Computation Of Hydrogen Bond Basicity As A Descriptor In Bioisosterism: A Quantum Chemical Topology Perspective.

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences

2013

Anthony James Green

School of Chemistry

List Of Contents.

Preliminary Information

Background To Research

Research Report

Word Count: 56,848

List Of Tables.

hydrogen bond acceptors with 2 potential sites of protonation that are housed in chemically different environments have been included. (159).

Table 3.8.4. The hydrogen bond acceptors listed in this table form the data set for a set of calculations in which protonated methanol is donating a hydrogen. The hydrogen bond acceptors in this table are the strongest proton acceptors and have pK_{BH}^{+} values of 9 or above. (162).

Table 3.8.5. All hydrogen bond acceptors used in computations to form hydrogen bond complexes with methanol used in this research along with their pK_{BHX}, $\Delta E(H)$, and Z values. The Z value is equated from a Grubb's test for an outlier. A description of the Grubb's test is given in the text above. No Z values are greater than the critical value of 3.75 at the 0.05 significance level so no outliers detected. (179).

Table 3.9.1. The hydrogen bond acceptors used in this section. The hydrogen bond acceptors are ranked in order of Brönsted basicity (pK_{BH}^{+}) from strongest to weakest. (185).

Table 3.10.1. The hydrogen bond acceptors listed in this table have been from the hardest to the softest. The hardness values stated here have been computed as described in the above text. (196).

Table 3.11.1. The hydrogen bond acceptors used in this section. Each hydrogen bond acceptor accepts two separate hydrogen bonds from methanol. The details of how the complexes have been set up are described in the text. (205).

Table 3.11.2. The data in this table relates to computations performed on hydrogen bond complexes where each hydrogen bond acceptor receives two separate hydrogen bonds from two methanol hydrogen bond donors. The subscripts 1 and 2 are put in place to distinguish between data relating to each individual hydrogen bond. Details of how predicted pK_{BHX} and Ki values wrer calculated are given in the text. (208).

Table 3.11.3. The data in this table considers each hydrogen bond separately. The optimised geometry is from the hydrogen bond complexes where each hydrogen bond acceptor receives two separate hydrogen bonds from two methanol hydrogen bond

donors. However this time each hydrogen bond has in turn been deleted and energy point calculations performed on each single hydrogen bond complex. (214).

List Of Figures.

Figure 3.6.1. Strychnine contains six potential hydrogen bond acceptor sites (123).

Figure 3.6.2. Strychnine can be fragmented into its major hydrogen bond acceptor sites. The three major hydrogen bond acceptor sites of strychnine are the amide oxygen, the ether oxygen and the amine nitrogen. (126).

8between pK_{BHX} and ΔE(H) for the set of 12 tertiary amines listed in table3.7.1. The data points are labelled by pK_{BHX} values and can therefore be identified with the tertiary amine they represent. (123).

Figure 3.7.2. Protonated tertiary amine diazabicyclooctane donating a hydrogen bond to the oxygen atom of methanol. (130).

Figure 3.7.3. The relationship between pK_{BHX} and ΔE(H) for the set of 12 protonated tertiary amines listed in table 3.7.1 The data points are labelled by pK_{BHX} values and can therefore be identified with the protonated tertiary amine they represent. The protonated tertiary amine acts as the hydrogen bond donor whereas the oxygen atom of methanol has assumed the role of the hydrogen bond acceptor. (131).

Figure 3.7.4. An example of a hydrogen bond complex where protonated trimethyl amine is the hydrogen bond donor. The NH^+ site of protonated trimethyl amine is donating a hydrogen bond to the oxygen atom in acetone. (134).

Figure 3.7.5. The relationship between pK_{BHX} and ΔE(H) for hydrogen bond complexes formed between the bases in table 3.7.2 and the hydrogen bond donor protonated trimethyl amine. (134).

Figure 3.7.6. The relationship between pK_{BHX} and ΔE(H) for hydrogen bond complexes formed between the bases in table 3.7.2 and the hydrogen bond donor methanol. (135).

Figure 3.7.7. An example of a starting geometry for a complex behaving as a zwitterion. Protonated diazabicyclooctane is donating a hydrogen bond to deprotonated methanol. (137).

Figure 3.7.8 The relationship between pK_{BHX} and ΔE(H) for hydrogen bond complexes formed between the tertiary amines in table 3.7.1 and methanol. The complex is model to behave as a zwitterion as the methanol is deprotonated and behaves as the hydrogen bond acceptor whereas the tertiary amine is protonated and behaves as the hydrogen bond donor. (137).

Figure 3.7.9 When modelling the hydrogen bond complex like a zwitterion, initially the tertiary amine is protonated and is donating a hydrogen bond to deprotonated methanol. During the optimisation step, a full proton transfer is observed from the tertiary amine to methanol. The resulting complex is one where methanol is donating a hydrogen bind to the tertiary amine. This is displayed here using diazabicyclooctane as an example of a tertiary amine. (139).

Figure 3.8.1. From the bases listed in table 3.8.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methanol is the hydrogen bond donor. (151).

Figure 3.8.2. The correlation between ΔE(H) and pK_{BHX} for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. (164).

Figure 3.8.3. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values. The method used to compose the subsets is described in the text above. Correlations between ΔE(H) and pK_{BHX} have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. (164).

Figure 3.8.4. The correlation between $\Delta E(H)_{(MeOH\text{-}HBD)}$ and pK_{BHX} for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. The meaning of $\Delta E(H)_{(MeOH\text{-}HBD)}$ is described in the text above. (166).

Figure 3.8.5. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values. The method used to compose the subsets is described in the text above. Correlations between $\Delta E(H)_{(MeOH\text{-}HBD)}$ and pK_{BHX} have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. The meaning of $\Delta E(H)_{(MeOH\text{-}HBD)}$ is described in the text above. (166).

Figure 3.8.6. The correlation between $\Delta E(H)_{(Methanol)}$ and pK_{BHX} for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. The meaning of $\Delta E(H)_{(Methanol)}$ is described in the text above. (169).

Figure 3.8.7. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values. The method used to compose the subsets is described in the text above. Correlations between $\Delta E(H)_{(Methanol)}$ and pK_{BHX} have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. The meaning of $\Delta E(H)_{(Methanol)}$ is described in the text above. (170).

Figure 3.8.8. The correlation between $\Delta E(H)_{(Complex)}$ and pK_{BHX} for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. The meaning of $\Delta E(H)_{(Complex)}$ is described in the text above. (171).

Figure 3.8.9. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values. The method used to compose the subsets is described in the text above. Correlations between $\Delta E(H)_{(Complex)}$ and pK_{BHX} have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. The meaning of $\Delta E(H)_{(Complex)}$ is described in the text above (169).

Figure 3.8.10. The relationship between experimental pK_{BHX} and computed $\Delta E(H)$ values for the 332 hydrogen bond acceptors listed in table 3.8.5. The hydrogen bond donor used in the computations was methanol. (178).

Figure 3.9.1. From the bases listed in table 3.9.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value of each base and the computed $\Delta E(H)$ value for each base in each 15 base subset. Computed $\Delta E(H)$ values are taken from computed hydrogen bond complexes where methanol is the hydrogen bond donor. (186).

Figure 3.9.2. From the bases listed in table 3.9.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value of each base and the computed $\Delta E(H)$ value for each base in each 15 base subset. Computed $\Delta E(H)$ values are taken from computed hydrogen bond complexes where methylamine is the hydrogen bond donor. (187).

Figure 3.9.3. A direct comparison between methanol and methylamine hydrogen bond donors. From the bases listed in table 3.9.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value of each base and the computed $\Delta E(H)$ value for each base in each 15 base subset. The blue diamonds represent the data series where methanol has been used as the hydrogen bond donor

whereas the red dots represent the data series where methylamine was used as the hydrogen bond donor. (188).

Figure 3.10.1. Hardness values have been computed for the hydrogen bond acceptors listed in table 3.10.1 and plotted against data base pK_{BHX} values. The r^2 value of 0.0065 indicates no relationship between hardness and pK_{BHX}. (197).

Figure 3.10.2. From the bases listed in table 3.10.1, 15 base subsets have been chosen according to the criteria in the above text. The mean hardness value is plotted for each 15 base subset. The r^2 value is taken from correlations between the hardness value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methanol is the hydrogen bond donor. (198).

Figure 3.10.3. From the bases listed in table 3.10.1, 15 base subsets have been chosen according to the criteria in the above text. The mean hardness value is plotted for each 15 base subset. The r^2 value is taken from correlations between the hardness value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methylamine is the hydrogen bond donor. (199).

Figure 3.10.4. The overall relationship between pK_{BHX} and ΔE(H) for all 197 hydrogen bond donors listed in table 3.10.1. Methanol is the hydrogen bond donor. (199).

Figure 3.10.5. The overall relationship between pK_{BHX} and ΔE(H) for all 197 hydrogen bond donors listed in table 3.10.1. Methylamine is the hydrogen bond donor. (201).

Figure 3.11.1. The interaction between the human RXH alpha ligand binding domain and docosa hexanoic acid [2]. The interaction image is taken from the protein data bank [3]. (203).

Figure 3.11.2. An example of a hydrogen bond acceptor accepting two separate hydrogen bonds from two separate hydrogen bond donors. In this example ethanol is accepting two separate hydrogen bonds from two methanol hydrogen bond donor molecules. (205).

Figure 3.11.3. The relationship between pK_{BHX} and ΔE(H) for the set of 15 hydrogen bond acceptors listed in table 3.11.1 Methanol is the hydrogen bond donor. A loose optimisation criteria has been used during the geometry optimisation phase. (206).

Figure 3.11.4. Optimised geometry of the 4-methylphenol complex. Two separate methanol hydrogen bond donors each donate a hydrogen bond to 4-methylphenol. As can be seen here the complex is virtually symmetrical, however the hydrogen bond on the left generates a much lower ΔE(H) value than the hydrogen bond on the right. (207).

Figure 3.11.5. The data plotted here is taken from table 3.11.2. The predicted total pK_{BHX} relates to the hydrogen bond complexes where two individual methanol donors each donate a separate hydrogen bond to a single hydrogen bond acceptor site. Data for 3-methylphenol and 4-methylphenol have been omitted from this plot for reasons explained in the above text. (210).

Figure 3.11.6. An example of a hydrogen bond acceptor accepting two separate hydrogen bonds from two separate hydrogen bond donors. In this example acetone is accepting two separate hydrogen bonds from two methanol hydrogen bond donor molecules. The hydrogen bond donor methanols are distributed symmetrically about the hydrogen bond acceptor site. (210).

Figure 3.11.7. An example of a hydrogen bond acceptor accepting two separate hydrogen bonds from two separate hydrogen bond donors. In this example acetonitrile is accepting two separate hydrogen bonds from two methanol hydrogen bond donor molecules. The hydrogen bond on the right has a gives a weaker estimated pK_{BHX} than the one on the left. This is possibly due to a secondary interaction between the methanol oxygen on the right hand side and the methyl hydrogen of acetonitrile. (211).

Abstract.

University of Manchester
Anthony James Green
Doctor of Philosophy
2013

Computation Of Hydrogen Bond Basicity As A Descriptor In Bioisosterism: A Quantum Chemical Topology Perspective.

Hydrogen bonding is a regularly occurring non covalent interaction in biological systems. Hydrogen bonding can influence a drug's interaction with its target. It is therefore important to practically measure the relative strengths of hydrogen bonds. Hydrogen bond basicity is a measure of a hydrogen bond acceptor's capacity to accept hydrogen bonds. There are many hydrogen bond basicity scales. However, the pK_{BHX} scale is claimed to be the most relevant to medicinal chemists because it gives a thermodynamically deducible values for each site in polyfunctional bases. A computed property, the change in energy of the hydrogen bond donor hydrogen bond atom ΔE(H), derived from the quantum theory of atoms in molecules has been found to correlate strongly with pK_{BHX} values for OH and NH hydrogen bond donors. In particular, R^2 values of 0.95 and 0.97 have been found when methanol and methylamine respectively are used as hydrogen bond donors. The property ΔE(H) has also been successfully used to predict the pK_{BHX} values of an external data set and the values of polyfunctional bases. The strength of the correlations are not dramatically affected by using scaled down fragments of bases, or by relaxing the convergence criteria during the geometry optimisation step of calculations. The relationship between ΔE(H) and pK_{BHX} has been found to break down for tertiary amines, and more generally for strong proton acceptors with pK_{BH}^{+} values greater than 6. The successful pK_{BHX} prediction model was, however, unsuccessful in predicting drug binding data and pK_{BHX} values of bases that accept two separate hydrogen bonds. At this moment in time both the reason why the relationship between pK_{BHX} and ΔE(H) is present and then breaks down for strong proton acceptors is unfortunately unknown.

Declaration.

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of the University of Manchester or any other university or institute of learning.

Copyright Statement.

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given the University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf, in any relevant thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on presentation of Theses.

Acknowledgement.

I would like to thank everybody who has helped me complete this thesis, both on an academic and personal level. Firstly, a special thanks Prof Paul Popelier for scientific advice, personal support and much encouragement and patience. To Nathan Kidley and John Delaney at Syngenta who provided insight that helped to drive the research forward. Thanks to Syngenta for providing funding. Thanks to my parents Howard and Carole. Without their constant help and support the completion of this thesis would not have been possible. To my sister Joanna. Thanks to all members of the Popelier group past and present for numerous discussions and technical support. Finally, thanks to all my friends.

List Of Abbreviations Used.

ADMET	Absorbtion-Distribution-Metabolism-Excretion-Toxicity
AIL	Atomic Interaction Line
AIM	Atoms In Molecules
BCP	Bond Critical Point
BSSE	Basis Set Superposition Error
CP	Critical Point
DFT	Density Functional Theory
FTIR	Fourier Transform Infrared Spectroscopy
GTO	Gaussian Type Orbital
LFER	Linear Free Energy Relationship
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HF	Hartree Fock
HSAB	Hard Soft Acid Base Principle
IAS	Inter Atomic Surface
PCA	Principal Component Analysis
QID	Quantum Isostere Database
QTAIM	Quantum Theory of Atoms In Molecules
SCF	Self Consistent Field
STO	Slater Type Orbital

Background To Research.

1 – Introduction

Hydrogen bonds are non-covalent intermolecular interactions that are found throughout biological systems. Hydrogen bonds are electrostatic interactions between a hydrogen atom bonded to an electronegative atom, and an electron rich area such as an electronegative species, a lone pair of electrons or a pi system. Hydrogen bonding is of particular interest due to its importance in drug distribution and interaction thermodynamics. Polar groups on drug molecules may form hydrogen bonds with biological solvents during absorption and distribution. Drug-solvent hydrogen bonds are broken and drug-target hydrogen bonds formed during drug binding. Therefore a measure of the relative strengths of hydrogen bonds is useful in drug design and development. The relative strengths of the electronegative species that form hydrogen bonds with an electron depleted hydrogen atoms are known as hydrogen bond basicities. As many drugs have multiple hydrogen bond sites, it is necessary for basicity values to be obtained for each hydrogen bond site.

Drug molecules are usually large chemical structures of which only a few small sites interact with the target. It is practical, especially in computational chemistry to consider these small sites individually rather than the molecule as a whole. The drug molecule may be chopped up into smaller sections called fragments. Fragments retain the chemical activity of the site of the drug they have been cut from, and vastly reduce the time taken to perform any calculations on them. Fragments may be altered to improve biological activity. Many properties may be altered in order to improve the biological activity of a drug. Hydrogen bonding is one of these properties. One of the concepts that is widely applied to drug discovery and development is bioisosterism.

1.1 Bioisosterism In Drug Design And Disovery.

Bioisosterism is a method of fragment based drug chemistry that substitutes an active fragment with one of similar chemical properties. One of the properties that may be considered when making a bioisosteric stbstitution is hydrogen bonding.

Bioisosterism is a concept widely applied in medicinal chemistry and drug discovery [4-6]. A bioisostere is a substituent of a fragment of an active compound with similar physical or chemical properties [7]. In the context of drug design and development, bioisosteres are used to enhance the desired biological effects such as target selectivity and administration/absorption, distribution, metabolism [8], excretion/elimination and toxicology (ADMET) profile [9] without drastically altering the initial chemical structure [10-12]. During the drug design and development process, a compound with desired biological activity is repeatedly modified to give new candidate drugs with similar but hopefully improved biological effects [13, 14]. This process is called lead optimisation [15-21].

Pairs of molecules can be compared to each other using similarity measurements [22, 23]. Molecules can be compared by global measures or local measures [24]. Global measures calculate how similar one chemical structure is to another whereas local similarity measures the similarity between fragments of molecules [22]. The information stored in a chemical structure can be written in the form of a vector as shown in equation 1.1,

$$X_A = \{P_{1A}, P_{2A}, P_{3A}, \ldots P_{iA}, \ldots P_{nA}\} \qquad (1.1)$$

where X_A represents molecular fragment A, and the P_{iA}s represents the chemical properties that describe fragment A. The Quantum Isostere Database (QID) [25] runs a similarity search between a query fragment and fragments stored in the database. The properties the QID compares include conformation, molecular electrostatic potential,

charge distribution, polarity, shape and hydrogen bonding [25]. Similarity scores are given in terms of Euclidean distances from the query fragment in an *n*-dimensional property space defined by the *n* properties that make up the fragment vector. Equation 1.2 shows how Euclidean distances are calculated between fragments *A* and *B* (D_{AB}),

$$D_{AB} = \sqrt{\sum_{i=1}^{n}\left(I_i\left(\frac{P_{iA} - P_{iB}}{P_{i(\max)} - P_{i(\min)}} \right)\right)^2} \qquad (1.2).$$

In equation 1.2, $P_{i(max)}$ and $P_{i(min)}$ refer to the upper and lower values of the ith property. This measure serves to normalise the Euclidean distances. The *I* term is a user-specified multiplier representing the importance of a search term. Equations 1.1 and 1.2 show how molecules are described in terms of certain properties, and how they may be compared by the properties by which they are described. However, there are no strict rules that dictate exactly how similar in activity replacements must be before they can be called bioisosteric. The following section discusses the criteria that must be met in order for a pair of substituents to be defined as bioisosteric.

The original definition of bioisosterism stated that replacements are bioisosteric if structurally related groups display similar biological activity [26]. The definition of bioisosterism is an extension of the definition of chemical isosterism by Langmuir [27]. Chemical isosterism describes similarities in physical properties between species. The similarities observed were between species with the same number of valence electrons, N_2O and CO_2 for example. However, there are molecules with the same number of valence electrons but containing a different number of atoms that also display similar physical properties. Grimm's Hydride Displacement Law [28] accounts for this observation. The Hydride Displacement Law illustrates the physical similarities between an atom with a certain number of valence electrons and the hydride of the atom adjacent to it in the periodic table in the previous column. For example O, NH and CH_2 are all isosteric.

Bioisosterism can be separated into classical bioisosterism and non-classical bioisosterism [29]. Classical bioisosteres satisfy the original criteria of Langmuir [27] and

Grimm [28]. Non-classical bioisosteres do not show the same structural, electronic or atomic properties as their classical relatives. Instead, non-classical bioisosteres account for a wider range of compounds that display similar biological properties. Therefore, with non-classical bioisosterism in mind, the definition of bioisosterism was extended to include a wide class of structurally unrelated molecules that display similar biological properties [30]. Given the diversity of the chemical properties of fragments that are considered to be bioisosterically equivalent, there are huge numbers of potentially bioisosteric replacements. The BIOSTER database [31] contains thousands of pairs of molecules that are considered to be bioisosteric replacements. It may therefore seem as though there are endless possibilities of bioisosteric replacements to consider in the process of lead optimisation. However, lead optimisation is a case specific process. Only the properties important to the biological activity in each specific case need to be enhanced. Therefore only bioisosteres relating to specific activity need to be considered. Unfortunately, this level of accuracy is often unknown. Instead, a number of broad properties have been outlined [30]. These properties relate to both the drug molecule and its environment, and serve as a starting point when considering bioisosteric substitutions. The properties are size, shape, electronic distribution, lipid solubility, water solubility, pK_a, chemical reactivity and hydrogen bonding capacity. Of the eight broad properties outlined by Thornber [30], water and lipid solubility, and hydrogen bonding capacity will be discussed in detail in the following text.

1.2 Hydrogen Bonding.

Hydrogen bonding is an important non-covalent interaction between a biological molecule and its local environment [32]. In biological systems the local environment may be a protein [33], enzyme [34], a lipid membrane, an aqueous solution [35], a carbohydrate or RNA or DNA. When considering a drug molecule interacting with its local environment it is therefore clear that hydrogen bonding influences both the interaction with the target [36-38], and its permeability and solubility through lipid membranes and

aqueous solutions, respectively [39] [40]. A link can be formed between the influences of hydrogen bonding on target selectivity and solvation effects [41]. The binding enthalpies and entropies of a drug interacting with its biological target are affected by hydrogen bonds formed between polar groups and the aqueous solution in which the drug molecule is carried [42, 43]. Unfavourable binding enthalpy changes are due to the desolvation of polar groups exposed on the drug molecule [42]. As the layer of solvent surrounding the drug molecule is shed, hydrogen bonds are broken resulting in an unfavourable enthalpy increase. However, there is a favourable entropy increase upon binding due to the layer of solvent being shed. Hydrogen bonds formed between the drug molecule and its target result in favourable enthalpy changes. Also as the drug binds the target there is a respective unfavourable loss of conformational entropy due to the drug becoming fixed in the binding site. This poses a problem in the optimisation of drug candidates because favourable thermodynamic changes must be maximised whilst unfavourable thermodynamic changes are minimised [44, 45]. Searching for bioisosteric replacements using hydrogen bonding as a descriptor could therefore lead to improved thermodynamic properties of the drug candidate.

A hydrogen bond is a non-covalent interaction between an electronegative species and a polar hydrogen atom bonded to an electronegative atom. The molecule in which the hydrogen atom is housed is known as a hydrogen bond donor (HBD) and the molecule containing the electronegative species that the hydrogen interacts with is called the hydrogen bond acceptor (HBA). During hydrogen bond formation electron density is transferred from non-bonding electron pairs or a π system on a HBA, towards the hydrogen of the HBD [46]. Lewis acids are acceptors of electron density whereas Lewis bases donate electron density. Therefore HBDs and HBAs can be thought of as special cases of Lewis acids and bases, where HBDs are Lewis acids and HBAs are bases [1]. The acidity of HBDs and the basicity of HBAs have been described in several different scales

1.3 The General Solvation Equation.

Hydrogen bonding is involved in the lipid and water solubility of biological molecules. Many biological processes involve the transport of solutes between different phases, for example the blood-brain barrier [47] in the body, or the cuticular matrix in plants [48]. The general solvation equations, given by equations 1.3.1 and 1.3.2, have been developed by Abraham to model the solubility of a series solutes in specific phases [40] [49] [50]

$$\log SP = c + rR_2 + s\pi_2^H + a\sum\alpha_2^H + b\sum\beta_2^H + \ell \log L \qquad (1.3.1)$$

$$\log SP = c + rR_2 + s\pi_2^H + a\sum\alpha_2^H + b\sum\beta_2^H + vV_x \qquad (1.3.2).$$

SP relates to a set of solute water-solvent partition coefficients in a given system, R_2 is the excess molar refraction [51], π_2^H is the dipolarity / polarisability [52], α_2^H [53] and β_2^H [54] are the hydrogen bond acidity and hydrogen bond basicity. V_x is McGowan's characteristic molecular volume (cm^3 mol^{-1} / 100) [55] and logL is the gas-hexadecane partition coefficient [56]. The coefficients c, r, s, a, b and l or v are system specific constants which characterise and contain chemical information about the phae. Equations 1.3.1 and 1.3.2 differ by only the final terms, the McGowan's molecular volume and the gas-hexadecane partition coefficients. These terms both describe solute size. Equation 1.3.1 is usually used for the transfer between the gas phase and a condensed phase, whereas equation 1.3.2 is typically used for transport processes involving two or more condensed phases [48]. Therefore the choice of whether to use equation 1.3.1 or 1.3.2 depends on the characteristics of the solvent.

The general solvation equations are part of the wider field of linear free energy relationships (LFERs) [57]. A LFER is a linear correlation between the logarithm of a rate constant or equilibrium constant for one series of reactions, and the logarithm of the rate

constant or equilibrium constant for a related series of reactions. Therefore the parameters in equations 1.3.1 and 1.3.2 were originally obtained experimentally [49]. There have been several attempts to predict these parameters theoretically by correlating them with a number of theoretical descriptors. It has been found that the polarisability parameter correlates successfully to the molecular dipole moment, the polarisability, CHelpG atomic charges and the frontier molecular orbital energies [58]. This has been found by using partial least squares models [59]. Partial least squares models are predictive statistical methods based on multiple linear regression. The hydrogen bond acidity correlates with the electrostatic potential at the nucleus of the hydrogen atom [60]. Less successful attempts have been made to correlate theoretical descriptors to the hydrogen bond basicity parameter [61]. The following section will discuss the construction of hydrogen bond basicity scales and their advantages and disadvantages. Reasons why it has proven more difficult to correlate theoretical descriptors to hydrogen bond basicity will also be discussed. A refinement for the model of the prediction of hydrogen bond basicity from theoretical descriptors will also be suggested.

1.4 Hydrogen Bond Basicity Scales.

The importance of hydrogen bonding in medicinal chemistry has been discussed in the previous section. However, despite the importance of hydrogen bonding in biological systems, there is a general lack of understanding of the relative strengths of hydrogen bond basicity [1]. The relative basicity or strengths of HBAs are often incorrectly correlated with the pK_{HB} scale based on the protonation of a basic site on a Brønsted base [1]. Although various family-dependent relationships can be observed, Brønsted proton basicity scales should not be used to reliably predict the relative strengths of HBAs. One of the main reasons for this is that the main protonation site of a HBA, is often different to the major hydrogen bond acceptor site [1]. The basicity scales that will be discussed here are based on equilibrium constants rather than binding

enthalpies [62]. By considering equilibrium constants rather than binding enthalpies entropic contributions to complex formation are accounted for. This allows the full thermodynamic profile to be obtained.

The first scales set up to measure hydrogen bond basicity were by Taft and co-workers [63] [64]. The initial scale set up investigated LFERs between 4-fluorophenol and a series of other reference (-OH) HBDs [63]. The experiments were set up to measure the equilibrium constants K_f for HBA…HBD complexes of the ratio 1:1 only. The solvent used was CCl_4, and the temperature was held at 25°C. Equation 1.4.1 shows how the K_f are calculated,

$$K_f / dm^3 mol^{-1} = \frac{[HBA..HBD]}{[HBA][HBD]} \qquad (1.4.1).$$

In equation 1.4.1, HBA…HBD represents the hydrogen bonded complex. The K_f's for a number of bases complexed against a series of (-OH) reference acids have been plotted against the K_f for the same reaction with 4-fluorophenol used consistently as the reference acid. In this way linear free energy relationships can be investigated using the LFER equation shown in equation 1.4.2,

$$\log K_f = m\left(\log K_f^{4\text{–}FC_6H_4OH}\right) + c \qquad (1.4.2).$$

In equation 1.4.2, K_f denotes the equilibrium constant for the formation of the complex of a HBA with one of the (-OH) reference acids, whereas $\left(K_f^{4\text{–}FC_6H_4OH}\right)$ is the equilibrium constant where 4-fluorophenol is used as the reference acid. It can be seen that equation 1.4.2 is in the general form of a straight line equation where *m* and *c* are the gradient and the intercept, respectively. The LFERs could then be used to set up the first scale to measure hydrogen bond basicity [63]. The scale uses the complexation of a HBA with 4-fluorophenol as the reference acid to measure hydrogen bond basicity. Using the LFER equation given in equation 1.4.2, it was possible to estimate $\left(K_f^{4\text{–}FC_6H_4OH}\right)$ where

the experimental value was unknown simply by solving the equation deriving *m* and *c* (see equation 1.4.4). Therefore the measure of hydrogen bond basicity was relative to complexation with 4-fluorophenol. This enabled a scale to be set up measuring the strengths of a series of hydrogen bond acceptors complexed to 4-fluorophenol. The scale was called the pK_{HB} scale and its mathematical meaning is shown in equations 1.4.3 and (1.4.4)

$$pK_{HB} = \log\left(K_f^{4-FC_6H_4OH}\right) \qquad (1.4.3)$$

$$pK_{HB} = \frac{\left(\log\left(K_f\right) - c\right)}{m} \qquad (1.4.4).$$

As mentioned above, equation 1.4.4 allows a pK_{HB} value to be obtained in the absence of primary data that would otherwise be used to calculate a pK_{HB} value from equation 1.4.3.

Family-independent relationships were observed for the complexation of bases with 4-fluorophenol. The *logK* values for a series of bases complexed against the various alcohols are linear. Therefore, some generality of the pK_{HB} scale had been observed. However, when the LFERs were repeated using 5-fluoroindole instead of 4-fluorophenol as the reference acid no family-independent behaviour was observed [65]. The advantage of the pK_{HB} scale is that it related to the Gibbs energy. The relationship between the pK_{HB} scale and Gibbs energy is shown in equation 1.4.5,

$$\log K_f = \frac{-\Delta G_{(T)}^{\circ}}{2.303RT} \qquad (1.4.5).$$

Although Taft's pK_{HB} scale has the advantage of being Gibbs energy related, little use has been made of it. One of the main criticisms of Taft's scale was that only one reference acid had been used; hence, the overall generality of the scale was never really

established. The pK_{HB} scale was later incorporated into the β scale [66]. The β scale was made up from a mixture of data including equilibrium constants and spectral shifts. However spectral shifts are subject to the Frank-Condon principle [67]. The spectral shifts are detected in less time than it takes the solvent to reorganise. Equilibrium constants are measured when both the enthalpic and entropic effects are balanced. At equilibrium the system has undergone entropic changes. Entropic contributions include the conformational entropy changes as fixed complexes are formed, and desolvation entropy as the layer of solvent surrounding binding sites is shed as complexes are formed. The contribution of desolvation entropy is not accounted for when data is taken from spectral shifts because the spectral shift is detected in less time than it takes for desolvation entropic effects to take place. Therefore data collected from spectral shifts are proposed to be enthalpy related [68]. As the β scale is made up from a mixture of enthalpy related spectral data and Gibbs energy related equilibrium constants, it is thought to be indefinable in thermodynamic terms [68].

A scale that would generally be considered more useful would have to combine the thermodynamic relevance of the Gibbs energy related pK_{HB} scale with the large number of reference acids used to construct the β scale. The pK_{HB} and β scales were later adapted by Abraham to set up the $logK_{\beta}$ scale [68]. The $logK_{\beta}$ scale was set up using a mixture of spectral data and equilibrium constants. The $logK_{\beta}$ scale was set up using a standard reference acid, 4-nitrophenol. There was still the need to establish a scale of general hydrogen bond basicity. In order to achieve this, the scale would have to be set up against a number of different acids. The $\log K_B^H$ scale was made using many reference acids [54]. This scale had advantages and disadvantages. The advantage of this scale is that it was set up by solving LFER equations for thirty-four acids, giving thirty-four $\log K_B^H$ values for each base. Equation 1.4.6 shows the LFER equations used to obtain the $\log K_B^H$ values,

$$\log K^{i}(A) = L_A \log K_B^{H_i} + D_A \qquad (1.4.6).$$

In equation 1.4.6 *i* refers to the base and A represents the acid. Therefore thirty-four equations for each base would be solved. For each base it was found that all thirty-four equations intersect at a point where *logK* = -1.1 [54]. This "magic point" where the equations intersect at *logK* = -1.1 allowed an equivalent but more convenient linear transform of the $\log K_B^H$ scale to be set up. This linear transform was called the β_2^H scale [54]. The β_2^H scale gives one value for each base as a measure of its hydrogen bond strength. Conveniently by adding the 1.1 *logK* units to the $\log K_B^H$ scale, and dividing by 4.636, the β_2^H scale ranges from 0 to 1. Therefore a β_2^H value of 0 represents no hydrogen bonding, and a β_2^H value of 1 represents the strongest hydrogen bonding. Equation 1.4.7 shows how Abrahams transformed the $\log K_B^H$ scale into the β_2^H scale,

$$\beta_2^H = \left(\log K_B^H + 1.1\right)/4.636 \qquad (1.4.7).$$

The β_2^H scale is Gibbs energy related as it is set up using formation constants. It is also related to Taft's pK_{BH} scale because 4-fluorophenol was used as one of the reference acids so any base with a pK_{BH} value will be included in the β_2^H scale. However, due to the statistical treatment of the thirty-four such $\log K_B^H$ values to obtain β_2^H values, the determination of secondary values not included in the original data set is not straightforward.

So far all the scales measuring hydrogen bond basicity that have been discussed here are for 1:1 acid:base ratios only. Many organic bases and likely drug molecules have multiple hydrogen bonding sites. These bases with multiple hydrogen bonding sites are known as polyfunctional hydrogen bond bases. As the β_2^H scale of hydrogen bond

basicity was considered the most general, it was adapted to account for polyfunctional hydrogen bond bases. The resulting scale was called the $\sum\beta_2^H$ scale [69]. The $\sum\beta_2^H$ scale was set up to account for the situation in which a base with multiple hydrogen bond acceptor sites is surrounded by an excess of hydrogen bond donors. In this case a ratio of 1(HBA) : n(HBD) is possible. The $\sum\beta_2^H$ scale can therefore be thought of as a measure of overall hydrogen bond basicity. The term $\sum\beta_2^H$ is seen in the general solvation equations stated in equations 1.3.1 and 1.3.2. It is from equations 1.3.1 and 1.3.2 that $\sum\beta_2^H$ values were calculated. By restricting the bases used as solutes to have a hydrogen bonding acidity value of zero, or monofunctional acids for which α_2^H can be used as an estimate of $\sum\alpha_2^H$, it is possible to determine $\sum\beta_2^H$ values from back calculation [69]. It should be mentioned here that α_2^H is a scale that measures hydrogen bond acidity [53] and is set up in the same way as the β_2^H scale. The $\sum\beta_2^H$ scale is often abbreviated to β_2^H and sometimes further abbreviated simply to β leading to confusion over which value is actually being reported. It is the $\sum\beta_2^H$ scale that is the most quoted hydrogen bond basicity scale for property based design in drug discovery [1, 70]. The $\sum\beta_2^H$ scale was set up to adapt the β_2^H scale to account for polyfunctional hydrogen bond bases. The β_2^H scale was designed for medicinal chemists. It was therefore claimed that the $\sum\beta_2^H$ scale of overall hydrogen bond basicity was more useful than the β_2^H scale when considering biological properties [69].

Despite being the most quoted scale of hydrogen bond basicity [1, 70], and despite the claims by its creators to be the most useful in the analysis of biochemical properties [69], there are problems in the interpretation of $\sum\beta_2^H$ values. The $\sum\beta_2^H$ scale assigns only one $\sum\beta_2^H$ value to the whole molecule. As the $\sum\beta_2^H$ scale is calculated mostly

from partition coefficients it is *logK* related. Equation 1.4.5 shows how *logK* values are related to Gibbs energy. However, equation 1.4.5 is based on the formation constant from a 1:1 complexation. The $\sum\beta_2^H$ scale is based on a single value for a 1(HBA):n(HBD) stoichiometry. It has therefore been argued that not only does the single $\sum\beta_2^H$ value give no useful information, but that it is also thermodynamically incorrect [1].

The thermodynamic problem of the $\sum\beta_2^H$ scale will be explained in the following text. Equations 1.4.8 and 1.4.9 show the formation constants for a 1(HBA):2(HBD) complexation, where C_1 and C_2 refer to the complexes formed at two different sites on the HBA.

$$K_1 = \frac{[C_1]}{[HBA][HBD]} \qquad (1.4.8)$$

$$K_2 = \frac{[C_2]}{[HBA][HBD]} \qquad (1.4.9).$$

Equations 1.4.8 and 1.4.9 represent a case where the concentration of HBD is in excess of HBA. Equations 1.4.8 and 1.4.9 show two separate formation constants for each complexation of acid at each basic site. However, the experimental methods used to determine the $\sum\beta_2^H$ scale are unable to determine $[C_1]$ and $[C_2]$ individually [1, 69]. Only the sum of $[C_1] + [C_2]$ and therefore the total equilibrium constant K_t are calculated as shown in equations 1.4.10 and 1.4.11,

$$K_t = \sum_{i=1}^{n} K_i \qquad (1.4.10)$$

$$K_t = \frac{[C_1]+[C_2]}{[HBA][HBD]} \qquad (1.4.11).$$

Equation 1.4.12 shows the relationship between Gibbs energy and formation constants.

$$\Delta G_i = -RT \ln K_i \qquad (1.4.12).$$

The term K_i in equation 1.4.12 cannot be identified with equation 1.4.10 as the logarithm of a sum is not equal to the sum of logarithms [1]. It follows that no thermodynamic data can be calculated from the $\sum \beta_2^H$ scale.

The scales of hydrogen bond basicity discussed so far are rarely used by medicinal chemists. It is clear that there is a need for a measure of hydrogen bond basicity that targets the needs of medicinal chemists. The requirements of this scale are that it should combine the simplicity of the pK_{HB} scale and its thermodynamic relevance with the analysis of polyfunctional bases as in the $\sum \beta_2^H$ scale. The pK_{BHX} database [1] seems to meet these requirements. The pK_{BHX} scale is set up similarly to the pK_{HB} scale in that it makes use of equilibrium constants of 1:1 complexes in the solvent CCl_4 where 4-fluorophenol is the reference acid. The pK_{BHX} values are therefore Gibbs energy related and have thermodynamic relevance. Equations 1.4.13, 1.4.14 and 1.4.15 show how pK_{BHX} values are obtained from equilibrium constants.

$$K_{BHX} = \frac{[HBA][HBD]}{[HBA\cdots HBX]} \qquad (1.4.13)$$

$$pK_{BHX} = -\log_{10} K_{BHX} \qquad (1.4.14)$$

$$K_{BHX} = 10^{-pK_{BHX}} \qquad (1.4.15).$$

The pK_{BHX} scale also accounts for polyfunctional bases. An important difference that separates the pK_{BHX} scale from the pK_{HB} and the $\sum \beta_2^H$ scales are the experimental methods used to calculate the equilibrium constants. The difference in experimental

methods allows polyfunctional bases to be included in the analysis. On one hand a Fourier transform infrared (FTIR) spectrometry methodology is used in the determination of the pK_{BHX} scale whereas on the other hand ^{19}F NMR, UV and dispersive IR techniques are mostly used for in the determination of pK_{HB} and $\sum\beta_2^H$ values [1]. The advantage of using FTIR is that multiple significant hydrogen bond acceptor sites of polyfunctional bases can be analysed [1]. Analysis of FTIR spectra shows that there is a peak correlating to the wavenumber of the O-H bond on the hydrogen bond acceptor. The shift in these wavenumbers upon hydrogen bond formation is found to correlate with pK_{BHX} values. Strong correlations of the shift in wavenumber and pK_{BHX} have been established for 1:1 complexes of 4-fluorophenol and various families of HBAs [71-83]. Therefore values for the shift in wave number for polyfunctional bases may be translated into pK_{BHX} values by using the straight line equation for the correlation of pK_{BHX} and shift in wavenumber for each chemical family of HBA. This allows a pK_{BHX} value to be assigned to each HBA site. It is clear that pK_{BHX} values are able to analyse polyfunctional bases as the $\sum\beta_2^H$ scale can, but as each basic site is considered separately, pK_{BHX} values are thermodynamically significant whereas $\sum\beta_2^H$ values are not. Also the pK_{BHX} database contains 1338 pK_{BHX} values related to 1164 bases compared to the 117 pK_{HB} values and only 90 β_2^H values [1]. The large data set is advantageous to medicinal chemists as most drugs contain several hydrogen bond acceptor sites consisting of mainly O, N, S, F, Cl and C-pi. It is therefore a possibility that the pK_{BHX} scale could be used as a descriptor characterising hydrogen bonding from which bioisosteric replacements could be found. The obvious way in which pK_{BHX} values could be used to find bioisosteric replacements is to search for atomic acceptor sites with similar strengths. However, it is a general misconception that hydrogen bond acceptors may be viewed as atomic sites [1]. The pK_{BHX} database classifies hydrogen bond acceptor sites on three levels. The three levels are: atomic site, functional group, and subfunction. For example, acetone is the hydrogen bond acceptor, the atomic site of the hydrogen bond acceptor is the oxygen

atom, the oxygen is in a carbonyl functional group and the subfunctional group of the carbonyl oxygen is a ketone. In doing so, the pK_{BHX} database allows for a quantitative analysis of hydrogen bond basicity for atomic sites, functional groups and subfunctional groups. Previously, there had been a good qualitative understanding of hydrogen bonds in drug–target interactions, with a strong need for quantitative data [84]. The quantitative data could be used in identifying bioisosteric replacements for functional groups or subfunctional groups with similar hydrogen bond acceptor properties.

There are a great number of increasing *in silico* methods in use to find bioisosteric replacements [85]. The hydrogen bond basicity parameters in the general solvation equations are calculated from experimental data. As experimental values are used, the reliability of the hydrogen bond basicity terms could be reduced due to experimental error and missing data [86]. This problem could be overcome by using reliable *in silico* estimates of hydrogen bond basicity.

1.5 Computational Models To Predict Hydrogen Bond Basicity.

There have been a number of computational models used to predict hydrogen bond basicity. The most negative value of the electrostatic potential located on a chosen surface of constant electron density on the hydrogen bond donor atom has been used to obtain a family-dependent model for the hydrogen bond basicity parameter in the general solvation equation [87]. The hydrogen bond basicity parameters used in this study were the β values obtained from linear solvation energy relationships [88]. This model was then applied to correlate the minimum electrostatic potential with pK_{HB} values, giving a reasonable linear relationship [89].

Later, Platts [61], developed a model to predict hydrogen bond basicity based on density functional theory (DFT) calculations on a series of bases taken from Abraham's list of $\sum \beta_2^H$ values [69]. Theoretical properties were calculated for both the isolated base, and the base complexed with Hartree-Fock (HF). These theoretical properties were

then correlated with $\sum\beta_2^H$ values. None of the properties of the isolated base correlated particularly well with $\sum\beta_2^H$. The hydrogen bond donor was changed to hydrogen fluoride (HF). The model was improved when properties of the base–HF complex were used. The best correlations were found with the hydrogen bond energy, the change in the length of the HF bond upon complexation, the change in electron density at the bond critical point of the HF bond upon complexation, and the electron density at the hydrogen bond critical point. Refinements to this model were made to allow pK_{BH} values to be predicted by combining the minimum electrostatic potential with the sum of the energy density at the hydrogen bond critical point [90].

There are some disadvantages to the models correlating properties of the hydrogen bond critical point with $\sum\beta_2^H$ values and the refinement of this model combining the minimum electrostatic potential and properties of the hydrogen bond critical point to predict pK_{BH} values. The first disadvantage is present in both models. The disadvantage is that hydrogen bond complexes must be used rather than properties of the isolated base. It is computationally expensive to calculate properties of the hydrogen bonded complex. The second disadvantage is present in the model that uses a correlation between properties of the hydrogen bonded complex and $\sum\beta_2^H$ values. In this model a single property is correlated with a $\sum\beta_2^H$ value, which can be thought of as a measure of overall hydrogen bond basicity. Therefore a single property from a single hydrogen bond is correlated to an overall value of hydrogen bond basicity where multiple hydrogen bonds can contribute to the $\sum\beta_2^H$ value. A model has been suggested [91] to eliminate both of these disadvantages by correlating properties of the free base with β_2^H values. Using properties of the free base eliminates the need for expensive calculations of hydrogen bonded complex. Also correlating with β_2^H would seem more appropriate than correlating with $\sum\beta_2^H$ because β_2^H values are from 1:1

complexes, as described above. However the results are found to be less accurate than Platts' model [61] correlating properties of the hydrogen bonded complex with $\sum\beta_2^H$ values. It must also be noted that the results in [91] consider nitrogen bases in a separate model. It follows that two models are needed to describe a data set of hydrogen bond acceptors. Therefore there are two major issues concerning the use of [91] in drug design. The first being that the model does not predict basicity values of polyfunctional bases, and the second being that the model is not universal as nitrogen bases must be considered separately.

It would therefore seem that the most accurate models used to predict hydrogen bond basicity must analyse properties of the hydrogen bonded complex. Also by using β_2^H values the model is limited to the 1:1 hydrogen bonded complexes only. It is likely that in drug–target interactions, 1(HBA):n(HBD) complexes are likely and an improvement to the model could be made by correlating to a measure of hydrogen bond basicity that accounts for all major hydrogen bond acceptor sites.

A further refinement of the models used to predict hydrogen bond basicity could be suggested. The refined model must keep the accuracy of Platts' model by considering properties of the hydrogen bonded complex, take into account polyfunctional hydrogen bond acceptors and be applicable to an entire set of hydrogen bond acceptors. It has been explained above that the pK_{BHX} scale was set up to analyse the strength of polyfunctional bases. The advantage of using pK_{BHX} values over $\sum\beta_2^H$ values is that a pK_{BHX} value is listed for each hydrogen bond acceptor site. This allows for the correlation of a single value of a particular property of a specific hydrogen bond, with a single value relating to the basicity of that hydrogen bond.

The next section describes a refined model for the prediction of hydrogen bond basicity. The model aims to find correlations between properties of hydrogen bonded complexes and pK_{BHX} values.

1.6 The Hard Soft Acid Base Principle.

The hard soft acid base principle (HSAB) was introduced into chemistry to categorise intuitive observations of Lewis acids and Lewis bases [92-96]. A Lewis acid is classified as so due to its electron accepting capacity. A Lewis base is considered to be an electron donor. In a Lewis acid base reaction electron density is transferred from a Lewis base towards a Lewis acid. However the concept of Lewis acid base reactions has much wider implications. The generalised description of a Lewis acid base reaction in terms of electron density can be extended to almost all chemical reactions including hydrogen bond complex formation. During a hydrogen bond formation reaction, electron density is transferred from the hydrogen bond acceptor towards the hydrogen bond donor. Natural bond order analysis of hydrogen bond formation reveals that electrons are transferred from the nonbonding lone pairs of the hydrogen bond acceptor to the antibonding σ^* orbital of the H-X bond [46]. Therefore hydrogen bond donors can be considered as a special class of Lewis acid and the hydrogen bond acceptor can be thought of as a special class of Lewis base.

The HSBA principle is of interest when considering the relative strengths of hydrogen bond complexes. As described above, hydrogen bond complexes can be considered special cases of acid-base reactions. Therefore the relative strengths of hydrogen bond acceptors could be related to their chemical hardness values. Hydrogen bonds are non–covalent interactions that are mainly electrostatic in nature. The following text will describe how acid-base reactions may be predominately either electrostatic, or covalent in character depending on the hardness values of the reacting speies.

The definition of a Lewis acid base reaction can also be applied to the reactions of metal ions in aqueous solutions of fluoride or iodide ions. It is the observations of metal ions in aqueous solutions of fluoride or iodide ions where the HSAB principle story begins. Certain metal ions such as Hg^{2+} and Pt^{2+} were found to form stable complexes in aqueous solutions of iodide ions and weak or no complexes with fluoride ions, whereas other metal ions such as Mg^{2+} and Al^{3+} were found to display the opposite behaviour

forming stable complexes in aqueous solutions of fluoride ions and weak or no complexes in aqueous solutions of iodide ions [92, 97]. Further comparisons between mercury and magnesium can be made by observing them in their naturally occurring ores. Mercury occurs naturally as a sulphide ore whereas magnesium occurs naturally as a carbonate ore. In fact magnesium actually occurs as magnesium oxide. However magnesium oxide reacts with carbon dioxide in the air to form magnesium carbonate. Therefore the complexes that were compared in early observations were the oxides and the sulphides. It would be expected that the oxide is the most stable complex for both the magnesium and mercury ores. So why is mercury found as a sulphate ore? The answer to this question is hidden in the formation enthalpies of oxides and sulphides for both magnesium and mercury. Equation 1.6.1 reveals magnesium oxide to be 54 kcal / mol more stable than magnesium sulphide. However equation 1.6.2 shows that mercury oxide is only 1 kcal / mol more stable than mercury sulphide. Therefore mercury may naturally occur as a sulphide whereas magnesium may not.

$$\left(\Delta H^{\circ}, MgO\right)-\left(\Delta H^{\circ}, MgS\right)=54 kcal/mol \qquad (1.6.1)$$

$$\left(\Delta H^{\circ}, HgO\right)-\left(\Delta H^{\circ}, HgS\right)=1 kcal/mol \qquad (1.6.2)$$

From the early observations described above, a trend in the reactivity of bases was noticed. It was noticed that Lewis bases would behave a certain way depending on their electronegativity. Bases would form complexes more readily with certain metal ions depending on which end of the electronegativity scale they are on. A series of metal ions were screened for their reactivity with Lewis bases and those with similar behaviour were classified based on which end of the electronegativity the bases they formed the most stable complexes were on. Metal ions that formed complexes most readily with the most electronegative bases were classified as hard. Metal ions that formed the most stable complexes with the least electronegative bases were classified as soft. Bases were also classified accordingly. Hard bases were classified as the most electronegative and

soft bases were classified as the least electronegative. From the observations described above it could be said that hard acids would bond more favourably to hard bases and soft acids would bond more favourably with soft bases and thus the HSAB principle was formed.

The acids and bases in each category of the HSBA principle were found to have certain common properties. In general hard acids are small electron acceptors, of high charge and have outer electrons that are not easily polarised. Hard bases are electron acceptors at the highest end of the electronegativity scale, are not easily polarised and have outer electrons that are not easily oxidised. Soft acids generally have a lower charge and are considered to be polarisable electron acceptors. Soft bases are electron acceptors at the lower end of the electronegativity scale, are polarisable and have low energy empty orbitals that are easily oxidised.

The description of the HSBA principle outlined above shows how the early classification of hard and soft acids and bases has some similarities with hydrogen bond basicity scales. The first thing to note is that the early HSBA principle allowed for an empirical ordering of hardness. The ordering of hardness was based on formation enthalpies of a series of acids or bases with a reference reciprocal acid or base. Hydrogen bond basicity scales are also relative to a common reference acid. Therefore the classification of hardness requires two references to be used, one hard and one soft. The tested reactants are classified as hard or soft depending on which reference they form the most stable complex with. Reactants of similar strength are grouped together. Just like in hydrogen bond basicity scales the ordering of hardness could change when the references are changed. The choice of the reference acid can be justified for hydrogen bond basicity scales. The reference acid, usually 4-fluorophenol, although chosen for historical reasons gives a scale that can be applied to the majority of biological interactions. The ordering of basicity will remain unchanged for all OH donors as well as NH and NH^+ donors, groups commonly found in biological solvents and amion acid side chains. Although the ordering of basicity will remain unchanged for all OH, NH

and NH^+ hydrogen bond donors, it is possible that the actual values of basicity could change. Therefore there is no absolute value of hydrogen bond basicity. Each value is case specific to its chemical environment. This is a potential problem when attempting to classify hydrogen bond basicity. However, each biological interaction is specific to its chemical environment. Therefore it would be impossible to quantify absolute hydrogen bond basicity. The empirical ordering of hardness relative to references used is an example of fuzzy logic [98]. Fuzzy logic is a mathematical term applied to situations where there is not enough information to quantify a concept or categorise observations into definitive sets. Prior to the quantification of the HSAB principle Lewis acids and bases were classed as either hard or soft only. Lewis acids and bases were classed based on their charge, size and polarisabilities. As hardness is not a specific chemical interaction and can be thought of as a resistance to chemical change; quantification of absolute hardness is desirable.

The quantification of chemical hardness makes use of a branch of quantum chemistry called density functional theory (DFT). DFT uses the electron density $\rho(\mathbf{r})$, rather than the wave function to describe the chemical information of a system. DFT is comparable to Schrödinger's wave theory. DFT uses the electron density rather than the more usual wave function. The ground state energy of a system may be expressed as a functional of the electron density. This is far more practical as DFT calculations are less expensive computationally.[99, 100].

A chemical system contains a number of electrons N, the atomic nuclei, and potential energy. For the purpose of this brief introduction into DFT the potential energy can be thought of as the external potential $v(\mathbf{r})$, due to the positions of the nuclei. The Hamiltonian is the operator that acts on this system. The Hamiltonian acts on a system to determine the energy of a system in a particular state. The energetic state of the system depends on N and $v(\mathbf{r})$. Therefore the state of a system depends on the Hamiltonian. The energy of a system is also a functional of the electron density and must

also be a functional of *N* and *v*(**r**). Therefore the energy of a system can be expressed in terms of *N* and *v*(**r**) [101]

$$E = E\left[N, v(\mathbf{r})\right] \qquad (1.6.3).$$

A branch of DFT known as density functional reactivity theory or simply conceptual DFT is based on perturbations made to a system by altering *N* or *v*(**r**). Chemical hardness was quantified using conceptual DFT. The following text will explain how the gap between the chemical intuition of the HSBA principle, and the quantum chemistry of conceptual DFT was bridged.

Absolute hardness, or global hardness as it is also known was quantified by Parr and Pearson by using equation 1.6.4 [102], where η is global hardness and μ is chemical potential

$$2\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_v = \left(\frac{\partial \mu}{\partial N}\right)_v \qquad (1.6.4).$$

Global hardness is a property of a whole molecule. Global hardness is the sensitivity of a molecule's total energy to perturbations in its electronic population. As described above the energy of a system can be expressed in terms of *N* and *v*(**r**). Conceptual DFT makes use of either perturbations to *N* or *v*(**r**). It is clear from equation 1.6.4 that perturbations are made to *N*, the electronic population of the molecule. Therefore the external potential *v*(**r**) must be constant. A constant external potential is shown by the inclusion of *v* outside the bracket in equation 1.6.4. The external potential is altered by changing the positions of the atomic nuclei. This is an important understanding when computing hardness. Any change to the electronic population of a molecule must be done under the condition that the geometry of the molecule is frozen. Therefore under a fixed geometry the global hardness is the second derivative of the total energy with respect to the

change in electronic population. The chemical potential of a system is defined as the first derivative of the change in total energy with respect to electronic population. The chemical potential μ, is the Lagrange multiplier of DFT. The inverse of the chemical potential is defined as the electronegativity χ [103]. It follows that hardness is the derivative of chemical potential with respect to electronic population. Although equation 1.6.4 is quoted and there is a multiple of 2 before the hardness value, the 2 has been dropped from the equation in recent years for convenience.

Quantum chemical topology (QCT) is described in detail in section 1.7. In brief QCT uses the electron density to partition the atoms in a molecule into individual topological objects. Using quantum QCT as an atomic partitioning method, it is possible to have a fractional electronic population for a particular atom. However, global hardness deals with the molecule in its entirety. It is not possible to add or subtract fractions of electrons from a system. Only whole electrons may be added or subtracted from a molecule. Due to the perturbations to *N* being restricted to integral numbers, the function E(N) produces discontinuities [104, 105]. As electrons are gained or lost from a system, electrons are transferred from one type of electronic orbital to another [106] causing discontinuities in the function E(N). The transition of electrons between atomic orbitals gives rise to a cusp in the function E(N). Derivatives may not be evaluated at a cusp. Therefore in order to calculate hardness from conceptual DFT, equation 1.6.4 must be approximated. Equation 1.6.4 may be approximated by applying the method of finite differences. The method of finite differences will now be described.

Take a molecule with *N* electrons. If one electron is removed, the molecule has *N-1* electrons. If the molecule gains an electron, it has *N+1* electrons. If the energies of the molecules *N*, *N-1* and *N+1* are known, the function E(N) can be plotted. The mean of the slope from *N* to *N-1* is the negative of the ionisation potential *A*. The mean of the slope from *N+1* to *N* is the negative of the electron affinity [107]. As a strict evaluation of derivatives involving *I* and *A* are not possible, a finite difference approximation must be applied to define hardness within the context of DFT. As it is the neutral molecule *N*,

that is of interest, the derivative must be approximated at the point *N*. The method used to approximate the derivative at *N* is known as a centred difference approximation. The equation for the first derivative of E(N), approximated at *N* using the centred difference approximation, is shown below

$$\chi = -\mu = -\left(\frac{\partial E}{\partial N}\right)_v \qquad (1.6.5)$$

$$\chi \approx \frac{E(N+h) - E(N-h)}{2h} \qquad (1.6.6)$$

$$\frac{E(N+h) - E(N-h)}{2h} = \frac{I + A}{2} \qquad (1.6.7).$$

In equations 1.6.6 *h* is equal to the number of electrons gained or lost from the system. Therefore, when applying the formula for the centred difference to hardness in the context of DFT, *h* is always equal to 1.

The equations above relate to the first derivative of E(N). The first derivative gives a quantification of electronegativity and chemical potential as described above. This is an example of a first order response function of DFT. Chemical hardness is defined as the second derivative of E(N). The quantification of chemical hardness is an example of a second order response function of DFT. The following equations shows how the second derivative of E(N), equation 1.6.4, may be approximated using a centred finite difference approach

$$\eta \approx \frac{E(N+h) - 2E(N) + E(N-h)}{2h^2} \qquad (1.6.8)$$

$$\frac{E(N+h)-2E(N)+E(N-h)}{2h^2}=\frac{I-A}{2} \qquad (1.6.9).$$

Until the introduction of equation 1.6.4, the HSAB principle was an example of fuzzy logic [98]. Equation 1.6.4 successfully bridged the gap between observational chemistry and quantum theory. The validity of equation 1.6.4 was strengthened as equation 1.6.4 recovers all observed trends known about the hard soft acid base principle [92, 102].

Chemical hardness is, in general terms, the resistance of a molecule to chemical change or reaction. Therefore chemical hardness can also be thought of as an indicator of chemical stability [94]. The equations stated above give a greater insight into the physical meaning of hardness [108]. Stability in the context of chemical hardness can be more rigorously defined as a resistance to a change in a molecule's electronic population. The harder a molecule is, the more energy it will take to change its electronic population by a defined amount.

Chemical hardness is of relevance when investigating hydrogen bond basicity. The formation of a hydrogen bond can be thought of as a type of reaction. During the formation of a hydrogen bond electron density is transferred from the hydrogen bond acceptor towards the hydrogen bond donor. Therefore the more resistant the hydrogen bond base is to loss of electron density the harder it is. It is well known that hard acids will react with hard bases and soft acids with soft bases. Based on the theory of the hard soft acid base principle, it is therefore likely that the most stable hydrogen bond complexes will be formed by acids and bases with similar hardness values. This poses a potential issue when developing basicity scales. The basicity scale is self-contained and not applicable to a general hydrogen bond situation. The ordering and basicity values will change when the reference hydrogen bond donor is changed. Most hydrogen bond basicity scales take either 4-fluorophenol or methanol to be the hydrogen bond donor. Based on the HSAB principle it can be assumed that the most stable hydrogen bond

complexes and therefore the most basic hydrogen bond acceptors will be those with hardness values closest to either 4-fluoropheol or methanol. In general it is the difference in electronegativities that drives electron density from the hydrogen bond acceptor towards the hydrogen bond donor, and chemical hardness that inhibits it. The inhibitory role hardness plays in the progression of a reaction is shown in equation 1.6.10

$$\Delta N = \frac{\left(\chi_Y^\circ - \chi_Z^\circ\right)}{2\left(\eta_Y + \eta_Z\right)} \qquad (1.6.10).$$

Δ*N* is the fractional number of electrons transferred from isolated reactant Z to isolated reactant Y as Z and Y come together from infinite distance of separation to form product YZ. The initial electronegativities of the reactants are χ°. As the product YZ is formed the electronegativities of Y and Z will change. One will increase as the other decreases until the reaction reaches equilibrium and the electronegativities of Y and Z are the same. This is an example of the electronegativity equalisation principle [109, 110].

The hardness and electronegativity values in equation 1.6.10 are for Y and Z in their entirety and not for specific atoms or functional sites within Y and Z. There is no known method for calculating either the hardness or electronegativity for atoms in molecules. However there is a method for estimating the reactivity for an atom in a molecule. The Fukui function [111, 112] is a second order response function of DFT and is widely used as a reactivity indicator [99, 101].

The Fukui function arises from the expansion of $E=E[N,v(\boldsymbol{r})]$ up to second order [101]. The second order expansion of $E=E[N,v(\boldsymbol{r})]$ yields a mixed derivative shown by equation 1.6.11. The mixed derivative arises from perturbations in *N and v(**r**)* that are caused by an attacking agent. The mixed derivative is known as the Fukui function *f(**r**)*

$$f(\mathbf{r}) = \left(\frac{\partial}{\partial N} \left(\frac{\delta E}{\delta v(\mathbf{r})} \right)_N \right)_V = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_v = \left(\frac{\delta \mu}{\delta v(\mathbf{r})} \right)_N \qquad (1.6.11).$$

By examining equation 1.6.11, it can be seen that the Fukui function can be described as either the change in electron density at a point corresponding to a change in electron number at a fixed external potential, or as the change in chemical potential with respect to a change in external potential at a point where the electron number is fixed. It is the Fukui function which represents the change in electron density at a point caused by perturbations to the electron number at a fixed external potential that is commonly used as a reactivity indicator and is therefore of interest in this research. The Fukui function of interest, like the hardness equation, is sensitive to change in electron number. Only whole electrons may be introduced into, or taken away from the system. The Fukui function is another example of a function where the variable is the electron number. Functions that have the electron number as the variable are discontinuous, as described above. Therefore the Fukui function must be approximated using the method of finite differences. The centred difference used to calculate chemical hardness is used to approximate the derivative at the point *N* on a plot with points *N*, *N+1* and *N-1*. The derivative may also be evaluated at points *N+1* and *N-1*. The derivative at *N+1* is known as the right derivative and the derivative at the point *N-1* is known as the left derivative. The left derivative of ρ(**r**)[N] measures susceptibility to electrophilic attack or electron loss and is indicated by the symbol $f(\mathbf{r})^-$. The right derivative of ρ(**r**)[N] measures susceptibility to nucleophilic attack or gain of electrons and is indicated by $f(\mathbf{r})^+$. It follows that the centred difference evaluated at *N* measures susceptibility to radical attack and the flow of electrons both into and away from the system. The centred difference is indicated by $f(\mathbf{r})^0$. Equations 1.6.12 – 1.6.14 show how each Fukui function is approximated using the method of finite differences

$$f(\mathbf{r})^- = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_v^- \cong \rho_N(\mathbf{r}) - \rho_{N-1}(\mathbf{r}) \qquad (1.6.12)$$

$$f(\mathbf{r})^{+} = \left(\frac{\partial\rho(\mathbf{r})}{\partial N}\right)_{v}^{+} \cong \rho_{N+1}(\mathbf{r}) - \rho_{N}(\mathbf{r}) \qquad (1.6.13)$$

$$f(\mathbf{r})^{0} = \left(\frac{\partial\rho(\mathbf{r})}{\partial N}\right)_{v}^{0} \cong \frac{\rho_{N+1}(\mathbf{r}) - \rho_{N-1}(\mathbf{r})}{2} \qquad (1.6.14).$$

The Fukui function of interest measures the sensitivity of the electron density to perturbations in the number of electrons in the system. The electron density is a scalar value that is evaluated a point in 3 dimensional space somewhere in the system. It follows that the Fukui function must too give rise to a scalar field surrounding the system. The Fukui function is evaluated at specific points and can be plotted in much the same way as the electron density. Given that the Fukui function can be plotted as a scalar field in 3-dimensional space, it follows that

$$\int f(\mathbf{r})\mathbf{dr} = 1 \qquad (1.6.15).$$

The Fukui function is a reactivity indicator. The most reactive point on a molecule is the one with the largest Fukui function. The type of reactivity depends on which of the 3 Fukui functions is used.

Local hardness may be evaluated for a system. The difference between global hardness and local hardness is that global hardness is a property of a whole molecule whereas local hardness is evaluated at a specific point. The point with the largest value of local hardness is the point in the system that is most resistance to change in electron density. The evaluation of local hardness again relies on being able to evaluate the electron density at specific points. Equation 1.6.16 is the equation for local hardness. It can be seen that the electron number in equation 1.6.4 has simply been replaced by electron density, allowing for a scalar field of local hardness to be plotted around the system

$$\eta(\mathbf{r}) = \left(\frac{\delta\mu}{\delta\rho(\mathbf{r})} \right)_v \qquad (1.6.16).$$

The Fukui functions may be used to derive local softness σ(**r**). Whereas local hardness is the point in a system most resistant to changes in electron density, local softness finds the point in the molecule most susceptible to changes in electron density. Therefore the point in the molecule with the largest local softness value is not necessarily the most reactive but the most likely to undergo chemical change. Chemical change may not necessarily be the breaking and forming of bonds as in a chemical reaction, but a change in shape and structure due to eaily polarisable electron density. Equation 1.6.17 shows how local softness, $\sigma(\boldsymbol{r})$ is calculated. *S* refers to global softness, the inverse of global hardness

$$\sigma(\mathbf{r}) = \left(\frac{\partial\rho(\mathbf{r})}{\partial N} \right)_v \times \left(\frac{\partial N}{\partial\mu} \right)_v = f \times S \qquad (1.6.17).$$

The relationship between global hardness and local hardness is given in equation 1.6.18

$$\eta = \int \eta(\mathbf{r}) f(\mathbf{r}) \qquad (1.6.18).$$

Unfortunately there is no known method to calculate atomic hardness. However it is possible to condense local hardness into atomic form. Local hardness can be condensed into atomic form by integrating equation 1.6.18 over topological atomic basins, *Ω* [113] as described by the quantum theory of atoms in molecules (see section 1.8) [114, 115] rather than integrating over the entire molecule. It can be seen from equation 1.6.18 that condensing local hardness into atomic form would be a lengthy calculation. In the context of hydrogen bonding it is perhaps more relevant to measure reactivity than susceptibility to change in electronic population. It is therefore the Fukui function that is of interest. Fortunately each Fukui function can be condensed into atomic form [116-118]. Once again, to obtain a condensed Fukui function integration over an atomic basin

Ω is required. Condensed Fukui functions for atom *A* are given in equations 1.6.19 to 1.6.21

$$f_A^- \cong \int_{\Omega_A} \rho(\mathbf{r})d\mathbf{r} - \int_{\Omega_A^-} \rho^-(\mathbf{r})d\mathbf{r} = N_A - N_A^- \qquad (1.6.19)$$

$$f_A^+ \cong \int_{\Omega_A^+} \rho^+(\mathbf{r})d\mathbf{r} - \int_{\Omega_A} \rho(\mathbf{r})d\mathbf{r} = N_A^+ - N_A \qquad (1.6.20)$$

$$f_A^\circ \cong \left(\frac{\int_{\Omega_A^+} \rho^+(\mathbf{r})d\mathbf{r} - \int_{\Omega_A^-} \rho^-(\mathbf{r})d\mathbf{r}}{2} \right) = \frac{N_A^+ - N_A^-}{2} \qquad (1.6.21).$$

It must be noted that the signs (+) and (-) in equations 1.6.19 – 1.6.21 are not to be confused with charges. The sign (+) implies an electron has been added to the system whereas the sign (-) indicates an electron has left the system. Equations 1.6.19 to 1.6.21 can be used to calculate atomic softness. Atomic softness is calculated by replacing the Fukui function in equation 1.6.17 with one of the condensed Fukui functions.

The above text has introduced the ideas of local hardness and local reactivity. The following text will discuss the application and interpretation of local hardness and local reactivity.

The text introducing the HSAB has described how a soft acid will preferentially form a complex with a soft base. The interaction between a soft acid and a soft base will be predominantly covalent. It has been shown that the site with the largest value of local softness is the same as the site with the largest Fukui function. It can be therefore concluded that the softest local site is also the most reactive. Intuitively, it would be expected that the interaction of the two soft species occurs through their softest sites.

However, the interaction of two soft species occurs not through the sites of maximum local softness, but through sites where the Fukui functions are approximately equal.

The HSAB principle described above also states that hard acids form complexes preferentially with hard bases, interactions that tend to be ionic in nature. The site of the hard interaction has been linked to the point of the lowest Fukui function. The ionic hard interactions are dominated by electrostatic effects while covalent factors are less important. Soft interactions are predominantly covalent. Covalent bonding is dictated by the shapes of the frontier orbitals. Interestingly the Fukui function condenses to the frontier molecular orbitals. Soft interactions are found at the point of the largest Fukui function, a finding that therefore highlights the covalent nature of soft interactions. Following the same logic, the low value of the Fukui function at the site of hard interactions demonstrates their lack of covalent character. The low value of the Fukui function for hard interactions can be misinterpreted because low values of the Fukui function mean the site is not very reactive. Clearly for a hard-hard complex to form a reaction is occurring. The Fukui function cannot account for electrostatic effects. As electrostatic effects dominate hard-hard interactions perhaps the Fukui function is not the best reactivity indicator to use when describing hydrogen bonding.

The Fukui function is also reaction specific. Three Fukui functions are needed to describe either electron loss, electron gain or radical attack. The hard soft acid base principle states that hardness inhibits the progression of a reaction. Equations 1.6.17 and 1.6.18 show how a Fukui function is needed to describe local hardness and softness. Therefore local hardness and softness are reaction specific. Results concerning local hardness and softness have produced some confusing results. In the case of electrophilic attack on benzocyclobutadiene and water, the hardest and softest local sites were both found to be located on the oxygen atom [113]. Intuitively, the oxygen atom would be expected to be hard. The finding that the oxygen atom is the softest local site is explained by the local abundance of global softness being largest for the oxygen atom,

and not that the oxygen atom is necessarily soft. This is mathematically correct as local softness integrates to global softness. However, local hardness does not integrate to global hardness. The equation for local softness contains a Fukui function. Therefore local softness appears to be a reaction specific indicator of reactivity whereas local hardness appears to be a general indicator of resistance to chemical change.

In the context of hydrogen bond basicity the hard soft acid base principle gives useful insight into the interpretation of results. The strongest complexes are expected to be formed when the hardness values of the hydrogen bond donor and the hydrogen bond acceptor are similar. Therefore if a hydrogen bond basicity scale uses methanol as the reference hydrogen bond donor, the scale will only be applicable to hydrogen bond donors with similar hardness values to methanol. Furthermore, a Fukui function to describe electrophilic attack could be useful to explain hydrogen bond basicity. It is not the most reactive site that would be expected to form the strongest hydrogen bonds, but the one which has a similar Fukui function value to the hydrogen bond donor. The HSAB principle again highlights how hydrogen bond basicity cannot be thought of as a general descriptor but as a descriptor specific to the donor and acceptor used. It would be expected that the strongest hydrogen bond complexes would be those where the acid and base have similar hardness values and those where the local softness values (for electrophilic attack) are also similar.

1.7 The pK_a Slide Rule.

The final factor to consider in acid base reactions is the acid dissociation constant pK_a. The acid dissociation constant can be calculated as follows when measurements are made in dilute aqueous solution.

$$AH + H_2O \leftrightarrow H_3O^+ + A^- \qquad (1.7.1).$$

The acid AH may also act as a hydrogen bond donor. The dissociation constant for the conjugate base pK_{BH+} can be calculated as follows when measurements are made in dilute aqueous solution.

$$BH^+ + H_2O \leftrightarrow H_3O^+ + B \qquad (1.7.2).$$

In equation 1.7.2 B may be a hydrogen bond acceptor as well as a proton acceptor. A hydrogen bond complex may be formed between AH and B to give AH•••B. However, proton transfer may occur in the reaction to give A^-•••H^+B. To calculate whether or not a proton transfer has taken place it is essential to know the acid dissociation constants for the hydrogen bond donors, and the dissociation constant for the conjugate base of the hydrogen bond acceptor. The difference between pK_a and pK_{BH+} indicates whether or not proton transfer has taken place. This is known as the pK_a difference. The equation to calculate pK_a difference is shown below.

$$\Delta pK_a = pK_a - pK_{BH+} \qquad (1.7.3).$$

If equation 1.7.3 returns a negative value, then proton transfer has taken place and the hydrogen bond complex will take the form of A^-•••H^+B. This type of hydrogen bond is a double charge assisted hydrogen bond and is stronger than an ordinary hydrogen bond. An ordinary hydrogen bond AH•••B does not undergo a proton transfer. Therefore the pK_a difference is positive. The strength of a hydrogen bond increases as the pK_a difference approaches zero. In general hydrogen bonds weaken as the pK_a difference increases above zero. As the pK_a difference increases above zero, the hydrogen bond complexes become increasingly less attractive as the AH is more likely to become dissociated than B is likely to pick up the dissociated proton. As the pK_a difference decreases below zero the conjugate acid has an increasingly greater attraction to its acquired proton than the conjugate base and therefore A^- becomes increasingly less attracted to H^+B. The chemistry described above for calculating hydrogen bond strength based on pK_a difference is collectively known as the pK_a slide rule. The pK_a slide rule

states that the strongest hydrogen bonds are formed as the pK_a difference approaches zero. The pK_a slide rule covers an impressive range of $-14 \leq pK_a \leq 53$ for acids and $-12 \leq pK_{BH+} \leq 16$ for bases.

The pK_a slide rule [119] is of relevance in the field of hydrogen bond basicity scales. Like the equalisation of reactivity indicators and hardness values taken from the hard soft acid base principle, the pK_a slide rule states that pK_a equalisation produces the strongest hydrogen bonds. Based on the pK_a slide rule it can be suggested that any hydrogen bond basicity scale can only be extrapolated for hydrogen bond donors with similar pK_a values to the reference hydrogen bond donor.

The strength of a base as a proton acceptor, or its Brönsted basicity is defined by pK_{BH}^+. pK_{BH}^+ values are obtained from equations 1.7.4, 1.7.5 and 1.7.6

$$B + H^+ \leftrightarrow BH^+ \qquad (3.7.4)$$

$$K_{BH^+} = \frac{[B][H^+]}{[BH^+]} \qquad (3.7.5)$$

$$pK_{BH^+} = -\log_{10} K_{BH^+} \qquad (3.7.6).$$

Equations 3.7.4 – 3.7.6 therefore describe the strength of a neutral base. It can be seen that pK_{BH}^+ can be thought of as the K_a of a conjugate acid. It follows that the strongest bases have the largest pK_{BH}^+ values and are the most likely to be protonated in solution. As the pK_{BH}^+ scale is based on equilibrium constants of acid base reactions, pK_{BH}^+ is a form of Brönsted basicity.

All hydrogen bonds are unique and it has proven difficult to make a general hydrogen bond basicity scale that accounts for the basicities of hydrogen bond acceptors independently of the hydrogen bond donor. Abraham [54] attempted to set up a general hydrogen bond basicity scale with 34 different reference hydrogen bond donors. The scale is described in detail in the previous section. The problem with the $logK_B^H$ scale is

that it is not possible to extrapolate data for bases outside the original data set without complicated mathematics. This renders the $logK_B^H$ scale unhelpful for medicinal chemists who require information on novel molecules.

Using the equalisation methods learned from the hard soft acid base principle and the pK_a slide rule, it might be possible to take a hydrogen bond basicity scale and use a scaling factor to determine basicity values for hydrogen bond complexes where the hydrogen bond donor has values of hardness, and or pK_a considerably different from the reference hydrogen bond donor.

1.8 The Quantum Theory Of Atoms In Molecules

This section will describe the basic principles of the quantum theory of atoms in molecules (QTAIM). QTAIM features heavily in this research. QTAIM is used to obtain the properties that will be used in this research in an attempt to predict hydrogen bond basicity values. The following section will describe how such properties are obtained and how they will be used throughout this research.

QTAIM essentially provides a method in which atoms in molecules can be partitioned into topologically bound objects that are representative of the molecular structure hypothesis. QTAIM was pioneered by Bader and co-workers [114, 115, 120]. Although QTAIM is a purely theoretical entity it has been found to recover all aspects of experimental chemistry [121] including pK_a prediction [122, 123], heats of formation [124], the stability of proteins [125, 126] the mechanical properties of metallic alloys [127] and even the brittleness of materials [128].

QTAIM makes utilises the electron density which can be obtained from the Schrödinger equation [129]. The Schrödinger equation is an eigenvalue equation. The Hamiltonian $\hat{H}$, is the quantum mechanical operator which contains all the operations relating to the kinetic and potential energies of the particles in a system. Solutions to the

time independent Schrödinger equation exist for only certain values of energy. The values of energy are called the eigenvalues. There is one eigenfunction that corresponds to each eigenvalue. The eigenfunction is the wave function. Therefore, the solution to the Schrödinger equation for a given energy value E_i is the wave function Ψ_i, and the Hamiltonian is the operator that acts on the eigenfunction Ψ_i to determine the eigenvalues E_i. The time independent or stationary state Schrödinger equation for a system in state *i* is given in equation 1.8.1

$$\hat{H}\Psi_i = E_i\Psi_i \qquad (1.8.1).$$

The wave function can be thought of as a crudely quantum analogue to classical Hamiltonian physics in which the state of a particle at any instance in time can be described by two vectors, **r** for position and **p** for momentum. The single function Ψ(**r**) replaces the vectors **r** and **p** when dealing with matter with a wave like form such as electrons. According to Heisenberg's uncertainty principle the exact position and momentum of an electron cannot be precisely determined simultaneously. All the information known about a system is contained within Ψ(**r**). Unfortunately Ψ(**r**) cannot be directly observed. Unlike the vectors **r** and **p** in classical physics, the function Ψ(**r**) contains a complex number. The way in which the wave function is handled to eliminate the complex number is to multiply it by its complex conjugate Ψ*, thus giving Ψ*Ψ or $|\Psi|^2$, the square of the modulus of Ψ. As stated above, the position and momentum of an electron cannot be precisely determined simultaneously. According to Born, the entire charge of an electron can however be found in a position between **r** and **r** + d**r** with a probability proportional to $|\Psi|^2$ as shown by equation 1.8.2. Given Born's statement, electrons appear to be thought of as point charges with wave-like form, hence exhibiting wave-particle duality as a probability density, $P(\boldsymbol{r})$

$$P(\mathbf{r}) = \Psi(\mathbf{r})\Psi^*(\mathbf{r})d\mathbf{r} \qquad (1.8.2).$$

According to the Pauli exclusion principle no two electrons may occupy the same quantum state at the same time. The first three quantum numbers relate to the spatial portion of the quantum state of an electron. The first three quantum numbers may be the same for two electrons. However to obey the Pauli exclusion principle the fourth quantum number must be different. The fourth quantum number relates to the spin of the electrons. Two electrons occupying the same spatial quantum state must have opposing spins. Therefore if two electrons are in quantum states differing by only spin *a* or spin *b* it might be thought that the total wave function is as follows

$$\Psi = \Psi_1(\mathbf{r}, a)\Psi_2(\mathbf{r}, b) \qquad (1.8.3).$$

According to Born's interpretation of $|\Psi|^2$ the wave function shown above in equation 1.8.3 gives the probability of finding electron 1 in space 1 with spin a and electron 2 in space 2 with spin b. This cannot be true since all electrons are both identical and indistinguishable. Equation 1.8.3 distinguishes between electrons 1 and 2 by spin. To account for this a linear combination of the possibilities of finding electrons 1 and 2 in spins a and b must be used. This is known as anti-symmetrising a wave function and is shown in equation 1.8.4

$$\Psi = \Psi_1(\mathbf{r}, a)\Psi_2(\mathbf{r}, b) - \Psi_1(\mathbf{r}, b)\Psi_2(\mathbf{r}, a) \qquad (1.8.4).$$

The meaning of $|\Psi|^2$ in equation 1.8.4 is that both spin states a and b are occupied by either electron 1 or electron 2.

It is now clear that the wave function lies in 4 dimensional space. Therefore a system with *N* electrons has a wave function with *4N* coordinates. The 3 spatial coordinates **r**, along with the spin coordinate a or b can be written as the vector **q** for each electron. The total set of *4N* coordinates describing the electrons of a system can

be written as **Q**. Equation 1.8.2 therefore becomes equation 1.8.5 for a many electronic system

$$P(\mathbf{q}) = \Psi(\mathbf{Q})\Psi^*(\mathbf{Q})d\mathbf{q}_1 d\mathbf{q}_2 \cdots d\mathbf{q}_N \qquad (1.8.5).$$

The meaning of equation 1.8.5 is the probability of finding the system in a certain configuration in which each electron is in a particular position with a particular spin.

The wave function described by equation 1.8.5 lies in 4 dimensional space and cannot be visualised. In order to readily visualise the 4 dimensional wave function it must be converted into a 3 dimensional function. This is done by eliminating the spin coordinates. The spin coordinates are eliminated by summing equation 1.8.5 over all spins of all electrons. Along with summing over all spins of all electrons, equation 1.8.5 is integrated over all spatial coordinates of all electrons except one. Which electrons spatial coordinates are not integrated over does not matter because all electrons are identical and indistinguishable as shown by equation 1.8.4. By integrating over all the spatial coordinates of all electrons except electron 1 for example, the probability of finding electron 1 in the volume element bound by $d\tau_1$ is obtained. The probability of finding electron 1 in the volume element bound by $d\tau_1$ is independent of the positions and spins of all other electrons in the system. The resulting probability per unit volume of finding one electron is known as the probability density P(**r**) and is shown in equation 1.8.6. The probability density P(**r**) ranges between 0 and 1.

$$P(\mathbf{r}) = \sum_{Spins} \left| \int d\tau_2 \int d\tau_3 \cdots \int d\tau_N \Psi^*(\mathbf{Q})\Psi(\mathbf{Q}) \right| d\tau_1 \qquad (1.8.6).$$

For an *N* electronic system the probability density P(**r**) is converted into the electron density ρ(**r**) by multiplying equation 1.8.6 by *N*, the number of electrons.

Equation 1.7.8 shows how the electron density is calculated. In equation 1.8.8, ∫dτ′ indicates that integration is over the coordinates of all but one electrons

$$\rho(\mathbf{r}) = N\int d\tau \Psi^{*}(\mathbf{Q})\Psi(\mathbf{Q}) \qquad (1.8.8).$$

It is the electron density ρ(**r**) that is used to partition atoms in molecules into topologically bound objects. The unit of electron density is e / Å^3. Electron density will, however, be represented in atomic units, au in this document.

The electron density ρ(**r**) maps to a number at an infinitesimally small volume of space. The electron density can therefore be evaluated at any point in space and therefore forms a scalar field. Figure 1.8.1 shows a 2 dimensional representation of the electron density taken from a cross section of a water molecule. Regions of pre determined constant electron density are linked by contour lines. The nucleus of the oxygen atom is taken as the origin point in space. The x axis is formed between the oxygen nucleus and the nucleus of one of the hydrogen atoms. The cross section is fixed in the plane of the oxygen nucleus and the other hydrogen nucleus.

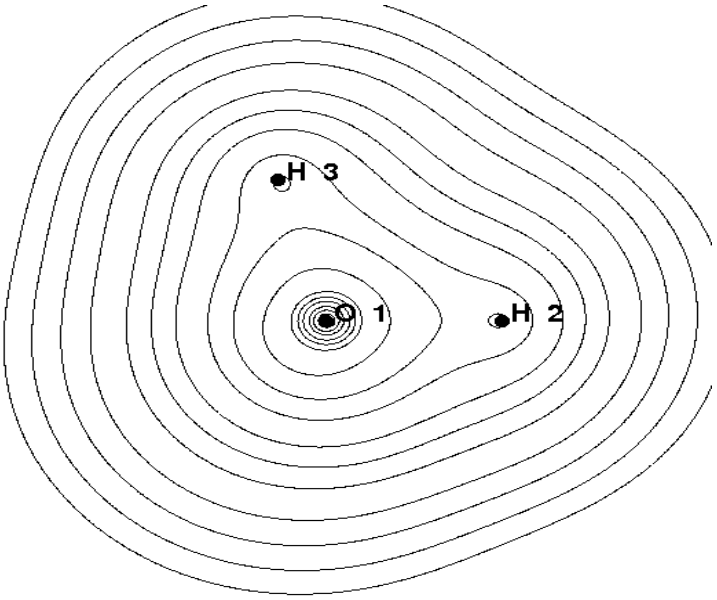


Figure 1.8.1. A contour map of the electron density of a water molecule. The contour lines join regions of constant electron density. The outer line represents an electron density of 0.001 au. The

electron density increases following the pattern 2 x 10^n, 4 x 10^n, 8 x 10^n where *n* increases by integers from -3 to 2.

The electron density approaches 0 as the distance from the nucleus approaches infinity. This means that the volume of an isolated atom is infinite. The electron density does however rapidly approach 0. Therefore a practical cut off point must be chosen in order to give atoms finite boundaries. Between 98% and 99% of the electron density is enclosed by the volume bound by 0.001 au. Throughout this research the outer surface of an atom will be bound by the region of electron density equal to 10^{-6} au.

The outer surfaces of atoms are bound by a practically chosen region of constant electron density. However, atoms in molecules may be bound by other atoms. QTAIM separates atoms in molecules into individual topological objects. So how are the boundaries between atoms formed? To answer this the gradient vector field of the electron density must be analysed. The gradient vector field of the electron density traces the gradient of the electron density from infinity along the path of steepest ascent towards the nucleus. In other words the gradient of the electron density can be thought of as the derivative of the electron density with respect to its position. Therefore gradient vectors are always orthogonal to contours. The vectors also form a set of discrete functions that never overlap. The gradient vectors approach the nucleus where they terminate. The gradient of the electron density cannot be evaluated at the nucleus. The electron density approaches infinity at the nucleus and therefore no derivative can be evaluated. The change in electron density reaches a cusp at the nucleus because the electron density approaches infinity. Figure 1.8.2 shows the gradient vector field of the electron density of the water molecule displayed in figure 1.8.1. Figure 1.8.3 is a combination of figure 1.8.1 and figure 1.8.2 where the gradient vector field is superimposed onto the contour map of the electron density.

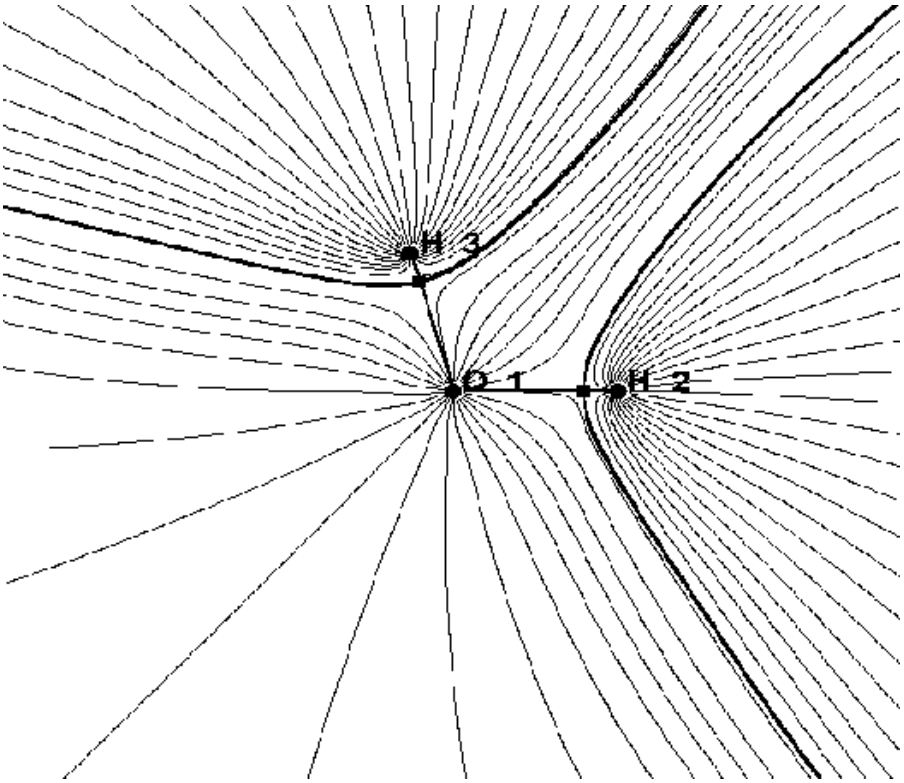


Figure 1.8.2. The gradient vector field of a water molecule. The lines represent gradient vector paths. The lines follow the gradient of the electron density along the path of steepest ascent from 0.001 au to the nucleus where they terminate.

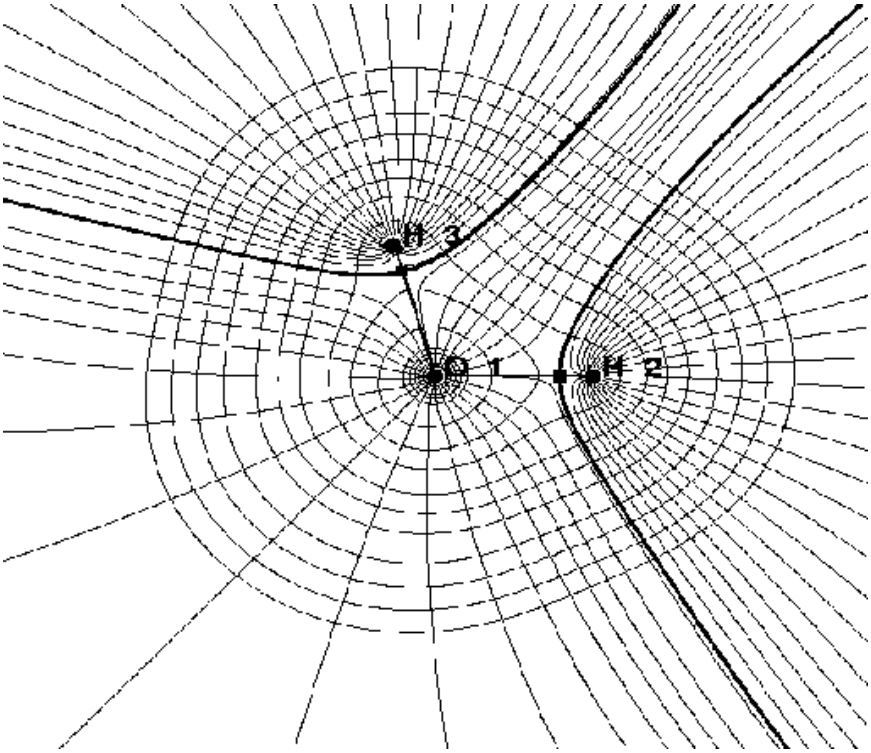


Figure 1.8.3. The gradient vector field the electron density superimposed on a contour map of the electron density. It can be seen that as the vectors trace paths in the direction of steepest ascent, they are always orthogonal to the contours they dissect.

At the nucleus the gradient of the electron density vanishes. A point at which the gradient of the electron density vanishes is known as a critical point. The nucleus, or more precisely a point which almost coincides with the nucleus as the gradient cannot be evaluated at the nucleus is a type of critical point. It can be seen from figure 1.8.2 and

figure 1.8.3 that there is a gradient path that does not terminate at a nucleus. This path is highlighted in bold. The path terminates at a point between two nuclei. This point is another type of critical point and is highlighted by a square point in figure 1.8.2 and figure 1.8.3. In fact two gradient vectors terminate at this point. The other path comes from the opposite direction and is present due to symmetry. The two paths do not intersect, they terminate at the critical point. The gradient of the electron density vanishes at this critical point. The critical point is a maximum in the electron density along the two gradient paths that terminate there and will be known from now on as the bond critical point (BCP). In fact, there are three other types of critical point found when analysing the full topology of atoms in molecules. Along with the BCP, there is also the ring critical point, cage critical point and non nuclear attractor. Each type of critical point is classified according to their eigenvalues of the Hessian matrix $\nabla\nabla\rho$. The number of non-zero eigenvalues donater the rank *r*, of the critical point. The signs of the eigenvalues are donated by the signature *s*. A negative eigenvalue is assigned a value of -1 and a positive eigenvalue is assigned a value of +1. The sum of the values of the eigenvalues equals *s*. Critical points are classified according to (*r*, *s*). Therefore a critical point with three non-zero eigenvalues, two of which are negative and one of which is positive is a (3, -1) critical point. The BCP is infact a (3, -1) critical point.

The two vectors that terminate at the BCP indicate the boundary between two atoms. The two paths form a special type of surface that partitions two atoms and is known as the interatomic surface. The atoms are not partitioned arbitrarily. The interatomic surface is a natural surface that is governed by the electron density gradient vector field. The interatomic surface differs from any arbitrary surface. The way in which the interatomic surface differs from any arbitrary surface is the way in which the surface is orientated in respect to the electron density gradient vector field. The interatomic surface is orientated in such a way that in all places the normal vector to the interatomic surface is perpendicular to the electron density gradient vector. One property of perpendicular vectors is that their dot product is 0. Equation 1.8.9 shows this

$$\mathbf{a}\cdot\mathbf{b}=\|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta \qquad (1.8.9)$$

where θ =90°, cosθ = 0, it follows that **a·b** = 0.

At any point on the interatomic surface, the normal vector to the surface will be perpendicular to the gradient vector of the electron density. This statement describes equation 1.8.10

$$\nabla\rho(\mathbf{r})\cdot\mathbf{n}(\mathbf{r})=0 \quad \forall\mathbf{r}\in\mathbf{IAS} \qquad (1.8.10).$$

The interatomic surface is a special type of surface known as a zero flux surface. Flux can be thought of as a measure of how much of something passes a surface. Total flux is a measure of the total amount of a property that crosses a surface at all points on the surface. The property that is of interest is the gradient vector. Therefore how much the gradient vector passes through a surface is known as the flux. In order to define flux mathematically for an arbitrary surface *S*, the gradient of the electron density at every infinitesimal portion of the surface *dS* must be known. At every portion on the surface *dS* the orientation of the surface to the vector field must be known. This is accounted for by the dot product between the gradient vector and the normal to the surface. In order to find the total flux all the infinitesimal contributions must then be added together. The double surface integral sums the infinitesimal contributions. Total flux is defined as equation 1.8.11

$$Total\ Flux=\iint_S \nabla\rho\cdot\mathbf{n}\,dS \qquad (1.8.11).$$

For an interatomic surface equation 1.8.11 becomes equation 1.8.12

$$Total\ Flux = \iint_{IAS} \nabla\rho \cdot \mathbf{n}\, dIAS = 0 \qquad (1.8.12).$$

Gauss' divergence theorem can be applied to equation 1.8.12. The theorem essentially equates a volume integral to a surface integral. The relationship between a volume integral and a surface integral is shown in equation 1.8.13

$$\iiint_{V} dV \nabla \cdot \mathbf{V} = \iint_{S} dS\ \mathbf{V} \cdot \mathbf{n} \qquad (1.8.13).$$

Differentiation indicated by the del operator cancels one of the triple volume integrals leaving a double surface integral. Equation 1.8.13 can be applied to the electron density field. Differentiation of the electron density generates the gradient vector field of the electron density. Second order differentiation generates the Laplacian of the electron density. The Laplacian of the electron density is an indicator of the local concentration or depletion of electron density. The pockets of local concentration and depletion of electron density will cancel each other over an entire atomic basin, giving equation 1.8.14.

$$\iiint_{\Omega} dV \nabla^2 \rho = 0 \qquad (1.8.14).$$

Applying Gauss' divergence theorem to equation 1.8.12 gives equation 1.8.15

$$\iiint_{\Omega} dV \nabla^2 \rho = \iint_{S} dS \nabla\rho \cdot \mathbf{n} \qquad (1.8.15).$$

Since the surface is an IAS, the left hand side of equation 1.8.15 becomes equivalent to equation 1.8.12. Therefore equation 1.8.13 equates to equation 1.8.15. It must also be noted that surfaces that are not IASs, but extend to infinity also give $\nabla\rho\cdot\mathbf{n} = 0$ due to $\nabla\rho$ vanishing at infinity.

Looking at figures 1.8.2 and 1.8.3 it can be seen that there is also a highlighted line that joins connected atoms. The line is known as the atomic interaction line (AIL). The AIL is a gradient path that originates at the BCP and terminates at a nucleus. In fact the AIL is two gradient vectors. The two gradient vectors both originate at the BCP. One gradient vector terminates at one nucleus and the other gradient vector terminates at the nucleus of the atom that the first nucleus is connected to. The AIL can therefore be thought of as the bond path. The AIL is perpendicular to the IAS. The BCP is a maximum in electron density along the IAS. The BCP represents a minimum in electron density in the AIL. The BCP is therefore a saddle point with respect to the IAS and the AIL.

The AIL joins two connected atoms. The two connected atoms do not necessarily have to be chemically bonded in order for an AIL to exist. In fact there will be an observed AIL between any two atoms at infinite distance, but it would be incorrect to assume that these two atoms are chemically bonded. In order for the AIL to be interpreted as a bond path between two chemically bonded atoms, the molecule must be at a local minimum on the potential energy surface. Therefore if the geometrical organisation of a molecule deviates from a local minimum to a higher energy level the interpretation of AILs can change.

The above text has described how the electron density is used to partition atoms in molecules up into individual objects with naturally occurring finite boundaries. The volume of space taken up by an atom is known as the atomic basin Ω. Properties can be attributed to individual atoms in molecules. These atomic properties are defined as volume integrals over Ω. The integrands have to be specified for each property. The atomic volume is defined by equation 1.8.16

$$v(\Omega) = \int_{\Omega} d\tau \qquad (1.8.16).$$

The electronic population of an atom can be found by following equation 1.8.17

$$N(\Omega) = \int_{\Omega} d\tau \rho(\mathbf{r}) \qquad (1.8.17).$$

The QTAIM charge *q* can be calculated once the electronic population is known by simply subtracting the atomic population from the atomic number of the atom in question

$$q(\Omega) = Z - N(\Omega) \qquad (1.8.18).$$

The atomic energy of an atom is defined as follows

$$E(\Omega) = \int_{\Omega} dr(K(\mathbf{r}) \cdot (1 - R)) \qquad (1.8.19)$$

where K is the kinetic energy density, which is defined by equation 1.8.20, and *R* represents a virial ratio correction factor.

$$K(\Omega) = \frac{1}{4} N \int_{\Omega} d\tau' \left[\Psi^* \nabla^2 \Psi + \Psi \nabla^2 \Psi^* \right] \qquad (1.8.20).$$

The above text has described how QTAIM can be used to generate properties relating to the electron density of a system, and atomic properties for atoms in molecules. The properties described will be used throughout the course of this research.

Research Report.

2 - Data Generation.

2.1 – Hydrogen Bond Complex Generation.

In this section the methods used to obtain wave functions for the hydrogen bonded complexes will be described. The methods used here will aim to replicate experiment as accurately as possible. Contrary to popular methodology used in modelling hydrogen bonding, the complexes will not be subjected to any geometrical constraints. Geometry optimisations that will be described in more detail below will be undertaken in order to find one minimum on the potential energy surface. There is only one hydrogen bond basicity value given for any hydrogen bond acceptor site in any given scale. However there are many minima for each complex. As the chosen geometry will be used to model basicity values it will be convenient to use a single geometry. This is because multiple geometries give multiple models. The aim is for one model arising from one geometry models one basicity value. Therefore only one of the minima on the potential energy surface will be used. It is important that the optimised geometry of the complex taken models the hydrogen bond in question accurately.

Geometry optimisations are run in the gas phase, and, in the most cases one donor complexed is with one acceptor. Examples of 2:1 complexes will be described also. However, these trimeric systems are as complicated as the modelling gets in this research. A trimeric complex in the gas phase does not resemble the intricate network of hydrogen bonding that one would expect to find in an experimental system in which a hydrogen bond acceptor is surrounded by an excess of donor. Given that the aim of this research is to provide industrial chemists with a quick estimate of hydrogen bond strength, generating optimised geometries to model the intricate network of hydrogen bonding in an experimental system would be far too costly. Generating an optimised geometry that accurately represents the hydrogen bond of interest, and minimises or

eliminates any other interactions that stabilise the complex is crucial to the success of this research. It follows that the starting geometry from which the optimisation is run must be carefully chosen. When selecting a starting geometry it is important to position the hydrogen bond donor around the hydrogen bond acceptor in such a way that secondary stabilising interactions are minimised. By allowing the geometry optimisation algorithm to run from such a carefully chosen starting point, we would hope to find a minimum that represents the strongest possible hydrogen bond. This might well not represent the global minimum as secondary stabilising interactions could lower the energy of the complex further. The aim of this research is to model the strength of the hydrogen bond and not the stability of the complex. It would be useful to medicinal chemists to have an accurate estimate of the strength of a particular hydrogen bond rather than the stability of an arbitrary complex.

Thermodynamic data, namely ΔG, can be obtained from the equilibrium constant of the hydrogen bond formation reacton. It is therefore important to comment on the assumptions made when linking an optimised geometry to an experimental equilibrium constant. The optimised geometry is in the gas phase at 0K. The experimental procedure allows a mixture of liquids to come to equilibrium at 298K. Therefore there is no contribution from vibrational entropy in the model. Also, in the equilibrium constant is taken from a dynamic equilibrium where atoms, are free to move and bonds free to rotate there is a contribution to the thermodynamic signature from translational and rotational entropy. Translational and rotational entropy is not accounted for in the model. In the experiment, the acid and base are surrounded by solvent. The solvent, CCl_4, is non-polar in order to eliminate enthalpic contributions to the thermodynamic signature from breaking and forming bonds between solvent and acid and base. There will be an entropic contribution from the change in arrangement of solvent around the acid and base during complex formation that is not accounted for in the model.

Thermodynamic data is important in drug design and discovery. Thermodynamic data can then be applied to relevant case studies to give an estimate of binding

properties. In the drug discovery process a candidate drug must bind the target with high potency before it is carried forward further down the drug development line. As there are contributions to ΔG from each individual hydrogen bond site, it is important to model each hydrogen bond as specifically and as purely as possible. If optimised complexes were to be stabilised by secondary interactions other than the primary hydrogen bond in question, inaccurate estimates of hydrogen bond strength are a possible consequence.

Complexes stabilised by the strongest hydrogen bond are of interest. When estimating binding data of systems that have multiple hydrogen bonds, any errors associated with predicted hydrogen bond strengths will be added, giving a large error when attempting to predict the total contribution of hydrogen bonding to the system. This can be minimised by again eliminating secondary stabilising interactions in our optimised complexes by selecting a good initial geometry. Of course ΔG of binding in drug–protein cases is far more complicated and involves an entropic contribution as discussed above.

Hydrogen bond formation contributes favourably to the enthalpic component of ΔG. Although it is possible to calculate ΔG from the equilibrium constant of a particular process, it is not possible to associate a ΔG calculated for a hydrogen bond formation reaction in an apolar solvent with a ΔG for a drug–protein binding reaction. This is because hydrogen bond formation in a drug-protein interaction involves thermodynamic processes with the biological solvent. No information about the desolvation enthalpy and desolvation entropy, the conformational entropy, or the hydrophobic effect is known. The only information we aim to provide the industrial chemist with is a quick way to accurately measure the relative strengths of hydrogen bond acceptor sites. Hydrogen bonding is only one piece in the complicated puzzle of obtaining ΔG of binding. However if a certain drug needs to be optimised to give a more favourable binding ΔG, knowing how a particular hydrogen bond can be strengthened could provide an intuitive starting point.

2.2 Computation Of Hydrogen Bond Complexes.

For reasons outlined above it is important to select a quality starting geometry. A quality starting geometry should consider possible secondary stabilising interactions that could occur during in the optimisation process. A recent study analysing the geometries of hydrogen bonds to alkyl chlorides illustrates how secondary stabilising interactions lower the energy of the complex [130]. For our purpose this is undesirable. The problem in [130] was overcome by fixing certain angles and running a geometry optimisation under these constraints. This is again undesirable for this research as the aim is to represent an experiment where the hydrogen bond donors and acceptors are free to react and reach equilibrium. The aim of this research is to model a complex stabilised by only the hydrogen bond in question. Therefore the goal is to optimise the geometry of the complex to a local minimum that minimises or eliminates any secondary stabilising interactions. The optimisation must be carried out without any constraints.

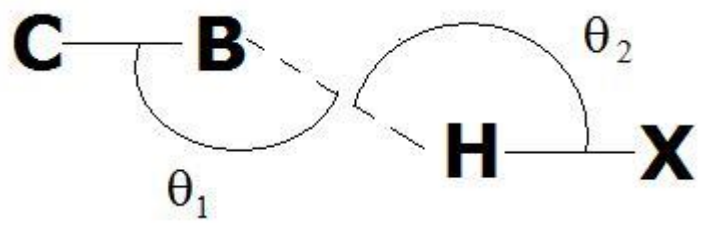


Figure 2.2.1. A typical hydrogen bond. The hydrogen bond is illustrated by the dashed line. B represents the hydrogen bond acceptor and X is the electronegative atom the hydrogen is bonded to. In this research X is limited to O, N, or F. Relevant angles are labelled θ_1 and θ_2.

Given the goals of the optimisation process, the initial geometry must be carefully selected. The hydrogen bond length and θ_1 in figure 2.2.1 are based on a method used by Platts to optimise the geometry of hydrogen bonded complexes [61]. The hydrogen atom donating the hydrogen bond will be placed 2 Å away from the hydrogen bond acceptor atom in the region of the lone pair. For carbonyl oxygen and sp2 nitrogen

acceptors θ_1 will be set around 120°, for alcoholic and ether oxygen and sp3 nitrogen acceptors θ_1 will be set to 109°. The hydrogen bond angle θ_2 will be set to 180° in order to minimise steric interference of the side chains of the hydrogen bond donor molecule with groups on the hydrogen bond acceptor molecule. The majority of hydrogen bond acceptors in this research involve either oxygen or nitrogen atoms. There are a few cases in which the acceptor site is either a sulphur or a halogen atom. The halogens used are fluorine, chlorine or bromine. Hydrogen bonds to sulphur or halogen acceptor sites are significantly weaker than those to oxygen or nitrogen sites. It is therefore reasonable to expect complexes containing sulphur of halogen acceptor sites to be further stabilised by secondary interactions. As discussed above, secondary interactions are undesirable for the purpose of this research. In order to minimise the chance of generating an optimised geometry that has secondary stabilising interactions, θ_1 will be set to 180° for starting geometries involving halogen acceptors, and 120° for sulphur acceptors. The dihedral angle C-B-H-X will vary depending on the system. The choice of the dihedral will rely on chemical intuition. A good starting dihedral angle will minimise steric interference by spacing groups attached to C and X as far apart as possible on the Newman projection along the B-H bond. Side chains must also be considered. Side chains must be arranged around the molecule to minimise steric interference too. By following all of the above steps, not only will the chance of finding the desired minimum be increased, but the time taken to compute the optimisation step will be greatly reduced.

The various different hydrogen bond donors used in this research are made up of OH donors, NH and NH^+ donors and HF. The OH donors used are water, methanol, 4-fluorophenol and the side chain of serine. The serine hydrogen bond donor will be capped with methyl at both the carbonyl and amino terminals. The NH donor used is methylamine. The NH^+ donors used were limited to protonated tertiary amines for reasons that will be explained in following sections. The hydrogen bond donors used in this research were chosen in line with those that the pK_{BHX} scale is thought to be

applicable to [1]. The pK_{BHX} scale is not said to be applicable to formation calculated where HF is the hydrogen bond donor. Therefore HF is used as a control in this research as we would not expect reliable predictions of pK_{BHX} values to be obtained from complexes generated where HF is the hydrogen bond donor. However HF is the most simple hydrogen bond donor to use. HF is small and there are therefore fewer primitive Gaussian functions to fit, leading to reduced computational time. HF is linear which reduces the complexity associated with side chain steric interference. HF is a weak hydrogen bond acceptor, much weaker than the majority of bases studied in this research. Because HF is a much weaker hydrogen bond acceptor than the majority of bases studied in this research there is less chance of the hydrogen bond being inverted in the optimisation phase. An inverted hydrogen bond is one in which the desired hydrogen bond donor becomes rotated during the optimisation process and the electronegative hydrogen atom bonded to the desired hydrogen bond donor atom accepts a hydrogen bond from a donor somewhere on the base. Inverted hydrogen bonds are problematic and will be discussed in more detail below.

Water is the most simple hydrogen bond donor that the pK_{BHX} scale can be applied to. It is also biologically relevant. Water may form hydrogen bonds with proteins and drug molecules as they are transported in solution to their target. The enthalpy is lost as hydrogen bonds are broken as water is shed from the binding sites of the drug and protein. Enthalpy is gained when hydrogen bonds are formed when the drug and protein bind. Not only is water small and computationally inexpensive, it is biologically relevant and the pK_{BHX} scale can be extrapolated to encompass its complexes. There are no real issues concerning the starting geometry of hydrogen bonded complexes where water is the donor. The starting geometry of complexes using water as the donor are generated following the above guidelines. The only real consideration is that the non-hydrogen-bonded hydrogen of the water molecule needs to be angled away from the base to reduce inverted hydrogen bonds.

Once optimised complexes have been generated for a series of bases using water as the donor, it is easy to extend the data set to include methanol and 4-fluorophenol donors. The optimised geometries for the complexes where water is used as the donor can be used as templates for methanol and 4-fluorophenol donors. By opening up the optimised water outputs in GaussView, and simply substituting the non hydrogen bonded hydrogen of water for methyl, a good starting geometry for the methanol set of donors is created.

The same principle can be applied to 4-fluorophenol too. No further modifications need to be made to generate the methanol input geometries. It is necessary to carefully rotate the phenol ring out of the plane of the hydrogen bond when dealing with the 4-fluorophenol initial geometries.

Generating an initial starting geometry for serine complexes requires more attention to detail. Serine is used as a hydrogen bond donor in this research. The side chain OH group acts as the hydrogen bond donor. However the oxygen of the hydrogen bond donor OH group may act as a hydrogen bond acceptor. There are three other separate sites on the serine molecule that may act as hydrogen bond acceptors. They are the carbonyl and alcoholic oxygen atoms at the carboxyl terminus, and the nitrogen atom at the amino terminus. The oxygen acceptors at the carboxyl terminus are situated five bonds away from the hydrogen bond donor. The amino nitrogen is four bonds away from the hydrogen bond donor. Therefore between the hydrogen bond acceptor on the base and the oxygen potential acceptors on the serine molecule there are five separate dihedral rotations that may occur. Between the hydrogen bond acceptor in the base and the nitrogen potential acceptor in serine there are four separate dihedral rotations that may occur. There is a much greater potential for conformational freedom in serine complexes compared to water methanol and even 4-fluorophenol complexes.

In many cases secondary interactions proved to be unavoidable. The secondary interactions usually include weaker stabilising hydrogen bonds between the above mentioned potential hydrogen bond acceptors in the serine molecule and a hydrogen

atom situated somewhere in the base. The only way to avoid secondary stabilising interactions from occurring in systems with larger molecules such as serine would be to freeze certain angles in the initial geometry and running a potential energy scan optimising the relaxed degrees of freedom. It has been described above that for the purpose of this research that it is undesirable to fix any geometrical parameters in the optimisation process. The guidelines outlined above have been followed in order to minimise secondary stabilising interactions.

Although many serine complexes have unavoidable secondary stabilising interactions the desired hydrogen bond is still the dominant attractive force in all of the geometries used. It was on occasions very difficult to generate an input geometry that optimised to a minimum where the desired hydrogen bond remained in place for serine complexes. The combination of the size of the serine molecule, the number of hydrogen bond donors and acceptors in serine and the totally relaxed optimisation procedure often resulted in minima being generated where the complex is stabilised by hydrogen bonds other than the desired hydrogen bond. At times this was a rotated hydrogen bond. It was also not uncommon to find structures stabilised by a dominant hydrogen bond between the above mentioned hydrogen bond acceptors in serine and a hydrogen atom somewhere on the base. Unusually, the hydrogen bond in these cases often involved a CH donor group.

Hydrogen bonds involving CH donor groups are weak. This is because, in the context of electronegativity, CH bonds are non polar. Therefore it would be expected that a stronger interaction would be formed between the desired OH donor on serine and the acceptor site in the base. The fact that on some occasions this was not found suggests that serine itself prefers to be in a geometry that effectively hides the desired hydrogen bond donor site from being able to form a hydrogen bond, an unfortunate consequence of the relaxed optimisation method. The optimised geometry that is stabilised by weak CH hydrogen bonds must be in a lower energy conformation than any geometry stabilised by the desired hydrogen bond along the potential energy surface for that

particular starting geometry. As described above, the conformation of the serine molecule could contribute more to the energy of the system than a strong hydrogen bond.

The best way to overcome optimising to a minimum that does not include the desired hydrogen bond is to start the optimisation again using a different starting geometry. An improved starting geometry can be carefully selected by visually following the stages of the failed optimisation. By following the stages of the failed optimisation it is possible to visualise where the attractive and repulsive interactions are taking place along the potential energy surface of that particular starting geometry. This gives good insight into selecting an improved starting geometry. Unwanted attractive and repulsive interactions can be minimised by rotating attractive and repulsive groups away from each other.

The most tricky desired optimised geometries to find were those where methylamine was used as the hydrogen bond donor. Unlike all the OH donors used, there are two possible hydrogen bond donors bonded to the nitrogen on methylamine. Also, amino nitrogen atoms are much stronger than any of the OH oxygen atoms used in this research. Due to possible steric interference resulting from the trivalent nature of the amino nitrogen, and the knowledge that amines are strong hydrogen bond acceptors the possibility of inverted hydrogen bonds being found increases. In fact that is what was found on numerous occasions.

It has been described above that a completely relaxed optimisation procedure was chosen to resemble experiment. In a computation chosen to resemble experiment inverted hydrogen bonds are found. This leads to the question of what is actually being measured in the experiment. In the experiment either methanol or 4-flouorphenol is used as the hydrogen bond donor. However when measuring the basicity of weaker hydrogen bond acceptors than methanol it is likely that inverted hydrogen bonds will be formed. The pK_{BHX} value given in the database is the logarithm of the formation constant between donor and acceptor. However, when measuring this, the concentration of the

complex is used. The concentration of the complex does not distinguish between desired and inverted hydrogen bonds. Therefore inaccurate basicity values could be given. The experimentalists try to overcome this in two ways. Firstly, they saturate the base by using a large excess of donor. Secondly, after measuring the concentration ratios needed to calculate the formation constant, family dependent correlations are set up between the pK_{BHX} values and the shift in wavenumber of the OH bond of the donor upon hydrogen bond formation.

It has been shown for many chemical families that family dependent correlations exist between the shift in wavenumber of the OH donor upon hydrogen bond formation, and the pK_{BHX} values calculated by measuring the hydrogen bond formation constant experimentally [71-83, 131-138]. Given such a strong relationships between the above mentioned shift in wave number and pK_{BHX} values, it is possible that any inverted hydrogen bonds found are insignificant. Even for the family of thioamides and thioureas there is a strong relationship between the shift in wavenumber of the OH bond of the donor group and the pK_{BHX} scale [79]. Thioamides and thioureas are weaker hydrogen bond acceptors than the oxygen site on methanol or 4-fluorophenol. It would be a reasonable hypothesis to suggest that inverted hydrogen bonds are possible in complexes with methanol or 4-fluorophenol and thioamides and thioureas. However given the strong relationship between the spectroscopic and thermodynamic data inverted hydrogen bonds appear either absent or insignificant.

Another factor to take into account when considering the possibility of inverted hydrogen bonds is the strength of the hydrogen bond donor. There is an entire reciprocal literature to hydrogen bond basicity devoted to hydrogen bond acidity [53, 70, 91, 139-141]. The principles remain the same though and scales of hydrogen bond acidity are based on thermodynamics, are relative, and use a standard common hydrogen bond acceptor. To my knowledge the issue of the direction of the hydrogen bond complex, in other words whether the desired hydrogen bond is formed or an inverted complex is found, has not been discussed in the literature. It can only be assumed that the

information found in the spectroscopic analysis is strong enough to reliably conclude that the desired hydrogen bond has been formed. However it is worth noting and discussing the possibility of alternative hydrogen bonds forming in the experimental procedure. Not only is it possible that inverted hydrogen bonds form; it is also possible that a network of hydrogen bonds will be formed in the experimental equilibrium reaction.

Given that it is possible for a network of hydrogen bonds to form, the cooperative effect should also be considered. The cooperative effect is the ability of one interaction to either induce or inhibit the formation of a second interaction. It could be that the desired hydrogen bond is affected by the cooperative effect. In the equilibrium reaction that takes place in the experimental procedure, the cooperative effect is not really of any concern. However the cooperative effect becomes relevant when applying hydrogen bond basicity to biological interactions. A pK_{BHX} basicity value that has been either greatly increased or reduced by a cooperative interaction at an additional site on the base might give misleading estimates of the equivalent basic site in a biological interaction. If the biological interaction involves a protein, for example, the additional site that is the site of the cooperative effect might no longer be available in the biological system. It could be that the site of the cooperative effect in the equilibrium experiment is buried away in the core of the protein in a biological interaction.

At this point it is also worth noting that each hydrogen bond is unique. The information that can be extracted from the pK_{BHX} database ranks the relative basicities of only the set of bases that data has been collected on. The internal structure of the database has not been discussed and it is therefore unknown whether pK_{BHX} values can be extrapolated beyond the data set. We are told that the pK_{BHX} database is applicable only to certain OH and NH donors and that the relative basicities are likely to change if another class of donor is used. Fortunately, the vast majority of relevant biological reactions involve OH or NH hydrogen bond donors. The database is also organised into chemical identities. Each base is classified primarily by its donor atom type, secondly by its functional group and thirdly by its subfunctional group. For example the base

ethylamine would have a nitrogen acceptor site, a functional group as an amine and a subfunctional group as a secondary amine. Therefore if a medicinal chemist requires information about a potential hydrogen bond site in a case study, knowing the functional group and subfunctional group of the acceptor site is necessary in order to obtain an estimate of its basicity from the pK_{BHX} database.

Although the pK_{BHX} database contains information about over 1100 bases and 1300 basic sites it is unlikely that an exact match for a relevant medicinal chemistry problem will be found. An exact match is required to obtain not only useful numerical information about the basic site but an exact replication of the experimental conditions. An exact replication involves the same donor, solvent pH temperature etc. Therefore the best information a medicinal chemist can hope to obtain from the pK_{BHX} database is whether an acceptor site is qualitatively considered weak or strong. This estimation of strength will be extrapolated from the nearest match to the case study from the database. Chemical insight and intuition can be applied to estimate the effects of any substitutions going from the nearest match to the case study. For example, electron withdrawing substituents such as electronegative groups or aromatic rings decrease the basicity whereas electron donating groups such as alkyl chains increase basicity. Further insight could be gained by analysing the internal structure of the pK_{BHX} database. The aim of this research is to develop a model to compute fast and accurate pK_{BHX} values for any given base. Predicting basicity values for any given base is useful to the medicinal chemistry community as it eliminates the need to extrapolate beyond the pK_{BHX} database data set to obtain basicity values. However, the basicity values computed are still relative and in fact only really predict the equilibrium of excess 4-fluorophenol and base in an aprotic solvent.

2.3 Ab Initio Quantum Chemistry.

The quantum theory of atoms in molecules has been described in section 1.8 which explains how that in order to obtain atomic properties a wave function is required.

The wave function is a solution to the Schrödinger equation. Therefore a solution to the Schrödinger equation is essential to carry out the research in this document.

The only experimental values that appear in the Schrödinger equation are fundamental constants. A computational method that uses only fundamental constants and no other experimental parameters is said to be *ab initio*. The term *ab initio* means from first principles. In fact, if the Schrödinger equation is given in atomic units the fundamental constants are eliminated. Ab initio computations can be very demanding. Due to computational limitations, an exact solution to the Schrödinger equation can only be obtained for the hydrogen atom. In order to obtain a reliable electron density from *ab initio* methods a number of approximations must be made.

The Born-Oppenheimer approximation is chosen to eliminate the problem of changing nuclear coordinates. As the electrons move much more quickly than the nuclei, the nuclei are considered to be clamped in a fixed position.

As described in section 1.8, the wave function is an *N*-electron solution to the Schrödinger equation. The wave function describes the spatial and spin coordinates of all electrons. However the *N*-electronic wave function may be approximated by using the product of one-electron wave functions. One electron wave functions describe the spin orbital for a single electron. This is known as the Hartree product. While the Hartree product allows for a simplified wave function, some crucial physical effects are lost. The Hartree product inaccurately describes electron-electron repulsion by assuming that the repulsion action on one electron is that of the average of all other electrons. This statement is incorrect as each electron directly repels another electron. As the electron-electron repulsion is seen as the average field from all other electrons, the repulsion acting on one electron depends on the solutions to each of the other one-electron wave functions. Therefore solutions to the Schrödinger equation based on the Hartree product must be solved iteratively until a self consistent field (SCF) is realised.

The Hartree product does not account that electrons are equivalent and indistinguishable. Section 1.7 describes how the Pauli exclusion principle must be obeyed

and therefore the correct wave function for an electronic system must be antisymmetric. For example, if electron 1 in orbital 1 has spin *a*, and electron 2 in orbital 1 has spin *b*, then it must be equally allowed that electron 1 in orbital 1 has spin *b* and electron 2 in orbital 1 has spin *a*. Interchanging the electron labels changes the sign of the wave function. The Hartree equations therefore become the Hartree-Fock equations. The possible spin orbitals for an *N*-electron system can be written as a matrix; such an example is given in equation 2.3.1

$$\Psi(X_1, X_2, \cdots X_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(X_1) & \chi_2(X_1) & \cdots & \chi_N(X_1) \\ \chi_1(X_2) & \chi_2(X_2) & \cdots & \chi_N(X_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(X_N) & \chi_2(X_N) & \cdots & \chi_N(X_N) \end{vmatrix} \qquad (2.3.1)$$

where the columns of the matrix refer to the electron and the rows to the spin orbital. It can be seen that the wave function is equivalent to the determinant of the matrix. This special type of determinant is called the Slater determinant.

For molecular systems a linear combination of atomic orbitals is used. This idea centres the orbitals on every nucleus in the molecule rather than just one. Basis sets are used to make this possible. Basis sets are sets of functions which are combined in linear combinations of atomic orbitals to generate molecular orbitals. Atomic orbitals are centred on the atomic nucleus of each atom. Atomic orbitals are well described by Slater type orbitals (STO). STOs can be approximated by a linear combination of Gaussian type orbitals (GTO). Various combinations of GTOs can be used to describe the electrons of each atom. These combinations of GTOs are called basis sets. The basis sets are therefore the set of functions introduced to an orbital in the Hartree-Fock equations.

Basis sets vary in level of theory and hence accuracy. The more GTOs fitted to a given molecule to construct a basis set the closer the energy is to the exact value. Large basis sets are accurate but consequently give rise to lengthy calculations. Minimal basis sets fit one STO to each nucleus. Each STO may be made up of n-Gaussian functions.

Minimal basis sets therefore take the form STO-nG, where G represents Gaussian functions. Larger basis sets are required to accurately predict the energy of larger systems. Split valence basis sets place fewer restrictions on the spatial coordinates of the electrons. Split valence basis sets fit more STOs to each nucleus. For example double zeta basis sets fit 2 STOs to each nucleus whereas triple zeta basis sets fit three STOs to each nucleus. By fitting more than one STO to each nucleus, the effects of surrounding atoms in a molecule on that nucleus are accounted for to a certain extent. The presence of other atoms could lead to bonding or weaker interactions. Split valence basis sets assume that the core electrons are less influenced by chemical environment than the valence electrons. Double zeta split variance basis sets therefore assume the form A-BCg, where A represents the number of GTO's fitted to core electrons and B and C indicate the valence electrons are represented by two linear combinations of GTOs, the first made up of B GTOs and the second of C GTOs. Triple zeta basis sets take the form of A-BCDg.

Polarisation functions can also be added to improve accuracy when considering bonded systems. Polarization functions add orbitals of higher angular momentum to atoms that are usually unoccupied in the ground state atom. For example the 1s orbital is usually considered for the hydrogen atom. Adding functions that are representative of p and d orbitals allows the electron density around hydrogen nuclei to be more spread out. This flexibility more accurately describes bonded systems.

Diffuse functions can also be added to represent the electron density of an atomic orbital that is most distant from nucleus. Diffuse functions are indicated by a + symbol. Two + symbols indicate that diffuse functions are also added to light atoms hydrogen and helium. Diffuse functions are important in modelling anionic and hydrogen bonded systems.

Hartree-Fock does not account for electron correlation. The ground state energy obtained in the SCF is always higher than the real ground state energy. The difference is known as the correlation energy. Density Functional Theory (DFT) has been used to

tackle the electron correlation energy problem. DFT allows the ground state energy of a multi electron system to be expressed as a functional of the electron density. Therefore DFT is a post Hartree-Fock method used to introduce electron correlation. The B3LYP is a popular functional used in *ab initio* chemistry.

2.4 Computational Methods And Statistics.

DFT calculations were performed using Gaussian 03 [142]. The hydrogen bonded structures were optimised at the B3LYP/6-311++G(2d,p) level. Properties of the complexes, based on those used by Platts [61] were analysed and correlated with their respective pK_{BHX} values. The properties include the hydrogen bond length r(X...H), change in the HBD bond length Δr(H-X), and the hydrogen bond energy ΔE. Quantum chemical topology [114, 115] related properties are those in which the electron density of the hydrogen bonded complexes are used to analyse the hydrogen bonded complex. Atomic properties are obtained by a volume integral over an atomic basin (Ω). The atomic properties used here are the change in the HBD energy $\Delta E(H)$, and the change in charge on the HBD $\Delta q(H)$. The atomic energies and charges are defined in equations 1.8.17 to 1.8.20. Atomic integrations used to analyse atomic properties were calculated with the AIMALL suite [143].

Bond properties were analysed at the critical points. The gradient vector field of the electron density displays bond critical points with (3,-1) critical points defined as the density saddle point between 2 nuclei. It is the (3,-1) critical points that are of interest here. Properties analysed at the hydrogen bond critical point are the electron density ρ(X...H), the Laplacian of the electron density $\nabla^2\rho(X...H)$, the kinetic energy density G(X...H), and the kinetic energy divided density divided by the electron density G/ρ(X...H). Properties of the HBD critical point are the change in the electron density $\Delta\rho(H-X)$, and the change in the laplacian of the electron density $\Delta\nabla^2\rho(H-X)$. Properties of the bond critical points were obtained from the MORPHY98 package [144]. The MORPHY

package was also used to generate figure 5 which shows an atomic basin of a HBD and all bond critical points of a hydrogen bonded complex.

Basis set superposition errors (BSSE) calculations were performed in order to accurately estimate the hydrogen bond binding energy for the water series only. As the BSSE of binding energy was very small, BSSE calculations were not performed to calculate the binding energy for the other series of complexes. The method used to perform the basis set superposition error calculations was based on the method used by Handy and co-workers [145]. Equations 2.4.1, 2.4.2 and 2.4.3 show the methods used to calculate the BSSE,

$$\Delta E = E_{complex} - E_{HBA} - E_{HBD} \qquad (2.4.1)$$

$$\Delta E(BSSE) = E^{*}_{HBA} + E^{*}_{HBD} - E^{CP}_{HBA} - E^{CP}_{HBD} \qquad (2.4.2)$$

$$\Delta E^{CP} = \Delta E - \Delta E(BSSE) \qquad (2.4.3).$$

Equation 2.4.1 shows how to calculate the hydrogen bond binding energy without considering the BSSE. Equation 2.4.2 calculates the BSSE using the counterpoise correction method [146]. In equation 2.4.2 the asterisk refers to the geometry of the HBA and HBD in the optimised complex, and CP stands for counterpoise and refers to the energy of the HBA and HBD in the presence of ghost HBD and HBA Gaussian functions, respectively. Equation 2.4.3 calculates the counterpoise corrected hydrogen bond binding energy from equations 2.4.1 and 2.4.2.

Criticism of the B3LYP functional led to several investigations showing how the MPWB1K functional consistently performs better for weak interactions such as hydrogen bonding [147-149]. The B3LYP functional has been criticised for unsatisfactory modelling of kinetics and hydrogen bond energies as well as poorly predicting the energetic of weak interactions. Taking this into account, a further set of calculations the MPWB1K/6-311++G(2d,p) level were performed for the training set of bases in the water complex.

The water complexes were chosen to reduce computational time as water is the smallest hydrogen bond donor used.

The Grubb's test is a statistical method used to test for outliers. The Grubb's test is performed following equation 2.4.5.

$$Z = \frac{|mean - value|}{SD} \qquad (2.4.5).$$

The mean value in equation 2.4.5 is the mean of all values of the property that is being analysed, and SD is the standard deviation. The Grubb's test is a test that looks for an outlier within a data set. The Grubb's test determines whether the most extreme value in the data set is a significant outlier from the rest. If an outlier is detected, the outlier can be removed from the data set and the test be performed again iteratively until no further outliers are found. The Grubbs test performed here is at the 0.05 significance level. If a Z value is found to be greater than the critical level, then there is less than 0.05% chance that that the point would be found so far from the rest by chance alone. Therefore it is a conclusive outlier. The Grubb's test also relies on a normal distribution of the data.

3 – Results And Discussion.

3.1 Principal Component Analysis Of Properties Of Hydrogen Bonded Complexes.

Principal component analysis (PCA) [150] is a statistical method used to analyse multi-dimensional data. PCA is essentially a method used to rotate and centre an axis system on the dimension with the largest varience. This is done by analysing the covariance matrix of the data set. The eigenvalues of the covariance matrix represent each principal component. The eigenvalue with the largest represents the first principal component, the one with the largest varience. The second largest eigenvalue represents the second prioncipal component, the one with the second largest varience, and so on. All principal components are orthogonal to each other. The corresponding eigenvector for each principal component is a column vector with *n* columns, where *n* represents the number of properties. The row vector with the values for each of the *n* properties, for eavh variable is combined with the eigenvector. The resulting dot product represents the scoreings for each variable. The scores for each variable for one principal component can be plotted against the scores for each variable for another principal component. The contribution each property makes to each principal component is known as the loading value. Loading values can be plotted in the same way as scoring values. Principal component analysis has been carried out for several properties of hydrogen bonded complexes modelled as the following text explains.

DFT calculations were performed using Gaussian 03 [142]. Hydrogen bonded complexes were generated by placing a hydrogen atom of a water molecule approximately 2.0 Å away from the primary hydrogen bond acceptor site. Water was chosen as the hydrogen bond donor as it is the simplest biologically relevant solvent. Also the order of the pK_{BHX} scale based on the reference donor 4-fluorophenol is unchanged when water is used as the hydrogen bond donor, therefore allowing the hydrogen bonded complexes used here to be identified with the pK_{BHX} scale. The

methods used to obtain optimised hydrogen bond complexes are described in section 2.2. The hydrogen bond donors and their pK_{BHX} values [1] used are listed in table 3.1.1.

Hydrogen Bond Acceptor	pK_{BHX}	Hydrogen Bond Acceptor	pK_{BHX}
3-Chloropyridine	1.31	Methylformate	0.65
4-Methylpyridine	2.07	N-Methylaniline	0.26
Acetamide	2.06	Phenol	0.07
Acetone	1.18	Pyridine	1.86
Acetonitrile	0.91	Pyrrolidine	2.59
Acrylonitrile	0.7	t-Butylamine	2.23
Aniline	0.46	Tetrahydropyran	1.23
Chloroacetnitrile	0.39		
Dimethyl sulfide	0.12		
Dimethylamine	2.26		
Ethanol	1.02		
Ethyl thiol	-0.16		
Ethylamine	2.17		
Formamide	1.75		
MeSCN	0.73		
Methanol	0.82		
Methyl acetate	1.00		
Methylamine	2.20		

Table 3.1.1. A list of the bases and their corresponding pK_{BHX} values [1] used in this study.

Quantum chemical topology [114, 115] related properties are those in which the electron density of the hydrogen bonded complexes are used to analyse the hydrogen bonded complex. Properties of the bond critical points were obtained through MORPHY98 [144], whereas atomic integrations used in population analysis were calculated with the AIMALL suite [143].

Table 3.1.2 gives a list of the properties of the hydrogen bonded complexes to be analysed here, and how they appear abbreviated in the following results.

Property	Abbreviation	Description Of Property
r(X...H)	r(X...H)	Hydrogen bond length.
Δr(H-O)	delta r(H-O)	Change in bond length of hydrogen bonded hydrogen and oxygen in HBD
ΔE	delta E	Hydrogen bond energy.
q(H2O)	q(H2O)	Charge on HBD
Δq(H)	delta q(H)	Change in charge of hydrogen bonded hydrogen in HBD
ΔE(H)	delta E(H)	Change in energy of hydrogen bonded hydrogen in HBD
ρ(X...H)	rho(X...H)	Electron density at the hydrogen bond critical point
∇^2 ρ(X...H)	lap(X...H)	Laplacian of the electron density at the hydrogen bond critical point
Δρ(H-O)	delta rho(H-O)	Change in electron density at the hydrogen bonded hydrogen and oxygen critical point in HBD
Δ ∇^2 ρ(H-O)	delta lap(H-O)	Change in laplacian of the electron density at the hydrogen bonded hydrogen and oxygen critical point in HBD
G(X...H)	G(X...H)	Kinetic energy density at the hydrogen bond critical point
G/ρ(X...H)	G/rho(X...H)	Kinetic energy density divided by the electron density at the hydrogen bond critical point

Table 3.1.2. A list of the properties to be analysed in the results section. The abbreviation column shows how the property will be displayed in figures.

PCA [150] was carried out on the set of bases in table 3.1.1 and the properties in table 3.1.2. The PCA calculations were carried out using SIMCA-P 10 [151]. PCA calculations were performed in order to clarify that the set of properties in table 3.1.2 separate the bases into clusters of same HBA acceptor atoms, and are uncorrelated.

Correlations between each property of the hydrogen bonded complex in table 3.1.2 and the pK_{BHX} value given in table 3.1.1 were performed by using the linear regression method in SIMCA-P 10 [151].

Figure 3.1.1 shows a scatter plot of the scores of the bases for the first principal component t[1] and the second principal component t[2]. The scores are a dot product of two vectors, one containing the properties of the hydrogen bonded complex and the other is the eigenvector relating to a certain eigenvalue from the covariance matrix of original data. The eigenvalue with the largest magnitude corresponds to the eigenvector in the direction of the most variance. The dot product of this eigenvector with the vector containing the original data of the hydrogen bonded complex gives the score of the first principal component. The score of the second principal component is found by taking the eigenvector from the covariance matrix with the second largest magnitude, and so on. The contribution of each property to each component can be seen in a loading plot such as in figure 3.1.2. The property with the highest loading magnitude for a particular component contributes most to that particular component.

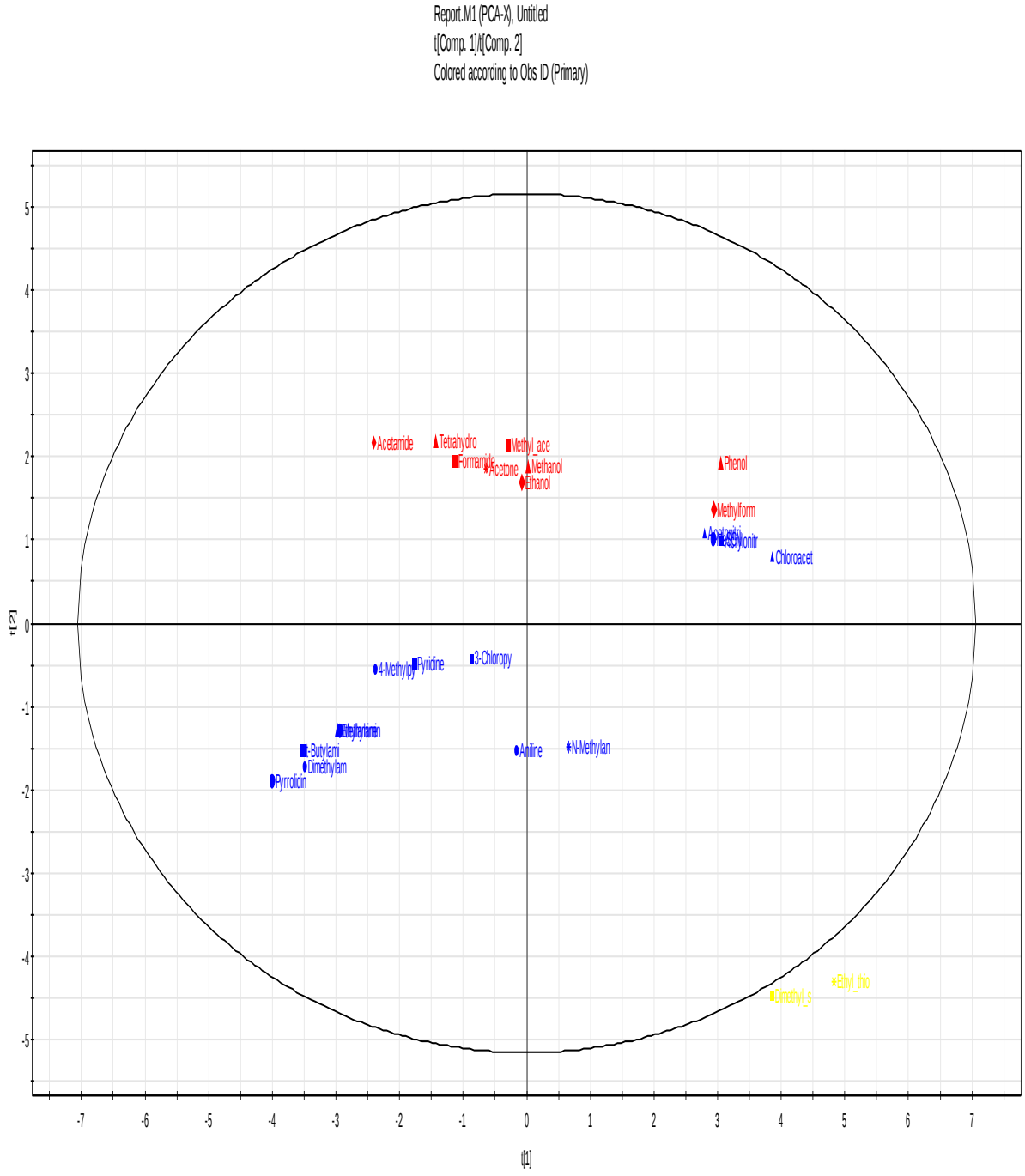


Figure 3.1.1. Scatter plot of the scorings of the first, t[1] and second, t[2] principal components complied from each of the properties in table 3.1.2. Base names in red are O acceptors, blue are N acceptors, and yellow are S.

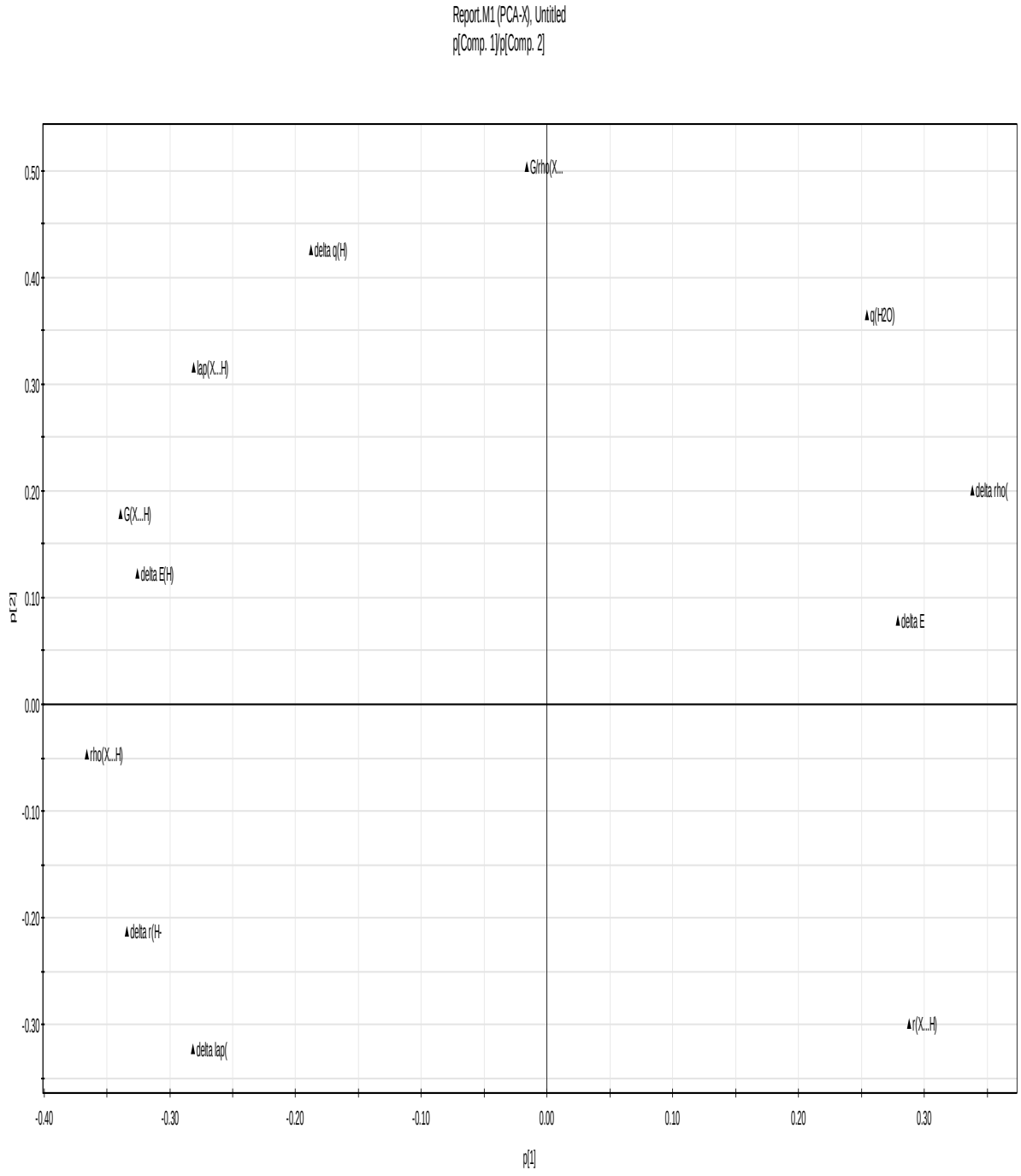


Figure 3.1.2. Scatter plot of the loadings of the first, p[1] and second, p[2] principal components. The properties displayed here have abbreviated descriptions. See table 3.1.2 for details.

Figure 3.1.1 shows a fairly good clustering of the bases by the atoms of their hydrogen bond acceptor sites. It could be concluded that the first principal component relates to the strength of the base. The stronger N acceptor bases have the most negative scores while weak S acceptors have the most positive scores. Intermediate O acceptors seem to be somewhere in between. From figure 3.1.2 we can see that the properties with the largest magnitude along the first principal component are the electron density, and the kinetic energy density at the hydrogen bond critical point. Therefore the topology of the hydrogen bond may relate to its strength. Figure 3.1.1 also shows a possible separation of N and O acceptors by the second principal component. By examining figure 3.1.2, it could be concluded that the charges on the HBD atom distinguish between N and O acceptors because the charge on the HBD and the change in charge on the hydrogen bond donor hydrogen load highly onto the second principal component.

The third principal component offers no useful information as can be seen in figure 3.1.3. Although no information regarding hydrogen bond basicity can be offered, figure 3.1.3 reveals aniline to be a possible outlier in the third principal component. Analysis of the loading plot of the third component (figure 3.1.4) shows how the hydrogen bond energy loads much higher than any other property.

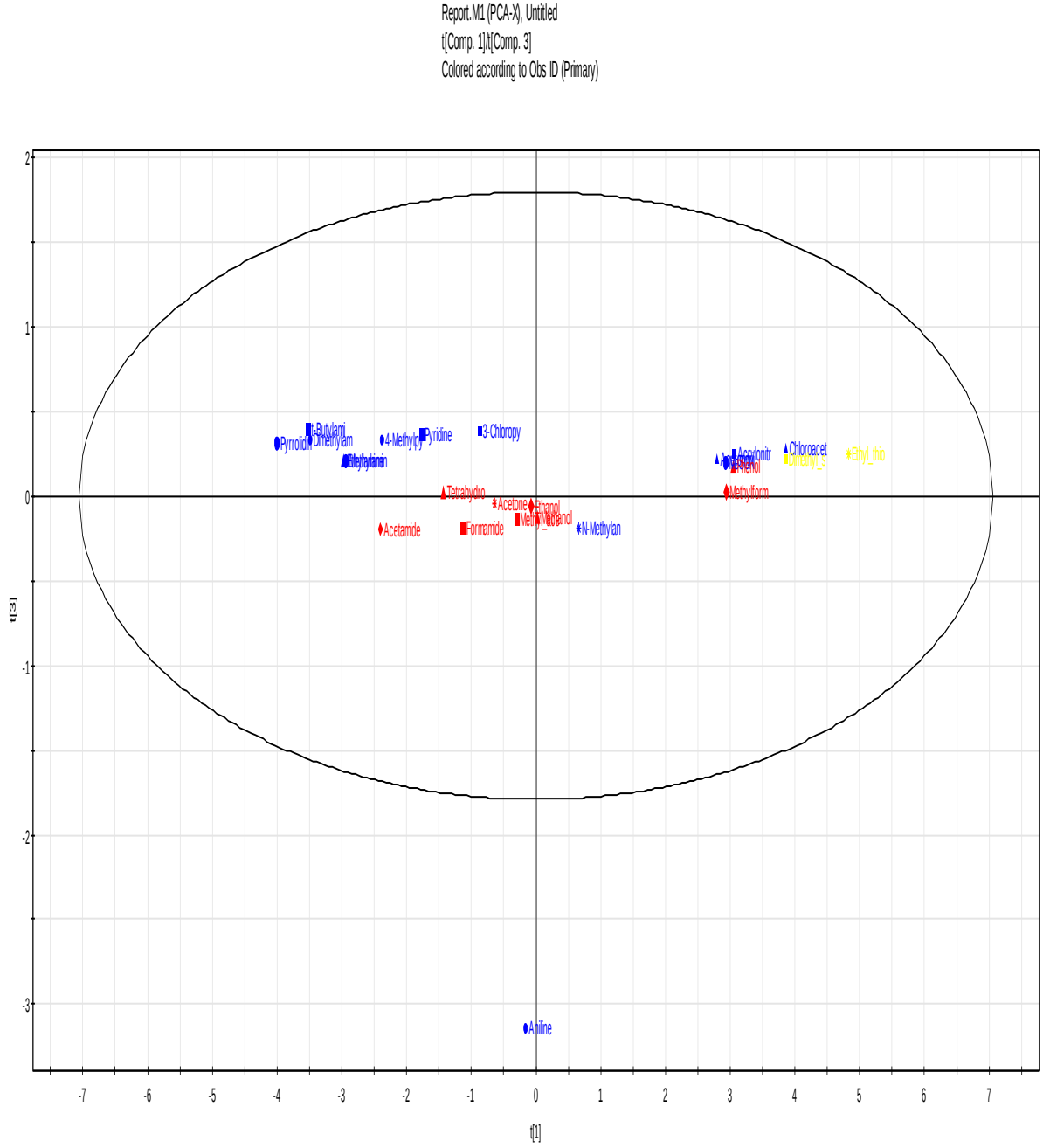


Figure 3.1.3. Scatter plot of the scorings of the first, t[1] and second, t[2] principal components complied from each of the properties in table 3.1.2. Base names in red are O acceptors, blue are N acceptors, and yellow are S.

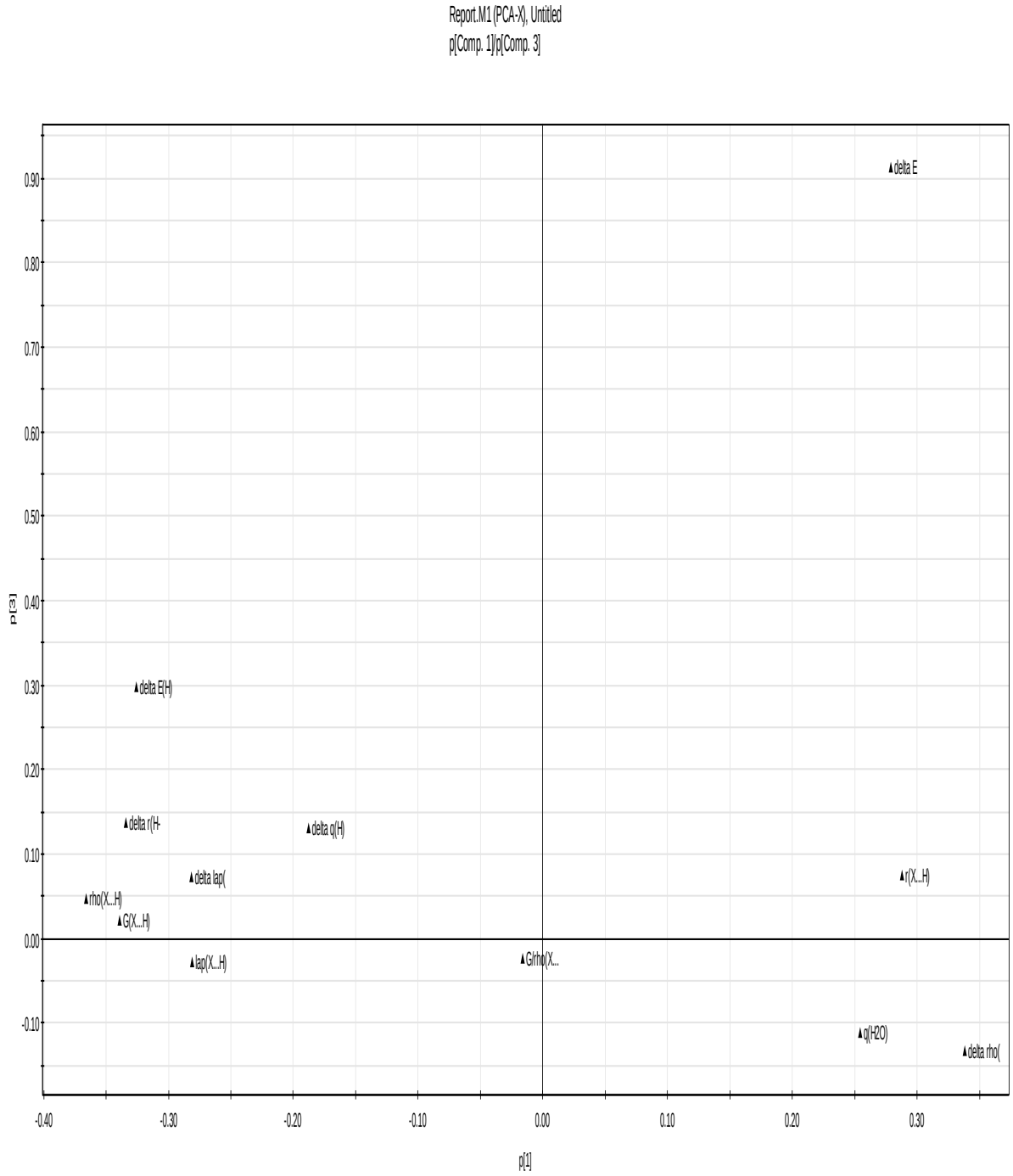


Figure 3.1.4. Scatter plot of the loadings of the first, p[1] and second, p[2] principal components. The properties displayed here have abbreviated descriptions. See table 3.1.2 for details.

Table 3.1.3 shows the R^2 values from the linear regression analysis for correlations of each property correlated with pK_{BHX} values.

Property	**Abbreviation**	**R^2**
r(X…H)	r(X…H)	0.4542
Δr(H-O)	delta r(H-O)	0.7138
ΔE	delta E	0.8752*
q(H2O)	q(H2O)	0.4154
Δq(H)	delta q(H)	0.2862
ΔE(H)	delta E(H)	0.8905
ρ(X…H)	rho(X…H)	0.7557
lap ρ(X…H)	lap(X…H)	0.4156
Δρ(H-O)	delta rho(H-O)	0.7323
Δlap ρ(H-O)	delta lap(H-O)	0.4598
G(X…H)	G(X…H)	0.6257
G/ρ(X…H)	G/rho(X…H)	0.000686

Table 3.1.3. Correlation results for each property with pK_{BHX} values. * - result obtained with the omission of aniline from the data set.

Table 3.1.3 reveals two possible properties that could be used in a model to predict hydrogen bond basicity. The change in atomic energy of the hydrogen atom acting as a hydrogen bond donor gives the best correlation with pK_{BHX} with an R^2 of 0.8905. The hydrogen bond energy itself correlates well with pK_{BHX} with an R^2 of 0.8752. Figure 3.1.3 shows how these properties load highly onto the first principal component.

This strengthens the idea that the first principle component may be linked to basicity. Also the kinetic energy density divided by the electron density at the hydrogen bond critical point contributes very little to the first principal component, and has an R^2 value close to zero when correlated with pK_{BHX}. The hydrogen bond energy loads highly onto the third principal component. The scoring plot of the third principal component shown in figure 3.1.4 reveals aniline to be an outlier. The correlation of hydrogen bond energy with pK_{BHX} also reveals aniline to be an outlier and for this reason it has been omitted from the results.

Other properties such as the electron density at the hydrogen bond critical point and the change in electron density at the hydrogen bonded hydrogen and oxygen critical point in the HBD give reasonable correlations with pK_{BHX}. However, they are not strong enough correlations to allow for a reliable predictive model of hydrogen bond basicity.

It can be concluded that both the change in atomic energy of the hydrogen atom acting as a hydrogen bond donor and the hydrogen bond energy could be used to model hydrogen bond basicity. However, as the BSSE must be considered the hydrogen bond energy calculation is time consuming and therefore probably not a viable option. The calculation of the atomic energy on the hydrogen bond donor site involves an inexpensive integration over the hydrogen's atomic basin. The change in atomic energy of the hydrogen atom acting as a hydrogen bond donor is therefore the most likely candidate for modelling hydrogen bond basicity. It could also be concluded that it may well be necessary to use properties of a hydrogen bonded complex rather than properties of the free base to predict hydrogen bond basicity. This is because the properties that correlate best with pK_{BHX} involve calculations on the HBD.

3.2 Computation Of Hydrogen Bond Basicity pK_{BHX}.

The results in section 3.1 showed how the property of those listed in table 3.2 correlate best with pK_{BHX} is ΔE(H). Following on from this finding, the data set was increased and the procedure outlined in section 3.1 repeated. The increased data set is listed in table 3.2.1. The increased data set of hydrogen bond acceptors was used to compute properties from hydrogen bond complexes formed with water, methanol, 4-fluorophenol, serine OH side chain, methylamine and hydrogen fluoride HF. An example of a hydrogen bond involving one of the hydrogen bond acceptors listed in table 3.2.1 and each of the hydrogen bond donors is shown in figure 3.2.1.

Hydrogen Bond Acceptor	**pK_{BHX}**
3-Chloropyridine	1.31
4-Methylpyridine	2.07
Acetamide	2.06
Acetone	1.18
Acetonitrile	0.91
Acrylonitrile	0.70
Aniline	0.46
Chloroacetnitrile	0.39
Dimethyl Sulfide	0.12
Dimethylamine	2.26
Ethanol	1.02
Ethyl Thiol	-0.16
Ethylamine	2.17
Formamide	1.75
MeSCN	0.73
Methanol	0.82
Methyl Acetate	1.00
Methylamine	2.2
Methylformate	0.65
N-Methylaniline	0.26
Phenol	-0.07
Pyridine	1.86
Pyrrolidine	2.59

t-Butylamine	2.23
Tetrahydropyran	1.23
2,6-Dimethylaniline	0.47
3-Chloroaniline	0.13
3-Fluoroaniline	0.20
3-Methylphenol	0.01
4-Methylphenol	0.03
Dimethyldisulfide	-0.49
Ethylmethylsulfide	0.18
p-Toluidine	0.56
4-Fluorophenol	-0.12
4-Bromo-N,N-dimethylaniline	-0.42
Oxydibenzene	-0.80
Diethyl Disulfide	-0.40
4-Aminopyridine	2.56
N,N,N-Trimethyl Ammoniopropanamidate	3.59
Triethylarsine Oxide	4.89
Trimethylamine Oxide	5.46

Table 3.2.1. The training set of bases to be used as hydrogen bond acceptors, and their pK_{BHX} values [1].

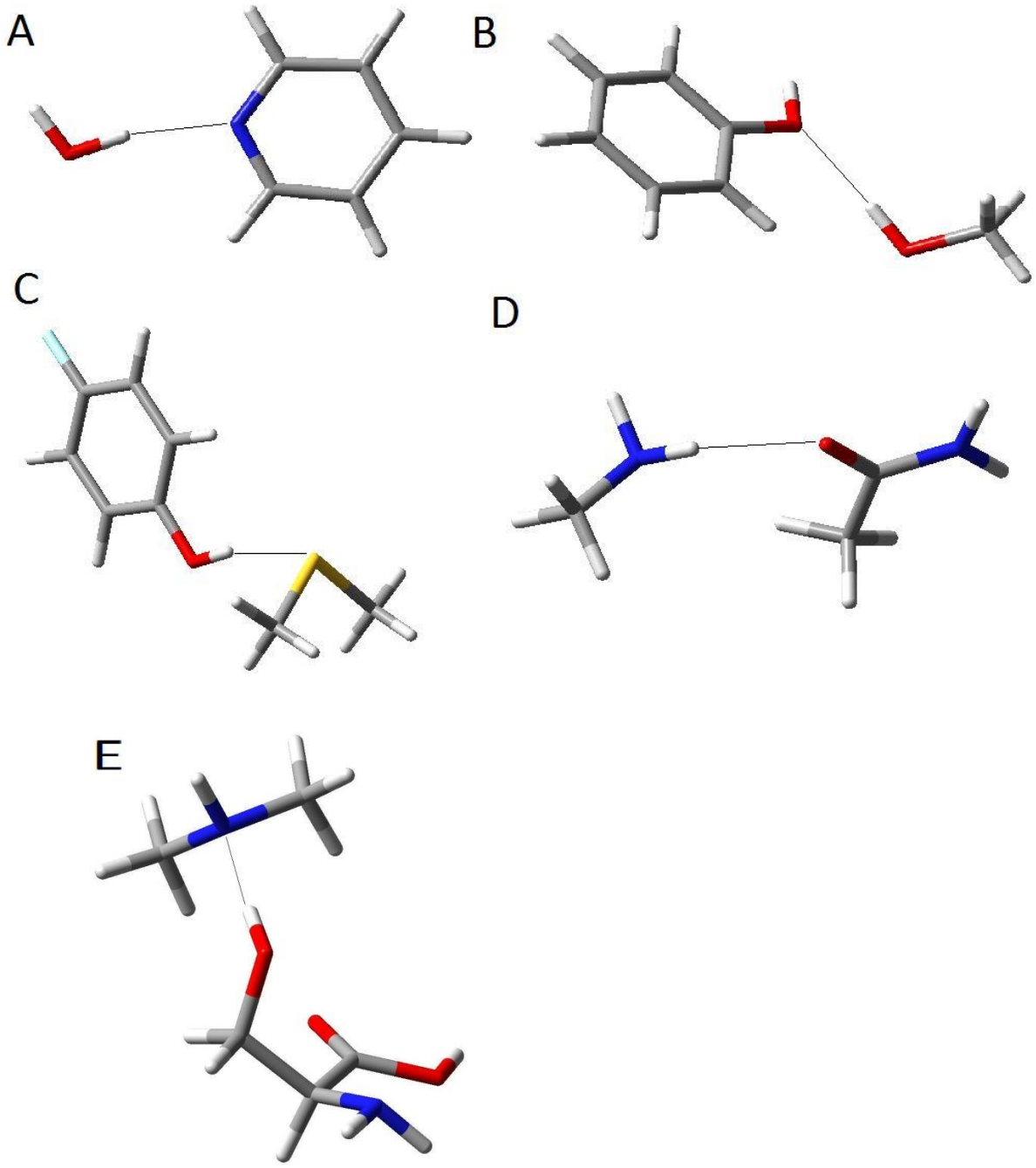


Figure 3.2.1. Hydrogen bonded complexes. A – water-pyridine, B – methanol-phenol, C – 4-fluorophenol-dimethyl sulphide, D – methylamine-acetamide and E – serine-dimethyl amine.

Hydrogen bond complexes were generated and properties obtained following the procedures outlined in section 2. The properties used are described in detail in section

3.1. The results of the correlation between the property and the pK_{BHX} value are listed in table 3.2.2 for each hydrogen bond donor.

	HYDROGEN BOND DONOR					
Property	**Water**	**Methanol**	**4-fluorophenol**	**Serine**	**Methylamine**	**Hydrogen Fluoride**
r(X…H)	0.46	0.49	0.49	0.46	0.52	0.41
Δr(H-O)	0.80	0.81	0.80	0.76	0.87	0.38
ΔE	0.32	0.33	0.52	0.29	0.048	0.60
Δq(H)	0.58	0.56	0.57	0.64	0.81	0.00
ΔE(H)	0.96 (0.97)	0.95	0.91	0.93	0.97	0.04
ρ(X…H)	0.81	0.86	0.88	0.79	0.90	0.74
$\nabla^2\rho$(X…H)	0.64	0.68	0.55	0.58	0.83	0.10
Δρ(H-X)	0.81	0.82	0.81	0.78	0.83	0.73
$\Delta\nabla^2$ ρ(H-X)	0.59	0.51	0.65	0.62	0.66	0.65
G(X…H)	0.80	0.84	0.83	0.79	0.83	0.71
G/ρ(X…H)	0.09	0.09	0.06	0.08	0.11	0.01

Table 3.2.2. Using table 3.2.1, the listed properties of each complex have been plotted against their corresponding pK_{BHX} value. The resulting R^2 values are displayed. The ΔE(H) value in brackets in the water set are taken from MPWB1K calculations. A slightly reduced data set of 35 HBAs was used for MPWB1K calculations. The data set is as table 1, minus 3-Chloropyridine, Dimethyl sulfide, Ethyl thiol, Ethylmethylsulfide, 4-Fluorophenol and Diethyldisulfide. These HBAs were omitted due to computational timing constraints. As a direct comparison, r^2 = 0.97 for ΔE(H) when the reduced data set is used for B3LYP calculations.

Table 3.2.2 summarises the correlations of calculated molecular properties with the experimental pK_{BHX} values. With the exception of the HF set, the ΔE(H) consistently outperforms all other properties, giving R^2 values of 0.96 for water, 0.95 for methanol, 0.91 for 4-fluorophenol, 0.93 for serine and 0.97 for methylamine. The plots of ΔE(H) against pK_{BHX} for water, methanol, 4-fluorophenol, serine and methylamine are shown in figure 2. However, pK_{BHX} values did not correlate with ΔE(H) when HF was used as the HBD, R^2 = 0.04. It is claimed that the pK_{BHX} scale is applicable to hydrogen bonded complexes, given that the HBD is of OH, or strong NH origin [1]. Therefore HF was chosen as a control probe as hydrogen bonded complexes where HF is the HBD have no link to the experimental pK_{BHX} scale. The relationship between ΔE(H) and pK_{BHX} values

gains credibility as it seems to hold only for HBDs that the pK_{BHX} scale may be applied to. Therefore a strong link between theory and experiment has been established through the partitioned atomic energy of the HBD. Essentially, the energy assigned to a theoretically derived atomic basin is related to a molecular free energy quantity that is derived from experiment. The atomic basin of a hydrogen atom in a hydrogen bond complex has been generated using MORPHY98 [152] and is shown in figure 3.2.2.

Table 3.2.1 shows the full list of all properties calculated. Particular attention was given to the electronic energy of the hydrogen bond ΔE. However poor correlations across the board of OH and NH donors were observed. The BSSE corrected energy offered no improvement to the correlations. However the plot of ΔE against pK_{BHX} where HF was used as the probe gave a correlation of 0.60. Although 0.60 is a weak correlation, it is significantly better than any of the values obtained for OH and NH donors. Lamarche and Platts [153] obtained a value or $R^2 = 0.91$ for the correlation of computed ΔE with Taft's pK_{HB} values using HF as the HBD. Taking into account the stronger correlations of ΔE with pK_{BHX} and the observations of Lamarche and Platts when HF is used as the HBD, it could be said that any relationship between ΔE and basicity is false due to the use of HF as a probe. Indeed the computed ΔE of a HF complex correlates reasonably well which basicity values but a link cannot be made between theory and experiment as the basicity scales are not applicable to HF complexes.

The MPWB1K density functional has been shown to perform better than B3LYP when computing the interaction energies of hydrogen bonded complexes [147-149]. However, as the binding energies of the complexes used here do not correlate with pK_{BHX} values, it seems to be adequate to use B3LYP. Also table 3.2.1 shows how $\Delta E(H)$ calculated from a MPWB1K wave function performs equally as well as $\Delta E(H)$ calculated from a B3LYP wave function when correlating $\Delta E(H)$ to pK_{BHX}.

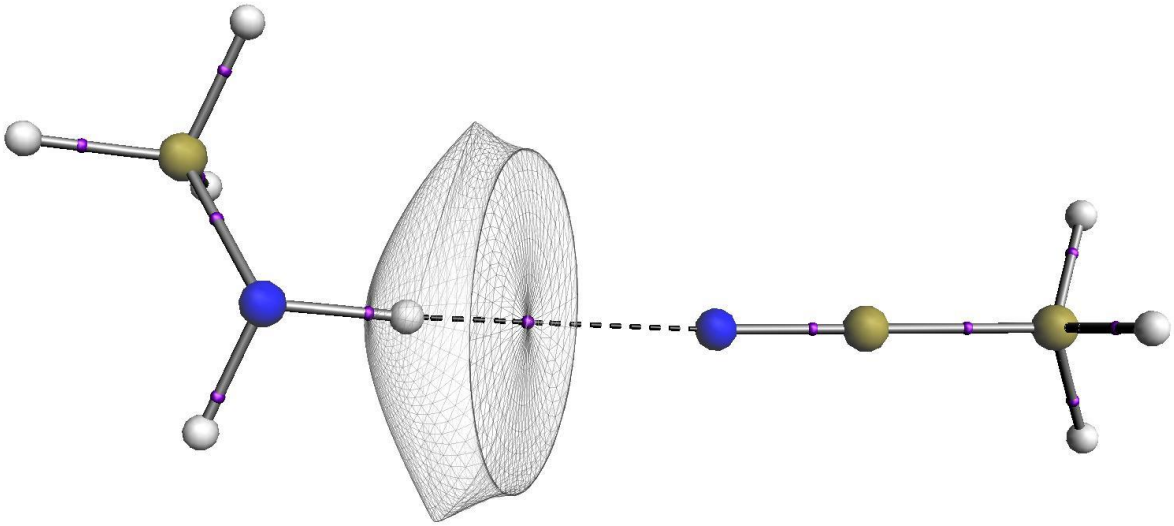


Figure 3.2.2. Quantum chemical topology of the acetonitrile-methylamine complex. The purple dots are the bond critical points and the area highlighted around the HBD is its atomic basin where the electron density has been cut off at $\rho=10^{-3}$ at the outer surfaces. The atomic basin is integrated over to give ΔE(H).

Plots displaying the correlation between computed ΔE(H) and experimental pK_{BHX} values are shown in figure 3.2.3.

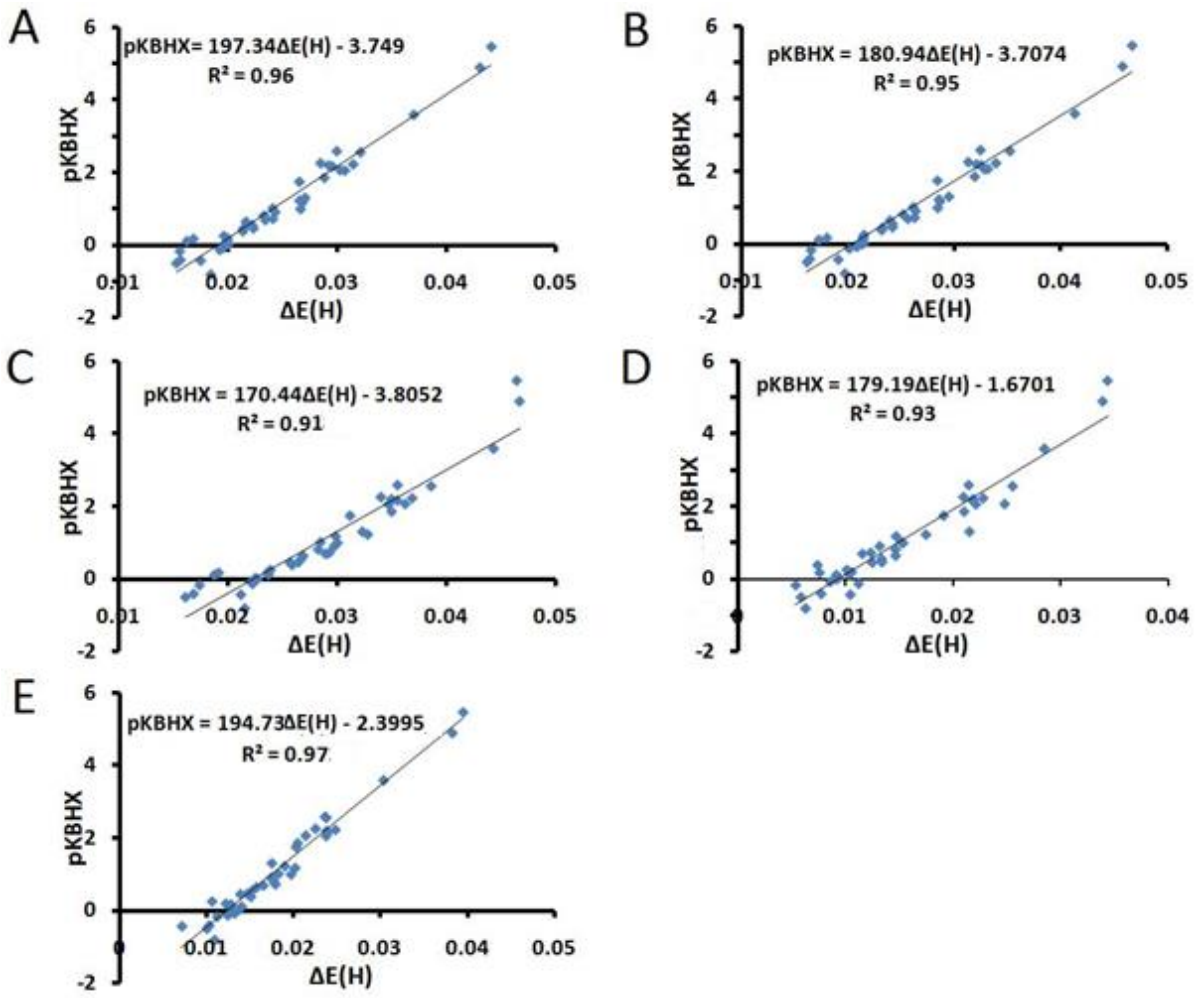


Figure 3.2.3. Correlation of the QCT ΔE(H) with pK_{BHX} values for the bases listed in table 3.2.1 with hydrogen bond donors, A – water, B – methanol, C – 4-fluorophenol, D – serine and E – methylamine.

The results in figure 3.2.3 and table 3.2.2 show that when methanol is used as a hydrogen bond donor, the r^2 value of correlations between pK_{BHX} and ΔE(H) is 0.95. When both water (0.96) and methylamine (0.97) are used as hydrogen bond donors, the r^2 values are slightly higher. However the pK_{BHX} database has been set up using 4-fluorophenol and methanol as hydrogen bond donors. Methanol gives a better correlation compared to 4-fluorophenol when used as a hydrogen bond donor to obtain computed ΔE(H) values. The r^2 value for correlations between pK_{BHX} and ΔE(H) when 4-fluorophenol is used as the hydrogen bond donor is 0.91 whereas methanol returns an

improved 0.95 r^2 value. Therefore methanol will be used as the hydrogen bond donor of choice for the rest of this research unless stated otherwise.

Using methanol as the hydrogen bond donor, hydrogen bond complexes were formed and ΔE(H) values computed following the procedures in section 2 for the hydrogen bond acceptors listed in table 3.2.3. The hydrogen bond acceptors listed in table 3.2.3 have formed a test data set. The test is to see if predicted pK_{BHX} values based on the straight line equation of figure 3.2.3 B correlate with actual database pK_{BHX} values. The straight line equation of figure 3.2.3 B is given again in equation 3.2.1

$$pK_{BHX} = 180.94\Delta E(H) - 3.7074 \qquad (3.2.1).$$

se = 0.0017, *rms* = 0.028, *F* = 33794.73.

HYDROGEN BOND ACCEPTOR	pK_{BHX}	pK_{BHX} Predicted	Error
4-Chloropyridine	1.54	1.84	-0.30
4-Ethylpyridine	2.07	2.30	-0.23
Benzaldehyde	0.78	0.92	-0.14
Benzylamine	1.84	2.09	-0.25
Butan-2-one	1.22	1.53	-0.31
Cyanamide	1.56	1.37	0.19
Dimethylformamide	2.1	1.91	0.19
Isopropanethiol	-0.1	-0.66	0.56
Isopropylamine	2.2	2.32	-0.12
Propanol	1	1.04	-0.04
Propionitrile	0.93	1.22	-0.29
Propiophenone	1.04	1.73	-0.69
trichloroacetonitrile	-0.26	-0.07	-0.19
2,4,6-Trimethylpyridine	2.29	2.59	-0.30
2-Aminopyridine	2.12	3.11	-0.99
2-Chloropyridine	1.05	1.17	-0.12
3-Aminopyridine	2.2	2.28	-0.08
4-Chloro-N-Methylaniline	0.05	-0.03	0.08
5-Bromopyrimidine	0.59	1.02	-0.43
Ammonia	1.74	1.95	-0.21
Anisolel	-0.07	0.22	-0.29
Benzylmethylsulfide	-0.02	-0.30	0.28
Isoquinoline	1.94	2.17	-0.23
N-Methylbenzamide	2.03	2.45	-0.42
N-Methylpropanamide	2.24	2.60	-0.36
1,3,5-Triazine	0.32	0.74	-0.42

4,6-Dimethylpyrimidine	1.47	2.00	-0.53
4-Chlorobutyronitrile	0.83	1.03	-0.20
Cyanicbromide	0.19	0.44	-0.25
Cyclooctanone	1.45	1.65	-0.20
Diisopropylether	1.11	1.69	-0.58
Ethyl Chloroacetate	0.67	1.03	-0.36
Ethylenesulfite	0.87	1.25	-0.38
Isoxazole	0.81	1.11	-0.30
N-Formylmorpholine	1.93	1.83	0.10
Phenylcyanate	0.77	1.04	-0.27
Phthalazine	1.97	2.11	-0.14
Pyridazine	1.65	1.77	-0.12
Pyrrolidine-1-Carbonitrile	1.66	2.07	-0.41
Progesterone	1.75	2.16	-0.41
Diphenylphosphinic Chloride	2.17	2.16	0.01
		Mean Absolute Error = -0.223	

Table 3.2.4. The test set of bases to be used as hydrogen bond acceptors, and their pK_{BHX} values [1].

It is found that $\Delta E(H)$ is able to predict the pK_{BHX} of an external data set. Figure 3.2.4 shows the plot of predicted against actual pK_{BHX} values for the bases listed in table 3.2.4. Predicted pK_{BHX} values are calculated from equation 3.2.1, the equation for the line of the plot of $\Delta E(H)$ against pK_{BHX} where methanol is the HBD, complexed to bases in table 3.2.1. Methanol was chosen as the probe to use in the model due to its simplicity, its use in obtaining secondary LFER derived experimental values and the strong correlation, $R^2 = 0.95$ between the $\Delta E(H)$ of these complexes and pK_{BHX} values.

There is a relatively strong correlation, $R^2 = 0.90$ for the plot of predicted against experimental pK_{BHX} values. This demonstrates impressive extrapolation of the original model. The only constraint put on the molecules in table 3.2.3 is that the HBA must be either an oxygen, nitrogen or sulphur atom. Not only does the model based on $\Delta E(H)$ show impressive external predictability, but also good generality. This is because all HBAs can be grouped together in one data set. There seems to be no need to separate the HBAs into atomic groups and build individual models for each HBA type.

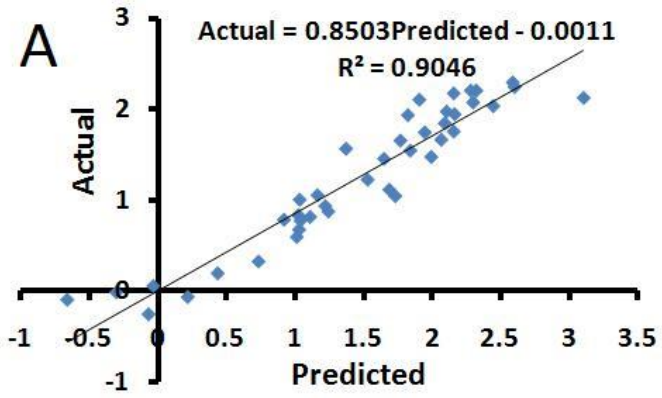


Figure 3.2.4. Correlation of the predicted pK_{BHX} values with actual pK_{BHX} values for the bases listed in table 3.2.3. The common HBD is methanol. Predictions are based on the straight line equation for methanol complexes (figure 3.2.3 B).

The results discussed so far are for 1:1 HBD:HBA complexes where the bases have only one major HBA site. The attractiveness of the pK_{BHX} scale to the medicinal chemist is that it includes data for the more realistic cases of polyfunctional bases with two or more major HBA sites. Table 3.2.4 shows a list of 11 bases, each with two major HBA sites. Two approaches were used to test the validity of the $\Delta E(H)$ model for polyfunctional bases. The first method simply involved generating a 1:1 complex for each HBA site. The second aimed to mimic experiment in which the bases is surrounded by an excess of acid allowing each HBA site to come to equilibrium with a HBD. The way in which the experiment was mimicked was to generate 2:1 HBD:HBA complexes where both HBA sites are occupied in the same calculation.

Hydrogen Bond Acceptor	**pK_{BHX}**
Methyl Nicotinate (N)	1.44
Methyl Nicotinate (O)	0.51
3-Benzoyl Pyridine (N)	1.42
3-Benzoyl Pyridine (O)	0.68
4-Acetylpyridine (N)	1.41
4-Acetylpyridine (O)	0.78
Ethyl 4-cyanobenzoate (N)	0.66
Ethyl 4-Cyanobenzoate (O)	0.53
S-Cotinine (N)	1.62

S-Cotinine (O)	2.16
4-Acetylbenzonitrile (N)	0.65
4-Acetylbenzonitrile (O)	0.6
3-Acetylpyridine (N)	1.39
3-Acetylpyridine-(O)	0.9
N,N-Diethylnicotinamide (N)	1.63
N,N-Diethylnicotinamide (O)	1.98
4-Cyanopyridine (Pyr)	0.92
4-Cyanopyridine (Nit)	0.47
3-Cyanopyridine (Pyr)	0.82
3-Cyanopyridine (Nit)	0.53
2-Cyanopyridine (Pyr)	0.48
2-Cyanopyridine (Nit)	0.61

Table 3.2.4. The set of polyfunctional bases to be used as hydrogen bond acceptors, and their pK_{BHX} values [1]. The symbol in brackets indicates the HBA site where (N) is a nitrogen atom, (O) an oxygen atom, (Nit) a nitrogen atom on a nitrile functional group and (Pyr) a nitrogen atom on a pyridine functional group.

Figure 3.2.5 shows the results of the two methods. The first method in which 1:1 complexes were generated marginally out performs the second method in which 2:1 complexes were computed. This may be surprising as 2:1 complexes are closer to the experimental method, however it is beneficial to the theoretical chemist as 1:1 complexes are computationally less expensive than the 2:1 complexes to compute.

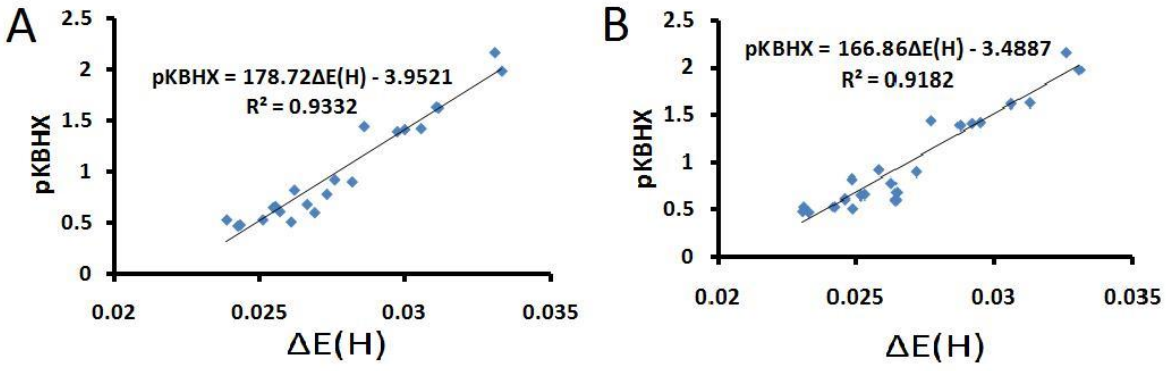


Figure 3.2.5. Correlation of the QCT ΔE(H) with pK_{BHX} values for the polyfunctional bases listed in data table 3.2.4 with the common HBD methanol. A – 1:1 complex, B – 2:1 complex.

3.3 – Effects Of Fragmentation On pK_{BHX} Computation.

The methods above describe how to obtain estimated pK_{BHX} values from computed properties. The method has been executed and tested for a series of small hydrogen bonded complexes. A property has been obtained, ΔE(H), that accurately correlates to database pK_{BHX} values. Unfortunately in order to obtain ΔE(H) it is necessary to obtain a wave function for geometrically optimised hydrogen bonded complexes. The complexes formed by drug–protein interactions are much larger than the ones described in the following section. The consequence of this is to increase the time it takes to compute a wave function. It is therefore necessary to reduce the time of the computation without reducing the accuracy of the pK_{BHX} calculation.

Fragment based drug design has gained popularity in recent years [154-172]. The method involves chopping up a large molecule into smaller fragments, the properties of which can be summed to give the molecular properties of the total molecule [173]. Therefore by chopping a ligand into fragments containing hydrogen bond sites, it should be possible to gain hydrogen bonding information about the whole ligand. When fragmenting molecules, the usual procedure is to leave the rings intact by chopping off all groups attached [154]. This is called the ring cleavage method. The ring cleavage method produces a series of rings, ring substituents – functional groups attached to the rings - and linkers – chains of atoms joining rings together.

The ring cleavage method will be applied as a starting point to fragment ligands. However, complications still arise when trying to apply this method as the fragments generated need to be capped appropriately. Also the fragments generated may still be too large to be used for simple calculations.

In order to analyse the sensitivity of hydrogen bonding to the proximity of functional groups a series of computations will be performed. Acetone and isopropylamine will form hydrogen bond complexes with water. DFT calculations were performed using Gaussian 03 [142]. Hydrogen bonded complexes were generated by placing a hydrogen atom of the water molecule approximately 2.0 Å away from the

primary hydrogen bond acceptor site, and following the guidelines outlined above in section 2.2. Atomic integrations used to generate ΔE(H) were calculated with the AIMALL suite [143]. The experiments will then be sequentially repeated by placing the functional group a further carbon atom away from the HBA each time. In other words, the carbon chain joining the hydrogen bond acceptor to the functional group is extended by one atom until the pK_{BHX} value equals the initial values calculated for acetone and isopropylamine. The functional groups tested are shown for acetone in figure 2.3.1.

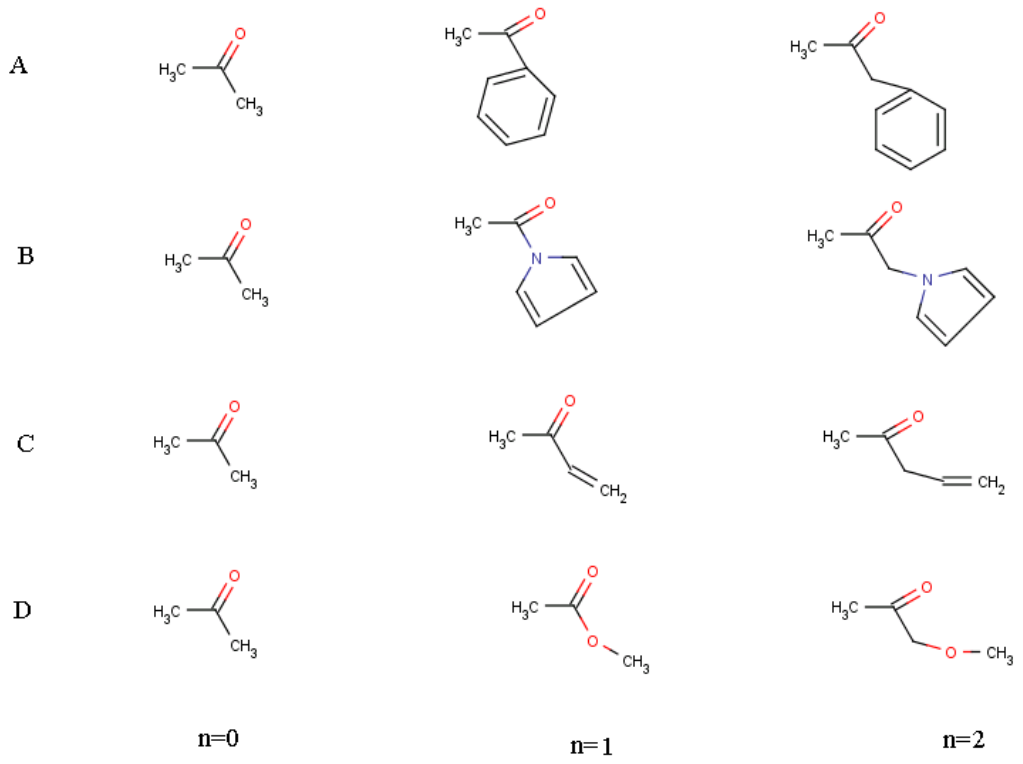


Figure 3.3.1. Substitutions made to acetone in order to assess the sensitivity of hydrogen bond basicity to the distance from a functional group. n refers to the number of carbon atoms added to the chain between the HBA and the functional groups A – benzene, B – pyrrole, C – Ethene and D – Ether.

Table 3.3.1 shows the results of how the calculated pK_{BHX} values of acetone and isopropylamine change as methyl groups are substituted for electron withdrawing functional groups.

It can be concluded that the carbon chain linking the functional group to the HBA need be extended by no more than one atom (n=2), in order for the calculated pK_{BHX} value to return to that of the unsubstituted base. Therefore it can be concluded that fragments may be capped with methyl if the functional group is two or more atoms away from the HBA.

The pK_{BHX} database itself has also been analysed to assess the influence of alkyl chains on basicity. The pK_{BHX} values of primary amines appear to be insensitive to the size of the alkyl chain. Methyl amine has a value of 2.2, ethylamine 2.17, propylamine 2.2 and hexadecylamine 2.26.

It is reasonable to suggest that linker groups containing long carbon chains may be represented by one terminal methyl group.

	Acetone			Isopropylamine		
	n			n		
Substitution	0	1	2	0	1	2
Benzene	1.54	1.72	1.54	2.29	2.21	
Ethene	1.54	1.59		2.29	2.12	2.20
Ether	1.54	1.50	1.55	2.29	1.72	2.48
Pyrrole	1.54	1.51		2.29	1.78	2.20

Table 3.3.1. The pK_{BHX} values are displayed for various substitutions of acetone and isopropylamine. Details are given in the text and visualised in figure 3.3.1.

3.4 – Effects Of Level Of Theory On pK_{BHX} Computation.

It is also possible to speed up the geometry optimisation step of the hydrogen bonded complexes. Two methods of achieving faster optimisation have been tested. The first method uses a lower level Hartree-Fock calculation, HF/6-31G(d). The second method relaxes the force and displacement convergence threshold values used to determine optimisation. A set of 21 HBAs evenly spanning the range of the pK_{BHX} scale have been selected from data sets used in previous sections. The bases used here are

listed in table 3.4.1. Using methanol as the HBD, and the methods outlined in the previous section, ΔE(H) values have been obtained for both B3LYP and HF wave functions using both the default and relaxed convergence criteria.

Hydrogen Bond Acceptor	pK_{BHX}
Ethyl thiol	-0.16
Phenol	-0.07
4-methylphenol	0.03
Dimethyl sulfide	0.12
Aniline	0.46
p-toluidine	0.56
Methylformate	0.65
Acrylonitrile	0.7
Methanol	0.82
Acetonitrile	0.91
Ethanol	1.02
Acetone	1.18
4-Acetylpyridine (N)	1.41
Pyridazine	1.65
Pyridine	1.86
Acetamide	2.06
Ethylamine	2.17
t-Butylamine	2.23
Pyrrolidine	2.59
Triethylarsine oxide	4.89

Table 3.4.1. The set of bases to be used as hydrogen bond acceptors, and their pK_{BHX} values [1]. The (N) after 4-Acetylpyridine indicates that the HBA site is a nitrogen atom on as 4-Acetylpyridine is a polyfunctional base with an additional HBA site on the oxygen atom.

Figure 3.4.1 shows that the correlation of pK_{BHX} with ΔE(H) is sensitive to DFT correlation methods but not at all sensitive to a slight relaxation of the force and displacement convergence thresholds. By relaxing the optimisation convergence criteria, it is therefore to speed up the optimisation step in the process of computing hydrogen bond basicity.

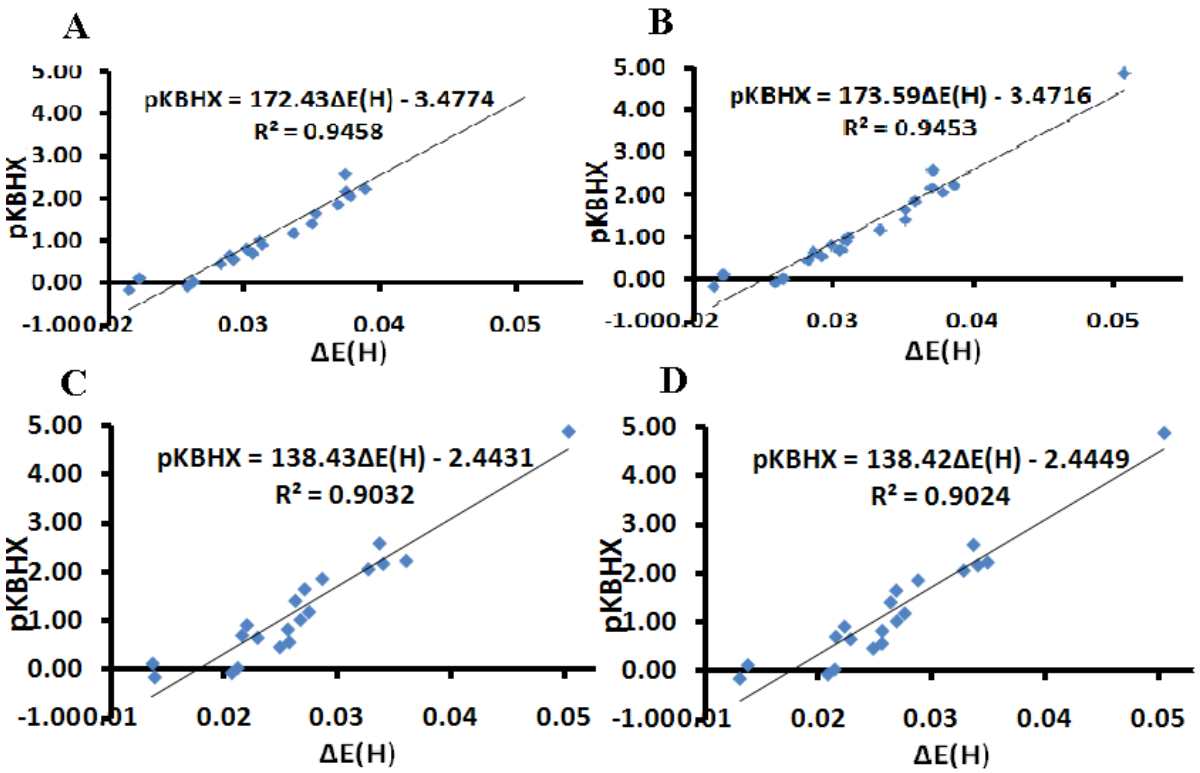


Figure 3.4.1. Correlation of the QCT ΔE(H) with pK_{BHX} values [1] for the bases listed in table 3.4.1 with the common HBD methanol. A – B3LYP default convergence. B - B3LYP relaxed convergence, C - HF default convergence, D – HF loose convergence.

3.5 – Computation Of Drug Binding Data.

Drug binding affinity is measured in terms of equilibrium constants. The equilibrium constant that relates to a drug target interaction is called *Ki*. The binding affinity of a drug is a thermodynamic measure of a drugs interaction with its target. The binding affinity value offers no information about the ability of the drugs ability to induce a response in the target once bound. The ability of the drug to induce a response in its target once bound is called efficacy [174]. The efficacy or intrinsic activity as it is also known is a relative measure of the strength of the response that the drug induces once bound to the target. Some drugs are more efficacious than others as they produce a

stronger response by occupying the same number or a lower number of receptors. The most potent drugs occupy the same number of receptors at lower concentrations.

For candidate lead compounds to become potential drugs they must have a high binding affinity, be efficacious and potent. Efficacy and potency are beyond the scope of this research. The following text will describe methods in which an attempt will be made to predict biding affinity through computed predicted hydrogen bond basicity values.

In order to estimate Ki binding values of hydrogen bonded complexes it is necessary to firstly obtain a reliable pK_{BHX} values for each HBA in the complex. Equations 1.4.13, 1.4.14 and 1.4.15 show how Ki values are calculated from pK_{BHX} values

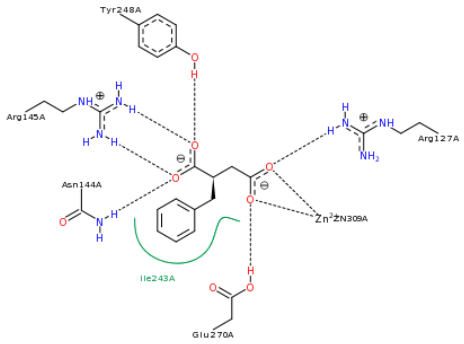
The experiments will be set up in order to reduce computational time. Water is chosen as the hydrogen bond donor for the purpose of reducing the computational time of the optimisation step. Water is the smallest hydrogen bond donor that the pK_{BHX} scale can be applied to. The methods used to compute optimised hydrogen bonded complexes in which water is the hydrogen bond donor have been described in section 2. At this point it is necessary to refer to figure 3.2.3 A from section 3.2. It has been shown in section 3.2 that ΔE(H) correlates well with pK_{BHX}. By taking the equation 3.2.1 for the correlation between pK_{BHX} for the bases in table 3.2.1, and the computed ΔE(H) values, equation 1.4.15 becomes equation 3.5.1

$$K_{BHX} = 10^{-(197.34\Delta E(H)-3.749)} \qquad (3.5.1).$$

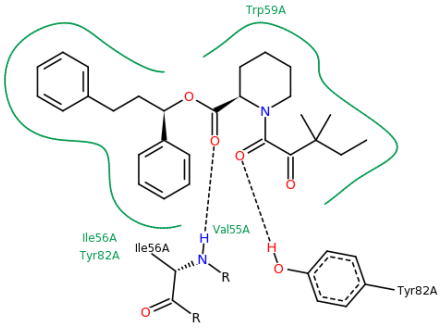
Ki is a dissociation constant that relates to reversible binding. In many drug interactions Ki is a total binding property of numerous interactions. It is therefore necessary to obtain Ki from equation 1.4.10

In this section an attempt will be made to compute Ki values via pK_{BHX} values for real drug-protein interactions. The original GOLD validation set [175] will serve as the source of protein-ligand complexes used in this work. Only complexes with Ki values

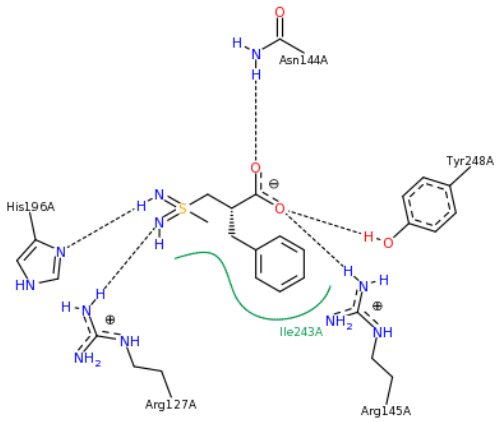
entered into the Binding MOAD [176], will be used in this test set. The following diagrams (figure 3.5.1) show the drug interactions obtained from the protein data bank (PDB) [177] that will form the test set for this study. Dashed black line indicate hydrogen bonds, solid green lines show hydrophobic interactions.



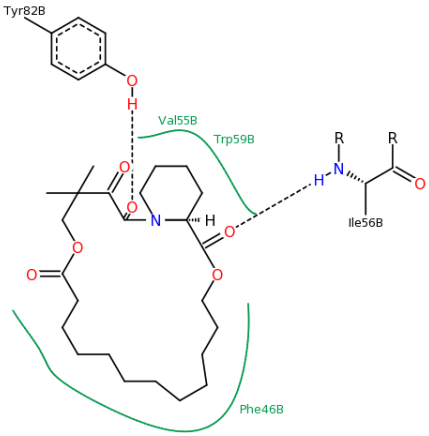
1CBX-BSZ



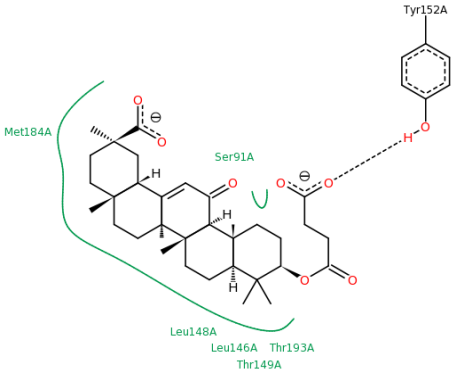
1FKG-SB3



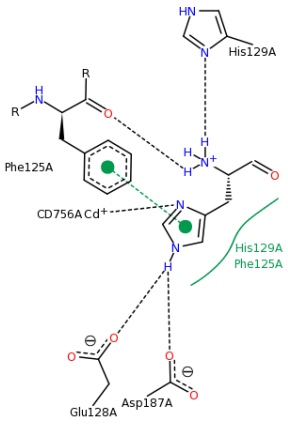
1CPS-CPM



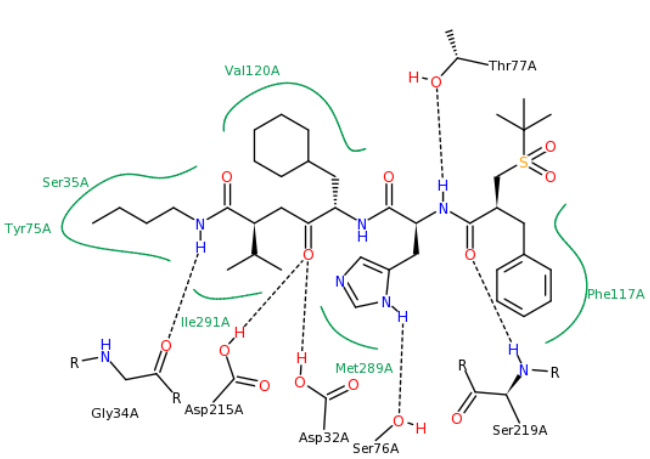
1FKI-SB1



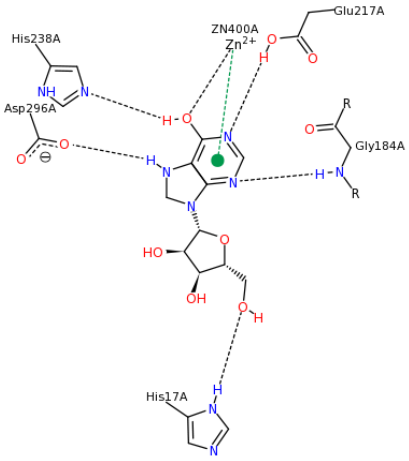
1HDC-CPO



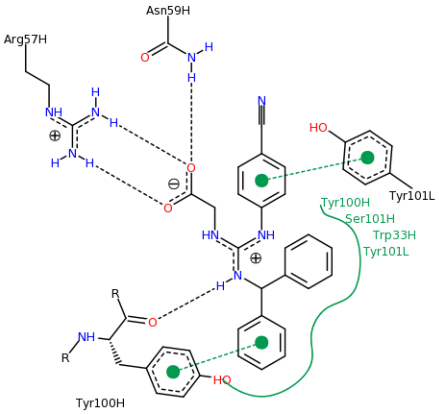
1HSL-HIS



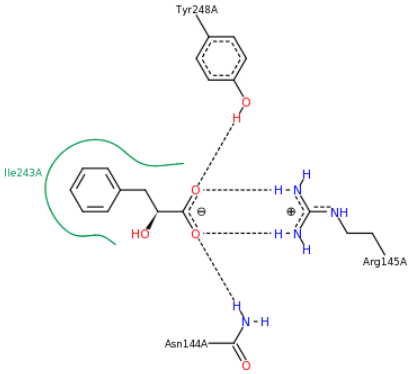
1RNE-C60



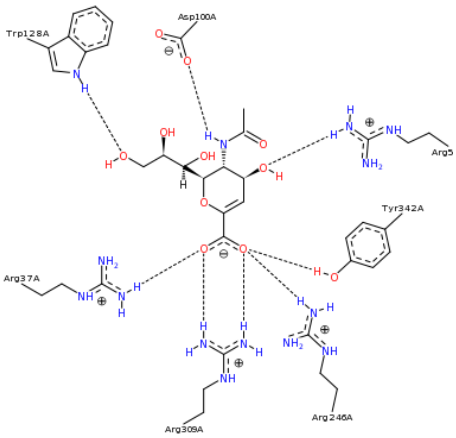
2ADA-HP3



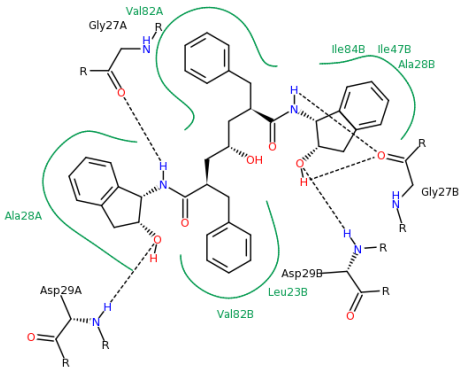
2CGR-GAS



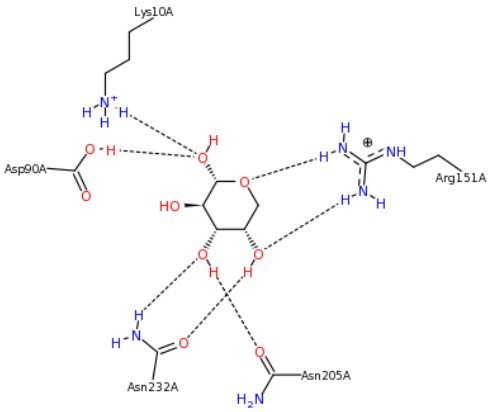
2CTC-HFA



2SIM-DAN



4PHV-VAC



6ABP-ARA

Figure 3.5.1. Drug binding interactions taken from the protein data bank. The interactions shown here will form the test set for this section.

It is desirable to fragment large molecules into smaller sections to avoid lengthy computation. Once again, to continue to understand the methods used in this section, it is necessary to state some important findings from section 3.3. The results of the experiments that test the sensitivity of the pK_{BHX} database to various functional groups show how the pK_{BHX} database is not particularly sensitive to the proximity of electron withdrawing groups. Given the finding that pK_{BHX} values are not particularly sensitive to the proximity of electron withdrawing groups, it is possible to outline a method for fragmenting the test set of hydrogen bonded complexes shown above.

Aromatic rings and ring substituents will remain intact. Non-aromatic rings can be thought of as carbon chains, unless a non aromatic ring is bonded to an aromatic ring, in which case they will remain intact. Carbon chains may be adequately capped with methyl groups in order to shorten the chain. Methyl groups may be used to cap carbon chains that have a functional group housed within the chain that is two or more atoms away from a HBD. If functional groups are closer to a HBA, then they must be retained in the fragment. Figure 3.5.2 shows the fragments generated by applying these methods to the test set. Table 3.5.1 shows the number of times each HBA of the fragments shown in figure 3.5.2 occurs in the full interactions shown in the diagrams above in figure 3.5.1. Using water as a HBD and the methods outlined in section 2.2, a pK_{BHX} and a K_{BHX} have

been computed for each of the fragmented HBA sites shown in figure 3.5.2. The results are shown in table 3.5.2. The actual *Ki* values have been extracted from the Binding MOAD [176] and compared with the computed values. Results are shown in table 3.5.3.

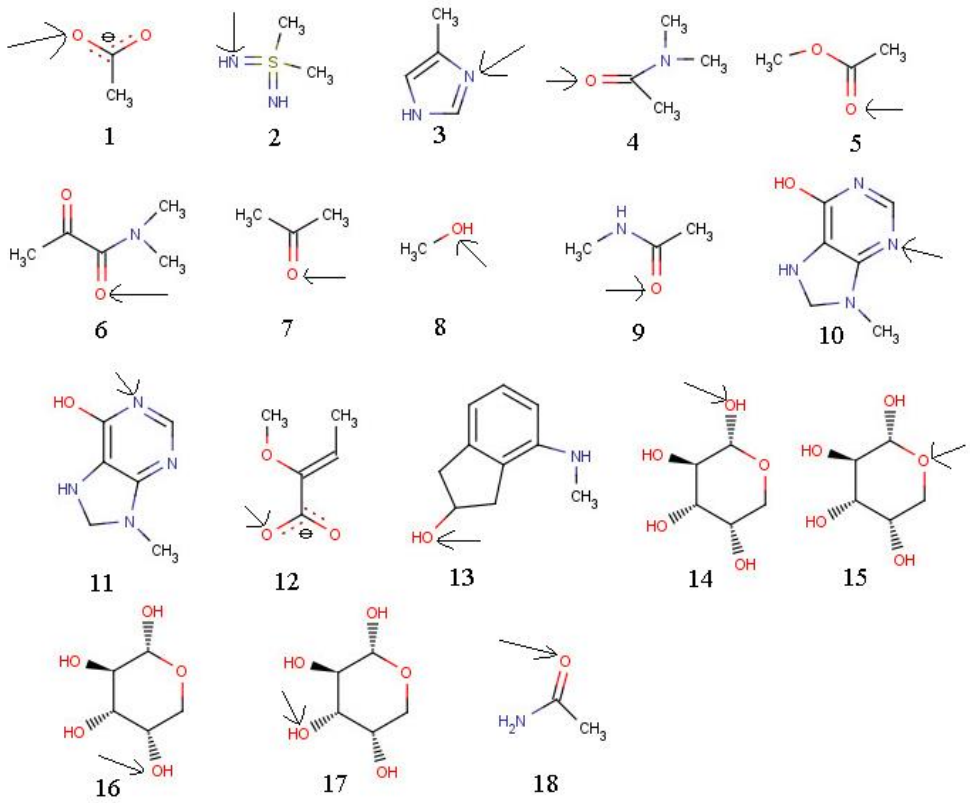


Figure 3.5.2. Fragments generated from the test set of hydrogen bonded complexes shown above. Details of how the fragments have been generated are given in the text. The arrow points to the HBA.

Complex	Frequency of HBA
1CBX-BSZ	4 x 1
1CPS-CPM	2 x 1, 2, 3
1FKG-SB3	4, 5
1FKI-SB1	5, 6
1HDC-CBO	1
1HSL-HIS	2 x 1, 7
1RNE-C60	2 x 7, 2 x 8, 9
2ADA-HP3	1, 8, 3, 10, 11
2CGR-GAS	2 x 1, 7
2CTC-HFA	2 x 1
2SIM-DAN	1, 2 x 8, 2 x 12
4PHV-VAC	2 x 7, 13
6ABP-ARA	14, 15, 16, 17, 2 x 18

Table 3.5.1. The interactions of each complex are shown in the above diagrams. The frequency of each HBA relates to the number of times the HBA in the fragments shown in figure 3.5.2 occur.

Fragment	pK_{BHX}	K_{BHX} (M)
1	7.48	0.000000033
2	3.17	0.00068
3	2.76	0.0017
4	2.70	0.002
5	1.50	0.032
6	1.81	0.015
7	1.54	0.029
8	0.84	0.14
9	2.61	0.0024
10	2.58	0.0027
11	4.45	0.000036
12	6.67	0.00000021
13	0.41	0.39
14	2.42	0.0038
15	0.36	0.43
16	1.27	0.053
17	2.68	0.0021
18	2.31	0.0048

Table 3.5.2. The pK_{BHX} [1] and K_{BHX} for each of the HBAs in the fragments shown in figure 3.5.2.

Complex	Computed *Ki* (M)	Actual *Ki* (M)
1CBX-BSZ	1.30E-07	4.50E-07
1CPS-CPM	2.40E-03	2.20E-07
1FKG-SB3	3.30E-02	1.00E-08
1FKI-SB1	4.70E-02	1.00E-07
1HDC-CBO	3.30E-08	1.00E-06
1HSL-HIS	2.90E-02	6.40E-08
1RNE-C60	3.50E-01	2.00E-09
2ADA-HP3	1.50E-01	1.00E-10
2CGR-GAS	2.90E-02	5.30E-08
2CTC-HFA	6.70E-08	1.30E-04
2SIM-DAN	2.90E-01	3.80E-04
4PHV-VAC	4.50E-01	6.70E-10
6ABP-ARA	5.00E-01	4.40E-07

Table 3.5.3. Computed and actual *Ki* values for each complex.

Unfortunately, it has not been possible to compute accurate *Ki* values using the methods outlined above. As it is possible to estimate pK_{BHX} values, there must be two reasons why this attempt to calculate Ki has been unsuccessful. Firstly, the methods used to fragment the drug molecule may not be sufficient, leading to inaccurate basicity values and loose Ki predictions. Secondly, it may not be possible to estimate Ki values of drug-protein interactions from experimental pK_{BHX} values. Binding is governed by the effects of enthalpy and entropy. Although results in the previous section do not show basicity correlating with binding energy, the experimental procedure would suggest that pK_{BHX} values are dominated by enthalpy. This is due to the use of a non-polar solvent, which cannot account for the effects of desolvation entropy seen in biological interactions. Basicity values do not seem to account for entropy driven hydrophobic interactions either. Hydrophobic interactions contribute to the binding of biological interactions, as the diagrams of the test set shown above clearly indicate. [1]

3.6 – Strychnine Case Study.

Strychnine is a highly toxic alkaloid that is used in agrochemistry as a pesticide. Strychnine is a glycine and acetylcholine antagonist that acts primarily in the spinal cord. The antagonism of glycine and acetylcholine receptors in the spinal cord by strychnine blocks neurotransmission in the motor neurons that control muscle contraction. Due to the action of strychnine on motor neurons, strychnine is said to be a neurotoxin. Strychnine poisoning in animals causes eventual death due to respiratory failure [178]. Strychnine, although used as a pesticide also affects humans in much the same way. The strychnine molecule is shown in figure 3.6.1 and contains six potential hydrogen bond acceptor sites. Therefore strychnine is a polyfunctional base.

Strychnine

Figure 3.6.1. Strychnine contains six potential hydrogen bond acceptor sites.

The six potential hydrogen bond acceptor sites on strychnine are the two nitrogen atoms and the two oxygen atoms as well as benzoic and ethylenic sites. It is not possible to obtain binding data for the individual hydrogen bond acceptor sites of strychnine by measuring binding constants directly. However, it is possible to obtain estimated basicity values for each binding site in strychnine. In order to obtain basicity values for polyfunctional bases such as strychnine, Fourier transform infrared spectroscopy (FTIR) must be used. FTIR measures the shift in wavenumber of the OH hydrogen bond donor bond upon complexation. It has been found that formation constants calculated directly for monofunctional bases in 1:1 complexes with a hydrogen bond donor correlate well with the values for shift in wavenumber of the OH bond upon complexation obtained via

FTIR. The correlations between the FTIR data and formation constants are only found for specific families and sub-families of chemicals. Once a correlation has been established for families it is possible to estimate binding data for polyfunctional bases. FTIR traces will display a peak for each hydrogen bond acceptor site. Therefore a shift in wavenumber of the OH hydrogen bond donor bond will be available for each hydrogen bond acceptor site on a polyfunctional base. By identifying each hydrogen bond acceptor site with its family of chemical classification, formation constants can be found by inserting the value of the shift in wavenumber of the OH bond into the equation of the straight line for the correlation between formation constants and the shift in wave number of the OH bond.

The FTIR output for strychnine complexed with 4-fluorophenol gives three peaks that correspond to three major hydrogen bond acceptor sites. The three hydrogen bond acceptor sites on strychnine are the ether oxygen, the amide oxygen and the amine nitrogen.

The equilibrium constant for the formation of a hydrogen bond complex between a hydrogen bond donor and a hydrogen bond accepter is shown below

$$K_t = \frac{[HBA\cdots HBD]}{[HBA][HBD]} \qquad (3.6.1).$$

In equation 3.6.1 K_t is the total equilibrium constant. The total equilibrium constant is calculated from the concentration of all hydrogen bonded complex regardless of the hydrogen bonding site. For polyfunctional bases K_t will be calculated from a [HBA···HBD] component that includes hydrogen bonds formed at different acceptor sites. Equation 3.6.2 shows how the concentrations of the different hydrogen bond complexes formed by polyfunctional bases add up to the total concentration of hydrogen bonded complex

$$[HBA\cdots HBD]=\sum_{i=1}^{n}[HBA\cdots HBD]_i \qquad (3.6.2).$$

Given that the summation of the concentrations of hydrogen bond complexes formed by polyfunctional bases add up to the concentration of the total hydrogen bonded complex, it follows that the summation of individual equilibrium constants add up to K_t as in equation 1.4.10. The value of K_t can be calculated directly and the global pK_{BHX} can be calculated from equation 1.4.14. The values of K_i can be estimated from FTIR data as described above. Therefore the validity of using FTIR as a tool for estimating K_i can be tested by comparing the measured K_i with a value calculated from adding individual estimated K_i values. In a study using strychnine as an example [179], the measured experimental global pK_{BHX} of 2.97 compared very well with the value estimated from the summation of FTIR data of 2.85.

It has been established in section 3.2 that ΔE(H) correlates well with pK_{BHX}. Given this relationship it is possible to compute estimated pK_{BHX} values for each hydrogen bond acceptor site in strychnine and compare the results with those of Laurence. The data set of hydrogen bond acceptors complexed with methanol used in section 3.2 was used to predict pK_{BHX} values for hydrogen bond acceptors. The relationship between pK_{BHX} and ΔE(H) is shown in equation 3.6.3

$$pK_{BHX}=197.34\Delta E(H)-3.749 \qquad (3.6.3).$$

Equation 3.6.4 is used to calculate equilibrium constants from pK_{BHX} values

$$K\left(mol^{-1}dm^3\right)=10^{pK_{BHX}} \qquad (3.6.4).$$

Methods to fragment hydrogen bond acceptors have been discussed in section 3.3. The methods of section 3.3 have been used to fragment strychnine into molecules containing

the three major hydrogen bond acceptor sites. Figure 3.6.2 shows the three hydrogen bond acceptor fragments of strychnine used.

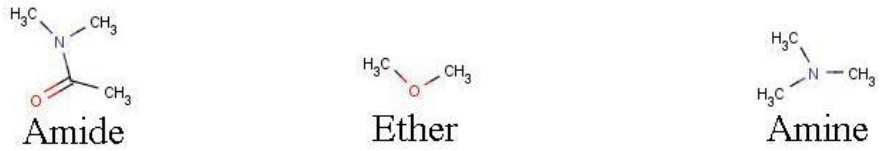


Figure 3.6.2. Strychnine can be fragmented into its major hydrogen bond acceptor sites. The three major hydrogen bond acceptor sites of strychnine are the amide oxygen, the ether oxygen and the amine nitrogen.

A geometry optimisation computation where methanol is the hydrogen bond donor has been performed for each fragment. The methods in section 2.2 describe the technical details of how ΔE(H) values are obtained from wave functions of optimised geometries. The results of these calculations are shown below in table 3.6.1.

Fragment	Computed pK_{BHX}	Actual pK_{BHX}	Computed K($mol^{-1}dm^3$)	Actual K($mol^{-1}dm^3$)
Amine	1.65	2.51	44.65	320
Ether	0.83	0.6	6.81	4
Amide	2.70	2.59	499.16	389

Table 3.6.1. A comparison between experimental [1] and computed data for each hydrogen bond acceptor fragment of strychnine.

The computed global $pK_{BHX} = \log_{10}(44.65 + 6.81 + 499.19) = 2.74$. This compares very well with the experimental value of 2.97. The computed predictions of pK_{BHX} and equilibrium constants compare very well with the actual experimental values for the ether and amide fragments. The computated prediction of pK_{BHX} and equilibrium constants for the amine fragment does not compare well with the actual experimental value. The poor prediction of the amine values contributes to a successful prediction of global pK_{BHX}. If the prediction of amine data could be improved the predicted value for the global pK_{BHX} would also be improved.

There are two reasons why the prediction of amine data is poor. The first reason is that the method of fragmentation is unsuccessful for tertiary amines. The second reason and is much more serious that the computational model for predicting pK_{BHX} values described in section 3.2 is unsuccessful for tertiary amines. The following section will investigate the validity of the computational model for predicting pK_{BHX} values of tertiary amines.

3.7 Tertiary Amines.

Following on from the strychnine case study, this section investigates the relationship between ΔE(H) and pK_{BHX}. The model used to predict the pK_{BHX} values in the strychnine case study was taken from section 3.2 where methanol was the hydrogen bond donor. The relationship between pK_{BHX} and ΔE(H) established in section 3.2 was used to compute predicted pK_{BHX} values for the hydrogen bond donors in strychnine. The model worked well, accurately predicting pK_{BHX} values for ether and amide fragments. However, the pK_{BHX} value for the tertiary amine fragment was not accurately predicted. In this section a set of twelve tertiary amines with pK_{BHX} values ranging between 0.76 and 2.33 have been used to investigate the relationship between pK_{BHX} and ΔE(H). Computations were performed for hydrogen bond complexes following the methods described in section two using methanol as the hydrogen bond donor. As described in section 2.2 ΔE(H) values were computed for the tertiary amines. Table 3.7.1 lists the tertiary amines and their pK_{BHX} values used in this section.

Tertiary Amine	pK_{BHX}
Diazabicyclooctane	2.33
Nicotine	1.11
N-methylpyrrolidine	2.19
N-methyltetrahydroisoquinoline	1.8
N,N-diisopropylethylamine	1.05
N,N-dimethylaminoacetonitrile	0.76
N,N-dimethylaminopropionitrile	1.15
N,N-dimethylethylamine	2.17
N,N-dimethylpropargylamine	1.6
Quinuclidine	2.17
Triallylamine	1.34
Triethyamine	1.98

Table 3.7.1. Tertiary amines used as hydrogen bond acceptors and their pK_{BHX} values.

Figure 3.7.1 shows the correlation between pK_{BHX} and ΔE(H) for the set of 12 tertiary amines listed in table 3.7.1. The points are labelled by their pK_{BHX} value so the tertiary amines can be identified with their correct data point.

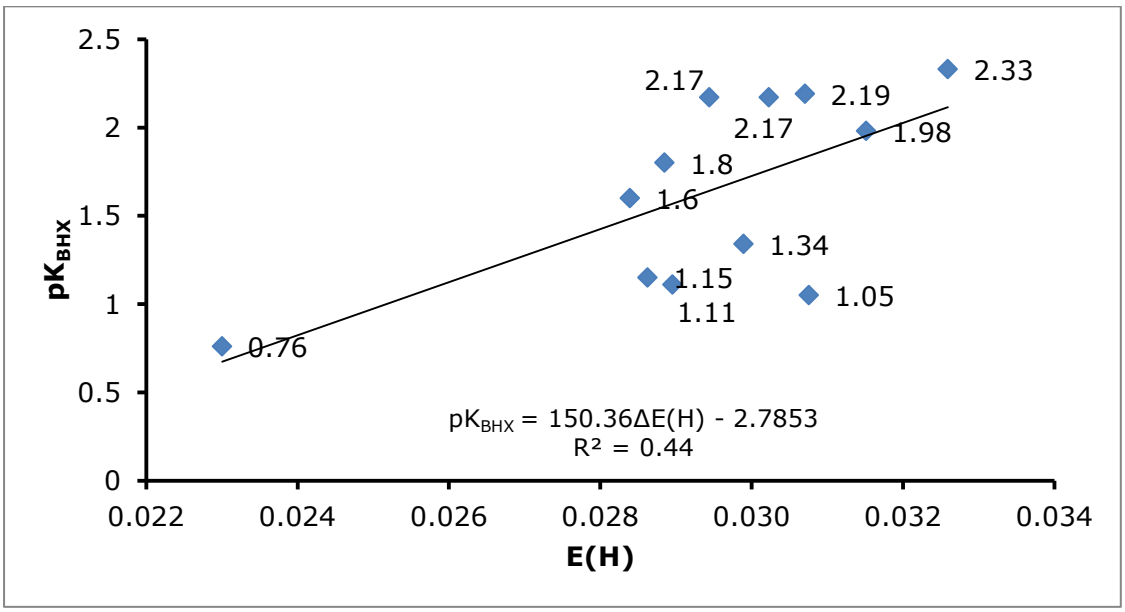


Figure 3.7.1. The relationship between pK_{BHX} and ΔE(H) for the set of 12 tertiary amines listed in table3.7.1. The data points are labelled by pK_{BHX} values and can therefore be identified with the tertiary amine they represent.

Figure 3.7.1 reveals how the relationship previously observed between pK_{BHX} and ΔE(H) has broken down. There is no relationship between pK_{BHX} and ΔE(H) for the 12 tertiary amine used in this study. The r^2 value of 0.44 for the correlation between pK_{BHX} and ΔE(H) for the data set of 12 tertiary amines listed in table 2 is low enough to suggest that there is no relationship. The set of 12 tertiary amines used to test the relationship between pK_{BHX} and ΔE(H) has been chosen from the pK_{BHX} database. The data set of 12 tertiary amines covers a wide range of pK_{BHX} values and is representative of the entire set of tertiary amines covered by the pK_{BHX} database. The 12 tertiary amines used in this section have pK_{BHX} values that are spread out within the range of 0.76 and 2.33 to avoid problems with clustering of data points. Therefore it can be concluded from the correlation seen in figure 3.7.1 that there is no relationship between pK_{BHX} and ΔE(H).

The value of pK_{BHX} for the tertiary amine fragment of strychnine was poorly predicted by a model set up with no tertiary amines in the training set. Given the poor correlation between pK_{BHX} and ΔE(H) for tertiary amines it follows that pK_{BHX} values are poorly predicted for tertiary amines from a model based on a correlation between pK_{BHX} and ΔE(H). Especially as the model included no tertiary amines in the training set.

Tertiary amines are strong proton acceptors. The pK_{BHX} values calculated for tertiary amines are based on equilibrium constants for the formation of a hydrogen bond complex. The hydrogen bonds are formed from a solution of hydrogen bond acceptor in the presence of excess acid. The acceptor–donor solution is bathed in an aprotic solvent CCl_4. Either methanol of 4-fluorophenol has been used as the hydrogen bond donor. Methanol and 4-fluorophenol have been added as aqueous solutions and therefore there will be a small amount of water in the experiment. As the experimental set up contains water, and tertiary amines are generally strong proton acceptors it is possible that the hydrogen bond acceptor site on the tertiary amine may have been protonated. A protonated tertiary amine can take the role of the hydrogen bond donor. This gives rise to the possibility of a hydrogen bond complex forming between the protonated tertiary

amine donor, and the oxygen of the methanol as the acceptor. An example of a protonated tertiary amine forming a hydrogen bonded complex with the oxygen of methanol is shown in figure 3.7.2 using protonated diazabicyclooctane as the hydrogen bond donor.

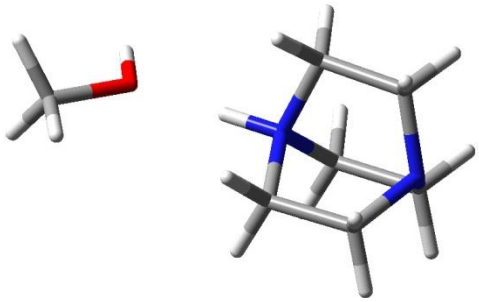


Figure 3.7.2. Protonated tertiary amine diazabicyclooctane donating a hydrogen bond to the oxygen atom of methanol.

Once again following the methods described in section 2 hydrogen bond calculations for the set of protonated tertiary amines were performed. The protonated tertiary amines are the hydrogen bond donors and the oxygen atom of methanol the hydrogen bond acceptors. The hydrogen atom used to calculate ΔE(H) is the protonated site of the tertiary amine rather than the hydrogen atom on methanol. In order to calculate ΔE(H) for the protonated tertiary amine values for E(H) were computed for the free optimised tertiary amine and the hydrogen bonded protonated tertiary amine. The values of ΔE(H) for the protonated tertiary amines have been correlated to pK_{BHX} values of the tertiary amines as listed in table 3.7.1. The result is shown below in figure 3.7.3. Each data point in figure 3.7.3 is labelled by the pK_{BHX} value of the tertiary amine it represents.

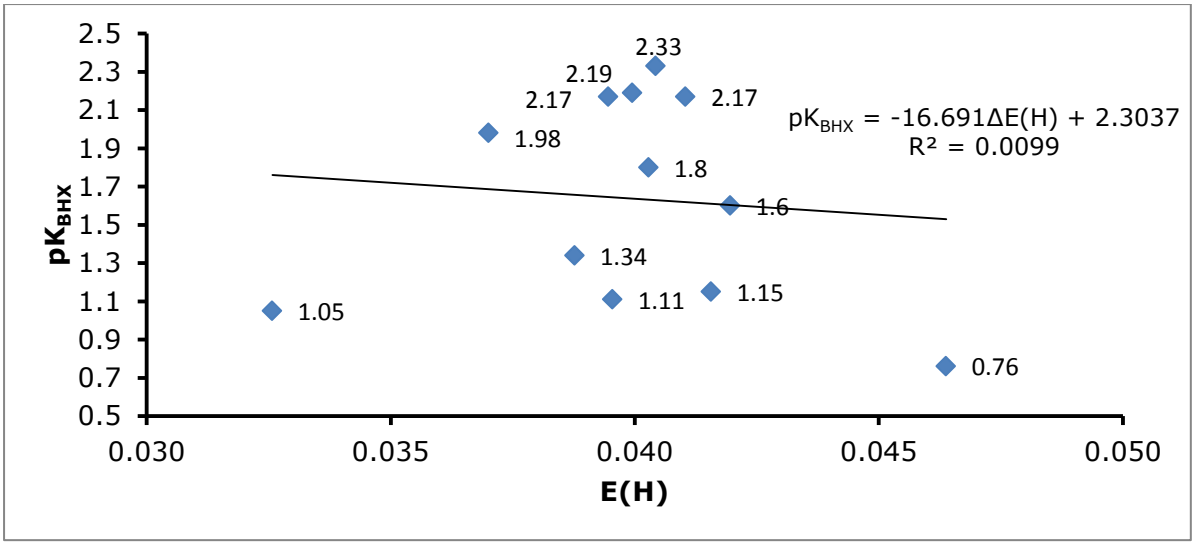


Figure 3.7.3. The relationship between pK_{BHX} [1] and ΔE(H) for the set of 12 protonated tertiary amines listed in table 3.7.1 The data points are labelled by pK_{BHX} values and can therefore be identified with the protonated tertiary amine they represent. The protonated tertiary amine acts as the hydrogen bond donor whereas the oxygen atom of methanol has assumed the role of the hydrogen bond acceptor.

It can be seen from figure 3.7.3 that there is no relationship between the pK_{BHX} value of a tertiary amine and the ΔE(H) value taken from the hydrogen bond complex formed between protonated tertiary amine and methanol. The r^2 value of 0.0099 confirms a random scattering of data points with no observed trend. Therefore protonated tertiary amines, when taking the role of hydrogen bond donors cannot be used to compute pK_{BHX} values from ΔE(H) values. The value of ΔE(H) cannot be used to predict the basicity of the molecule it is housed in, but as we know from section 3.2, ΔE(H) values correlate generally well to pK_{BHX} values of bases included in the pK_{BHX} database with the exception of tertiary amines. Although the pK_{BHX} database has been set up using methanol or 4-fluorophenol as the reference hydrogen bond donor, the pK_{BHX} values listed are applicable when the hydrogen bond donor is of the NH^+ form. Therefore the values of ΔE(H) used to form figure 3.7.3 should be capable of predicting the pK_{BHX} of, in this case methanol, or any other hydrogen bond acceptor they form a complex with.

The next step of this research was to conexamine whether the pK_{BHX} database is applicable to situations where protonated tertiary amines are hydrogen bond donors. In order to do so, it must firstly be assumed that the model to predict pK_{BHX} values established in section 3.2 will hold up in agreement with the hypothesis that protonated tertiary amines can be used successfully as hydrogen bond donors. In other words a strong correlation between pK_{BHX} and ΔE(H) for hydrogen bond complexes where the hydrogen bond donor is a protonated tertiary amine confirms that the R_3NH^+ group of tertiary amines can successfully be used as hydrogen bond donors. It has already been established that pK_{BHX} correlates strongly with ΔE(H) for a set of 41 bases used in section 3.2. The correlation between pK_{BHX} and ΔE(H) is observed for the set of 41 bases used in section 3.2 when the hydrogen bond donor is either an OH or NH group. For the data set of bases used in section 3.2 a strong correlation is observed between pK_{BHX} and ΔE(H) for OH and NH hydrogen bonds as stated by Laurence et al [1]. The relationship between pK_{BHX} and ΔE(H) has already been established (section 3.2) for a set of bases with pK_{BHX} values distributed as they are in the entire data base. Therefore it is not only suitable, less time consuming to use data set containing a reduced number of bases that are taken from section 3.2 where strong correlations are seen. Therefore it is to be expected that a strong correlation between pK_{BHX} and ΔE(H) when a protonated tertiary amine is used as the hydrogen bond donor will be observed when a reduced data set of bases is used.

A set of 19 bases from the pK_{BHX} database have been selected. The 19 bases are chosen from data sets used in section 3.2 where successful correlations between pK_{BHX} and ΔE(H) have been established. The 19 base have been chosen to have pK_{BHX} values that are distributed as per the pK_{BHX} database. The bases selected contain the same distribution of nitrogen, oxygen and sulphur bases as table 3.2.1 which formed the original test set. The 19 bases used in this section are listed in table 3.7.2.

Base	pK_{BHX}
Ethyl thiol	-0.16
Phenol	-0.07
4-methylphenol	0.03
Dimethyl sulfide	0.12
Dimethyl sulfide	0.46
Aniline	0.56
p-toluidine	0.65
Methylformate	0.7
Acrylonitrile	0.82
Methanol	0.91
Acetonitrile	1.02
Ethanol	1.18
Acetone	1.41
4-Acetylpyridine (N)	1.65
Pyridazine	1.86
Pyridine	2.06
Acetamide	2.17
t-Butylamine	2.23
Pyrrolidine	2.59
Triethylarsine oxide	4.89

Table 3.7.2. The 19 bases listed in this table are used as hydrogen bond acceptors in complexes with a protonated hydrogen bond donor. The bases are chosen from data sets used in section 3.2. The pK_{BHX} values are distributed as per the pK_{BHX} database [1]. The spread of pK_{BHX} values and number of oxygen, nitrogen and sulphur acceptors are chosen to replicate table 3.2.1.

Following the instructions in section 2, ΔE(H) values were computed for hydrogen bond complexes where protonated trimethyl amine was the hydrogen bond donor. The bases listed in table 3.7.2 were the hydrogen bond acceptors. An example of such a hydrogen bond is illustrated in figure 3.7.4 by the protonated trimethyl amine–acetone complex. The computed ΔE(H) have been plotted against the pK_{BHX} values from the database listed in table 3.7.2. The results of the correlation between pK_{BHX} and ΔE(H) are shown in figure 3.7.5 below.

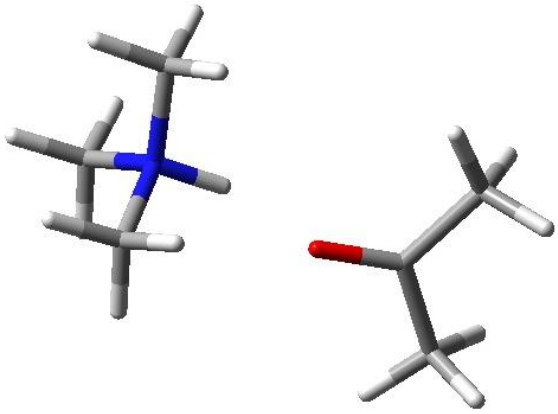


Figure 3.7.4. An example of a hydrogen bond complex where protonated trimethyl amine is the hydrogen bond donor. The NH^+ site of protonated trimethyl amine is donating a hydrogen bond to the oxygen atom in acetone.

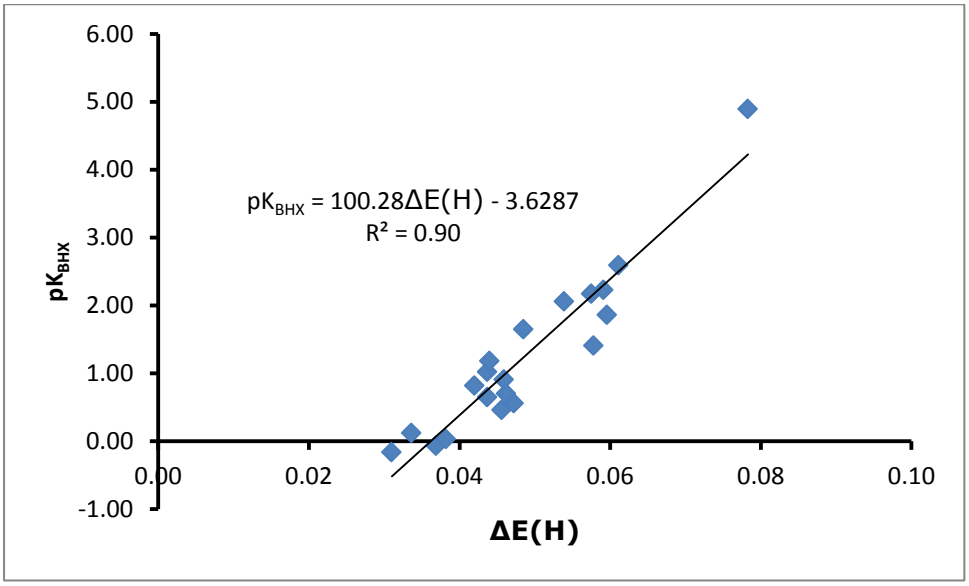


Figure 3.7.5. The relationship between pK_{BHX} [1] and ΔE(H) for hydrogen bond complexes formed between the bases in table 3.7.2 and the hydrogen bond donor protonated trimethyl amine.

There is a fairly strong correlation between pK_{BHX} and ΔE(H) where protonated trimethyl amine is the hydrogen bond donor. Figure 3.7.5 shows an r^2 value of 0.90 for the correlation between pK_{BHX} and ΔE(H) where protonated trimethyl amine is the hydrogen bond donor. The finding of a strong correlation between pK_{BHX} and ΔE(H) where protonated trimethyl amine is the hydrogen bond donor confirms the hypothesis

that pK_{BHX} values can be predicted by computing ΔE(H) where the hydrogen bond donor is the protonated site in a NH^+ group. The finding that pK_{BHX} values may be predicted by ΔE(H) values taken from a NH^+ group ties in with the authors statement that the pK_{BHX} database is applicable to NH^+ groups and therefore provides a strong link between computed ΔE(H) and experimental values. A strong link between experimental pK_{BHX} values and ΔE(H) was observed in section 3.2 where methanol was used as the hydrogen bond donor. Figure 3.7.6 displays the correlation between pK_{BHX} and ΔE(H) for the bases listed in table 3.7.2, but this time methanol has been used as the hydrogen bond donor. It can be seen from figure 3.7.6 that there is a slightly stronger correlation when methanol was used as the hydrogen bond donor than when protonated trimethyl amine was the hydrogen bond donor. The same data set gave an r^2 value of 0.95 when methanol was used as the hydrogen bond donor compared to an r^2 value of 0.9 when protonated trimethylamine was used as the hydrogen bond donor.

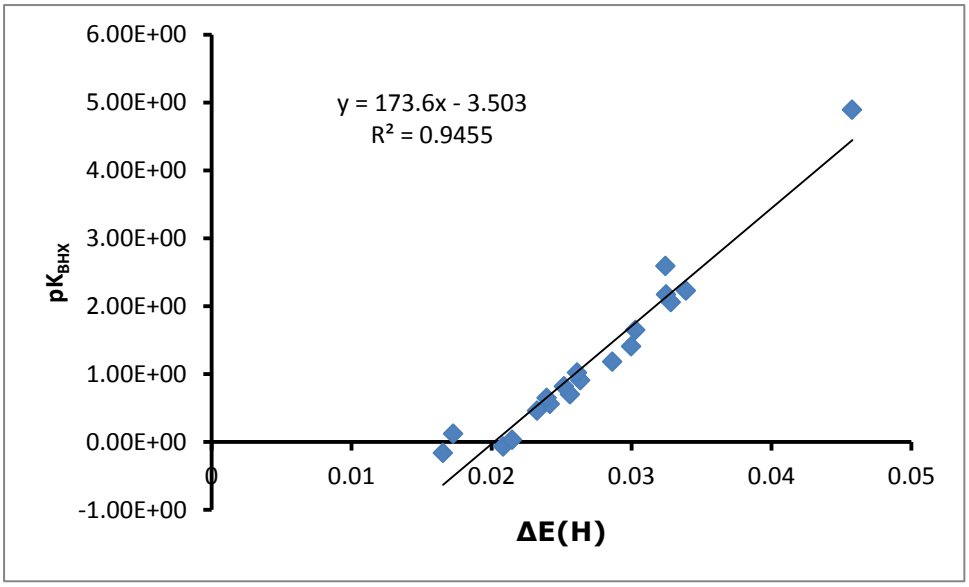


Figure 3.7.6. The relationship between pK_{BHX} and ΔE(H) for hydrogen bond complexes formed between the bases in table 3.7.2 and the hydrogen bond donor methanol.

Data taken from a complex where a protonated tertiary amine behaves as the hydrogen bond donor cannot be used to predict pK_{BHX} values of the unprotonated form. Computed ΔE(H) values for protonated tertiary amines in hydrogen bond complexes do

not correlate with pK_{BHX} values of the unprotonated form. However, ΔE(H) values for protonated tertiary amines in hydrogen bond complexes where the protonated tertiary amine is the hydrogen bond donor do correlate to pK_{BHX} values of a set of bases. Therefore if a protonated tertiary amine is chosen as the reference, hydrogen bond donor ΔE(H) values can be used to predict pK_{BHX} values.

The breakdown of the relationship between pK_{BHX} and ΔE(H) for tertiary amines requires further investigation. It has been explained above how it could be possible for the tertiary amine to become protonated in the experimental set up. The protonation could arise from water present in the experiment. Tertiary amines are generally strong proton acceptors. However the presence of water in the experimental set up arises from reactants being added in aqueous solution and therefore the extent of protonation that is possible is uncertain. It is possible that the tertiary amine could become protonated disregarding water as a source. The source of the proton could be the hydrogen bond donor itself. The hydrogen bond donor could transfer a proton to the hydrogen bond donor instead of forming a hydrogen bond complex. In this instance the hydrogen bond complex could be formed between a protonated tertiary amine donor, and a deprotonated methanol or 4-fluorophenol acceptor. The hydrogen bond could therefore form between the positive NH^+ group and the OH^- group. A hydrogen bond formed between a NH^+ group and a OH^- group causes the complex to behave in a zwitterion like manner.

The next hypothesis why the relationship between pK_{BHX} and ΔE(H) breaks down for tertiary amines is that the complex behaves as zwitterion in solution. In order to link this hypothesis to computed ΔE(H), further calculations have been performed. Following the guidelines in section 2, ΔE(H) values have been computed for the hydrogen bond formed between the protonated tertiary amines in table 3.7.1 and deprotonated methanol. An example of a starting geometry for such computations is shown in figure 3.7.7 for the complex between protonated diazabicyclooctane and deprotonated methanol.

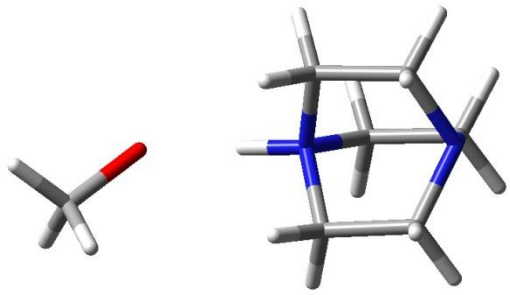


Figure 3.7.7. An example of a starting geometry for a complex behaving as a zwitterion. Protonated diazabicyclooctane is donating a hydrogen bond to deprotonated methanol.

Once again no relationship was found between pK_{BHX} and ΔE(H) for hydrogen bond complexes behaving as zwitterions.

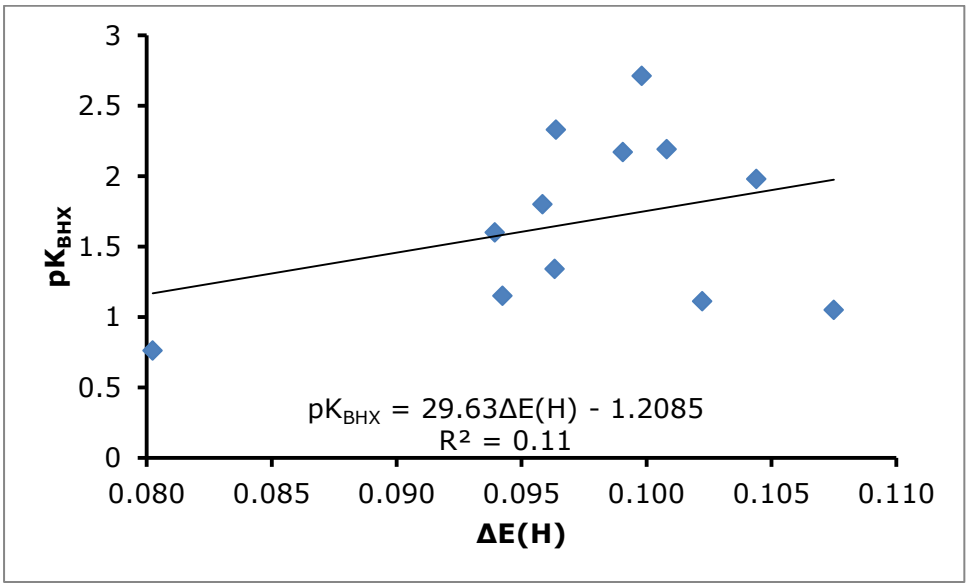


Figure 3.7.8 The relationship between pK_{BHX} and ΔE(H) for hydrogen bond complexes formed between the tertiary amines in table 3.7.1 and methanol. The complex is modelled to behave as a zwitterion as the methanol is deprotonated and behaves as the hydrogen bond acceptor whereas the tertiary amine is protonated and behaves as the hydrogen bond donor.

It can be seen from figure 3.7.8 that ΔE(H) values taken from complexes behaving like zwitterions cannot be used to predict pK_{BHX} values of tertiary amines. The r^2 value of 0.11 indicates a random scattering of points for the plot of pK_{BHX} against ΔE(H). It can therefore be concluded that pK_{BHX} values of tertiary amines cannot be predicted by calculating ΔE(H) values of hydrogen bond complexes modelled as

zwitterions. It is already known from results displayed in figure 3.7.3 that pK_{BHX} values of tertiary amines cannot be predicted from ΔE(H) values taken from complexes where the hydrogen bond donor is a protonated tertiary amine and the hydrogen bond acceptor is methanol. It has also been shown (figure 3.7.1) that pK_{BHX} values of tertiary amines cannot be predicted from ΔE(H) values computed from hydrogen bond complexes where the tertiary amine is the hydrogen bond acceptor and methanol is the hydrogen bond donor. There is also no correlation between pK_{BHX} and ΔE(H) when the hydrogen bond complex is modelled as a zwitterion. In fact when modelling the complex as a zwitterions, the starting geometry resembles the experiments where the hydrogen bond donor is a tertiary amine. The only difference is that the methanol is deprotonated.

It has been shown that the ΔE(H) values taken from complexes involving a protonated tertiary amine donor can only be used to predict the pK_{BHX} value of the hydrogen bond acceptor and not the pK_{BHX} value of the tertiary amine. During the geometry optimisation step in the set of calculations where the complex is modelled like a zwitterion, there is a full proton transfer between the protonated tertiary amine and the deprotonated methanol. After a full proton transfer between the protonated tertiary amine and the deprotonated geometry, an optimised geometry of the complex illustrates a situation where the neutral tertiary amine accepts a hydrogen bond from methanol.

The proton transfer is displayed in figure 3.7.9. As a full proton transfer has occurred, the complex resembles those used to generate the initial model where methanol is the hydrogen bond donor and the tertiary amines are the hydrogen bond acceptors. It has already been shown that ΔE(H) values taken from complexes between tertiary amines and methanol cannot be used to successfully predict pK_{BHX} values of the tertiary amines. The substantial difference between the ΔE(H) values taken from complexes between methanol donating a hydrogen bond to tertiary amines and the ΔE(H) values taken from complexes behaving like a zwitterion is the initial E(H) values used. When methanol is donating a hydrogen bond to the tertiary amine the initial E(H) value is taken from free methanol. When the complex is behaving like a zwitterion the

initial E(H) value is taken from the protonated tertiary amine, as the initial hydrogen bond donor is the protonated tertiary amine. As the proton is transferred from the protonated tertiary amine to the deprotonated methanol a hydrogen bond complex is formed identical to the situation where methanol is donating a hydrogen bond to the tertiary amines.

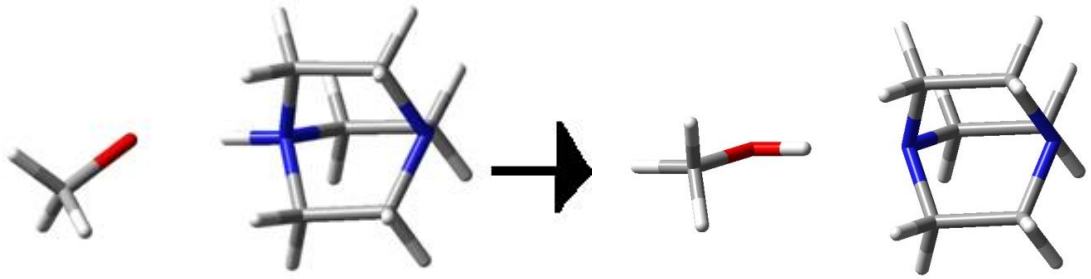


Figure 3.7.9 When modelling the hydrogen bond complex like a zwitterion, initially the tertiary amine is protonated and is donating a hydrogen bond to deprotonated methanol. During the optimisation step, a full proton transfer is observed from the tertiary amine to methanol. The resulting complex is one where methanol is donating a hydrogen bond to the tertiary amine. This is displayed here using diazabicyclooctane as an example of a tertiary amine.

It has been found that the relationship between pK_{BHX} and $\Delta E(H)$ that has been established in section 3.2 has broken down for tertiary amines. Following the logic that tertiary amines are strong proton acceptors and are likely protonated in solution, it has also been found that $\Delta E(H)$ values for protonated tertiary amines acting as hydrogen bond donors do not correlate to pK_{BHX} values of tertiary amines. Also, pK_{BHX} values of tertiary amines cannot be predicted from $\Delta E(H)$ values of hydrogen bonded complexes where the complex behaves like a zwitterion. In order to understand why the relationship between pK_{BHX} and $\Delta E(H)$ breaks down for tertiary amines, the experimental method used to obtain pK_{BHX} will be discussed, bearing in mind the hypothesis of the likely protonated state of the tertiary amines in the experimental solution.

Experimental pK_{BHX} values are calculated from equilibrium constants of a hydrogen bond acceptor surrounded by excess hydrogen bond donor in an aprotic solution. The aprotic solution is CCL_4 and the hydrogen bond donor is usually 4-

fluorophenol or methanol. The following equations show the equilibrium between a tertiary amine R_3N and a hydrogen bond donor HX

$$R_3N + HX \xleftrightarrow{K_1} R_3N \cdots HX \qquad (3.7.1)$$

$$K_1 = \frac{\left[R_3N \cdots HX\right]}{\left[R_3N\right]\left[HX\right]} \qquad (3.7.2).$$

The equilibrium constants for monofunctional bases may be calculated directly be measuring the concentrations of each free acid and base, and the complex. The values of pK_{BHX} calculated directly for monofunctional bases have been found to correlate very strongly to the shift in wave number of the OH bond of hydrogen bond donor upon complexation. The correlation between pK_{BHX} and shift in wave number of the OH bond for monofunctional bases allows pK_{BHX} values of polyfunctional bases to be obtained. To obtain pK_{BHX} values for polyfunctional bases, the shift in wave number of the OH bond must be known. Fourier transform infrared spectroscopy is used to calculate the shift in wave number of the OH bond. The value of the shift in wave number of the OH bond is substituted into a straight line equation of the correlation between the shift in wave number of the OH bond and pK_{BHX} in order to generate a pK_{BHX} value. However, if the bases are protonated in the experimental solution, an additional hydrogen bonded complex independent from the desired one may be formed. The following equations describe how an additional hydrogen bond complex could form

$$R_3NH^+ + XH \xleftrightarrow{K_2} R_3NH^+ \cdots XH \qquad (3.7.3)$$

$$K_2 = \frac{\left[R_3NH^+ \cdots XH\right]}{\left[R_3NH^+\right]\left[XH\right]} \qquad (3.7.4).$$

If protonation of the hydrogen bond acceptor occurs, then there are potential issues surrounding the interpretation of the equilibrium constant. The complex

concentration taken to calculate the equilibrium constant will be made up of a combination of the complex in equation 3.7.1 and equation 3.7.3. However, only the non protonated hydrogen bond complex (equation 3.7.1) is of interest when calculating hydrogen bond basicity.

The concentration of complex measured is that of the total complex made up of the protonated and non-protonated form. It is assumed by the experimentalists that the total concentration of the complex is made up of that of only the desired hydrogen bond complex. In the case of protonation the experimentalists will actually over-estimate the concentration of the desired hydrogen bond complex. Over-estimating the concentration of the hydrogen bond complex will lead to a larger equilibrium constant being calculated. A larger equilibrium constant leads to a larger pK_{BHX} value and an over estimated hydrogen bond basicity. The effects of overestimating the hydrogen bond basicity of a hydrogen bond acceptor can be extrapolated to account for possible errors in estimating the hydrogen bond basicity of polyfunctional bases. The possible error in estimating hydrogen bond basicity of polyfunctional bases arises from shifts in wave number of the OH bond upon complexation correlating with untrue equilibrium constants.

In summary, the above text has described a possible flaw in the experimental technique used to calculate hydrogen bond basicities that could possibly lead to over-estimating the basicity of a given hydrogen bond acceptor. However, the hydrogen bond acceptor must be highly basic in order to become protonated in the experimental set up. It can therefore be hypothesised that bases that are strong proton acceptors could have pK_{BHX} values that have been estimated to be higher than they actually are. If this is the case then the above hypothesis could be used to explain why the model to compute pK_{BHX} values described in section 3.2 under predicted the literature value of the tertiary amine fragment of strychnine. The hypothesis could also explain why no relationship exists between pK_{BHX} and $\Delta E(H)$.

To test the hypothesis that experimental pK_{BHX} values for hydrogen bond acceptors that are strong proton acceptors are over-estimated, the model established in

section 3.2 will be used. The model relies on a strong correlation between ΔE(H) and pK_{BHX}. Assuming the relationship between ΔE(H) and pK_{BHX} is universally accurate, it would be expected that no correlation between ΔE(H) and pK_{BHX} exists for the strongest proton acceptors, whereas a strong correlation is seen between ΔE(H) and pK_{BHX} for weaker proton acceptors. The following section investigates how the relationship between pK_{BHX} and ΔE(H) is affected by the proton acceptor strength of the hydrogen bond acceptors.

3.8 Relationship Between pK_{BHX} Prediction And Brönsted Basicity.

The research in this section will attempt to explain why the correlation between ΔE(H) and pK_{BHX} breaks down for tertiary amines. It is understood that tertiary amine are strong proton acceptors. Strong proton acceptors are defined by large pK_{BH}^{+} values, as defined in equation 3.7.4.

The following research is based on the hypothesis that hydrogen bond acceptors with large positive pK_{BH}^{+} values will not generate computed ΔE(H) values that correlate with their experimental pK_{BHX} values. It has been hypothesised in the previous section that the database pK_{BHX} values could possibly be over-estimated for hydrogen bond acceptors that are also strong proton acceptors. The reason that database pK_{BHX} values could possibly be overestimated for strong proton acceptors lies in the speculation that the hydrogen bond acceptors could be present in the form of conjugate acids in the experimental set up. In other words the hydrogen bond acceptors could well be protonated in the experimental solution. It is therefore possible that protonated hydrogen bond acceptors could form an additional hydrogen bond complex in the experimental solution.

The possible additional hydrogen bond complex is one where the conjugate acid donates a hydrogen bond to the oxygen atom of methanol or 4-fluorophenol. Methanol

or 4-fluorophenol are used as reference desired hydrogen bond donors. Methanol or 4-fluorophenol are therefore supposed to donate a hydrogen bond to the neutral base. The role of methanol or 4-fluorophenol is reversed, instead of donating hydrogen bonds they are accepting them. The concentration of total hydrogen bond complex measured in the calculation of pK_{BHX} values could therefore quite possibly be too high due to the presence of the additional complex where the conjugate acid is donating a hydrogen bond. This is because the total concentration of hydrogen bond complex will be made up of desired complex where the neutral base is accepting a hydrogen bond, and the undesired complex where the conjugate acid is donating a hydrogen bond. As only hydrogen bond basicity of the neutral base is of interest, the complex formed by the conjugate acid is undesired. However the undesired complex contributes to a possible over-estimation of pK_{BHX} values.

It is expected that the hydrogen bond acceptors with the largest positive pK_{BH}^{+} values will not form complexes with computed $\Delta E(H)$ values that correlate with pK_{BHX}. To further this hypothesis, knowledge about the pK_a slide rule can be drawn upon. The pK_a slide rule states that the strongest hydrogen bond complexes will be formed by acids and bases with similar dissociation constants. The greater the difference there is in dissociation constants, the more ionic character the interaction will portray. As the interaction becomes increasingly ionic, the more likely a complete proton transfer is likely to occur.

If a proton transfer or protonation reaction occurs the resulting hydrogen bond will be either very weak or in some cases non-existent. As the dissociation constants of the hydrogen bond donor and hydrogen bond acceptor become increasingly similar, the covalent character of the hydrogen bond interaction increases. The strongest hydrogen bonds formed are those with the most covalent character. Therefore it would be expected that the strongest hydrogen bond acceptors in the pK_{BHX} database are those that have dissociation constants similar to the reference hydrogen bond donor.

It would be expected that methanol and 4-flouorphenol form weak hydrogen bonds with hydrogen bond acceptors with large positive pK_{BH}^{+} values. Following from this it would be expected that hydrogen bond acceptors with large positive pK_{BH}^{+} values such as tertiary amines, would have low pK_{BHX} values. However pK_{BHX} and pK_{BH}^{+} are known to be uncorrelated [1]. It is possible that hydrogen bond acceptors with large positive pK_{BH}^{+} values become protonated in the experimental solution and donates a strong hydrogen bond to methanol or 4-fluorophenol. If the situation where the protonated hydrogen bond acceptor donates a strong hydrogen bond to methanol or 4-fluorophenol arises then results would display an over estimated pK_{BHX} database value. An overestimated pK_{BHX} value is what the model described in section 3.6 suggests the pK_{BHX} database shows. It has also been shown that there is no correlation between hydrogen bond basicity and Brönsted basicity [1].

It is known that there is no relationship between pK_{BHX} and pK_{BH}^{+} for 217 bases used in this section. In certain cases the dominant hydrogen bond acceptor site of a molecule is different from the first protonation site. Therefore it would not be expected that pK_{BH}^{+} should correlate with pK_{BHX}. However, given the pK_a slide rule [119], it would be expected that hydrogen bond acceptors with similar pK_{BH}^{+} values should have similar pK_{BHX} values. This is also not the case. For example, N,N-diethylaniline and 4-methoxypyridine have pK_{BH}^{+} values of 6.61 and 6.58 respectively. However, the pK_{BHX} value of N,N-diethylaniline is 0.05 whereas the value for 4-methoxypyridine is 2.13. The following text will explain how the hypothesis that high ranking Brönsted bases will not give rise to hydrogen bonded complexes where the computed ΔE(H) values correlate strongly with experimental pK_{BHX} values.

A set of 217 bases has been selected. The 217 bases along with their pK_{BH}^{+} [1] and pK_{BHX} values [1] are listed in table 3.8.1. The bases in table 3.8.1 are ranked in order of descending pK_{BH}^{+} values. The 217 bases chosen are based on those chosen to illustrate how there is no relationship between pK_{BH}^{+} and pK_{BHX} in [1]. The bases and their pK_{BHX} and pK_{BH}^{+} values are listed in the supplementary information of [1]. Using

methanol as the hydrogen bond donor, hydrogen bond complex computations have been performed in the usual way as described in section 2. Once again ΔE(H) values have been calculated.

HYDROGEN BOND ACCEPTOR	pK_{BH}^{+}	pK_{BHX}
1,1,2,3,3-Pentamethylguanidine	13.8	3.16
1,1,3,3-Tetramethylguanidine	13.6	3.21
N,N-Dimethyl-N'-propylimidoformamide	11.46	2.59
Pyrrolidine	11.306	2.59
N'-Isobutyl-N,N-dimethylimidoformamide	11.3	2.52
Azetidine	11.29	2.59
Dibutylamine	11.25	2.11
Diisopropylamine	11.2	2
1,2,2,6,6-Pentamethylpiperidine	11.19	1.23
Quinuclidine	11.1516	2.71
Piperidine	11.1236	2.38
Azepane	11.1	2.24
2,2,6,6-Tetramethylpiperidine	11.0725	1.88
N-Methylcyclohexylamine	11.04	2.24
Diethylamine	11.0151	2.25
N,N-Dimethyl-N'-(4-methylbenzyl)imidoformamide	10.91	2.36
N-methylbutylamine	10.9	2.24
Dimethylamine	10.7788	2.26
N,N-Dimethylcyclohexylamine	10.72	2.15
Triethylamine	10.7174	1.98
Tert-butylamine	10.6837	2.23
Ethylamine	10.6784	2.17
Isopropylamine	10.67	2.2
Tripropylamine	10.66	1.47
Methylamine	10.6532	2.2
N'-Benzyl-N,N-dimethylimidoformamide	10.65	2.35
n-Butylamine	10.6385	2.19
1,6-Diaminohexane	10.63	2.21
n-Octylamine	10.61	2.27
n-Hexadecylamine	10.61	2.26
c-Hexylamine-MeOH	10.58	2.29
n-Propylamine	10.5685	2.2
NN-Dimethylisopropylamine-MeOH	10.47	2.11

N-Methylpyrrolidine	10.46	2.19
N-Butylpyrrolidine	10.36	2.04
1,4-Diaminobutane	10.35	2.21
N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	10.32	2.12
1,3-Diaminopropane	10.17	2.31
N,N-Dimethylethylamine	10.16	2.17
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	10.15	2.1
N-Methylallylamine	10.11	2
N-Methylpiperidine	10.08	2.11
N,N'-dimethylethylenediamine	9.94	2.29
Tributylamine	9.93	1.55
3-Methoxypropylamine	9.92	2.22
N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	9.86	2
2-Phenylethanamine	9.83	2.16
Trimethylamine	9.7977	2.13
Ethylenediamine	9.626	2.25
2-Phenylpyrrolidine	9.6	1.93
4-N,N-Dimethylaminopyridine	9.58	2.8
N-Methylbenzylamine-MeOH	9.56	1.82
Allylamine	9.52	1.93
2-Methoxyethylamine	9.44	2.26
Piperazine	9.432	2.11
1,2,3,4-Tetrahydroisoquinoline	9.41	2.04
Benzylamine	9.34	1.84
Ammonia	9.244	1.74
Diallylamine	9.24	1.7
4-Aminopyridine	9.12	2.56
c-Propylamine	9.1	1.72
N,N-Dimethylbenzylamine	8.91	1.59
N,N,N',N'-Tetramethylethylenediamine	8.85	2.02
N-Methyl-2-phenylpyrrolidine	8.8	1.38
N,N-Dimethylallylamine	8.64	1.92
Diazabicyclooctane	8.52	2.33
Dibenzylamine	8.52	1.34
Morpholine	8.4918	1.78
N,N-Dimethyl-N'-(4-methylphenyl)imidoformamide	8.45	2.07
Triallylamine	8.28	1.34
N,N-Dimethyl-N'-(2-methylphenyl)imidoformamide	8.27	1.63
N,N-Dimethyl-N'-phenylimidoformamide	8.15	1.9
Propargylamine	8.15	1.56
3-Aminopropionitrile	7.8	1.33
Hexamethylenetetramine	7.4696	1.33
N,N-Dimethylpropargylamine	7.45	1.6
2,4,6-Trimethylpyridine	7.43	2.29
N-Methylmorpholine	7.41	1.56

N'-(4-Bromophenyl)-N,N-dimethylimidoformamide	7.4	1.65
1,1-Diphenylmethanimine	7.18	1.8
1-Methyl-1H-imidazole	7.12	2.72
N'-(4-Acetylphenyl)-N,N-dimethylimidoformamide	7.02	1.52
2-N,N-Dimethylaminopyridine	6.99	1.61
2,6-Dimethylpyridine	6.72	2.14
2-Aminopyridine	6.71	2.12
N'-(2-Bromophenyl)-N,N-dimethylimidoformamide	6.71	1.37
2-Methyl-N-(phenylmethylene)propan-2-amine	6.7	1.29
N,N-Diethylaniline	6.61	0.05
4-Methoxypyridine	6.58	2.13
3,4-Dimethylpyridine	6.47	2.24
N'-(4-Cyanophenyl)-N,N-dimethylimidoformamide	6.44	1.32
3,5-Dimethylpyridine	6.34	2.21
Thiazolidine	6.22	1.1
N,N,N',N'-Tetramethylbenzene-1,4-diamine	6.05	1.13
3-Aminopyridine	6.03	2.2
4-Methylpyridine	6.03	2.07
4-Ethylpyridine	6.02	2.07
N,N-Dimethyl-N'-(4-nitrophenyl)imidoformamide	6.02	1.2
4-Tert-butylpyridine	5.99	2.11
2-Methylpyridine	5.96	2.03
2-Ethylpyridine	5.89	1.94
2-Isopropylpyridine	5.83	1.76
2-Tert-butylpyridine	5.76	1.42
3-Ethylpyridine	5.73	2.01
3-Methylpyridine	5.66	2
N,N,4-Trimethylaniline	5.63	0.69
4-Vinylpyridine	5.62	1.95
2,2,2-Trifluoroethylamine	5.61	0.71
Isoquinoline	5.4	1.94
N,4-Dimethylaniline	5.36	0.43
4-Phenylpyridine	5.35	1.96
Acridine	5.24	1.95
Pyridine	5.2	1.86
p-Toluidine	5.08	0.56
N,N-Dimethylaniline	5.07	0.39
1,2,3,4-Tetrahydroquinoline	5.03	0.7
2-Vinylpyridine	4.92	1.65
N-Methylaniline	4.85	0.26
Quinoline	4.85	1.89
2-Phenylpyridine	4.72	1.43
Aniline	4.61	0.46
1-Phenylpyrrolidine	4.3	0.16
4-Bromo-N,N-dimethylaniline	4.23	0.17

Dimethylaminoacetonitrile	4.2	0.67
2,6-Dimethylaniline	3.95	0.47
7,8-Benzoquinoline	3.95	1.16
4-Chloro-N-methylaniline	3.9	0.05
4-Chloropyridine	3.83	1.54
1-Acetyl-1H-imidazole	3.6	1.86
3-Fluoroaniline	3.59	0.2
2-Aminopyrimidine	3.54	1.85
3-Chloroaniline	3.52	0.13
4-Acetylpyridine	3.51	1.41
Phthalazine	3.17	1.97
3-Dimethylamino-5,5-dimethylcyclohexenone	3.14	2.92
Tripropargylamine	3.09	0.83
Methylnicotinate	3.08	1.44
2-Methoxypyridine	3.06	0.99
3-Fluoropyridine	3	1.35
3-Bromopyridine	2.84	1.31
3-Chloropyridine	2.8	1.31
1,3-Thiazole	2.518	1.37
1-Methyl-1H-pyrazole	2.06	1.84
4-Methoxypyridine 1-oxide	2.05	3.7
Pyridazine	2	1.65
4-Cyanopyridine	1.86	0.92
3-Cyanopyridine	1.34	0.82
Isoxazole	1.3	0.81
4-Methylpyridine 1-oxide	1.29	3.12
Pyrimidine	0.93	1.07
3-Methylpyridine 1-oxide	0.93	2.92
2-Bromopyridine	0.9	1.03
Phenazine	0.9	1.22
4-Phenylpyridine 1-oxide	0.83	2.85
1,3-Oxazole	0.8	1.3
Diphenylamine	0.79	-1.05
Pyridine 1-oxide	0.79	2.72
2-Chloropyridine	0.72	1.05
3,5-Dichloropyridine	0.66	0.85
Pyrazine	0.37	0.92
4-Chloropyridine 1-oxide	0.33	2.44
1,1,3,3-Tetramethylurea	-0.14	2.44
N,N-Dimethylacetamide	-0.21	2.44
2-Cyanopyridine	-0.26	0.48
2-Fluoropyridine	-0.44	0.95
Tetramethylene sulfone	-0.65	1.17
Acetamide	-0.66	2.06
N-Methylpropanamide	-0.7	2.24

1-Methyl-2-pyrrolidone	-0.71	2.38
3,5-Dichloropyridine 1-oxide	-0.94	1.56
Hexamethylphosphoramide	-0.97	3.6
N-Methylformamide	-1.1	1.96
N,N-Dimethylformamide	-1.13	2.1
Formamide	-1.47	1.75
4-Nitropyridine 1-oxide	-1.51	1.05
Dimethyl sulfoxide	-1.54	2.54
N,N,2-Trimethylpropanamide	-1.61	2.26
Ethanol	-1.94	1.02
Propan-2-ol	-2.02	1.06
N,N,2,2-Tetramethylpropanamide	-2.03	2.1
Methanol	-2.05	0.82
Propan-1-ol	-2.12	1
N,N-Dimethylthioacetamide	-2.25	1.22
Methyl phenyl sulfoxide	-2.27	2.24
Diethyl ether	-2.39	1.01
2-Bromethanol	-2.41	0.54
2-Chloroethanol	-2.45	0.5
3-Methyl-5,5-dimethylcyclohexenone	-2.5	1.74
Diphenyl sulfoxide	-2.54	2.04
4-Methoxyacetophenone	-3.02	1.33
Acetone	-3.06	1.18
Butan-2-one	-3.06	1.22
Methyl 4-nitrophenyl sulfoxide	-3.11	1.58
3-Methylbutan-2-one	-3.29	1.2
3-Chloro-5,5-dimethylcyclohexenone	-3.36	1.21
3,3-Dimethylbutan-2-one	-3.37	1.17
Ethyl acetate	-3.45	1.07
Acetophenone	-3.46	1.11
Dimethyl sulfone	-3.5	1.1
4-Fluoroacetophenone	-3.53	1
4-Isopropylacetophenone	-3.69	1.21
4-Ethylacetophenone	-3.71	1.25
4-Methylacetophenone	-3.78	1.24
3,5-Dimethylheptan-4-one	-3.86	1.07
Nonan-5-one	-3.89	1.21
Methyl acetate	-3.9	1
4-Tert-butylacetophenone	-3.97	1.25
4-Trifluoromethylacetophenone	-4.13	0.78
4-Nitroacetophenone	-5.02	0.57
Ethyl benzoate	-6.16	0.94
t-Butyl methyl sulfide	-6.68	0.25
2,2,4,4-Tetramethylpentan-3-one	-6.8	0.96
Dimethyl sulfide	-6.95	0.12

Methyl benzoate	-7.05	0.89
Acetonitrile	-10.1	0.91
Benzonitrile	-10.4	0.8
Chloroacetonitrile	-12.8	0.39

Table 3.8.1. The hydrogen bond acceptors used in this section. The hydrogen bond acceptors are ranked in order of Brönsted basicity (pK_{BH}^{+}) [1] from strongest to weakest.

Once again correlations have been set up between ΔE(H) and pK_{BHX}. In this section the correlations are taken from 15 base susets from within the data set of 217. The first subset includes the 15 bases with the highest pK_{BH}^{+} values. The first subset therefore includes the first 15 bases listed in table 3.8.1. The second subset is made up of the bases ranked between 2 and 16 in table 3.8.1. In other words there is a sliding window containing 15 bases. Each new window replaces the base with the largest pK_{BH}^{+} value in the previous window with the base with the next largest pK_{BH}^{+} value outside that of the previous window. For each subset r^2 values have been calculated. The r^2 values for each subset have been plotted against the mean pK_{BH}^{+} value of each subset. The results are displayed in figure 3.8.1.

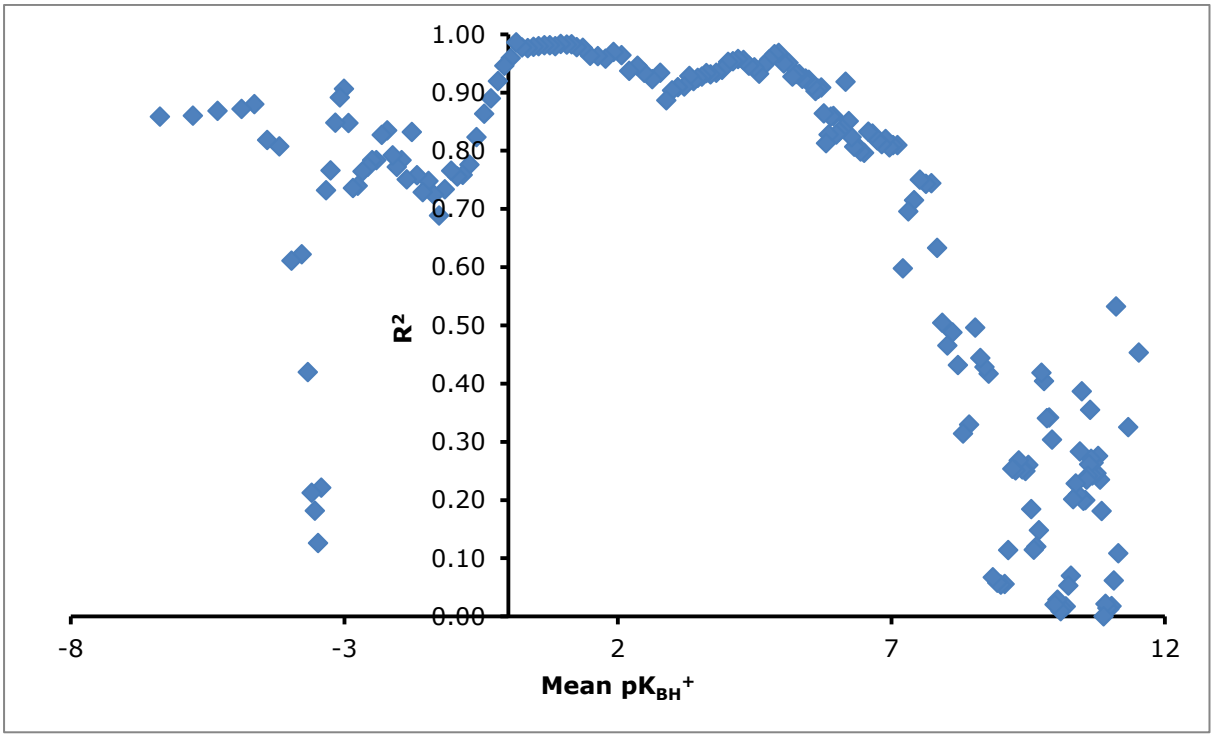


Figure 3.8.1. From the bases listed in table 3.8.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value [1] is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value [1] of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methanol is the hydrogen bond donor.

Figure 3.8.1 reveals an interesting relationship between the pK_{BH}^{+} of a base and the relationship of the bases pK_{BHX} value and its computed ΔE(H). In general it would appear that bases that have a pK_{BH}^{+} value of greater than 6 form hydrogen bond complexes that produce ΔE(H) values that do not correlate with pK_{BHX} values. Figure 3.8.1 shows that as the mean pK_{BH}^{+} value of the subsets rises above 6, the r^2 value rises to approximately 0.9. Any 15 base subset with an r^2 value of less than 0.9 can be said to display no relationship between pK_{BHX} values and computed ΔE(H). In this research the r^2 value of 0.9 has been chosen to indicate a strong correlation. The r^2 value of 0.9 has been chosen arbitrarily. However, correlations with an r^2 value of 0.9 or above can be generally thought of as strong. It is strong correlations that are of interest in this research. Strong correlations are of interest because it is important that the data sets can be accurately extrapolated and also used to accurately predict unknown values. Therefore any r^2 value of less than 0.9 can instantly be discarded. However, figure 3.8.1 shows an interesting climb in r^2 values as the mean pK_{BH}^{+} value approaches 6. It appears as if the r^2 value is improving as the mean pK_{BH}^{+} value approaches 6. An improving r^2 value implies that the relationship between ΔE(H) and pK_{BHX} for each subset as the mean pK_{BH}^{+} value approaches 6. It is very difficult to conclude that the relationship between ΔE(H) and pK_{BHX} is improving as the mean pK_{BH}^{+} value for each subset approaches 6. It is difficult to conclude an improving relationship between ΔE(H) and pK_{BHX} because any r^2 value of less than 0.9 indicates no relationship. Therefore it would matter not what the r^2 value is and only whether or not the r^2 value is above the threshold of 0.9.

It can also be seen in figure 3.8.1 that as the mean pK_{BH}^{+} value is less than -1, the r^2 value drops below 0.9. Once again r^2 values of less than 0.9 indicate that there is no relationship between computed ΔE(H) and pK_{BHX}. Subsets with a mean pK_{BH}^{+} value of less than -1 display no relationship between pK_{BHX} and ΔE(H). Therefore as the proton acceptor strengths of the hydrogen bond acceptors decrease, there is less chance of a relationship between pK_{BHX} and ΔE(H). It follows that pK_{BHX} predictions from computed ΔE(H) will be less accurate for proton acceptors with pK_{BH}^{+} values of less than -1.

It appears that there is a window of pK_{BH}^{+} values from which pK_{BHX} values can accurately be predicted from computed ΔE(H) values. The window of pK_{BH}^{+} values which allows accurate pK_{BHX} predictions ranges from -1 to 6. It can be seen from figure 3.8.1 that in the pK_{BH}^{+} range of -1 to 6, the r^2 value consistently remains above 0.9.

The poor correlations between ΔE(H) and pK_{BHX} for subsets with a mean pK_{BH}^{+} value of 6 or above is exactly what was initially hypothesised. The initial model to predict pK_{BHX} values from computed ΔE(H) values is described in section 3.2. The model broke down for tertiary amines in a case study described in section 3.6. The observation that tertiary amines are strong proton acceptors gave rise to the hypothesis that hydrogen bond acceptors with large positive pK_{BH}^{+} values will have pK_{BHX} values that are poorly predicted by computed ΔE(H) values. The hypothesis also speculated on the validity of the experimental pK_{BHX} values of strong proton acceptors. The speculation about the validity of experimental pK_{BHX} values lead to the suggestion that a protonated hydrogen bond complex could form in the experimental solution for very basic hydrogen bond acceptors. The presence of a protonated hydrogen bond acceptor could generate undesired hydrogen bond complexes that would contribute to an over-estimated pK_{BHX} value. Although the hypothesis that pK_{BHX} values would be poorly predicted for hydrogen bond acceptors with large positive pK_{BH}^{+} values has been proven to be true in this section, it is not possible to conclude why at this point in the research. To conclude that the experimental values are over estimated would be to assume that the relationship between ΔE(H) and formation constants is universal. Without a mathematical

explanation it cannot be said that there is a definitive relationship between pK_{BHX} and ΔE(H). At this stage a mathematical link between pK_{BHX} and ΔE(H) has not been looked for. In other words the observed relationship between pK_{BHX} and ΔE(H) cannot at this stage be mathematically justified. The presence of additional hydrogen bond complexes in the experimental solution of hydrogen bond acceptors with large positive pK_{BH}^{+} values must be confirmed analytically by experiment. The hypothesis that the pK_{BHX} values are overestimated due to the presence of additional protonated hydrogen bond complexes in the experimental solution offers. At this stage only a suggestion as to why the relationship between ΔE(H) and pK_{BHX} breaks down for strong proton acceptors.

Further computations can, however, be performed on the hydrogen bond acceptors with high positive pK_{BH}^{+} values. There is no relationship between pK_{BHX} and ΔE(H) for hydrogen bond acceptors with large positive pK_{BH}^{+} values. It is likely that hydrogen bond acceptors with large positive pK_{BH}^{+} values will be protonated in the experimental solution because they are very strong proton acceptors. The computed hydrogen bond complexes in the above computations were for methanol donating a hydrogen bond to one neutral hydrogen bond acceptor. As it is likely that hydrogen bond acceptors with large positive pK_{BH}^{+} values will be protonated in the experimental solution, further computations must be performed on protonated hydrogen bond acceptors.

Following the procedures outlined in section 2, hydrogen bond computations were performed. Methanol was once again used as the hydrogen bond donor. The hydrogen bond acceptors, in this case have been protonated at the hydrogen bond donor site. In the case where two possible protonation sites are present on a base, the most electronegative atom is chosen. This may or may not be the hydrogen bond acceptor site. In the case of symmetrical molecules such as 1,3-diaminopropane, the site of protonation is not chosen to be the hydrogen bond acceptor site. Figure 3.8.1 shows that the relationship between pK_{BHX} and ΔE(H) breaks down for hydrogen bond acceptors with pK_{BH}^{+} values greater than 6. The data set chosen in for this section consists of the

hydrogen bond acceptors with pK_{BH}^{+} values of 6 and above. The hydrogen bond acceptors chosen for this research are listed in table 3.8.2 along with their pK_{BH}^{+} and pK_{BHX} values. Following the methods also listed in chapter 2, ΔE(H) values have been computed.

HYDROGEN BOND ACCEPTOR	**pK_{BH}^{+}**	**pK_{BHX}**
1,1,2,3,3-Pentamethylguanidine	13.8	3.16
1,1,3,3-Tetramethylguanidine	13.6	3.21
N,N-Dimethyl-N'-propylimidoformamide	11.46	2.59
Pyrrolidine	11.306	2.59
N'-Isobutyl-N,N-dimethylimidoformamide	11.3	2.52
Azetidine	11.29	2.59
Dibutylamine	11.25	2.11
Diisopropylamine	11.2	2
1,2,2,6,6-Pentamethylpiperidine	11.19	1.23
Quinuclidine	11.1516	2.71
Piperidine	11.1236	2.38
Azepane	11.1	2.24
2,2,6,6-Tetramethylpiperidine	11.0725	1.88
N-Methylcyclohexylamine	11.04	2.24
Diethylamine	11.0151	2.25
N,N-Dimethyl-N'-(4-methylbenzyl)imidoformamide	10.91	2.36
N-methylbutylamine	10.9	2.24
Dimethylamine	10.7788	2.26
N,N-Dimethylcyclohexylamine	10.72	2.15
Triethylamine	10.7174	1.98
Tert-butylamine	10.6837	2.23
Ethylamine	10.6784	2.17
Isopropylamine	10.67	2.2
Tripropylamine	10.66	1.47
Methylamine	10.6532	2.2
N'-Benzyl-N,N-dimethylimidoformamide	10.65	2.35
n-Butylamine	10.6385	2.19
1,6-Diaminohexane	10.63	2.21
n-Octylamine	10.61	2.27
n-Hexadecylamine	10.61	2.26
c-Hexylamine-MeOH	10.58	2.29
n-Propylamine	10.5685	2.2
NN-Dimethylisopropylamine-MeOH	10.47	2.11
N-Methylpyrrolidine	10.46	2.19
N-Butylpyrrolidine	10.36	2.04
1,4-Diaminobutane	10.35	2.21

N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	10.32	2.12
1,3-Diaminopropane	10.17	2.31
N,N-Dimethylethylamine	10.16	2.17
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	10.15	2.1
N-Methylallylamine	10.11	2
N-Methylpiperidine	10.08	2.11
N,N'-dimethylethylenediamine	9.94	2.29
Tributylamine	9.93	1.55
3-Methoxypropylamine	9.92	2.22
N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	9.86	2
2-Phenylethanamine	9.83	2.16
Trimethylamine	9.7977	2.13
Ethylenediamine	9.626	2.25
2-Phenylpyrrolidine	9.6	1.93
4-N,N-Dimethylaminopyridine	9.58	2.8
N-Methylbenzylamine-MeOH	9.56	1.82
Allylamine	9.52	1.93
2-Methoxyethylamine	9.44	2.26
Piperazine	9.432	2.11
1,2,3,4-Tetrahydroisoquinoline	9.41	2.04
Benzylamine	9.34	1.84
Ammonia	9.244	1.74
Diallylamine	9.24	1.7
4-Aminopyridine	9.12	2.56
c-Propylamine	9.1	1.72
N,N-Dimethylbenzylamine	8.91	1.59
N,N,N',N'-Tetramethylethylenediamine	8.85	2.02
N-Methyl-2-phenylpyrrolidine	8.8	1.38
N,N-Dimethylallylamine	8.64	1.92
Diazabicyclooctane	8.52	2.33
Dibenzylamine	8.52	1.34
Morpholine	8.4918	1.78
N,N-Dimethyl-N'-(4-methylphenyl)imidoformamide	8.45	2.07
Triallylamine	8.28	1.34
N,N-Dimethyl-N'-(2-methylphenyl)imidoformamide	8.27	1.63
N,N-Dimethyl-N'-phenylimidoformamide	8.15	1.9
Propargylamine	8.15	1.56
3-Aminopropionitrile	7.8	1.33
Hexamethylenetetramine	7.4696	1.33
N,N-Dimethylpropargylamine	7.45	1.6
2,4,6-Trimethylpyridine	7.43	2.29
N-Methylmorpholine	7.41	1.56
N'-(4-Bromophenyl)-N,N-dimethylimidoformamide	7.4	1.65
1,1-Diphenylmethanimine	7.18	1.8
1-Methyl-1H-imidazole	7.12	2.72

N'-(4-Acetylphenyl)-N,N-dimethylimidoformamide	7.02	1.52
2-N,N-Dimethylaminopyridine	6.99	1.61
2,6-Dimethylpyridine	6.72	2.14
2-Aminopyridine	6.71	2.12
N'-(2-Bromophenyl)-N,N-dimethylimidoformamide	6.71	1.37
2-Methyl-N-(phenylmethylene)propan-2-amine	6.7	1.29
N,N-Diethylaniline	6.61	0.05
4-Methoxypyridine	6.58	2.13
3,4-Dimethylpyridine	6.47	2.24
N'-(4-Cyanophenyl)-N,N-dimethylimidoformamide	6.44	1.32
3,5-Dimethylpyridine	6.34	2.21
Thiazolidine	6.22	1.1
N,N,N',N'-Tetramethylbenzene-1,4-diamine	6.05	1.13
3-Aminopyridine	6.03	2.2
4-Methylpyridine	6.03	2.07
4-Ethylpyridine	6.02	2.07
N,N-Dimethyl-N'-(4-nitrophenyl)imidoformamide	6.02	1.2

Table 3.8.2. The hydrogen bond acceptors listed in this table have pK_{BH}^{+} values [1] of 6 or above. The hydrogen bond acceptors in this table have been found to form hydrogen bond complexes with methanol where the computed ΔE(H) value has no relationship with experimental pK_{BHX} values.

Unfortunately, there are no results to discuss for this experiment. This is due to the vast majority of the complexes failing in the geometry optimisation step. After several attempts using multiple starting geometries, it has not been possible to obtain a set of results. The protonated site of the hydrogen bond acceptor did not accept hydrogen bonds from methanol. During the optimisation step, the complex would rotate and the intended protonated hydrogen bond acceptor site became the hydrogen bond donor. It was not the intention of the experiment to investigate such a complex.

However, it can be concluded that, as a protonated hydrogen bond acceptor complex is not formed during optimisation, it is unlikely that one will form in the experimental solution. This is due to electronic reasons. For example, a R_3N hydrogen bond acceptor site becomes protonated at the site of the lone pair on the nitrogen atom. The lone pair is no longer free to accept a hydrogen bond.

Additional calculations were also performed on the protonated hydrogen bond acceptors listed in table 3.8.2. Hydrogen bond computations were performed as above.

The focus in this section is the site of protonation. In the case where there are two potential protonation sites the most electronegative atom was previously chosen. In this section one data set takes the hydrogen bond acceptor site to be the major site of protonation whereas the other data set takes the site that is not the hydrogen bond acceptor site as the site of protonation. The non-symmetrical hydrogen bond acceptors with 2 potential hydrogen bond acceptor sites are listed in table 3.8.3.

HYDROGEN BOND DONOR	**pK_{BH}^{+}**	**pK_{BHX}**
N,N-Dimethyl-N'-propylimidoformamide	11.46	2.59
N'-Isobutyl-N,N-dimethylimidoformamide	11.3	2.52
N,N-Dimethyl-N'-(4-methylbenzyl)imidoformamide	10.91	2.36
N'-Benzyl-N,N-dimethylimidoformamide	10.65	2.35
N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	10.32	2.12
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	10.15	2.1
3-Methoxypropylamine	9.92	2.22
N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	9.86	2
4-N,N-Dimethylaminopyridine	9.58	2.8
Piperazine	9.432	2.11
4-Aminopyridine	9.12	2.56
Morpholine	8.4918	1.78
N,N-Dimethyl-N'-(4-methylphenyl)imidoformamide	8.45	2.07
N,N-Dimethyl-N'-(2-methylphenyl)imidoformamide	8.27	1.63
N,N-Dimethyl-N'-phenylimidoformamide	8.15	1.9
3-Aminopropionitrile	7.8	1.33
N-Methylmorpholine	7.41	1.56
1-Methyl-1H-imidazole	7.12	2.72
N'-(4-Acetylphenyl)-N,N-dimethylimidoformamide	7.02	1.52
2-Aminopyridine	6.71	2.12
N'-(2-Bromophenyl)-N,N-dimethylimidoformamide	6.71	1.37
4-Methoxypyridine	6.58	2.13
N'-(4-Cyanophenyl)-N,N-dimethylimidoformamide	6.44	1.32
N,N,N',N'-Tetramethylbenzene-1,4-diamine	6.05	1.13
3-Aminopyridine	6.03	2.2
N,N-Dimethyl-N'-(4-nitrophenyl)imidoformamide	6.02	1.2

Table 3.8.3. The hydrogen bond acceptors listed in this table have two possible sites of protonation. One of the sites is the hydrogen bond acceptor site; the other site is independent of the hydrogen bond acceptor site. The hydrogen bond acceptors listed in this table have been taken from table 3.8.2. Symmetrical molecules with two possible sites of protonation such as 1,3-dipropylamine have been excluded from this table. Only hydrogen bond acceptors with two potential sites of protonation that are housed in chemically different environments have been included. pK_{BHX} and pK_{BH}^{+} values taken from [1].

Once again it has not been possible to obtain a set of results for this experiment. The reasons are the same as for the last experiment and discussed above. Reasons are largely concerned with the complex rotating in the geometry optimisation step.

Therefore, as in the last experiment, it is unlikely that this type of complex will form in the experimental solution.

It is also possible that the protonated site on the hydrogen bond acceptor could donate a hydrogen bond to the oxygen atom of methanol. It has been shown that for a data set consisting of protonated tertiary amines that the pK_{BHX} value of the neutral tertiary amines cannot be predicted from ΔE(H) values calculated from hydrogen bond complexes where the oxygen atom of methanol accepts a hydrogen bond from a protonated tertiary amine (section 3.7). In fact the ΔE(H) value can be used to predict only the pK_{BHX} value of methanol. In other words, protonated tertiary amines can be used as hydrogen bond donors in hydrogen bond complexes where computed ΔE(H) values successfully correlate with pK_{BHX} values. It would therefore not be expected that pK_{BHX} values of protonated hydrogen bond acceptors would relate to ΔE(H) values calculated from complexes where the protonated hydrogen bond acceptor is donating a hydrogen bond.

The relationship between pK_{BHX} and ΔE(H) breaks down for hydrogen bond acceptors with pK_{BH}^{+} values of 6 or above. The finding that pK_{BHX} values do not correlate with ΔE(H) for hydrogen bond acceptors that are strong proton acceptor supports the initial hypothesis. The initial hypothesis suggested that strong proton acceptors would

form hydrogen bond complexes that will produce ΔE(H) values from which pK_{BHX} values cannot be successfully predicted. The hypothesis also further speculated the possibility of protonated hydrogen bond acceptors forming reversed hydrogen bond complexes in which the protonated hydrogen bond acceptors donate hydrogen bonds. A link between ΔE(H) and pK_{BHX} has not been found for such complexes.

The final set of calculations on the topic of protonated hydrogen bond complexes examines the possibility of protonation of the hydrogen bond donor. It is possible that the hydrogen bond donor itself could become protonated in the experimental solution. It is also possible that a protonated hydrogen bond donor could become a proton donor to the hydrogen bond acceptors. The hydrogen bond donors commonly used experimentally are methanol and 4-fluorophenol. The hydrogen bond acceptors of interest in this section are strong proton acceptors with pK_{BH}^{+} values of 9 or above. As the hydrogen bond acceptors used in this section are much stronger than methanol which has a pK_{BH}^{+} value of -2.03, any proton carried by methanol would be transferred to the hydrogen bond acceptor in the experimental solution. A different type of hydrogen bond complex would be formed in which the complex is protonated and the proton is shared between the oxygen of methanol and the hydrogen bond acceptor site. The proton takes on the role of the hydrogen bond donor. The source of the hydrogen bond donor may be either methanol or the protonated hydrogen bond acceptor site. Since previous attempts to correlate pK_{BHX} values with ΔE(H) values taken from hydrogen bond complexes where a protonated hydrogen bond acceptor donates a hydrogen bond have failed, the source of the hydrogen bond in this section will be protonated methanol. Only the strongest proton acceptors have been investigated in this section. The proton acceptors chosen have pK_{BHX} values of 9 or above and have been listed in table 3.8.4.

HYDROGEN BOND ACCEPTOR	**pK_{BH}^{+}**	**pK_{BHX}**
1,1,2,3,3-Pentamethylguanidine	13.8	3.16
1,1,3,3-Tetramethylguanidine	13.6	3.21
N,N-Dimethyl-N'-propylimidoformamide	11.46	2.59
Pyrrolidine	11.306	2.59

N'-Isobutyl-N,N-dimethylimidoformamide	11.3	2.52
Azetidine	11.29	2.59
Dibutylamine	11.25	2.11
Diisopropylamine	11.2	2
1,2,2,6,6-Pentamethylpiperidine	11.19	1.23
Quinuclidine	11.1516	2.71
Piperidine	11.1236	2.38
Azepane	11.1	2.24
2,2,6,6-Tetramethylpiperidine	11.0725	1.88
N-Methylcyclohexylamine	11.04	2.24
Diethylamine	11.0151	2.25
N,N-Dimethyl-N'-(4-methylbenzyl)imidoformamide	10.91	2.36
N-methylbutylamine	10.9	2.24
Dimethylamine	10.7788	2.26
N,N-Dimethylcyclohexylamine	10.72	2.15
Triethylamine	10.7174	1.98
Tert-butylamine	10.6837	2.23
Ethylamine	10.6784	2.17
Isopropylamine	10.67	2.2
Tripropylamine	10.66	1.47
Methylamine	10.6532	2.2
N'-Benzyl-N,N-dimethylimidoformamide	10.65	2.35
n-Butylamine	10.6385	2.19
1,6-Diaminohexane	10.63	2.21
n-Octylamine	10.61	2.27
c-Hexylamine-MeOH	10.58	2.29
n-Propylamine	10.5685	2.2
NN-Dimethylisopropylamine-MeOH	10.47	2.11
N-Methylpyrrolidine	10.46	2.19
N-Butylpyrrolidine	10.36	2.04
1,4-Diaminobutane	10.35	2.21
N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	10.32	2.12
1,3-Diaminopropane	10.17	2.31
N,N-Dimethylethylamine	10.16	2.17
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	10.15	2.1
N-Methylallylamine	10.11	2
N-Methylpiperidine	10.08	2.11
N,N'-dimethylethylenediamine	9.94	2.29
Tributylamine	9.93	1.55
3-Methoxypropylamine	9.92	2.22
N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	9.86	2
2-Phenylethanamine	9.83	2.16
Trimethylamine	9.7977	2.13
Ethylenediamine	9.626	2.25
2-Phenylpyrrolidine	9.6	1.93

4-N,N-Dimethylaminopyridine	9.58	2.8
N-Methylbenzylamine-MeOH	9.56	1.82
Allylamine	9.52	1.93
2-Methoxyethylamine	9.44	2.26
Piperazine	9.432	2.11
1,2,3,4-Tetrahydroisoquinoline	9.41	2.04
Benzylamine	9.34	1.84
Ammonia	9.244	1.74
Diallylamine	9.24	1.7
4-Aminopyridine	9.12	2.56
c-Propylamine	9.1	1.72

Table 3.8.4. The hydrogen bond acceptors listed in this table form the data set for a set of calculations in which protonated methanol is donating a hydrogen. The hydrogen bond acceptors in this table are the strongest proton acceptors and have pK_{BH}^{+} [1] values of 9 or above.

Hydrogen bond calculations have been set up following the methods outlined in chapter 2. The hydrogen bond donor is protonated methanol. The ΔE(H) values have been computed for the hydrogen bond donor also following the methods outlined in chapter 2.

Correlations have been set up between pK_{BHX} and ΔE(H). It can be seen from figure 1 that there is no relationship between pK_{BHX} and ΔE(H) for hydrogen bond complexes formed between protonated methanol and hydrogen bond acceptors with pK_{BH}^{+} values of 9 or above. The r^2 value of 0.0016 indicates that there is not only no relationship between pK_{BHX} and ΔE(H), but that the data is randomly scattered. There is no general trend in the data. It can, however, be seen that there are clusters of trends in the data. Sub-correlations within the data set appear visible. The sub-correlations appear particularly noticeable in the ΔE(H) ranges of 0.3 – 0.34, and 0.36 – 0.40. In order to further investigate the possibility of correlating subsets within the data set, further analysis has been performed.

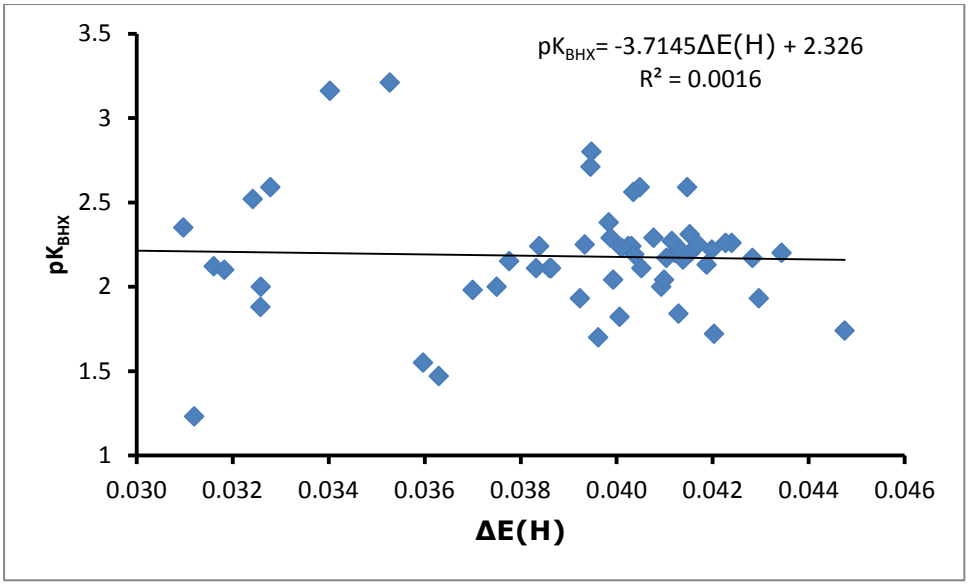


Figure 3.8.2. The correlation between ΔE(H) and pK_{BHX} [1] for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4.

The analysis follows the methods described in previously to distribute the data set into subsets. The subsets consist of 15 bases. The first set contains the 15 bases in table 3.8.4 with the largest positive pK_{BH}^+ values. The next set is made by sliding the 15 base window down table 3.8.4 by one base replacing the base with the largest pK_{BH}^+ value with the base with the highest pK_{BH}^+ value outside the first set. Correlations between ΔE(H) and pK_{BHX} have been set up for each 15 base subset. The r^2 values from the correlations of each 15 base subset have been plotted against the mean pK_{BH}^+ value of the subset. The results are displayed in figure 3.8.3.

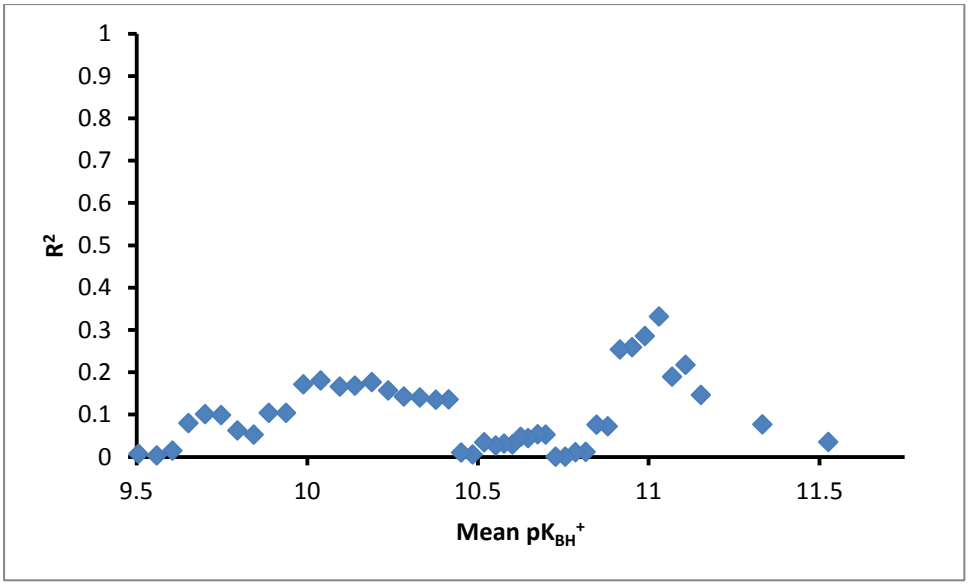


Figure 3.8.3. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values. The method used to compose the subsets is described in the text above. Correlations between ΔE(H) and pK_{BHX} have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset.

Figure 3.8.3 shows clearly that there is no recovery in the relationship between ΔE(H) and pK_{BHX} as the pK_{BH}^{+} value of the hydrogen bond acceptors decreases. The potential sub correlations observed in figure 3.8.3 cannot be found by ranking the hydrogen bond acceptors in order of pK_{BH}^{+} and distributing the data set into 15 base subsets. The possible sub-correlations observed in figure 3.8.3 were found in certain ΔE(H) ranges. The fact that pK_{BHX} and pK_{BH}^{+} are not related explains why the sub-correlations described above cannot be recovered by ranking the data in order of pK_{BH}^{+}.

The relationship between ΔE(H) and pK_{BHX} cannot be recovered by modelling the hydrogen bond complex with the hydrogen bond donor methanol in a protonated state. The values used to compute the ΔE(H) value used in the above analysis is taken as the difference between the energy of the proton on free protonated methanol and the energy of the hydrogen bonded hydrogen in the complex. However, as the protonated methanol forms a hydrogen bond complex with each hydrogen bond acceptor listed in table 3.8.4, the proton is transferred to the hydrogen bond acceptor. The hydrogen bond acceptor

assumes the role of the hydrogen bond donor by donating a hydrogen bond to methanol. It can therefore be said that methanol, although the initial hydrogen bond donor, assumes the role of the hydrogen bond acceptor. Therefore ΔE(H) values could be calculated by subtracting the energy of the hydrogen atom forming the hydrogen bond in the complex with the energy of a hydrogen atom in free neutral energy. The form of ΔE(H) calculated in the way described above will be called $\Delta E(H)_{(MeOH\text{-}HBD)}$. The values of $\Delta E(H)_{(MeOH\text{-}HBD)}$ have been substituted into figures 3.8.2 and 3.8.3 and the results are shown in figures 3.8.4 and 3.8.5.

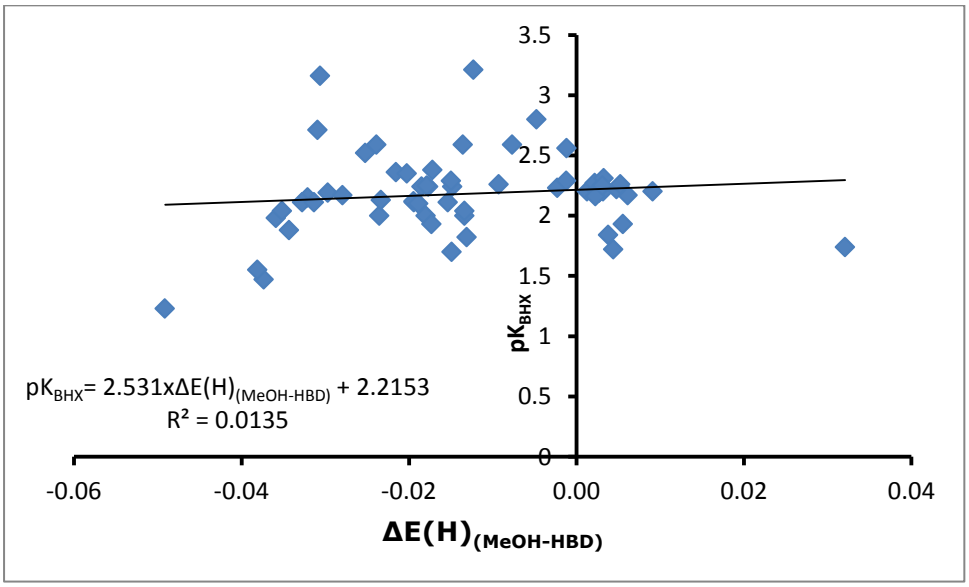


Figure 3.8.4. The correlation between $\Delta E(H)_{(MeOH\text{-}HBD)}$ and pK_{BHX} [1] for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. The meaning of $\Delta E(H)_{(MeOH\text{-}HBD)}$ is described in the text above.

It can be seen in figure 3.8.4 that there is no relationship between $\Delta E(H)_{(MeOH\text{-}HBD)}$ and pK_{BHX} for the bases listed in table 3.8.4. Once again the r^2 value of 0.0135 indicates a random scattering of the data points.

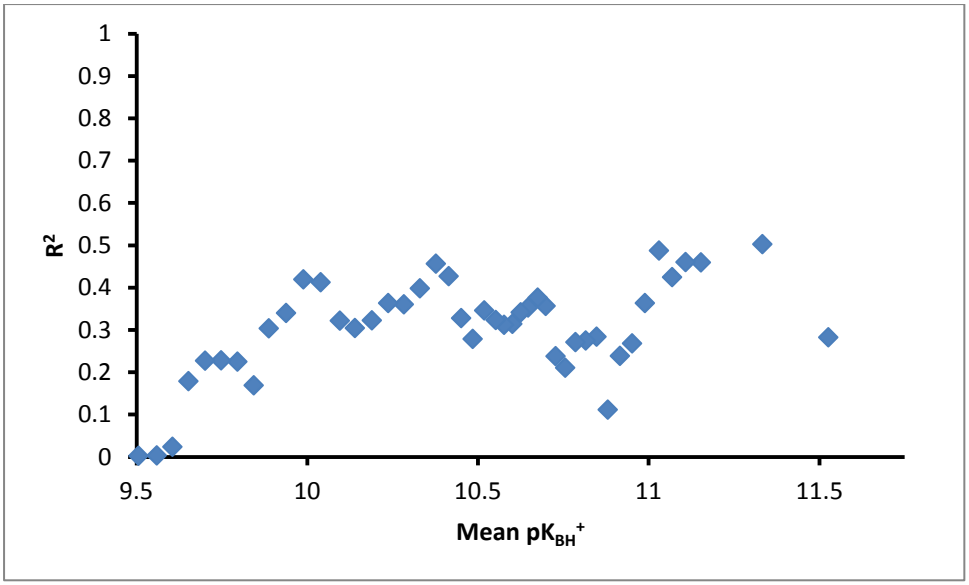


Figure 3.8.5. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values [1]. The method used to compose the subsets is described in the text above. Correlations between $\Delta E(H)_{(MeOH\text{-}HBD)}$ and pK_{BHX} [1] have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. The meaning of $\Delta E(H)_{(MeOH\text{-}HBD)}$ is described in the text above.

There is no developing relationship between pK_{BHX} and $\Delta E(H)_{(MeOH\text{-}HBD)}$ for hydrogen bond acceptors with decreasing proton acceptor strength. Figure 3.8.5 shows that as the mean pK_{BH}^{+} value of each 15 base subset decreases the r^2 value of the correlation between pK_{BHX} and $\Delta E(H)_{(MeOH\text{-}HBD)}$ remains well below the threshold value of 0.9.

Using the same set of hydrogen bond donors listed in table 3.8.4, it would be interesting to see if two further variations of ΔE(H) can be substituted into figures 3.8.2 and 3.8.3 in an attempt to find a relationship between ΔE(H) and pK_{BHX}. The first variation of ΔE(H) is taken to be the difference in energy between the hydrogen atom in free methanol and the hydrogen atom in methanol in the complex. The hydrogen atom in methanol in the complex is not the hydrogen bond donor. This variation of ΔE(H) will be called $\Delta E(H)_{(Methanol)}$. The results of substituting $\Delta E(H)_{(Methanol)}$ into figures 3.8.2 and 3.8.3 are displayed in figures 3.8.6 and 3.8.7.

The second variation of ΔE(H) is taken to be the difference in energy between the hydrogen in methanol in the complex and the hydrogen bond donor in the complex. This version of ΔE(H) will be called $\Delta E(H)_{(Complex)}$. The results of substituting $\Delta E(H)_{(Complex)}$ into figures 3.8.2 and 3.8.3 are shown in figures 3.8.8 and 3.8.9.

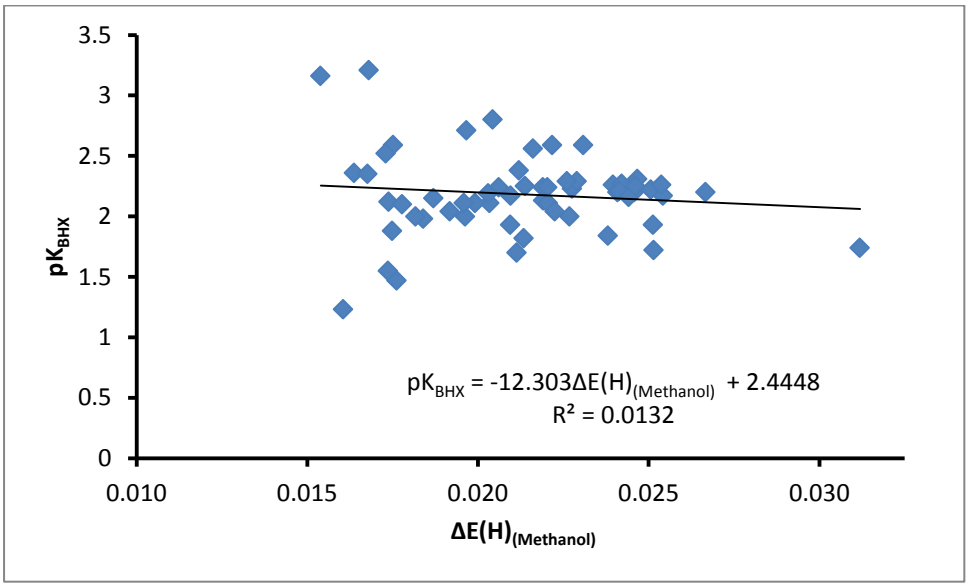


Figure 3.8.6. The correlation between $\Delta E(H)_{(Methanol)}$ and pK_{BHX} [1] for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. The meaning of $\Delta E(H)_{(Methanol)}$ is described in the text above.

It can be seen in figure 3.8.6 that there is no relationship between $\Delta E(H)_{(Methanol)}$ and pK_{BHX} for the bases listed in table 3.8.4. Once again the r^2 value of 0.0132 indicates a random scattering of the data points.

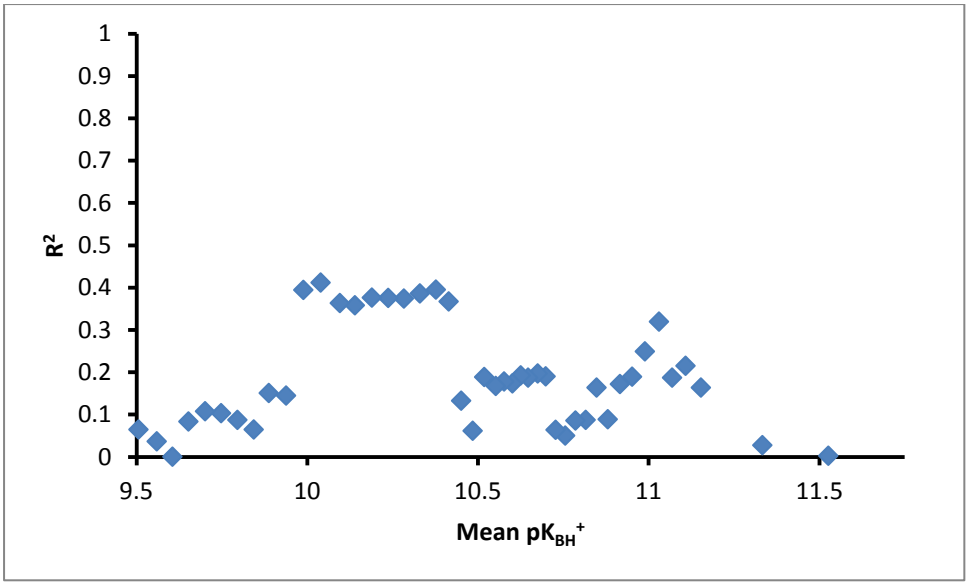


Figure 3.8.7. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values [1]. The method used to compose the subsets is described in the text above. Correlations between $\Delta E(H)_{(Methanol)}$ and pK_{BHX} [1] have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. The meaning of $\Delta E(H)_{(Methanol)}$ is described in the text above.

There is no developing relationship between pK_{BHX} and $\Delta E(H)_{(Methanol)}$ for hydrogen bond acceptors with decreasing proton acceptor strength. Figure 3.8.7 shows that as the mean pK_{BH}^{+} value of each 15 base subset decreases; the r^2 value of the correlation between pK_{BHX} and $\Delta E(H)_{(Methanol)}$ remains well below the threshold value of 0.9.

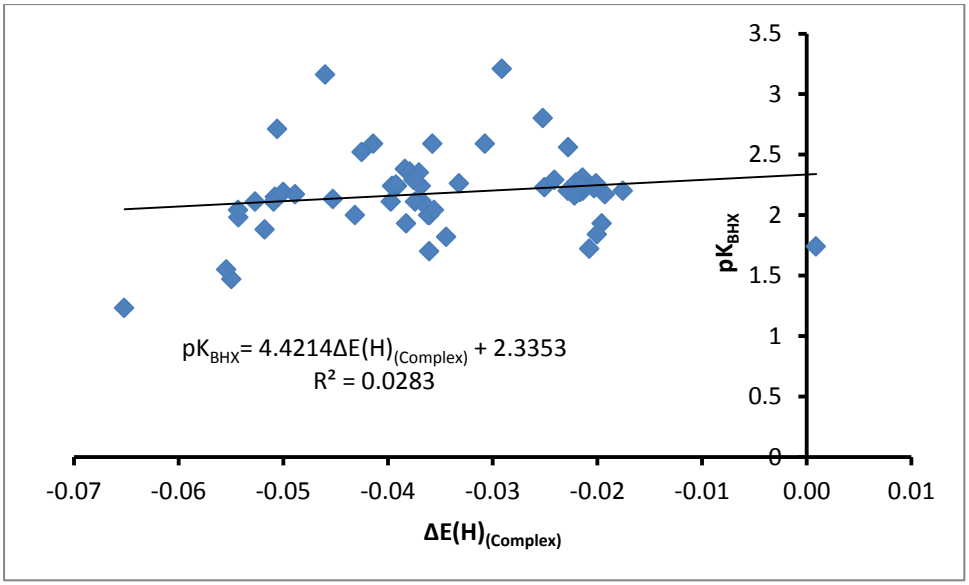


Figure 3.8.8. The correlation between $\Delta E(H)_{(Complex)}$ and pK_{BHX} [1] for hydrogen bond complexes where protonated methanol donates a hydrogen bond to the hydrogen bond acceptors listed in table 3.8.4. The meaning of $\Delta E(H)_{(Complex)}$ is described in the text above.

It can be seen in figure 3.8.8 that there is no relationship between $\Delta E(H)_{(Complex)}$ and pK_{BHX} for the bases listed in table 3.8.4. Once again the r^2 value of 0.0283 indicates a random scattering of the data points.

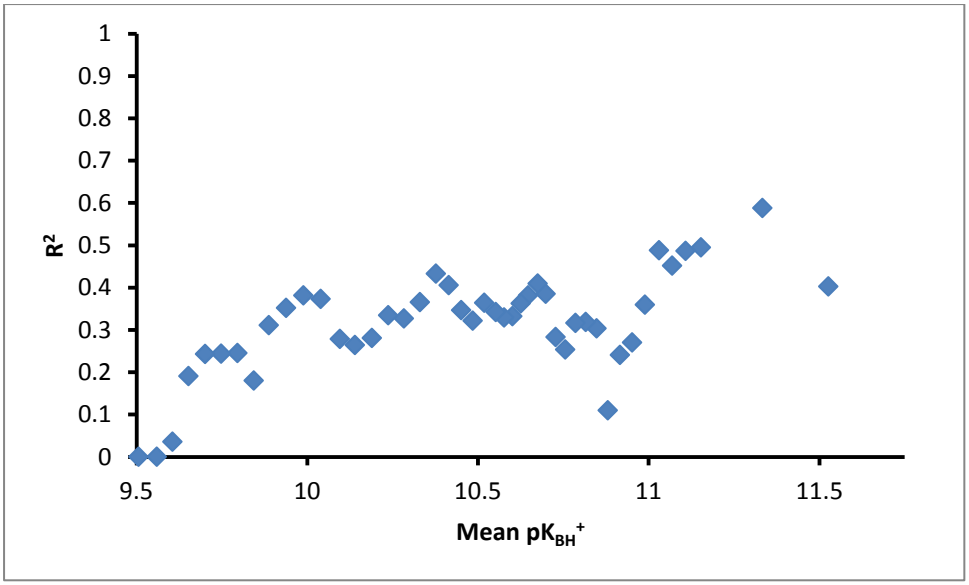


Figure 3.8.9. The hydrogen bond acceptors listed in table 3.8.4 have been distributed into 15 base subsets according to their pK_{BH}^{+} values [1]. The method used to compose the subsets is described in the text above. Correlations between $\Delta E(H)_{(Complex)}$ and pK_{BHX} [1] have been set up for each subset and the resulting r^2 value is plotted against the mean pK_{BH}^{+} value of the subset. The meaning of $\Delta E(H)_{(Complex)}$ is described in the text above.

There is no developing relationship between pK_{BHX} and $\Delta E(H)_{(Complex)}$ for hydrogen bond acceptors with decreasing proton acceptor strength. Figure 3.8.9 shows that as the mean pK_{BH}^{+} value of each 15 base subset decreases, the r^2 value of the correlation between pK_{BHX} and $\Delta E(H)_{(Complex)}$ remains well below the threshold value of 0.9.

The research in this section has provided strong evidence to support the hypothesis that the relationship between pK_{BHX} and $\Delta E(H)$ will break down when the hydrogen bond acceptors have large positive pK_{BH}^{+} values. However, it is unfortunate that at this stage the reason why the relationship between pK_{BHX} and $\Delta E(H)$ breaks down for hydrogen bond acceptors with large positive pK_{BH}^{+} values is unknown. The hypothesis that possible protonation of the hydrogen bond acceptor in the experimental solution may contribute to the breakdown of the relationship between pK_{BHX} and $\Delta E(H)$ for hydrogen bond donors with large positive pK_{BH}^{+} values has been extensively tested in this section. Not one of the computations involving protonated hydrogen bond complexes have given results that recovered the relationship between pK_{BHX} and $\Delta E(H)$. However,

the observation that the relationship between pK_{BHX} and ΔE(H) will break down for hydrogen bond acceptors with large positive pK_{BH}^{+} values is conclusive.

It can also be noticed from figure 3.8.1 that the relationship between pK_{BHX} and ΔE(H) appears to break down when the pK_{BH}^{+} values of the hydrogen bond donors are lower (more negative) than -0.5. Therefore the relationship between pK_{BHX} and ΔE(H) appears to be valid for hydrogen bond donors with pK_{BH}^{+} values between -0.5 and 6.0. Throughout the course of this research 332 hydrogen bond donors have been used to accept hydrogen bonds from methanol in computations. The 332 hydrogen bond donors are listed in table 3.8.5. Also listed in table 3.8.5 are the pK_{BHX} values of each hydrogen bond acceptor, the ΔE(H) values and the Z values. The Z values are the values taken from a Grubbs test on the relationship between pK_{BHX} and ΔE(H) for the entire data set. The Z values are calculated as per equation 2.4.5.

The critical value for the data set in table 3.8.5 (N=332) = 3.75. There were no significant outliers detected. The lack of outliers is reflected in the correlation between pK_{BHX} and ΔE(H) for the 332 bases listed in table 2. Figure 3.8.10 reveals an R^2 value of 0.85 for the correlation between pK_{BHX} and ΔE(H) for all 332 hydrogen bond donors. Therefore, despite the presence of hydrogen bond donors in the data set with pK_{BH}^{+} values of less than -0.5 and greater than 6.0, a generally good relationship is observed between pK_{BHX} and ΔE(H).

HYDROGEN BOND DONOR	ΔE(H)	pK_{BHX}	Z
3-Chloropyridine	0.02948	1.31	0.01146
4-Methylpyridine	0.03314	2.07	0.73235
Acetamide	0.03282	2.06	0.66691
Acetone	0.02864	1.18	0.18346
Acetonitrile	0.02635	0.91	0.64907
Acrylonitrile	0.02562	0.7	0.79738
Aniline	0.02327	0.46	1.27475
Chloroacetnitrile	0.02317	0.39	1.2956
Dimethyl sulfide	0.01727	0.12	2.49558
Dimethylamine	0.03129	2.26	0.35684
Ethanol	0.02613	1.02	0.69325
Ethyl thiol	0.01653	-0.16	2.64636
Ethylamine	0.03249	2.17	0.59981

Formamide	0.0284	1.75	0.23243
MeSCN	0.02625	0.73	0.66923
Methanol	0.02518	0.82	0.8859
Methyl acetate	0.02843	1	0.22493
Methylamine	0.03207	2.2	0.51477
Methylformate	0.02394	0.65	1.13841
N-Methylaniline	0.02151	0.26	1.63421
Phenol	0.02084	-0.07	1.76988
Pyridine	0.0319	1.86	0.48005
Pyrrolidine	0.03243	2.59	0.58931
t-Butylamine	0.03389	2.23	0.88566
Tetrahydropyran	0.02855	1.23	0.20052
2,6-dimethylaniline	0.02418	0.47	1.09074
3-chloroaniline	0.02141	0.13	1.65447
3-fluoroaniline	0.02145	0.2	1.64691
3-methylphenol	0.02125	0.01	1.68736
4-methylphenol	0.02149	0.03	1.63724
Dimethyldisulfide	0.01616	-0.49	2.72225
Ethylmethylsulfide	0.01803	0.18	2.34277
p-toluidine	0.02419	0.56	1.08823
4-Fluorophenol	0.02013	-0.12	1.91391
4-Bromo-NN-dimethylaniline	0.01909	0.17	2.12613
Oxydibenzene-MeOH	0.01971	-0.8	2.00051
Diethyldisulfide	0.01643	-0.4	2.66831
4-Aminopyridine	0.0352	2.56	1.15295
NNN-trimethylammoniopropanamidate	0.0413	3.59	2.3928
Triethylarsine oxide	0.04578	4.89	3.3054
Trimethylamine oxide	0.04668	5.46	3.48905
Methyl nicotinate_N	0.02859	1.44	0.19205
Methyl nicotinate_O	0.02608	0.51	0.70309
3-Benzoyl pyridine-N	0.03056	1.42	0.20747
3-Benzoyl pyridine-O	0.02663	0.68	0.59186
4-Acetylpyridine-N	0.02999	1.41	0.09254
4-Acetylpyridine-O	0.02731	0.78	0.45345
Ethyl 4-cyanobenzoate-N	0.02554	0.66	0.8134
Ethyl 4-cyanobenzoate-O	0.02386	0.53	1.15538
S-cotinine-N	0.03116	1.62	0.32928
S-cotinine-O	0.0331	2.16	0.72489
4-Acetylbenzonitrile-N	0.02547	0.65	0.82844
4-Acetylbenzonitrile-O	0.02689	0.6	0.53875
3-Acetylpyridine-N	0.02974	1.39	0.04038
3-Acetylpyridine-O	0.02818	0.9	0.2759
NN-diethylnicotinamide-N	0.03108	1.63	0.31434
NN-diethylnicotinamide-O	0.03334	1.98	0.77385
4-Cyanopyridine-Pyr	0.02758	0.92	0.39824

4-Cyanopyridine-Nit	0.02425	0.47	1.07676
3-Cyanopyridine-Pyr	0.02619	0.82	0.68063
3-Cyanopyridine-Nit	0.02511	0.53	0.90172
2-Cyanopyridine-Pyr	0.02432	0.48	1.06223
2-Cyanopyridine-Nit	0.02569	0.61	0.7829
4-Chloropyridine	0.03067	1.54	0.23127
4-Ethylpyridine	0.03321	2.07	0.74633
Benzaldehyde	0.02559	0.78	0.80397
Benzylamine	0.03207	1.84	0.51459
Butan-2-one	0.02896	1.22	0.11858
Cyanamide	0.02809	1.56	0.29506
Dimethylformamide	0.03104	2.1	0.30568
Isopropanethiol	0.01685	-0.1	2.58168
Isopropylamine	0.03334	2.2	0.77303
Propanol	0.02622	1	0.67548
Propionitrile	0.02725	0.93	0.46543
Propiophenone	0.03007	1.04	0.10852
trichloroacetonitrile	0.02013	-0.26	1.91469
2,4,6-Trimethylpyridine	0.0348	2.29	1.07149
2-Aminopyridine	0.03766	2.12	1.65289
2-Chloropyridine	0.02693	1.05	0.53005
3-Aminopyridine	0.03311	2.2	0.72757
4-Chloro-N-Methylaniline	0.02032	0.05	1.87607
5-Bromopyrimidine	0.0261	0.59	0.69908
Ammonia	0.03126	1.74	0.35049
Anisolel	0.02171	-0.07	1.59361
Benzylmethylsulfide	0.01883	-0.02	2.17938
Isoquinoline	0.03246	1.94	0.59479
N-Methylbenzamide	0.03402	2.03	0.91215
N-Methylpropanamide	0.03488	2.24	1.08648
1,3,5-Triazine	0.02456	0.32	1.01283
4,6-Dimethylpyrimidine	0.03152	1.47	0.40311
4-Chlorobutyronitrile	0.02616	0.83	0.68702
Cyanicbromide	0.02291	0.19	1.348
Cyclooctanone	0.02962	1.45	0.01636
Diisopropylether	0.02983	1.11	0.05952
Ethyl Chloroacetate	0.02619	0.67	0.68086
Ethylenesulfite	0.02737	0.87	0.44035
Isoxazole	0.02664	0.81	0.59003
N-Formylmorpholine	0.03058	1.93	0.21159
Phenylcyanate	0.02625	0.77	0.66869
Phthalazine	0.03214	1.97	0.52942
Pyridazine	0.03028	1.65	0.15039
Pyrrolidine-1-Carbonitrile	0.03192	1.66	0.48561
Progesterone	0.03242	1.75	0.58741

1,1,2,3,3-Pentamethylguanidine	0.04004	3.16	2.1381
1,1,3,3-Tetramethylguanidine	0.04033	3.21	2.19598
N,N-Dimethyl-N'-propylimidoformamide	0.03735	2.59	1.58922
Pyrrolidine	0.03243	2.59	0.58931
N'-Isobutyl-N,N-dimethylimidoformamide	0.03698	2.52	1.51418
Azetidine	0.03281	2.59	0.66677
Dibutylamine	0.03295	2.11	0.69524
Diisopropylamine	0.03302	2	0.70835
1,2,2,6,6-Pentamethylpiperidine	0.03079	1.23	0.25381
Quinuclidine	0.02944	2.71	0.01961
Piperidine	0.03269	2.38	0.64094
Azepane	0.03162	2.24	0.42379
2,2,6,6-Tetramethylpiperidine	0.03175	1.88	0.44988
N-Methylcyclohexylamine	0.03273	2.24	0.64916
Diethylamine	0.03223	2.25	0.54731
N,N-Dimethyl-N'-(4-methylbenzyl)imidoformamide	0.03573	2.36	1.25954
N-methylbutylamine	0.0321	2.24	0.52195
Dimethylamine	0.03129	2.26	0.35684
N,N-Dimethylcyclohexylamine	0.03207	2.15	0.51485
Triethylamine	0.03152	1.98	0.40234
Tert-butylamine	0.03389	2.23	0.88566
Ethylamine	0.03249	2.17	0.59981
Isopropylamine	0.03346	2.2	0.79743
Tripropylamine	0.03165	1.47	0.43072
Methylamine	0.03207	2.2	0.51477
N'-Benzyl-N,N-dimethylimidoformamide	0.03531	2.35	1.1742
n-Butylamine	0.03303	2.19	0.70986
1,6-Diaminohexane	0.03314	2.21	0.7332
n-Octylamine	0.03285	2.27	0.67316
n-Hexadecylamine	0.03289	2.26	0.68127
c-Hexylamine-MeOH	0.03372	2.29	0.85126
n-Propylamine	0.03295	2.2	0.69348
NN-Dimethylisopropylamine-MeOH	0.03121	2.11	0.3406
N-Methylpyrrolidine	0.03071	2.19	0.23807
N-Butylpyrrolidine	0.03049	2.04	0.19355
1,4-Diaminobutane	0.03323	2.21	0.75207
N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	0.0343	2.12	0.96948
1,3-Diaminopropane	0.03298	2.31	0.70076
N,N-Dimethylethylamine	0.03023	2.17	0.14114
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	0.03348	2.1	0.80239
N-Methylallylamine	0.03116	2	0.32934
N-Methylpiperidine	0.03136	2.11	0.3701
N,N'-dimethylethylenediamine	0.03234	2.29	0.57007
Tributylamine	0.03195	1.55	0.49007
3-Methoxypropylamine	0.03182	2.22	0.4642

N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	0.03272	2	0.648
2-Phenylethanamine	0.03257	2.16	0.61667
Trimethylamine	0.0301	2.13	0.11436
Ethylenediamine	0.03296	2.25	0.69715
2-Phenylpyrrolidine	0.03211	1.93	0.52421
4-N,N-Dimethylaminopyridine	0.03654	2.8	1.42517
N-Methylbenzylamine-MeOH	0.03184	1.82	0.46865
Allylamine	0.03151	1.93	0.40138
2-Methoxyethylamine	0.03202	2.26	0.50502
Piperazine	0.03146	2.11	0.39138
1,2,3,4-Tetrahydroisoquinoline	0.02326	2.04	1.27732
Benzylamine	0.03273	1.84	0.64876
Ammonia	0.03126	1.74	0.35049
Diallylamine	0.02986	1.7	0.06646
4-Aminopyridine	0.0352	2.56	1.15295
c-Propylamine	0.03154	1.72	0.40723
N,N-Dimethylbenzylamine	0.03026	1.59	0.14779
N,N,N',N'-Tetramethylethylenediamine	0.03042	2.02	0.17856
N-Methyl-2-phenylpyrrolidine	0.03015	1.38	0.12441
N,N-Dimethylallylamine	0.02964	1.92	0.02012
Diazabicyclooctane	0.0334	2.33	0.78507
Dibenzylamine	0.03251	1.34	0.60416
Morpholine	0.03097	1.78	0.29073
N,N-Dimethyl-N'-(4-methylphenyl)imidoformamide	0.03323	2.07	0.75185
Triallylamine	0.02989	1.34	0.07259
N,N-Dimethyl-N'-(2-methylphenyl)imidoformamide	0.03199	1.63	0.49945
N,N-Dimethyl-N'-phenylimidoformamide	0.03258	1.9	0.61872
Propargylamine	0.0294	1.56	0.02905
3-Aminopropionitrile	0.02688	1.33	0.54175
Hexamethylenetetramine	0.03004	1.33	0.10262
N,N-Dimethylpropargylamine	0.02839	1.6	0.23297
2,4,6-Trimethylpyridine	0.0348	2.29	1.07149
N-Methylmorpholine	0.02973	1.56	0.03806
N'-(4-Bromophenyl)-N,N-dimethylimidoformamide	0.03227	1.65	0.55575
1,1-Diphenylmethanimine	0.03456	1.8	1.02162
1-Methyl-1H-imidazole	0.0365	2.72	1.41679
N'-(4-Acetylphenyl)-N,N-dimethylimidoformamide	0.02952	1.52	0.00345
2-N,N-Dimethylaminopyridine	0.03142	1.61	0.38202
2,6-Dimethylpyridine	0.03384	2.14	0.87502
2-Aminopyridine	0.03766	2.12	1.65289
N'-(2-Bromophenyl)-N,N-dimethylimidoformamide	0.0312	1.37	0.33873
2-Methyl-N-(phenylmethylene)propan-2-amine	0.03278	1.29	0.65895
N,N-Diethylaniline	0.01522	0.05	2.91417
4-Methoxypyridine	0.03388	2.13	0.88369
3,4-Dimethylpyridine	0.03371	2.24	0.84883

N'-(4-Cyanophenyl)-N,N-dimethylimidoformamide	0.028	1.32	0.31321
3,5-Dimethylpyridine	0.03352	2.21	0.80994
Thiazolidine	0.02762	1.1	0.3909
N,N,N',N'-Tetramethylbenzene-1,4-diamine	0.02621	1.13	0.67699
3-Aminopyridine	0.03311	2.2	0.72757
4-Methylpyridine	0.03027	2.07	0.14819
4-Ethylpyridine	0.03324	2.07	0.75341
N,N-Dimethyl-N'-(4-nitrophenyl)imidoformamide	0.02709	1.2	0.49903
4-Tert-butylpyridine	0.03351	2.11	0.80795
2-Methylpyridine	0.03308	2.03	0.71992
2-Ethylpyridine	0.03325	1.94	0.75518
2-Isopropylpyridine	0.0334	1.76	0.78639
2-Tert-butylpyridine	0.03014	1.42	0.12289
3-Ethylpyridine	0.03297	2.01	0.69829
3-Methylpyridine	0.03279	2	0.66175
N,N,4-Trimethylaniline	0.02075	0.69	1.78813
4-Vinylpyridine	0.03261	1.95	0.6249
2,2,2-Trifluoroethylamine	0.02667	0.71	0.58316
Isoquinoline	0.03246	1.94	0.59479
N,4-Dimethylaniline	0.02244	0.43	1.44489
4-Phenylpyridine	0.0328	1.96	0.66342
Acridine	0.034	1.95	0.90871
Pyridine	0.0319	1.86	0.48005
p-Toluidine	0.02419	0.56	1.08823
N,N-Dimethylaniline	0.01971	0.39	1.99988
1,2,3,4-Tetrahydroquinoline	0.02326	0.7	1.27732
2-Vinylpyridine	0.03247	1.65	0.59712
N-Methylaniline	0.02151	0.26	1.63421
Quinoline	0.03324	1.89	0.75431
2-Phenylpyridine	0.0305	1.43	0.19647
Aniline	0.02327	0.46	1.27475
1-Phenylpyrrolidine	0.01839	0.16	2.26885
4-Bromo-N,N-dimethylaniline	0.01909	0.17	2.12613
Dimethylaminoacetonitrile	0.02474	0.67	0.97552
2,6-Dimethylaniline	0.02418	0.47	1.09074
7,8-Benzoquinoline	0.02995	1.16	0.08462
4-Chloro-N-methylaniline	0.02032	0.05	1.87607
4-Chloropyridine	0.03067	1.54	0.23127
1-Acetyl-1H-imidazole	0.03218	1.86	0.53694
3-Fluoroaniline	0.02145	0.2	1.64691
2-Aminopyrimidine	0.03587	1.85	1.28865
3-Chloroaniline	0.02141	0.13	1.65447
4-Acetylpyridine	0.02999	1.41	0.09254
Phthalazine	0.03208	1.97	0.51767
3-Dimethylamino-5,5-dimethylcyclohexenone	0.03742	2.92	1.60471

Tripropargylamine	0.02452	0.83	1.02066
Methylnicotinate	0.02863	1.44	0.18416
2-Methoxypyridine	0.02883	0.99	0.14374
3-Fluoropyridine	0.02952	1.35	0.00408
3-Bromopyridine	0.02986	1.31	0.06618
3-Chloropyridine	0.02948	1.31	0.01146
1,3-Thiazole	0.03003	1.37	0.10107
1-Methyl-1H-pyrazole	0.03378	1.84	0.86422
4-Methoxypyridine 1-oxide	0.03972	3.7	2.07233
Pyridazine	0.03028	1.65	0.15039
4-Cyanopyridine	0.02758	0.92	0.39824
3-Cyanopyridine	0.02619	0.82	0.68063
Isoxazole	0.02664	0.81	0.59003
4-Methylpyridine 1-oxide	0.03781	3.12	1.68418
Pyrimidine	0.02813	1.07	0.28616
3-Methylpyridine 1-oxide	0.03711	2.92	1.54077
2-Bromopyridine	0.02726	1.03	0.46421
Phenazine	0.03088	1.22	0.27351
4-Phenylpyridine 1-oxide	0.03696	2.85	1.5101
1,3-Oxazole	0.02965	1.3	0.02312
Diphenylamine	0.01805	-1.05	2.33699
Pyridine 1-oxide	0.03663	2.72	1.44281
2-Chloropyridine	0.02693	1.05	0.53005
3,5-Dichloropyridine	0.02717	0.85	0.48248
Pyrazine	0.02782	0.92	0.34877
4-Chloropyridine 1-oxide	0.03646	2.44	1.40904
1,1,3,3-Tetramethylurea	0.03483	2.44	1.07599
N,N-Dimethylacetamide	0.0352	2.44	1.15176
2-Cyanopyridine	0.02432	0.48	1.06223
2-Fluoropyridine	0.02675	0.95	0.56804
Tetramethylene sulfone	0.0288	1.17	0.14975
Acetamide	0.03282	2.06	0.66691
N-Methylpropanamide	0.03488	2.24	1.08648
1-Methyl-2-pyrrolidone	0.0387	2.38	1.86387
3,5-Dichloropyridine 1-oxide	0.03359	1.56	0.82484
Hexamethylphosphoramide	0.04001	3.6	2.13085
N-Methylformamide	0.03008	1.96	0.11004
N,N-Dimethylformamide	0.03104	2.1	0.30567
Formamide	0.0284	1.75	0.23243
4-Nitropyridine 1-oxide	0.03112	1.05	0.32181
Dimethyl sulfoxide	0.03358	2.54	0.82325
N,N,2-Trimethylpropanamide	0.03467	2.26	1.04407
Ethanol	0.02613	1.02	0.69325
Propan-2-ol	0.02685	1.06	0.54784
N,N,2,2-Tetramethylpropanamide	0.03118	2.1	0.33511

Methanol	0.02518	0.82	0.8859
Propan-1-ol	0.02622	1	0.67548
N,N-Dimethylthioacetamide	0.02379	1.22	1.16941
Methyl phenyl sulfoxide	0.03257	2.24	0.61677
Diethyl ether	0.02651	1.01	0.61642
2-Bromethanol	0.02421	0.54	1.08399
2-Chloroethanol	0.02345	0.5	1.23957
3-Methyl-5,5-dimethylcyclohexenone	0.03218	1.74	0.53671
Diphenyl sulfoxide	0.03491	2.04	1.09302
4-Methoxyacetophenone	0.03128	1.33	0.354
Acetone	0.02864	1.18	0.18346
Butan-2-one	0.02896	1.22	0.11858
Methyl 4-nitrophenyl sulfoxide	0.03001	1.58	0.09663
3-Methylbutan-2-one	0.02906	1.2	0.09711
3-Chloro-5,5-dimethylcyclohexenone	0.03041	1.21	0.17791
3,3-Dimethylbutan-2-one	0.02936	1.17	0.03711
Ethyl acetate	0.02767	1.07	0.37929
Acetophenone	0.02968	1.11	0.02819
Dimethyl sulfone	0.02849	1.1	0.21285
4-Fluoroacetophenone	0.02935	1	0.03919
4-Isopropylacetophenone	0.03044	1.21	0.18296
4-Ethylacetophenone	0.03039	1.25	0.17314
4-Methylacetophenone	0.03042	1.24	0.17857
3,5-Dimethylheptan-4-one	0.0299	1.07	0.07441
Nonan-5-one	0.02821	1.21	0.2705
Methyl acetate	0.02843	1	0.22493
4-Tert-butylacetophenone	0.03052	1.25	0.19883
4-Trifluoromethylacetophenone	0.02783	0.78	0.34731
4-Nitroacetophenone	0.02671	0.57	0.57614
Ethyl benzoate	0.02868	0.94	0.17505
t-Butyl methyl sulfide	0.01993	0.25	1.95561
2,2,4,4-Tetramethylpentan-3-one	0.02711	0.96	0.49421
Dimethyl sulfide	0.01727	0.12	2.49558
Methyl benzoate	0.0289	0.89	0.12961
Acetonitrile	0.02635	0.91	0.64907
Benzonitrile	0.02697	0.8	0.52305
Chloroacetonitrile	0.02317	0.39	1.2956
Diazabicyclooctane	0.03259	2.33	0.62188
Nicotine	0.02896	1.11	0.11842
N-methylpyrrolidine	0.03071	2.19	0.23807
N-methyltetrahydroisoquinoline	0.02885	1.8	0.1398
N,N-diisopropylethylamine	0.03076	1.05	0.24838
N,N-dimethylaminoacetonitrile	0.023	0.76	1.32978
N,N-dimethylaminopropionitrile	0.02863	1.15	0.18556
N,N-dimethylethylamine	0.03023	2.17	0.14114

N,N-dimethylpropargylamine	0.02839	1.6	0.23297
Quinuclidine	0.02944	2.71	0.01961
Triallylamine	0.02989	1.34	0.07259
Triethyamine	0.03152	1.98	0.40234

Table 3.8.5. All hydrogen bond acceptors used in computations to form hydrogen bond complexes with methanol used in this research along with their $pK_{BHX\ [1]}$, ΔE(H), and Z values. The Z value is equated from a Grubb's test for an outlier. A description of the Grubb's test is given in the text above. No Z values are greater than the critical value of 3.75 at the 0.05 significance level so no outliers are detected.

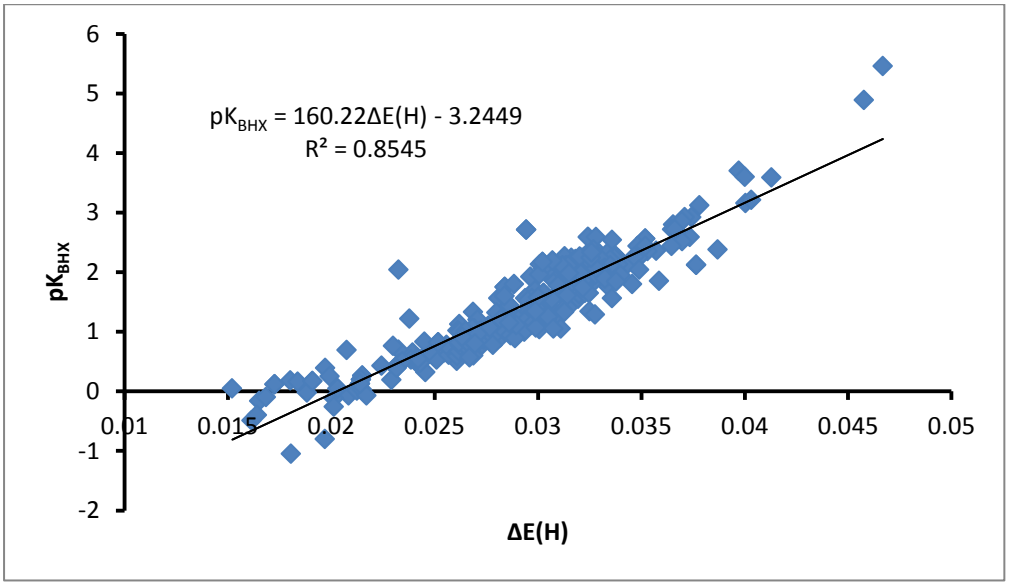


Figure 3.8.10. The relationship between experimental pK_{BHX} and computed ΔE(H) values for the 332 hydrogen bond acceptors listed in table 3.8.5. The hydrogen bond donor used in the computations was methanol.

3.9 Relationship Between pK_{BHX} Prediction And Brönsted Basicity Using Methylamine As The Hydrogen Bond Donor

It has been stated that the pK_{BHX} scale, although set up using 4-fluorophenol and methanol as reference hydrogen bond donors, is applicable to amine hydrogen bond donors. It has also been shown in section 3.2 that there is a strong correlation between pK_{BHX} and ΔE(H) when methylamine is used as the hydrogen bond donor. The following

section will aim to find out whether the relationship between pK_{BHX} and ΔE(H) is pK_{BH}^{+} dependent when methylamine is used as the hydrogen bond donor.

The pK_{BHX} database has been stated to be applicable to both alcohol and amine hydrogen bond donors. However amines are much stronger bases than alcohols. Methylamine is a much stronger proton donor than methanol. The pK_{BH}^{+} value of methanol is -2.05 whereas methylamine has a value of 10.65. Methanol has a pK_a value of around 15 whereas methylamine has a value of around 10.5. Therefore methylamine is a much stronger base than methanol. It is known from the pK_a slide rule [119] that the strongest hydrogen bond complexes are formed by hydrogen bond donors and acceptors with similar pK_a values. Therefore methanol will form stronger hydrogen bond complexes with weaker acids whereas methylamine will form the strongest hydrogen bond complexes with stronger acids. Given what the pK_a slide rule states, although the pK_{BHX} scale may be applicable to both methanol and methylamine hydrogen bond donors, it is unlikely that the pK_{BHX} values would be similar.

The ΔE(H) values are computed from a hydrogen bond complex in the donor : acceptor ratio of 1 : 1. Solvent effects are not modelled. The influence of hydrogen bond networks or multiple hydrogen bonds are also not modelled. Both solvent and hydrogen bond networks will be present in the experimental calculation of pK_{BHX} value. However, it can be hypothesised that the strongest hydrogen bond complexes will be subjected to the least solvent and hydrogen bond network effects and will be therefore most like the model. Therefore it can be hypothesised that the model will return a relationship between pK_{BHX} and ΔE(H) that has the largest r^2 values where the pK_{BH}^{+} values are similar. The results in the previous section showed that the relationship between pK_{BHX} and ΔE(H) is strongest in the pK_{BH}^{+} value range of -1 to 6. Methanol has a pK_{BH}^{+} value of -2.05 which is outside the range of -1 to 6 and therefore means the above mentioned hypothesis is not strictly true. However, the effects on changing the hydrogen bond donor to the more basic methylamine are interesting and could lead to an explanation as to why pK_{BHX} and ΔE(H) do not correlate well for very high and very low pK_{BH}^{+} values.

A set of 215 bases has been selected. The 215 bases are listed along with their pK_{BH}^{+} and pK_{BHX} values are listed in table 3.9.1. The bases in table 3.9.1 are ranked in order of descending pK_{BH}^{+} values. The 215 bases chosen are based on those chosen to illustrate how there is no relationship between pK_{BH}^{+} and pK_{BHX} in [1]. The bases and their pK_{BHX} and pK_{BH}^{+} values are listed in the supplementary information of [1]. Using methanol as the hydrogen bond donor, hydrogen bond complex computations have been performed in the usual way as described in chapter 2. Once again ΔE(H) values have been calculated. The process has been repeated substituting the hydrogen bond donor methanol for methylamine.

HYDROGEN BOND ACCEPTOR	pK_{BH}^{+}	pK_{BHX}
1,1,2,3,3-Pentamethylguanidine	13.8	3.16
1,1,3,3-Tetramethylguanidine	13.6	3.21
N,N-Dimethyl-N'-propylimidoformamide	11.46	2.59
Pyrrolidine	11.306	2.59
N'-Isobutyl-N,N-dimethylimidoformamide	11.3	2.52
Azetidine	11.29	2.59
Dibutylamine	11.25	2.11
Diisopropylamine	11.2	2
1,2,2,6,6-Pentamethylpiperidine	11.19	1.23
Quinuclidine	11.1516	2.71
Piperidine	11.1236	2.38
Azepane	11.1	2.24
2,2,6,6-Tetramethylpiperidine	11.0725	1.88
N-Methylcyclohexylamine	11.04	2.24
Diethylamine	11.0151	2.25
N,N-Dimethyl-N'-(4-methylbenzyl)imidoformamide	10.91	2.36
N-methylbutylamine	10.9	2.24
Dimethylamine	10.7788	2.26
N,N-Dimethylcyclohexylamine	10.72	2.15
Triethylamine	10.7174	1.98
Tert-butylamine	10.6837	2.23
Ethylamine	10.6784	2.17
Isopropylamine	10.67	2.2
Tripropylamine	10.66	1.47
Methylamine	10.6532	2.2
N'-Benzyl-N,N-dimethylimidoformamide	10.65	2.35
n-Butylamine	10.6385	2.19

1,6-Diaminohexane	10.63	2.21
n-Octylamine	10.61	2.27
n-Hexadecylamine	10.61	2.26
c-Hexylamine-MeOH	10.58	2.29
n-Propylamine	10.5685	2.2
NN-Dimethylisopropylamine-MeOH	10.47	2.11
N-Methylpyrrolidine	10.46	2.19
N-Butylpyrrolidine	10.36	2.04
1,4-Diaminobutane	10.35	2.21
N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	10.32	2.12
1,3-Diaminopropane	10.17	2.31
N,N-Dimethylethylamine	10.16	2.17
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	10.15	2.1
N-Methylallylamine	10.11	2
N-Methylpiperidine	10.08	2.11
N,N'-dimethylethylenediamine	9.94	2.29
Tributylamine	9.93	1.55
3-Methoxypropylamine	9.92	2.22
N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	9.86	2
2-Phenylethanamine	9.83	2.16
Trimethylamine	9.7977	2.13
Ethylenediamine	9.626	2.25
2-Phenylpyrrolidine	9.6	1.93
4-N,N-Dimethylaminopyridine	9.58	2.8
N-Methylbenzylamine-MeOH	9.56	1.82
Allylamine	9.52	1.93
2-Methoxyethylamine	9.44	2.26
Piperazine	9.432	2.11
1,2,3,4-Tetrahydroisoquinoline	9.41	2.04
Benzylamine	9.34	1.84
Ammonia	9.244	1.74
Diallylamine	9.24	1.7
4-Aminopyridine	9.12	2.56
c-Propylamine	9.1	1.72
N,N-Dimethylbenzylamine	8.91	1.59
N,N,N',N'-Tetramethylethylenediamine	8.85	2.02
N-Methyl-2-phenylpyrrolidine	8.8	1.38
N,N-Dimethylallylamine	8.64	1.92
Diazabicyclooctane	8.52	2.33
Dibenzylamine	8.52	1.34
Morpholine	8.4918	1.78
N,N-Dimethyl-N'-(4-methylphenyl)imidoformamide	8.45	2.07
Triallylamine	8.28	1.34
N,N-Dimethyl-N'-(2-methylphenyl)imidoformamide	8.27	1.63
N,N-Dimethyl-N'-phenylimidoformamide	8.15	1.9

Propargylamine	8.15	1.56
3-Aminopropionitrile	7.8	1.33
Hexamethylenetetramine	7.4696	1.33
N,N-Dimethylpropargylamine	7.45	1.6
2,4,6-Trimethylpyridine	7.43	2.29
N-Methylmorpholine	7.41	1.56
N'-(4-Bromophenyl)-N,N-dimethylimidoformamide	7.4	1.65
1,1-Diphenylmethanimine	7.18	1.8
1-Methyl-1H-imidazole	7.12	2.72
N'-(4-Acetylphenyl)-N,N-dimethylimidoformamide	7.02	1.52
2-N,N-Dimethylaminopyridine	6.99	1.61
2,6-Dimethylpyridine	6.72	2.14
2-Aminopyridine	6.71	2.12
N'-(2-Bromophenyl)-N,N-dimethylimidoformamide	6.71	1.37
2-Methyl-N-(phenylmethylene)propan-2-amine	6.7	1.29
N,N-Diethylaniline	6.61	0.05
4-Methoxypyridine	6.58	2.13
3,4-Dimethylpyridine	6.47	2.24
N'-(4-Cyanophenyl)-N,N-dimethylimidoformamide	6.44	1.32
3,5-Dimethylpyridine	6.34	2.21
Thiazolidine	6.22	1.1
N,N,N',N'-Tetramethylbenzene-1,4-diamine	6.05	1.13
3-Aminopyridine	6.03	2.2
4-Methylpyridine	6.03	2.07
4-Ethylpyridine	6.02	2.07
N,N-Dimethyl-N'-(4-nitrophenyl)imidoformamide	6.02	1.2
4-Tert-butylpyridine	5.99	2.11
2-Methylpyridine	5.96	2.03
2-Ethylpyridine	5.89	1.94
2-Isopropylpyridine	5.83	1.76
2-Tert-butylpyridine	5.76	1.42
3-Ethylpyridine	5.73	2.01
3-Methylpyridine	5.66	2
N,N,4-Trimethylaniline	5.63	0.69
4-Vinylpyridine	5.62	1.95
2,2,2-Trifluoroethylamine	5.61	0.71
Isoquinoline	5.4	1.94
N,4-Dimethylaniline	5.36	0.43
4-Phenylpyridine	5.35	1.96
Acridine	5.24	1.95
Pyridine	5.2	1.86
p-Toluidine	5.08	0.56
N,N-Dimethylaniline	5.07	0.39
1,2,3,4-Tetrahydroquinoline	5.03	0.7
2-Vinylpyridine	4.92	1.65

N-Methylaniline	4.85	0.26
Quinoline	4.85	1.89
2-Phenylpyridine	4.72	1.43
Aniline	4.61	0.46
1-Phenylpyrrolidine	4.3	0.16
4-Bromo-N,N-dimethylaniline	4.23	0.17
Dimethylaminoacetonitrile	4.2	0.67
2,6-Dimethylaniline	3.95	0.47
7,8-Benzoquinoline	3.95	1.16
4-Chloro-N-methylaniline	3.9	0.05
4-Chloropyridine	3.83	1.54
1-Acetyl-1H-imidazole	3.6	1.86
3-Fluoroaniline	3.59	0.2
2-Aminopyrimidine	3.54	1.85
3-Chloroaniline	3.52	0.13
4-Acetylpyridine	3.51	1.41
Phthalazine	3.17	1.97
3-Dimethylamino-5,5-dimethylcyclohexenone	3.14	2.92
Tripropargylamine	3.09	0.83
Methylnicotinate	3.08	1.44
3-Fluoropyridine	3	1.35
3-Bromopyridine	2.84	1.31
3-Chloropyridine	2.8	1.31
1,3-Thiazole	2.518	1.37
1-Methyl-1H-pyrazole	2.06	1.84
4-Methoxypyridine 1-oxide	2.05	3.7
Pyridazine	2	1.65
4-Cyanopyridine	1.86	0.92
3-Cyanopyridine	1.34	0.82
Isoxazole	1.3	0.81
4-Methylpyridine 1-oxide	1.29	3.12
Pyrimidine	0.93	1.07
3-Methylpyridine 1-oxide	0.93	2.92
2-Bromopyridine	0.9	1.03
Phenazine	0.9	1.22
4-Phenylpyridine 1-oxide	0.83	2.85
1,3-Oxazole	0.8	1.3
Pyridine 1-oxide	0.79	2.72
2-Chloropyridine	0.72	1.05
3,5-Dichloropyridine	0.66	0.85
Pyrazine	0.37	0.92
4-Chloropyridine 1-oxide	0.33	2.44
1,1,3,3-Tetramethylurea	-0.14	2.44
N,N-Dimethylacetamide	-0.21	2.44
2-Cyanopyridine	-0.26	0.48

2-Fluoropyridine	-0.44	0.95
Tetramethylene sulfone	-0.65	1.17
Acetamide	-0.66	2.06
N-Methylpropanamide	-0.7	2.24
1-Methyl-2-pyrrolidone	-0.71	2.38
3,5-Dichloropyridine 1-oxide	-0.94	1.56
Hexamethylphosphoramide	-0.97	3.6
N-Methylformamide	-1.1	1.96
N,N-Dimethylformamide	-1.13	2.1
Formamide	-1.47	1.75
4-Nitropyridine 1-oxide	-1.51	1.05
Dimethyl sulfoxide	-1.54	2.54
N,N,2-Trimethylpropanamide	-1.61	2.26
Ethanol	-1.94	1.02
Propan-2-ol	-2.02	1.06
N,N,2,2-Tetramethylpropanamide	-2.03	2.1
Methanol	-2.05	0.82
Propan-1-ol	-2.12	1
N,N-Dimethylthioacetamide	-2.25	1.22
Methyl phenyl sulfoxide	-2.27	2.24
Diethyl ether	-2.39	1.01
2-Bromethanol	-2.41	0.54
2-Chloroethanol	-2.45	0.5
3-Methyl-5,5-dimethylcyclohexenone	-2.5	1.74
Diphenyl sulfoxide	-2.54	2.04
4-Methoxyacetophenone	-3.02	1.33
Acetone	-3.06	1.18
Butan-2-one	-3.06	1.22
Methyl 4-nitrophenyl sulfoxide	-3.11	1.58
3-Methylbutan-2-one	-3.29	1.2
3-Chloro-5,5-dimethylcyclohexenone	-3.36	1.21
3,3-Dimethylbutan-2-one	-3.37	1.17
Ethyl acetate	-3.45	1.07
Acetophenone	-3.46	1.11
Dimethyl sulfone	-3.5	1.1
4-Fluoroacetophenone	-3.53	1
4-Isopropylacetophenone	-3.69	1.21
4-Ethylacetophenone	-3.71	1.25
4-Methylacetophenone	-3.78	1.24
3,5-Dimethylheptan-4-one	-3.86	1.07
Nonan-5-one	-3.89	1.21
Methyl acetate	-3.9	1
4-Tert-butylacetophenone	-3.97	1.25
4-Trifluoromethylacetophenone	-4.13	0.78
4-Nitroacetophenone	-5.02	0.57

Ethyl benzoate	-6.16	0.94
t-Butyl methyl sulfide	-6.68	0.25
2,2,4,4-Tetramethylpentan-3-one	-6.8	0.96
Dimethyl sulfide	-6.95	0.12
Methyl benzoate	-7.05	0.89
Acetonitrile	-10.1	0.91
Benzonitrile	-10.4	0.8
Chloroacetonitrile	-12.8	0.39

Table 3.9.1. The hydrogen bond acceptors used in this section. The hydrogen bond acceptors are ranked in order of Brönsted basicity (pK_{BH^+}) [1] from strongest to weakest.

Once again correlations have been set up between ΔE(H) and pK_{BHX}. In this section the correlations are taken from 15 base subsets from within the data set of 215. The first subset includes the 15 bases with the highest pK_{BH^+} values. The first subset therefore includes the first 15 bases listed in table 3.9.1. The second subset is made up of the bases ranked between 2 and 16 in table 3.9.1. In other words there is a sliding window containing 15 bases. Each new window replaces the base with the largest pK_{BH^+} value in the previous window with the base with the next largest pK_{BH^+} value outside that of the previous window. For each subset r^2 values have been calculated. The r^2 values for each subset have been plotted against the mean pK_{BH^+} value of each subset. The results are displayed in figures 3.9.1 and 3.9.2.

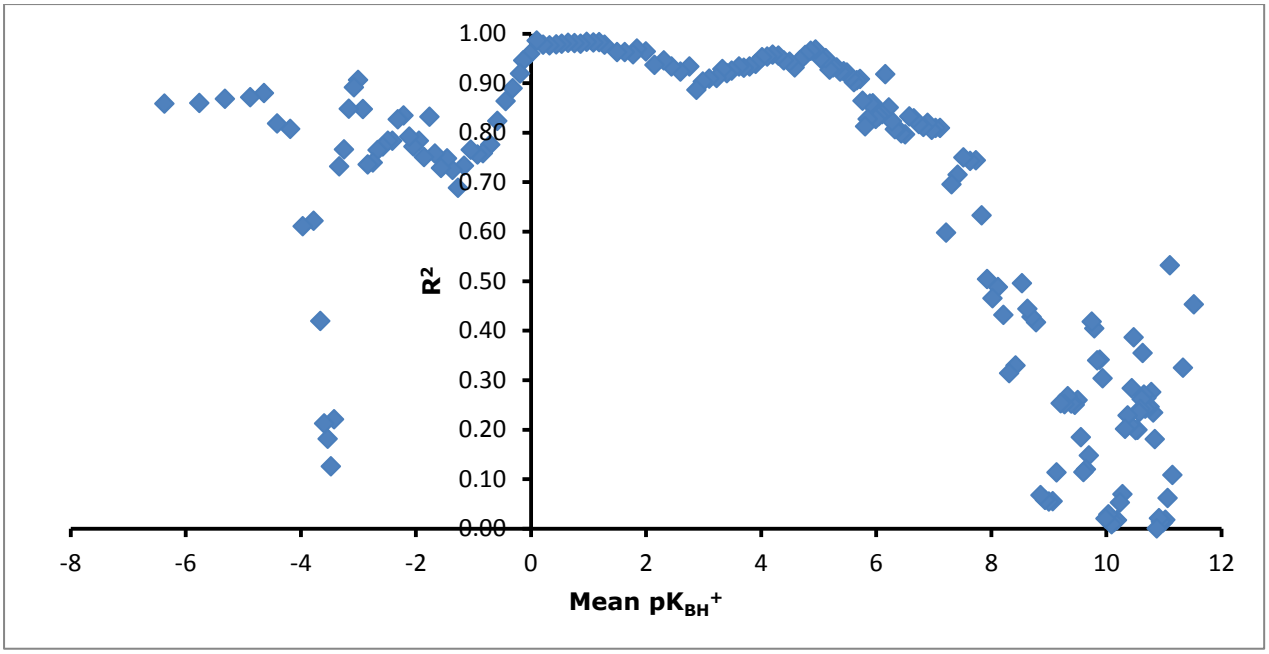


Figure 3.9.1. From the bases listed in table 3.9.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value [1] is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methanol is the hydrogen bond donor.

-

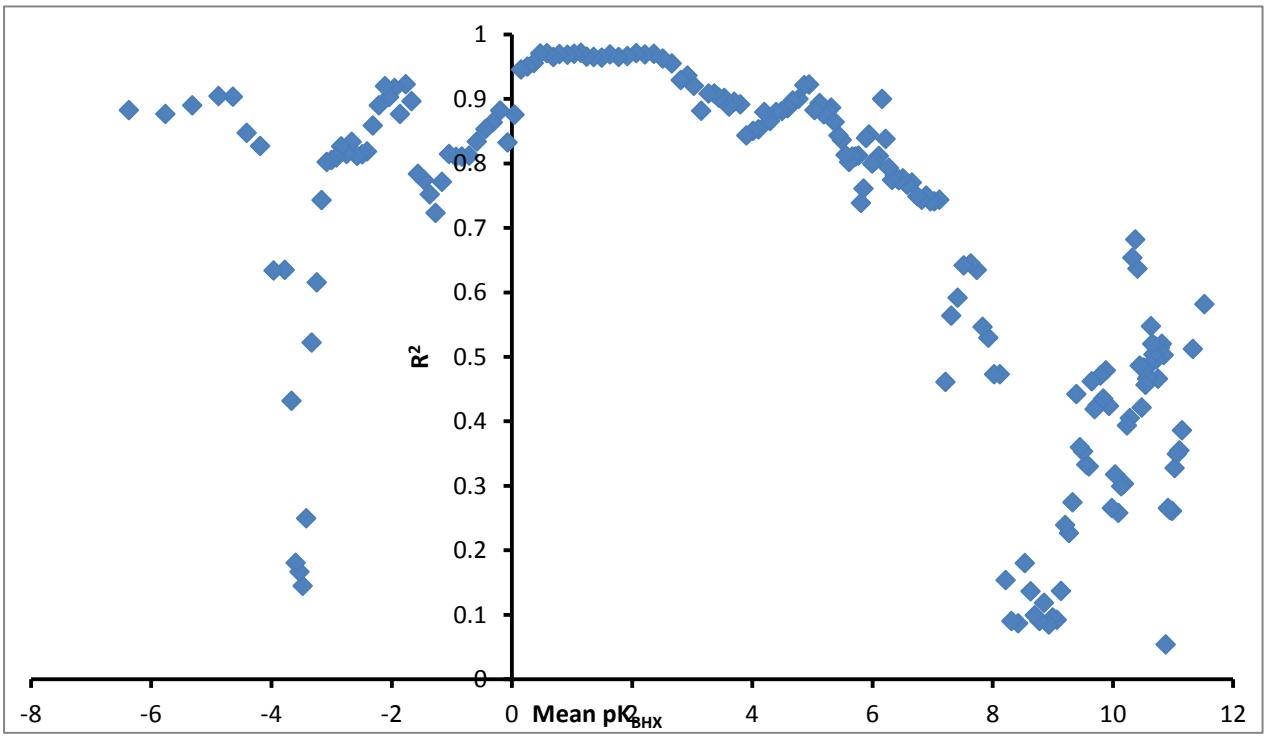


Figure 3.9.2. From the bases listed in table 3.9.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value [1] is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methylamine is the hydrogen bond donor.

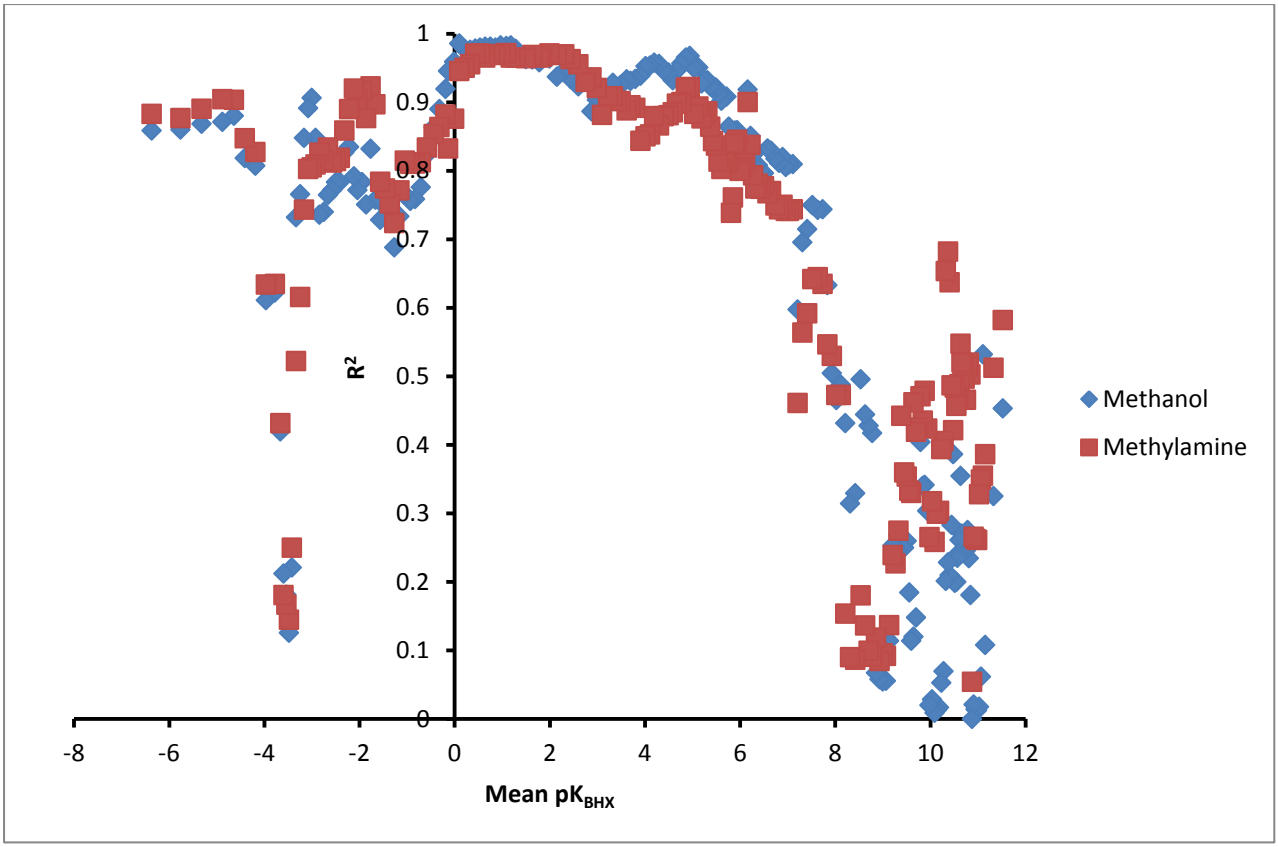


Figure 3.9.3. A direct comparison between methanol and methylamine hydrogen bond donors. From the bases listed in table 3.9.1, 15 base subsets have been chosen according to the criteria in the text. The mean pK_{BH}^{+} value [1] is plotted for each 15 base subset. The r^2 value is taken from correlations between the pK_{BHX} value [1] of each base and the computed ΔE(H) value for each base in each 15 base subset. The blue diamonds represent the data series where methanol has been used as the hydrogen bond donor whereas the red dots represent the data series where methylamine was used as the hydrogen bond donor.

Figure 3.9.1 confirms the findings in the previous section that the relationship between pK_{BHX} and ΔE(H) only returns r^2 values of greater than 0.9 in the pK_{BH}^{+} value range of -1 to 6. This finding was stated in the previous section but has been repeated here using a slightly reduced data set in order to allow a direct comparison between methanol and methylamine hydrogen bond donors to be made. Figure 3.9.2 shows a similar finding when methylamine is used as the hydrogen bond donor. However, when methylamine is used as the hydrogen bond donor the pK_{BH}^{+} range in which the relationship between pK_{BHX} and ΔE(H) returns an r^2 value of greater than 0.9 appears to

be slightly reduced. The pK_{BH}^+ range in which the relationship between pK_{BHX} and ΔE(H) returns an r^2 value of greater than 0.9 appears to be between 0 and 3. This is shown clearly in figure 3.9.3 as both the data for methanol and methylamine is plotted on the graph. Therefore the best hydrogen bond donor to use to predict pK_{BHX} values from is methanol because methanol can be used to reliably predict a broader range of basicities.

Bases that have high pK_{BH}^+ values have pK_{BHX} values that are poorly predicted when methylamine is used as the hydrogen bond donor as figure 3.9.2 shows. The pK_a slide rule suggests that the strongest complexes should be formed between methylamine and the bases with similarly high pK_{BH}^+ values. It is therefore possible that the computation of ΔE(H) values for complexes where methylamine is donating a hydrogen bond to a strong base overestimates the pK_{BHX} value. The pK_{BHX} value is possibly overestimated because methanol or 4-fluorophenol would form a much weaker complex in the experiment than the computed complex involving methylamine.

It has been hypothesised inconclusively in previous sections that the relationship between pK_{BHX} and ΔE(H) breaks down for hydrogen acceptors with high positive pK_{BH}^+ values due to possible protonations taking place in the experiment. The possible protonations give rise to many different types of hydrogen bond complexes forming in the experiment as discussed in the previous section. The pK_a slide rule [119] teaches that the strongest hydrogen bond complexes formed by acids and bases with similar pK_a values have the most covalent character. The weakest hydrogen bond complexes have less covalent character and sre mainly electrostatic in nature. Stronger hydrogen bond complexes are a mixture of electrostatic and covalent interactions. Any protonation taking place in the experiment would be at the site of the strongest proton acceptor, the base with the largest positive pK_{BH}^+ value. However, as methylamine has a much larger pK_{BH}^+ value than methanol and the majority of bases in table 3.9.1, it is likely that the site of protonation would be the nitrogen atom of methylamine. Given the large positive pK_{BH}^+ value of methylamine it would be interesting to repeat the procedure in this section using protonated methylamine as the hydrogen bond donor. Modelling hydrogen

bonds where protonated methylamine is donating a hydrogen bond to a series of hydrogen would serve as a further attempt to solve the presently unknown problem of why the relationship between pK_{BHX} and ΔE(H) breaks down for bases with large positive pK_{BH}^{+} values. However this would have to be done as a future experiment.

It can be concluded from the result shown in this section that the relationship between pK_{BHX} and ΔE(H) appears to be, to a certain extent independent of the strength of the hydrogen bond donor. However the relationship between pK_{BHX} and ΔE(H) is only valid in the pK_{BH}^{+} range of 0 to 3 when methylamine is used as the hydrogen bond donor compared to a range of -1 to 6 when methanol is the hydrogen bond donor. Therefore the pK_{BHX} values of stronger proton acceptors are poorly predicted when methylamine is used as the hydrogen bond donor. This is possibly due to a stronger hydrogen bond in the model than is actually present in the experiment where methanol or 4-fluorophenol are the hydrogen bond donors. However it has previously been proven in section 3.2 that pK_{BHX} and hydrogen bond enthalpy are not strongly correlated. Therefore the reasons why the relationship between pK_{BHX} and ΔE(H) breaks down for hydrogen bond acceptors with large positive pK_{BH}^{+} values is unfortunately still unknown.

3.10 Relationship Between Chemical Hardness And pK_{BHX} Prediction.

The purpose of the research in this section is to determine whether a link between hardness and hydrogen bond basicity exists. The relationship between pK_{BHX} and ΔE(H) does not hold for hydrogen bond acceptors with large positive pK_{BH}^{+} values. Despite the efforts reported in previous sections the reasons why are still unknown. However it does appear that methylamine does not predict the pK_{BHX} values for hydrogen bond donors in the pK_{BH}^{+} of 3 to 6 as well as methanol when used as the hydrogen bond donor. As the hard soft acid base principle states the strongest hydrogen bond

complexes should be formed from hydrogen bond donors and acceptors with similar hardness values.

Results in sections 3.8 and 3.9 have shown that the relationship between pK_{BHX} and ΔE(H) is not present for bases with large positive pK_{BH}^{+} values. It has been hypothesised that possible protonation effects could take place during the experimental calculation of pK_{BHX} values, thus affecting the computation of pK_{BHX} values. This was, however, not proven in any of the computations described in previous sections. A further hypothesis based on the pK_a slide rule also failed to explain why the relationship between pK_{BHX} and ΔE(H) breaks down for hydrogen bond acceptors with large positive pK_{BH}^{+} values. However, a knowledge of the pK_a slide rule leads to the finding that methanol returns ΔE(H) values that correlate with pK_{BHX} values better than methylamine in the pK_{BH}^{+} region of 3 to 6. The reason for this quite possibly lies in the theory that the strongest hydrogen bonds are formed by hydrogen bond donors and hydrogen bond acceptors with similar pK_a values. It is also known from the HSAB principle that the strongest acid–base complexes are formed when the hardness values of the acid and bases are similar in value. The theory behind the hard soft acid base principle is discussed in detail in section 1.6.

The isolated bases listed in table 3.10.1 have undergone a geometry optimisation following the procedure described in chapter 2. Once the optimised geometry has been obtained, energy point calculations were performed on the frozen optimised geometry for the singly positive and negatively charged base. This allows hardness values to be calculated from equation 1.6.9. Equation 1.6.9 shows how hardness values are computed from the energy of the neutral and the singly positive and negatively charged bases listed in table 3.10.1.

HYDROGEN BOND ACCEPTOR	pK_{BHX} [1]	$pK_{BH}^{\pm}$ [1]	Hardness (Ev)
Acetonitrile	0.91	-10.1	6.358587
Ammonia	1.74	9.244	5.836667
Methanol	0.82	-2.05	5.783936

2-Chloroethanol	0.5	-2.45	5.698482
Formamide	1.75	-1.47	5.635184
Ethanol	1.02	-1.94	5.581954
Methyl acetate	1	-3.9	5.497632
Propan-1-ol	1	-2.12	5.477271
2,2,2-Trifluoroethylamine	0.71	5.61	5.472866
Dimethyl sulfone	1.1	-3.5	5.464994
Propan-2-ol	1.06	-2.02	5.419857
Ethyl acetate	1.07	-3.45	5.40689
Isoxazole	0.81	1.3	5.370516
2-Bromethanol	0.54	-2.41	5.306732
1,3-Thiazole	1.37	2.518	5.223932
4-Chloropyridine	1.54	3.83	5.206105
1,3-Oxazole	1.3	0.8	5.199484
2-Fluoropyridine	0.95	-0.44	5.192351
3-Fluoropyridine	1.35	3	5.171497
N-Methylformamide	1.96	-1.1	5.157663
Methylamine	2.2	10.6532	5.149756
Acetamide	2.06	-0.66	5.126162
Diethyl ether	1.01	-2.39	5.097369
3-Aminopropionitrile	1.33	7.8	5.096907
Acetone	1.18	-3.06	5.07812
Pyrimidine	1.07	0.93	5.068691
Ethylamine	2.17	10.6784	4.993795
n-Propylamine	2.2	10.5685	4.993689
2-Chloropyridine	1.05	0.72	4.98566
Propargylamine	1.56	8.15	4.983704
c-Propylamine	1.72	9.1	4.979502
3-Chloropyridine	1.31	2.8	4.971814
4-Methylpyridine	2.07	6.03	4.950102
n-Butylamine	2.19	10.6385	4.945078
N-Methylpropanamide	2.24	-0.7	4.944831
Isopropylamine	2.2	10.67	4.943933
Allylamine	1.93	9.52	4.942447
Butan-2-one	1.22	-3.06	4.93466
4-Cyanopyridine	0.92	1.86	4.918792
4-Ethylpyridine	2.07	6.02	4.909632
Pyrazine	0.92	0.37	4.889365
3-Cyanopyridine	0.82	1.34	4.888314
2-Cyanopyridine	0.48	-0.26	4.873452
3-Methylpyridine	2	5.66	4.872583
3-Methylbutan-2-one	1.2	-3.29	4.871363
3-Bromopyridine	1.31	2.84	4.868531
2-Methylpyridine	2.03	5.96	4.867638
4-Tert-butylpyridine	2.11	5.99	4.866569

2-Bromopyridine	1.03	0.9	4.861887
Tert-butylamine	2.23	10.6837	4.848179
2-Methoxyethylamine	2.26	9.44	4.847582
N,N-Dimethylformamide	2.1	-1.13	4.846237
Benzonitrile	0.8	-10.4	4.842618
c-Hexylamine-MeOH	2.29	10.58	4.83404
1-Methyl-1H-pyrazole	1.84	2.06	4.819747
Ethylenediamine	2.25	9.626	4.810824
2-Ethylpyridine	1.94	5.89	4.806484
3,3-Dimethylbutan-2-one	1.17	-3.37	4.795337
3-Ethylpyridine	2.01	5.73	4.790116
Methyl benzoate	0.89	-7.05	4.779295
2-Isopropylpyridine	1.76	5.83	4.773462
Azetidine	2.59	11.29	4.765762
3,4-Dimethylpyridine	2.24	6.47	4.760774
3,5-Dichloropyridine	0.85	0.66	4.755196
Ethyl benzoate	0.94	-6.16	4.752356
Dimethylamine	2.26	10.7788	4.749358
4-Methoxypyridine	2.13	6.58	4.74519
Dimethylaminoacetonitrile	0.67	4.2	4.744676
n-Octylamine	2.27	10.61	4.744387
Methylnicotinate	1.44	3.08	4.738418
Nonan-5-one	1.21	-3.89	4.736038
1-Acetyl-1H-imidazole	1.86	3.6	4.727055
Dimethyl sulfide	0.12	-6.95	4.725206
2-Aminopyrimidine	1.85	3.54	4.712055
Dimethyl sulfoxide	2.54	-1.54	4.708379
Pyridazine	1.65	2	4.70606
3,5-Dimethylpyridine	2.21	6.34	4.697237
3-Methoxypropylamine	2.22	9.92	4.696778
N,N-Dimethylacetamide	2.44	-0.21	4.680588
2,6-Dimethylpyridine	2.44	-0.21	4.668606
Pyrrolidine	2.59	11.306	4.647733
1,3-Diaminopropane	2.31	10.17	4.645784
Morpholine	1.78	8.4918	4.634608
N-Methylallylamine	2	10.11	4.629292
Diethylamine	2.25	11.0151	4.610463
N-methylbutylamine	2.24	10.9	4.603748
2,4,6-Trimethylpyridine	2.29	7.43	4.598358
Tripropargylamine	0.83	3.09	4.597007
N,N,2-Trimethylpropanamide	2.26	-1.61	4.595227
3,5-Dimethylheptan-4-one	1.07	-3.86	4.582552
Acetophenone	1.11	-3.46	4.577077
1-Methyl-1H-imidazole	2.72	7.12	4.571684
2,2,4,4-Tetramethylpentan-3-one	0.96	-6.8	4.569544

1,4-Diaminobutane	2.21	10.35	4.557651
4-Fluoroacetophenone	1	-3.53	4.548555
N,N,2,2-Tetramethylpropanamide	2.1	-2.03	4.544797
Benzylamine	1.84	9.34	4.512156
N,N-Dimethylpropargylamine	1.6	7.45	4.510861
3-Methyl-5,5-dimethylcyclohexenone	1.74	-2.5	4.505536
3-Chloro-5,5-dimethylcyclohexenone	1.21	-3.36	4.503423
N-Methylcyclohexylamine	2.24	11.04	4.500648
Diallylamine	1.7	9.24	4.493543
4-Vinylpyridine	1.95	5.62	4.492293
Thiazolidine	1.1	6.22	4.491865
2-Phenylethanamine	2.16	9.83	4.484085
Trimethylamine	2.13	9.7977	4.482651
4-Aminopyridine	2.56	9.12	4.482591
N-Methylmorpholine	1.56	7.41	4.476799
Piperidine	2.38	11.1236	4.451421
4-Methylacetophenone	1.24	-3.78	4.424426
2-Methyl-N-(phenylmethylene)propan-2-amine	1.29	6.7	4.42341
2-Vinylpyridine	1.65	4.92	4.42286
Methyl phenyl sulfoxide	2.24	-2.27	4.421387
4-Acetylpyridine	1.41	3.51	4.418016
N-Methylpyrrolidine	2.19	10.46	4.415055
1,6-Diaminohexane	2.21	10.63	4.413736
t-Butyl methyl sulfide	0.25	-6.68	4.413709
1,1,3,3-Tetramethylurea	2.44	-0.14	4.406255
N,N'-dimethylethylenediamine	2.29	9.94	4.371661
4-Isopropylacetophenone	1.21	-3.69	4.36765
N-Methylpiperidine	2.11	10.08	4.362638
Diisopropylamine	2	11.2	4.359083
4-Ethylacetophenone	1.25	-3.71	4.356509
2-Aminopyridine	2.12	6.71	4.356167
Dibutylamine	2.11	11.25	4.350972
Azepane	2.24	11.1	4.350549
4-Tert-butylacetophenone	1.25	-3.97	4.349934
Pyridine 1-oxide	2.72	0.79	4.342307
3-Aminopyridine	2.2	6.03	4.330382
Hexamethylenetetramine	1.33	7.4696	4.324662
4-Phenylpyridine	1.96	5.35	4.322886
N,N-Dimethylallylamine	1.92	8.64	4.318964
1,1,3,3-Tetramethylguanidine	3.21	13.6	4.30878
3-Methylpyridine 1-oxide	2.92	0.93	4.284378
N,N-Dimethylbenzylamine	1.59	8.91	4.28251
4-Methoxyacetophenone	1.33	-3.02	4.278768
N,N-Dimethylcyclohexylamine	2.15	10.72	4.275389
3-Fluoroaniline	0.2	3.59	4.26106

NN-Dimethylisopropylamine	2.11	10.47	4.250485
2,2,6,6-Tetramethylpiperidine	1.88	11.0725	4.236453
Dibenzylamine	1.34	8.52	4.229633
Phthalazine	1.97	3.17	4.224847
Triallylamine	1.34	8.28	4.224269
Quinoline	1.89	4.85	4.222822
3-Chloroaniline	0.13	3.52	4.221987
Diphenyl sulfoxide	2.04	-2.54	4.217616
Isoquinoline	1.94	5.4	4.199105
3,5-Dichloropyridine 1-oxide	1.56	-0.94	4.197612
Aniline	0.46	4.61	4.178811
N'-Isobutyl-N,N-dimethylimidoformamide	2.52	11.3	4.177585
Diazabicyclooctane	2.33	8.52	4.173403
N,N-Dimethyl-N'-propylimidoformamide	2.59	11.46	4.173218
4-Chloropyridine 1-oxide	2.44	0.33	4.155738
2-Phenylpyridine	1.43	4.72	4.154358
Piperazine	2.11	9.432	4.151018
1,1-Diphenylmethanimine	1.8	7.18	4.147196
4-Methylpyridine 1-oxide	3.12	1.29	4.138369
N,N,N',N'-Tetramethylethylenediamine	2.02	8.85	4.133586
N-Butylpyrrolidine	2.04	10.36	4.129758
4-Nitroacetophenone	0.57	-5.02	4.093495
Hexamethylphosphoramide	3.6	-0.97	4.091293
N-Methyl-2-phenylpyrrolidine	1.38	8.8	4.083132
4-N,N-Dimethylaminopyridine	2.8	9.58	4.075243
N,N-Dimethylthioacetamide	1.22	-2.25	4.067354
1,1,2,3,3-Pentamethylguanidine	3.16	13.8	4.056973
Tripropylamine	1.47	10.66	4.049973
N'-(3,5-Dichlorobenzyl)-N,N-dimethylimidoformamide	2	9.86	4.048016
p-Toluidine	0.56	5.08	4.030496
N'-(3-Chlorobenzyl)-N,N-dimethylimidoformamide	2.1	10.15	4.0118
N'-(4-Chlorobenzyl)-N,N-dimethylimidoformamide	2.12	10.32	4.001279
1,2,2,6,6-Pentamethylpiperidine	1.23	11.19	3.984855
N-Methylaniline	0.26	4.85	3.979263
2,6-Dimethylaniline	0.47	3.95	3.975037
N'-(2-Bromophenyl)-N,N-dimethylimidoformamide	1.37	6.71	3.961597
2-N,N-Dimethylaminopyridine	1.61	6.99	3.958203
N'-(4-Cyanophenyl)-N,N-dimethylimidoformamide	1.32	6.44	3.926699
3-Dimethylamino-5,5-dimethylcyclohexenone	2.92	3.14	3.915717
7,8-Benzoquinoline	1.16	3.95	3.884363
Methyl 4-nitrophenyl sulfoxide	1.58	-3.11	3.877986
N,N-Dimethyl-N'-phenylimidoformamide	1.9	8.15	3.864911
1,2,3,4-Tetrahydroquinoline	0.7	5.03	3.863794
N,4-Dimethylaniline	0.43	5.36	3.862258
N,N-Dimethylaniline	0.39	5.07	3.831168

N,N-Dimethyl-N'-(2-methylphenyl)imidoformamide	1.63	8.27	3.822072
N'-(4-Bromophenyl)-N,N-dimethylimidoformamide	1.65	7.4	3.813336
4-Nitropyridine 1-oxide	1.05	-1.51	3.786389
4-Bromo-N,N-dimethylaniline	0.17	4.23	3.779347
N,N-Dimethyl-N'-(4-methylphenyl)imidoformamide	2.07	8.45	3.762499
N,N,4-Trimethylaniline	0.69	5.63	3.731618
N'-(4-Acetylphenyl)-N,N-dimethylimidoformamide	1.52	7.02	3.720334
N,N-Diethylaniline	0.05	6.61	3.70255
N,N-Dimethyl-N'-(4-nitrophenyl)imidoformamide	1.2	6.02	3.699603
1-Phenylpyrrolidine	0.16	4.3	3.676997
4-Phenylpyridine 1-oxide	2.85	0.83	3.666292
Phenazine	1.22	0.9	3.412773
Acridine	1.95	5.24	3.41205
N,N,N',N'-Tetramethylbenzene-1,4-diamine	1.13	6.05	3.404818

Table 3.10.1. The hydrogen bond acceptors listed in this table have been from the hardest to the softest. The hardness values stated here have been computed as described in the above text.

Figure 3.10.1 shows how there is no relationship between hardness and pK_{BHX}. Therefore pK_{BHX} values cannot be predicted by hardness values alone. The r^2 value of 0.0065 indicates a random scattering of data points on the scatter plot of hardness and pK_{BHX}. The finding that hardness and pK_{BHX} are not related was to be expected. However it is expected that the strength of the complex is related to the relative hardnes of the hydrogen bond donors and acceptors. It is shown in table 3.10.1 that methanol has a hardness value of 5.784 eV. Therefore methanol is ranked as the third hardest molecule in the group listed in table 3.10.1. Based on the HSAB principle when methanol serves as the hydrogen bond donor, the strongest hydrogen bond complexes should be formed with hydrogen bond acceptors with hardness values close to 5.784 eV. It could therefore be hypothesised that pK_{BHX} will correlate most strongly with $\Delta E(H)$ where the hydrogen bond donors have hardness values similar to 5.784.

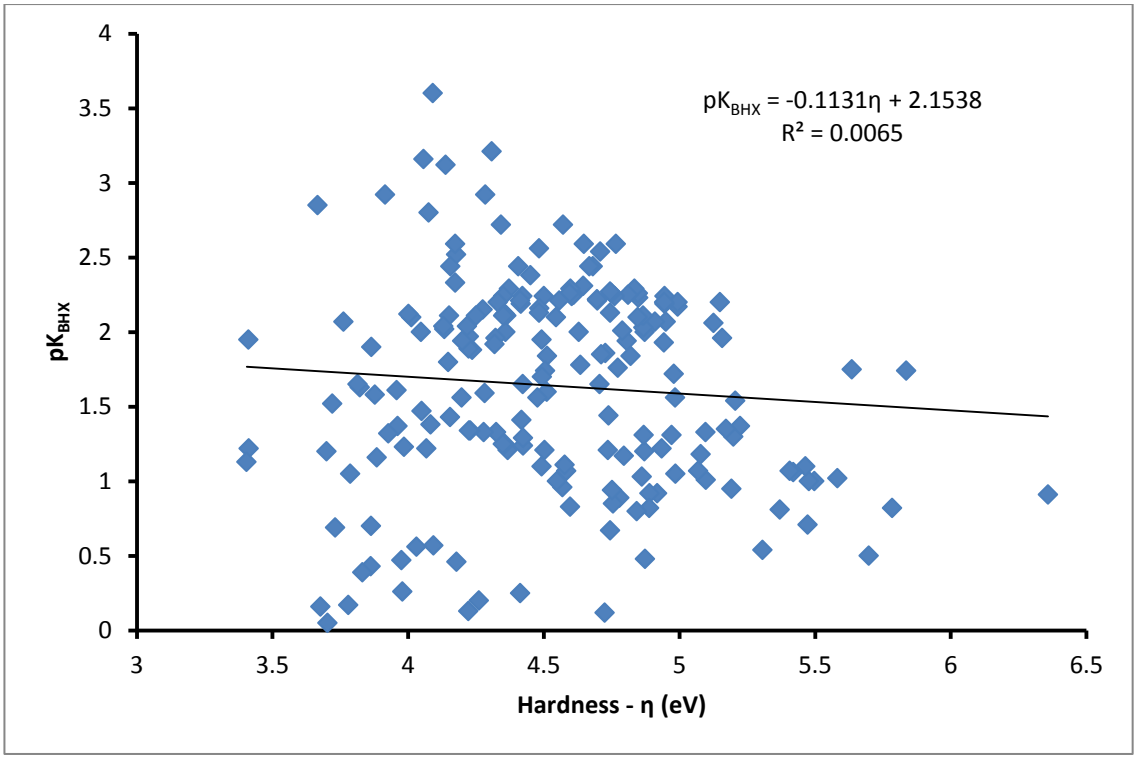


Figure 3.10.1. Hardness values have been computed for the hydrogen bond acceptors listed in table 3.10.1 and plotted against data base pK_{BHX} values [1]. The r^2 value of 0.0065 indicates no relationship between hardness and pK_{BHX}.

The results in figure 3.10.1 show the correlations that have been set up between ΔE(H) and pK_{BHX} where methanol is the hydrogen bond donor. In this section the correlations are taken from 15 base subsets from within the data set of 197 hydrogen bond donors listed in table 3.10.1. The first subset includes the 15 bases with the highest hardness values. The first subset therefore includes the first 15 bases listed in table 3.10.1. The second subset is made up of the bases ranked between 2 and 16 in table 3.10.1. In other words there is a sliding window containing 15 bases. Each new window replaces the base with the largest hardness value in the previous window with the base with the next largest hardness value outside that of the previous window. For each subset r^2 values have been calculated. The r^2 values for each subset have been plotted against the mean hardness value of each subset.

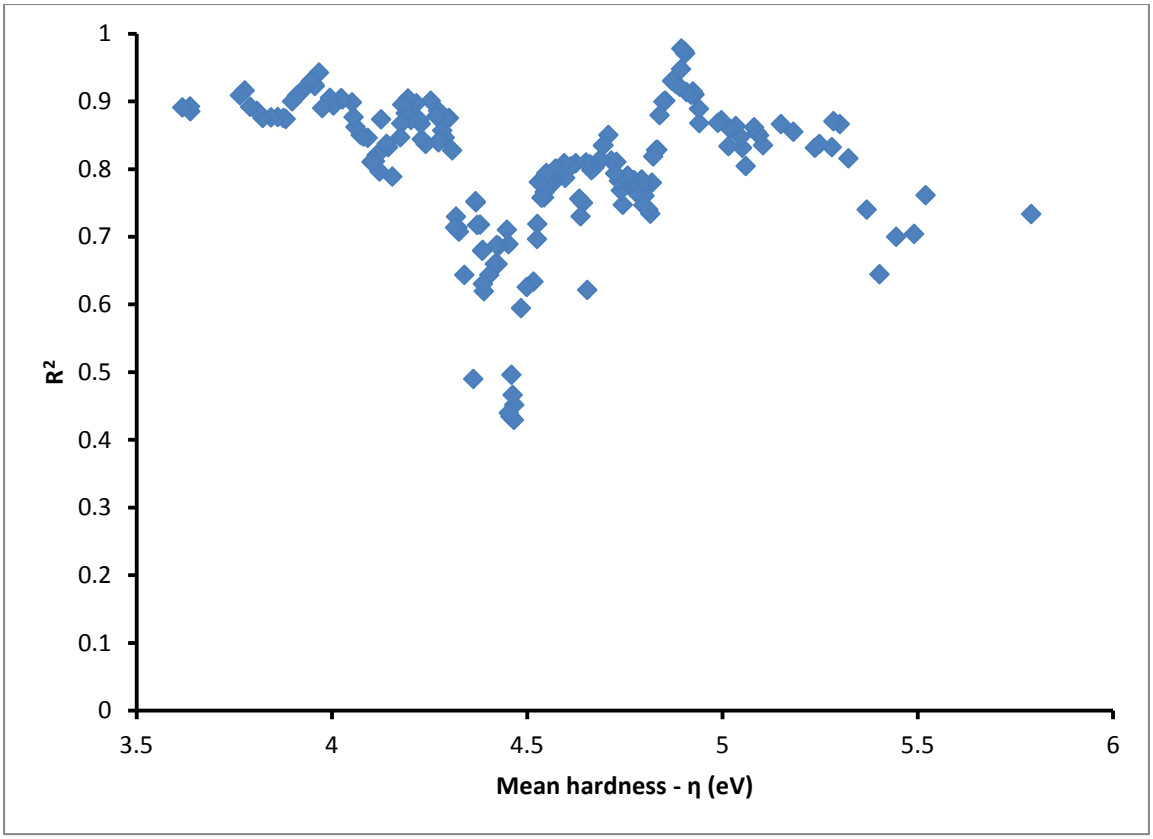


Figure 3.10.2. From the bases listed in table 3.10.1, 15 base subsets have been chosen according to the criteria in the above text. The mean hardness value is plotted for each 15 base subset. The r^2 value is taken from correlations between the hardness value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methanol is the hydrogen bond donor.

It can be seen from figure 3.10.2 that the strength of the correlation between pK_{BHX} and ΔE(H) is not dependent on the hardness of the hydrogen bond acceptors. The correlation between pK_{BHX} and ΔE(H) is generally high, r^2 = 0.8 to 0.9 across the range of hardnesses with the exception being a drop in the region of η = 4.2 to 4.7 eV. There also appears to be a slight drop in r^2 values when the hardness values rise above 5 eV. Methanol, the hydrogen bond donor used to compute ΔE(H) values, has a hardness value of 5.78 eV. The HSAB principle states that methanol should form the strongest hydrogen bond complexes with hydrogen bond acceptors with hardness values close to 5.78 eV. Once again it is possible that the computation of methanol complexed with

similarly hard bases over-estimates the strength of the hydrogen bond. The possible overestimation could be due to the fact that the acid : base ratio of 1 : 1 in the computation is void of experimental effects that could possibly weaken the hydrogen bond in question. The experimental effect that could possibly weaken the hydrogen bond in question could be hydrogen bond networks forming in the experimental solution that distort the hydrogen bond from that of the computational hydrogen bond.

The above procedure has been repeated substituting methanol as the hydrogen bond donor for methylamine. Methylamine is slightly softer than methanol and has a computed hardness value of 5.15 eV. The results are shown in figure 3.10.3.

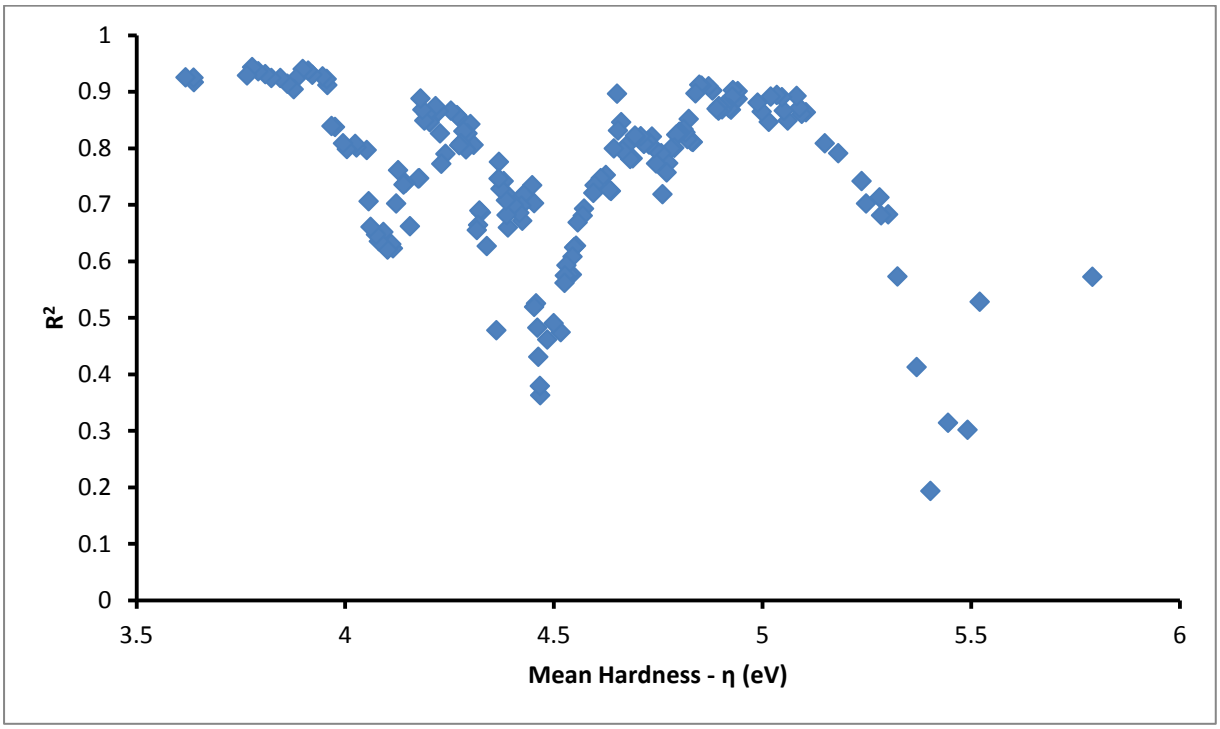


Figure 3.10.3. From the bases listed in table 3.10.1, 15 base subsets have been chosen according to the criteria in the above text. The mean hardness value is plotted for each 15 base subset. The r^2 value is taken from correlations between the hardness value of each base and the computed ΔE(H) value for each base in each 15 base subset. Computed ΔE(H) values are taken from computed hydrogen bond complexes where methylamine is the hydrogen bond donor.

Figure 3.10.3 reveals a very similar pattern to figure 3.10.2. Therefore the correlation between pK_{BHX} and ΔE(H) is not only not obviously dependent on hardness

but also appears independent of the hydrogen bond donor despite its pK_{BH}^{+} or hardness value. One slight difference in the pattern of data between figures 3.10.2 and 3.10.3 is that the r^2 values seem to have a sharper decline as the mean hardness value rises above 5eV when methylamine is used as the hydrogen bond donor. One possible conclusion that can be drawn from this is that pK_{BHX} does not correlate well with ΔE(H) when the hardness value of the hydrogen bond acceptor is greater than that of the hydrogen bond donor. The possible reasons as to why the pK_{BHX} does not correlate well with ΔE(H) for these bases are described above in the context of a methanol hydrogen bond donor. The trend in the results shown in figure 3.10.2 and 3.10.3 is not obvious enough to draw any firm conclusions other than that there is a generally strong correlation between pK_{BHX} and ΔE(H) across the range of hardness values of bases used in this research. This is highlighted in figures 3.10.4 and 3.10.5 which show the relationship between pK_{BHX} and ΔE(H) for the entire 197 hydrogen bond acceptors used in this section. The r^2 values of 0.83 and 0.79 where methanol and methylamine are used as hydrogen bond donors respectively indicate a general correlation between pK_{BHX} and ΔE(H). This helps to explain the consistently high r^2 values observed in figure 3.10.4 and figure 3.10.5.

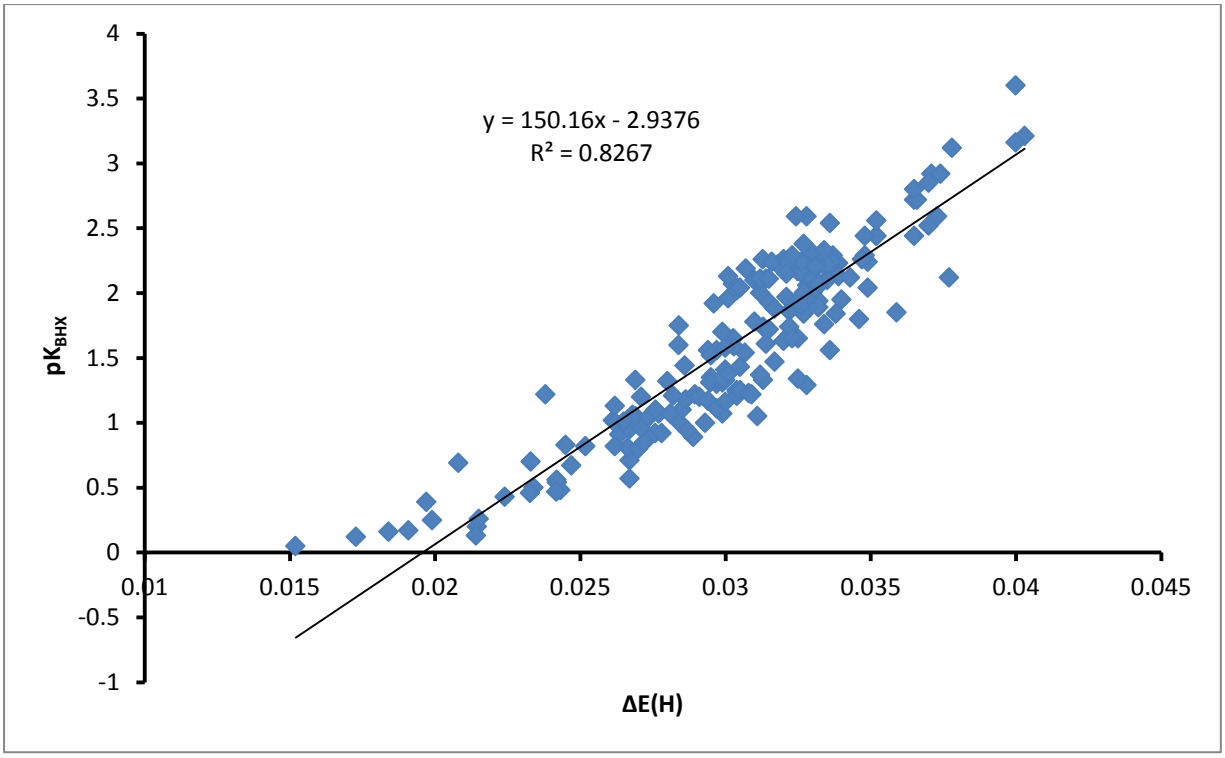


Figure 3.10.4. The overall relationship between pK_{BHX} [1] and ΔE(H) for all 197 hydrogen bond donors listed in table 3.10.1. Methanol is the hydrogen bond donor.

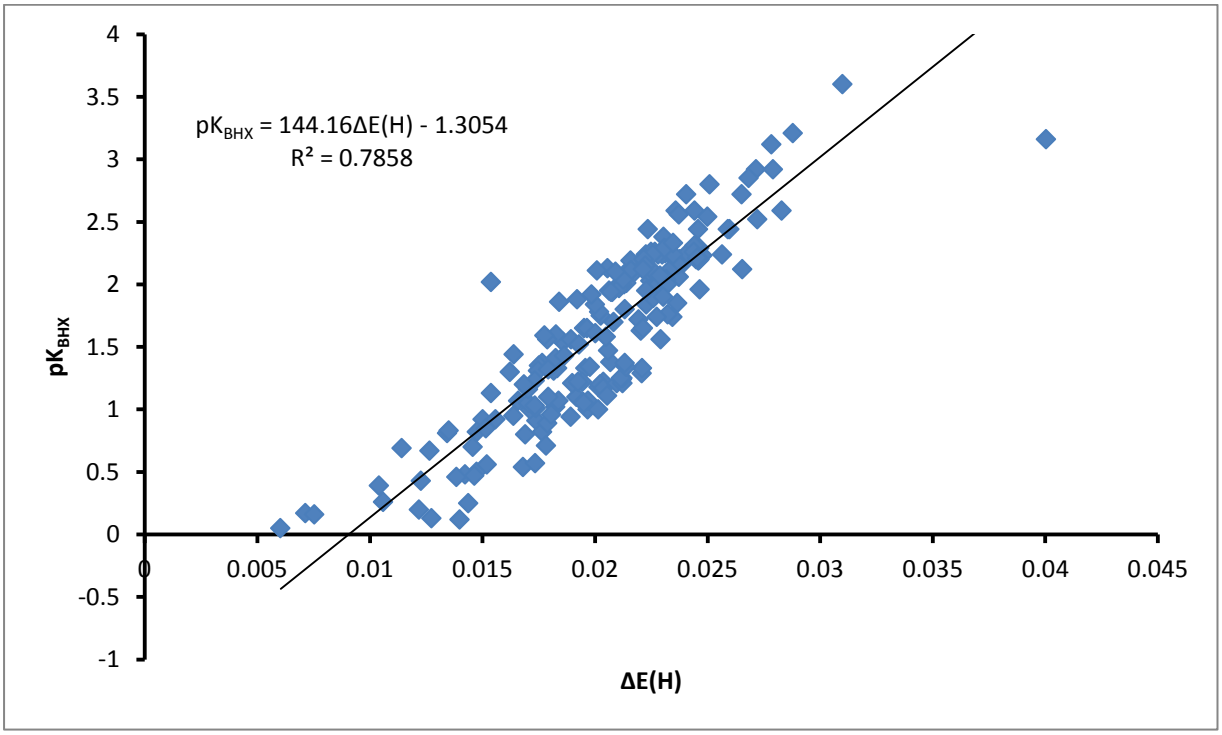


Figure 3.10.5. The overall relationship between pK_{BHX} [1] and ΔE(H) for all 197 hydrogen bond donors listed in table 3.10.1. Methylamine is the hydrogen bond donor.

3.11 Prediction Of pK_{BHX} Values Where The Hydrogen Bond Acceptor Accepts Two Separate Hydrogen Bonds

There are many types of hydrogen bond complexes occurring in biological interactions. Each hydrogen bond complex is unique and will have a unique set of thermodynamic data associated with it that no general hydrogen bond scale will be able to predict. The pK_{BHX} scale serves as a good system that allows hydrogen bond acceptors to be ranked in order of strength. Therefore a hydrogen bond acceptor site can be altered accordingly to either strengthen or weaken a hydrogen bond interaction even though no thermodynamic data can be obtained. Despite the potential problems associated with the practical applications of the pK_{BHX} scale stated in the above text, the pK_{BHX} scale still has a place in medicinal chemistry in ranking hydrogen bond basicity strength.

The strength of the hydrogen bond acceptors that can be ranked by the pK_{BHX} scale are based on the principle that a hydrogen bond acceptor site accepts one hydrogen bond from one hydrogen bond donor. Hydrogen bonds in drug binding interactions are unfortunately often not quite so simple. The three centred hydrogen bond [180] where two acceptor sites each receive a hydrogen bond from one hydrogen bond donor has been discussed in the literature. During this research much time has been spent searching through the protein data bank for drug interaction case studies to investigate (see section 3.5). It was found that on several occasions that the hydrogen bond interactions of drug – target complexes often involved two hydrogen bond donors donating a hydrogen bond each to the same hydrogen bond acceptor site. An example of this are illustrated in figure 3.11.1.

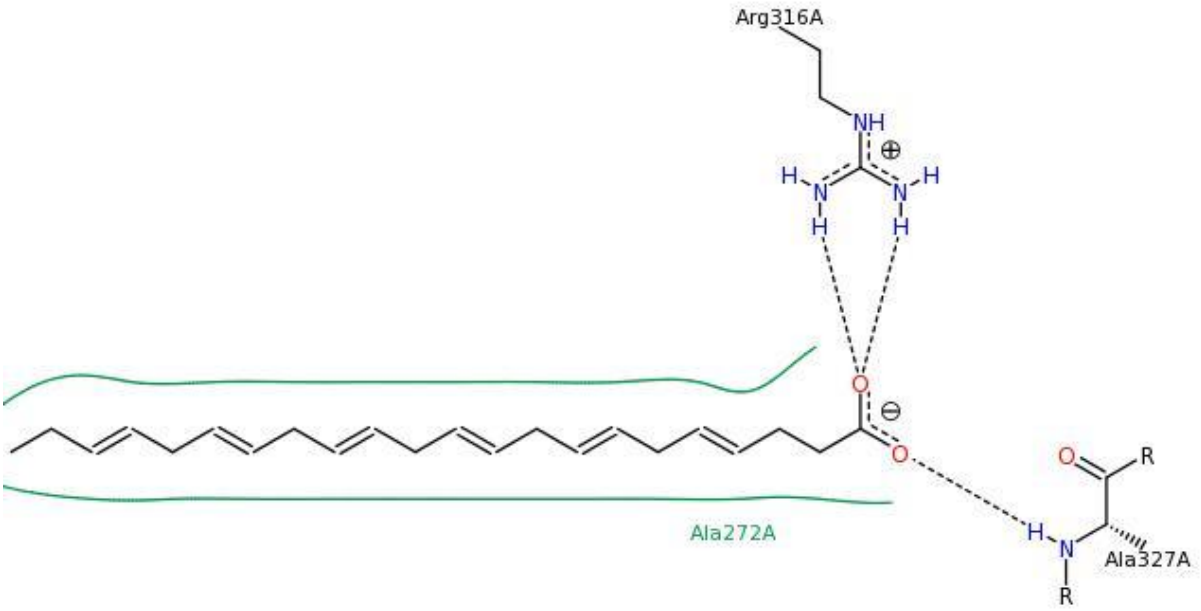


Figure 3.11.1. The interaction between the human RXH alpha ligand binding domain and docosa hexanoic acid [2]. The interaction image is taken from the protein data bank [3].

The research in this section will focus on cases where a single hydrogen bond acceptor site accepts two hydrogen bonds from two hydrogen bond donors. The hydrogen bond acceptor site will have one pK_{BHX} value. It has already been demonstrated in section 3.2 how pK_{BHX} values correlate well with computed ΔE(H) values. These correlations have been set up using ΔE(H) values obtained from optimised complexes in the donor : acceptor ratio of 1 : 1. It has been shown in section 3.2 how computed ΔE(H) values can be used to predict pK_{BHX} values. The following research investigates the effect on the correlations between pK_{BHX} and ΔE(H), and the predictability of pK_{BHX} values from computed ΔE(H) values obtained from a model where two separate hydrogen bond acceptors each donate a hydrogen bond to a single hydrogen bond acceptor site.

Hydrogen bond complexes have been set up for the hydrogen bond acceptors listed in table 3.11.1. The hydrogen bond acceptors chosen are based on the original test set used in section 3.2 and have been reduced in such a way as to maintain the

distribution of pK_{BHX} values as per the pK_{BHX} database. Each hydrogen bond acceptor listed in table 3.11.1 accepts two hydrogen bonds from two separate hydrogen bond donors. The hydrogen bond donor that has been used is methanol. An example of such a complex is shown in figure 3.11.2 using ethanol as an example.

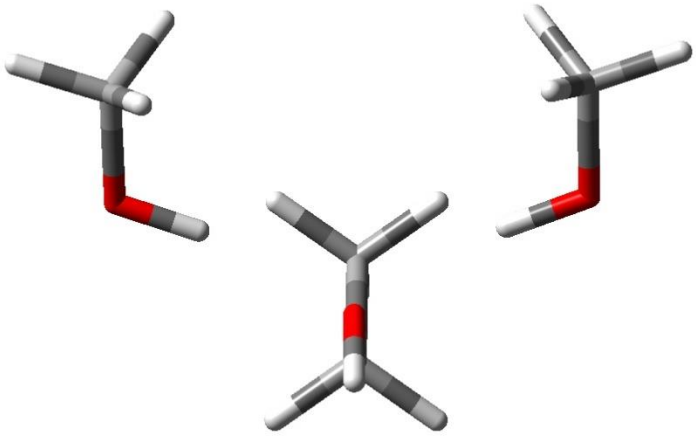


Figure 3.11.2. An example of a hydrogen bond acceptor accepting two separate hydrogen bonds from two separate hydrogen bond donors. In this example ethanol is accepting two separate hydrogen bonds from two methanol hydrogen bond donor molecules.

The hydrogen bond complex geometries have been optimised as described in chapter 2. However in this section a loose optimisation criteria has been chosen to save computational time. It has been shown in section 3.4 that a loose optimisation criteria does not reduce the strength of the correlation between pK_{BHX} and $\Delta E(H)$. As described in chapter 2 $\Delta E(H)$ values have been computed.

Hydrogen Bond Acceptor	pK_{BHX} [1]
Ethyl thiol	-0.16
Phenol	-0.07
4-methylphenol	0.03
Dimethyl sulfide	0.12
p-toluidine	0.56
Methylformate	0.65
Acrylonitrile	0.7
Methanol	0.82
Acetonitrile	0.91
Ethanol	1.02
Acetone	1.18
Pyridine	1.86
Acetamide	2.06
Triethylarsine oxide	4.89
3-methylphenol	0.01

Table 3.11.1. The hydrogen bond acceptors used in this section. Each hydrogen bond acceptor accepts two separate hydrogen bonds from methanol. The details of how the complexes have been set up are described in the text.

Figure 3.11.3 shows the correlation between pK_{BHX} and $\Delta E(H)$ for the hydrogen bond acceptors listed in table 3.11.1 where a single methanol hydrogen bond donor donates one hydrogen bond. The straight line equation shown in figure 3.11.3 will be used in the remainder of this section. The straight line equation will be used to predict

pK_{BHX} values of hydrogen bond acceptors in complexes where each acceptor receives two separate hydrogen bonds from two separate methanol hydrogen bond donors.

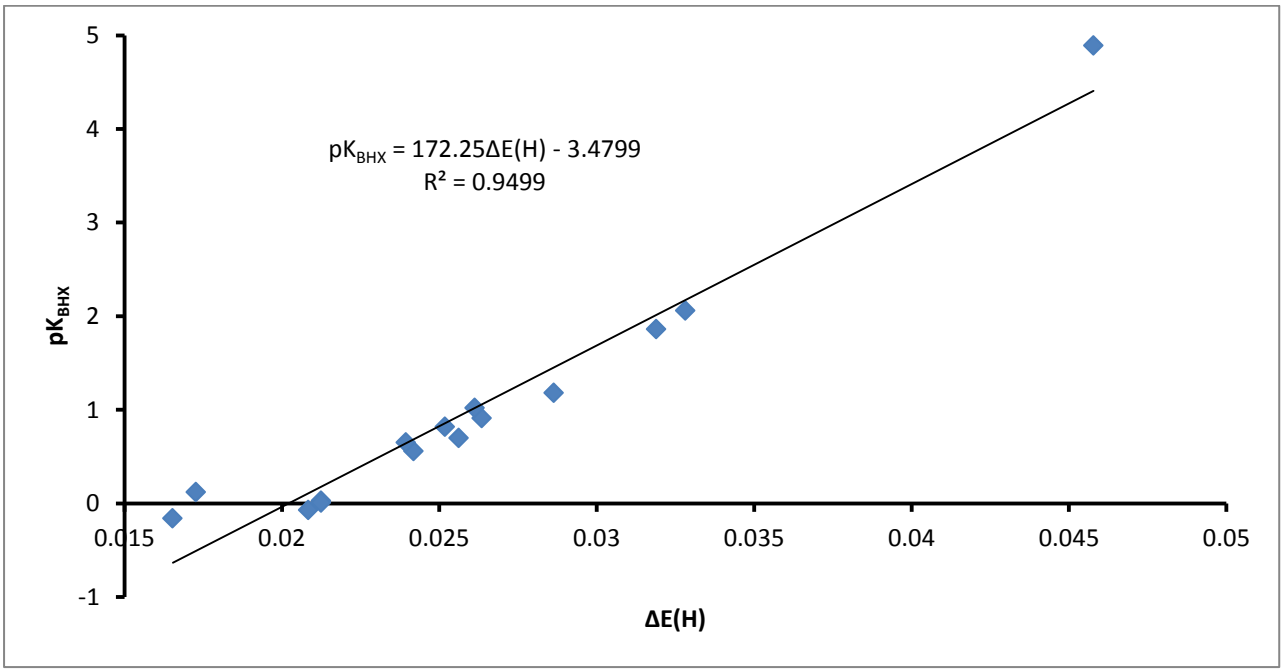


Figure 3.11.3. The relationship between pK_{BHX} [1] and ΔE(H) for the set of 15 hydrogen bond acceptors listed in table 3.11.1 Methanol is the hydrogen bond donor. A loose optimisation criteria has been used during the geometry optimisation phase.

The relationship between pK_{BHX} and ΔE(H) for the hydrogen bond acceptors can be seen in figure 3.11.3. The r^2 value of 0.95 indicates a strong correlation between pK_{BHX} and ΔE(H) for the hydrogen bond acceptors listed in table 3.11.1 even though a loose optimisation criteria has been used.

The equation for the linear relationship between pK_{BHX} and ΔE(H) can be used to predict unknown pK_{BHX} values for hydrogen bond acceptors. The equation for the linear relationship between pK_{BHX} and ΔE(H) is stated in again in equation 3.11.1. The equation can also be used to extrapolate beyond the data set of its origin and accurately predict pK_{BHX} values

$$pK_{BHX} = 172.25\Delta E(H) - 3.4799 \qquad (3.11.1).$$

The aim of the research in this section is to investigate the relationship between pK_{BHX} and ΔE(H) when there are two separate hydrogen bond donors each donating a hydrogen bond to a single hydrogen bond acceptor. Each hydrogen bond acceptor has one pK_{BHX} value, however in this instance there will be two ΔE(H) values. Each ΔE(H) value can be inserted into equation 3.11.1. Substituting each ΔE(H) value into equation 3.11.1 will generate two predicted pK_{BHX} values, each value representing one of the two separate hydrogen bonds. However each hydrogen bond acceptor has only one pK_{BHX} value. Each predicted pK_{BHX} value can be converted into a predicted equilibrium constant value following equation 1.4.15.

A single predicted pK_{BHX} value may be calculated by summing each predicted formation constant and taking the logarithm of the total equilibrium constant. Equations 3.11.3 and 3.11.4 demonstrate how this is possible. The subscript *t* in equations 1.4.10 and 3.11.2 indicate the total values are to be used

$$pK_{BHX_t} = \log_{10} K_t \qquad (3.11.2).$$

Table 3.11.2 shows the results of such calculations for the 15 hydrogen bond acceptors listed in table 3.11.1.

Hydrogen Bond Acceptor	pK_{BHX}	$\Delta E(H)_1$	$\Delta E(H)_2$	Predicted pK_{BHX1}	Predicted pK_{BHX2}	Predicted Ki_1	Predicted Ki_2	Total pK_{BHX}
Ethyl thiol	-0.16	0.014224	0.013986	-1.0299	-1.07082	0.093347	0.084952	-0.74885
Phenol	-0.07	0.005638	0.007285	-2.50877	-2.22511	0.003099	0.005955	-2.04315
4-methylphenol	0.03	-3.68452	0.010022	-638.138	-1.75361	0	0.017635	-1.75361
Dimethyl sulfide	0.12	0.015101	0.015364	-0.87883	-0.83341	0.13218	0.146754	-0.5545
p-toluidine	0.56	0.000972	0.019534	-3.3125	-0.11521	0.000487	0.766982	-0.11494
Methylformate	0.65	0.021586	0.023456	0.238374	0.560353	1.731307	3.63373	0.729573
Acrylonitrile	0.7	0.014459	0.01546	-0.98931	-0.81691	0.102492	0.152436	-0.59358
Methanol	0.82	0.020766	0.020724	0.097015	0.089739	1.250303	1.22953	0.394422
Acetonitrile	0.91	0.014324	0.016631	-1.01261	-0.6152	0.097138	0.242547	-0.46892
Ethanol	1.02	0.021925	0.021982	0.296734	0.306478	1.980316	2.025247	0.602664
Acetone	1.18	0.025795	0.025819	0.963308	0.967375	9.189851	9.276309	1.266377
Pyridine	1.86	0.027345	0.013551	1.230288	-1.14568	16.99369	0.071502	1.232111
Acetamide	2.06	0.036519	0.031799	2.810564	1.997464	646.4939	99.41774	2.872687
Triethylarsine oxide	4.89	0.043532	0.043198	4.018406	3.960994	10432.93	9141.006	4.291678
3-methylphenol	0.01	-34.2391	0.008066	-5901.16	-2.09055	0	0.008118	-2.09055

Table 3.11.2. The data in this table relates to computations performed on hydrogen bond complexes where each hydrogen bond acceptor receives two separate hydrogen bonds from two methanol hydrogen bond donors. The subscripts 1 and 2 are put in place to distinguish between data relating to each individual hydrogen bond. Details of how predicted pK_{BHX} [1] and *Ki* values were calculated are given in the text.

The first comment to make about the data in table 3.11.2 is that there appears to be an anomaly concerning $\Delta E(H)_1$ for 4-methylphenol and 3-methylphenol. The $\Delta E(H)_1$ value of 4-methylphenol and 3-methylphenol is much lower than expected. The optimised geometry for the 4-methylphenol is shown in figure 3.11.4.

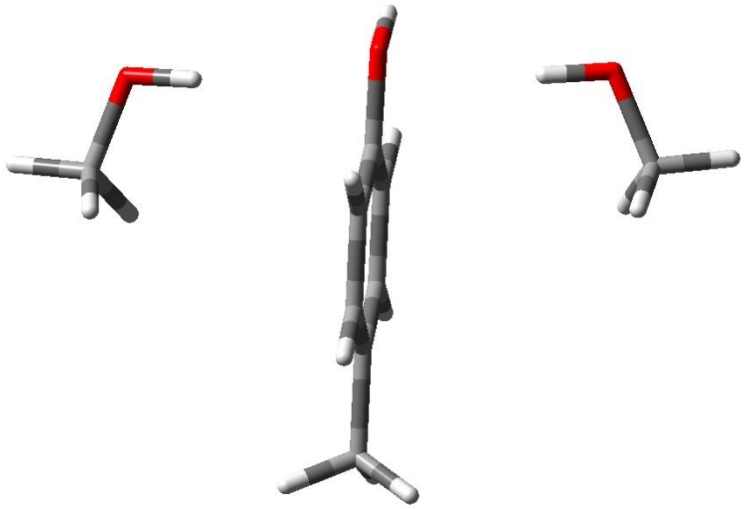


Figure 3.11.4. Optimised geometry of the 4-methylphenol complex. Two separate methanol hydrogen bond donors each donate a hydrogen bond to 4-methylphenol. As can be seen here the complex is virtually symmetrical, however the hydrogen bond on the left generates a much lower ΔE(H) value than the hydrogen bond on the right.

It can be seen from figure 3.11.4 that the optimised geometry of the 4-methylphenol complex is fairly symmetrical about the hydrogen bond acceptor. However, the hydrogen bond on the left in figure 3.11.4 generates a much lower ΔE(H) value than the hydrogen bond on the right. The reason, as shown in figure 3.11.4, is not due to any steriochemical effect. The reason must therefore be attributed to computation, either in the wave function or the integration phase of obtaining ΔE(H) values. As no reason can be attributed to the anomalies concerning the 4-methylphenol and 3-methylphenol complexes, they will be omitted from the following analysis.

The actual pK_{BHX} values and the predicted total pK_{BHX} values for all the hydrogen bond acceptors in table 3.11.2 have been plotted against each other. The plot is shown in figure 3.11.5. For reasons explained above 4-methylphenol and 3-methylphenol have been omitted from the plot.

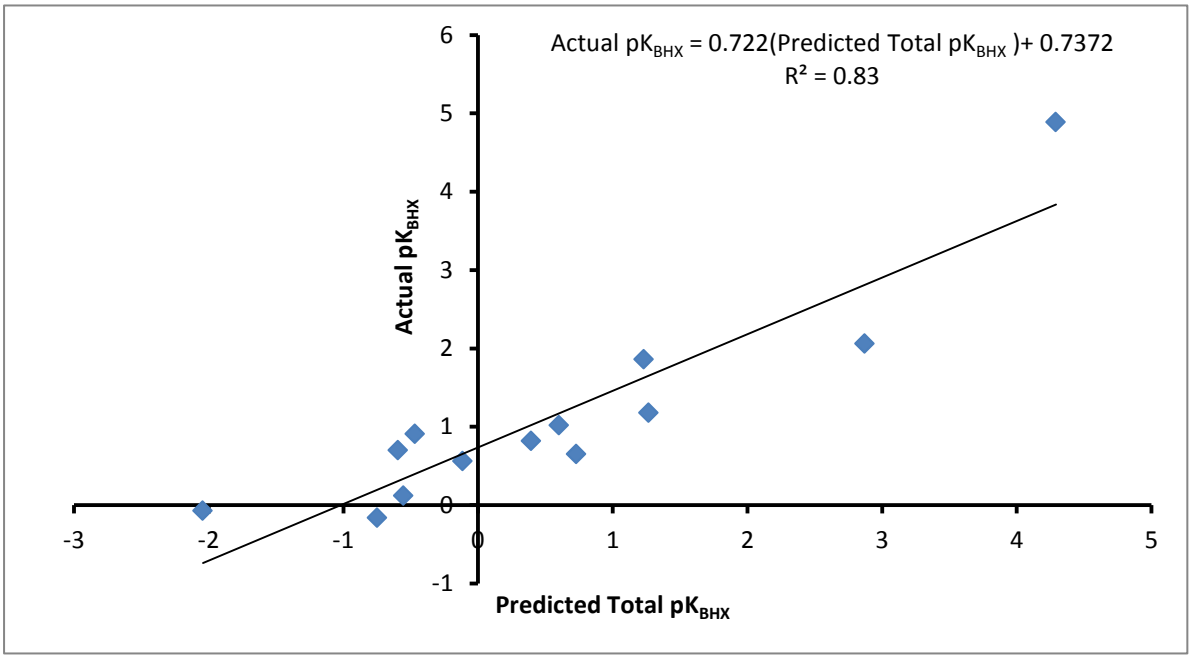


Figure 3.11.5. The data plotted here is taken from table 3.11.2. The predicted total pK_{BHX} relates to the hydrogen bond complexes where two individual methanol donors each donate a separate hydrogen bond to a single hydrogen bond acceptor site. Data for 3-methylphenol and 4-methylphenol have been omitted from this plot for reasons explained in the above text.

It can be seen from figure 3.11.5 that there is reasonably strong correlation between the computed predicted total pK_{BHX} values and the actual pK_{BHX} values taken from the database. The r^2 value is 0.83. However, closer inspection reveals that there are often large discrepancies between the actual and predicted pK_{BHX} values. Therefore, even though a relationship exists between the predicted and actual pK_{BHX} values, computed total pK_{BHX} values taken from hydrogen bond complexes where each hydrogen bond acceptor receives two separate hydrogen bonds from two separate methanol donors do not compare well to actual pK_{BHX} values.

To understand why the predicted total pK_{BHX} values in table 3.11.2 do not compare well with actual values, the optimised geometries of two examples will be highlighted. The first example to be discussed is the acetone complex. Figure 3.11.6 shows the optimised geometry of the acetone complex.

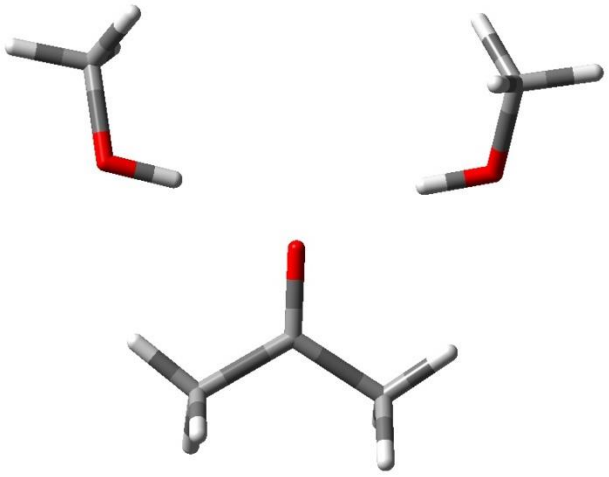


Figure 3.11.6. An example of a hydrogen bond acceptor accepting two separate hydrogen bonds from two separate hydrogen bond donors. In this example acetone is accepting two separate hydrogen bonds from two methanol hydrogen bond donor molecules. The hydrogen bond donor methanols are distributed symmetrically about the hydrogen bond acceptor site.

It can be seen in figure 3.11.6 that the acetone complex displays symmetry about the hydrogen bond acceptor oxygen atom. This symmetry is reflected in the results as the predicted pK_{BHX} values for acetone relating to the first and second methanol molecules are 0.96 and 0.97 respectively. The symmetry is also reflected in that the total predicted pK_{BHX} of 1.27 compares well with the actual value of 1.18. It can be concluded that when a hydrogen bond acceptor accepts two separate hydrogen bonds from two separate hydrogen bond donors that are positioned symmetrically about the acceptor site, each hydrogen bond contributes equally to the total pK_{BHX} value. It follows that the summation of $\Delta E(H)$ for each hydrogen bond yields an accurate predicted total pK_{BHX} value. However, when there is an absence of symmetry of hydrogen bond donors about the acceptor site, predicted total pK_{BHX} values do not compare well with actual

pK_{BHX} values. The acetonitrile complex is an example of a complex with no symmetry and is shown in figure 3.11.7.

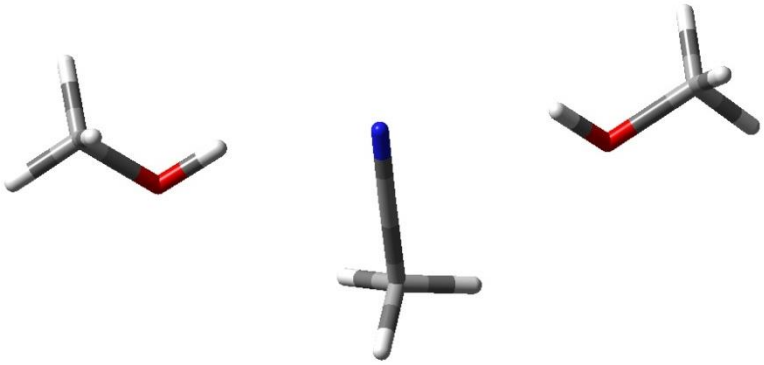


Figure 3.11.7. An example of a hydrogen bond acceptor accepting two separate hydrogen bonds from two separate hydrogen bond donors. In this example acetonitrile is accepting two separate hydrogen bonds from two methanol hydrogen bond donor molecules. The hydrogen bond on the right has a gives a weaker estimated pK_{BHX} than the one on the left. This is possibly due to a secondary interaction between the methanol oxygen on the right hand side and the methyl hydrogen of acetonitrile.

It can be seen from figure 3.11.7 that there is a possible secondary stabilising interaction acting on the methanol on the right hand side. This interaction is between the oxygen atom of the right hand side methanol and the hydrogen atom of the methyl group of acetonitrile. It is possible that the methanol accepts a hydrogen bond from the methyl hydrogen of acetonitrile. This possible hydrogen bond could weaken the hydrogen bond inbetween that methanol donates to the nitrogen acceptor site of acetonitrile. This is reflected in the results as table 3.11.2 shows that the hydrogen bond on the right, that is indicated by subscript 1, is weaker than the hydrogen bond on the left, indicated by subscript 2. The predicted pK_{BHX} values are -1.01 and -0.61 respectively. This leads to

a predicted total pK_{BHX} value of -0.47 which is much lower than that of the actual value 0.91.

Previous models to estimate pK_{BHX} values have investigated one hydrogen bond donor donating one hydrogen bond to one hydrogen bond acceptor. As the predicted total pK_{BHX} values obtained in this section do not compare well with actual pK_{BHX} values, it is possible that the presence of a second hydrogen bond interferes with the first in such a way that pK_{BHX} values cannot be predicted accurately. To investigate this, two energy point computations will be performed on the optimised complexes used to generate the results in table 3.11.2. The first energy calculation will be performed on the optimised geometry with one methanol molecule deleted. The second energy calculation will be performed on the optimised geometry with the other methanol deleted and the first methanol present. Therefore a $\Delta E(H)$ value will be obtained for each methanol in a 1:1 complex with the hydrogen bond acceptor fixed in the optimised geometry of the 2:1 complexes discussed in the above text. The results are shown in table 3.11.3.

Hydrogen Bond Acceptor	pK_{BHX}	$\Delta E(H)_1$	$\Delta E(H)_2$	Estimated pK_{BHX1}	Estimated pK_{BHX2}	Ki_1	Ki_2	Total pK_{BHX}
Ethyl thiol	-0.16	0.016087	0.015917	-0.70885	-0.73814	0.195501	0.182753	-0.42222
Phenol	-0.07	-5.24846	0.008875	-907.527	-1.95113	0	0.011191	-1.95113
Dimethyl sulfide	0.12	0.017015	0.017364	-0.54904	-0.48897	0.282465	0.324364	-0.21693
p-toluidine	0.56	0.003138	0.020773	-2.93937	0.098185	0.00115	1.253676	0.098583
Methylformate	0.65	0.024375	0.024597	0.718677	0.756911	5.232111	5.713611	1.039244
Acrylonitrile	0.7	0.014919	0.016647	-0.91018	-0.6124	0.122975	0.244118	-0.43522
Methanol	0.82	0.023573	0.023556	0.58057	0.577661	3.806891	3.78147	0.880148
Acetonitrile	0.91	0.015353	0.017101	-0.83541	-0.53427	0.146079	0.292236	-0.35821
Ethanol	1.02	0.024668	0.024692	0.769088	0.773243	5.876078	5.932573	1.0722
Acetone	1.18	0.028275	0.028297	1.39046	1.394243	24.57311	24.78807	1.693386
Pyridine	1.86	0.03064	0.016598	1.797871	-0.62088	62.7872	0.239397	1.799524
Acetamide	2.06	0.039184	0.034171	3.269609	2.40606	1860.413	254.7183	3.325337
Triethylarsine oxide	4.89	0.045911	0.045865	4.428246	4.420422	26806.83	26328.25	4.725381

Table 3.11.3. The data in this table considers each hydrogen bond separately. The optimised geometry is from the hydrogen bond complexes where each hydrogen bond acceptor receives two separate hydrogen bonds from two methanol hydrogen bond donors. However this time each hydrogen bond has in turn been deleted and energy point calculations performed on each single hydrogen bond complex.

Table 3.11.3 reveals another possible anomaly concerning $\Delta E(H)_1$ of phenol. The issues regarding this anomaly can be associated with those of the 3-methylphenol and 4-methylphenol complexes as discussed above. The data obtained for phenol will be omitted from the rest of this discussion. However, the general conclusion that can be drawn from table 3.11.3 is that predicted total pK_{BHX} values do not compare to actual pK_{BHX} any better when two 1:1 donor acceptor complexes are used in the optimised geometry of the 2:1 complex, than the single 2:1 complex itself. As neither two 1:1 complexes nor one 2:1 complex generate total pK_{BHX} values that compare well with the actual pK_{BHX} value, the geometry of the complex must be important. The geometry of an optimised single 1:1 complex must be key to the relationship between $\Delta E(H)$ and pK_{BHX}.

Unfortunately accurate pK_{BHX} values have not been predicted from summing the $\Delta E(H)$ values of each hydrogen bond in complexes where two separate hydrogen bond donors each donate a hydrogen bond to a single hydrogen bond acceptor. In previous sections it has been shown that there is a strong connection between $\Delta E(H)$ and pK_{BHX}. These accurate predictions are for an optimised 1:1 complex that models a 1:1 experimental complex. Therefore the actual pK_{BHX} value of a hydrogen bond acceptor does not change when the hydrogen bond acceptor becomes involved in a 2:1 complex as modelled here. The pK_{BHX} value does not relate to the summation of each Ki value in a 2:1 complex. Therefore the equilibrium constant for the base in an excess of acid as in the experimental set up cannot be related to the base in a biological environment as in figure 3.11.1 where two hydrogen bonds are present. The steric effects of the second hydrogen bond complex must interfere with the strength of the complex in such a way that the hydrogen bond acceptor in the 2:1 complex does not have the same basicity as the hydrogen bond acceptor in a 1:1 complex.

Once again it has been found that the pK_{BHX} scale offers little thermodynamic information when related to realistic biological situations. Although it has been stated that the pK_{BHX} scale can be used to generate ΔG values, these ΔG values are only

specific to the experimental situation and cannot be extrapolated to predict binding data of biological interactions.

4 Conclusions.

The aim of this research has been to find a way to accurately and quickly predict hydrogen bond basicity values. In order to do so, a suitable hydrogen bond basicity scale had to be chosen from the many available. The pK_{BHX} scale was chosen. The advantages of the pK_{BHX} scale are that an individual thermodynamically viable basicity value is given for each site on a polyfunctional base. Also the pK_{BHX} database [1] contains over 1100 hydrogen bond acceptors. Both of these stated advantages lead the authors to claim that the pK_{BHX} scale is the most applicable to medicinal chemists.

A property obtained by applying the quantum theory of atoms in molecules was found to correlate well with pK_{BHX} values. The property that was found to correlate with pK_{BHX} values is the change in the atomic energy of the hydrogen bond donor hydrogen atom upon binding ΔE(H). Not only was ΔE(H) found to correlate to pK_{BHX} values, it was also found that data could be used to predict the pK_{BHX} values of an external test set of hydrogen bond acceptors.

The property ΔE(H) also correlated well to pK_{BHX} values of polyfunctional bases. It was found that the correlations of values taken from 1 : 1 complexes slightly outperformed the correlations taken from 2 : 1 complexes.

Polyfunctional bases are ample in medicinal chemistry. There are also often lots of incidents of 3 centred hydrogen bonds and hydrogen bond complexes where the two individual hydrogen bond donors each donate to a single hydrogen bond acceptor. Unfortunately this research has been unable to find a link between ΔE(H) and the latter case. It must be noted that pK_{BHX} values are taken from formation constants of the base in an excess of acid. The basicity of a hydrogen bond acceptor in biological interaction does not mimic the experimental setup. Therefore it can be concluded that although the relative basicity of the hydrogen bond acceptor will be unchanged in the biological interaction, other factors will influence the strength of the hydrogen bond actually

formed. This is further highlighted by the unsuccessful attempts to predict known Ki values from predicted hydrogen bond basicity values.

It could be that the strength of the hydrogen bond in a biological interaction is different from what it would be in the experimental set up. It could also be that other thermodynamic effects that are absent from the experimental setup and specific to each biological interaction play an important role in drug binding data. The entropic and enthalpic effects have been discussed in detail in the main body of the report.

In order for the link between ΔE(H) and pK_{BHX} to become usable, ΔE(H) values must be obtained quickly. Unfortunately there are two limiting factors to the generation of ΔE(H) values. The first being the geometry optimisation step that leads to wave function generation. The second being the atomic integration of the hydrogen atom from the wave function. Two ways have been found to speed up the geometry optimisation. The first way reduces the size of the geometry to be optimised by fragmenting the hydrogen bond acceptor. A successful method of fragmenting the hydrogen bond acceptors has been obtained. It was also found that the geometry optimisation time can be reduced by enforcing a loose optimisation criteria. The ΔE(H) values obtained from wave functions generated in such a way correlate almost as well to pK_{BHX} as the ΔE(H) values generated using standard optimisation criteria.

Following on from finding a successful way of fragmenting hydrogen bond acceptors without altering the accuracy of the predictions of basicity, a case study was carried out. Strychnine was chosen as the example. Strychnine was fragmented into its three hydrogen bond acceptor sites. It was found that the ΔE(H) of the tertiary amine fragment did not correlate well with the pK_{BHX} value. It was found that tertiary amines in general do not give ΔE(H) values that correlate well with pK_{BHX} values.

It was hypothesised that tertiary amines are strong proton acceptors and would perhaps become protonated in the experiment leading to overestimated basicity values. Although it was found that the relationship between pK_{BHX} and ΔE(H) did break down for strong proton acceptors in general, the reason why is still unknown.

It was also hypothesised that the hardness values of hydrogen bond acceptors could affect the correlations between pK_{BHX} and ΔE(H). Given that hard bases form stronger complexes with hard acids and vice versa it was hypothesised that only hydrogen bond acceptors with hardness values in the region of that of methanol would give ΔE(H) values that correlate well with pK_{BHX}. This was not found to be true.

It has been found that ΔE(H) correlates well with pK_{BHX} values. The reason why this is so is unknown. However, this only holds true for hydrogen bond acceptors with pK_{BH}^{+} values between 0 and 6. The reason why is also unknown.

The pK_{BHX} data base is aimed at medicinal chemists and is claimed to be the most valid by its creators. The thermodynamic information obtained from pK_{BHX} values relates only to the binding situation of the complex formation in the experimental setup. It cannot be extrapolated to biological interactions. The strength of hydrogen bond interactions in biological interactions is largely governed by steric effects. Therefore the pK_{BHX} scale is a scale of relative basicities the cannot be extrapolated beyond the experimental environment of the scale itself. This is because each hydrogen bond interaction is unique in source and environment.

6 Future Work.

The most important discovery of this research is that computed ΔE(H) values accurately predict pK_{BHX} values of mid range proton acceptor strength. However, why this is so is still unknown. At present a mathematical link between ΔE(H) and pK_{BHX} has not been found. Nor has the reason why the relationship between pK_{BHX} and ΔE(H) breaks down as pK_{BH}^{+} values increase above 6 or decrease below 0. In order to tackle both of the stated unknowns, it is essential to tackle the first and find a mathematical proof that pK_{BHX} and ΔE(H) are indeed linked. In doing so a potential explanation as to why the relationship between pK_{BHX} and ΔE(H) is valid for only hydrogen bond acceptors with mid range proton acceptor strength could be offered. Various hypothesis have been proposed as to why the relationship between pK_{BHX} and ΔE(H) break down in sections 3.7 -3.10.

The hypotheses were based on potential protonations occurring in the experimental solution which could lead to a complex formed from interactions other than the one in question. The computation of hydrogen bonded complexes has been performed in such a way to minimise secondary interactions. Therefore if there are secondary interactions present in the experimental solution, then the experiment and the computation are inconsistent. Secondary interactions can be further analysed to consider hydrogen bond networks. Hydrogen bond networks are the series of hydrogen bond interactions in a solution that influence the conformation and hence strength of a complex. If hydrogen bond networks and secondary interactions are present in the experimental solution then the experiment and computation are inconsistent. The inconsistency occurs as multiple hydrogen bonds will be present in the experimental solution whereas the computation models only the hydrogen bond of interest. Future work should aim to understand the experimental procedure, the possibility of secondary interactions and hydrogen bond networks occurring and the effect they have on the hydrogen bond of interest compared to the isolated hydrogen bond of interest.

When planning future work it is important to consider the bigger picture. It must be considered how a model that predicts pK_{BHX} values benefits the scientific

community. The pK_{BHX} scale was designed to benefit the medicinal chemistry community as a useful tool in drug design and development. However, as high throughput technologies advance and chemical knowledge improves, it must be questioned whether a relative scale of hydrogen bond basicities valid for only select hydrogen bond donors is of any use to the drug design process. It is at present not possible to design a drug from scratch with the optimum desired pharmacological profile. From a hydrogen bonding point of view, to design a drug from scratch, specific thermodynamic information must be available for each hydrogen bond present in the drug target interaction. The model established here is not capable of generation accurate thermodynamic information that relates to binding affinities. This is because each hydrogen bond is unique with a unique thermodynamic profile. Also other interactions other than hydrogen bonding affect binding affinities. Future work should consider the possibility of predicting binding affinities of drug target interactions.

6 References.

1. Egea, P.F., A. Mitschler, and D. Moras, *Molecular recognition of agonist Ligands by RXRs.* Mol Endocrinol, 2002. **16**: 987-997.
2. Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**: 235-242.
3. Olesen, P.H., *The use of bioisosteric groups in lead optimization.* Curr. Opin. Drug Discovery Dev., 2001. **4**: 471-8.
4. Clark, R.D., A.M. Ferguson, and R.D. Cramer, *Bioisosterism and molecular diversity.* Perspect. Drug Discovery Des., 1998. **9-11**: 213-224.
5. Burger, A., *Isosterism and bioisosterism in drug design.* Prog Drug Rese, 1991. **37**: 287-371.
6. Lima, L.M. and E.J. Barreiro, *Bioisosterism: a useful strategy for molecular modification and drug design.* Curr. Med.Chem., 2005. **12**: 23-49.
7. Stepan, A.F., V. Mascitti, K. Beaumont, and A.S. Kalgutkar, *Metabolism-guided drug design.* Med. Chem. Comm., 2013. **4**(4): 631-652.
8. Bowen, J.P. and O.F. Guner, *A Perspective on Quantum Mechanics Calculations in ADMET Predictions.* Curr. Top. Med. Chem., 2013. **13**: 1257-1272.
9. de Aquino, R.A.N., L.V. Modolo, R.B. Alves, and A. de Fatima, *Design of New Drugs for the Treatment of Alzheimer's Disease Based on Tacrine Structure.* Curr. Drug Targets, 2013. **14**: 378-397.
10. Showell, G.A. and J.S. Mills, *Chemistry challenges in lead optimization: silicon isosteres in drug discovery.* Drug Discov. Today, 2003. **8**: 551-556.
11. Garcia-Sosa, A.T., U. Maran, and C. Hetenyi, *Molecular Property Filters Describing Pharmacokinetics and Drug Binding.* Curr. Med.Chem., 2012. **19**: 1646-1662.
12. Mugnaini, C., S. Pasquini, and F. Corelli, *The Bioisosteric Concept Applied to Cannabinoid Ligands.* Curr. Med.Chem., 2012. **19**: 4794-4815.
13. Ziarek, J.J., Y. Liu, E. Smith, G.L. Zhang, F.C. Peterson, J. Chen, Y.P. Yu, Y. Chen, B.F. Volkman, and R.S. Li, *Fragment-Based Optimization of Small Molecule CXCL12 Inhibitors for Antagonizing the CXCL12/CXCR4 Interaction.* Curr. Top. Med Chem., 2012. **12**: 2727-2740.
14. Ghate, M. and S.V. Jain, *Structure Based Lead Optimization Approach in Discovery of Selective DPP4 Inhibitors.* Mini-Rev. Med. Chem., 2013. **13**: 888-914.
15. Zhang, L.T., M.E. Cvijic, J. Lippy, J. Myslik, S.L. Brenner, A. Binnie, and J.G. Houston, *Case study: technology initiative led to advanced lead optimization*

screening processes at Bristol-Myers Squibb, 2004-2009. Drug Discov. Today, 2012. **17**: 733-740.

16. Badrinarayan, P. and G.N. Sastry, *Rational Approaches Towards Lead Optimization of Kinase Inhibitors: The Issue of Specificity.* Curr. Pharm. Des., 2013. **19**: 4714-4738.
17. Ou-Yang, S.S., J.Y. Lu, X.Q. Kong, Z.J. Liang, C. Luo, and H.L. Jiang, *Computational drug discovery.* Acta Pharmacol. Sin., 2012. **33**: 1131-1140.
18. Cruz-Monteagudo, M., M. Cordeiro, E. Tejera, E.R. Dominguez, and F. Borges, *Desirability-Based Multi-Objective QSAR in Drug Discovery.* Mini-Rev. Med. Chem., 2012. **12**: 920-935.
19. Xiang, M.L., Y. Cao, W.J. Fan, L.J. Chen, and Y.R. Mo, *Computer-Aided Drug Design: Lead Discovery and Optimization.* Comb. Chem. High Throughput Screening, 2012. **15**: 328-337.
20. Guido, R.V.C., G. Oliva, and A.D. Andricopulo, *Modern Drug Discovery Technologies: Opportunities and Challenges in Lead Discovery.* Comb. Chem. High Throughput Screening, 2011. **14**: 830-839.
21. Willett, P., J.M. Barnard, and G.M. Downs, *Chemical similarity searching.* J. Chem. Inf. Comput. Sci., 1998. **38**: 983-996.
22. Zhang, Y.J., *In Silico Technologies in Drug Design, Discovery and Development.* Curr. Top. Med Chem., 2010. **10**: 617-618.
23. Holliday, J.D., S.P. Jelfs, and P. Willett, *Calculation of Intersubstituent Similarity Using R-Group Descriptors.* J.Chem.Inf. Comp.Sc., 2003. **43**: 406-411.
24. Devereux, M., P.L.A. Popelier, and I.M. McLay, *The Quantum Isostere Database: a web-based tool using Quantum Chemical Topology to predict bioisosteric replacements for drug design.* J.Chem.Inf.&Modell., 2009. **49**: 1497-1513.
25. Friedman, H.L., *Influence of isosteric replacements upon biological activity.* NASNRS, 1951. **206**: 295-358.
26. Langmuir, I., *Isomorphism, Isosterism and Covalence.* J. Am. Chem. Soc., 1919. **41**: 868,1543.
27. Grimm, H.G., *Structure and size of the Non-Metallic Hydrides.* Z. Elektrochem., 1925. **31**: 474.
28. Patani, G.A. and E.J. LaVoie, *Bioisosterism: A rational approach in drug design.* Chem.Rev., 1996. **96**:3147-3176.
29. Thornber, C.W., *Isosterism and Molecular modification in Drug Design.* Chem.Soc.Rev., 1979. **8**: 563-580.
30. Ujvary, I., *BIOSTER-A Database of Structurally Analogous Compounds.* Pestic. Sci., 1997. **51**: 92-95.

31. Aurora, R. and G.D. Rose, *Helix capping.* Protein Science, 1998. **7**(1): 21-38.
32. Nimmrich, V. and A. Eckert, *Calcium channel blockers and dementia.* Br. J. Pharmacol., 2013. **169**: 1203-1210.
33. Lonsdale, R., J.N. Harvey, and A.J. Mulholland, *A practical guide to modelling enzyme-catalysed reactions.* Chem. Soc. Rev., 2012. **41**: 3025-3038.
34. Pang, X.F., *Properties of proton transfer in hydrogen-bonded systems and its experimental evidences and applications in biology.* Prog. Biophs. Mol. Bio., 2013. **112**: 1-32.
35. Fersht, A.R., J.P. Shi, J. Knilljones, D.M. Lowe, A.J. Wilkinson, D.M. Blow, P. Brick, P. Carter, M.M.Y. Waye, and G. Winter, *Hydrogen-Bonding And Biological Specificity Analyzed By Protein Engineering.* Nature, 1985. **314**: 235-238.
36. Berezovsky, I.N., *Thermodynamics of allostery paves a way to allosteric drugs.* BBA-Proteins Proteom., 2013. **1834**: 830-835.
37. Sinko, W., S. Lindert, and J.A. McCammon, *Accounting for Receptor Flexibility and Enhanced Sampling Methods in Computer-Aided Drug Design.* Chem. Biol. Drug. Des., 2013. **81**: 41-49.
38. Raevsky, O.A., *Descriptors of molecular structure in computer-aided design of biologically active compounds.* Uspekhi Khimii, 1999. **68**: 555-576.
39. Kamlet, M.J., M.H. Abraham, R.M. Doherty, and R.W. Taft, *Solubility Properties in Polymers and Biological Media .4. Correlation of Octanol Water Partition-Coefficients with Solvatochromic Parameters.* J. Am. Chem. Soc., 1984. **106**: 464-466.
40. Huggins, D.J., W. Sherman, and B. Tidor, *Rational Approaches to Improving Selectivity in Drug Design.* J. Med. Chem., 2012. **55**: 1424-1444.
41. Freire, E., *A thermodynamic approach to the affinity optimization of drug candidates.* Chem.Biol.Drug Des., 2009. **74**: 468-472.
42. Garbett, N.C. and J.B. Chaires, *Thermodynamic studies for drug design and screening.* Expert Opin. Drug Discov., 2012. **7**: 299-314.
43. Freire, E., *Do enthalpy and entropy distinguish first in class from best in class?* Drug Discov. Today 2008. **13**: 861-866.
44. Guo, Z., *Enthalpy and entropy in drug optimization.* Chinese J. Med. Chem., 2012. **22**: 310-322.
45. Reed, A.E., L.A. Curtiss, and F. Weinhold, *Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint.* Chem. Rev., 1988. **88**: 899-926.
46. Laurence, C., K.A. Brameld, J. Graton, J.-Y. Le Questel, and E. Renault, *The pKBHX Database: Toward a Better Understanding of Hydrogen-Bond Basicity for Medicinal Chemists.* J.Med.Chem., 2009. **52**: 4073-4086.

47. Deli, M.A., C.S. Abraham, Y. Kataoka, and M. Niwa, *Permeability studies on in vitro blood-brain barrier models: Physiology, pathology, and pharmacology.* Cell. Mol. Neurobiol., 2005. **25**: 59-127.
48. Platts, J.A. and M.H. Abraham, *Partition of volatile organic compounds from air and from water into plant cuticular matrix: An LFER analysis.* Environ. Sci. Technol., 2000. **34**: 318-323.
49. Abraham, M.H., *Scales of Solute Hydrogen-Bonding - Their Construction and Application to Physicochemical and Biochemical Processes.* Chem.Soc.Rev., 1993. **22**: 73-83.
50. Oumada, F.Z., M. Roses, E. Bosch, and M.H. Abraham, *Solute-solvent interactions in normal-phase liquid chromatography: a linear free-energy relationships study.* Anal. Chim. Acta, 1999. **382**: 301-308.
51. Ghose, A.K. and G.M. Crippen, *Atomic Physiochemical Parameters For 3-Dimensional-Structure-Directed Quantitative Structure-Activity-Relationships .2. Modeling Dispersive And Hydrophobic Interactions.* J. Chem. Inf. Comput. Sci., 1987. **27**: 21-35.
52. Abraham, M.H. and G.S. Whiting, *Hydrogen Bonding XVI. A new solute solvation parameter, pi2H, from gas chromatographic data.* J.of Chromatography, 1991. **587**: 213-228.
53. Abraham, M.H., P.L. Grellier, D.V. Prior, P.P. Duce, J.J. Morris, and P.J. Taylor, *Hydrogen-Bonding .7. A Scale of Solute Hydrogen-Bond Acidity Based on Log K-Values for Complexation in Tetrachloromethane.* J.Chem.Soc.-Perkin Trans. 2, 1989: 699-711.
54. Abraham, M.H., P.L. Grellier, D.V. Prior, J.J. Morris, and P.J. Taylor, *Hydrogen-Bonding.10. A Scale of Solute Hydrogen-Bond Basicity Using Log K Values for Complexation in Tetrachloromethane.* J.Chem.Soc.-Perkin Trans. 2, 1990: 521-529.
55. Abraham, M.H. and J.C. McGowan, *The use of characteristic volumes to mesaure cavity terms in reversed phase liquid chromatography.* Chromatographia, 1987. **23**: 243-246.
56. Acree, J., William E. and M.H. Abraham, *Solubility predictions for crystalline polycyclic aromatic hydrocarbons (PAHs) dissolved in organic solvents based upon the Abraham general solvation model.* Fluid Phase Equilib., 2002. **201**: 245-258.
57. Platts, J.A., D. Butina, M.H. Abraham, and A. Hersey, *Estimation of molecular linear free energy relation descriptors using a group contribution approach.* J. Chem. Inf. Comput. Sci., 1999. **39**: 835-845.

58. Lamarche, O., J.A. Platts, and A. Hersey, *Theoretical prediction of the polarity/polarizability parameter piH2.* Phys.Chem.Chem.Phys., 2001. **3**: 2747-2753.
59. Geladi, P. and B.R. Kowalski, *Partial Least-Squares: A tutorial.* Analyt.Chim.Acta, 1986. **185**: 1-17.
60. Platts, J.A., *Theoretical Prediction of Hydrogen Bond Donor Capacity.* Phys.Chem.Chem.Phys., 2000. **2**: 973-980.
61. Platts, J.A., *Theoretical Prediction of Hydrogen Bond Basicity.* Phys.Chem.Chem.Phys., 2000. **2**: 3115-3120.
62. Laurence, C., J. Graton, M. Berthelot, F. Besseau, J.Y. Le Questel, M. Lucon, C. Ouvrard, A. Planchat, and E. Renault, *An Enthalpic Scale of Hydrogen-Bond Basicity. 4. Carbon pi Bases, Oxygen Bases, and Miscellaneous Second-Row, Third-Row, and Fourth-Row Bases and a Survey of the 4-Fluorophenol Affinity Scale.* J. Org. Chem., 2010. **75**(12): 4105-4123.
63. Taft, R.W., D. Gurka, L. Joris, P.V. Schleyer, and J.W. Rakshys, *Studies of Hydrogen-Bonded Complex Formation with P-Fluorophenol .V. Linear Free Energy Relationships with Oh Reference Acids.* J. Am. Chem. Soc., 1969. **91**: p4801-&.
64. Joris, L., J. Mitsky, and R.W. Taft, *EFFECTS OF POLAR APROTIC SOLVENTS ON LINEAR FREE-ENERGY RELATIONSHIPS IN HYDROGEN-BONDED COMPLEX FORMATION.* J. Am. Chem. Soc., 1972. **94**: 3438-&.
65. Mitsky, J., R.W. Taft, and L. Joris, *Hydrogen-bonded complex formation with 5-fluoroindole - applications of PkHB scale.* J. Am. Chem. Soc., 1972. **94**: 3442-&.
66. Kamlet, M.J. and R.W. Taft, *Solvatochromic comparison method .1. beta-scale of solvent hydrogen-bond acceptor (HBA) basicities.* J. Am. Chem. Soc., 1976. **98**: 377-383.
67. Ahmed, F., *A Good Example Of The Franck-Condon Principle.* J. Chem. Educ., 1987. **64**: 427-428.
68. Abraham, M.H., P.P. Duce, D.V. Prior, D.G. Barratt, J.J. Morris, and P.J. Taylor, *Hydrogen Bonding. Part 9.t Solute Proton Donor and Proton Acceptor Scales for Use in Drug Design.* J.Chem.Soc. Perkin II, 1989: 1355-1376.
69. Abraham, M.H., *Hydrogen-bonding .31. construction of a scaleof solute effective or summation hydrogen-bond basicity.* J. Phys. Org. Chem., 1993. **6**: 660-684.
70. Abraham, M.H., A. Ibrahim, A.M. Zissimos, Y.H. Zhao, J. Comer, and D.P. Reynolds, *Application of hydrogen bonding calculations in property based drug design.* Drug Discov. Today, 2002. **7**: 1056-1063.
71. Berthelot, M., M. Helbert, C. Laurence, and J.Y. Lequestel, *Hydrogen-Bond Basicity Of Nitriles.* J. Phys. Org. Chem., 1993. **6**: 302-306.

72. Berthelot, M., C. Laurence, M. Safar, and F. Besseau, *Hydrogen-bond basicity pK(HB) scale of six-membered aromatic N-heterocycles.* J.Chem.Soc. Perkin II, 1998: 283-290.
73. Besseau, F., C. Laurence, and M. Berthelot, *Hydrogen-Bond Basicity Of Esters, Lactones And Carbonates.* J.Chem.Soc. Perkin II, 1994: 485-489.
74. Besseau, F., M. Lucon, C. Laurence, and M. Berthelot, *Hydrogen-bond basicity pK(HB) scale of aldehydes and ketones.* J.Chem.Soc. Perkin II, 1998: 101-107.
75. Graton, J., M. Berthelot, and C. Laurence, *Hydrogen-bond basicity pK(HB) scale of secondary amines.* J.Chem.Soc. Perkin II, 2001: 2130-2135.
76. Graton, J., F. Besseau, M. Berthelot, E. Raczynska, and C. Laurence, *The pK(HB) scale for hydrogen bonds in aliphatic tertiary amines.* Can. J. Chem., 2002. **80**: 1375-1385.
77. Graton, J., C. Laurence, M. Berthelot, J.Y. Le Questel, F. Besseau, and E.D. Raczynska, *Hydrogen-bond basicity pK(HB) scale of aliphatic primary amines.* J.Chem.Soc. Perkin II, 1999: 997-1001.
78. Laurence, C., M. Berthelot, M. Helbert, and K. Sraidi, *1st Measurement Of The Hydrogen-Bond Basicity Of Monomeric Water, Phenols, And Weakly Basic Alcohols.* J. Phys. Chem., 1989. **93**: 3799-3802.
79. Laurence, C., M. Berthelot, J.Y. Lequestel, and M.J. Elghomari, *Hydrogen-Bond Basicity Of Thioamides And Thioureas.* J.Chem.Soc. Perkin II, 1995: 2075-2079.
80. Laurence, C., M. Berthelot, M. Lucon, and D.G. Morris, *Hydrogen-Bond Basicity Of Nitrocompounds.* J.Chem.Soc. Perkin II, 1994: 491-493.
81. Laurence, C., M. Berthelot, E. Raczynska, J.Y. Lequestel, G. Duguay, and P. Hudhomme, *Hydrogen-Bond Basicity Of Cyanamide, Amide, Thioamide, And Sulfonamide Iminologues.* J. Chem. Res., Synop., 1990: 250-251.
82. Raczynska, E.D. and C. Laurence, *Hydrogen-Bonding Basicity Of Aetamidines And Benzamidines.* J. Chem. Res., Synop., 1989: 148-149.
83. Raczynska, E.D., C. Laurence, and P. Nicolet, *HYDROGEN-BONDING BASICITY OF AMIDINES.* J.Chem.Soc. Perkin II, 1988: 1491-1494.
84. Banerjee, A.K.a.M., U., *Extracting the significant descriptors by 2D QSAR and docking efficiency of NRTI drugs: A Molecular Modeling Approach .* Internet J. Genomics Proteomics., 2007. **2**(2).
85. Devereux, M. and P.L.A. Popelier, *In Silico Techniques for the Identification of Bioisosteric Replacements for Drug Design.* Curr. Top. Med Chem., 2010. **10**: 657-668.

86. Ertl, P., *World Wide Web-based system for the calculation of substituent parameters and substituent similarity searches.* J.Mol.Graphics Modell., 1998. **16**: 11-13.
87. Murray, J.S., S. Ranganathan, and P. Politzer, *Correlations between the Solvent Hydrogen-Bond Acceptor Parameter Beta and the Calculated Molecular Electrostatic Potential.* J.Org.Chem., 1991. **56**: 3734-3737.
88. Kamlet, M.J., J.L.M. Abboud, M.H. Abraham, and R.W. Taft, *Linear solvation energy relationships .23. A Comprehensive collection of the solvatochromic parameters, Pi-star Alpha and Beta, and some methods for simplifying the general solvatochromatic equation.* J. Org. Chem., 1983. **48**: 2877-2887.
89. Hagelin, H., J.S. Murray, T. Brinck, M. Berthelot, and P. Politzer, *Family-Independent Relationships between Computed Molecular-Surface Quantities and Solute Hydrogen-Bond Acidity Basicity and Solute-Induced Methanol O-H Infrared Frequency-Shifts.* Can.J.Chem., 1995. **73**: 483-488.
90. Lamarche, O. and J.A. Platts, *Atoms in molecules investigation of the pKHB basicity scale: electrostatic and covalent effect in hydrogen bonding.* Chem.Phys.Lett., 2003. **367**: 123-128.
91. Devereux, M., P.L.A. Popelier, and I.M. McLay, *A refined model for prediction of hydrogen bond acidity and basicity parameters from quantum chemical molecular descriptors.* Phys.Chem.Chem.Phys., 2009. **11**: 1595–1603.
92. Pearson, R.G., *Chemical Hardness*1997, John Wiley-VCH, Weinheim, Germany.
93. Pearson, R.G., *Hard And Soft Acids And Bases.* J. Am. Chem. Soc., 1963. **85**: 3533-&.
94. Pearson, R.G., *Acids And Bases.* Science, 1966. **151**: 172-&.
95. Pearson, R.G. and J. Songstad, *Application Of Principle Of Hard And Soft Acids And Bases To Organic Chemistry.* J. Am. Chem. Soc., 1967. **89**: 1827-&.
96. Pearson, R.G., *Recent advances in the concept of hardness and softness.* J.Chem.Educ., 1987. **64**: 561-568.
97. Pearson, R.G., *Chemical hardness and density functional theory.* J. Chem. Sci., 2005. **117**: 369-377.
98. Zadeh, L.A., *Fuzzy Sets.* Information and Control, 1965. **8**: 338-&.
99. Geerlings, P. and F. De Proft, *Conceptual DFT: the chemical relevance of higher response functions.* Phys. Chem. Chem. Phys., 2008. **10**: 3028-3042.
100. Geerlings , P., F. DeProft, and W. Langenaeker, *Conceptual Density Functional Theory.* Chem.Rev., 2003. **103**: 1793-873.
101. Liu, S.B., *Conceptual Density Functional Theory and Some Recent Developments.* Acta Phys. Chim. Sin., 2009. **25**: 590-600.

102. Parr, R.G. and R.G. Pearson, *Absolute Hardness - Companion Parameter to Absolute Electronegativity.* J. Am. Chem. Soc., 1983. **105**: 7512-7516.
103. Parr, R.G., R.A. Donnelly, M. Levy, and W.E. Palke, *Electronegativity - Density Functional Viewpoint.* J. Chem. Phys., 1978. **68**:3801-3807.
104. Ayers, P.W., *The dependence on and continuity of the energy and other molecular properties with respect to the number of electrons.* J. Math. Chem., 2008. **43**: 285-303.
105. Perdew, J.P., R.G. Parr, M. Levy, and J.L. Balduz, *Density-Functional Theory For Fractional Particle Number - Derivative Discontinuities Of The Energy.* Phys. Rev. Lett., 1982. **49**: 1691-1694.
106. Iczkowski, R. and J.L. Margrave, *Electronegativity*. J. Am. Chem. Soc., 1961. **83**(: 3547-&.
107. De Luca, G., E. Sicilia, N. Russo, and T. Mineva, *On the hardness evaluation in solvent for neutral and charged systems.* J. Am. Chem. Soc., 2002. **124**: 1494-1499.
108. Gilman, J.J., *Chemical and physical hardness.* Mat Res Innovat, 1997. **1**: 71-76.
109. Sanderson, R.T., *Chemical Bonds and Bond Energy*1976, NY, USA: Academic Press.
110. Mortier, W.J., S.K. Ghosh, and S. Shankar, *Electronegativity Equalization Method For The Calculation Of Atomic Charges In Molecules.* J. Am. Chem. Soc., 1986. **108**: 4315-4320.
111. Parr, R.G. and W.T. Yang, *Density Functional-Approach to the Frontier-Electron Theory of Chemical-Reactivity.* J. Am. Chem. Soc., 1984. **106**: 4049-4050.
112. Yang, W., R.G. Parr, and R. Pucci, *Electron-Density, Kohn-Sham Frontier Orbitals, And Fukui Functions.* J. Chem. Phys., 1984. **81**: 2862-2863.
113. Torrent-Sucarrat, M., F. De Proft, P. Geerlings, and P.W. Ayers, *Do the Local Softness and Hardness Indicate the Softest and Hardest Regions of a Molecule?* Chem. Eur. J., 2008. **14**: 8652-8660.
114. Popelier, P.L.A., *Atoms in Molecules. An Introduction.*2000, London, Great Britain: Pearson Education.
115. Bader, R.F.W., *Atoms in Molecules. A Quantum Theory.*1990, Oxford, Great Britain: Oxford Univ. Press.
116. Cioslowski, J., M. Martinov, and S.T. Mixon, *Atomic Fukui Indexes From The Topological Theory Of Atoms In Molecules Applied To Hartree-Fock And Correlated Electron-Densities.* J. Phys. Chem., 1993. **97**: 10948-10951.

117. Fradera, X. and M. Sola, *Second-order atomic Fukui indices from the electron-pair density in the framework of the atoms in molecules theory.* J. Comput. Chem., 2004. **25**: 439-446.
118. Fitzgerald, G., *On the use of fractional charges for computing Fukui functions.* Mol. Simul., 2008. **34**: 931–936.
119. Gilli, P., L. Pretto, V. Bertolasi, and G. Gilli, *Predicting Hydrogen-Bond Strengths from Acid-Base Molecular Properties. The pK(a) Slide Rule: Toward the Solution of a Long-Lasting Problem.* Acc. Chem. Res., 2009. **42**: 33-44.
120. Matta, C.F. and R.J. Boyd, *The Quantum Theory of Atoms in Molecules. From Solid State to DNA and Drug Design.*2007, Weinheim, Germany: Wiley-VCH.
121. Matta, C.F. and R.F.W. Bader, *An experimentalist's reply to "What is an atom in a molecule?".* J. Phys. Chem. A, 2006. **110**: 6365-6371.
122. Harding, A.P., D.C. Wedge, and P.L.A. Popelier, *pKa Prediction from "Quantum Chemical Topology" Descriptors.* J.Chem.Inf.Mod., 2009. **49**: 1914–1924.
123. Adam, K.R., *New Density Functional and Atoms in Molecules Method of Computing Relative pKa Values in Solution.* J. Phys.Chem.A, 2002. **106**: 11963-11972.
124. Wiberg, K.B., R.F.W. Bader, and C.D.H. Lau, *Theoretical-Analysis of Hydrocarbon Properties .2. Additivity of Group Properties and the Origin of Strain-Energy.* J. Am. Chem. Soc., 1987. **109**: 1001-1012.
125. Matta, C.F. and R.F.W. Bader, *Atoms-in-Molecules Study of the Genetically Encoded Amino Acids. III. Bond and Atomic Properties and their Correlations with Experiment including Mutation-Induced Changes in Protein Stability and Genetic Coding.* Prot.Struct.Func.Gen., 2003. **52**(3): 360-399.
126. Popelier, P.L.A. and F.M. Aicken, *Atomic Properties of selected Biomolecules : Quantum Topological Atom Types of Carbon occurring in natural Amino Acids and derived Molecules.* J.Am.Chem.Soc., 2003. **125**: 1284-1292.
127. Eberhart, M.E., D.P. Clougherty, and J.M. McLaren, *Bonding-property relationships in intermetallic alloys.* J.Mater.Res., 1993. **8**: 438-448.
128. Eberhart, M.E., *Why Things Break.* Scientific American, 1999. **281**: 66-74.
129. Schrodinger, E., *Quantisation as an eigen value problem.* Ann. Phys., 1926. **79**: 361-U8.
130. Nadas, J., S. Vukovic, and B.P. Hay, *Alkyl chlorides as hydrogen bond acceptors.* Comp. Theor. Chem., 2012. **988**: 75-80.
131. Lequestel, J.Y., C. Laurence, A. Lachkar, M. Helbert, and M. Berthelot, *Hydrogen-Bond Basicity Of Secondary And Tertiary Amides, Carbamates, Ureas And Lactams.* J.Chem.Soc. Perkin II, 1992: 2091-2094.

132. Raczynska, E.D., C. Laurence, and M. Berthelot, *Basicity Of The Hydrogen-Bonds Of Formamides Substituted On The Imino Nitrogen Atom.* Can. J. Chem., 1992. **70**: 2203-2208.
133. Chardin, A., C. Laurence, M. Berthelot, and D.G. Morris, *Hydrogen-bond basicity of the sulfonyl group. The case of strongly basic sulfonamidates RSO(2)N(-)N(+)Me(3).* J.Chem.Soc. Perkin II, 1996: 1047-1051.
134. Besseau, F., C. Laurence, and M. Berthelot, *The pK(HB) scale of pi bases.* Bull. Soc. Chim. Fr., 1996. **133**: 381-387.
135. Chardin, A., C. Laurence, M. Berthelot, and D.G. Morris, *The pK(HB) scale of dimethylnitramine, nitramine vinylogues and nitramidates.* Bull. Soc. Chim. Fr., 1996. **133**: 389-393.
136. Ouvrard, C., M. Berthelot, and C. Laurence, *The first basicity scale of fluoro-, chloro-, bromo- and iodo-alkanes: some cross-comparisons with simple alkyl derivatives of other elements.* J.Chem.Soc. Perkin II, 1999: 1357-1362.
137. Berthelot, M., F. Besseau, and C. Laurence, *The hydrogen-bond basicity pK(HB) scale of peroxides and ethers.* European J. Org. Chem., 1998: 925-931.
138. Marquis, E., J. Graton, M. Berthelot, A. Planchat, and C. Laurence, *Hydrogen body of arylamines - Competition with pi and N sites.* Can. J. Chem., 2004. **82**: 1413-1422.
139. Abraham, M.H. and Y.H. Zhao, *Determination of Solvation Descriptors for Ionic Species: Hydrogen Bond Acidity and Basicity.* J.Org.Chem., 2004. **69**: 4677-4685.
140. Alkorta, I., N. Campillo, I. Rozas, and J. Elguero, *Ring strain and hydrogen bond acidity.* J. Org. Chem., 1998. **63**: 7759-7763.
141. Lamarche, O. and J.A. Platts, *Complementary Nature of Hydrogen Bond Basicity and Acidity Scales from electrostatic and Atoms in Molecules Properties.* Phys.Chem.Chem.Phys., 2003. **5**: 677-684.
142. GAUSSIAN03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A. J.; Vreven, J., T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A.

D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. In Gaussian, Inc., Pittsburgh PA, 2003.

143. Keith, T.A., *AIMAll (Version 10.07.25)*, 2010.
144. Popelier, P., L.A. and R.G.A. Program MORPHY with a contribution from Bone, *MORPHY98*, 1998, UMIST: Manchester, UK.
145. Tuma, C., A.D. Boese, and N.C. Handy, *Predicting the binding energies of H-bonded complexes: A comparative DFT study.* Phys. Chem. Chem. Phys., 1999. **1**: 3939-3947.
146. Boys, S.F. and F. Bernardi, *The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors.* Mol. Phys., 1970. **19**: 553-566.
147. Zhao, Y. and D.G. Truhlar, *Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: The MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions.* J. Phys. Chem. A, 2004. **108**: 6908-6918.
148. Zhao, Y. and D.G. Truhlar, *Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions.* J. Phys. Chem. A, 2005. **109**: 5656-5667.
149. Zhao, Y. and D.G. Truhlar, *Benchmark databases for nonbonded interactions and their use to test density functional theory.* J. Chem. Theory Comput., 2005. **1**: 415-432.
150. Livingstone, N.J., *Data Analysis for Chemists, 1st ed,.* Oxford Univ. Press, Oxford,, 1995.
151. UMETRICS, *SIMCA-P 10.0*, 2002, info@umetrics.com: www.umetrics.com, Umeå, Sweden.
152. MORPHY98, a program written by P.L.A. Popelier with a contribution from R.G.A. Bone, UMIST, Manchester, England, EU (1998).
153. Lamarche, O. and J.A. Platts, *Theoretical prediction of the hydrogen-bond basicity pK(HB).* Chem.Eur.J., 2002. **8**: 457-466.
154. Wang, L.R., Z.J. Xie, P. Wipf, and X.Q. Xie, *Residue Preference Mapping of Ligand Fragments in the Protein Data Bank.* J. Chem. Inf. Model., 2011. **51**: 807-815.

155. Devereux, M. and P.L.A. Popelier, *In Silico Techniques for the Identification of Bioisosteric Replacements for Drug Design.* Curr. Top. Med Chem., 2008. **10**: 657-668.
156. Devereux, M., P.L.A. Popelier, and I.M. McLay, *Toward an Ab Initio Fragment Database for Bioisosterism: Dependence of QCT Properties on Level of Theory, Conformation, and Chemical Environment.* J. Comput. Chem., 2009. **30**: 1300-1318.
157. Ren, J., J. Li, F. Shi, X. Wang, J. He, Y. Xu, N. Zhang, B. Xiong, J. Shen, æ□Žå□¥, çŸ³å³°, çŽ‹æ˜•, ä½•å»ºå□Ž, è®¸å□¶æ˜¥, å¼ ä¹ƒéœž, ç†Šå…µ, and æ²ˆç«žå°·, *Progress in the fragment-based drug discovery.* Acta Pharm. Sin., 2013. **48**: 14-24.
158. Murray, C.W., M.L. Verdonk, and D.C. Rees, *Experiences in fragment-based drug discovery.* Trends Pharmacol. Sci., 2012. **33**: 224-232.
159. Stamford, A. and C. Strickland, *Inhibitors of BACE for treating Alzheimer's disease: a fragment-based drug discovery story.* Curr. Opin. Chem. Biol., 2013. **17**: 320-328.
160. Sheng, C.Q. and W.N. Zhang, *Fragment Informatics and Computational Fragment-Based Drug Design: An Overview and Update.* Med. Res. Rev., 2013. **33**: 554-598.
161. Kumar, A., A. Voet, and K.Y.J. Zhang, *Fragment Based Drug Design: From Experimental to Computational Approaches.* Curr. Med. Chem., 2012. **19**: 5128-5147.
162. Mortier, J., C. Rakers, R. Frederick, and G. Wolber, *Computational Tools for In Silico Fragment-Based Drug Design.* Curr. Top. Med Chem., 2012. **12**: 1935-1943.
163. Speck-Planche, A., F. Luan, and M. Cordeiro, *Role of Ligand-Based Drug Design Methodologies toward the Discovery of New Anti-Alzheimer Agents: Futures Perspectives in Fragment-Based Ligand Design.* Curr. Med. Chem., 2012. **19**: 1635-1645.
164. DesJarlais, R.L., *Using Computational TechniquesS In Fragment-Based Drug Discovery*, in *Fragment-Based Drug Design: Tools, Practical Approaches, and Examples*2011, Elsevier Academic Press Inc: San Diego. 137-155.
165. Lanter, J., X.Q. Zhang, and Z.H. Sui, *Medicinal Chemisry Inspired Fragment-Based drug Discovery*, in *Fragment-Based Drug Design: Tools, Practical Approaches, and Examples*2011, Elsevier Academic Press Inc: San Diego. 421-445.

166. Hubbard, R.E. and J.B. Murray, *Experiences In Fragment-Based Lead Discovery*, in *Fragment-Based Drug Design: Tools, Practical Approaches, and Examples*2011, Elsevier Academic Press Inc: San Diego. 509-531.
167. Konteatis, Z.D., *In silico fragment-based drug design.* Expert Opin. Drug Discov., 2010. **5**: 1047-1065.
168. Murray, C.W. and T.L. Blundell, *Structural biology in fragment-based drug design.* Curr. Opin. Struct. Biol., 2010. **20**: 497-507.
169. Salum, L.B. and A.D. Andricopulo, *Fragment-based QSAR strategies in drug design.* Expert Opin. Drug Discov., 2010. **5**: 405-412.
170. Tanaka, D., *Fragment-based Drug Discovery: Concept and Aim.* Yakugaku Zasshi, 2010. **130**: 315-323.
171. Takahashi, O., Y. Masuda, A. Muroya, and T. Furuya, *In Silico Approaches for Fragment-based Drug Design.* Yakugaku Zasshi, 2010. **130**: 349-354.
172. Loving, K., I. Alberts, and W. Sherman, *Computational Approaches for Fragment-Based and De Novo Design.* Curr. Top. Med Chem., 2010. **10**: 14-32.
173. Hajduk, P.J. and J. Greer, *A decade of fragment-based drug design: strategic advances and lessons learned.* Nat. Rev. Drug Discovery, 2007. **6**: 211-219.
174. Clarke, W.P. and R.A. Bond, *The elusive nature of intrinsic efficacy.* Trends Pharmacol. Sci., 1998. **19**: 270-276.
175. Jones, G., P. Willett, R.C. Glen, and A.R. Leach, *Development and Validation of a Genetic Algorithm for Flexible Docking.* J. Mol. Biol., 1997. **267**: 727-748.
176. Hu, L., M.L. Benson, R.D. Smith, M.G. Lerner, and H.A. Carlson, *Binding MOAD (Mother Of All Databases).* Proteins: Struct., Funct., Bioinf., 2005. **60**: 333-340.
177. Bernstein, F.C., T.F. Koetzle, G.J.B. Williams, E.F. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, *Protein Data Bank - Computer-Based Archival File For Macromoleculr Structures.* J. Mol. Biol., 1977. **112**: 535-542.
178. Mishra, S.D., D. Bose, S.K. Shukla, A. Durgabanshi, and J. Esteve-Romero, *Monitoring strychnine and brucine in biochemical samples using direct injection micellar liquid chromatography.* Anal. Methods, 2013. **5**: 1747-1754.
179. Laurence, C. and M. Berthelot, *Observations on the strength of hydrogen bonding.* Perspect. Drug Discov. Des., 2000. **18**: 39-60.
180. Adcock, W., J. Graton, C. Laurence, M. Lucon, and N. Trout, *Three-centre hydrogen bonding in the complexes of syn-2,4-difluoroadamantane with 4-fluorophenol and hydrogen fluoride.* J. Phys. Org. Chem., 2005. **18**: 227-234.