Where does charge reside in amino acids? The effect of side-chain protonation state on the atomic charges of Asp, Glu, Lys, His and Arg

ARTICLE *in* COMPUTATIONAL AND THEORETICAL CHEMISTRY · AUGUST 2014

Impact Factor: 1.55 · DOI: 10.1016/j.comptc.2014.07.020

2 AUTHORS:

Timothy J Hughes
The University of Manchester
4 PUBLICATIONS **26** CITATIONS

SEE PROFILE

Paul L A Popelier
The University of Manchester
190 PUBLICATIONS **7,101** CITATIONS

SEE PROFILE

Computational and Theoretical Chemistry 1053 (2015) 298–304

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

Where does charge reside in amino acids? The effect of side-chain protonation state on the atomic charges of Asp, Glu, Lys, His and Arg

Timothy J. Hughes, Paul L.A. Popelier *

Manchester Institute of Biotechnology (MIB), 131 Princess Street, Manchester M1 7DN, United Kingdom
School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

ARTICLE INFO

Article history:
Received 3 July 2014
Received in revised form 31 July 2014
Accepted 31 July 2014
Available online 8 August 2014

Keywords:
Quantum Theory of Atoms in Molecules (QTAIM)
Amino acids
Quantum Chemical Topology (QCT)
Atomic charge
Protonation

ABSTRACT

Quantum topological atomic charges have been calculated at B3LYP/apc-1 level to identify where the charge is located on amino acid residues when the side-chain has been either protonated (Arg, Lys, His) or deprotonated (Glu, Asp). All energy local energy minima in the Ramachandran map of each (neutral) amino acid were populated with a number of distorted molecular geometries, summing up to a thousand geometries for each amino acid. The majority of the molecular charge is found on the side-chain (81–100%), with a large percentage of the charge located on the functional group undergoing protonation/deprotonation. Each side-chain (or residue) methylene group was found to act as an insulator between an amino acid's backbone and its side-chain because it accepts the majority of charge not located on the side-chain. As a result there is no significant charge on backbone atoms relative to the neutral molecule. In the case of His^+ and Arg^+, where the charge is spread over a large number of atoms due to resonance, the influence of the positive charge on the backbone atoms is reduced.

1. Introduction

The complex mechanisms of enzymatic catalysis have been studied intensively for decades. A common feature in these mechanisms is the protonation and deprotonation of the active site amino acid side-chains involved in the catalysis. For example, the rate limiting step in the conversion of $CO_2 + H_2O \rightarrow HCO_3^- + H^+$ by the enzyme carbonic anhydrase is a proton transfer involving the residue His64 [1,2]. Similarly, a proton transfer mechanism involving a Glu residue in the active site is found to be the rate-determining step in the mechanism of the enzyme glutaminylcyclase [3]. The subtle changes in the electronic charge of the active site atoms of glutaminylcyclase play a role in determining the path that the reaction follows. This effect arises through strengthening of hydrogen bonds within the active site upon proton transfer. The mechanism employed by enzyme horseradish peroxidase includes a nucleophilic attack by the hydroxyl oxygen of Ser195. However, this step requires activation through the deprotonation of the hydroxyl group [4]. Deprotonation results in the charge of the oxygen atom becoming more negative and hence more nucleophilic.

The above examples show that when developing a computational model to describe enzymatic reactions, any changes in electronic structure must be captured. Early potentials that enabled the modelling of reactions include the empirical valence bond approach [5] and the "ReaxFF" force field [6]. The popularity of QM/MM approaches is increasing in the study of such systems due to increases in computer power [7]. Currently under development in our lab is the quantum chemical topological force field (QCTFF). This is a novel approach to building a molecular mechanics force field, in which machine learning is used to map quantum mechanical properties (such as atomic multipole moments [8–10], kinetic energy [11] and exchange-repulsion) directly to the coordinates of the system. Preliminary work has shown that this methodology enables the modelling of changes in atomic charge as a reaction path is followed.

There is a perhaps surprising lack of literature detailing the changes in the atomic charges of amino acids upon a change of the side-chain protonation state, with studies [12–15] typically focusing on the zwitterionic states of amino acids. To address this gap in the literature, a thousand geometries for each of a total of five amino acids that most commonly undergo changes in protonation state (Asp, Glu, His, Lys, Arg) have been sampled for both the protonated and deprotonated state, and the changes in average atomic charges have been compared. In this work, charges have been obtained from the Quantum Theory of Atoms in Molecules

* Corresponding author. Tel.: +44 161 3064511; fax: +44 161306 4559.
E-mail address: pla@manchester.ac.uk (P.L.A. Popelier).

http://dx.doi.org/10.1016/j.comptc.2014.07.020
2210-271X/

(QTAIM) [16–18]. The extensive QTAIM work [19–21] of Matta and Bader on all natural amino acids, provides a rich background to the current work but does not specifically address the question of where an excess or depletion of a formal unit charge resides compared to the neutral amino acid.

There are many methods of obtaining atomic charges but the question of which protocol produces the "best" atomic charges is contentious. Arguments for and against the different charge methods typically fall into one of two competing schools of thought. The first supports a belief that the atomic charge should be capable of reproducing the electrostatic potential around an atom. The second asserts that the charge should correctly describe the charge transfer within a molecule. We subscribe to the latter assertion. Indeed, an atomic charge should do no more than what the name says: describe, and indeed correctly represent, the charge on an atom. Demanding that the single number that is the atomic charge also reproduces the electrostatic potential generated by the atom, is problematic because it ignores the non-spherical features of an atom in a molecule, which are prominent at short (and even medium) range [22]. Atomic charges that are fitted to best reproduce an electrostatic potential are just numbers [23], one of many possible solutions, and by no means guaranteed to capture true charge transfer effects. QTAIM charges fall into the second category of charges: they do reproduce well the charge transfer in a molecule, even if they have been criticised for being "unrealistically" high [24]. At the other end of the spectrum, the Hirshfeld population analysis produces very small charges, which ironically become more QTAIM-like, when corrected for the arbitrariness of the promolecule, as done in Hirsheld-I [25]. The criticism that QTAIM charges do not reproduce the electrostatic potential is remedied by performing a multipolar expansion (of which the QTAIM charge is the first term of the expansion, the monopole moment) where it was shown [26] a long time ago that reproduction of the *ab initio* electrostatic potential was achieved at a so-called interaction rank of $L = 5$. To quote from this work, "*This work makes clear that the atomic population (or rank zero multipole moment) is just one term of the expansion of a physically observable quantity, namely the electrostatic potential. Hence, QTAIM populations (and thus charges) cannot be judged on their reproduction of the electrostatic potential. Instead, they must be seen in the context of a multipolar expansion of the exact electrostatic potential of a topological atom*".

2. Background and computational details

2.1. Topological atoms

The electron density of a system partitions itself, without the need for setting any parameter values, or calculation through an iterative procedure. The only concept required is that of the gradient vector of the electron density, denoted $\nabla\rho$, which traces paths of steepest ascent. The vast majority of these so-called gradient paths terminate at a nuclear position, thereby carving out one subspace for each nucleus. These subspaces are identified with topological atoms. More details can be found in a very recent, didactic, historic and refreshing account [27] of QTAIM. The central idea of partitioning a quantum system by means of gradient paths was first carried out [16] on ρ by the group of Bader, and later [28,29] on the Laplacian of ρ. Meanwhile several other three-dimensional quantum property density functions were investigated (for a list of examples see Box 8.1 in [27]) justifying [30] the overarching name Quantum Chemical Topology (QCT) [31].

Fig. 1 shows an example, relevant to the current work, of a protonated lysine falling apart into topological atoms, as generated [32,33] by in-house software. The latter can be seen as bubbles, touching each other without overlapping or leaving gaps; they are malleable boxes that change their shape in response to a change in the nuclear skeleton.

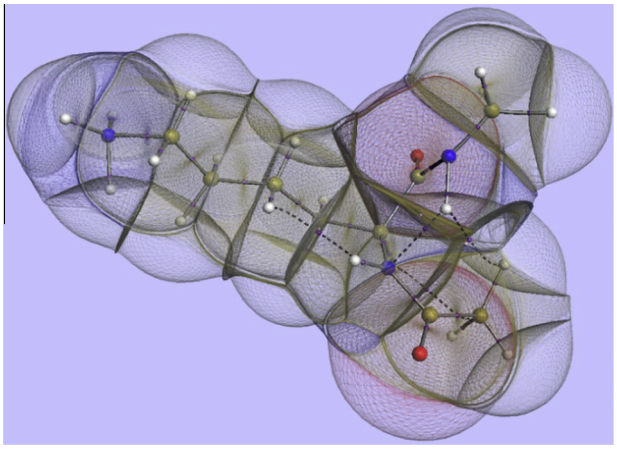

Fig. 1. Finite-element representation of a molecular geometry of protonated lysine.

2.2. Geometry generation

Each amino acid was capped by a [$CH_3C{=}O$] group at the N-terminus, and by a [$NHCH_3$] group at the C-terminus to create the so-called "dipeptide". The minimum energy geometries for each neutral amino acid were obtained through a comprehensive search of the potential energy surface [34]. The number of energetic minima for each amino acid is given in Table 1.

A thousand geometries for each amino acid were obtained by distributing energy into the normal vibrational modes of each local energy minimum for each neutral amino acid. These geometries were generated by the in-house computer program EROS. Quasi-random amounts of energy are put into the normal vibrational modes, which then spreads evenly over all vibrational modes. The motion of the vibrational modes is modelled harmonically, and 'snapshots' of the nuclear coordinates at random points during the vibrational motion formed the set of geometries. Bonds may dissociate and atoms fly apart if the input energy is too high, which is corrected by an iterative process where initially a maximum input energy is defined, typically about 200 kJ mol^{-1} and geometries are generated. If broken bonds are present then the maximum energy is lowered and the process is repeated until no broken bonds are present. More details can be found in Ref. [35].

All charged amino acid residues except Arg were obtained by direct addition or removal of a proton on the side-chain of the distorted geometries. For each of the thousand sampled neutral Asp and Glu residues, the acidic proton was removed in order to obtain the geometries of the Asp^- and Glu^-, respectively. A similar approach was also taken in the case of Lys^+, where a proton was added to the primary amine to give the positively charged tetrahedral ammonium group. His^+ was similarly obtained by protonating the lone pair position of N29 to give the positively charged imidazolium group. Due to the more complex structural changes that take place in Arg upon protonation a different approach was taken

Table 1
Number of local energy minima for each amino acid studied in this work.

Amino acid	No. minima
Asp	36
Glu	36
His	24
Lys	39
Arg	61

in obtaining the Arg^+ geometries. In particular, the neutral Arg has a guanidine system with two pyramidal nitrogens (N19 and N16) and one planar nitrogen (N34). However, in Arg^+ this group formally becomes a guanidinium group, which has three planar nitrogens. The addition of a proton to N34 causes the geometrical change between guanidine and guanidinium. Therefore, an alternative approach was taken; a proton was added to each of the minimum energy geometries and then the guanidinium group alone ($[-NH-C(NH_2)_2]^+$) was allowed to relax by partial geometry optimisation. These new "minima" were then input to EROS to sample the thousand distorted Arg^+ geometries.

2.3. Computational details

Normal modes sampling was performed by the in-house FORTRAN code EROS. All *ab initio* calculations were performed by GAUSSIAN09 [36] at the B3LYP/apc-1 [37] level of theory, taking advantage of a basis set with polarisation and diffuse functions optimised for use with density functionals. QTAIM charges for all atoms were calculated with the program AIMAll [38], and are listed in the Supporting Information, as averages over all configurations (corresponding to all local energy minima), along with ranges and standard deviations. An atomic charge is an atomic property the least sensitive to integration error [39].

3. Results and discussion

Numbered geometries for all five protonated capped amino acids (Asp, Glu, Lys^+, His^+ and Arg^+) are provided in Fig. 2. For convenience, both the protonated and deprotonated amino acids share a common numbering system. The following discussion refers to the amino acids as consisting of both side-chain atoms and backbone atoms. The set of side-chain atoms consist of all atoms starting with C_β (including its methylene hydrogens), whereas the backbone corresponds to the C_α, the two peptide groups and the methyl caps, as well as all associated hydrogen atoms.

Tables containing the average value, range and standard deviation of all atomic charges for both the protonated and deprotonated systems studied in this work is provided as Supporting Information. With the exception of Arg/Arg^+ (due to the different method of obtaining the structures), patterns in both the standard deviation and the range of atomic charge for similar atom types are observed.

A number of general observations can be made, across the various systems. For an atom in a neutral amino acid, both the range of the atomic charge and its standard deviation maintain a similar value in the charged amino acid. Within an amino acid, there is also no clear distinction in the behaviour of the range of the charge or the standard deviation between the side chain atoms and backbone atoms. Peptide nitrogen atoms have a range between 0.65 and 0.80 a.u. around the average value. Peptide oxygen atoms have a smaller range, between 0.45 and 0.65 a.u. around the average. Peptide hydrogen atoms have the smallest range in atomic charge of the peptide group atoms, with a range of 0.30–0.40 a.u. around the average value. Peptide carbon atoms show the largest range of atomic charges with a range often over 1 a.u. around the average value. Alpha carbon atoms have a smaller range of roughly 0.65 a.u. The side chain methylene carbon atoms exhibit smaller ranges in atomic charge than both the peptide and C_α atoms. The standard deviation of the atomic charge does not show any correlation with the magnitude of the charge. However, a larger range of charges does correlate to a larger standard deviation. The standard deviation of the hydrogen atoms is always less than 0.05 a.u., which is considerably smaller than most carbon, nitrogen and oxygen atoms. Hydrogen atoms within the functional groups of positively charged amino acid side chains (for example H32, H33 or H34 of Lys^+) exhibit a decrease in the range of charges and the standard deviation relative to the same hydrogen atoms when present in the neutral side chain. This is due to less charge being available to these atoms (overall charge of +1). The standard deviation of the carbon, oxygen and nitrogen typically lie in the range of 0.5–1.2 a.u.

3.1. Acidic amino acids (Asp and Glu)

The atomic charge (averaged over all thousand geometries) for all atoms of both Asp and Asp^- can be seen in Fig. 3. The difference between the atomic charges in the neutral and in the charged amino acid is also plotted. Atom H25 is the acidic proton that is removed upon deprotonation. In the neutral molecule, the acidic proton has a charge of +0.56 a.u. (see Fig. 3), which means that upon deprotonation a charge of $(-1) + 0.56 = -0.44$ a.u. is left over

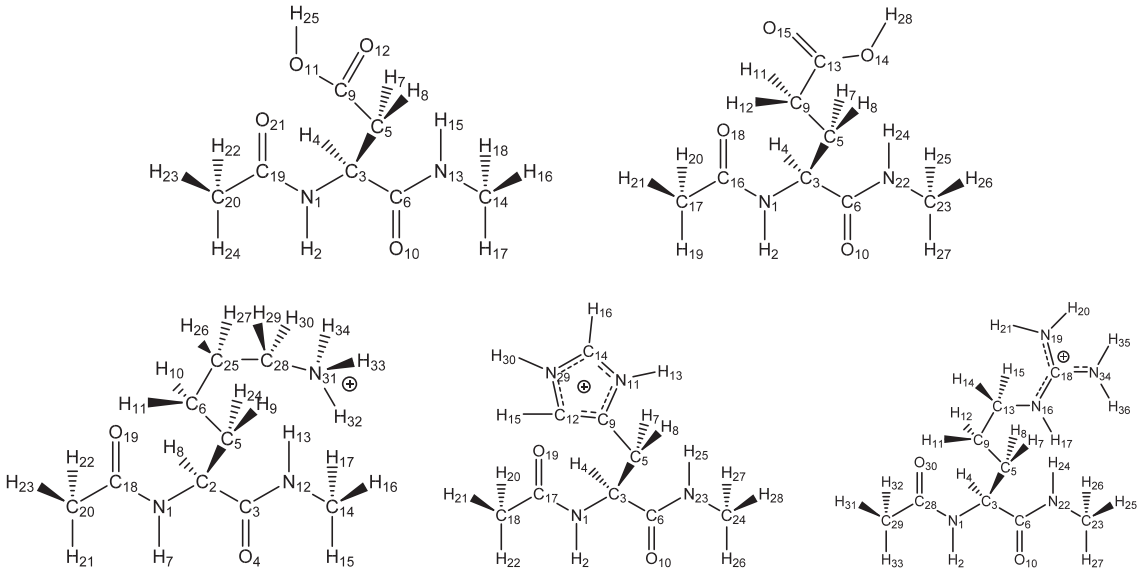

Fig. 2. Numbered geometries for capped amino acids Asp (top left), Glu (top right), Lys^+ (bottom left), His^+ (bottom middle) and Arg^+ (bottom right). The numerical labels of the atoms ("atom number") of the deprotonated geometries are the same. In all five cases the proton removed upon deprotonation is the highest numbered proton.

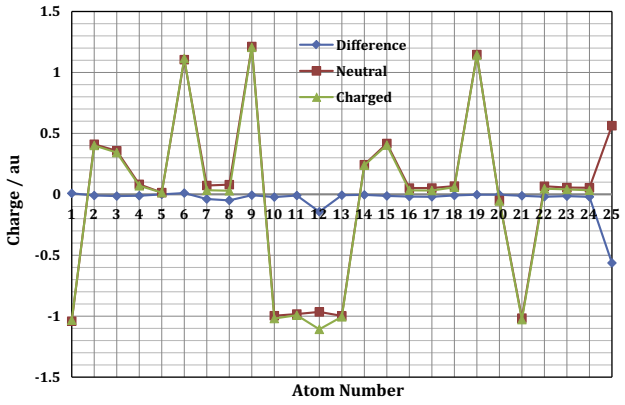

Fig. 3. The averaged atomic charges of both Asp (green) and Asp^- (red) and the difference (blue) between the neutral and charged atomic charges. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to be distributed over the remaining geometry. Of this 0.44 of an electron, 57% (−0.25 a.u.) moves onto the side-chain atoms. The remaining 43% (−0.19 a.u.) is found on the backbone atoms.

Despite the even spread of H25's charge over the whole molecule upon deprotonation, the total molecular charge is highly concentrated on the side-chain of the molecule of Asp^-. Upon deprotonation the sum of all side-chain atomic charges (including H25 for the neutral side-chain) decreases from −0.01 a.u. to −0.81 a.u. meaning that 81% of the total molecular charge is found on the side-chain atoms. The carboxylate group of Asp^- has a summed charge of −0.89 a.u. (89% of the molecular charge). The methylene group of the side-chain increases in charge from Asp to Asp^-, with a summed (group) atomic charge of 0.08 a.u. (= |−0.89 − (−0.81)|). There are no chemically significant changes in atomic charges of the backbone atoms. Curiously, one of the most significant changes in backbone charge is that the hydrogen atoms on the methyl capping groups undergo a difference in summed charge of −0.10 a.u. when going from Asp to Asp^-.

Similar results are found for the deprotonation of Glu to Glu^-. The differences in average atomic charge over a thousand conformations are shown in Fig. 4. Atom H28 corresponds to the acidic proton that is removed when going from Glu to Glu^-. The charge of H28 in Glu is 0.55 a.u. meaning that in Glu^- only −0.45 a.u. of additional negative charge is available to the molecule for redistribution. A value of −0.33 a.u. of the additional charge (73%) remains on the side-chain atoms, and the remaining −0.12 a.u. is shared by the backbone atoms.

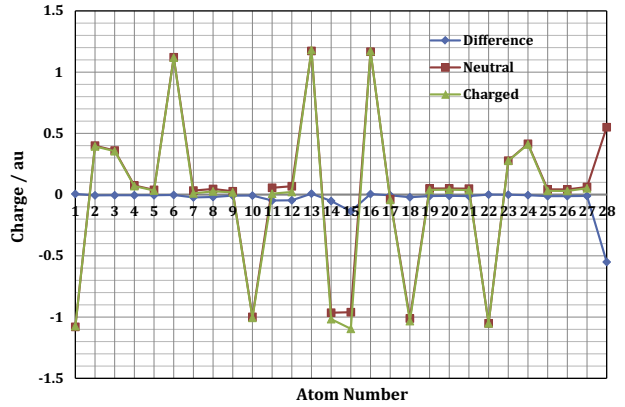

Fig. 4. The averaged atomic charges of both Glu (green) and Glu^- (red) and the difference (blue) between the neutral and charged atomic charges. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Similar to Asp^-, it is apparent that the majority of the negative molecular charge of Glu^- is found on the side-chain atoms (−0.88 a.u., 88% of total molecular charge). A similar situation to that of Asp^- arises where the majority of the side-chain charge of Glu^- is concentrated on the carboxylate group. In Glu^- the carboxylate atoms have a summed charge of −0.93 a.u., which is an increase in charge of −0.73 a.u. relative to the summed charge of the neutral carboxylic acid group. There is no significant change in backbone atom charges. The methyl hydrogens increase in summed charge by −0.7 a.u., which is less than in the case of Asp^-.

There are differences between the changes seen in atomic charges for the two systems Asp^- and Glu^-. Eight percent more charge is located on the side-chain of Glu^- than on the side-chain of Asp^-. Also, less of the additional charge available upon deprotonation is found on the backbone atoms for Glu^- (26%) compared to Asp^- (43%). This observation has led to the idea of a "buffering" methylene group. Methylene groups are neutral fragments in the side-chain that act to separate the polar carboxylic acid/carboxylate group from the rest of the amino acid. The additional methylene group in the side-chain of Glu^- creates a more insulating buffer between the charged carboxylate group and the amino acid backbone. This buffering is responsible for the increased localisation of the charge on the side-chain in Glu^- than in Asp^-.

In summary, the deprotonation of the acidic hydrogen in Asp and Glu, causes the newly available negative charge to predominantly reside on the side-chain atoms (81% and 88% for Asp^- and Glu^-, respectively). In particular, the charge is localised on the three carboxylate atoms (COO^-). Changes in the charge of backbone atoms, when going from the protonated to the deprotonated state, are insignificant due largely to "buffering" methylene groups. The buffering effect is greater in the case of Glu^- where there are two methylene groups.

3.2. Basic amino acids (Lys, His and Arg)

Fig. 5 shows the atomic charges of Lys and Lys^+. The acidic proton in Lys^+ (H33) has a charge of 0.48 a.u. This means that the atoms present in Lys undergo a sum increase in positive charge of 0.52 a.u. when going from neutral Lys to protonated Lys^+ (because 0.52 of an electron has moved onto H33). A positive charge of 0.44 a.u. (85%) is generated on side-chain atoms. As one would expect, the backbone atoms of Lys^+ remain relatively unaffected by the protonation of the amine group due to the four methylene groups "buffering" the ammonium group from the backbone. This explains the summed charge of the backbone atoms increasing by only (0.52–0.44 =) 0.08 a.u. upon protonation.

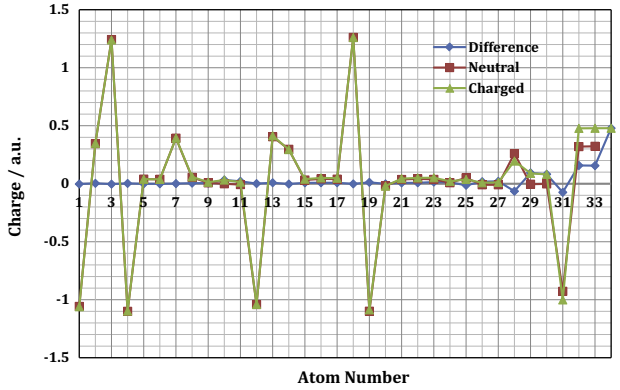

Fig. 5. The averaged atomic charges of both Lys (red) and Lys^+ (green) and the difference (blue) between the neutral and charged atomic charges. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

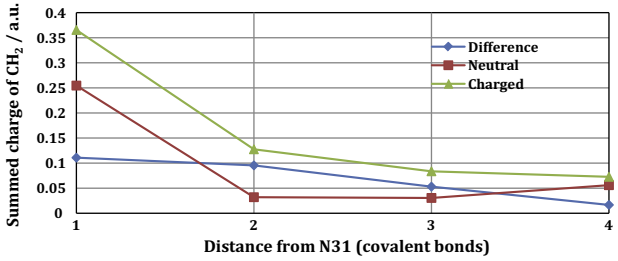

Fig. 6. Summed charges of the methylene groups of Lys (red), Lys^+ (green) and their difference (blue) against the number of covalent bonds from the side-chain nitrogen atom (N31) (1 = C_ε, 2 = C_δ, 3 = C_γ and 4 = C_β). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

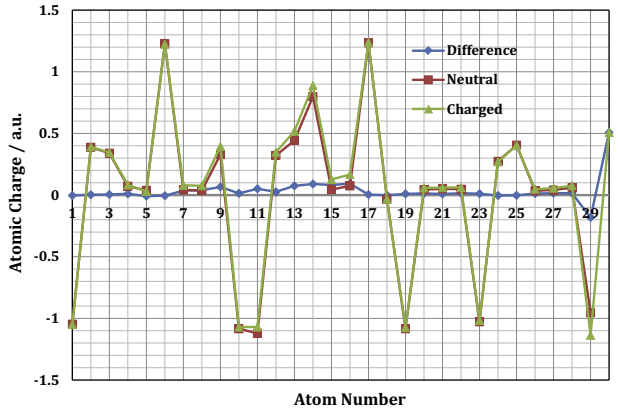

Fig. 7. The averaged atomic charges of both His (red) and His^+ (green) and the difference (blue) between the neutral and charged atomic charges. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

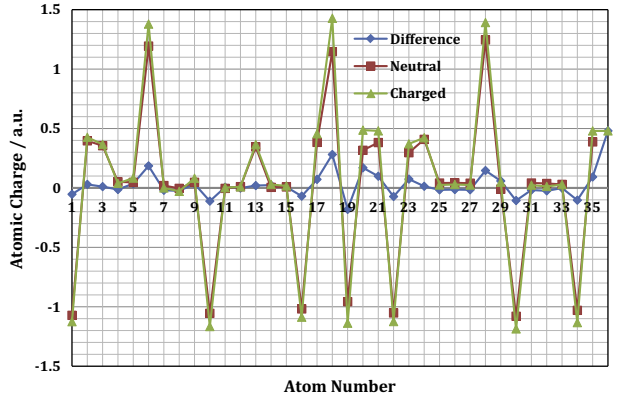

Fig. 8. The averaged atomic charges of both Arg (red) and Arg^+ (green) and the difference (blue) between the neutral and charged atomic charges. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fragmenting the molecule into side-chain atoms and backbone atoms and summing the atomic charges gives a clear illustration of the buffering effect. The summed charge of all side-chain atoms in Lys is 0.09 a.u., whereas in Lys^+ the summed charge is 1.01 a.u. (an increase of 0.92 a.u.), whereas the backbone atoms have a summed charge of −0.01 a.u. This shows that all of the positive molecular charge is found on the side-chain. The ammonium atoms ($[-NH_3]^+$) of Lys^+ have a summed charge of 0.43 a.u., which is the largest contribution to the molecular charge. The remaining charge resides on the methylene groups. The summed charge of

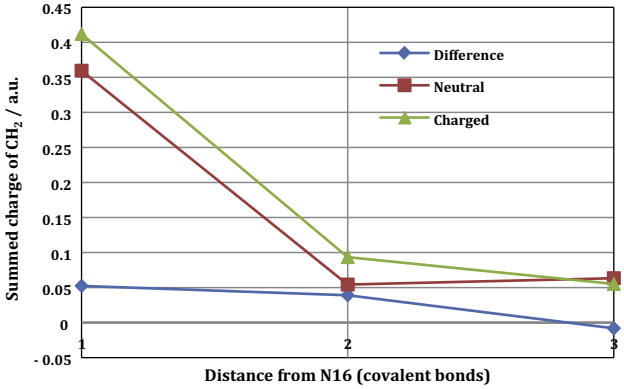

Fig. 9. Summed charges of the methylene groups of Arg (red), Arg^+ (green) and their difference (blue) against the number of covalent bonds counting from the side-chain nitrogen atom (N16) (1 = C_δ, 2 = C_γ and 3 = C_β). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

each methylene group is plotted in Fig. 6 against the number of covalent bonds between the carbon atom and the ammonium nitrogen. The summed charge of the methylene atoms decreases as the number of covalent bonds away from the ammonium nitrogen increases. The summed charge of the methylene groups in the neutral Lys molecule are also plotted in Fig. 6. From left to right, the gap between the neutral and charged values narrows, and by the fourth methyl carbon the difference between the charged and neutral methylene groups is only a summed charge of 0.02 a.u. This illustrates clearly the "buffering" effect of the methylene groups; the backbone atoms are almost unaware of the protonation of the amine group.

The atomic charges of His and His^+ can be seen in Fig. 7. The acidic proton of His^+ (H30) has a charge of 0.51 a.u. meaning that 0.49 a.u. of positive charge much be built up on the atoms present only in His (0.49 of an electron has moved onto H30). Of this charge, 76% (0.37 a.u.) lies on the side-chain atoms. The summed charge of the side-chain atoms is 0.93 a.u, which is 0.88 a.u more positive than the neutral side-chain. This again shows that the molecular charge is predominantly located on the side-chain, with the backbone atoms of His^+ undergoing a change in summed charge of 0.12 a.u. The only other amino acid that only has a single methylene group to protect the side-chain from the effects of side-chain protonation is Asp/Asp^-. The backbone atoms of Asp^- experience a greater change in summed charge (−1.9 a.u.). An incorrect assumption would be at the methylene (C5H7H8, Fig. 2) in Asp^- is a worse "buffer" than the methylene (C5H7H8) in His^+. This is not true. Instead, in His^+ the positive charge is delocalised over the imidazolium and therefore its methylene group is no longer directly bonded to a charged atom but rather a group of atoms charged to a lesser extent. Thus, the methylene group in His^+ is only 0.07 a.u. more positive than the methylene in the neutral His, compared to a difference of −0.16 a.u. for the methylene of Asp and Asp^-.

The atomic charges of Arg and Arg^+ can be seen in Fig. 8. The acidic proton of Arg^+ (H36) has a charge of 0.48 a.u. meaning that 0.52 a.u. of positive charge is built up on the atoms present in the neutral Arg molecule. The side-chain atoms of Arg increase by a total summed charge of 0.44 a.u. when the proton is added, which accounts for 86% of the charge build up. The small contribution to this charge by the backbone atoms is due to a combination of the factors previously discussed. Firstly, there are three buffering methylene groups to separate the protonated guanidinium group from the backbone. The summed charge of the methylene groups can be seen in Fig. 9. By the second methyl group (C_γ) the

difference between the charged and neutral methylene groups is less than 0.05 a.u. The second reason for the low increase in backbone charge is that the positive charge is stabilised by the delocalised – system of the guanidinium group. The eight guanidine atoms present in Arg account for 81% of the total side-chain increase in summed charge of Arg^+.

The side-chain atoms of Arg^+ have a summed charge of 1.01 a.u., accounting for all of the positive charge of the molecule. The backbone atoms have a summed charge of −0.01 a.u. due to the three "buffering" methylene groups and the spread of the charge over the guanidinium group. The guanidinium group has a summed charge of 0.45 a.u, which is the largest contribution to the molecular charge. The next largest contributor to the molecular charge is the methylene group adjacent to the guanidinium group, with summed charge of 0.41 a.u.

4. Conclusions

The atomic charges of five amino acids that undergo protonation (Lys, His and Arg) and deprotonation (Asp and Glu) have been studied. The QCT atomic charges of all atoms, averaged over a thousand conformations, for both charged and neutral amino acids have been compared. For Asp and Glu, which are deprotonated to form Asp^- and Glu^-, the majority of the negative charge is located on the side-chain atoms (81% and 88% respectively). Less charge is found on the backbone of Glu^- than Asp^- due to the additional side-chain methylene group "buffering" the charge. The buffering effect of methylene groups is more apparent in the positively charged amino acids Lys^+, His^+ and Arg^+ due to the large number of methylene groups in Lys^+ and Arg^+. By the third methylene group counting from the site of protonation, the summed charge of the CH_2 group is comparable to that of the neutral molecule. Spread of the charge over multiple side-chain atoms (such as in the imidazolium ring of His^+ and the guanidinium group of Arg^+) also reduces the effect of the charge on backbone atoms.

Acknowledgements

TJH is grateful to AstraZeneca for the provision of top-up funding his BBSRC CASE Ph.D. studentship.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.comptc.2014.07.020.

References

[1] C.M. Maupin, G.A. Voth, Proton transport in carbonic anhydrase: insights from molecular simulation, Biochim. Biophys. Acta (BBA) – Proteins Proteomics 1804 (2010) 332–341.
[2] R.L. Mikulski, D.N. Silverman, Proton transfer in catalysis and the role of proton shuttles in carbonic anhydrase, Biochim. Biophys. Acta (BBA) – Proteins Proteomics 1804 (2010) 422–426.
[3] M. Calvaresi, M. Garavelli, A. Bottoni, Computational evidence for the catalytic mechanism of glutaminyl cyclase, DFT Invest. Proteins: Struct., Funct., Bioinformatics 73 (2008) 527–538.
[4] H.B. Dunford, Mechanisms of horseradish peroxidase and alpha-chymotrypsin, Prog. React. Kinet. Mech. 38 (2013) 119–129.
[5] A. Warshel, R.M. Weiss, An empirical valence bond approach for comparing reactions in solutions and in enzymes, J. Am. Chem. Soc. 102 (1980) 6218–6226.
[6] A.C.T. van Duin, S. Dasgupta, F. Lorant, W.A. Goddard, ReaxFF: a reactive force field for hydrocarbons, J. Phys. Chem. A 105 (2001) 9396–9409.
[7] M.W. van der Kamp, A.J. Mulholland, Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology, Biochemistry 52 (2013) 2708–2728.
[8] M.J.L. Mills, P.L.A. Popelier, Polarisable multipolar electrostatics from the machine learning method Kriging: an application to alanine, Theor. Chem. Acc. 131 (2012) 1137–1153.
[9] M.J.L. Mills, G.I. Hawe, C.M. Handley, P.L.A. Popelier, Unified approach to multipolar polarisation and charge transfer for ions: microhydrated Na+, PhysChemChemPhys 15 (2013) 18249–18261.
[10] S.M. Kandathil, T.L. Fletcher, Y. Yuan, J. Knowles, P.L.A. Popelier, Accuracy and tractability of a Kriging model of intramolecular polarizable multipolar electrostatics and its application to histidine, J. Comput. Chem. 34 (2013) 1850–1861.
[11] T.L. Fletcher, S.M. Kandathil, P.L.A. Popelier, The prediction of atomic kinetic energies from coordinates of surrounding atoms using Kriging machine learning, Theor. Chem. Acc. 133 (1499) (2014). 1491-1410.
[12] P.I. Nagy, B. Noszal, Theoretical study of the tautomeric/conformational equilibrium of aspartic acid zwitterions in aqueous solution, J. Phys. Chem. A 104 (2000) 6834–6843.
[13] M. Remko, D. Fitz, B.M. Rode, Effect of metal ions (Li+, Na+, K+, Mg2+, Ca2+, Ni2+, Cu2+ and Zn2+) and water coordination on the structure and properties of L-histidine and zwitterionic L-histidine, Amino Acids 39 (2010) 1309–1319.
[14] E. Deplazes, W. van Bronswijk, F. Zhu, L.D. Barron, S. Ma, L.A. Nafie, K.J. Jalkanen, A combined theoretical and experimental study of the structure and vibrational absorption, vibrational circular dichroism, Raman and Raman optical activity spectra of the L-histidine zwitterion, Theor. Chem. Acc. 119 (2008) 155–176.
[15] F. Weixin, K.R. Amareshwar, K. Rai, Z. Lu, Z. Lin, Structural stabilities of metallated histidines in gas phase and existence of gaseous zwitterionic histidine conformers, J. Mol. Struct. 895 (2009) 65–71.
[16] R.F.W. Bader, Atoms in Molecules. A Quantum Theory, Oxford Univ. Press, Oxford, Great Britain, 1990.
[17] P.L.A. Popelier, Atoms in Molecules. An Introduction, Pearson Education, London, Great Britain, 2000.
[18] P.L.A. Popelier, in: D.J. Wales (Ed.), Structure and Bonding. Intermolecular Forces and Clusters, vol. 115, Springer, Heidelberg, Germany, 2005, pp. 1–56.
[19] C.F. Matta, 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, Proteins: Struct., Funct. Genet. 52 (2003) 360–399.
[20] C.F. Matta, R.F.W. Bader, Atoms-in-molecules study of the genetically encoded amino acids. II. Computational study of molecular geometries proteins: structure, Funct. Genet. 48 (2002) 519–538.
[21] C.F. Matta, R.F.W. Bader, An atoms-in-molecules study of the genetically-encoded amino acids: I. Effects of conformation and of tautomerization on geometric, atomic, and bond properties, Proteins: Struct. Funct. Genet. 40 (2000) 310–329.
[22] S. Cardamone, T.J. Hughes, P.L.A. Popelier, Multipolar electrostatics, Phys. Chem. Chem. Phys. 16 (2014) 10367–10387.
[23] P.L.A. Popelier, New insights in atom-atom interactions for future drug design, Curr. Top. Med. Chem. 12 (2012) 1924–1934.
[24] C. Fonseca Guerra, J.W. Handgraaf, E.J. Baerends, F.M. Bickelhaupt, Voronoi deformation density (VDD) charges: assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD methods for charge analysis, J. Comp. Chem. 25 (2004) 189–210.
[25] P. Bultinck, C. Van Alsenoy, P.W. Ayers, R. Carbó-Dorca, Critical analysis and extension of the Hirshfeld atoms in molecules, J. Chem. Phys. 126 (2007) 144111.
[26] D.S. Kosov, P.L.A. Popelier, Atomic partitioning of molecular electrostatic potentials, J. Phys. Chem. A 104 (2000) 7339–7345.
[27] P.L.A. Popelier, in: G. Frenking, S. Shaik (Eds.), The Nature of the Chemical Bond Revisited, Wiley-VCH, 2014, pp. 271–308 (Chapter 8).
[28] N.O.J. Malcolm, P.L.A. Popelier, An algorithm to delineate and integrate topological basins in a three-dimensional quantum mechanical density function, J. Comp. Chem. 24 (2003) 1276–1282.
[29] N.O.J. Malcolm, P.L.A. Popelier, The full topology of the Laplacian of the electron density: scrutinising a physical basis for the VSEPR model, Faraday Discuss. 124 (2003) 353–363.
[30] P.L.A. Popelier, F.M. Aicken, Atomic properties of amino acids: computed atom types as a guide for future force field design, ChemPhysChem 4 (2003) 824–829.
[31] P.L.A. Popelier, É.A.G. Brémond, Geometrically faithful homeomorphisms between the electron density and the bare nuclear potential Int, J. Quant. Chem. 109 (2009) 2542–2553.
[32] M. Rafat, M. Devereux, P.L.A. Popelier, Rendering of quantum topological atoms and bonds, J. Mol. Graph. Modell. 24 (2005) 111–120.
[33] M. Rafat, P.L.A. Popelier, Visualisation and integration of quantum topological atoms by spatial discretisation into finite elements, J. Comput. Chem. 28 (2007) 2602–2617.
[34] Y. Yuan, M.J.L. Mills, P.L.A. Popelier, F. Jensen, Comprehensive analysis of energy minima of the twenty natural amino acids, J. Phys. Chem. A, in press, http://dx.doi.org/10.1021/jp503460m.
[35] T.J. Hughes, S.M. Kandathil, P.L.A. Popelier, Accurate prediction of polarised high order electrostatic interactions for hydrogen bonded complexes using the machine learning method Kriging, Spectrochim. Acta Part A (2014), http://dx.doi.org/10.1016/j.saa.2013.10.059.
[36] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M.

Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian Inc., Wallingford, CT, 2009.

[37] F. Jensen, Polarization consistent basis sets. III. The importance of diffuse functions, J. Chem. Phys. 117 (2002) 9234–9240.

[38] T.A. Keith, AIMAll (Version 10.07.25), 2010, aim.tkgristmill.com.

[39] F.M. Aicken, P.L.A. Popelier, Atomic properties of selected biomolecules. Part 1. The interpretation of atomic integration errors, Can. J. Chem. 78 (2000) 415–426.