Bull. Chem. Soc. Ethiop. **2008**, 22(3), 433-440. ISSN 1011-3924
Printed in Ethiopia

TOWARDS A GENERIC MODEL OF CATALYSIS

Martin Grayson* and Stefan J. Janusz

Department of Chemistry, the University of Sheffield, Sheffield S3 7HF, UK

(Received February 4, 2008; revised March 22, 2008)

ABSTRACT. A generic method of modelling catalysis using fundamental experimental bond based data is described. This model is a form of electrostatically driven catalysis which has some considerable literature. It is proposed that it can be used in conjunction with molecular modelling visualization tools. Chemical concepts are applied without direct calculation of any electronic structure. Related to the electrically distorted bond model however is the geometrical distortion which forms the basis of the entasis effect. The current state of modelling entasis is reviewed but the model calculations presented here are of the electrical strain induced in a molecule prior to reaching the transition state. We consider polarizabilities and hardness/softness parameters to see how local polarizations of the electron density may also be responsible for activation of a localised area of a large molecule.

KEY WORDS: Electrostatic catalysis, Geometrical strain, Environment strain, Entasis, Polarizability, Hardness and softness

INTRODUCTION

The methodology presented here is simple, almost to the extreme, but contains within it tried and trusted experimental data, which in itself contains the quantum mechanical values by the inclusion of polarizabilities, and some values from density functional theory by the implicit use of electronegativity and hardness/softness parameters.

The basic idea is that to assist the breaking of a bond you must create a field which opposes the bond dipole of that bond leading to a localized increase in energy. The chemical bond then is in an electronic state more like the transition state than the initial state and the reaction can move forward. (If a single molecule were to be fixed in a hypothetical homogeneous electric field the energy induced in the molecule by that field would not be evenly distributed. A polar bonds would interact only at second order in electric field via the polarizability whereas polar bonds would either be enhanced in strength or have energy pumped into them by the field according to their orientation.)

It is apparent in this model that a high polarizability will allow a bond to become more reactive regardless of the direction of local fields it experiences. This also accords with the criteria in organic mechanisms for a good leaving group, which requires a high polarizability and a low value of the hardness/softness parameter. However we have to consider the balance of polarization energies across the whole reaction profile. It is likely that the polarization energy of reactants and products is similar. However for the reactive bonds polarization will increase at the transition state as the bond length is stretched or alternatively more excited states are mixed in a quantum mechanical sum over states picture. The effect of a high polarizability is to give a lower transition state energy.

The criteria for a good leaving group is usually dominated by the solvation of the leaving particle. This in itself is a large area of study and some aspects of it are mentioned later.

This model was invented independently before the work on electrostatically driven catalysis was found. Náray-Szabó and Ferenczy have for several years advanced this theory with respect to the large electric fields found near some solid surfaces and in zeolites [1, 2]. Table 1 illustrates the magnitudes of electric fields found in various important chemical environments.

*Corresponding author. E-mail: M.Grayson@sheffield.ac.uk

Table 1. The magnitudes of fields found in nature.

	Field (au)	Field (Vm^{-1} x 10^{-10})	Reference
In natural transitions in ice	0.000002-0.000018	0.0001-9	[3]
At the surface in an earthquake	0.00008	0.004	[4, 5]
Inside a semiconductor	0.0002	0.010	[6]
Typical finite field calculation	0.002	0.103	
Protein environment	0.006- > 0.008	0.309- > 0.411	[7, 8]
Strong zeolite	0.019	1	[1, 2]
Valence electrons	0.117	6	[1, 2]

A description of catalysis using essentially a quantum mechanical sum over states and perturbation theory methodology has been expounded by Tchougréeff [9]. Though this kind of approach is potentially exact it cannot be generic, because so much complexity is folded into the potential surface and the reaction coordinate. A great deal of information about the energetics of the system, if not the whole reaction, must be known or calculated before one knows this trajectory (or one may even have to sum over multiple trajectories), and so the solution is complex and different for every reaction. This leads to an explosion of information which makes it very difficult for the practical chemist to do an on the spot scrap paper analysis of reaction mechanisms.

Inorganic biological systems

A particularly interesting area of chemical catalysis which is largely unexplored by computational chemistry is the study of inorganic biomolecules. That inorganic elements are ubiquitous in living things is often overlooked, but they are actually instrumental in many vital processes [10]. Light-harvesting chlorophyll contains magnesium. The iron in haemoglobin carries a dioxygen ligand through mammalian blood. Copper-containing haemocyanin does the same thing in crustaceans. Ca^{2+} also has many roles in the soft tissues as well as the skeletons of living organisms.

Metals and their ions are also crucial in many catalytic processes, at the active site of enzymes. Because of the manner in which certain amino acid residues along the polypeptide ligate to the metal, it was originally thought that the metal's role was to provide a geometrical template to which the polypeptide could conform. However, crystallographic data have revealed that the geometry at metal centres in many metalloenzymes is 'non-standard'. Since 'standard' geometries are optimized for the best electronic interaction (and subsequent energy minimization through mutual stabilization) between the metal and its ligands it follows that 'non-standard' geometries imply that the system is energetically strained. At the metal centre at least, the enzyme is in a non-equilibrium state. In these so-called 'entatic' complexes (from the Greek entasis, meaning strain), the tertiary structure of the polypeptide influences the geometry at the metal centre, rather than (or perhaps, as well as) the metal influencing the conformation of the protein. Geometrical effects can compress or stretch bonds, entasis, but electrical effects can also weaken bonds by moving electron density away from the bonding region.

The entatic state and catalysis

Since the 1950s, it has been proposed that the entatic state of enzymes might account for their tremendous power as catalysts [11]. Any system in a configuration that lies away from that at

equilibrium must have an energy greater than the equilibrium, so it is possible that certain catalytic pathways could be opened up or made more favourable by an intrinsic strain. Many non-enzymatic catalysts involving d-block metals require certain geometric or electronic strains to be formed at some intermediate stage of the catalysis. An example is the ring-opening metathesis polymerization (ROMP) reaction [12], whereby a thermodynamically strained hydrocarbon chelate is briefly formed, causing the complex to fall apart. The strain gives the system an incentive to exit the intermediate chelate stage and so propagate the catalysis. Other systems, which attempt to mimic the strains inside enzymes with working 'small molecule' models (i.e. non-enzymatic), have been developed and investigated [13].

The above complexes are several orders of magnitude more proficient catalysts than anything seen in the traditional realm of d-block coordination complexes. Enzymes can catalyze chemical reactions whose natural (i.e. non-catalyzed) half-lives are of the order of the age of the universe, and effect these changes in time frames of less than a second [14]. A major difference between enzymes and 'traditional' (that is, artificial; non-biological) catalysts is that the protein component imparts a non-equilibrium geometry at the active site, and on the coordination sphere of the metal it contains. Only very exotic and unstable artificial complexes display anything approaching this amount of strain (note that enzymes, though technically quite strained, are also remarkably stable). So it seems plausible that the novel, 'strained' geometries seen at the active sites in metalloenzymes could be a factor in producing this remarkable catalytic proficiency.

Investigation of these systems with computer codes is desirable because attempts to synthesize working models of the active sites of enzymes have usually failed. The metal centres in these small molecule mimics, lacking the 'rigid' backbone of a polypeptide, tend to adopt the standard equilibrium (i.e. non-entatic) geometries described earlier [15]. This can be rationalized in several ways. Most likely is that it is far more energetically favourable for the ligands in an artificial catalyst to rearrange to positions in the coordination sphere that maximize overlap with the d-orbitals of the metal, so promoting a more thermodynamically stable structure. In the case of the enzyme, optimization of the geometry at the metal centre is a consideration of far less importance than the energetically unfavourable rearrangement of the bulk of the protein that such an action would necessitate. Generally, a protein's conformation is one that maximizes intramolecular interactions between domains within the structure. The number of thermodynamically favourable, so-called 'weak interactions', vastly outweigh the disfavoured mis-fitting at the metal centre, and it is conceivable that the folding of the protein is used to pay for the raising of the potential energy of the active site.

The active site in an enzyme can be thought of as existing in isolation. The polypeptide effectively buffers it from the external environment, and so it can exist in a strained state. In a calculation the model molecule is in isolation, unless the program is explicitly told to model the system as though there is a solvent, or an array of similar molecules, surrounding it. While real small molecule models of active sites, through energetic exchanges with their environment, break down - virtual models, effectively in isolation, do not. Computational theoretical work therefore offers a good way to study the entatic state.

A brief review of existing work on entasis

There exists a wealth of literature that has relevance to this problem, that all attack it from a unique perspective. Chemists, biochemists, biologists and biophysicists all have their own particular views, many contradictory to others; it seems that the contention of the matter has inspired much work, but papers offering a real insight into the phenomenon are rare because a simple solution is so elusive.

An early attempt to rationalize the problem was outlined in a paper by Vallee and Williams [15]. It was here that the term 'entatic' was first invoked to describe an enzymatic system "in a stretched state or under tension... (implying) a catalytically poised state intrinsic to the active

site". When bonds are stretched they are more polarizable and therefore more reactive. Because of the crudity of crystallographic analysis at the time, the nature of the active site was not known directly, but inferred from spectroscopic data. Metalloenzymes were deemed "exceptionally well suited for the examination of the physicochemical basis of [enzymes]...", because the physical (i.e. spectroscopic) properties of metals, "...constitute intrinsic probes." It was understood from the analyses of these data, that the natures of the metals in their coordination spheres were quite unlike those seen in model complexes (i.e. d-block coordination compounds). What was proposed, in light of the known proficiency of enzymes as catalysts, was that the active site has a configuration "...closer to that of a unimolecular transition state than to that of a conventional, stable molecule, thereby constituting an energetically poised domain."

In computational terms, much of the work on metalloenzymes has been tentative, but now the technical problems associated with mixing quantum mechanics and molecular mechanics, as a mixed QM/MM calculation have become routine, (it is now available in several computational chemistry packages), QM/MM can be used to calculate properties of the active site and polypeptide, respectively, in one calculation [16]. There have also been advances in terms of modelling proton tunnelling in enzymes [17]. Fariselli *et al.* [18] have modelled the electron correlation in haemocyanin. Most of the work in the field is of this nature: a thorough study of one particular system, and often of one particular phenomenon to which that system gives rise. Most of the time, answers have already been determined experimentally, so that theoretical work, especially that of a computational nature, often confirms facts rather than predicting new data.

The geometrical strain in the activated complex is best modelled by quantum mechanical calculation or QM/MM methods. However when thinking about the characteristics required of an active site, and as a back of the envelope thinking estimation, suitable for display in a molecular graphics system, bond polarization considerations can be a help. There is suggested a model with some predictive power, extrapolated from the nature of the bonds in the active site. The general idea is this: that a chemical bond's strength is at a maximum when the local dipole is the natural bond dipole, and by applying an electrostatic field, particularly, in opposition to the direction of that dipole, the bond is weakened. The model draws upon a dataset of bond based polarizabilities [19] to determine the likelihood of bond breaking (the more polar or polarizable the 'easier' energetically that it should be). The stability of the leaving group (more polarizable species tend to be better leaving groups) is also partially predicted by this model. The charge transfer component of the polarizability can also be influenced in the course of the reaction by the Pearson hardness factor b. In particular how much greater negative charge and how much increase in polarizability an atom or functional group undergoes is determined by the value of b. The value of a, the traditional electronegativity determines the unperturbed bond dipole.

CALCULATIONS

Electrical strain in a simple model

We will now make a digression to looking at one of the simplest reaction systems: how the methyl halides undergo SN2 hydrolysis and we will consider how the CH_3-X bond becomes broken.

From the modelling of protein NMR it is known that the sort of fields experienced in the equilibrium environment are on average 0.006 au rising to 0.008 au for the more perturbed atoms [7, 8] (1 au = 5.14220 x 10^{11} V m^{-1}). It is clear that the model here requires a field of about a half to one orders of magnitude greater than these equilibrium fields to put in as much energy as the bond energy. It is possible that a catalytic site can generate a change of electrical environment one order of magnitude greater than these normal perturbed environment. (Even

greater fields than these might be possible in zeolites [20].) Therefore fields of 0.04 and 0.08 au have been used in the model calculations, even though they will be higher than the fields involved in the transition states of the SN2 hydrolysis.

It should be noted that the electric field at the nucleus of an atom in a molecule is always zero at a minimum or stationary point geometry, otherwise there would be a force on the nucleus. This means the effect of the field is to slightly modify the electronic structure from the free molecule.

Table 2. Contributions to the induced energy at applied electric fields 0.04 and 0.08 au.

Bond	Energy from μ (kJ mol^{-1})	Energy from α (kJ mol^{-1})	Total energy (kJ mol^{-1})	Bond energy (kJ mol^{-1})
C-H (0.04)	53.72	8.10	61.81	414.22
C-H (0.08)	107.43	33.76	139.82	414.22
C-F	187.60	8.49	196.08	485.34
C-F	375.19	33.94	409.13	485.34
C-Cl	202.47	29.48	254.49	338.90
C-Cl	404.94	208.08	613.02	338.90
C-Br	195.86	43.66	261.77	284.51
C-Br	391.72	263.64	655.36	284.51
C-I	172.31	66.05	268.83	217.57
C-I	344.62	386.11	730.72	217.57

It is hoped that this model can be combined with a model of solvation effects in enzyme reactions, such as described by Warshel [21]. Since these reactions take place in aqueous conditions, the Sheffield solvation model, recently developed by the Pickup group specifically for high throughput computations, is potentially useful [22].

In the calculations in Table 2 an electric field is placed along the C-X bond in an energetically unfavourable direction with the bond dipole. The first column is the 1^{st} order interaction with the dipole, then the 2^{nd} order with the polarizability tensor elements with the correct symmetry, i.e. the parallel diagonal element. The total is compared with the bond energy. The bond dipoles and polarizabilities are taken from references [23] and [19] and given in Tables 3 and 4. For the smaller field 0.04 the energy enhancement only exceeds the bond energy for the reactive C-I bond. However for the larger field 0.08 once we are passed the unreactive C-H and C-F bonds all the bonds can be readily broken. In this limited case the simple model predicts chemical common sense.

Table 3. Bond dipole moments x 10^{30} (cm).

Carbon sp^3	Hydrogen	-1.300
Carbon sp^3	Fluorine	4.540
Carbon sp^3	Chlorine	4.900
Carbon sp^3	Bromine	4.740
Carbon sp^3	Iodine	4.170

Table 4. Polarizability ellipsoids x 10^{41} ($C^2m^2J^{-1}$).

		Parallel	Perpendicular
Carbon sp^3	Hydrogen	6.356	6.356
Carbon sp^3	Fluorine	6.661	6.661
Carbon sp^3	Chlorine	40.834	23.143
Carbon sp^3	Bromine	51.738	34.270
Carbon sp^3	Iodine	75.772	51.850

It has been hypothesised that when the polarizability of a system increases due to geometrical distortion more of the polarizability increase goes with the anion than with the rest of the molecule. This seems like common sense and the following ab initio calculation of distributed polarizabilities [24] as a function of bond length (Table 5) shows this is indeed the case for the alkyl halides at least.

Table 5. Distributed polarizabilities and gradients for alkyl halides.

	Polarizabilities ($C^2m^2J^{-1}$)			Polarizability gradients ($C^2m^2J^{-1}$ per Ångstrom)		
	C	Hal	H	C	Hal	H
CH_3F	11.12	11.12	3.48	4.87	6.68	-0.98
	(¶) 39.21 %	24.02 %	12.26 %	(§) 14.51	7.09	4.95
CH_3Cl	10.41	27.13	3.62	5.11	12.58	0.15
	21.52 %	56.06 %	7.47 %	10.55	17.31	2.31
CH_3Br	8.91	14.56	2.79	5.94	7.83	0.7
	27.97 %	45.71 %	8.77 %	4.78	8.9	0.16
CH_3I	9.3	22.67	3.09	6.62	8.29	0.99
	22.56 %	54.98 %	7.49 %	5.11	0.9	0

(¶) - The following 3 numbers are the percentage of the total molecular polarizability. (§) - The following 3 numbers are the 2^{nd} gradients of the polarizability.

The calculations are inconsistent in that it is not possible to have a uniform quality of basis set whilst going down a column of the periodic table from F to I. CH_3F and CH_3Cl are calculated using Sadlej's medium polarized basis set [25]. CH_3Br and CH_3I use only 3-21G [26]. However we see the polarizability gradient is always largest for the halogen and is more or less constant for the carbon in CH_3F and CH_3Cl. If we had a better Hartree-Fock basis for CH_3Br and CH_3I all gradients would presumably be larger and the carbon begins to own more of the polarizability because it has some of the iodine's polarization in its partitioning polyhedra because the iodine is so much larger and more polarizable. The integration over the Voronoi polyhedra is not an exact partitioning whatever exact means in this context. These simple calculations confirm what might be expected. As the bond which creates the leaving group is stretched, the leaving group becomes more polarizable relative to the rest of the molecule. At a transition state the whole complex has become more polarizable, (only one gradient is negative, that corresponding to the unreactive C-H bond in CH_3F). As expected in CH_3F the carbon atom owns the largest polarizability but from CH_3Cl onwards to the iodide the halogen owns the largest polarizability.

It can be noticed that chlorine has both a large 1st and 2nd gradient of the polarizability. This is analogous to the known behaviour where in some reactions chlorine behaves as though it were almost as electronegative as fluorine due to its much smaller value of the hardness parameter b. This is particularly apparent if one uses the more recent values of a and b from Politzer *et al.* [27] (Table 6) rather than the earlier values which are in many textbooks.

Table 6. Values of electronegativity parameters *a* and *b, units of Volts / electron* units of Volts / electron.

	a	*b*
Fluorine	12.18	17.36
Chlorine	9.38	11.30
Bromine	8.40	9.40
Iodine	8.10	9.15
The above data is from Huheey [23]		
Fluorine	10.41	14.03
Chlorine	8.29	9.35
Bromine	7.59	8.48
Iodine	6.76	7.41
Neon	10.78	21.55
Argon	7.88	15.75
Krypton	7.00	14.01
Xenon	6.07	12.14
The above data is from Politzer *et al.* [27]		

CONCLUSIONS

We have suggested a model which may be useful in visualising catalysis but can be used for either computation or quick estimation. It uses bond parameterizations which can be directly fitted to experiment or come from calculations on small model systems. However the ubiquity of additive and transportable concepts in chemistry such as functional groups, bond energies and spectroscopic tables indicates that bond based properties, as used here, are translatable from molecule to molecule.

We hope by suggesting a merging of ideas from entasis and electrostatic catalysis we might have added to the tools available for rationalizing reactions. It is intended to develop this model further so that the derived concepts such as polarizability, hardness/softness, electronegativity, bond and functional group properties can be applied to the large molecule problems of enzyme activity and inorganic catalysis.

AKNOWLEDGEMENTS

We would like to thank the following people for their help and support during the project which lead to this communication: Prof. David E. Fenton, Dr. Nicholas H. Williams and Prof. Barry T. Pickup. We would also like to thank the developers of the CADPAC program for making the distributed polarizability computations possible. We wish to acknowledge the use of the EPSRC's Columbus Service at the Rutherford Laboratory.

REFERENCES

1. Náray-Szabó, G.; Ferenczy, G.G. *Chem. Rev.* **1995**, 95, 829.
2. Náray-Szabó, G. *Recent Res. Devel. Physical Chem.* **2002**, 6, 567.
3. Choi, E.M.; Yoon, Y.H.; Lee, S; Hang, H. *Phys. Rev. Lett.* **2005**, 95, 085701.
4. Enriquez, A. *New Scientist* **2003**, 179, 2402, 26.
5. Losseva, T.V.; Menchinov, I.V. *Natural Hazards and Earth System Sciences* **2005**, 5, 649.
6. Pachioni, G.; Lomas, J.R.; Illas, F. *J. Mol. Catalysis* **1997**, 119, 263.
7. Horsley, W.J.; Sternlicht, H. *J. Amer. Chem. Soc.* **1968**, 90, 3738.
8. Augspurger, J.; Pearson, J.G.; Oldfield, E.; Dykstra, C.E.; Deok Park, K.; Schwartz, D. *J. Magn. Reson.* **1992**, 100, 342.
9. Tchougréeff, A.L. *Intl. J. Quantum Chem.* **1996**, 58, 67.
10. Fenton, D.E. *Biocoordination Chemistry*, Oxford University Press: Oxford; **1995**.
11. Williams, R.J.P. *Eur. J. Biochem.* **1995**, 234, 363.
12. Trnka, T.M.; Grubbs, R.H. *Acc. Chem. Res.* **2001**, 34, 18.
13. Williams, N.H.; Wyman, P. *J. Chem. Soc. Perkin Trans.* **2001**, 2, 2068.
14. Wolfenden, R.; Snider, M.J. *Acc. Chem. Res.* **2001**, 34, 938.
15. Vallee, B.L.; Williams, R.J.P. *Proc. Natl. Acad. Sci. USA* **1968**, 59, 498.
16. Gao, J.; Amara, P; Alhambra, C.; Field, M.J. *J. Phys. Chem. A* **1998**, 102, 4714.
17. Truhlar, D.G.; Gao, J.; Alhambra, C.; Garcia-Viloca, M.; Corchado, J.; Luz Sanchez, M.; Villa, J. *Acc. Chem. Res.* **2002**, 35, 341.
18. Fariselli, P.; Bottoni, A.; Bernadi, F.; Casadio, R. *Protein Sci.* **1999**, 8, 1546.
19. Grayson, M. *The Estimation of Molecular Polarizability and Magnetizability* in *Recent Research Developments in Physical Chemistry*, Pandalai, S.G. (Ed.), Vol. 6, Transworld Research Network: Trivandrum, India; **2002**; pp 437-455.
20. Khaliullin, R.Z.; Bell, A.T.; Kazansky, V.B. *J. Phys. Chem. A* **2001**, 105, 10454.
21. Warshel, A. *Acc. Chem. Res.* **1981**, 14, 284.
22. Grant, J.A.; Pickup, B.T.; Sykes, M.J.; Kitchen, C.A.; Nicholls, A. *Chem. Phys. Lett.* **2007**, 441, 163.
23. Huheey, J.E.; *Inorganic Chemistry*, 3rd ed., Harper International: New York; **1983**.
24. Stone, A.J. *The Theory of Intermolecular Forces*, Clarendon: Oxford; **1996**.
25. Sadlej, A.J. *Collect. Czech. Chem. Commun.* **1998**, 53, 1995.
26. Foresman, J.B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed., Gaussian, Inc.: Pittsburgh, PA, 15106 USA; **1993**.
27. Politzer, P.; Murray, J.S.; Grice, M.E. *Structure and Bonding, 80, Chemical Hardness, Springer*: Berlin; **1993**; p 101.