Supplementary information for 'QM/MM simulations predict a covalent intermediate in the hen egg white lysozyme reaction with its natural substrate'

Anna L. Bowman, Ian M. Grant & Adrian J. Mulholland

Energy corrections

The QM region in the QM/MM simulations was treated with the PM3 method. The reliability of semiempirical methods such as PM3 can vary for different systems and they have known limitations, for example, in the calculation of reaction energetics in many cases.[1] To overcome the limitations of the semiempirical method, energy corrections were applied based on high-level quantum chemical calculations. The structures of several small molecules were optimized in the gas phase using the MOLPRO program,[2] at various levels of theory, and a comparison of reaction energies was carried out, and corrections applied to the QM/MM free energies based on this comparison.

Starting from the effective Hamiltonian used in the QM/MM simulations:

$$\hat{H}_{eff} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM}$$

where $\hat{H}_{QM}$ is the Hamiltonian of the QM region, $\hat{H}_{MM}$ is the Hamiltonian of the MM region and $\hat{H}_{QM/MM}$ is the hybrid Hamiltonian that accounts for QM-MM interactions. The free energy profile calculated with a QM/MM potential can be separated into a solvent independent (gas phase) term and a solute-solvent interaction component:[3]

$$W_{tot}(\text{protein}) = G_{QM}(\text{gas}) + \Delta G_{Xs} + W_{MM}$$

where $G_{QM}(\text{gas})$ is the gas-phase free energy of the QM region, ΔG_{Xs} is the QM/MM interaction free energy and W_{MM} is the potential of mean force (PMF) for the MM region. Assuming ΔG_{Xs} is similar at high and low levels of theory it follows that:

$$W_{tot}^{HL}(\text{protein}) - W_{tot}^{LL}(\text{protein}) = G^{HL}_{QM}(\text{gas}) - G^{LL}_{QM}(\text{gas})$$

where $G^{HL}_{QM}(\text{gas})$ and $G^{LL}_{QM}(\text{gas})$ are the gas-phase energies of the QM region at a high (e.g. MP2/6-311+G(d)) and low (PM3 or AM1) level of theory, respectively. $W_{tot}^{LL}(\text{protein})$ is the free energy calculated with the umbrella sampling simulations at the semiempirical level and $W_{tot}^{HL}(\text{protein})$ is the corrected free energy.

The differences in entropy contributions to the free energy and zero point energy corrections were assumed to be small and so the corrected free energy profile of the reaction is:

$$W_{tot}^{HL}(\text{protein}) = W_{tot}^{LL}(\text{protein}) + E^{HL}_{QM}(\text{gas}) - E^{LL}_{QM}(\text{gas})$$

where E^{HL}_{QM}(gas) and E^{LL}_{QM}(gas) are points on the potential energy surface of the reaction in the gas phase, corresponding to the appropriate values of the reaction coordinate.

As the small molecules represent stationary points along the reaction pathway, the energy corrections were interpolated to give corrections along the whole reaction path. The difference in energy difference between two stationary points was divided by the number of points separating them and this value was added incrementally applied to each point, as described in the example below.

Example. To explain this procedure an example is given here for the proton transfer from Glu35 Oε1 to O1 of the D site NAM. To correct this PM3/CHARM22 free energy profile with B3LYP/6-311G+(2d), the small model corrections in Figure S1 were used. The molecules representing protonated glutamate (Glu35H) and NAM on the left hand side (acetic acid and tetrahydropyran-2-ol respectively) represent a point around r_1 = −0.7 Å (the minimum) on the free energy profile. The small models of the right hand side, Glu35 and NAMH (the acetic acid anion and tetrahydropyran-2-yl-oxonium respectively) represent the minimum at r_1 = ~ +0.75 Å (i.e. the protonated sugar and deprotonated Glu35). The energy difference between products and reactants (i.e. the reaction energy) when the molecules were optimized with PM3 was 176.2 kcal mol^{-1}. The energy difference between products and reactants when the models were optimized with B3LYP/6-311G+(2d) was 150.0 kcal mol^{-1}. This means that PM3 overestimates the reaction energy by 26.2 kcal mol^{-1}. There are 30 data points on the free energy profile between r_1 = −0.7 Å and +0.75 Å inclusive. −26.2 kcal mol^{-1} is divided by 30 to give −0.87 kcal mol^{-1}, hence −0.87 kcal mol^{-1} must be incrementally added to each point from r_1 = −0.7 Å to +0.75 Å i.e. −0.87 kcal mol^{-1} is added to the free energy at point r_1 = −0.7 Å, −1.74 is added to the next data point, −2.61 to the next point and so on until r_1 = +0.75 Å is reached, where –26.2 is added to the PM3/CHARMM22 free energy value, giving the B3LYP/6-311G+(2d)-PM3/CHARM22 corrected free energy profile. A similar approach was followed for corrections at the MP2 level (and for AM1/CHARMM22 QM/MM free energies), and for the second reaction step, formation of the covalent intermediate.

Small model results

Proton transfer from Glu35 to O1 of the D site NAM. The energies of optimized small molecules were used to correct the proton transfer from Glu35 to O1 of the D site NAM. Acetic acid was used to model the protonated Glu35 (Glu35H) and tetrahydropyran-2-ol was used to model the D site NAM sugar (NAM). The acetic acid anion was used to model the deprotonated Glu35 residue and tetrahydropyran-2-yl-oxonium was used to model the protonated D site NAM (NAMH) (see Figure S1). The NAM small model was in the $^{O}S_2$ skew conformation and the NAMH small model was in the $^{4}C_1$ chair conformation.

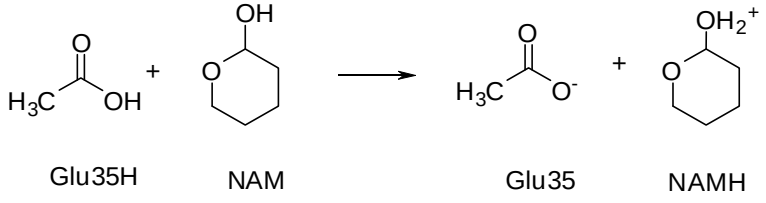


Figure S1. Small models used to correct calculated free energies of the proton transfer from Glu35 to the D site NAM O1

Table S1. Calculated energies (in kcal mol^{-1}) of the small molecules shown in Figure S1 (AM1 and PM3 energies are heats of formation; B3LYP and MP2 energies are electronic energies).

	PM3	AM1	B3LYP/6-311+G(2d)	MP2/6-311+G(2d)
Glu35H	−100.2	−100.0	−143726.0	−143439.9
NAM	−99.3	−111.2	−217668.3	−217187.9
Glu35	−118.4	−113.7	−143375.7	−143092.5
NAMH	95.2	71.0	−217868.6	−217385.0
Products − Reactants	176.2	168.4	150.0	150.2

Table S2. The discrepancies in kcal mol^{-1} between semiempirical heats of formation and *ab initio* electronic energies for the scheme shown in Figure S1.

	PM3	AM1
B3LYP/6-311+G(2d)	−26.2	−18.4
MP2/6-311+G(2d)	−26.0	−18.2

From the results for the small models shown in Figure S1, it is clear that the two higher level methods (B3LYP/6-311+G(2d) and MP2/6-311+G(2d)) are in good agreement for the energetics of this proton transfer (see Table S1). Both semiempirical methods significantly overestimate the energy difference between the reactant (e.g. **A**) and products (protonated D site NAM and deprotonated Glu35) . All the methods predict the transfer of a proton from a carboxylic acid to tetrahydropyran-2-ol (in the direction shown in Figure S1), to be unfavourable in the gas phase.

Free energy calculations at the PM3/CHARMM22 level showed the protonation of the product glycosidic oxygen to have an uncorrected barrier of $\Delta^{\ddagger}G$ ~35 kcal mol^{-1}. Correction at the B3LYP/6-311+G(2d) or MP2/6-311+G(2d) level both yielded similar corrected barrier heights of $\Delta^{\ddagger}G$ ~16 kcal mol^{-1}.

Breaking the D site NAM C1 – O1 bond: Gas phase optimization of the small models in Figure S2 was performed with each method. Tetrahydropyran-2-yl-oxonium was used to model the protonated D site NAM (NAMH) and tetrahydropyranium was used to model the D site NAM oxocarbenium ion (see Figure S2). The ion small model was in the ^{4}E envelope conformation.

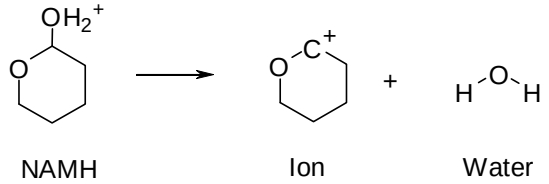


Figure S2. Small molecules used to used to correct calculated free energies of breaking the D site NAM C1 – O1 bond

Table S3. Calculated energies (in kcal mol^{-1}) of the small models shown in Figure S2 (AM1 and PM3 energies are heats of formation; B3LYP and MP2 energies are electronic energies).

	PM3	AM1	B3LYP/6-311+G(2d)	MP2/6-311+G(2d)
NAMH	95.2	71.0	−217868.6	−217385.0
Ion	140.7	133.8	−169922.8	−169522.7
Water	−53.2	−59.0	−47950.1	−47861.1
Products − Reactants	−7.7	3.9	−4.3	1.2

Table S4. The discrepancies (in kcal mol^{-1}) between semiempirical heats of formation and *ab initio* electronic energies for the scheme shown in Figure S2

	PM3	AM1
B3LYP/6-311+G(2d)	3.4	−8.2
MP2/6-311+G(2d)	8.9	−2.7

The energies of the small models following minimization with PM3 predict the reaction (in the direction shown in Figure S2) to be unfavourable, this is in agreement with the energies from the B3LYP/6-311+G(2d) optimizations. AM1 predicts the reaction to be favourable and the MP2/6-311+G(2d) energies also indicate a small decrease in energy. PM3 is found to underestimate the energy difference for the reaction of the small models whereas AM1 overestimates the energy difference.

Free energy calculations at the PM3/CHARMM22 level showed breaking the glycosidic bond to have an uncorrected barrier moving from the minimum between **A** and **B** to the covalent intermediate **C** of $\Delta^{\ddagger}G$ ~9 kcal mol^{-1}. Applying the corrections to this free energy profile yields a barrier of ~12 kcal mol^{-1} (B3LYP/6-311+G(2d)) and ~17 kcal mol^{-1} (MP2/6-311+G(2d)).

The corrected free energy profiles for the two reaction steps (i.e. for the proton transfer from Glu35 Oε1 to O1 of the D site NAM using reaction coordinate r_1, and that for breaking the D site NAM C1 – O1 bond using reaction coordinate r_2) were combined to give the whole free energy profile for forming the covalent intermediate **C** from the product **A**.

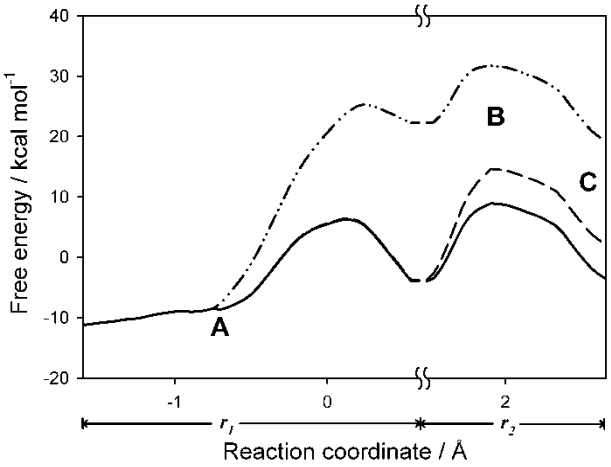


Figure S3. Combined free energy profile for the proton transfer from Glu35 to the D site NAM O1 and subsequent breaking of the D site NAM C1 – O1 bond in HEWL modeled with PM3/CHARM22 (dashed and dotted line), B3LYP/6-311+G(2d)-PM3/CHARM22 (solid line) and MP2/6-311+G(2d)-PM3/CHARM22 (dashed line).

Oxocarbenium ion vs covalent intermediate: Acetic acid tetrahydropyran-2-yl ester was used to model the covalent intermediate (Int). Acetate anion was used to model the Asp52 residue (Asp52) and tetrahydropyranium was used to model the D site NAM oxocarbenium ion (Ion) (see Figure S4)

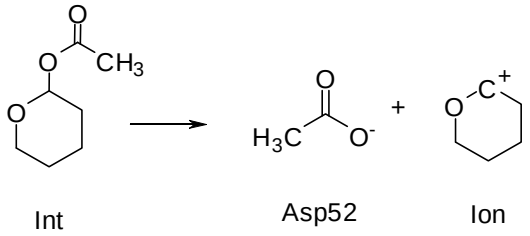


Figure S4. Small molecules used for energy calculations to correct calculated free energies for the breaking of the bond between the D site NAM C1 and the Asp52 Oδ2 in the covalent intermediate

Table S5. The energies in (kcal mol^{-1}) of the small molecules shown in Figure S3

	PM3	AM1	B3LYP/6-311+G(2d)	MP2/6-311+G(2d)
Int	−134.2	−140.8	−313437.9	−312764.2
Ion	140.7	133.8	−169922.8	−169522.7
Asp52	−118.4	−113.7	−143375.7	−143092.5
Products − Reactants	156.5	160.9	139.4	149.0

Table S6. The discrepancies (in kcal mol^{-1}) between semiempirical heats of formation and *ab initio* electronic energies for the scheme shown in Figure S3.

	PM3	AM1
B3LYP/6-311+G(2d)	−17.1	−21.5
MP2/6-311+G(2d)	−7.5	−11.8

Calculations with all methods indicate that the reaction is unfavourable in the direction shown in Figure S3. AM1 and PM3 both overestimate the energy difference between the covalent and ionic species in comparison to the higher level methods. The energy difference between the small models show the energies given by B3LYP/6-311+G(2d) to be ~10 kcal mol^{-1} more favorable towards the ionic species than those given by MP2/6-311+G(2d).

Umbrella sampling molecular dynamics free energy calculations on HEWL at the PM3/CHARMM22 level showed the oxocarbenium ion to lie ~46 kcal mol^{-1} above the covalent intermediate. Applying corrections to this free energy profile reduces this energy difference to ~30 kcal mol^{-1} (B3LYP/6-311+G(2d)) and ~38 kcal mol^{-1} (MP2/6-311+G(2d)); clearly the ionic form is predicted to be significantly higher in energy than the covalent intermediate.

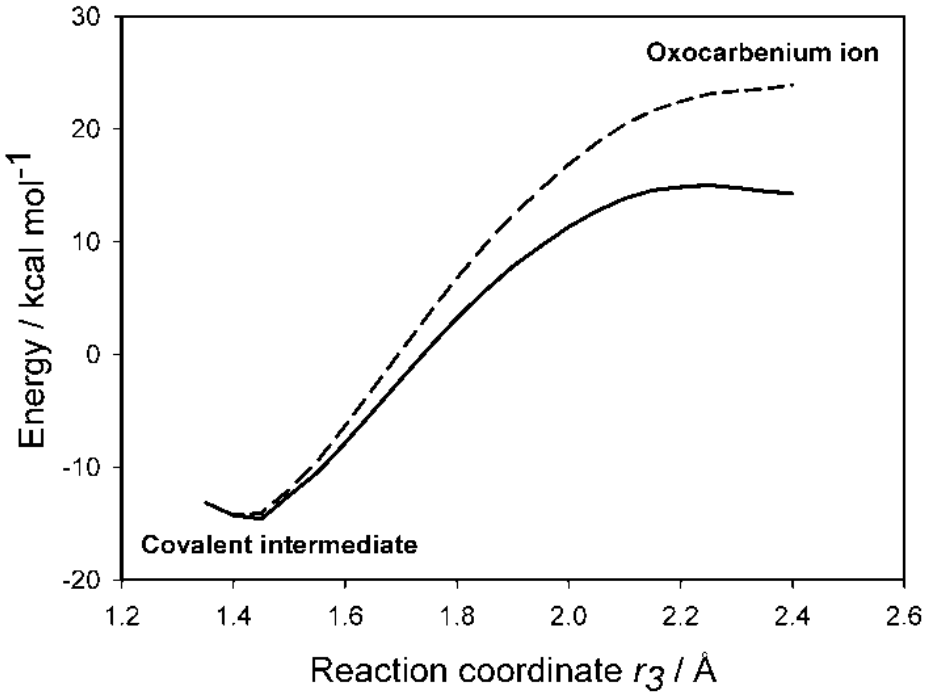
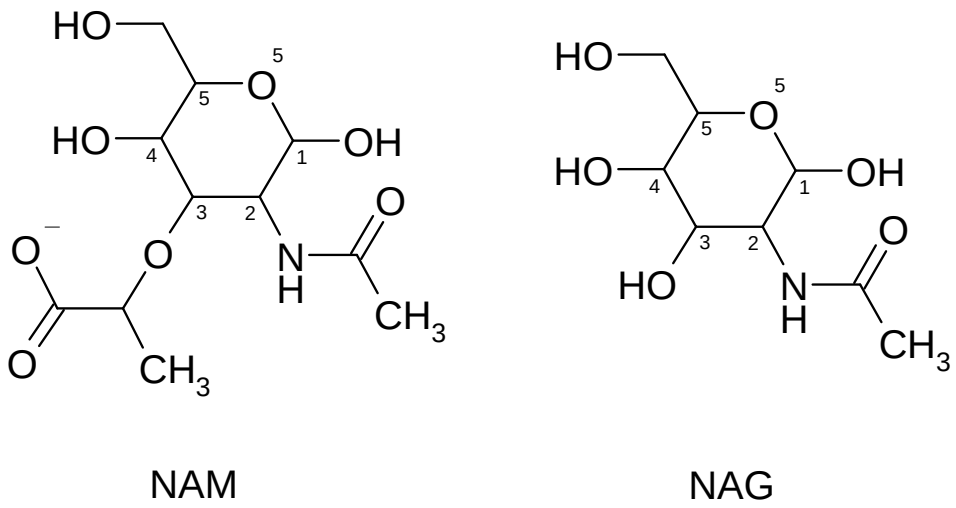


Figure S5. Corrected QM/MM free energy profiles for the breaking of the bond between the D site NAM C1 and Asp52 Oδ2 (i.e. for conversion of **C** (left hand side) to **B** (right hand side)): B3LYP/6-311+G(2d)-PM3/CHARM22 (solid line), MP2/6-311+G(2d)-PM3/CHARM22 (dashed line). The reaction coordinate here is defined as: $r_3 = d$(NAMD C1–Asp52 Oδ2)



NAM NAG

Figure S6. N-acetylmuramate (NAM) and N-acetylglucosamine (NAG), showing atom numbering for ring atoms.

References

1. W. Thiel, in *Modern Methods and Algorithms of Quantum Chemistry*, ed. J. Grotendorst, John von Neumann Institut fur Computing, Forschungszentrum Jülich, 2000, pp. 261-283.
2. H.-J. Werner, P. J. Knowles, R. Lindh, M. Schütz, P. Celani, T. Korona, F. R. Manby, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni and T. Thorsteinsson, Birmingham, UK, 2003.
3. K. Byun, Y. Mo and J. Gao, *J. Am. Chem. Soc.*, 2001, **123**, 3974-3979.