  Bull. Chem. Soc. Jpn., 61, 3129—3139 (1988)

Biomolecules Conform to the Rule of Topological Charge Stabilization

Jun-ichi AIHARA
Department of Chemistry, Faculty of Science, Shizuoka University, Oya, Shizuoka 422
(Received April 12, 1988)

Almost all biomolecules conform to the rule of topological charge stabilization. Heteroatoms are placed at the right sites in proteins, nucleoside bases, porphyrins, and related metabolites. It is noteworthy that typical redox-mediating coenzymes obey this rule both in the oxidized and reduced states. Thus, it is very likely that the rule of topological charge stabilization has been one of the most important guiding principles not only in chemical evolution but also in general biosynthesis.

Molecules containing one or more heteroatoms occur widely in nature. Nitrogen, oxygen, and sulfur are typical heteroatoms commonly found in biomolecules. These atoms are more electronegative than carbon and hydrogen.[1] Nucleic acids, proteins, coenzymes, and porphyrins all contain one or more π-electron subsystems containing heteroatoms. Considering that these molecules are biologically indispensable, it is a very important matter in chemistry and biochemistry to seek for the reason why they were chosen to construct living organisms. Why do the biomolecules contain many heteroatoms? What factors stabilize these molecules? Of course, it is never easy to answer these questions. Even partial answers will mark an important progress in the study of biosynthesis and molecular evolution.

In 1983 Gimarc pointed out that the pattern of charge densities in a molecule is determined mainly by the connectivity or the topology of the molecule.[2] Nature tends to place heteroatoms of great electronegativity in those positions where the isostructural, iso-π-electronic hydrocarbon has large charge densities.[2-4] Gimarc called the effect the rule of topological charge stabilization or the TCS rule. He showed that generally a molecule conforming to this rule is stable in our chemical sense.[2-4] In previous papers[5,6] the TCS rule was successfully used to explain the stability of so-called quasi-aromatic molecules and various organic molecules formed in primitive-Earth-simulation experiments.

This paper aims at rationalizing the stability of nucleoside bases and other important biomolecules in terms of Gimarc's TCS rule. We will show below that very many biologically important molecules are stable owing to the presence of heteroatoms at the right sites. As will be seen, most heteroatoms in every biomolecule are situated in such a manner that they attain the maximum thermodynamic stability of the whole conjugated system. Ketone oxygen and imine nitrogen contriburte one π electron to the conjugated system, whereas amine nitrogen and alcohol (or ether) oxygen contribute two π electrons instead of one.

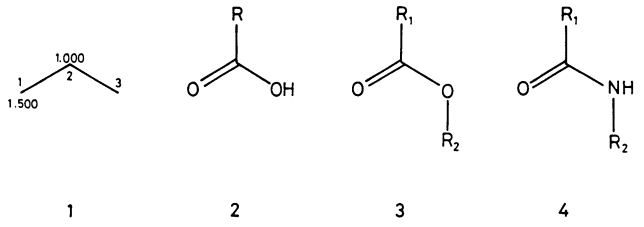
Results and Discussion

Charge densities calculated for a molecule containing heteroatoms depend on the choice of semiempirical parameters, but in the iso-π-electronic hydrocarbon all Coulomb integrals are uniform and charge densities are determined only by the topology of the molecule and the number of π electrons.[7] To emphasize this fact, Gimarc[2-4] referred to the isostructural, iso-π-electronic hydrocarbon as the uniform reference frame (URF). It is a (hypothetical) conjugated hydrocarbon obtained by replacing all the heteroatoms in a given molecule by sp^2 carbons and then by donating the same number of π electrons to the resulting conjugated system. The URF is, of course, not identical with a real molecule. It is charge densities in such URF's that the TCS rule adopts to estimate the stability of biomolecules containing heteroatoms. First-order perturbation theory has given a justification for the TCS rule.[3]

All charge densities given below will be those calculated for the π-electron systems of the URF's. We used the Hückel molecular orbital (HMO) model in its simplest form unless otherwise stated. The simple HMO model is known to yield exaggerated charges, but experience shows that they are adequate for the present purpose which requires only a qualitative pattern of charge distributions.[2] We assumed that the Coulomb integral has the same value for all carbons and that the resonance integral has the same value for all π bonds. All alkyl substituents will be disregarded for simplicity because they are scarcely conjugated with the parent π-electron framework.

Carboxylates and Peptides. Consider the allyl anion $C_3H_5^-$ (**1**), in which a central carbon is connected to two edge carbons. The conjugated system is composed of four π electrons moving through the three constituent carbons. The calculated charge density is greater at the edge carbons than at the central carbon. This anion represents the URF of the carboxyl group (**2**) not conjugated with the remainder of the molecule (R in **2**). Many carboxylic acids (**2**), carboxylic esters (**3**), and naturally occurring amino acids have this type of carboxyl groups. Here and hereafter R and R_i (i=1, 2, 3, ···) represent the substituents not conjugated with the π-electron system in question.

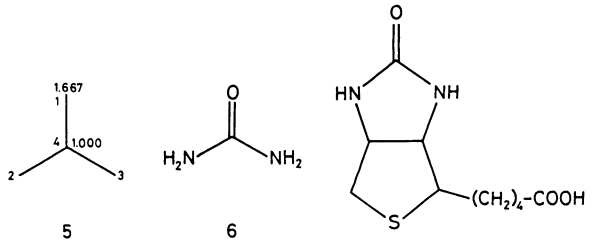
Topological charge stabilization says that electronegative heteroatoms prefer to occupy the sites where



1 2 3 4

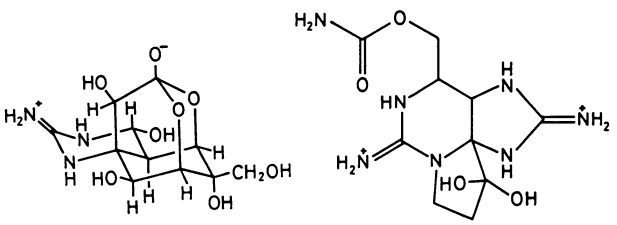
electron density is already very great in the iso-π-electronic URF.[2-4] Oxygen is much more electronegative than carbon.[1] Stability of carboxylic acids and carboxylates is attributable to the two constituent oxygens because these atoms occupy the two sites of greatest charge density in the URF. Individual peptide bonds in protein (**4**) are not conjugated with each other. They are also iso-π-electronic with the allyl anion (**1**). Stability of all proteins can thus be explained in terms of the TCS rule. This is possibly the primary reason why insulating methylenes intervene between adjacent peptide bonds.

Urea and Toxins. Gimarc considered the trimethylenemethane dianion $C(CH_2)_3^{2-}$ (**5**) as a URF of such common inorganic species as CO_3^{2-}, NO_3^- and SO_3.[2] The conjugated system of **5** is composed of six electrons moving through four carbons. The calculated charge density is greater at the peripheral carbons than at the central carbon. Naturally occurring CO_3^{2-}, NO_3^-, and SO_3 are very stable because all oxygens are located at the peripheral positions to achieve the maximum energy stabilization. Note that oxygen is more electronegative than nitrogen and sulfur.[1]

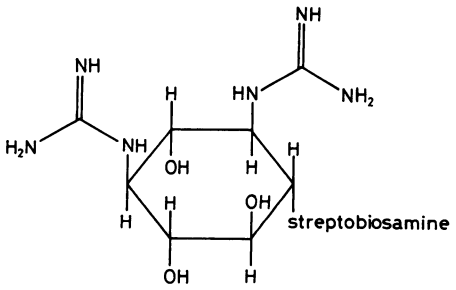


5 6 7

The trimethylenemethane dianion (**5**) also constitutes the URF's of urea (**6**) and biotin (**7**, growth factor). Analogous π-electron systems are found in many other biomolecules. For example, tetrodotoxin (**8**, swellfish toxin), saxitoxin (**9**, shellfish toxin), and streptomycin (**10**, antibiotic) have one, three, and two conjugated subsystems, respectively, which are all iso-π-electronic with urea. In these biomolecules nature puts oxygens and/or nitrogens at the peripheral positions of the trimethylenemethane framework where topology tends to distribute more charge density. These examples nicely illustrate the theory that the conformity to the TCS rule gives a stable molecular system.[2-4] Urea, biotin, tetrodotoxin, streptomycin, and saxitoxin are all Y-aromatic species.[8]

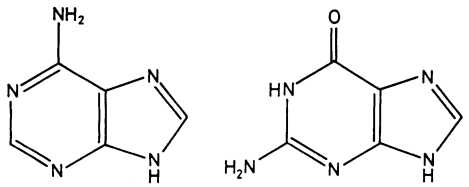


8 9

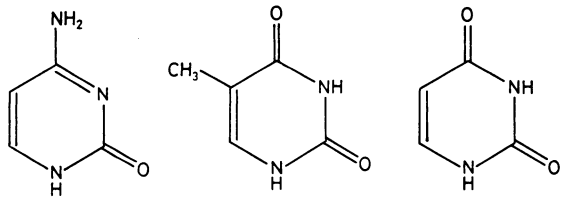


10

Nucleoside Bases. The purine and pyrimidine bases are key components of nucleic acids. Adenine (**11**), guanine (**12**), cytosine (**13**) thymine (**14**), and uracil (**15**) are principal nucleoside bases. The structural formulas **11**—**15** represent the bases in the tautomeric form which predominates according to spectroscopic analyses: the keto and amino form.[9,10] This form is favored in nucleic acids over the enol and imino form. In the present study, there is no need to discriminate between these two forms since both give the same URF. Even if a given base undergoes tautomerization, positions of the heteroatoms remain unchanged.



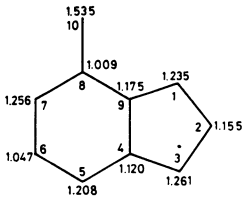
11 12



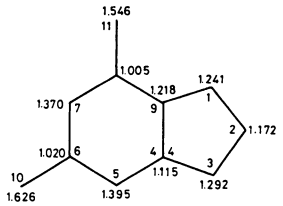
13 14 15

The URF's of two purine bases **11** and **12** are shown in **16** and **17**, respectively. Heteroatoms in these bases, without exceptions, are situated at the sites of large

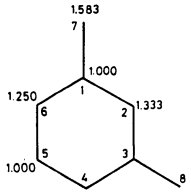
charge density in the corresponding URF's. It is noteworthy that all exomethylene carbons in these URF's (i.e., carbon 10 in **16** and carbons 10 and 11 in **17**) have high charge density. All these sites are occupied by amine nitrogens or ketone oxygens in the purine bases. All positions of high charge density in the URF's are occupied by heteroatoms in the corresponding purine bases. Three pyrimidine bases **13**, **14**, and **15** happen to have the same URF (**18**) with ten π electrons. It is clear that heteroatoms in these bases occupy the sites of high charge density in their common URF. Therefore it seems very likely that nature selected these nucleoside bases as parts of gene because they are not only stable but also easy to form.



16



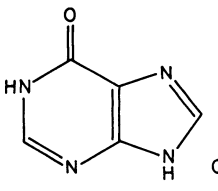
17



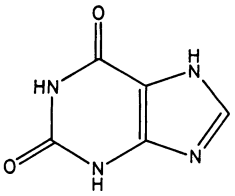
18

The nucleoside bases in the keto and amino form fit the base-pairing scheme of Watson and Crick for double-stranded DNA, whereas those in the enol and imino form do not. However, the charge density distribution in the URF of each base is not consistent with the keto and amino form. All exomethylene carbons in the URF are high in charge density as compared with ring carbons. The enol and imino form is favorable from this point of view since alcohol (or ether) oxygen attracts much more charge than ketone oxygen. This discrepancy might represent the limitation of the TCS rule.

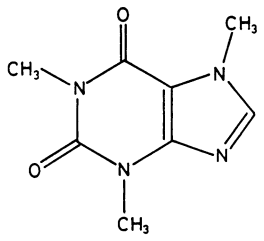
There are many biomolecules which have either a purine or pyrimidine nucleus. Among purine bases and alkaloids are hypoxanthine (**19**), xanthine (**20**), caffeine (**21**), theophylline (**22**), theobromine (**23**), and uric acid (**24**). The URF of hypoxanthine is the same as that of adenine (**16**). The URF's of xanthine, caffeine, theophylline, and theobromine are all identical with that of guanine (**17**). The URF of uric acid is shown in **25**. Among biological pyrimidines are 5-methylcytosine (**26**), 5-(hydroxymethyl)cytosine (**27**), and orotic acid (**28**). The URF's of **26** and **27** are identical with that of cytosine. The URF of orotic acid is shown in **29**. All these molecules conform to the TCS rule.



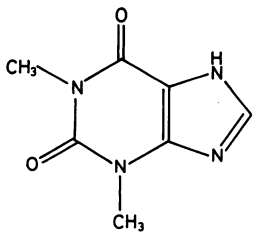
19



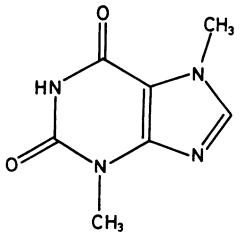
20



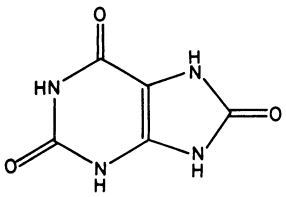
21



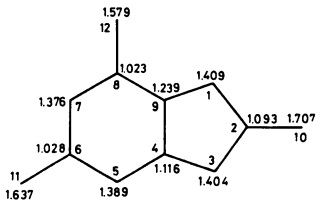
22



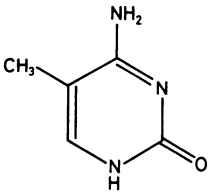
23



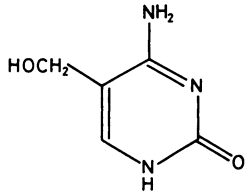
24



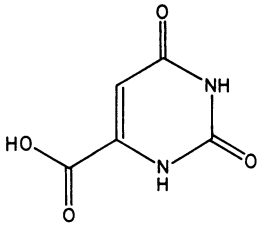
25



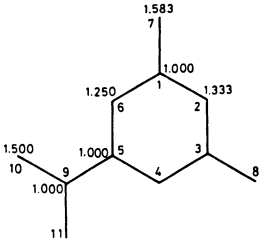
26



27

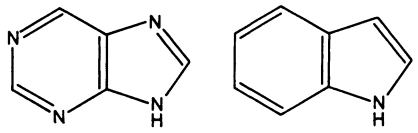


28

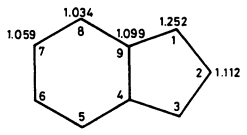


29

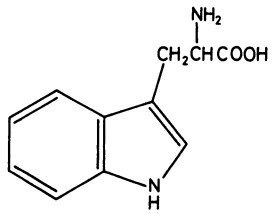
Unsubstituted purine and pyrimidine are not biomolecules. The calculated charge densities of the URF's of these molecules are not useful for rationalizing the positions of heteroatoms in nucleoside bases. The sites of high charge density in the URF's change by substitution. Furthermore, valence tautomerization sometimes occurs by substitution.[9,10] Unsubstituted purine (**30**) is iso-π-electronic with indole (**31**). Therefore, both have the same URF (**32**). A single indole nitrogen is located at one of the sites of greatest charge density in the URF. Thus, indole is stabilized by topological charge distribution, so are tryptophan (**33**) and many indole alkaloids. Indole is found in coal tar.



30 31



32

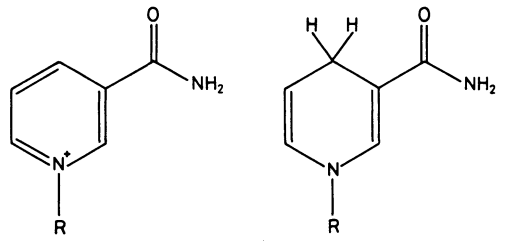


33

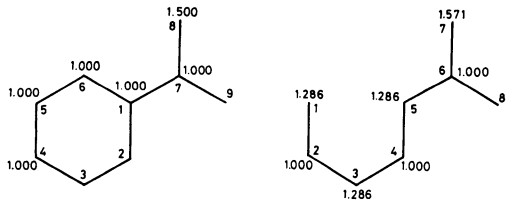
According to the original TCS rule,[2] one does not need to place heteroatoms at all sites of high charge density in the URF to stablize molecules. Indole indeed has one amine nitrogen at one of the sites of largest charge density in the URF. One nitrogen is enough to stabilize the indole conjugated system although there is one more site of largest charge density in the URF. Position 6 in the URF of pyrimidine bases (**18**) has high charge density, but these bases has no heteroatom there. In contrast, heteroatoms in **2**—**4**, **6**—**12**, and **19**—**24** are located at all sites of high charge density in the corresponding URF's. We might say that these biomolecules are extremely faithful to the TCS rule.

Coenzymes. Biological oxidation and reduction are catalyzed by enzymes, each of which functions in conjunction with a coenzyme which acts as an electron carrier. There are a number of oxidative enzymes but relatively few coenzymes.[11] These coenzymes are involved in many biological oxidation-reduction reactions. Biotin (**7**) acts as a coenzyme responsible for the transfer of carbon dioxide in carboxylation reactions.

There are two coenzymes which contain nicotinamide as an active center: namely, nicotinamide-adenine dinucleotide (NAD^+) and nicotinamide-adenine dinucleotide phosphate ($NADP^+$). NAD^+ ($NADP^+$) is reduced during the enzymatic reaction to give the dihydropyridine derivative NADH (NADPH). Their oxidized and reduced forms have the general formulas shown in **34** and **35**, respectively. The presence of a carbamoyl group ($-CONH_2$) in nicotinamide (**34**) is consistent with the charge density distribution in the URF (**36**, ten π electrons), whereas the ring nitrogen does not correspond to the site of high charge density in the URF. However, it seems unlikely that **34** is greatly destabilized by the ring nitrogen of the imine type.



34 35

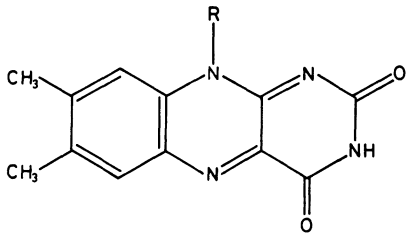


36 37

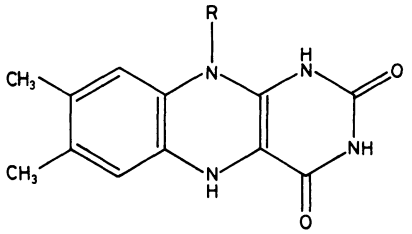
Let us consider the relative electron-attracting abilities of heteroatoms concerned. Amine nitrogen is a combination of N^{2+} and two π electrons, whereas imine nitrogen is a combination of N^+ and one π electron. Likewise, alcohol oxygen is a combination of O^{2+} and two π electrons, whereas ketone oxygen is a combination of O^+ and one π electron. The electron-attracting ability is then in the order: alcohol oxygen>amine nitrogen$\gg$ketone oxygen>imine nitrogen.[1] Considering that imine nitrogen is least electronegative, it seems unlikely that the imine nitrogen not situated at the right site destabilizes the π-electron system to a great extent. It is presumably for this reason that pyridine is as stable as benzene. The situation is essentially the same for nicotinamide (**34**). However, if alcohol oxygens and/or amine nitrogens are introduced into a given π-electron system, it is highly desirable for these atoms to occupy the sites of high charge density in the URF.

The URF of the reduced form of nicotinamide (**37**, ten π electrons) rationalizes the existence of three heteroatoms without any reservation. The ring nitrogen is now converted to amine nitrogen. It is then very desirable for this atom to occupy the site of high charge density in the URF. As shown in **37**, the amine nitrogen really occupys one of the sites of high charge density in the URF. Thus, three heteroatoms are required to stabilize both NAD^+ ($NADP^+$) and its reduced species.

There are two coenzymes which contain dimethylisoalloxazine as an active center: flavin mononucleotide (FMN) and flavin-adenine dinucleotide (FAD). These flavin coenzymes can be formulated as shown in **38**. They are formed in vivo from riboflavin (vitamin B_2), which can also be formulated as **38**. The isoalloxazine nucleus of a flavin coenzyme (oxidized form) is reduced during the enzymatic reaction to give the dihydro derivative or the reduced form (**39**). The yellow color of the flavin coenzyme disappears on reduction since the reduced form is colorless.

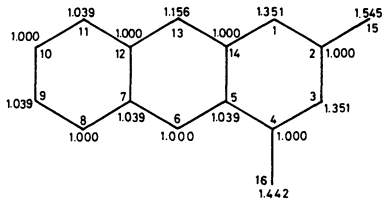


38

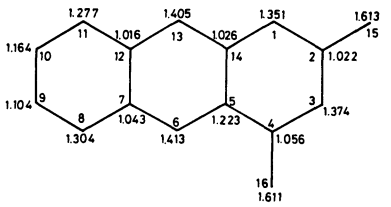


39

The URF of isoalloxazine (**40**, 18 π electrons) has large charge density at positions 1, 3, 13, 15, and 16. The corresponding positions in isoalloxazine are all occupied by electronegative heteroatoms. Position 6 in the URF is also occupied by a heteroatom (imine nitrogen) although it does not have very high charge density. This can again be rationalized by examining the charge density distribution in the URF of the reduced species (**41**, 20 π electrons). On reduction two imine nitrogens in isoalloxazine are converted into amine nitrogens. These amine nitrogens are assuredly placed at the sites of high charge density in the URF of the reduced form. Thus, imine nitrogen at position 6 in isoalloxazine is needed to stabilize its dihydro derivative. As mentioned above, imine nitrogens presumably influence the isoalloxazine conjugated system to a lesser extent. All sites of high charge density in **41** are occupied by heteroatoms in reduced isoalloxazine (**39**).

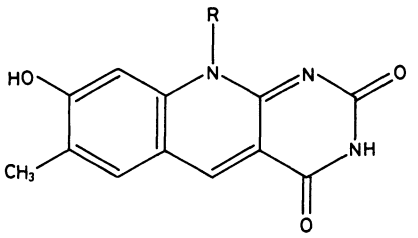


40

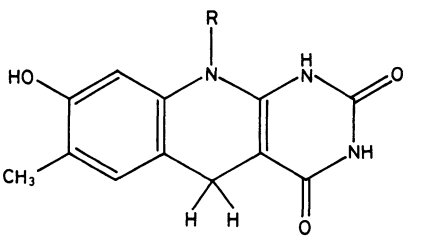


41

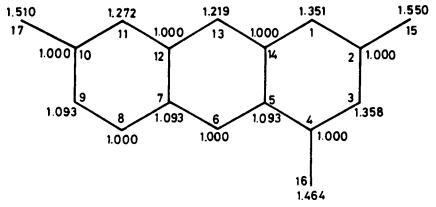
It is interesting to note that there are deazaflavin coenzymes,[12] which are structurally analogous to flavin coenzymes. It is a predominant cellular constituent in anaerobic methanogenic bacteria which biosynthesize methane. These bacteria are thought to be descendants of primordial organisms. Deazaflavin coenzymes (**42**) are likewise reduced during the enzymatic reaction to give the dihydro derivatives (**43**). The URF's of the oxidized and reduced forms are shown in **44** and **45**, respectively. Both forms have 20 π electrons each. All heteroatoms in **42** and **43** occupy the sites of high charge density in **44** and **45**, respectively. One carbon corresponding to position 6 in **44** is excluded from the conjugated system on



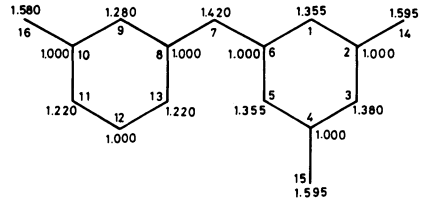
42



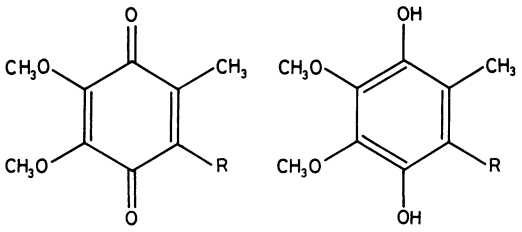
43



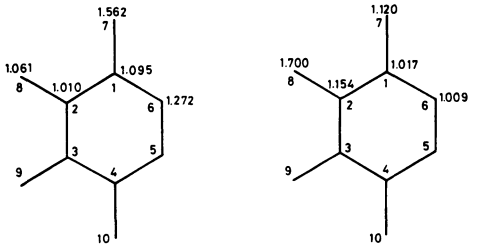
44



45



46 47



48 49

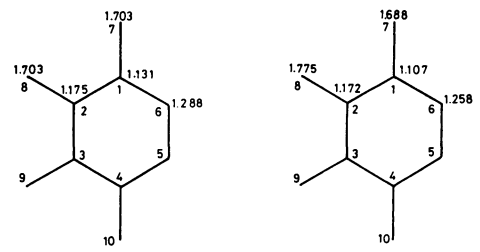
reduction. Therefore, there is no need to replace carbon 6 by nitrogen to stabilize the reduced form. Thus, the deazaflavin conjugated system is a functional hybrid between monocyclic nicotinamide and tricyclic flavin coenzymes.[12]

Mitochondria contain a quinone called ubiquinone which has the general structure **46**. The biological function of ubiquinone in redox reactions depends upon the reversible reduction to the corresponding hydroquinone (**47**). The URF of ubiquinone (**48**, 12 π electrons) has high charge density at positions 7 and 10. Two ketone oxygens are placed at the corresponding positions in the coenzyme. However, the positions of two methoxyl oxygens do not conform to the TCS rule. Charge density is not high at the corresponding positions (i.e, carbons 8 and 9) in the URF. Ubiquinone apparently constitutes an exceptional biomolecule, in the sense that highly electronegative methoxyl oxygens occupy the sites of low charge density in the URF.

This difficulty may be resolved in the following manner. Remember that carbons 8 and 9 in the URF of ubiquinone are to be replaced by methoxyl oxygens, but not by ketone oxygens. In general, a resonance integral between ether oxygen and carbon is relatively small.[7] Therefore, it might be meaningful to decrease the resonance integrals between carbons 2 and 8 and between carbons 3 and 9 in the URF of ubiquinone tentatively to 80% of the standard value. Such modification of the URF can be justified by the fact that ubiquinone has no possibility of valence tautomerization. The charge densities in the modified URF is shown in **49**. Then, we have largest charge density at carbons 8 and 9, which can now be replaced reasonably by methoxyl oxygens to give the ubiquinone conjugated system. Such a drastic change in charge densities was caused by the exchange of the sixth and seventh molecular orbitals. These two orbitals are accidentally very close in energy to each other in the unmodified URF.

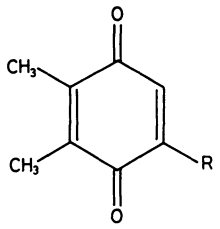
Unfortunately, another diffuculty arises with the modified URF of ubiquinone (**49**). Charge densities at carbons 7 and 10 are greatly reduced in this URF, and it seems that the presence of two ketone oxygens is not so effective for stabilizing the conjugated system. In this context, it is well-known that exomethylene carbons in an organic molecule generally are very reactive with high reactivity indices. For example, 3,6-dimethylene-1,4-cyclohexadiene cannot be isolated, whereas *p*-benzoquinone is fairly stable and isolable. Ketone oxygens are obviously helpful for stabilizing exocyclic doube bonds. The presence of two ketone oxygens in ubiquinone can be rationalized in this manner.

Methoxyl oxygens in ubiquinone play an important role when it is converted to the reduced form. The URF of the reduced form (**50**, 14 π electrons) has high charge density at positions 7, 8, 9, and 10. Thus, the four alcohol (or ether) oxygens are placed at these positions to stabilize the reduced form. The modified URF of the reduced form (**51**) was defined using the same resonance integrals as those in the modified URF of ubiquinone (**49**). It has a charge density distribution similar to **50**.

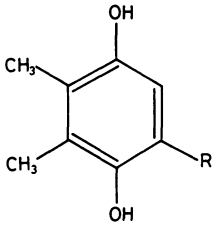


50 51

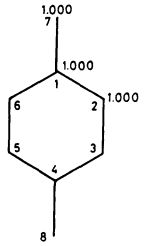
In photosynthesis plastoquinone (**52**) participates in reversible oxidation-reduction reactions. When plastoquinone is reduced during the reaction, the corresponding hydroquinone (**53**) is formed. Charge density is uniform over the URF of plastoquinone (**54**, eight π electrons). For the same reason as that given for ubiquinone, two exomethylene carbons in the URF are replaced by ketone oxygens in plastoquinone. The URF of hydroquinone (**55**, ten π electrons) has high charge density at positions 7 and 8. Two alcohol oxygens are placed at the corresponding positions in **53**. It follows that two oxygens are needed to stabilize both the oxidized and reduced forms of plastoquinone.



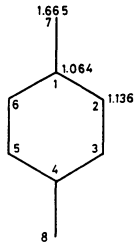
52



53



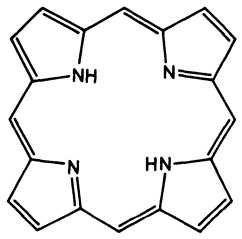
54



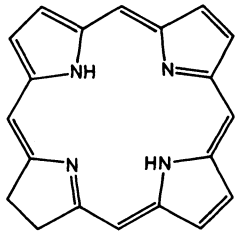
55

As has been seen above, it is very interesting that heteroatoms in representative redox-mediating coenzymes are arranged in such a manner that both the oxidized and reduced forms fulfill the TCS rule. Nature must have designed these molecular systems, taking into account the fact that amine nitrogen and alcohol (or ether) oxygen attract much more negative charge than imine nitrogen and ketone oxygen. The latter two atoms are sometimes introduced at sites of low charge density in the URF.

Porphyrins and Chlorophylls. All forms of life rely on the unique redox and electron-transfer abilities of hemoproteins and/or chlorophylls.[13,14] Active sites of hemoproteins and chlorophylls are the porphine (**56**) and chlorin (**57**) macrocycles, respectively, each acting as a ligand for a particular metal ion. Exactly speaking, chlorin represents a molecule formed by adding two hydrogens to the peripheral positions in the D-ring of any porphyrin. However, the term chlorin is here restricted to the unsubstituted chlorin nucleus (**57**). Although various substituents are usually attached to these molecules, we first disregard them all and evaluate the charge densities in metal-free porphine (**56**) and chlorin (**57**).

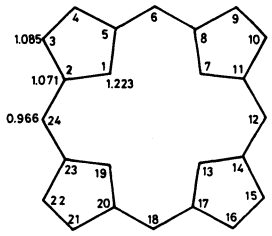


56

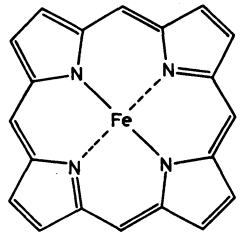


57

Pyrrole nitrogens in porphine (**56**) are located at those positions which have greatest charge density in the URF (**58**, 26 π electrons). Exactly the same observation was made over 35 years ago by Longuett-Higgins et al.[15] Historically, they were the first to introduce the idea of topological charge stabilization.[2] They noted that the great stability of porphine is largely due to the presence of nitrogens at positions in the skeleton where, for purely geometrical reasons, the π-electrons tend to congregate.[15] The TCS rule holds for **56**. Hemoproteins have an iron(II) ion at the center of every porphine nucleus. The metal ion is scarcely conjugated with the π-electron system of the porphine.[16] Therefore, there is no need to devise a different URF for the metal-porphine complex (**59**). In biosythesis the iron(II) ion is inserted after the formation of the porphine nucleus, and tightly bound to it.[13]



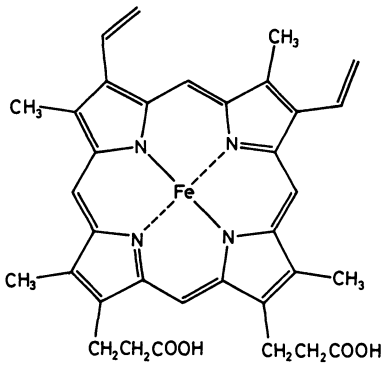
58



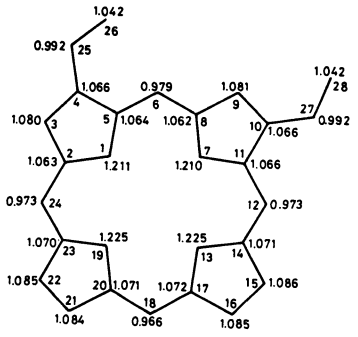
59

Substituted porphines have been referred to as porphyrins. All naturally occurring porphyrins have various substituents, most of which are not conjugated with the porphine nucleus. Heme (**60**) is an iron(II) complex of a porphyrin molecule called protoporphyrin. Human myoglobin and hemoglobin contain one and four heme molecules, respectively. The heme molecule not bound to protein has eight substituents, only two of which are conjugated with the porphine nucleus. As exemplified by **60**, vinyl and/or carbonyl groups in porphyrin are conjugated with the porphine nucleus. However, these substituents never displace the sites of large charge density in the URF. Compare the charge densities in the URF of porphine (**58**) with those in the URF of heme (**61**).

In genertal, a vinyl substituent is fairly reactive. Therefore, an alkyl group or the like is ofter attached to its β position. Otherwise, it is often replaced by the carbonyl substituent. When such a heme molecule is incorporated in protein, unsaturated substituents are used to bond the protein and become saturated ones.

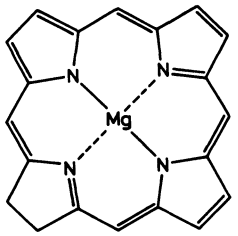


60

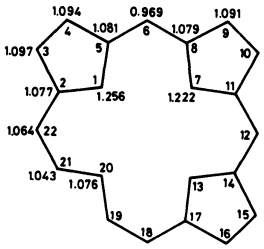


61

Chlorophylls have a magnesium-chlorin nucleus (**62**) with no carbon-carbon double bond in the D-ring. The calculated electron densities in the URF of chlorin (**63**, 24 π electrons) are highest at carbons 1, 7, and 13. There are nitrogens in the corresponding positions of chlorin. However, this molecule has one more nitrogen at position 20, the presence of which is not expectable from the URF (**63**). In the chlorophyll biosynthesis, the magnesium ion is inserted into the porphine nucleus in protoporphyrin.[13] After that, the reduction of the D-ring double bond occurs. Nitrogen at position 20 is thought to have been introduced to stabilize the metal-free porphine nucleus. After the reduction of the D-ring double bond all the four nitrogens are used to bind the magnesium ion tightly. If the magnesium ion is removed from chlorophyll, nitrogen at position 20 will be of the imine type. The magnesium ion is not conjugated with the π-electron system of chlorin.[17] Therefore, there is no need to devise a URF for the magnesium-chlorin complex (**62**).

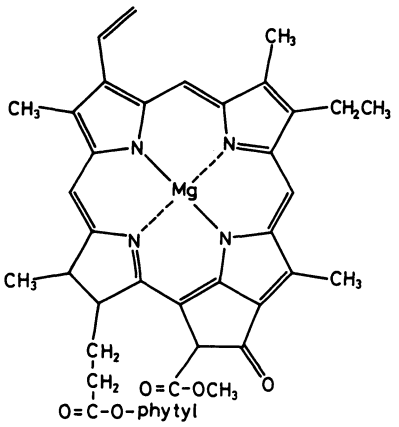


62

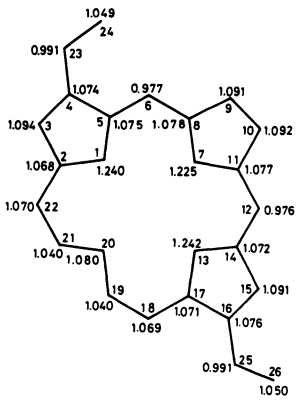


63

Naturally occurring chlorophylls have vinyl and/or formyl substituents in addition to alkyl substituents. Although photosynthetic organisms contain a variety of photosynthetic pigments, only chlorophyll *a* (**64**) performs a primary role in the photosynthetic act. This molecule has one vinyl and one carbonyl substituent attached to the chlorin nucleus. However, charge density is not large at all vinyl carbons of the URF (**65**). As in the case of porphyrins, vinyl substituents never change the sites of great charge density in the chlorin nucleus.



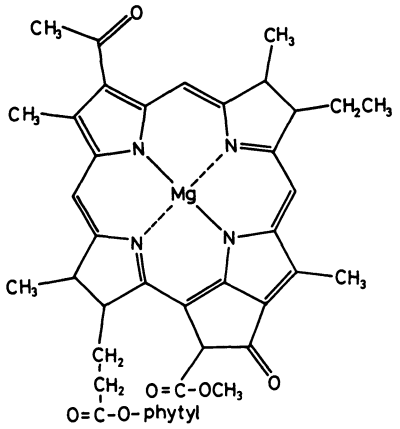
64



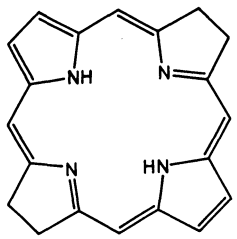
65

Photosynthetic bacteria produce a variety of photosynthetic pigments known collectively as bacte-

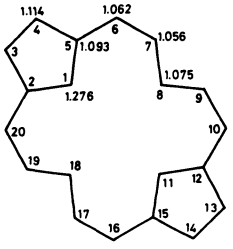
riochlorophylls.[13] Most bacteriochlorophylls have the same chlorin nucleus as those of chlorophylls. However, bacteriochlorophyll *a* (**66**) has a slightly different macrocycle (**67**), which has two additional hydrogens at the B-ring.[13] Two pyrrole nitrogens in **67** are located at sites of high charge density in the URF (**68**, 22 π electrons), but the other two are not. This again suggests that hydrogen addition at the B-ring must have occurred after the insertion of the magnesium ion. The URF of bacteriochlorophyll *a* itself is shown in **69**. The sites of high charge density in the URF are not displaced by the introduction of two unsaturated substituents. If the metal ion is removed from bacteriochlorophyll *a*, the two pyrrole nitrogens will be of the amine type, and the remaining two will be of the imine type.



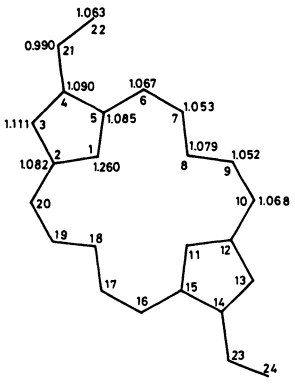
66



67

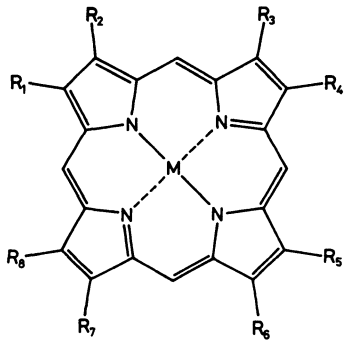


68

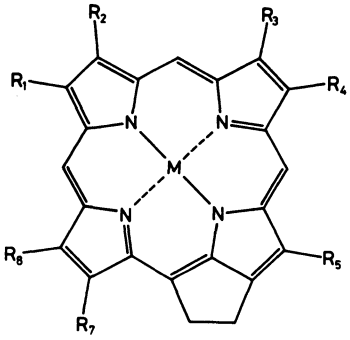


69

Porphyrins in geological samples[18] occur as complex mixtures of two major series: the etio and deoxophylloerythro types (**70** and **71**, respectively). In **70** and **71**, M represents an appropriate metal ion. These molecules are detectable in petroleum, shale, and coal. They have saturated substituents only, and the URF of porphine (**58**) applies. They were presumably derived from defunctionalization and full aromatization of naturally occurring chlorophylls.[18] This implies that unsaturated substituents attached to the original chlorophylls are not so stable as the porphine nucleus.



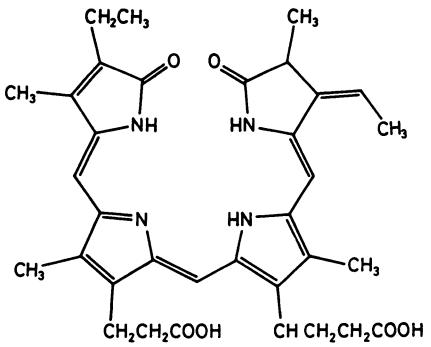
70



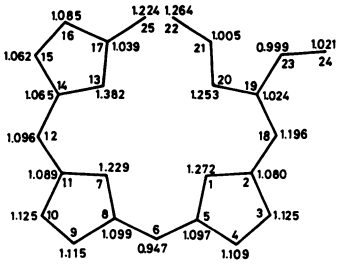
71

Phycobilins and bile pigments are formed in vivo by degradation of porphyrins.[14] Red and blue-green algae contain, in addition to chlorophyll *a*, blue or red pigments known as phycobilins, open-chain tetrapyrroles related to the chlorophylls but lacking the metal ion. These pigments absorb the sunlight effectively, and transfer its energy to chlorophyll *a*. Phycocyanobilin (**72**) is a phycobilin having the largest π-electron system. This blue compound has six heteroatoms except the isolated carboxyl groups. All heteroatoms occupy the sites of high charge density in the URF (**73**). Biliverdin (**74**) is one of the bile pigments, which is again an open-chain tetrapyrrole. A number of open tetrapyrroles are formed from biliverdin by reduction and oxidation reactions. The URF of biliverdin is shown in **75**. The locations of all heteroatoms in this molecule can be predicted from

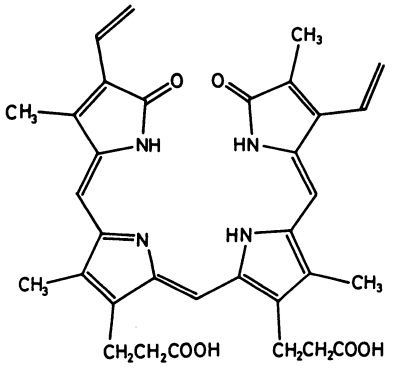
the charge density distribution in the URF. Thus, it is noteworthy that not only porphyrins but also their metabolites conform to the TCS rule very faithfully although their molecular structures are far from being simple.



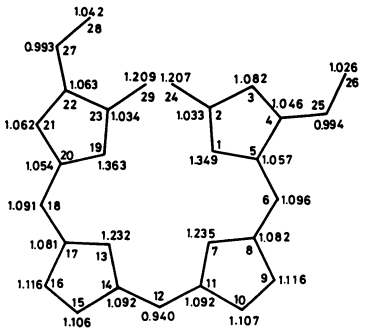
72



73



74



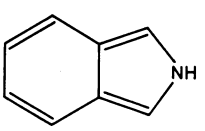
75

Concluding Remarks

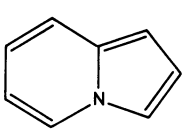
Gimarc pointed out that the TCS rule has potential value as a guide for synthetic efforts to be made by organic chemists.[2] We have shown that many biologically important molecules conform to the TCS rule. Most heteroatoms in any biomolecule are placed at the sites of large charge density in its URF. In all examples shown above, highly electronegative alcohol (or ether) oxygens and amine nitrogens always occupy the sites of high charge density in the corresponding URF's. Nature obviously adopts the TCS rule as a general guide for biosynthesis. It should be noted that various coenzymes have been designed so that both the oxidized and reduced forms satisfy the TCS rule.

Chemical and biochemical evolution[19] must have occurred in harmony with the TCS rule.[6] As has been seen above, nucleoside bases, amino acids, and many other biomolecules are fully consistent with the TCS rule. This coincides with the fact that many of these molecules can be prepared by rather simple reactions intended to simulate chemical evolution.[17] Therefore, it is very probable that nature picked up molecules available readily in prebiotic chemistry to construct living organisms.[6] Some of the molecules must have exerted important functions incidentally in cellular systems. Conversely, nature has been so skillful as to choose biologically important molecules from the products of chemical evolution. For example, nucleoside bases are best suited for pairing in DNA and RNA. Such bases were chosen from a group of rather simple molecules which satisfy the TCS rule.

From a different point of view, we can say that living organisms adopted energetically very accessible ways of biosynthesis. In general, a stable molecule is associated with low activation energy of formation.[20] The resulting biomolecules are thermodynamically very stable as compared to their isomeric species. Many derivatives of indole (**31**) are prepared in vivo, but no structural isomers of indole, such as isoindole (**76**) and indolizine (**77**), are scarcely found in cellular systems. These isomers do not conform to the TCS rule, and are much less stable than indole. It is obvious that living organisms do not make much effort to prepare high-energy conjugated systems. Molecules with important biological functions must have been chosen from synthetically very accessible molecules. Thus, topological charge stabilization serves as a unifying principle for the systemization of biosynthesis.



76



77

Helpful discussions with Prof. Yasuaki Okumura and Prof. Koichi Yoshinaga are gratefully acknowledged. This work was supported by the grant from the Ministry of Education, Science and Culture, administered by President of Shizuoka University.

References

1) L. Pauling, "The Nature of the Chemical Bond," Cornell University Press, Ithaca, NY (1960), Chap. 3.
2) B. M. Gimarc, *J. Am. Chem. Soc.*, **105**, 1979 (1983).
3) B. M. Gimarc and J. J. Ott, *J. Am. Chem. Soc.*, **108**, 4298 (1986).
4) J. J. Ott and B. M. Gimarc, *J. Am. Chem. Soc.*, **108**, 4303 (1986).
5) J. Aihara, *Bull. Chem. Soc. Jpn.*, **60**, 2268 (1987).
6) J. Aihara, *Bull. Chem. Soc. Jpn.*, **61**, 2309 (1988).
7) A. Streitwieser, Jr., "Molecular Orbital Theory for Organic Chemists," Wiley, New York (1961).
8) P. Gund, *J. Chem. Educ.*, **49**, 100 (1972).
9) J. Elguero, C. Marzin, A. R. Katritzky, and P. Linda, "The Tautomerism of Heterocycles (*Adv. Heterocycl. Chem., Suppl. 1*)," Academic Press, New York (1976), Chap. 5.
10) J. Elguero, C. Marzin, A. R. Katritzky, and P. Linda, "The Tautomerism of Heterocycles (*Adv. Heterocycl. Chem., Suppl. 1*)," Academic Press, New York (1976), Chap. 2.
11) H. C. S. Wood, "Comprehensive Organic Chemistry: The Synthesis and Reactions of Organic Compounds," ed by D. Barton and W. D. Ollis, Vol. 5, Pregamon, Oxford (1979), Part 24.3.
12) C. Walsh, *Acc. Chem. Res.*, **19**, 216 (1986).
13) H. Akhtar and P. M. Jordan, "Comprehensive Organic Chemistry: The Synthesis and Reactions of Organic Compounds," ed by D. Barton and W. D. Ollis, Vol. 5, Pregamon, Oxford (1979), Part 30.2.
14) K. M. Smith, "Rodd's Chemistry of Carbon Compounds: A Modern Comprehensive Treatise," ed by S. Coffey, Vol. IV, Part B, Elsevier, Amsterdam (1977), Chap. 12.
15) H. C. Longuet-Higgins, C. W. Rector, and J. R. Platt, *J. Chem. Phys.*, **18**, 1174 (1950).
16) H. Kashiwagi, personal communication.
17) F. Hirota, personal communication.
18) E. W. Baker and S. E. Palmer, "The Porphyrins," ed by D. Dorphin, Vol. 1, Academic Press, New York (1978), pp. 486—552.
19) S. L. Miller and L. E. Orgel, "The Origins of Life on the Earth," Prentice-Hall, Englewood Cliffs, NJ (1974).
20 W. J. le Noble, A. R. Miller, and S. D. Hamann, *J. Org. Chem.*, **42**, 338 (1977).