On the solvation of L-aspartic acid

A. T. PAXTON[1] and J. B. HARPER[2]

[1] Atomistic Simulation Group, School of Mathematics and Physics
Queen's University, Belfast, BT7 1NN, UK

[2] School of Chemical Sciences, University of New South Wales
Sydney NSW 2052, Australia

We use molecular statics and dynamics to study the stability of L-aspartic acid both *in vacuo* and solvated by polar and non polar molecules using density functional theory in the generalised gradient approximation. We find that structures stable *in vacuo* are unstable in aqueous solution and *vice versa.* From our simulations we are able to come to some conclusions about the mechanism of stabilisation of zwitterions by polar protic solvents, water and methanol.

1. Introduction

When α-amino acids occur in nature in neutral aqueous environments they occur predominantly as zwitterions; the amino group is *protonated* having the formula $-NH_3^+$ and the carboxyl group is *deprotonated* having the formula $-COO^-$. Whether a particular group is protonated depends on the pH in aqueous solution and for each group one defines a pK_a-value—the pH at which the group is protonated on average *half the time.* The situation in L-aspartic acid is complicated slightly by there being an additional ionisable $\beta-$carboxyl group terminating the side chain. Each of the three ionisable groups has its own pK_a [1] so that the structure one might find as a function of pH can be displayed as in figure 1. At very low pH the predominant species is that in which all the functional groups are protonated in which case the molecule has a charge of +1 in units of the electronic charge. Above the pK_a of the $\alpha-$carboxyl group that group tends to become deprotonated and the predominant molecule is neutral. In this work we will refer to this zwitterion as "z_1." In a wide range of pH including the range of biological interest both $\alpha-$ and $\beta-$carboxyl

groups tend to be deprotonated and the predominant species in solution has an overall charge of –1: here we will refer to this zwitterion as "z_0^-." Only above pH 9.8 is the predominant species that with amino group deprotonated, as illustrated in figure 1.

It has already been pointed out [2] that calculations *in vacuo* (gas phase) find that the zwitterion z_1 is unstable with respect to the non charge-separated form, which is illustrated schematically at the top left in figure 2 and which we will refer to as the "n" form of L-aspartic acid.† In this paper we make a preliminary study of the stability of neutral and negatively charged L-aspartic acid using electronic structure calculations based in the local density approximation to density functional theory. We confirm that the n form is the only stable structure *in vacuo* of the neutral molecule; furthermore stable and metastable forms of the negatively charged molecule *in vacuo* are also not the structure found in aqueous solution: the situation is summarised in figure 2 and is discussed below. We find that the effect of solvent stabilisation can be reproduced with rather a small number of solvent molecules in a simulation and that these reveal some interesting aspects of the underlying chemistry.

2. Methods

Our calculations are based in density functional theory (DFT) in the local density approximation (LDA) with additional non local corrections to the exchange and correlation functional for which we use the parameterisation of Perdew *et al.* [3]. A complete description of the theory can be found in Parr and Yang [4]. In the LDA the exchange and correlation potential at point $\mathbf{r}$ is a function only of the electron density at that point; in the generalised gradient approximation (GGA) it is a function of both the density and its gradients at the point $\mathbf{r}$. In this sense it is a non local correction. The principal reason for employing non local corrections is to allow a description of the hydrogen bond, which is strongly overbound in the LDA alone. The reason for this is simply that the LDA can give a poor description of situations where the electron density is very small, because of the incorrect asymptotic form of the exchange and correlation potential. Furthermore the GGA generally gives a better account of barrier heights in chemical reactions [5].

The Kohn–Sham equations which emerge from the DFT are solved here in a basis of linear muffin-tin orbitals (LMTO) using a method devised by Methfessel and van Schilfgaarde [6]. This is an all electron method in which the core is treated as spherical, but is allowed to relax, and is by construction orthogonal to all valence states. There is no pseudopotential or frozen core aproximation made thereby. Total energy and interatomic force are obtained in real space using an isolated cluster of atoms which may be neutral or charged. All the molecules and clusters considered here have an even number of electrons and hence we have

† It is usual in the literature to refer to this form as "neutral." We prefer the term "non charge-separated" which, while also non precise, emphasises that both n and z_1 are charge neutral compared to the singly deprotonated molecules that we also consider here.

not admitted spin polarisation in what follows. LMTO's have the exponential asymptotic form of Hankel functions rather than squared exponential of Gaussians and are therefore expected to be more readily complete. The choice of basis set however follows the usual "double energy plus polarisation" philosophy, hence hydrogen is described with two *s*-orbitals with different asymptotic decay and one *p*-orbital. Carbon, nitrogen and oxygen are described with two *s*- and *p*-orbitals and one *d*-orbital. The charge density is also represented in a Hankel function basis set which, at least in principle, is better suited than a Gaussian to describe the exchange and correlation potential at vanishing densities. Molecular statics is done using energy minimisation based in the Fletcher–Powell variable metric algorithm [7]. Molecular dynamics is performed in an NpT ensemble using a Nosé–Hoover thermostat [8] to control the temperature. Note, however that the temperature fluctuates strongly in simulations of small clusters and the energy is sometimes found to become locked into rotational modes.

We can confirm briefly that the most important structural properties of water are reproduced by the LMTO method. In table 1 we show results for the water monomer and dimer, confirming that we have a realistic description of the hydrogen bond.

TABLE 1: Calculated properties of the water molecule and water dimer. For the single molecule, we used a "triple energy plus polarisation" basis. Pseudopotential and observed values are taken from Sprik et al. *[9], except the observed O–O vibrational frequency which is taken from Braly* et al. *[10].*

		This Work	Pseudopotential	Observed
H_2O	bond length(Å)	0.97	0.97	0.96
	bond angle	104.6°	104.4°	104.5°
	dipole moment (D)	1.88	1.83	1.86
$(H_2O)_2$	O–O bond length(Å)	2.98	2.94	2.98
	O–H–O bond angle	172	174	174
	O–O vibrational frequency (cm^{-1})	∼170	223	150

In addition to our density functional computer programs we found two classical simulation packages extremely useful. Amino acids were solvated in the first instance using the program SOLVATE [11] which generates a solvation cloud in a strictly statistical way. The program Ghemical (Gnu-Chemical) [12] was used both for building molecules and modifiying the side groups and for geometry relaxation using a molecular mechanics force field [13]. Finally, the program Jmol [14] was used for creating the molecular images in the figures.

3. Gas phase molecular forms

The neutral and negatively charged forms of L-aspartic acid are displayed schematically in figure 2. Each of these is metastable or stable *in vacuo* (that is, there is a minimum in the zero temperature potential energy surface) when the structural energy is minimised in a valence force field. Such relaxed structures were used as the starting point for all our subsequent calculations. Using DFT-GGA we then find just one stable neutral form and one stable and one metastable singly deprotonated form. The structures we find for the stable molecules are shown in the lower half of figure 3.

3.1 Neutral molecules

Of the three charge neutral molecules, only the n form is stable. Both z_1 and z_2 are unstable in the sense that they are not even close to a local minimum in the GGA–energy surface. During the variable metric relaxation (which does not correspond to any physical process) the ionised groups move together until they are close enough to accomplish a proton exchange; they subsequently move apart. The energy gained in the relaxation of z_1 and z_2 are 213 kJ/mol and 246 kJ/mol respectively, however since the starting forms are not local minima in the energy surface these numbers are useful only in the context of molecular mechanics in which these forms are metastable. We have also performed molecular dynamics using the same starting structures for z_1 and z_2. Essentially the same geometrical changes occur, the zwitterions surviving for less than 0.1 ps at a target temperature of 335 K.

3.1 Charged molecules

A similar agreement between molecular statics and dynamics occurs for the three singly deprotonated ions at the right of figure 2. In this case we find that the z_2^- ion is metastable both in statics and dynamics. We have simulated z_2^- at 320–340K for 10 ps and it has not reverted to z_1^- over that time. One might speculate that there is not enough thermal energy to overcome the steric hindrance that prevents the two carboxyl groups getting close enough to complete proton transfer; indeed it is geometrically not possible for proton transfer to be accomplished in this case by the formation of the preferred *linear* hydrogen bond. On the other hand z_0^- reverts to z_1^- both in molecular statics and dynamics where the transfer takes some 0.15 ps, about twice as long as in the neutral molecules.

We explain the stability of n and z_1^- and the metastability of the z_2^- ion as arising from the apparent stability of the neutral $–NH_2$ amine group *in vacuo*. This is quite contrary to its behaviour in aqueous solution where it is stabilised only under extremely alkaline conditions as seen from figure 1. One might say that $–NH_3^+$ is stabilised in solvent if the solvent molecules are able to align their dipoles to screen the electric field of the ionised group. Another way to say this is that solvent may form hydrogen bonds to $–NH_3^+$. This matter is now addressed in the next section.

4. Solvated molecules

Using the program SOLVATE we have created a solvent cloud of some 25 water molecules around the molecules n, z_1 and z_0^-. The structures were then relaxed in the force field of the program Ghemical. In DFT-GGA molecular statics energy minimisations, we found that the neutral form n is unstable in its water shell and it reverts to z_1. We replaced the solvent around z_1 with methane molecules, using the program Ghemical, and relaxed this using a force field. When the energy of this configuration of molecules was minimised using DFT-GGA the molecule reverted to the n-form, suggesting that the stabilisation of z_1 requires a *polar protic* solvent. We have also confirmed in molecular dynamics runs that z_1 and z_0^- are stable in aqueous solvent, at least for times up to 1 ps.

These results confirm that when solvated the DFT-GGA correctly predicts the structures of L-aspartic acid and singly deprotonated L-aspartate ions to be those found in real life, as indicated in figure 1. Now, in order to extract the essential molecular physics attending solvation, we have made simulations of L-aspartic acid surrounded by just a few strategically placed solvent molecules. It turns out that the presence of just six water molecules will stabilise z_1 and z_0^- both in molecular statics and dynamics. This is consistent with previous molecular statics calculations for glycine [15]. In the following two sections we describe our results using water and methanol as solvents.

4.1 Water

We have removed all but six of the solvating water molecules from the z_1 zwitterion and relaxed this structure using the valence force field. This was used as the starting structure for a molecular dynamics simulation of 7 ps. The zwitterion did not revert to the n form indicating that in this situation it behaves as solvated not as in the vacuum. Looking at the simulation it seems that the $-NH_3^+$ and $\alpha-COO^-$ groups provide an electric field that attracts the water molecules which then align themselves so as to oppose this. The water molecules are too large to intersperse themselves between the two ionised groups; indeed an alternative explanation is illustrated in the snapshot taken at 0.997 ps in figure 3 (top left). This is not a typical configuration in that over the time of the simulation hydrogen bonds shorter than 1.5 Å appear rather rarely. It is seen though that in this instant both ionised groups have hydrogen bonded to a water molecule which presumably has the effect of reducing the range of the dipole electric field and stabilising the ionised group.

In the case of z_0^- we decided to retain 9 water molecules on the principle that they were needed to screen the fields of two ionised carboxyl groups in addition to the $-NH_3^+$ group. We first relaxed this cluster using molecular statics and confirmed that the z_0^- is stable as in the solvated rather than the vacuum state. We then made a molecular dynamics simulation over 2.8 ps without the zwitterion reverting to the stable vacuum structure, z_1^-. A similar picture of the dynamical stabilisation is seen as with z_1, namely that the water molecules orient themselves so as to oppose the attraction of the ionised groups and thus

preventing proton transfer. Again the occurence of hydrogen bonds shorter than 1.5 Å was rare, but at the top right of figure 3 we show the configuration at 0.2605 ps in which, in contrast to the left hand picture of z_1, we see a single water molecule participating in a hydrogen bond to *two* ionised groups, the $-NH_3^+$ and the β–COO^-. This is a nice illustration of the property of a polar protic solvent to act as both a hydrogen bond donor and a hydrogen bond acceptor. Again we emphasise that although this is a rare event in these simulations it does not invalidate the conventional wisdom that stabilisation of ionised groups in water is through hydrogen bonding. It depends on how one defines the length of a hydrogen bond (we use the cut off defined in Jmol at 1.5 Å whereas strong hydrogen bonds up to 3 Å are commonly thought to occur). Also, because of the very small number of water molecules in the simulations, it is possible that the low pressure is keeping the water rather further from the amino acid that might be expected in a bulk sample.

4.2 Methanol

We have seen that methane is unable to stabilise the zwitterion as a solvent. It is also interesting to consider another polar protic solvent in addition to water, but one with only one hydrogen bond donor. One reason for this is that in the calculations presented up to now, it has not been possible to follow the transitions between non charge-separated molecules and zwitterions. This means it is not possible to establish energy differences between the two forms or to estimate the energy barriers between them. This is because the stabilisation by water happens very rapidly due to the ease of reorientation of the water molecules in the dynamics. To a limited extent, as we now show, methanol being larger and having a smaller dipole moment than water (1.7 D, compared to 1.9 D for water) acts rather less rapidly to stabilise the zwitterion.

We used Ghemical to create a structure consisting of three methanol molecules and the neutral zwitterion z_1. This was relaxed in the valence force field to provide a starting structure for molecular dynamics. Figure 4 (a), at the top, shows the potential energy as a function of time during a simulation over 6 ps. An expanded section over time is shown in the graph below. Clearly the events appearing here are occuring during the "equilibration time" of the dynamics and so have limited thermodynamic significance. However viewed as non equilibrium events, based on an arbitrary initial condition, they do at least show a sequence of transitions between the n and z_1 forms separated by energy barriers. The configurations labelled "i" are intermediate structures defined as occuring when the proton in transfer is equidistant from the nitrogen and α-carboxyl oxygen atoms. Within the sequence shown by arrows, the structure with the lowest potential energy is in fact the n form, not the z_1 which is finally stabilised after about 1 ps. What we can learn from figure 4 (a) is approximate energy barriers and switching times for the n to z_1 transition. In figure 4 (b) we show snapshots of the structures at some of the times indicated by arrows. The configuration at $t = 0$ is unstable, although it is the zwitterion

z_1, and goes through a series of transitions to n and back during which the methanol molecules, and especially the one seen at the top left in each diagram, reorient themselves in order effectively to screen the dipole field and stabilise, finally, the z_1 zwitterion. In conclusion methanol, like water, is able to stabilise z_1, but the process takes a great deal longer in time. As far as hydrogen bonding goes the situation is much as described in section 4.1. The two shortest bonds between amino hydrogen and methanolic oxygen are 1.66 Å and 1.63 Å at the $t = 471$ fs structure shown in figure 4 (b). The fact that they are of very similar length and linear argues that these are hydrogen bonds and one may attribute the stabilisation by methanol to this in the same way as in water.

5. Discussion and Conclusions

DFT-GGA implemented in the LMTO method provides a fast and accurate scheme for the study of molecular physics. Indeed previously the instability of non charge-separated α-amino acids *in vacuo* has been found theoretically only using either particularly complete Gaussian basis sets within the Hartree–Fock approximation or through the use of Møller–Plesset correlation corrections [2]. The scheme is sufficiently fast to allow molecular dynamics simulations of some 20–30 molecules, taking roughly 1–5 minutes per time step on a single computer processor.

In brief, our conclusions are these, and we expect these to be generally true of α-amino acids.

1. In the gase phase charged and neutral forms are only stable if the amino group is deprotonated. The instability of the $-NH_3^+$ group *in vacuo* is very strong in the sense that there is not even a local minimum in the potential energy surface.
2. The $-NH_3^+$ group is readily stabilised by the presence of just a handful of molecules of a polar protic solvent. We confirm that neutral L-aspartic acid and its singly deprotonated ion when simulated in the presence of water are stable in those forms which are found in nature.
3. The stabilisation mechanism, as expected, is through hydrogen bonding of the ionised group to solvent molecules. One may also view this as a cancellation of the dipole electric field of the zwitterion through alignment of the solvent molecules. This process happens rather more slowly in methanol than in water.

These studies show just how small a change in the environment around L-aspartic acid and its deprotonated forms is necessary to change the position of a proton, and hence the nucleophilicity of the sites in question. The trends of solvation from methane through methanol to water shows the importance of both the dielectric constant and the ability of the solvent to stabilise the charge-separated species through hydrogen bonding. These small differences indicate just how finely balanced situations involving solvation of amino acids can be and this has specific implications for enzyme catalysis, in which the quantity of solvent present in the active site may be limited.

References

[1] GARRET, R. H., and GRISHAM, C. M., 1999, *Biochemistry* (Orlando: Harcourt Brace).

[2] NAGY, P. I., and NOSZÁL, B., 2000, *J. Phys. Chem. A*, **104**, 6834.

[3] PERDEW, J. P., BURKE, K., and ERNZERHOF, M., 1996, *Phys. Rev. Letters*, **77**, 3865; *Erratum*, 1997 *ibid.*, **78**, 1396.

[4] PARR, R. G., and YANG, W., 1989, *Density Functional Theory of Atoms and Molecules* (Oxford: The Clarendon Press).

[5] PERDEW, J., P., ERNZERHOF, M., ZUPAN, A and BURKE, K., 1998, *Adv. Quantum Chem.*, **33**, 1.

[6] METHFESSEL, M., and VAN SCHILFGAARDE, M., 1993, *Phys. Rev. B*, **48**, 4937; METHFESSEL, M., VAN SCHILFGAARDE, M. and SCHEFFLER, M., 1993, *Phys. Rev. Letters*, **70**, 29; BOTT, E., METHFESSEL, M., KRABS, W., and SCHMIDT, P. C., 1998, *J. Math. Phys.*, **39**, 3393.

[7] FLETCHER, R., and POWELL, M. J. D., 1964, *Comp. J.*, **6**, 163.

[8] NOSÉ, J., 1984, *J. Chem. Phys.*, **81**, 511; NOSÉ, J., 1984, *Mol. Phys.*, **52**, 255; HOOVER, W. G. , 1985, *Phys. Rev. A*, **31**, 1695.

[9] SPRIK, M., HUTTER, J. and PARRINELLO, M., 1996, *J. Chem. Phys.*, **105**, 1142

[10] BRALY, L. B., CRUZAN, J. D., LIU, K., FELLERS, R. S., and SAYKALLY, R. J., 2000, *J. Chem. Phys.*, **112**, 10293.

[11] GRUBMÜLLER, H., 1996, `http://www.mpibpc.gwdg.de/abteilungen/071/ solvate/docu.html` (Göttingen: Max-Planck-Institut für biophysikalische Chemie).

[12] HASSINEN, T., and PERÄKYLÄ, M., 2001, *J. Comput. Chem.*, **22**, 1229; `http://bioinformatics.org/ghemical`

[13] CLARK, M., CRAMER, R. D., and VAN OPDENBOSCH, N., 1989, *J. Comput. Chem.*, **10**, 982.

[14] `http://jmol.sourceforge.net`

[15] JENSEN, J. H., and GORDON, M. S., 1995, *J. Am. Chem. Soc.*, **117**, 8159.

Figure Captions

Figure 1: Showing, as a function of pH, the predominant species present in aqueous solutions containing L-aspartic acid.

Figure 2: The possible forms of L-aspartic acid *in vacuo*, showing L-aspartic acid molecules on the left, and singly deprotonated ions on the right. They are marked *stable*, *metastable* or *unstable* according to our findings using DFT-GGA molecular statics and dynamics. This figure also serves to establish our notation for the various forms we consider in the text.

Figure 3: Some snapshots of molecular dynamics simulations of the stable forms of L-aspartic acid (left) and singly deprotonated ions (right). Upper figures are solvated and lower *in vacuo*.

Figure 4: (**a**) The potential energy (with respect to isolated, spin polarised atoms) of the z_1 zwitterion in a cluster of three methanol molecules as a function of time in a molecular dynamics simulation. The lower graph shows an expanded view from the equilibration time in which the molecule goes through a number of activated transitions between the n and z_1 forms. The graph provides an estimate of barrier heights and transition rates although note that the temperature fluctuates in the range 300–400 K. (**b**) Snapshots of the cluster described in (a) at the times shown. See the text for a discussion.

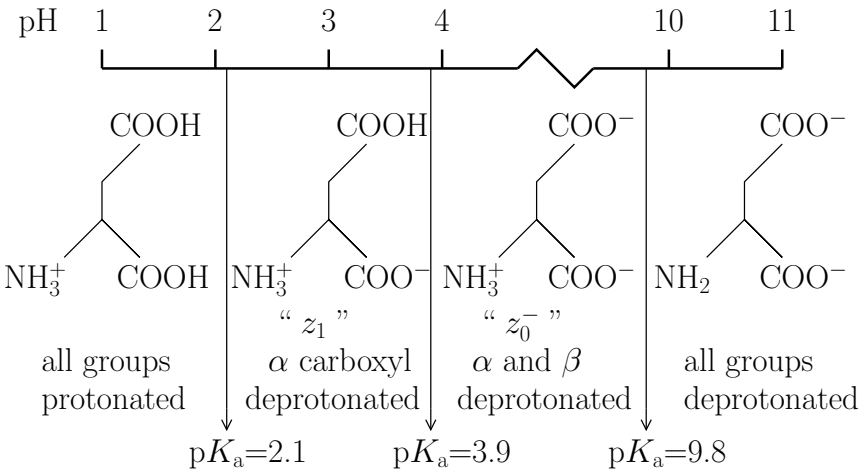


Figure 1

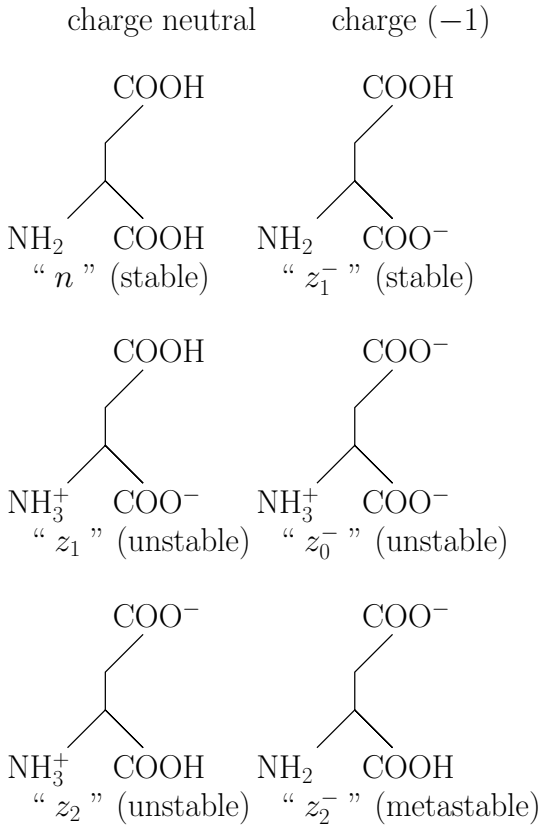
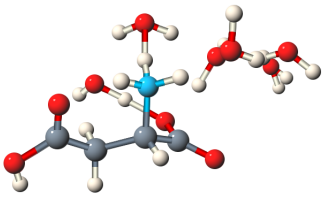
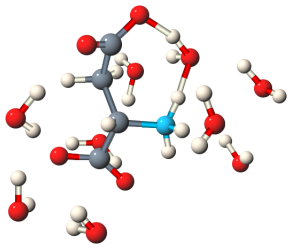


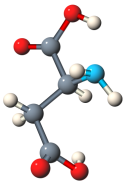
Figure 2



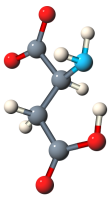
z_1 + 6 H_2O



z_0^- + 9 H_2O

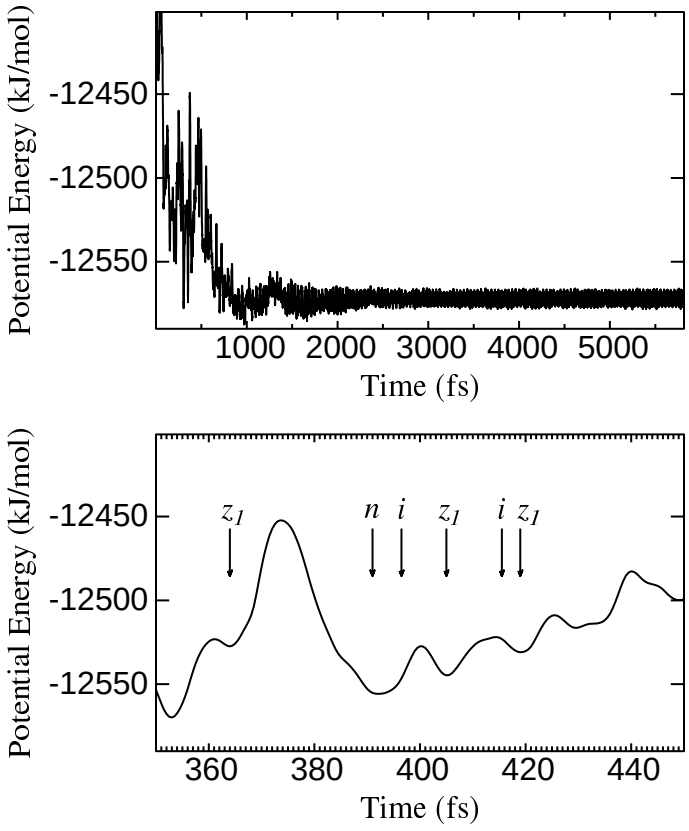


n

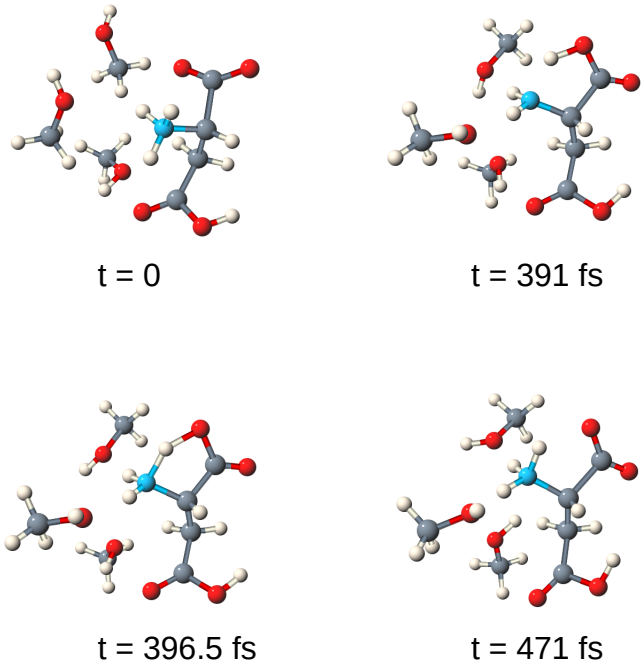


z_1^-

Figure 3


Potential Energy (kJ/mol)
-12450
-12500
-12550
1000
2000
3000
4000
5000
Time (fs)
Potential Energy (kJ/mol)
-12450
-12500
-12550
z1
n
i
z1
i
z1
360
380
400
420
440
Time (fs)

(a)


t = 0
t = 391 fs
t = 396.5 fs
t = 471 fs

(b)

Figure 4