{q^2 \alpha}{2 \epsilon^2 R^4}$$

The polarization energy in the case of a unit charge and a CH_2 group assumed spherical with $\alpha=1.84\text{\AA}^3$, is $\sim 0.002\text{ kcal/mol}$ at $R=5\text{\AA}$ and $\epsilon=15$.

(4) Dispersion (London) forces

These forces have a purely quantum mechanical nature, are always attractive for molecules or atoms in their electronic ground states, and arise even between neutral, apolar molecules or atoms. They are due to the average interactions of an instantaneous moment on one molecule brought about by charge density fluctuations and the electric moment induced on the other atom or molecule. A pair of isotropic groups of polarizability α have a dispersion interaction energy given by

$$E(R) \text{ (dispersion)} = \frac{3\Delta E \alpha^2}{4R^6}$$

Where ΔE is an average electronic excitation energy, often equated to the ionization potential of the

interacting systems. In the case of two CH₂ groups,

$$E(R) = -\frac{1340}{R^6}$$

If the CH₂ groups are 5 Å apart, the dispersion energy is about 0.1 kcal/mol.

(5) Mulliken charge transfer forces

The approach of an electron donor molecule (D:) to an electron acceptor (A) can lead to the transfer of an electron from the former to the latter with the concomitant appearance of a characteristic, intense absorption band in the visible or UV region of the spectrum. The Mulliken theory describes the ground state energy of the complex formed in term of resonance between a non-bonded form (:D, A) of the donor-acceptor pair, and a contribution (D⁺-A⁻) held together through the dative bond resulting from transfer of an electron from D to A. The total wave function for the ground state of the system will rely on the coefficients *a* and *b* which indicates the weighting factors for the two different contributions

$$\Psi_N = a \Psi(:D, A) + b \Psi(D^+ - A^-)$$

The thermodynamic parameters for the charge transfer complexes, that is, their free energy, enthalpy, and entropy of formation, are of the same magnitudes as for the hydrogen-bonded complexes.

(6) Hydrogen bonds

The distinguish feature of hydrogen bond is the involvement of a specific H atom of a proton donor group with a localized site of high electron density in the same or another molecule. The last two situations correspond to the formation of intramolecular and intermolecular hydrogen bonds, respectively. Since repulsion forces are always present to some extent in any molecular system, the overall net attractive or repulsive interaction depends on the sum of all the interactions. An important feature of hydrogen bond and of other weak attractive interactions in solution is that, at ordinary temperatures, only a fraction of the molecules are generally associated. At equilibrium, while a certain number of new complexes are continually formed, an equal number of complexes are continually broken, due to the kinetic energy of motion of the interacting molecules.

1-1-3 The Criteria for Hydrogen Bonding

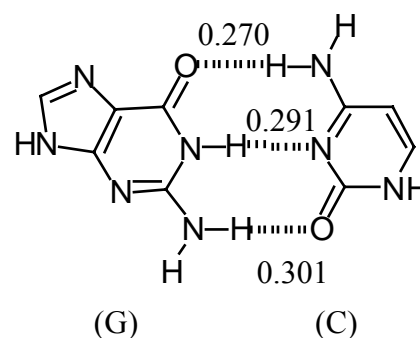
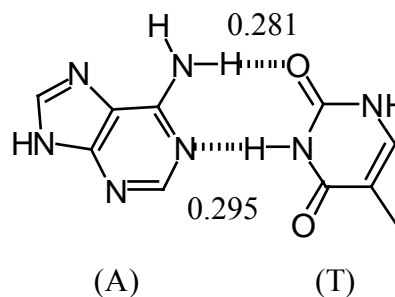
The criteria for the existence of hydrogen bonds are somewhat more clear cut than for other intermolecular interactions. No single criterion or single physical manifestations, however, can establish the presence of hydrogen bond in a given system beyond any reasonable doubt. Some convenient operational criteria for the existence of hydrogen bonds are listed below.

- (1) hydrogen bond occurs between a proton donor group A-H and proton acceptor group B, where A is an electronegative atom, O, N, S, X(F, Br, I, Cl) or C, and the acceptor group is a lone electron pair of an electronegative atom, or p-electron orbital of a multiple bond (unsaturated) system. In general, a hydrogen bond can be characterized as a proton shared by two lone electron pairs.

- (2) Hydrogen bond is a distinctly directional and specific interaction. It is more localized than any other type of weak interactions. Hydrogen bonds are linear, but appreciable variation in the angle A-H-B can occur.
- (3) The total hydrogen bond length R is equal to or less than the sum of the van der Waals radii of atoms A and B, that is, the total bond length contraction caused by hydrogen bond formation is equal to or is greater than twice the van der Waals radius of the hydrogen atom, as shown in **Table 1-1**.
- (4) The enthalpy of hydrogen bond generally falls in the range of 1 to 10 kcal/mol, Intermolecular interactions other than hydrogen bond also fall within range.
- (5) Hydrogen bond is an association phenomenon. It causes a decrease in the total number of free molecules and an increase in the average molecular weight.
- (6) In the hydrogen bond, a specific covalent A-H group interacts with a specific acceptor site. The A-H bond is thereby weakened but not broken, and the properties of the acceptor group are also affected.

Table 1-1 The bond radius of the basic hydrogen bonds

hydrogen bond	bond radius [nm]
OH·····O	0.272
OH·····N	0.279
OH·····S	0.331
OH·····F	0.272
OH·····Cl	0.312
OH·····Br	0.328
NH·····O	0.289
NH·····N	0.298
NH·····S	0.342
NH·····F	0.292
NH·····Cl	0.323
NH·····Br	0.337
FH·····F	0.244

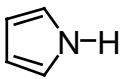


1-1-4 The Participants in hydrogen bond

Hydrogen bonding is an interaction between a covalently bound H atom, with some tendency to be donated, and a region of high electron density on an electronegative atom or group of atoms, which can accept the proton. Typical proton donor groups include hydrogen bound covalently to electronegative atoms such as O, N, S, X(halogen), and in special cases, carbon. It is possible that some

Si-H and P-H group form weak hydrogen bond, as summarized in **Table 1-2**. The nature of the atom or group to which the proton donor group A-H is attached can be much more variable than that of A itself. Thus the hydroxyl group, O-H can be attached to a variety of atoms or groups ranging from hydroperoxides R-O-O-H to silanols Si-O-H to phosphoric acids, and still be able to form hydrogen bonds. The atom or group to which A-H is attached affects the proton donor strength, that is to say, the acidity of A-H. Fluoroalcohols are appreciably more acidic than unsubstituted alcohols, and form stronger hydrogen bonds with proton acceptors.

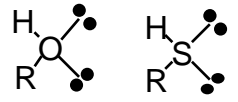
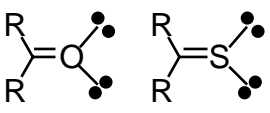
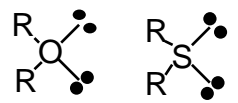
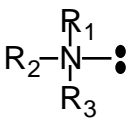
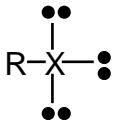
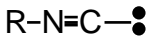
Table 1-2 Hydrogen bonding donors

donor group A-H	examples
O-H	H-O-H R-O-H $\text{R}_2\text{-}\overset{\text{R}_1}{\underset{\text{R}_3}{\text{C}}}\text{-Si-OH}$ $\text{R}_1\text{O}-\overset{\text{O}}{\underset{\text{OR}_2}{\text{P}}}-\text{O-H}$ $\text{R}-\overset{\text{O}}{\parallel}\text{C}-\text{O-H}$ $\text{R}_1\text{-}\overset{\text{R}_1}{\underset{\text{R}_2}{\text{C}}}=\text{N-O-H}$ $\text{R}-\overset{\text{O}}{\parallel}\text{C}-\text{O-OH}$
N-H	$\text{R}-\overset{\text{H}}{\underset{\text{H}}{\text{N}}}$ $\text{R}_1-\overset{\text{H}}{\underset{\text{R}_2}{\text{N}}}$  $\text{R}_1-\overset{\text{O}}{\parallel}\text{C}-\overset{\text{H}}{\underset{\text{H}}{\text{N}}}-\text{R}_2$ $\text{R}_1\text{-O}-\overset{\text{O}}{\parallel}\text{C}-\overset{\text{H}}{\underset{\text{H}}{\text{N}}}-\text{R}_2$ $\text{S}=\text{C}=\text{N-H}$
S-H	R-S-H
X-H	F-H
C-H	$\text{Cl}-\overset{\text{Cl}}{\underset{\text{Cl}}{\text{C}}}-\text{H}$ $\text{R}-\text{C}\equiv\text{H}$ $\text{N}\equiv\text{H}$
P-H	$\text{R}_1\text{-}\overset{\text{R}_1}{\underset{\text{R}_2}{\text{P}}}-\text{H}$ $\text{R}_1-\overset{\text{O}}{\parallel}\text{C}-\text{H}$ R_2

The H atom acceptors are unshared electron pair of an electronegative atom or the π electrons of a multiple bond system. **Table 1-3** lists most of the known hydrogen bonds acceptor groups, together with examples of each type. As with proton donors, the proton acceptor groups or atoms can be attached to an atom other than carbon and still participate in hydrogen bonds. Oxygen is a good proton acceptor whether it is attached to phosphorous, to sulfur, or to nitrogen. The oxygen atoms of metal β -ketoenolate

complexes can also participate in hydrogen bond as acceptors. The anions of electronegative atoms form strong hydrogen bonds. They have not, however, been studied thoroughly. The nitro and cyclopropyl groups not listed in table are borderline proton acceptors. The former is a multiple bond system similar to the nitrile group, while the cyclopropyl group is probably similar to a conjugated π -electron system, such as benzene.

Table 1-3 Hydrogen bonding acceptors

acceptor group B	examples
	water, alcohols, mercaptans
	ketones, esters, carboxylic acids, sulfoxides
	ethers, sulfides
	primary amines, secondary amines tertiary amines, pyrroles, pyridines
	organic halides, hydrogen fluoride
	isonitriles (isocyanides)
π -electron system	acetylenes, benzenes, nitriles, cyanides

The occurrence of hydrogen bond in systems constructed from the proton donors and acceptors listed in tables cannot be always taken for granted. A weak proton donor group, for example, P-H, is unlikely to form hydrogen bonds with a weak proton acceptor group, such as benzene.

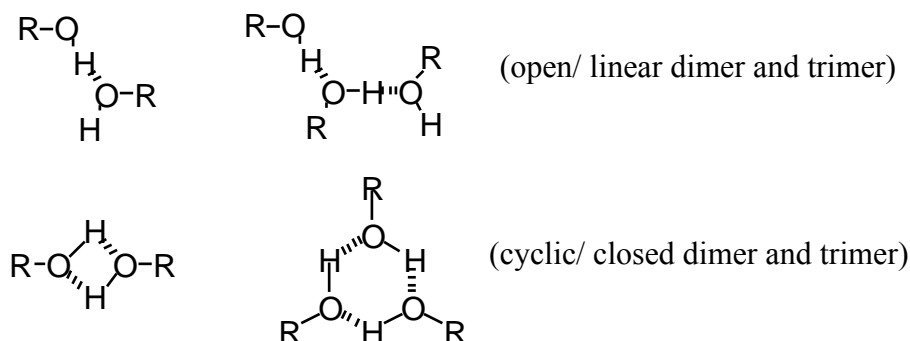
All molecules can be conveniently classified into four types with respect to their ability to participate in hydrogen bond (**Table 1-4**). Hydrogen bonding molecules are divided into types I through III, while molecules incapable of hydrogen bond form type IV. The latter include compounds which are used as the so-called inert solvents in studies of hydrogen bonded molecules.

Table 1-4 Classification of molecules according to Pimentel and McClellan

type	description	examples
I	molecules with one or more donor groups (acids) and no acceptor group	haloforms, highly halogenated compounds, acetylenes
II	molecules with one or more acceptor groups (bases) and no donor groups	ketones, ethers, esters, olefins, aromatics, tertiary amines, nitriles, isonitriles
III	molecules with both donor and acceptor groups	alcohols, water, phenols, inorganic and carboxylic acids, primary and secondary amines
IV	molecules with neither donor nor acceptor groups	saturated hydrocarbons, carbon tetrachloride, carbon disulfide

Various hydrogen-bonded systems occur when one or more of the three types of molecules are present. The bulk of studies dealing with hydrogen bonded systems have been performed in the liquid phase, and unless otherwise specified we will restrict ourselves to solutions. It is instructive to consider some of the types of hydrogen-bonded systems formed in solutions of various types of molecules.

The type I plus type II molecules form hydrogen bonded complexes, frequently in a simple one to one ratio. The strength of hydrogen bond relies on primarily the relative acidity of I and the basicity of II. Type III molecules can self-associate by hydrogen bond with themselves. Two types of hydrogen bonded complexes may be formed: (1) intermolecular involving two or more separate molecules; and (2) intramolecular, involving donor and acceptor sites within the same molecule. The strength of the hydrogen bond depends on the relative acidities and basicities of the donor and acceptor sites and in case of intramolecular hydrogen bond, on the spatial arrangement present. Self-association through intermolecular hydrogen bond can form a large variety of open/linear and cyclic/closed polymers, as indicated below.



When molecules of either type I or II added to type III molecules, several hydrogen bond can exist. In the case of type I plus type III molecules, the donor group of the type I molecule must compete with that of the type III molecule for the acceptor sites in the latter. This competition generally leads to

decrease in the extent of the self-association of the type III molecules, and the formation of one or more new hydrogen bonded complexes between the two types of molecules. This is also true for the system consisting of type II plus III molecules, except that in this case the competition is between the acceptor sites in the two types of molecules for the donor sites of the type III molecule. Because of the number and complexity of the molecular species possible in such systems, experimental results are difficult to interpret. Great simplification can generally be achieved through the usage of an inert solvent (type IV molecule). Dilution of a solution of a self-associated (type III) molecule by an inert solvent shifts the self-association equilibria to the left, towards the monomeric species. Sufficiently low concentrations the type III molecules may be completely unassociated. Therefore addition of an excess molecules of either type I or II to a dilute solution of type III molecules provides a convenient experimental technique for the investigation of many different hydrogen bond interactions. For example, addition of pyridine to a dilute solution of EtOH in CCl₄ is a convenient way of studying O-H-N bonding, due to the preponderance of the hydrogen bonded 1:1 complex over other hydrogen bonded species composed of self-associated EtOH molecules.

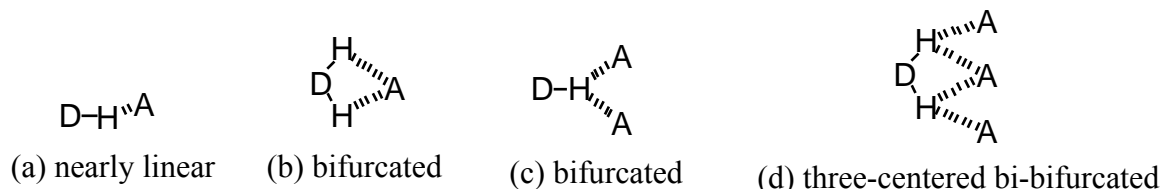
1-1-5 Strength of Hydrogen Bonded Complexes

Probably the most important non-covalent interaction for designing the supramolecular architecture or synthetic receptor is hydrogen bond, which is occurred between a proton donor D and acceptor A. Therefore those who intend to construct supramolecules must consider the strength of hydrogen bonds, taking into account the fundamental factors such as angle, the number of hydrogen bonds, flexibility and the distortions (conformation) of the host (guest) molecules, the solvent effects, and so on, which absolutely affect the strength of the hydrogen bonds, although the strength of the hydrogen bonds cannot be defined by each of individual factors. In this section, these factors will be explained.

(a) Angle and distances between donors and acceptors

As long as the charges of A and D are centered in these atoms one would predict for an optimal interaction a linear D-H-A arrangement with D-H-A angle close to 180 degree. In solid state structures, which offer a rich source of corresponding geometries, this is, however, rarely found²: typical D-A distances between hydroxyl oxygen atoms are for example 2.72 Å, whereas the sum of the van der Waals radius would be for O-O 3.04Å, and for a linear O-H-O bond even 5.44Å. Statistical analyses of directionality in organic crystals give an average angle of 167° for O-H-O bonds and 161° for N-H-O bonds. One hydrogen atom H can also interact with two acceptor A atoms in a three-center, often called bifurcated, arrangement (type b), less frequent are other types (type c and d) Unfortunately the description 'bifurcated' is in use for both configuration (b) and (c) and (d) in combination with the term 'three center'. Variety type of hydrogen bonds thus exist and their flexibility allows the tolerance of the different distances or angle from theoretically expected value. If the hydrogen bonds do occur with only special distances or angle (if the hydrogen bonds are very strict interaction), the regulation of the

rotation (conformation) for each CH₂ group is necessary, which results in the unexpectedly larger loss of not only entropy but also enthalpy, which is caused by the increase in *Gauche* (distorted) conformation, *vide infra*.



(b) *Donor and acceptor abilities (relationship with pKa values of donor and acceptor)*

A practically useful empirical approach for the quantification of donor and acceptor abilities in hydrogen bond rests on the analysis of thousands of binding constants between simple solute molecules, usually taken from spectroscopic measurements in solvents like CCl₄. The thus derived factor values allow one to estimate the binding free energies, or log *K* values, for all kinds of bonding between A and D functions in non-protonic solvents, based on a multiplicative combination of the factors *C_A* and *C_D*, or acidity and basicity constants, α₂^H, β₂^H in equations, with R=0.9956 and SD=0.09.

$$\begin{aligned}\Delta G [\text{kJ/mol}] &= 2.43 C_A C_D + 5.70 \text{ (for CCl}_4\text{)} \\ \Delta G [\text{kJ/mol}] &= 1.93 C_A C_D + 5.70 \text{ (for CHCl}_3\text{)} \\ &\text{or} \\ \log K [\text{M}^{-1}] &= 7.354 \alpha_2^{\text{H}} \beta_2^{\text{H}} - 1.094 \text{ (for CCl}_4\text{)}\end{aligned}$$

The factor values relate to other, less comprehensive, scales of electron accepting and donating power. For structurally related series of compounds, one observes linear correlations between complexation free energies and p*K_a* values of hydrogen bond donors and acceptors. Apart from well established classical steric and electronic substituent effects, there are neighboring group effects, which can exert an enormous influence on the basicity or acidity of donor or acceptor functions. The dissociation free energies of benzoic acid drop by 33kJ/mol upon introducing of one *ortho*-hydroxyl group, a second one has an almost additive effect of 59kJ/mol.

Related strong hydrogen bond exist, such as between neighboring carboxyl group; in these recently-called low-barrier hydrogen bond (LBHB) the proton H is in the time average centered between the A and D atoms, but separate by relatively low barriers. As a consequence, the hydrogen atom H is particularly depleted of shielding electrons and characterized by very large downfield NMR shift. Only for the symmetric LBHB profile does one expect a substantial difference between the small hydrogen H, and the larger deuterium atom D, and a corresponding isotope effect on the NMR shielding. LBHB associations are intermediates between the above-mentioned weaker bonds (with up to 35kJ/mol in the gas phase) with a distinct double-well potential, and very strong, close to single well bonds. Such NMR characteristics are also observed for intermolecular hydrogen bonds and in the

catalytic triad of serine proteases. However, the existence of LBHB in these systems was seriously questioned on the basis of more detailed studies.

Obviously, the symmetric potential profile depicted in LBHB bonds is expected only if A and D atoms possess similar basicities, and to reach its maximum if the pK_a values of donor and acceptor unit are matching. It has been proposed that such LBHB not only show spectral characteristics, but also are anomalously strong and their formation explains, in particular, extremely efficient transition state stabilization in active sites of some enzymes. Of course for a given hydrogen bond donor D-H the most stable hydrogen bond will be formed with acceptor A possessing the same basicity as D because interaction free energies correlate with pK_a values of A-H increasing on increase in pK_a until $pK_a = 0$; after that the proton transfer occurs from more acidic D-H to more basic A. However no derivations in linear correlations between $\log K$ and pK_a values at $\Delta pK_a = 0$ were reported indicating that there are no additional energy effects associated with symmetrical bonds of the LBHB type, although they do show specific NMR and IR spectroscopic features. Also, recent measurements of model compounds indicate that there is only a relatively small, if any, additional gain in free energies in such systems. Thus, the association free energies of the symmetric free acid in the Kemp acid derivatives is only higher by 1.4-2.4 kJ/mol than that of corresponding asymmetric amide-carboxylate interaction.

(c) *Preorganization and rigidification of host (guest) molecules*

The preorganization principle states that the more highly hosts and guests are organized for binding and low solvation prior to their complexation, the more stable will be their complexes. Multi-site interactions between host and guest molecules are almost invariably accompanied by some energetically unfavorable contributions such as strain, restrictions of internal rotations, desolvation. Evidently, both molecules will be optimally preorganized if (a) all complementary binding sites geometrically match; (b) in the complexed state they are in the same lowest free energy conformation as in the free state; and (c) polar binding sites need not to change solvation. In this case, all distortions will be negligible and the complexation will be energetically most favorable, including only the sum of intrinsic binding free energy. One also may expect the interactions of such preorganized molecules to be highly selective since guests of different structures, such as cations of different sizes, will need different optimal host conformations.

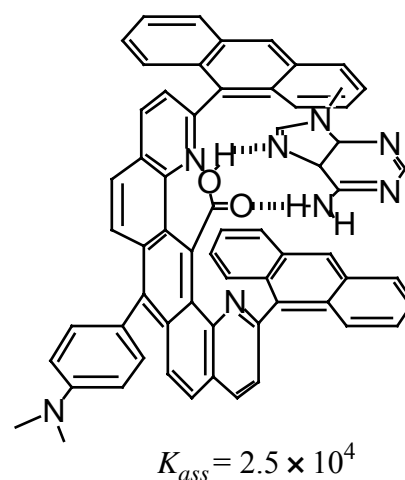
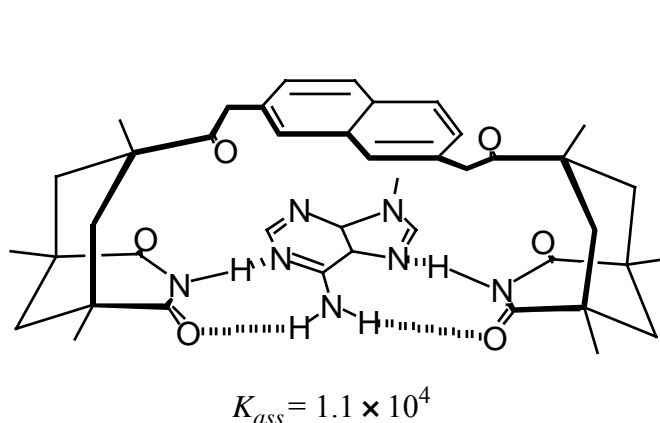
Although the general concept seems to be quite simple, it is not easy to materialize it for every particular case and there is still only a limited number of highly efficient preorganized hosts. An often applied strategy to preorganized host molecule is rigidification or introduction of macrocycles with purpose to obtain ideally a single conformation optimal for complexation. As a rule, creation of properly rigidified macrocycles requires complex synthetic routes. On the other hand, effective complexation requires primarily a sufficient number of complementary interaction sites, *vide infra*, which in view of usually minor entropy effects can also be linked by several flexible bonds. It is, however, essential, that the interaction sites are sterically matching without build-up of considerable strain, like too many gauche-conformations. Most non-covalent interactions tolerate a considerable degree of flexibility. The

tolerance depends very much on the type of interaction forces; so Coulomb in contrast to dispersive interactions require a much less tight fit. In addition, complexation with a rigid host can be kinetically slow, which is undesirable for many applications. Perfect biological protein receptors possess considerable conformational mobility. Fischer's lock-and-key hypothesis originally viewed protein-binding sites as rigid structures. Only later was the conformational mobility of proteins discovered, and the induced-fit hypothesis was proposed (This phenomena were detailed below.): structures of free and bound proteins often endure considerable conformation changes upon binding.

By the way, the entropic contributions in complexation process can be obtained by the sum of the ΔS_{rot} , ΔS_{trans} , $\Delta S_{\text{intern.rot}}$, and ΔS_{vib} .

$$\Delta S_{\text{ass}}(\text{gas phase}) = \Delta S_{\text{rot}} + \Delta S_{\text{trans}} + \Delta S_{\text{intern.rot}} + \Delta S_{\text{vib}}$$

In this equation, the contribution of ΔS_{rot} , ΔS_{trans} , $\Delta S_{\text{intern.rot}}$ is major. (The ΔS_{rot} is rotational entropy, which depends on the product of molecular moments of inertia, and therefore affected by the molecular weight and the shape. The ΔS_{trans} is translational entropy, which relies on solely the molecular weight.) Thus as discussed above, the flexibility in hydrogen bond can cancel out the negative contribution of entropy. Therefore, a question would be raised, why not the synthetically-accessible open chain host is used, whereas the macrocycles are widely used^{3,4} (see below), which is accompanied by the struggle of synthetic effort?



There are two reasons. One is, as previously discussed, the minimization of the irregular conformations. The second is that the solvation inside the cavity is absent or weaker than with hosts where the binding sites are more exposed to the solvent. The free energies of the complexations ΔG can be described by the sum of the gas phase free energy ($\Delta G(\text{g})_{\text{HG}}$), solvation free energy of the complex ($\Delta G(\text{solv})_{\text{HG}}$), and desolvation of the host and guest molecules ($\Delta G(\text{disolve})_{\text{H}}$, $\Delta G(\text{disolve})_{\text{G}}$).

$$\Delta G = \Delta G(\text{g})_{\text{HG}} + \Delta G(\text{solv})_{\text{HG}} + \Delta G(\text{disolve})_{\text{H}} + \Delta G(\text{disolve})_{\text{G}}$$

Not only can the desolvation term thus be minimized, the restricted cavity size can also produce

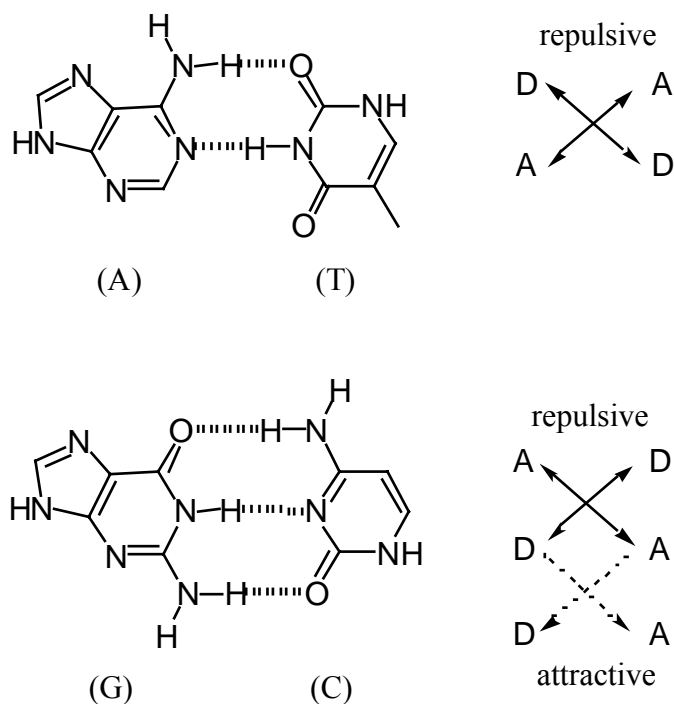
high-enthalpy solvent inside, which by uptake of a guest will then be liberated, and thus enhance the association constant.

(d) *The number of the hydrogen bonds*

Synthetic host-guest chemistry has produced a large array of effective hosts and guest molecules with peptide-like partial structures. And when we intend to realize more effective molecular recognition, the functional groups for more pairwise of interactions (hydrogen bonds) are introduced. This approach is, in a sense, correct, since in the biological systems, molecular recognition was occurred via multi-point interactions, as summarized in the next section. Indeed, in the simple solute-solute associations these are usually well-ordered, and allow to assign energetic values to the participating hydrogen bonds on firmer grounds. If we analyze the ΔG values of systems as a function of the possible number n of hydrogen bonds, we observe, however, a fairly linear correlation⁵. The few outliers can be assigned to unfavorable steric interactions, *vide ante*, or to too-flexible frameworks.

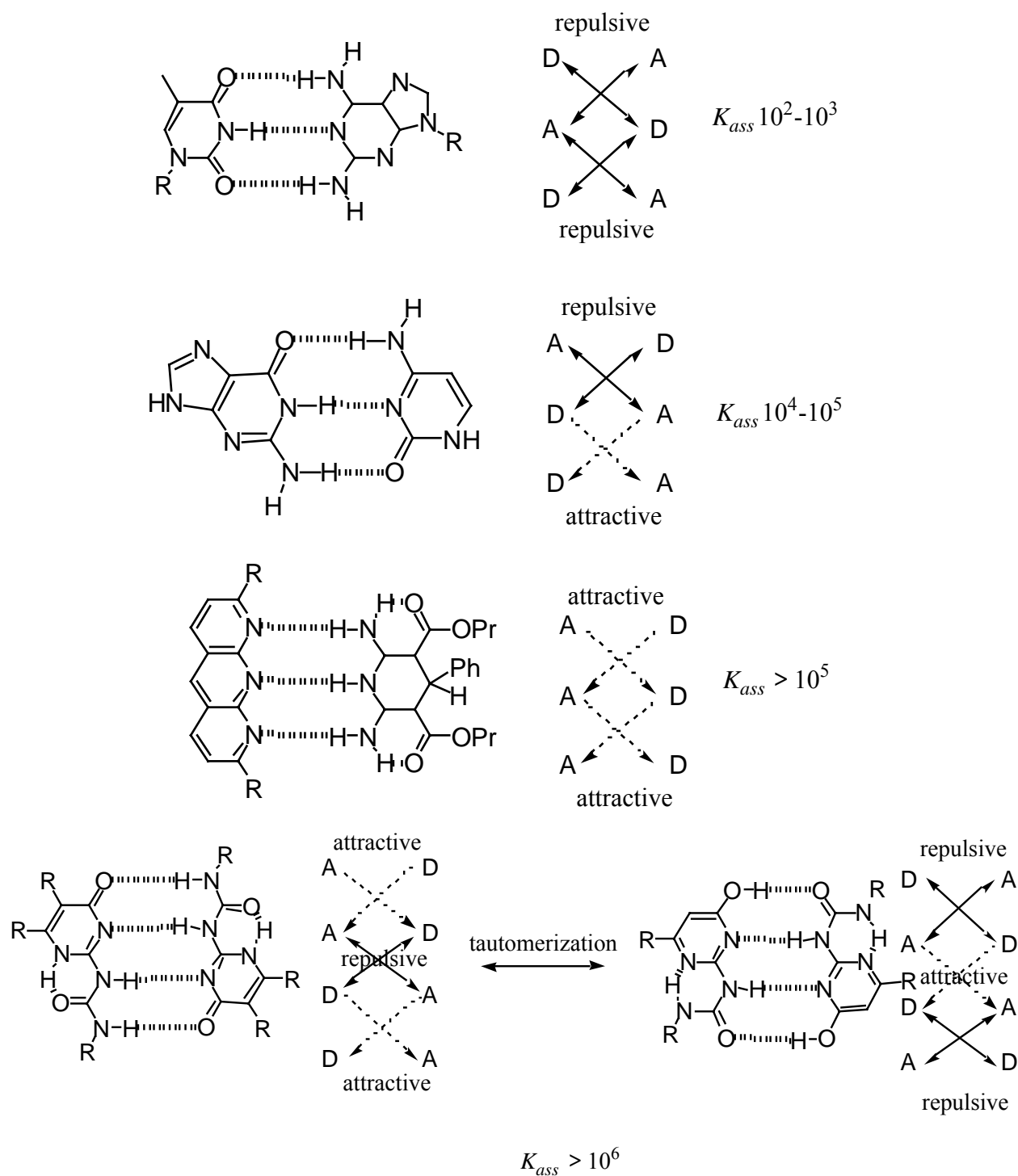
(e) *Secondary interaction*

Associations of amide-type structures are often accompanied by secondary interactions, which have been evaluated first on the basis of charge density calculations. These explained, why the Watson-Crick-base pairing free energy between A and T is only one third of that of the G-C analogue, although G-C has only one more hydrogen bond than A-T. The reason is, that in A-T pairs there are only repulsive secondary interactions between the atoms vicinal to the primary hydrogen bond, whereas in the G-C combination there are attractive secondary interactions⁶. Systematical analysis of 58 complexes with amide-like partial structures reveals, that one can describe, the secondary interactions by a common increment of 2.9kJ/ mol and hydrogen bond, irrespective whether they are repulsive or attractive. The increment then for primary amide-type hydrogen bond increases to 7.9kJ; it must be higher than the 5kJ mentioned above, as the value is derived neglecting the more often repulsive secondary forces. All values are given for the weak hydrogen bond donor chloroform as solvent; they approximately double if one uses CCl₄ instead.



In shortened word, the particular arrangement of neighboring donor and acceptor sites is an additional factor which significantly affects the strength of the complexation. In other examples, this phenomenon was recognized for the association of linear arrays of three hydrogen bonding sites; whereas complexes between common ADA-DAD motif exhibit an association constant of around 10^2 M^{-1} in chloroform⁷, this value is around 10^4 M^{-1} in complexes with DAA-ADD motif, while AAA and DDD arrays exhibit association constants 10^5 M^{-1} ⁸. Detailed calculations showed that this effect is due to differences in secondary interactions between same motifs. In the complexes, diagonally opposed sites repel each other electrostatically when they are of the same kind, whereas disparate sites attract each other. In the DDD-AAA motif the number of attractive secondary interactions is maximized, and in the ADA-DAD motif the number of repulsive interactions is at its largest.

Recently, very stable complexes can be obtained by introducing the quadruple hydrogen bonding units. Self-complementary quadruple hydrogen bonding units based on mono-ureido derivatives of amino triazines (DADA-array) with a dimerization constant $2 \times 10^4 \text{ M}^{-1}$ and hydrogen bonding units based on 2-ureido-4[1H]-pyrimidinones (DDAA), which dimerize in chloroform with an association constant of $6 \times 10^7 \text{ M}^{-1}$ were reported. Zimmermann and co-workers have reported a very stable self-complementary quadruple hydrogen bonding unit, in which all tautomers can dimerize via quadruple hydrogen bonds. In contrast, aspects of multiple hydrogen bonding units are the self-complementarity of DADA and DDAA arrays and the possibility of tautomerism. The latter may lead to loss of complexation when complementarity is lost, or when a DDAA array tautomerizes to a DADA array with a higher number of repulsive secondary interactions^{9 10}.



1-2 Molecular Recognition in Biological Systems

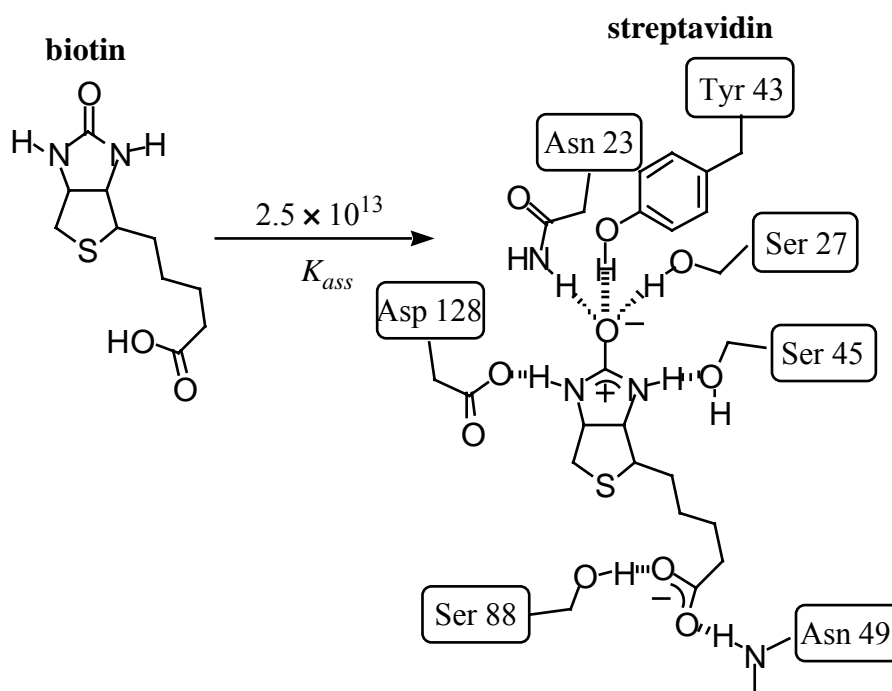
1-2-1 Introduction

Associations between host and guest molecules, H and G, are usually based on simultaneous non-covalent interactions between single binding sites, A (acceptor) and D (donor), which can be combinations like cation-anion, hydrogen bond-related acceptor-donor, lone pair acceptor-donor in the coordination bonds, and so on. Exceptions are solvent-driven equilibria and enforced guest

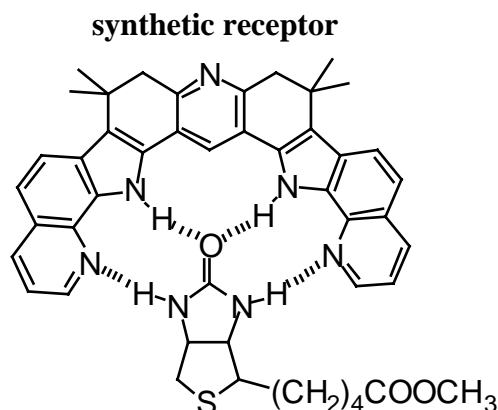
encapsulations within the closed host cavity. The need for several binding sites is quite evident: non-covalent interactions are usually weak, and concerted interplay between many sites is the only way to achieve strong and specific complexation (recognition) of a guest molecule. The principle multi-site complexation is very general in living system, where it ensures the efficiency of replication of enzyme-substrate and antigen-antibody interactions, as well as of other important biological functions. On the other hand, one can view multi-site complexation as a generalized chelate effect, which is of course well-known from coordination chemistry. An important requirement for multi-site binding is complementarity between binding sites of host and guest molecules. In another word, complexation will be most efficient when the shapes and arrangements of binding sites in host and guest molecules fit each other. This is the general lock and key principle of Fischer, who had already explained the remarkable specificity of the enzyme catalysis a century ago. In this section, the mechanism of the molecular recognitions in biological systems will be mentioned, by taking the biotin-recognition of streptavidin, or collagen or oligopeptides-recognition of integrin families.

1-2-2 Biotin recognition ability of the streptavidin

A good and traditional example of extremely efficient biological multi-site binding of a low molecular weight guest by a protein host is biotin and streptavidin, which have an equilibrium constant as ($K=2.5 \times 10^{13} \text{M}^{-1}$)¹¹, as indicated below.



A very large binding free energy of -76kJ/mol results here from simultaneous action of more than 10 weak pairwise van der Waals, electrostatic, hydrogen bonds and some hydrophobic interactions. A synthetic receptor for biotin uses only hydrogen bonds for the guest recognition and therefore exhibits only $9.3 \times 10^3 \text{M}^{-1}$ even in the less polar solvent CDCl_3 , which favors hydrogen bonds¹².



$$K_{ass} = 9.3 \times 10^3$$

1-2-3 General aspect of the functions of Integrin family

Integrins are the major receptors for cell adhesion to extracellular matrix proteins and cell-cell adhesions. In addition to mediating cell adhesion, integrins make transmembrane connections to the cytoskeleton and activate many intracellular signaling pathways. Since the recognition of the integrin receptor family was studied around 15 years ago, they have become the best-understood cell adhesion receptors (More than thousand of literatures dealing the integrin family are published per year.) Integrins and their ligands play key roles in development, immune responses, leukocyte traffic, hemostasis, and cancer and are at the heart of many human diseases genetic, autoimmune, and others (Table 1-5). Moreover the progress of the NMR or mass spectrometries X-ray analyses allowed the rapid advances in understanding integrin structure and function because of the elucidation of the 3D structures of one integrin and parts of others. These structural analyses have revealed some surprises but are also beginning to make sense of an enormous body of prior data on integrins¹³.

In Figure 1-1, a brief overview of the integrin was summarized. One set (blue in Figure 1-1) recognizes the tripeptide sequence, RGD, in molecules such as fibronectin and vitronectin in vertebrates, whereas the other set (purple in Figure 1-1) mediates adhesion to basement membrane laminins.

Figure 1-1 shows the complete mammalian set, comprising 8 β and 18 α subunits, so far known to assemble into 24 distinct integrins. Orthologs of more than half these subunits have been found only in chordates, including most of the β subunits and all the nine α subunits that have an extra inserted domain, known as an I or A domain. In addition to the ancient RGD and laminin receptor subfamilies, vertebrates have a set of collagen receptors with inserted I/A domains ($\alpha 1$, $\alpha 2$, $\alpha 10$, $\alpha 11$) and a pair of related integrins ($\alpha 4 \beta 1$, $\alpha 9 \beta 1$), which recognize both ECM proteins such as fibronectin and Ig-super-family cell surface counter-receptors. Most integrins recognize relatively short peptide motifs such as RGD sequences and, in general, an essential constituent residue is an acidic amino acid (Glu, Asp). And more precise ligand specificities are determined by the combination of both subunits of a given a heterodimer. For example, $\alpha 2 \beta 1$ integrin recognized well the Type I collagen, whereas the $\alpha 1 \beta 2$ integrin recognizes well the Type IV integrin.

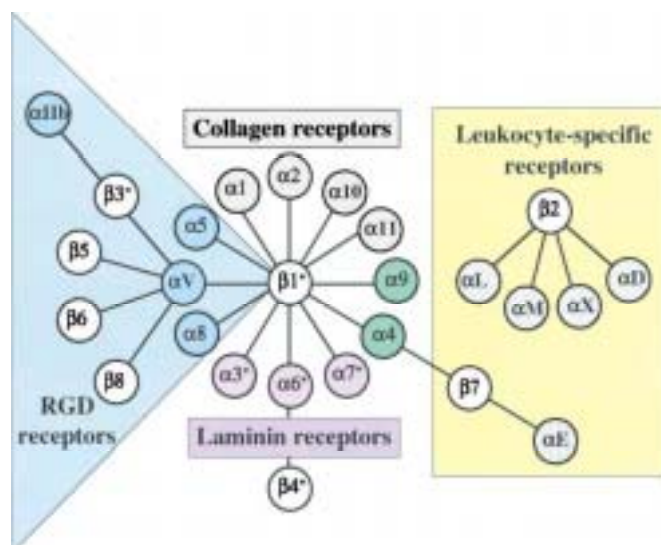


Figure 1-1 The Integrin Receptor Family Integrins are $\alpha\beta$ heterodimers; each subunit crosses the membrane once, with most of each polypeptide (> 1600 amino acids in total) the extracellular space and two short cytoplasmic domains (20-50 amino acids). The figure depicts the mammalian subunits and their $\alpha\beta$ associations; 8 β subunits can assort with 18 α subunits to form 24 distinct integrins. These can be considered in several subfamilies based on evolutionary relationships (coloring of α subunits), ligand specificity and, in the case of $\beta 2$ and $\beta 7$ integrins, restricted expression on white blood cells. α subunits with gray hatching or stippling inserted I/A domains. Such α subunits are restricted to chordates, as are $\alpha 4$ and $\alpha 9$ (green) and subunits $\beta 2$ - $\beta 8$. In contrast, α subunits with specificity for laminins (purple) or RGD (blue) are found throughout the metazoa and are clearly ancient. Asterisks denote alternatively spliced cytoplasmic domains. (from *cell*, 110, 2002, P674, figure 1)

The characteristic of the recognition ability of the integrin is to depend on the existence of the divalent cation such as Ca, Zn, Mn, and Mg. For example, the recognition of the collagen by the $\alpha 2\beta 1$ depends on the Mg, or Mn, which is suppressed by the addition of the RKK peptide¹⁴, which has cationic domain. Moreover, mutagenesis in Asp151, Asp254, Tyr221 result the loss of the recognition ability of the integrin. Now, this region was determined as the metal ion binding site of the integrin, which play a key role for the collagen recognition ability. The region, which contains generally the DXSYS sequences, is called as metal-ion dependent adhesion site (MIDAS). This is the essential domain of the recognition ability as overviewed the examples indicated in the next section.

1-2-4 Structural basis of collagen recognition by integrin $\alpha 2\beta 1$

The $\alpha 2\beta 1$, or $\alpha 1\beta 2$ integrin is known to recognize the collagen triple helix. Especially the $\alpha 1\beta 2$ integrin was included in the surface of the platelet, and plays an important role for the recognition of the wound part of the blood vessels, via recognition of the exposed collagen. The primary structure of the $\alpha 2\beta 1$ was revealed by Kamata et al and the critical site for the collagen recognition was determined as Asp151-Glu360. Especially the point mutation at Asp151, Asp254, and Tyr221 resulted the loss of the function, suggesting these residues were essential for the metal-ion mediated recognition of the collagen¹⁵. This site is now called MIDAS. The crystal structure of the $\alpha 2\beta 1$ was first reported in 1997. The crystal structure suggested the Asn154, Asp159, Leu220, Glu256, His258, Tyr285, Asp289, Leu291, Asn295, and Lys298 would contact the collagen triple helix.

On the other hand, a specific collagen motif recognized by the $\alpha 2$ -I domain has recently been identified as the hexapeptide GFOGER or GFOGVEGPOGPA (O; hydroxyproline) sequence within the context of a collagen triple helix. This finding has allowed the design of a collagen fragment suitable for co-recrystallization.

Very recently the shortest peptide folding into a stable triple helix, such as 21-residue peptide with the sequence [Ac(GPO)₂GFOGER(GPO)₃-NH₂], was synthesized as model compound of a collagen. This peptide was structurally identical to the native collagen, therefore co-recrystallization with integrin $\alpha 2\beta 1$ enabled to depict the mechanism for collagen, as follows¹⁶.

The collagen peptide binds across the front upper edge of the I domain, with a footprint 25 Å long and 10 Å wide. The interface buries a total of 1230 Å² of the collagen and I domain surfaces (Figures 1-2). The three MIDAS loops that coordinate the metal ion also provide the side chains that engage the collagen, forming a complementary surface with the contours of the triple helix. The majority of interactions are with the middle strand of the collagen, with fewer from the trailing strand and none from the leading strand. The middle strand glutamate coordinates the metal, while the middle strand arginine salt bridges to D219 from loop 2. The middle strand phenylalanine sits on a surface dimple formed by the I domain side chains Q215 and N154. The trailing strand phenylalanine makes hydrophobic contacts with Y157 and L286, while the arginine lies in an acidic pocket close to conserved E256, but is not salt bridged. Direct H-bonding interactions with the collagen main chain are from N154 and Y157 in loop 1, and from H258 in loop 3 (Figure 1-2B).

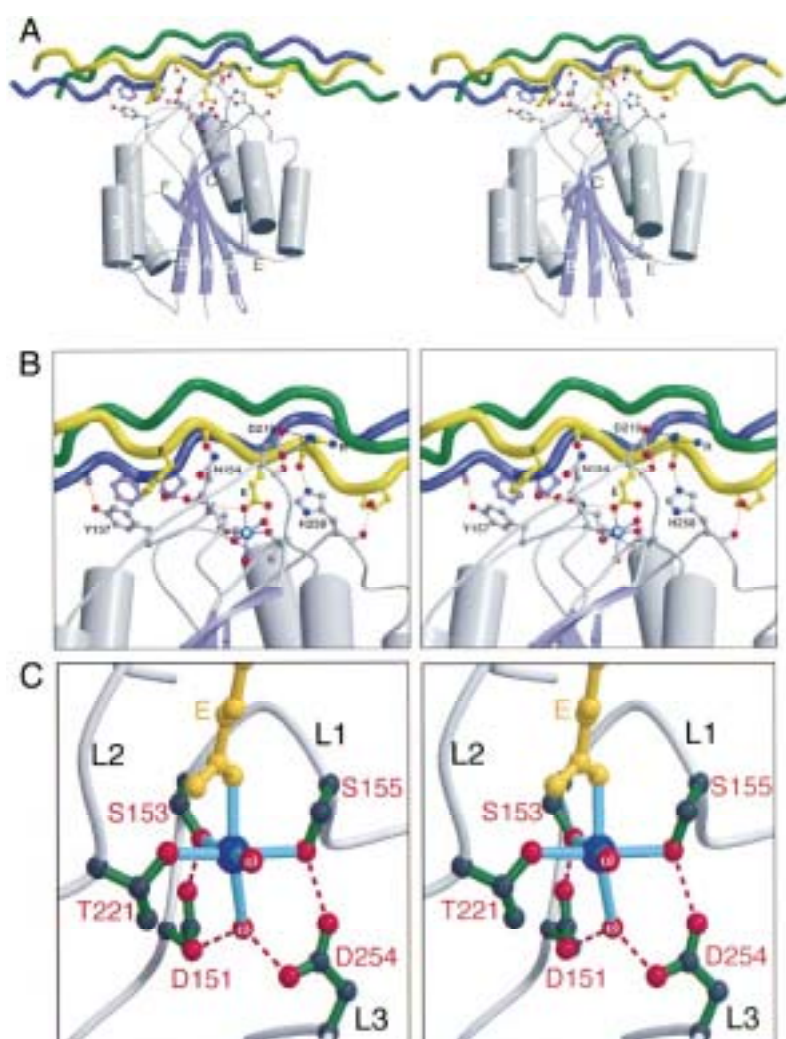


Figure 1-2 Structure of the I domain: collagen complex

(A) Stereo diagram of the $\alpha 2$ -I domain in complex with the collagen peptide. The I domain helices are shown as cylinders, β strands as arrows. The three strands of the collagen triple helix are shown as colored ribbons: leading strand in green, middle strand in yellow, and trailing strand in blue. Figures drawn with MOLSCRIPT, RENDER, and RASTER3D

(B) Close-up of Figure 1-6A, showing details of the I domain collagen interface. Selected side chains are shown as ball-and-stick, with H bonds as dotted lines. The metal ion is shown as a blue ball labeled "M." The principal interactions with the middle strand GFOGER motif (yellow) are: phenylalanine makes van der Waals contacts with side chains of N154 and Q215; the hydroxyproline carbonyl hydrogen bonds to N154; the glutamate bonds to the metal and H-bonds to T221; the arginine side chain salt bridges to D215, while its carbonyl H-bonds to H258. The principal interactions with the trailing strand GFOGER motif (blue) are: the main chain carbonyl preceding the GFOGER motif H-bonds to Y157; the phenylalanine makes van der Waals contacts with L286 and Y157; the hydroxyproline H-bonds to N154 main chain; the arginine makes weak ionic interactions with E256. The leading strand (green) makes no contacts with the I domain.

(C) Stereo diagram of the MIDAS motif. The metal ion is shown as a blue ball. Coordinating side chains are shown as ball-and-stick, with oxygen atoms in red, carbon in black. Water molecules are labeled "v"; the collagen glutamate is in gold. The three loops (L1, L2, and L3) coordinating the metal are shown schematically as gray ribbons. E256 from L3, which forms an indirect bond via the equatorial water, has been removed for clarity. The figure is rotated about a vertical axis by 180° relative to (B). (from *cell*, 101, 2000, P49, figure 1)

1-2-5 RGD tripeptide recognition by Integrin $\alpha V\beta 3$

RGD recognition ability is very important especially for the hemostasis, where the platelet co-aggregate with fibrinogen, via recognition of the RGD sequences by the $\alpha_{IIb}\beta_3$ integrin. The RGD recognition ability also depends on the MIDAS motif likely to the collagen recognition ability by $\alpha 2\beta 1$. The structural basis for the divalent cation dependent binding of heterodimeric $\alpha\beta$ integrins to their ligands, which contain the prototypical Arg-Gly-Asp sequence, was unknown¹⁷ till a few years ago, although for the α_{IIb} integrin, the essential location for the recognizing the HHLGGAKQAGDV sequence in fibrinogen was identified as Leu²⁴³-Val²⁵⁴, Asp³⁰⁷-Val³¹⁸, Asp³⁶⁵-Val³⁷⁶, and Asp⁴⁵⁷-Val⁴⁶⁸ by the mutagenesis experiments. This region includes have DXDXD(N)GXXDXXD or DXSXS sequences which are characteristically seen in MIDAS motif. The oligopeptides recognition ability of integrin was not completely known except: they all use an acidic residue (MIDAS site) and the recognition is divalent-cation dependent and the specificity for a particular ligand is then determined by additional contacts with the integrin.

Recently the technology of the peptide synthesis becomes sufficiently improved that various oligopeptide which has a macrocyclic structure and stable β -turn at Gly residue, which is essential for the integrin recognition ability, have been synthesized. Especially Arg-Gly-Asp-{D-Phe}-{N-methyl-Val-}¹⁸, called cyclo(RGDF5N{Me}V), which have stabilized β -turn at Gly residue eased recrystallization of the integrin-RGD complexes. This complex has depicted the structural basis for RGD recognition ability by the integrin $\alpha V\beta 3$, as follows¹⁹.

(a) contact of the Arg residue

The Arg side chain inserts into a narrow groove at the top of the propeller domain. The arginine guanidinium group is held in place by a bidentate salt-bridge to Asp218 at the bottom of the groove and by an additional salt bridge to Asp150 at the rear. The contacts leave most of the upper portion of the Arg side chain exposed to solvent, whereas the spacious rear of the groove probably contains water molecules that may provide additional contacts to the Arg guanidinium group.

(b) contact of the Asp residue

Unlike the ligand Arg, the ligand Asp side chain is completely buried in the complex. Contacts between the ligand Asp and βA primarily involve the Asp carboxylate group, which protrudes into a cleft between the βA loops A9- $\alpha 1$ and C9- $\alpha 3$ and forms the center of an extensive network of polar interactions. One of the Asp carboxylate oxygens contacts a Mn21 ion at MIDAS in βA . The second Asp carboxyl oxygen forms hydrogen bonds with the backbone amides of Tyr122 and Asn215, and also contacts the aliphatic portion of the Arg214 side-chain. Additional contacts involve the hydrophobic portion of the Asp side chain and the beta carbon atom of Asn215.

(c) contact of the Gly residue

The glycine residue lies at interface between the α and β subunits. It makes several hydrophobic interactions with αV , the most critical of which appears to be the contact with the carbonyl oxygen of Arg216.

The RGD-recognition ability of the integrin $\alpha V\beta 3$ is very strange comparing to another integrin family. Since a ligand-associated metal binding site (LIMBS) consisted of oxygen of Glu220; the side chains of Asp158, Asn215, and Asp217; and the carbonyl oxygens of Asp217 and Pro219, controls the recognition ability. When the divalent cation such as Mn was complexed with the LIMBS, the structure of the MIDAS region or ligand-contact region was optimized for the ligation.



Figure 1-3 The ligand-integrin binding site. Interactions between ligand and integrin. The peptide (yellow) and residues interacting with the ligand or with Mn2+ ions are shown in ball-and-stick representation. αV and $\beta 3$ residues are labeled blue and red, respectively. Oxygen and nitrogen atoms are in red and blue, respectively. The three Mn2+ ions in $\beta 3$ at MIDAS, ADMIDAS, and LIMBS are also shown. Hydrogen bonds and salt bridges (distance cutoff, 3.5 Å) are represented with dotted lines. (from *science*, 296, 2002, P154, Figure 2B)

1-2-6 The mechanism of the vWf-recognition ability of platelet GPIb α

Transient interactions of platelet-receptor glycoprotein Ib α (GPIb α) and the plasma protein von Willebrand factor (VWF) reduce platelet velocity at sites of vascular damage and play a role in haemostasis and thrombosis. Transient interactions of platelet-receptor glycoprotein Ib α (GPIb α) and immobilized von Willebrand factor (VWF) mediate the rolling of platelets at sites of vascular damage. Rolling reduces platelet velocity and prolongs the contact time with reactive components of the cell matrix. This facilitates platelet activation and subsequent integrin-mediated firm attachment. When the GPIb α □ □ □ □ □ □ □ □ were lacked (loss of function) or emphasized (gain of function), the Bernard-Soulier syndrome and type 2M von Willebrand disease comes arise. Although the crystal structure of VWF-A1 is known and there is a large body of mutagenesis data, the precise interactions between GPIb α and A1 of vWF, the mechanism of shear-induced activation, and the molecular basis of related bleeding disorders are, however, poorly understood²⁰. Very recently, using the

gain-of-function mutants of GpIb α (M239V) $\square\square\square$ co-crystallized with the A1 domain of the vWF. The crystal structure of the complex revealed the structural basis of the vWF recognition ability of GpIb α , as follows²¹.

The crystal structure of the VWF-binding domain of GpIb α displays an elongated, curved shape (Figure 1-4) that is typical for proteins containing leucine-rich repeats. Eight short leucine-rich repeats, seven of which were predicted on the basis of the amino acid sequence, make up the central region of the molecule.

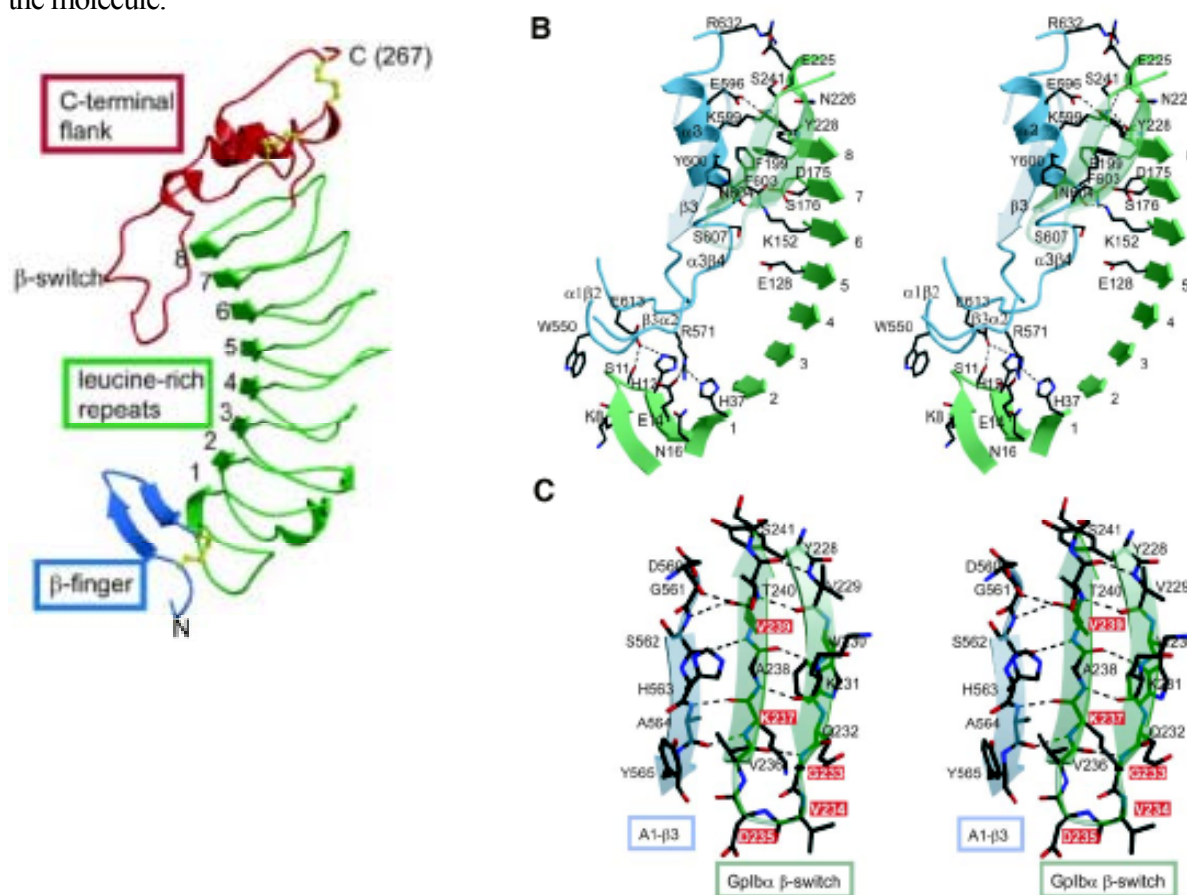


Figure 1-4 (A) Structures of the VWF-binding domain of GpIb α and the complex of GpIb α with the A1 domain of VWF. Ribbon representation of GpIb α . The NH₂-terminal α hairpin, called β finger is colored red and contains a disordered loop (residues 227 to 241) called β switch. Disulfide bridges are indicated in yellow ball-and-stick representation. Structures of the VWF-binding domain of GpIb α and the complex of GpIb α with the A1 domain of VWF. (B) Stereo view of a ribbon representation of the GpIb α -A1 complex. GpIb α de bridge in A1 shown in yellow ball-and stick representation. The β -switch of GpIb α adopts a β -hairpin structure that aligns with the central β sheet of A1. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (C) Stereo representation of residues at the A1-GpIb α interface, excluding residues located in the β sheets (from *science*, 297, 2002, P1177, Figure 1)

The crystal structure of the complex GpIb α -VWF-A1 shows that the globular A1 domain interacts with the concave face of GpIb α (Figure 1-4). The interaction surface is extended but discontinuous, comprising two distinct areas of tight interactions (Figure 1-5B). The first and most

extensive contact site is located near the top of A1 and buries a solvent-accessible surface of $\sim 1700 \text{ \AA}^2$. At this contact site, leucine-rich repeats 5 to 8 and the COOH-terminal flank of GpIb α interact with A1 helix $\alpha 3$, loop $\alpha 3\beta 4$, and strand $\beta 3$ (Figure 1-4B and C). In this region, electrostatic interaction play a key role for the complexation.

The second and smaller contact site buries a surface of $\sim 900 \text{ \AA}^2$ and involves interactions of the NH₂ terminal β -finger and the first leucine-rich repeat of GpIb α with loops $\alpha 1\beta 2$, $\beta 3\alpha 2$, and $\alpha 3\beta 4$ at the bottom face of the A1 domain. The flexible loop from 227 to 241 (called as β -switch) in the COOH-terminal flank of GpIb α undergoes a conformational change on complex formation. In the complex, this loop forms a 16-residue β -hairpin that extends from residues 227 to 242 and aligns with the $\alpha 3$ strand of A1 (residues 562 to 566), forming a continuous β -sheet shared between the two molecules (Figure 1-4C). The specific mutation M239V enhances binding ability via two direct intermolecular main-chain hydrogen-bonds. Overall, the conformational change of the β -switch in this region controls the affinity of the interacting partners.

1-3 Supramolecules Using Hydrogen Bonds

Supramolecular chemistry is concerned with the coalescence of molecules into noncovalent arrays, in another word, molecular recognition. To architect the supramolecules, weak, noncovalent interactions such as hydrogen bonds, coordination bonds, arene-arene interactions, and hydrophobic interactions are usually used. In this section, mainly molecular capsules using hydrogen bonds will be described as typical examples for the supramolecules because the same methodology is applicable for designing the hemoglobin model.

In order to exert the proteins' functions, the cavity structure is very often required, to stabilize the substrate-enzyme complex, or to protect the enzyme from outer environment. Therefore the molecular capsule is good target for mimicking these functions. Rebek's, and Rheinhoudt's groups are the pioneers of this field and have designed and synthesized a lot of molecular capsules with full of originality, as listed below.

(i) calix[4]arene based molecular capsule

Calix[4]arenes, having basket-shaped structure provides a versatile binding pocket for a number of guest, such as metal ions or aromatic molecules, by cation-arene, arene-arene interaction. The most typical example is **tetracarboxylcalix[4]arene-tetrapyridylcalix[4]arene duplex I**²² synthesized by Rheinhoudt although the complex ability with the guest molecule are not yet reported. Similar **calix[4]arene-based molecular capsules II, III**^{23, 24} are also reported by Rebek, which are consisted of the upper-rim substituted calix[4]arene and possesses the small molecule encapsulation ability. this calix[4]arene is substituted to four urea groups and dimerize via sixteen-fold hydrogen bonds. The dimerization constants are over 10^6 M^{-1} . Interestingly, all eight ureas are fixed in the same direction, and the sixteen-fold hydrogen bonds slows rotation about the calix[4]arene-urea bond resulting in an isomer of D_{4d} symmetry. The existence of the cavity structure between the two calix[4]arenes has been proven by the detection of the encapsulation of the solvent, which exchanges from/ to outer solvent sufficiently slowly to detect separated signals in ^1H NMR spectra. The suitable guest for this capsule is smaller than benzene, for instance, benzene, pyridine, fluorobenzene, pyridine, *p*-difluorobenzene. As written above, the size of the cavity of the calix[4]arene-based molecular capsule is too small for mimicking the enzyme activity, since the at least two molecules must embedded inside the cavity. Therefore to enlarge the size of the cavity, another molecular design is required.

(ii) resorcinol cyclic tetramer (resorcinarene) based molecular capsule

Resorcinarene resembles to calix[4]arene in its structure, therefore we can expect to construct the molecular capsule by dimerizing the resorcinarene. In fact, the **resorcinarene dimer IV**²⁵ based molecular capsule is reported, which is consisted of resorcinarene bearing four urea groups on the upper rim and stabilized via sixteen-fold hydrogen bonds. The suitable guest is xylenes, or norbornene derivatives, which is slightly larger than benzene. However for the enzyme model, the size of the cavity

inside the two resorcinarene is insufficient. Recently to form the larger cavity, the very unique molecular design is proposed. The **self folding cavitands**²⁶ **V**, **VI**, **VII**, have the largest cavity size among the unimolecular capsules. The cavity structure was formed by the U-shaped structure which is stabilized twelve-fold hydrogen bonds. The size was approximately 800\AA^3 which corresponds to the size of adamantane or cyclohexyl derivatives whose length is approximately 15\AA . Furthermore, Zinc porphyrin based self-folding cavitand has largest cavity (dimensions $10 \times 25\text{\AA}$). This host can encapsulates the adamantane and pyridine derivatives due to the sufficient size and coordination ability between zincporphyrin-pyridine. This is the first example for the enzyme mimetic, since we can expect two molecules are encapsulated inside the cavity.

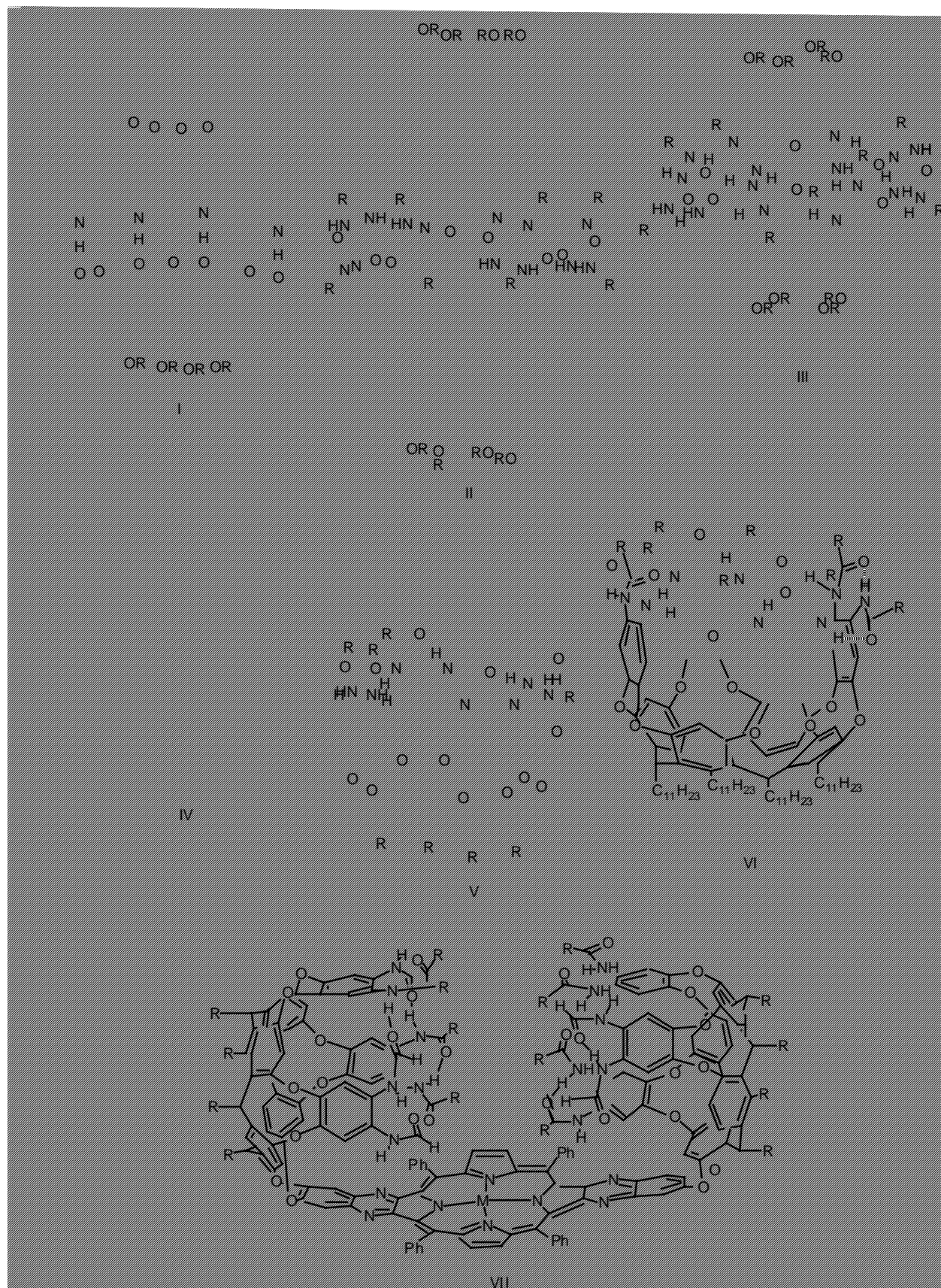
(iii) Glycouril-based molecular capsule

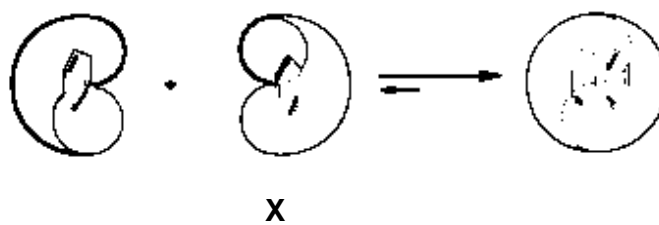
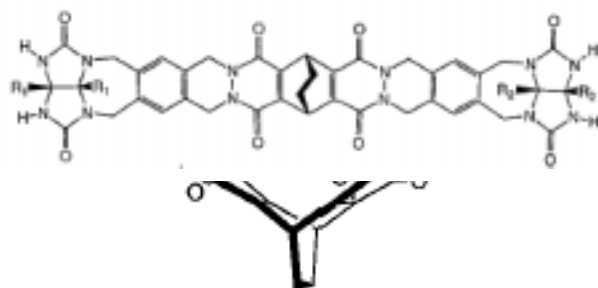
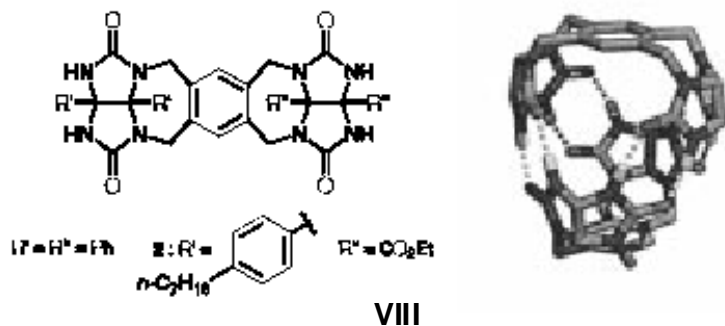
There are two types of glycouril-based molecular capsule. The one is cucurbituril and the other is seamed spheres.

Cucurbituril is a rigid **cyclic glycouril hexamer** with a central hallow core of diameter 5.5\AA . The upper and lower rims of the macrocycle are lined with inwardly pointing carbonyl groups. By virtue of strong electrostatic interactions, cucurbuturil is capable of binding strongly to aliphatic and aromatic ammonium ions.

There are many examples of seamed spheres type of glycouril-based molecular capsule, which are designed by Rebek. The “**tennis ball**²⁷” **VIII** and “**base ball**” are consisted of two diphenylglycouril units linked by a durene spacer. The size of the cavity corresponded to the small molecules such as ethylene or methane. the original tennis balls derivatives are also reported which is enlarged by changing the length of the spacer. The “**softballs IX**²⁸” is designed for the encapsulation of the larger molecules, by expanding the spacer between the two glycouril units while still keeping an appropriate shape to allow dimerization to occur. The size of the cavity corresponds to the 1-adamantanecarboxylic acid or 1-ferrocenecarboxylic acid. Very surprisingly, Rebek and co-workers succeeded to act this molecular capsule as a reaction chamber. When *p*-quinone and cyclohexadiene were embedded inside the cavity together, a Diels-Adler reaction has taken place as indicated **X**, nonetheless the efficiency, and turn-over are now desirable. To my best knowledge, this is, however the first example to demonstrate the chemical reactions inside the cavity fabricated by the synthetic molecular capsule.

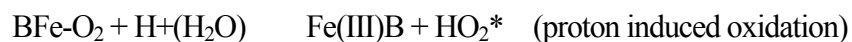
As overviewed above, it is the manifest destiny that we will be able to control the size of the cavity, and apply to the reaction chamber within the molecular capsule. These series of the molecular capsule gave me many impressive intimations for designing the hemoglobin model, since if I can control the cavity size inside the porphyrin-based molecular capsule, I will be able to adjust the oxygen binding ability of the porphyrinatoiron(II)!





1-4 Cavity Construction on the Porphyrin Plane via Covalent Bonds

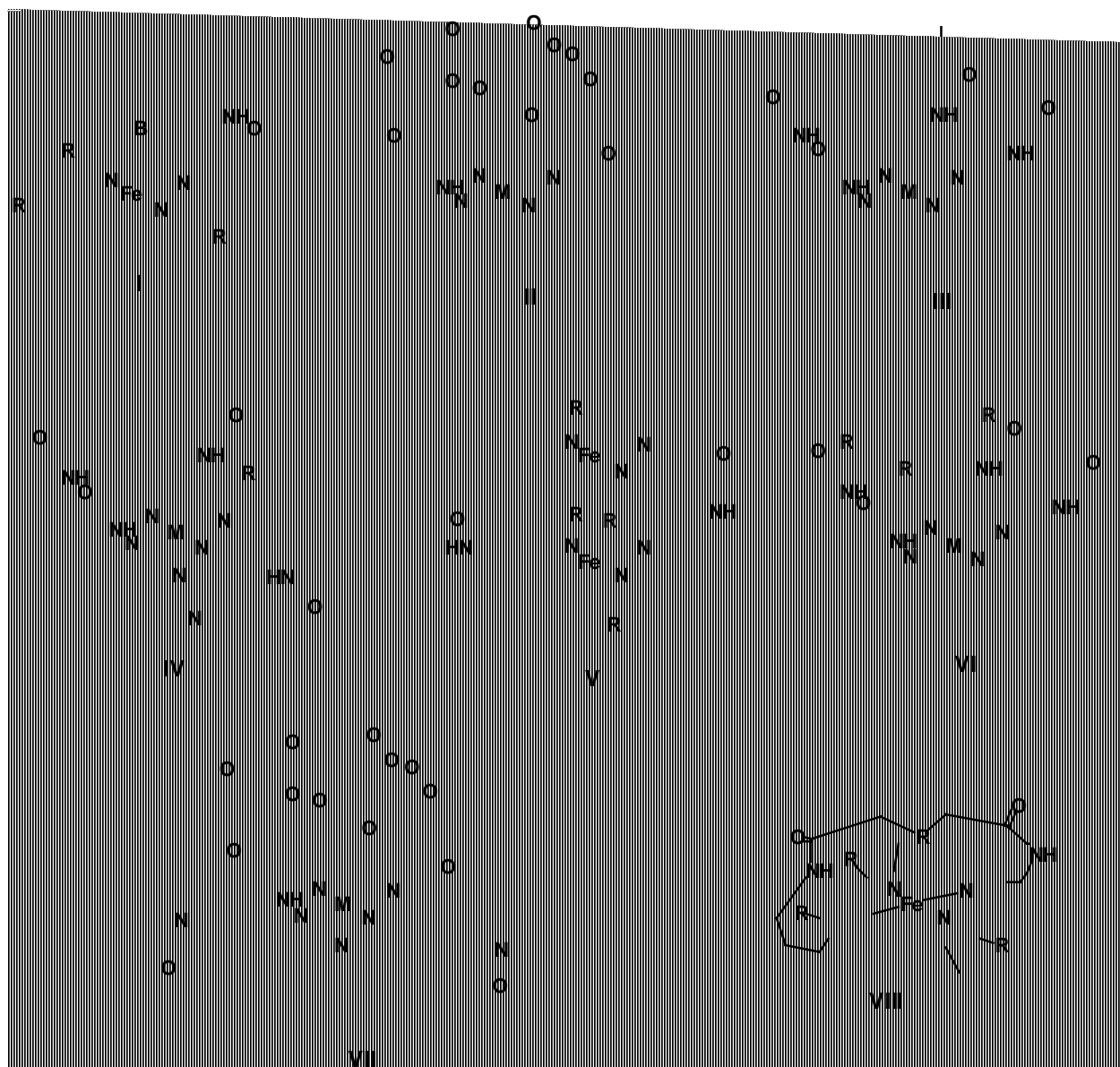
The functional center of the reversible dioxygen binding ability of Hb is iron(II) complex of protoporphyrin IX (PPIX), which ruled by the surrounding environment. If the porphyrinatoiron(II) exposed to the outer environment following two reactions would occur, which plunder the oxygen-binding ability from porphyrin by oxidation to iron(III).

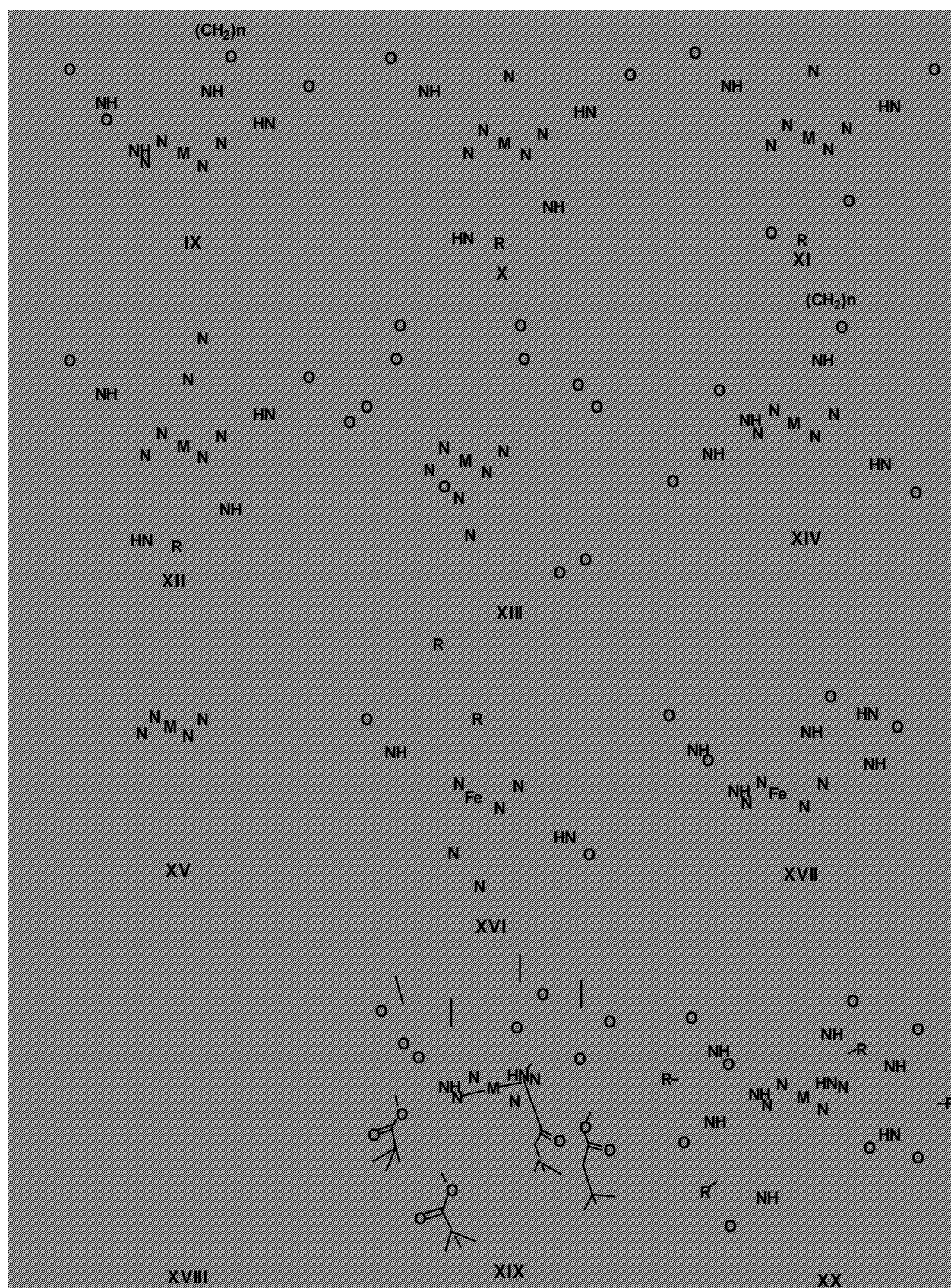


To prevent these reactions, the porphyrinatoiron(II) must be embedded in the hydrophobic and isolated microenvironment. In fact, the four PPIX s are included in the heme-pocket inside the globin, to exhibit reversible oxygen-binding ability.

To mimic these functions, we have only to provide the same environment for the porphyrin. Much effort has been directed to mimic this function mainly by modifying the tetraphenylporphyrin (TPP) with the hydrophobic and bulky substituents, and realized reversible dioxygen binding. The first

example of the oxygenation of the synthetic heme **I** was reported by Chang and co-workers in 1973.²⁹, simultaneous with the report of oxygenation of another iron-complex (**II**)³⁰. And the reversible dioxygen binding ability was first demonstrated by Collman's picket-fence porphyrin (**III**)³¹. The substituents can be so easily introduced that various picket-fence porphyrin analogues and derivatives are synthesized, as listed (**IV-XV**)³². Generally the TPP is more comfortable starting compound than PPIX, because the TPP by itself is bulky due to the phenyl groups standing perpendicularly to the porphyrin plane.





1-5 The design of the Supramolecular Hemoglobin Model

The reversible oxygen binding ability of iron(II) complex protoporphyrin IX (PPIX) is exhibited by being embedded inside the isolated and hydrophobic microenvironment of the heme-pocket in the hemoglobin. The isolated environment is necessary for the prevention of the μ -oxo dimerization of the porphyrinatoiron(II). And the hydrophobic reaction field is essential for avoiding of the proton-driven oxidation of the five-coordinate porphyrinatoiron(II) to the porphyrinatoiron(III). Therefore, the reported synthetic hemoprotein model possesses the covalently-bound hydrophobic and bulky substituents. As typical example, meso-tetraphenylporphyrin (TPP) derivatives sterically-encumbered by peripheral substitutes have been widely exploited, as listed above.

On the other hand, porphyrinatoiron(II) included in hemoglobin indicates a very complexed function, namely, cooperative oxygen-binding ability (allosteric effect) grounding the sigmoidal oxygen equilibrium curve of hemoglobin, which is allowed by the drastic structural changes of the subunits triggered and amplified by the minute distortion of a porphyrin plane induced by the oxygen-binding and to the heme-iron(II). Seeing this phenomena more microscopically, the diminish of the radius of the iron(II), induced by the dioxygen binding to deoxy-form iron(II) enable to locate the iron(II) at almost parallel to porphyrin plane. These structural changes of hemoprotein absolutely rely on the supramolecular structure, since such rearrangement of the heme-pocket must be occurred by a minute stimuli. Therefore, constructing a real hemoprotein model must bear the supramolecular structure, in another word, must be based on non-covalent architectures.

The supramolecular architecture based on the porphyrin, the preorganized (macrocyclic) molecules are appropriate. As the candidate, macrocyclic peptide³³, calix[4]arene, and crown-ether³⁴ are very potent. However, for macrocyclic peptides, there is a disadvantage for their solubility, and easily self-assembled to form the tubal supramolecule. In another word, they would not form the duplex with porphyrin but self-assembled. The crown-ether is so flat that it might not indicate the cavity structure. Therefore I want to propose hydrogen bonded porphyrin-calix[4]arene duplex. The calix[4]arene is the cavernous molecule whose size is almost identical to the porphyrin (15Å) and bears the properly flexible structure that would be able to change together with the minute distortion changes of porphyrin plane, by the coordination of dioxygen. Furthermore calix[4]arene have eight potent positions to be substituted, namely four hydroxyl groups at lower rim, and four para-position of phenyl ring with high electro-density at the upper rim. The size of the cavity is expected to be small but sufficient to involve the dioxygen molecule. And the last advantage of the calix[4]arene is that it is very synthetically accessible macrocycles due to the selective macrocyclization under the Zincke-Cornforth's conditions, although generally the syntheses of the macrocycle often suffers the low yield.

LITERATURE CITED

1. hydrogen bonding vinogradov/Linnell
2. Principles and methods in supramolecular chemistry WILEY, edited by H-J, Schneider, A. Yatsimirsky
3. J. Rebek Jr, *Angew. Chem. Int. Ed. Engl.* **1990**, 29, 245
4. S. C. Zimmerman, W. Wu, Z. Zeng, *J. Am. Chem. Soc.* **1991**, 113, 196
5. (a) H-J. Schneider, *Chem. Soc. Rev.* **1994**, 23, 227, (b) H-J. schneider, *Angew. Chem. Int. Ed. Engl.* **1991**, 30, 1417
6. J. Pranata, S. G. Wierschke, W.L.Jorgensen, *J. Am. Chem. Soc.* **1991**, 113, 2810
7. L. Brunsveld, B. J.B. Folmer, E.W.Meijer, and R. P. Sijbesma, *Chem. Rev.*, **2001**, 101, 4071
8. (a) W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, 112, 2008, (b) J. Pranata, S. G. Wierschke, W.L.Jorgensen, *J. Am. Chem. Soc.* **1991**, 113, 2810
9. For example, (a) H. Zeng, R.S.Miller, R. A. Flowers, B. Gong, *J. Am. Chem. Soc.*, **2000**, 122, 2635, (b) B.J.B. Folmer, R.P.Sijbesma, H. Kooijman, A. L. Spek, E.W.Meijer, *J. Am. Chem. Soc.*, **1999**, 121, 2001, (c) J.B. Brigittee, B.J.B. Folmer, R.P.Sijbesma, E.W.Meijer, *J. Am. Chem. Soc.*, **2001**, 123, 2093, (d) J.B. Brigittee, R.P.Sijbesma, A. J. M. Vekemans, E.W.Meijer, H. Kooijman, A.L.Spek, *J. Org. Chem.* **1996**, 61, 6371
10. (a) F. H. Beijer, r. P. Sijbesma, H. Kooijman, A.L.Spek, and E.W.Meijer, *J. Am. Chem. Soc.* **1998**, 120, 6761, (b) F. H. Beijer, H. Kooijman, A.L.Spek, R. P. Sijbesma, and E.W.Meijer, *Angew. Chem. Int. Ed. Engl.* **1998**, 37, 75, (c) S. H. M.Sontjens, R. P.Sijbesma, M. H. P. van Genderen, E.W.Meijer, *J. Am. Chem. Soc.*, **2000**, 122, 7487
11. (a) P. C. Weber, J. J. Wendoloski, M.W. Pantliano, F. R. Salemme, *J. Am. Chem. Soc.*, **1992**, 114, 3197, (b) T. Sano, C. R. Cantor, *Proc. Natl. Acad. Sci. USA*, **1995**, 92, 3180
12. V. Hedge, C-Y. Hung, P. Madhukar, R. Cunningham, T. Hopfner, R. P. Thummel, *J. Am. Chem. Soc.* **1993**, 115, 872
13. (a) R. O. Hynes, *Cell*, **2002**, 110, 673-687, (b) R. O. Hynes, *Cell*, **1987**, 48, 549-554
14. Ivaska, J., Kapyła, J., Pentikainen, O., Hoffren, A-M., Hermonen, J., Huttunen, P., Johnson, MS., and Heino, J. *J. Biol. Chem.* 274, **1999** 3513-3521
15. There are numerous number of literature discussing the collagen recognition ability, for example, (a) Knight, CG, Morton, LF., Onley, DJ., Peachey, AR., Messent, AJ., Smethurst, PA., Tuckwell, DS., Farndale, RW., and Barnes, MJ. *J. Biol. Chem.* 273 **1998** 33287-33294, (b) Kamata, T., and Takada, YJ. *Biol. Chem.* 269 **1994**, 26006-26010, (c) Kamata, T., Puzon, W., and Takada, YJ. *Biol. Chem.* 269 **1994**, 9659-9663, (d) Nolte, M., Pepinsky, B., Venyaminov, SY., Koteliansky, V., Gotwals, PJ., and Karpusas, M. *FEBS Letters* 452 **1999** 379-385, (e) Emsley, J., King, SL., Bergelson, JM., and Liddington, RC. *J. Biol. Chem.* 272 **1997** 28512-28517
16. J. Emsley, C. G. Knight, R. W. farndale, M. J. Barnes, R. C. Liddington, *Cell*, **2000**, 47-56
17. There are numerous number of literature discussing the RGD recognition ability, for example, (a)

- Kirchhofer, D., Gilit, J., Ruoslahti, E., Grzesiak, J., and Pierschbacher, MD. *J. Biol. Chem.* 265 **1990** 18525-18530, (b) Calvete, JJ., Schafer, W., Mann, K., Henschen, A., and Gonzalez-Rodriguez, J. *Eur. J. Biochem.* 206 **1992** 759-765, (c) Kamata, T., Irie, A., Tokuhira, M., and Takada, Y. *J. Biol. Chem.* 271 **1996** 18610-18615, (d) Hu, DD., White, CA., Panzer-Knodle, S., Page, JD., Nicholson, N., and Smith, JW. *J. Biol. Chem.* 274 **1999** 4633-4639, (e) Tozer, EC., Liddington, RC., Sutcliffe, MJ., Smeeton, AH., and Loftus, JC. *J. Biol. Chem.* 271 **1996** 21978-21984
18. (a) A. C. Bach, C. J. Eyermann, J. D. Gross, M. J. Bower, R. L. Harlow, P. C. Weber, W. F. Degrado, *J. Am. Chem. Soc.*, **1994**, 116, 3207, (b) H-H, Zhang, J. A. Pesti, T. D. Costello, P. J. Sheerman, R. Ueda, P. Ma, G. S. Kaufmann, R. Ward, J. L. McMillan, *J. Org. Chem.* **1996**, 61, 5180, (c) R. Haubner, R. Gratias, B. Diefenbach, S. L. Goodman, A. Jonczyk, H. Kessler, *J. Am. Chem. Soc.*, **1996**, 118, 7461, (d) J. wermuth, S. L. Goodman, A. Jonczyk, H. Kessler, *J. Am. Chem. Soc.*, **1997**, 119, 1328, L-H, Zhang, J. C. Chung, C. I. Valvis, P. Ma, S. kaufmann, R. Ward, *J. Org. Chem.* **1997**, 62, 2466, (e) R. Haubner, W. Schmitt, G. Holtzmann, S. L. Goodman, A. Jonczyk, H. Kessler, *J. Am. Chem. Soc.*, **1996**, 118, 7881
19. J-P, Xiong, T. Stehle, R. Zhang, A. Joachimiak, M. Frech, S.L.Godman, M. A. Arnaout, *science*, **2002**, 151
20. There are many literatures concerning the GPIIb. For example, (a) Du, X., L. Beutler, CH. Ruan, PA. Castaldi, MC Berndt *Blood* **69** (1987), 1524-1527, (b) Modderman, PW., LG. Admiraal, A. Sonnenberg, AEGK von dem Borne. *J. Biol. Chem.* **267** (1992) 364-369, (c) Lopez, JA., Leung, B., Reynolds, CC., Li, CQ., and Fox, J-E, *B.J. Biol. Chem.* **267** (1992) 12851-12859, (d) Calverley, DC., Yagi, M., Stray, SM., and Roth, GJ. *Blood* **86** (1995) 1361-1367, (e) Berndt, MC., C. Gregory, A. Kabral, H. Zola, D. Fournier, PA. Castaldi. *Eur. J. Biochem.* **151** (1985) 637-649, (f) J. A. Lopez, D. W. Chung, K. Fujikawa, F. S. Hagen, T. Papayannopoulou, and G. J. Roth, *Proc. Natl. Acad. Sci USA.* **84** (1987) 5615-5619 , (g) Kenneth., J. Clemetson and Jeannine, M. Clemetson *Seminars in Thrombosis and Hemostasis* **21** (1995) 130-136 , (g) Wicki, A. N., and Clemetson, K. *J. Eur. J. Biochem.* **153** (1985) 1-11, (h) Handa M., Titani, Holland, LZ., Roberts, JR., Ruggeri, ZM *J. Biol. Chem.* **261** (1986) 12579-12585 , (i) Kobe, B. , J. Deisenhofer, *Nature* **366** (1993) 751-756, (j) Michelson AD, Loscalzo, J., Melnick, B., Collar BS., Handin RI. *Blood*, **67** (1986) 19-26 , (k) Dong, JF., Li, CQ., Hyun, W., Afshar-Kharghan , V., Lopez, JA. *Biochemistry* **36** (41) (1997) 12421-12427, (l) Titani, K., Takio, K., Handa, M., and GuggeriZ. M. *Proc. Natl. Acad. Sci. USA.* **84** (1987) 5610-5614, (m) Vicente, V., Houghten, R. A., and Ruggeri, M. *J. Biol. Chem.* **265** (1990) 274-280, (n) Murata, M., Ware, J., and Ruggeri, Z. M. *J. Biol. Chem.* **266** (1991) 15474-15480 , (o) Cruz, MA., Peterson, E., Turci, SM., Handin, RI. *J. Biol. Chem.* **267** (1992) 1303-1309, (p) Schumpp-Vonach, B., Kresbach, G., Schlaeger, EJ., Steiner, B. *Cytotechnology* **17** (1995) 133-141, There are many literatures concerning the vWf. For example, (a) Titani, K., Kumar, S., Takio, K., Ericsson, LH, Wade, RD., Ashida, K., Walsh, KA., Chopek, MW., Sadler, JE., and Fujikawa, K. *Biochemistry* **25** (1986) 3171-3184, (b) Mohri, H., Fujimura, Y., Shima, M., Yoshioka, A., Houghten, RA., Ruggeri, ZM., and Zimmerman, TS. *J. Biol. Chem.* **263** (1988) 17901-17904, (c) Cruz, MA., Yan, H, Lee, RJ., and

- Handin, R.I. , *J. Biol. Chem.* **270** (1995) 10822-10827, (d) Sadler, J.E., *J. Biol. Chem.* **266** (1991) 22777-22780, (e) Bienkowska, J., Cruz, M., Atiemo, A., Handin, R., and Liddington, R. , *J. Biol. Chem.* **272** (1997) 25162-25167, (f) Mohri, H., Ohkubo, T., Ruggeri, ZM., Zimmerman, TS. ,*Rinsyo Byori* **40** (1992) 639-644, (g) Gralnick, HR., Willimams, S., , *Proc. Natl. Acad. Sci. USA* **89** (1992) 7880-7884 , (h) Fujimura, Y., Titani, K., Holland, LZ., Russel, SR., Roberts, JR., Elder, JH., Ruggeri, ZM., and Zimmerman, TS. ,*J. Biol. Chem.* **261** (1986) 381-385 , (i) Miyata, S., Goto, S., Federici, AB., Ruggeri, ZM. , *J. Biol. Chem.* **271** (1996) 9046-53, (j) Berndt MC, Ward CM, Booth WJ, Castaldi PA, Mazurov AV, Andrews , *Biochemistry* (1992) **31** 11144-11151, (k) Knott, HM, Berndt, MC., Kralicek, AV., O'Donoghue, SI., and King, GF. , *Biochemistry* **31** (1992) 11152-11158, (l) Kroner, PA., and Frey, AB. , *Biochemistry* **35** (1996) 13460-13468
21. (a) Huizinga, E. G., Tsuji, S., Romijn, R. A. P., Schiphorst, M. E., de Groot, P. G., Sixma, J. J. and Gros, P., *Science*, **2002**, 297, 1176-1179, (b) Sadler, J. E., *Science*, **2002**, 297 , 1128-1129
22. R. H. Vreekamp, W. Verboom and D. N. Reinhoudt, *J. Org. Chem.*, **1996**, 61, 4282
23. Y.L. Cho, D. M. Rudkevich, and Julius Rebek, Jr. *J. Am. Chem. Soc.* **2000**, 122, 9868-9869
24. R. K. Castellano, B.H. Kim, and Julius Rebek, Jr., *J. Am. Chem. Soc.* **1997**, 119, 12671-1267
25. (a) S Ma, D. M. Rudkevich, and J. Rebek, Jr. *J. Am. Chem. Soc.* **1998**, 120, 4977-4981, (b) F C. Tucci, D. M. Rudkevich, and J. Rebek, Jr. *J. Am. Chem. Soc.* **1999**, 121, 4928-4929
26. (a) U. Lulcking, F. C. Tucci, D. M. Rudkevich, and J Rebek, Jr. *J. Am. Chem. Soc.* **2000**, 122, 8880-8889, (b) S. D. Starnes, D. M. Rudkevich, and J. Rebek, Jr. *J. Am. Chem. Soc.* **2001**, 123, 4659-4669, (c) U. Lulcking, J. Chen, D. M. Rudkevich, J. Rebek, Jr. *J. Am. Chem. Soc.*, , **2001**, 123, 9930, (d) D. M. Rudkevich, G. Hilmersson, and J. Rebek, Jr. *J. Am. Chem. Soc.* **1998**, 120, 12216-12225
27. T. Szabo, G. Hilmersson, and J. Rebek, Jr. *J. Am. Chem. Soc.* **1998**, 120, 6193-6194
28. (a) Y. Tokunaga and J. Rebek, Jr. *J. Am. Chem. Soc.* **1998**, 120, 66-69, (b) J. M. Rivera, T. MartiLn, and J. Rebek, Jr. *J. Am. Chem. Soc.* **2001**, 123, 5213-5220
29. C. K. Chang, T. G. Trayler, *Proc. Natl. Acad. Sci. USA*, **1973**, 70, 2647
30. J. E. Baldwin, J. Huff, *J. Am. Chem. Soc.*, **1973**, 95, 5757
31. (a) J. P. Collman, R. R. Gagne, C. A. Reed, T. R. Halbert, G. Lang and W. T. Robinson, *J. Am. Chem. Soc.*, **1975**, 97, 1427, (b) J. P. Collman, *Acc. Chem. Res.*, **1977**, 10, 265
32. for example, (a) M. Momenteau, C. A. Reed, *Chem. Rev.*, **1994**, 94, 659 (b) T. Komatsu, E. Hasegawa, S. Kumamoto, H. Nishide, and E. Tsuchida *J. Chem. Soc., Dalton Trans.*, **1991**, 3281, (c) F. Tani, M. Matsu-ura, S. Nakayama, M. Ichimura, N. Nakamura, and Y. Naruta, *J. Am. Chem. Soc.*, **2001**, 123, 1133, (d) M. Momenteau and C. A. Reed, *Chem. Rev.*, **1994**, 94, 659, (e) R. D. Jones, D. A. Summerville and F. Basolo, *Chem. Rev.*, **1979**, 79, 139, (f) E. Tsuchida and H. Nishide, *Top. Curr. Chem.*, **1986**, 132, 64, (g) T. Komatsu, E. Hasegawa, H. Nishide and E. Tsuchida, *J. Chem. Soc., Chem. Comm.*, **1990**, 66, (h) J. P. Collman, P. C. Herrmann, L. Fu, T. A. Eberspacher, M. Eubanks, B. Boitrel, P. Hayoz, X. Zhang, J. I. Brauman and V. W. Day, *J. Am. Chem. Soc.*, **1997**, 119, 3481, (i) T. Komatsu, K. Sano and E. Tsuchida, *Chem. Commun.*, **1998**, 977, (j) F. Tani, M. Matsu-ura, S.

- Nakayama, M. Ichimura, N. Nakamura, and Y. Naruta, *J. Am. Chem. Soc.* 2001, 123, 1133
33. (a) M. R. Ghadiri, J. R. Granja, L. K. Bueher, *Nature*, **1994**, 369, 301, (b) J.R.Granja, M.R. Ghadiri, *J. Am. Chem. Soc.* **1994**, 116, 10785, (c) M.D.Yoder, N. T. Keen, F. Journak, *Science*, **1993**, 260, 1503, (d) M.R. Ghadiri, K. Kobayashi, J. R. Chadha, D. E. McRee, *Angew. Chem. Int. Ed. Engl*, **1995**, 270, 997
34. C. J. Pederson, *J. Am. Chem. Soc.*, **1967**, 89, 7017