Insights on the origin of catalysis on glycine N-methyltransferase from computational modelling

Katarzyna Świderek,[1,2] Iñaki Tuñón,[3,*] Ian H. Williams,[2] Vicent Moliner.[1,2,*]

1. Departament de Química Física i Analítica; Universitat Jaume I, 12071 Castellón (Spain)
2. Department of Chemistry, University of Bath, Bath BA2 7AY (United Kingdom)
3. Departament de Química Física, Universitat de València, 46100 Burjasot (Spain)

Corresponding authors:
I. Tuñón: tunon@uv.es
V. Moliner: moliner@uji.es

ABSTRACT

The origin of enzyme catalysis remains a question of debate despite much intense study. We report a QM/MM theoretical study of the S_N2 methyl transfer reaction catalyzed by a glycine *N*-methyltransferase (GNMT) and three mutants to test whether recent experimental observations of rate-constant reductions and variations in inverse secondary α-^{3}H kinetic isotope effects (KIEs) should be attributed to changes in the methyl donor−acceptor distance (DAD): is catalysis due to a compression effect? Semiempirical (AM1) and DFT (M06-2X) methods were used to describe the QM subset of atoms, while OPLS-AA and TIP3P classical force fields were used for the protein and water molecules, respectively. The computed activation free energies and KIEs are in good agreement with experimental data, but the mutations do not meaningfully affect the DAD: compression cannot explain the experimental variations on KIEs. On the contrary, electrostatic properties in the active site correlate with the catalytic activity of wild type and mutants. The plasticity of the enzyme moderates the effects of the mutations, explaining the rather small degree of variation in KIEs and reactivities.

INTRODUCTION

Glycine *N*-methyltransferase (GNMT) is an *S*-adenosyl-L-methionine (SAM)-dependent enzyme that catalyzes the transformation of glycine into sarcosine (see Figure 1a).[1] SAM-dependent methyltransferases are inhibited by *S*-adenosyl-L-homocysteine (SAH), and it is thought that GNMT can be a regulator of the cellular SAM/SAH ratio, playing a key role in other methyl transfer reactions.[2] Thus, the activity of GNMT has been related with the concentrations of SAH and homocysteine in plasma which, in turn, has been linked to dementia and to Alzheimer's disease.[3] Nevertheless, although the catalytic mechanism of GNMT is well known based on kinetic and crystallographic studies,[4,5,6] the origin of the rate enhancement in the chemical step, a S_N2 methyl transfer, is still a question of debate. The most accepted hypothesis of the origin of the rate enhancement in this (or any other) enzyme is the relative stabilization of the transition state (TS) with respect to the reactant state (RS), as originally proposed by Pauling.[7] In other words, the reaction catalyzed by an enzyme has a lower free-energy barrier than the counterpart reaction in aqueous solution. Based on molecular simulations, this reduction of the free-energy barrier has been attributed to electrostatic preorganization of the protein that stabilizes the TS with respect to the RS.[8] Nevertheless, for enzyme catalyzed S_N2 reactions, it has been suggested that specific protein fluctuations might reduce the donor–acceptor distance (DAD), thus diminishing the potential-energy barrier height and/or width and enhancing the rate by increasing the number of reactive trajectories over and through the barrier.[9,10,11,12,13,14] Obviously, in reactions where the transferred moiety is not a light particle, as in the case of GNMT, a large contribution from tunneling is not expected. The idea of reduction in the DAD has its origins in Schowen's "compression hypothesis"[15] for enzymic methyl transfer. This was based on experimental observations that the secondary kinetic isotope effect (2° KIE) $k(CH_3)/k(CD_3)$ for enzyme-catalyzed methyl transfer was much more inverse than for the uncatalyzed reaction in solution, suggesting a "tighter" TS for the enzymic reaction. Mechanical compression by the enzyme destabilizes the RS more than the TS, resulting in a lower free-energy barrier. Whether or not this idea is valid has since been argued based on theoretical[16,17,1819,20] and experimental studies.[21] Thus, the controversial contribution of protein motions to modulate the DAD, originally proposed in enzymatic hydride transfer reactions,[9,10,11,12,13] has been extended to GNMT[22] and other related methyltransferases such as catechol *O*-methyltransferase (COMT).[23,24]

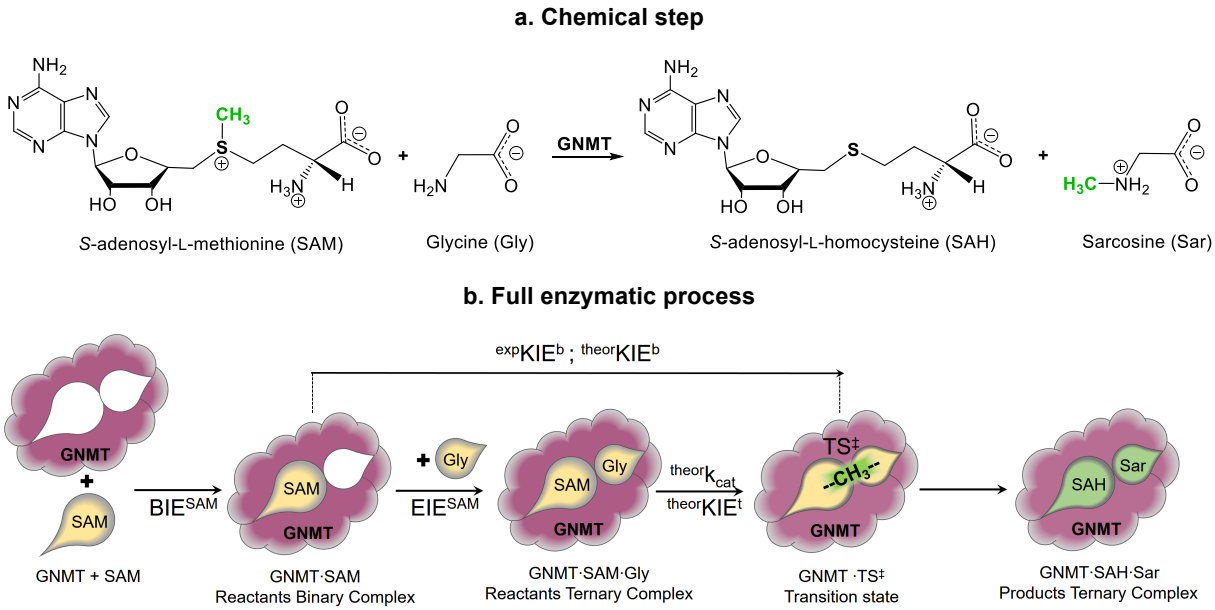


Figure 1. a) Schematic representation of the chemical step of the GNMT catalyzed reaction. b) Full catalyzed mechanism for the transformation from SAM into SAH catalyzed by GNMT. The particular isotope effect is indicated on each step.

In order to quantify and rationalize the origin of catalysis in enzymes, calculations can be carried out in the framework of Transition State Theory (TST).[25,26] Within this context, and including corrections for possible tunneling contributions and dynamic effects,[27,28] computed rate constants can be directly compared with those from kinetic experiments. In particular, computational QM/MM methods that combine quantum mechanics (QM) and molecular mechanics (MM) potentials have emerged as the most appropriate strategy for study of chemical reactions in very large flexible systems such as enzymes.

Several studies based on hybrid QM/MM methods have been carried out in our laboratories on enzymatic methyl transfer reaction catalyzed by COMT and, according to our studies, the origin of catalysis should be attributed to the electrostatic stabilization of the TS,[17,29,30] with a low coupling between the protein motions and the reaction coordinate on the chemical step of COMT.[31] In addition, determination of 2º KIEs for COMT catalyzed reaction did not support compression effects.[17] Regarding GNMT, our previous study on the rat-liver wild type enzyme (WT) showed how the binding energy is maximal for the TS[32] and the averaged DAD for the computed TS structures was actually smaller in solution than in the enzyme. These results rule out the compression along the donor-acceptor axis for these two methyl transferases. However, these conclusions have subsequently been questioned. The

experimentally measured reduction of the catalytic efficiency of GNMT when active-site residue Tyr21 was replaced by a series of mutants, as well as the change in the 2º α-^{3}H KIEs, was interpreted by Zhang and Klinman as a catalytically relevant reduction in the methyl DAD.[22]

Here we study the methyl transfer reaction catalyzed by WT rat-liver GNMT and its Y21F, Y21A and Y21G mutants by means of variational TST, and compare with the counterpart reaction in aqueous solution in order to shed light on the source of catalysis.

COMPUTATIONAL METHODS.

Set up of the models. The molecular system was prepared based on a crystal structure of WT rat-liver GNMT (PDB ID 1NBH) at 2.00Å resolution.[5] Since the standard pK_a values of ionizable groups can be shifted by local protein environments,[33] an accurate assignment of the protonation states of all these residues at pH = 7 was carried out. Recalculation of the pK_a values of the titratable amino acids has been done using the empirical PROPKA 2.0 program of Jensen *et al*.[34] Accordingly, Asp62 was protonated in OD2 position, His142, His264 and His286 were protonated in δ position, His58, His153, His174, His214, His245 were protonated in ε position, and finally His55 was protonated in both δ and ε position respectively.

Subsequently, in order to neutralize the system, one counterion (Cl^-) was placed into the optimal electrostatic position around the protein (that where the potential reaches maximum positive value). Afterwards, a series of optimization algorithms (steepest descent, conjugate gradient, and L-BFGS-B)[35] were applied. To avoid a denaturation or loss of structure prior to solvation of the protein, all the heavy atoms of the protein were restrained by means of a Cartesian harmonic umbrella with a force constant of 1000 kJ·mol^{-1}·Å^{-2}.

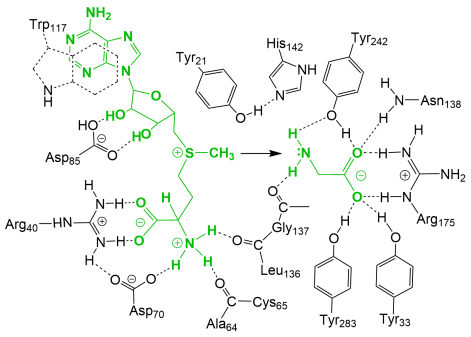


Figure 2. Schematic representation of the S_N2 methyl transfer reaction in the active site of GNMT. The entire SAM and glycine molecules (in green) were treated by QM potentials during the simulations, while classical force fields were used to describe the protein and solvent water molecules.

The protein was placed in a box of pre-equilibrated water molecules (100 × 80 × 80 $Å^3$), using the principal axis of the protein-inhibitor complex as the geometrical center. Any water with an oxygen atom lying in a radius of 2.8 Å from a heavy atom of the protein was deleted. The geometries of the remaining water molecules were then optimized. Later, an initial 500 ps classical MD simulation (at temperature 293 K), carried out to relax the system without significant changes in its geometry, was performed using the AMBER[36] and TIP3P[37] force fields for the protein and water molecules, respectively, as implemented in the fDYNAMO library.[38,39] The same force fields were used to describe the MM region in the subsequent hybrid QM/MM MD simulations. In this case, a small part of the system, consisting of the full SAM and glycine as depicted in Figure 2, was described by QM using either the AM1[40] semiempirical Hamiltonian or the M06-2X hybrid functional[41,42] with the standard 6-31+G(d,p) basis set. Due to the large number of degrees of freedom, any water molecule 20 Å away from any of the atoms of the substrate and cofactor was kept frozen in the remaining calculations. Cutoffs for the non-bonding interactions were applied using a force-switching scheme, within a range radius from 14.5 to 16 Å.

The same computational protocol was followed to prepare systems corresponding to the Y21F, Y21G and Y21A mutants. In these cases, the replacement of the original tyrosine residue on position 21 by Phe, Gly and Ala was performed on the X-ray structure and, consequently, the systems were equilibrated by hybrid QM/MM optimizations and MD simulations. Finally, a reference system in aqueous solution was prepared. The full SAM and substrate Gly molecules were embedded in a box of water molecules (100 × 80 × 80 $Å^3$,

19818 solvent water molecules in total) treated also by the TIP3P force field, as implemented in the fDYNAMO library. The same non-binding interaction conditions were applied to prepare the system in solution. Water molecules were equilibrated by means of hybrid AM1/MM MD in the reactant complex.

Long MD simulations of 10 ns were finally done for each of GNMT variants under NVT conditions with T = 293 K. In order to run long MD simulations NAMD software was used.[43] All models were described at MM level of theory using AMBER force field for enzyme and Gly substrate and contraions, TIP3P force field for water molecules. Missing parameters of SAM, determined using antechamber module available in AmberTools,[44,45] are reported in the Table S1 of Supporting Information. Because compression is a key aspect of the present problem, it is interesting to wonder whether the results are likely to be sensitive to employing an isobaric ensemble rather than the NVT ensemble. Simulations at higher pressure have shown that donor-acceptor distances in the hydride transfer reaction catalyzed by Morphinone Reductase can be reduced by 0.1-0.2 Å when the pressure is increased from 1 bar to 2000 bar.[46] However, the impact on the free energy barrier is expected to be small because the force constant associated to this distance at the reactants state is quite small. In the case of GNMT a similar trend could be expected. Instead, the effect of a similar pressure increase on the TS geometry must be quite limited considering that the force constant associate to the donor-acceptor distance increases from reactants to the TS. In fact, simulations of the reactant ternary complex of the wild-type enzyme under NPT conditions, at 1 and 1000 bar, rendered average structures with DAD statistically indistinguishable from the values derived from the NVT MD simulations (see Table S2 in Supporting Information).

Free energy surfaces. In order to get the free-energy landscape of the catalyzed reaction, potentials of mean force (PMFs) were traced along the selected reaction coordinate using the weighted histogram analysis method (WHAM)[47] combined with the umbrella sampling approach,[48] as implemented in fDYNAMO. In the present study, two–dimensional (2D) PMFs were generated by explicitly controlling the antisymmetric combination of distances defining the methyl transfer from the sulfur atom of SAM to the nitrogen atom of Gly, d(S-C) – d(C-N), ξ_1, and the distance between the donor and acceptor atoms, S and C, as a second reaction coordinate, ξ_2. The values of the variables sampled during the simulations are then pieced together to construct a distribution function from which the PMF is obtained as a function of the distinguished reaction coordinate (W(ξ)). Then, the activation free energy of a chemical step can be expressed as:[49]

$$\Delta G^{\ddagger}(\xi) = \left[W(\xi^{\ddagger}) + G_{\xi}(\xi^{R})\right] \quad (1)$$

where the superscripts indicate the value of the reaction coordinate at the reactants (R), and at the TS (‡), and G_{ξ} (ξ^{R}) is the free energy associated with setting the reaction coordinate to a specific value at the reactant state. Then, the quasiclassical activation free energy ΔG_{act}^{QC}, as appearing in Eq. 1, is calculated along the reaction coordinate ξ,[50,51] as the sum of $\Delta G^{\ddagger}(\xi)$ and the correction term due to the quantized nature of molecular vibrations (mainly zero-point energies).[50,52]

Because the generation of PMFs requires generation of a very large number of structures from QM/MM MD calculations, inevitably, we are restricted to the use of a semiempirical Hamiltonian (AM1) in this work. However, in order to improve the quality of our free-energy surfaces, based on the work of Truhlar and co-workers,[53,54,55] a spline under tension[56,57] is used to interpolate a correction term at any value of the reaction coordinates ξ_1 (and of ξ_2 in the case of 2D PMFs) selected to generate the free-energy surfaces.[58,59,60] In this way a new continuous energy function is constructed in order to obtain corrected PMFs:

$$E = E_{LL/MM} + S\left[\Delta E_{LL}^{HL}(\xi)\right] \quad (2)$$

where *S* refers to the spline function, and its argument $\Delta E_{LL}^{HL}(\xi)$ is a correction term evaluated from the single-point energy difference between a high-level (HL) and a low-level (LL) calculation of the QM subsystem. Herein the the M06-2X functional with the 6-31+G(d,p) basis set, as suggested by Truhlar and co-workers,[41,61] have been employed as the HL method. As our simulations demonstrate, the corrections to the energy can be significant, while the geometries generated from the sampling with LL calculations can be considered as representative of the HL geometries, thus giving credit to the selected technique. These calculations were carried out using the Gaussian09 program.[62]

Isotope Effects. Quasiclassical KIEs and equilibrium isotope effects (EIEs) were computed for isotopic substitutions of key atoms from ten couples of stationary structures, TSs and the enzyme:SAM binary complexes, at AM1/MM level of theory. Standard deviations on these KIEs were computed based on the averages over all possible combinations of 10 RC and 10 TS structures. KIEs were computed also at M06-2X/MM level of theory from 3 structures of the enzyme:SAM binary complex and 3 optimized TS, in each environment. The binding isotope effects (BIE)[63] were computed for the binding of SAM from aqueous solution to the active site of the protein to form the enzyme:SAM binary complex. Then, the ratio between

the rate constants, or equilibrium constants, corresponding to the light atom "L" and the heavier isotope "H" can be computed using the Transition State Theory (TST), as:

$$IE = \frac{\left(\frac{Q_b}{Q_a}\right)_L}{\left(\frac{Q_b}{Q_a}\right)_H} e^{-1/RT(\Delta ZPE_L - \Delta ZPE_H)} \quad (3)$$

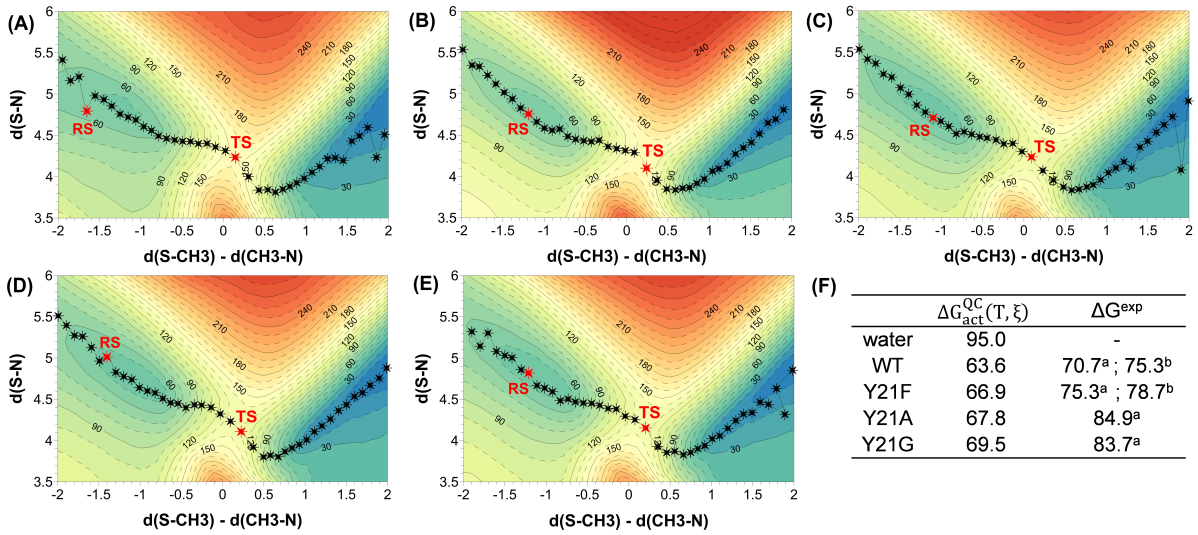
In eq. 3, the subscripts *H* and *L* refers to heavy and light isotopologs, respectively. The total partition function, *Q*, was computed as the product of the translational, rotational, and vibrational partition functions for the isotopologs in the two stationary structures under comparison, *a* and *b*. Thus, for KIE calculations *a* refers to RC and *b* to TS, and for BIE calculation a and *b* refer to the SAM in water and in the enzyme, respectively. Δ*ZPE* refers to the difference in the zero point energies between *a* and *b*. The subset of atoms used to define the Hessian for these IE calculations were those of the QM region, consistent with the "cut-off rule" and the local nature of isotope effects.[64] The contributions of translations, rotations and vibrations were assumed to be separable and were evaluated within the Born−Oppenheimer, rigid-rotor, and harmonic oscillator approximations. Keeping in mind that because both involved states, reactants and TS, are in a condensed media (the active site of a protein), contribution of translation and rotation to KIEs are negligible. Nevertheless, the full 3N × 3N Hessians have been subjected to a projection procedure to eliminate translational and rotational components, which give rise to small nonzero frequencies, as previously described.[64] Thus, it has been assumed that the 3N − 6 vibrational degrees of freedom are separable from the 6 translational and rotational degrees of freedom of the substrate.

Finally, as mentioned in the introduction, a qualitative change in KIEs is not expected if quantum tunneling effects were incorporating into the calculations, taking into account the mass of the transferring particle (a methyl group). A test was done with the Winger method confirming this prediction (average nuclear quantum correction of 1.003 was obtained for both isotopologs, cancelling out the effect on the KIE).

RESULTS AND DISCUSSION

The free-energy surfaces (FESs) generated from AM1/MM simulations for the reaction in water, WT GNMT and its Y21F, Y21A and Y21G mutants are shown in Figure 3, together with quasiclassical activation free energies $\Delta G_{act}^{QC}(T, \xi)$ computed at M06-2X:AM1/MM level. The first conclusion that can be derived from the FESs is that the free-energy barriers obtained for the reaction in the four proteins are close to each other, while that for the reaction in aqueous solution is much higher, by about 33 kJ·mol^{-1}: this is in agreement with the experimental observations that show how all variants of GNMT catalyze the reaction. The second important conclusion is that the mutations have only a small detrimental effect on the rate constant. An increase of 3.3 kJ·mol^{-1} in the free energy barrier is calculated for Y21F vs. WT, in very good agreement with the experimentally derived values of 4.6 and 3.3 kJ·mol^{-1} of Zhang and Klinman[22] and of Takata *et al*.,[5] respectively. The increases for Y21A and Y21G are higher (4.2 and 5.9 kJ·mol^{-1} relative to WT) but less than the experimental differences (14 and 13 kJ·mol^{-1}, respectively).[22] At this point, it is important to note that the experimental rate constants correspond to the catalytic turnover number under saturation of both the cofactor and substrate,[22] while our calculations correspond to the intrinsic chemical step from the ternary complex to the TS (see Figure 1b). Therefore, the predicted values can represent a lower limit, since other steps of the full catalytic process are not considered in the calculations and, consequently, a small underestimation with respect to experiment can be expected.

According to Figure 3, it is interesting to observe that although DAD diminishes from RS to TS, the minimum value is not achieved at the TS, in contrast to other reactions such as hydride transfer.[65] The DAD continues to decrease as the TS is passed, and only reaches its minimum value beyond the quadratic region of the FES. All the TS structures, optimized at M06-2X/MM (see Figure 4) or at AM1/MM show similar DAD values and the same conclusion is derived from the analysis of RS DADs (see Table S3 in Supporting Information).



	$\Delta G_{act}^{QC}(T,\xi)$	ΔG^{exp}
water	95.0	-
WT	63.6	70.7[a] ; 75.3[b]
Y21F	66.9	75.3[a] ; 78.7[b]
Y21A	67.8	84.9[a]
Y21G	69.5	83.7[a]

Figure 3. Summary of QM/MM kinetic data for the methyl transfer from SAM to glycine. QM/MM AM1/MM FESs for the S_N2 methyl transfer reaction in (A) aqueous solution, and catalyzed by (B) WT GNMT, (C) Y21F GNMT, (D) Y21A GNMT and (E) Y21G GNMT. Averaged values of the coordinates obtained along monodimensional FESs (Figure S3 of Supporting Information) are shown on the surfaces as black stars, with the RS and TS location marked in red. Distances are in Å and values of isoenergetic contours are in kJ·mol^{-1}. (F) Summary of kinetic data: quasiclassical activation free energies $\Delta G_{act}^{QC}(T,\xi)$ computed at M06-2X:AM1/MM level, together with the corresponding free energy barriers ΔG^{exp} deduced from experimental rate constants within the TST; [a] data from ref 22, and [b] data from ref 5. All barriers energies (in kJ·mol^{-1}) were obtained at 293 K.

The free energy of activation does not correlate directly with the TS DAD distance, nor with the change in DAD between the RS and TS. In addition, the average values of the S-C-N angle in RS and TS in the different proteins (see Table S3 in Supporting Information) show negligible differences, with no clear correlation between the free energy of activation of wild type and mutants with the deviations from linear trajectory. Nevertheless, it is noticeable how despite the angle in the TS of the different media are coincident (S-C-N angle ranging from 170 to 172 degrees), in RS all enzymes preorganize the donor and acceptor moieties in a significant more linear orientation (S-C-N angle ranging from 162 to 168 degrees) than the aqueous solution (S-C-N angle ca. 129 degrees) which, obviously facilitates the S_N2 reaction. Moreover, analysis of the position of Tyr21 in WT GNMT, in either the simulated or the X-ray crystallographic structure, reveals that this residue cannot be responsible for any compression effect (see Figure 4). The O-S-N angle (measured between the O atom of the amino-acid hydroxyl group, the S atom of SAM, and the N atom of the glycine) computed at M06-2X/MM level for WT GNMT is 67.5° in the X-ray structure, and 72.8° and 60.1° in the RS and TS structures, respectively. These values are far from the ca. 180° angle required for Tyr21 to be positioned behind the SAM cofactor in the active site in order to exert

compression on the DAD. It appears that the catalytic role of the enzyme cannot be associated with a mechanical compression effect along the donor-acceptor axis.

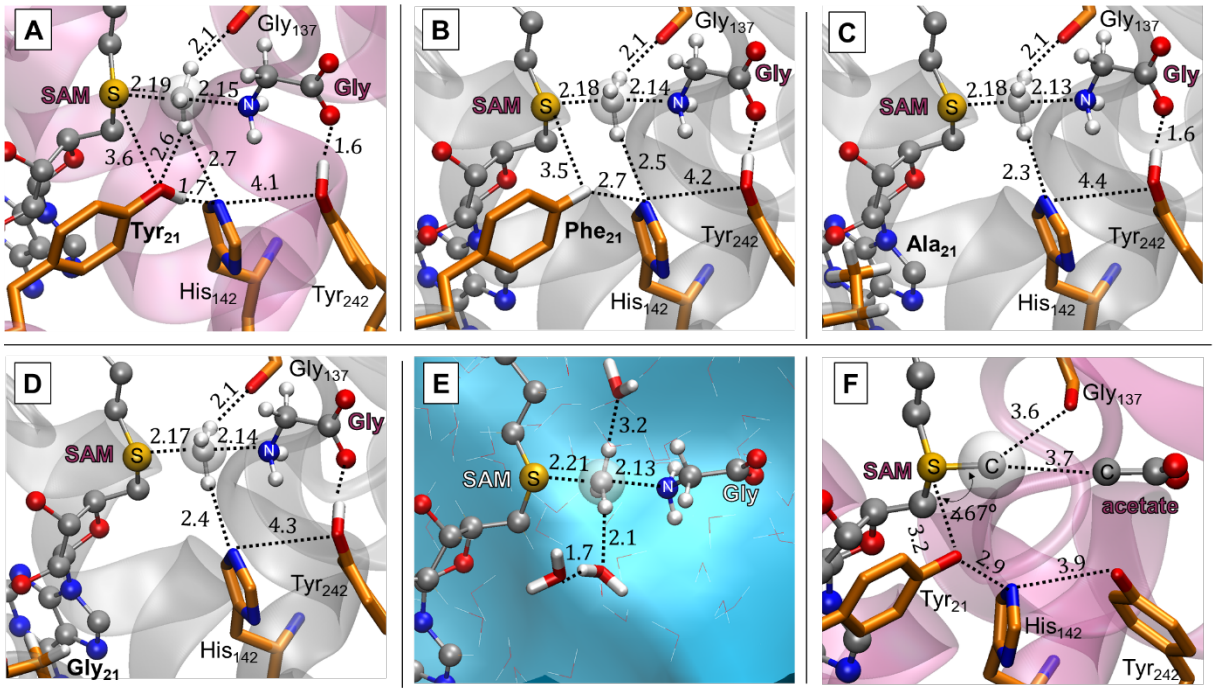


Figure 4. Depiction of the catalytic residues at the active site from structures of the TSs optimized at M06-2X/MM level in A) WT GNMT; B) Y21F GNMT; C) Y21A GNMT; D) Y21G GNMT and (E) aqueous solution. Panel (F) corresponds to the active site obtained from the X-ray structure of rat GNMT (PDB 1NBH). DAD values are equal to 4.34, 4.31, 4.30, 4.30 and 4.33 Å for the WT, Y21F, Y21A, Y21G and aqueous solution, respectively. Distances are in Å and the angle in degrees.

Isotope Effects.

Our results show a good qualitative agreement between the experimental and computational $^{1}H_3/^{3}H_3$ binding isotope effects (BIEs) for all four GNMT variants (see Table 1). These normal BIEs (values larger than unity) derive from stiffer vibrational modes associated with these H atoms in aqueous solution than in the enzyme active site. Geometrical analysis of the optimized structures shows that these H atoms make more and tighter interactions with water molecules in solution than with the residues in the GNMT:SAM binary complex of any of the protein variants (see Figures S4 and S5 in Supporting Information). This is reasonable since the active site must present a cavity before it can accommodate the binding of glycine.

Table 1. 1º $^{12}C/^{14}C$ KIEs and 2º $^{1}H_3/^{3}H_3$ KIEs between the enzyme:SAM binary complex and the TS ($^{theor}KIE^{b}$ in Figure 1b). $^{12}C/^{14}C$ BIEs and $^{1}H_3/^{3}H_3$ BIEs for the binding of SAM into the active site of the WT enzyme and the mutants Y21F, Y21A and Y21G. All values computed at AM1/MM level of theory at 293 K, as an average of all pairwise combinations of 10 RS structures with 10 TS structures. Results are shown as mean ±SD. [a]Experimental data, with errors, are taken from ref 22.

	2º $^{1}H_3/^{3}H_3$ KIE[b]		1º $^{12}C/^{14}C$ KIE[b]	
	Experiment[a]	Theory	Experiment[a]	Theory
water	ND	0.758 ± 0.027	ND	1.110 ± 0.008
WT	0.787 ± 0.005	0.749 ± 0.016	1.117 ± 0.008	1.105 ± 0.003
Y21F	0.809 ± 0.006	0.746 ± 0.017	1.119 ± 0.013	1.106 ± 0.003
Y21A	0.832 ± 0.013	0.755 ± 0.017	1.132 ± 0.007	1.106 ± 0.005
Y21G	0.867 ± 0.013	0.728 ± 0.018	1.146 ± 0.023	1.099 ± 0.004

	$^{1}H_3/^{3}H_3$ BIE		$^{12}C/^{14}C$ BIE	
	Experiment[a]	Theory	Experiment[a]	Theory
WT	1.029 ± 0.013	1.026 ± 0.036	ND	0.999 ± 0.008
Y21F	1.041 ± 0.018	1.037 ± 0.037	ND	0.999 ± 0.008
Y21A	1.031 ± 0.012	1.026 ± 0.038	ND	0.996 ± 0.008
Y21G	1.021 ± 0.044	1.043 ± 0.041	ND	1.000 ± 0.008

The calculations do indeed predict inverse 2º $^{1}H_3/^{3}H_3$ KIEs for all the proteins and slightly less inverse in aqueous solution as measured experimentally, but no meaningful trend can be observed (within the estimated 1σ uncertainty) between the GNMT variants. The origin of the inverse 2º $^{1}H_3/^{3}H_3$ KIEs must be related with an increase of the force constants associated to the methyl group hydrogen atoms from the GNMT:SAM binary complex to the TS (see Table S7 in Supporting Information), which is also reflected in the evolution of the C-H bond lengths diminishing from the GNMT:SAM:Gly ternary complex (RS) to the TS (see Figure S7 in Supporting Information). Among the four proteins, the WT stabilizes the TS better than the mutants (as reflected by the predicted and experimentally deduced free energy barriers) which can be related *a priori* with the loss of the stabilizing interaction between the methyl group and the hydroxyl group of Tyr21 in the WT after mutation. However, as shown in Figure 4, the interaction between a H atom of the methyl group and His124 appears to be stronger in mutants (shorter inter-atomic distance in the three mutants). In addition, a second H atom of the transferring methyl group interacts with the backbone carbonyl group of Gly137 in all the optimized TS structures with an O-C distance in general slightly shorter than in the GNMT:SAM binary complex (see Figure S5 in Supporting Information). Apparently these interactions are thus compensating for the loss of the interaction with Tyr21 and

explaining why the $^{1}H_3/^{3}H_3$ BIEs and 2º KIEs do not differ sensibly between WT and the mutants, within the standard deviations.

The computed 1º $^{12}C/^{14}C$ KIEs are normal and in an almost quantitative agreement with the experimental results, and the $^{12}C/^{14}C$ substitution produces no BIEs; but no experimental data were reported for this substitution to be compared with.

Electrostatic effects. The charges on the key atoms for donor, acceptor and the transferring methyl group (computed with the CHELPG method)[66] change in the same way along the reaction coordinate (r.c.) in aqueous solution and in WT GNMT (Figure 5). The charge on the donor S atom varies from slightly positive in the RS to slightly negative in the TS, in both media. The acceptor N atom is always more negatively charged in the enzyme than in solution, and the sum of the methyl-group atomic charges is always more positive in the enzyme than in solution. So it seems that the protein environment favours a TS with larger separation of charges between the transferring methyl group and the accepting N atom. The evolution of the charges on the three mutants along the r.c., deposited in the Supporting Information (see Figure S8 in the Supporting Information), are almost coincident with the WT shown in Figure 5. It is interesting that the methyl-group charge does not reach its maximum in the TS but after passing this structure. This can be related not only to the evolution of the DAD discussed above, but to the evolution of the electrostatic properties of the different environments. Thus, as shown in Figure 6, the averaged electrostatic potential created by each of the different environments at the three key atoms changes from the RS (r.c. = -1.5 Å) to the point where DAD reaches its minimum (r.c. ≈ +0.5 Å). All environments appear to be equilibrated from this point on. Consider the three key atoms. At the N atom the potential due to each environment changes from positive in the RS to negative in the PS. Since this atom is itself negatively charged in the RS and the magnitude of its charge decreases as the reaction proceeds, this means that all environments, from the point of view of this atom, stabilize the RS rather than the TS or PS. Similarly, since the charge on the S atom evolves from slightly positive in RS to slightly negative in PS, and the potential becomes less negative as the reaction proceeds, all environments also stabilize the RS. However, the negative potential created by each environment at the C atom stabilizes the positive charge developed on the transferring methyl group in the TS. Interestingly, there is a linear correlation between this potential and the calculated free energy barriers (more negative potentials correspond to smaller barriers) which corroborates the electrostatic origin of catalysis in GNMT (see Figure 6F). The WT enzyme is the environment that generates the most negative potential at the TS

thus stabilizing better the positive charge developed on the methyl group. This means that WT provides the most favourable environment within which to transfer the methyl group from S to N (see Table S8 in the Supporting Information for the full list of averaged electrostatic potentials). On the other hand, analysis of the contributions to the total electrostatic potential by residues supports the small differences observed in the four proteins. While the potential generated by residue 21 is less favourable to stabilize the positive charge developed in the methyl group in the TS after mutations, the potential generated by His142 is more favorable in the mutants than in the WT.

Comparison of the plots presented in Figure 6 reveals an important point. Whereas the change from RS to PS in the electrostatic potential due to the environment is quite different in water as compared to any of the proteins, these changes are all very similar for the WT enzyme and the three mutants. This accords with a higher energy of reorganization from RS to TS in aqueous solution, and with only small differences in enzyme reactivity consequent upon mutation of Tyr21. Thus, *a priori,* the catalytic effect of GNMT could be partially attributed to the generation of a non-polar cavity: since the reaction catalyzed by GNMT implies annihilation of charges from reactants to products, a polar medium (like water) would preferentially stabilize the reactants and consequently higher barriers should be expected. Nevertheless, our calculations reveal that this simplified picture is not complete, since the enzyme is generating a potential that facilitates the positively charged methyl transfer and the stabilization of the TS. Catalysis in GNMT is not only due to the absence of a strong reaction field (as the one felt in water by the reacting system) but also to the existence of a preorganized electric field favouring the methyl transfer. Indications of this catalytic effect have been previously detected in the methyl transfer reaction catalyzed by COMT,[30] and in other completely different proteins, such as the HIV1 PR[67] or *de novo* designed enzymes.[68,69]

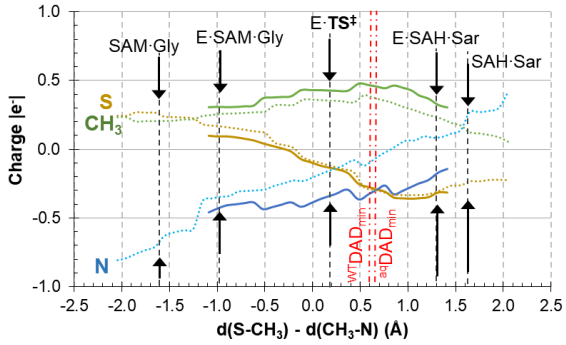


Figure 5. Evolution of charges (in a.u.) of the donor atom, S, acceptor atom, N, and the sum of the charges of the transferring methyl group, CH_3, along the reaction path. Results derived from the reaction in aqueous solution (dotted lines) and in the WT GNMT (solid lines). Vertical black dashed lines indicate the position of the reactant complex in aqueous solution (SAM-Gly) and in the enzyme (E-SAM-Gly), the TS and the products complex in the enzyme (E-SAH-Sar) and in aqueous solution (SAM-Sar). The position of the minimum DAD in aqueous solution (0.63 Å) and in the wild-type GNMT (0.59 Å) are presented as dashed red lines.

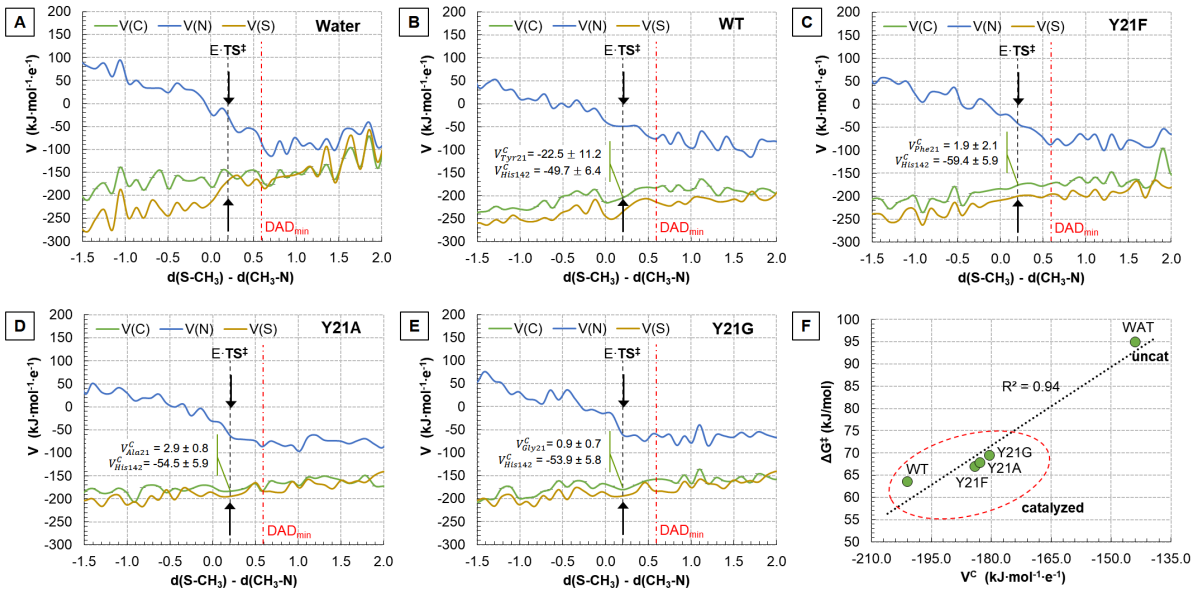


Figure 6. Evolution of electrostatic potential, V (in $kJ \cdot mol^{-1} \cdot e^{-1}$), created by the different environments on the donor atom, S (orange lines), acceptor atom, N (blue lines), and on the C atom of the transferring methyl group (grey lines), along the reaction path. A) in aqueous solution; B) WT; C) Y21F; D) Y21A); and E) Y21G. Average values of the contribution of residue 21 and 142 to the electrostatic potential created by the proteins on the C atom in the TS are reported in $kJ \cdot mol^{-1} \cdot e^{-1}$. F) Relationship between activation free energies and the averaged electrostatic potential created by the environment (protein or water molecules) on the carbon atom of the transferring methyl group at the TS structures (R^2 = 0.94 for linear regression).

CONCLUSIONS

The first important conclusion derived from our QM/MM study is that the activation free energy of the reaction in solution is significantly higher than in any of the enzymes, which reflects the catalytic role of the proteins. Although the differences are small, the WT enzyme shows the lowest activation free energy barrier while Y21A and Y21G mutants lead to the highest values. These results are in very good agreement with the values that can be derived from the experimentally measured rate constants by Takata *et al*.[5] and by Zhang and Klinman.[22]

The analysis of the average DAD in RS and TS does not provide any indication about the compression effect as a source of catalysis since similar values are obtained in all environments. In fact, structural analysis of the active site shows that Tyr21 cannot push the S donor atom towards the N acceptor atom.

The more inverse 2º $^{1}H_3/^{3}H_3$ KIEs observed for the WT GNMT than for the mutants has been interpreted as being due to a "compactness/tightness" of its active site as compared with the mutated variants, which have diminished ability to achieve a constrained DAD.[22] In contrast, our calculations point to another scenario. The hydroxyl O atom of Tyr21, together with an N atom of His142, interact with the H atoms of the methyl group in the TS of WT GNMT. In the mutants the loss of the interaction with the hydroxyl group of Tyr21 is compensated by stronger interactions established between the H atoms of the transferring methyl group and His142 that is occupying the cavity. This plasticity of the enzyme active site explains why the differences between variants of GNMT are so small.

Finally, it is possible to rationalize the reactivity of the different environments from the electrostatic point of view. Our simulations show how the methyl-group charge reaches a maximum after passing the TS and decreases again in PS. This is related not only to the evolution of the DAD, that reaches the minimum value at this point, but to the evolution of the averaged electrostatic potential created by the different environments at the S donor, N acceptor and C methyl atoms. The potentials change from the RS to the point with a value of the r.c. around +0.5 Å and it remains stable from this point on in all environments. The changes from RS to TS are more dramatic in water. The WT and mutants behave similarly, in agreement with the close values of computed activation free energies. The WT enzyme is the protein that generates the most favorable electrostatic environment to stabilize the charge developed on the methyl group in the TS, and thus is the most favourable environment to catalyze the reaction. The detrimental effect of substitution of Tyr21 on the electrostatic potential is partially compensated by His142 that, by approaching to the methyl group,

generates a higher potential. The changes on the electrostatic potential exerted by the different proteins on the substrate reflect that they are not static but dynamic macromolecules that have to evolve structurally from the RS to the TS, although at a low energy cost. The plasticity of the WT protein is responsible for the differences in the KIEs and rate constants after mutations being as small as they are.

SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website at DOI:

Parameters of SAM; plots reporting details of molecular dynamic simulations, 1D PMFs computed at AM1/MM level; additional tables and figures containing structural data; Figures of the change in C-hybridization along reaction coordinate; additional tables with KIEs, BIEs and force constants; additional tables and figures with evolution of charges along the reaction and electrostatic potential; and coordinates of TSs.

Acknowledgements

This work was supported by the Spanish Ministerio de Economía y Competitividad and FEDER funds (project CTQ2015-66223-C2), Generalitat Valenciana (PROMETEOII/2014/022), Universitat Jaume I (project UJI·B2017-31) and the USA National Institute of Health (Ref No. NIH R01 GM065368). V.M. is grateful to the University of Bath for the award of a David Parkin Visiting Professorship and to the Spanish Ministerio de Educación Cultura y Deporte for travel financial support (project PRX17/00166). Authors acknowledge computational resources from the Servei d'Informàtica of Universitat Jaume I.

Competing Financial Interests Statement

The authors declare no competing financial interests.

REFERENCES

1 Blumenstein, J.; Williams, G. R. *Biochem. Biophys. Res. Commun.* **1960**, *3*, 259–263.
2 Kerr, S. J.; Heady, J. E. *Adv. Enzyme Regu.l* **1974**, *12*, 103–117.
3 Seshadri, S.; Beiser, A.; Selhub, J.; Jacques, P. F.; Rosenberg, I. H.; D'Agostino, R. B.; Wilson, P. W.; and Wolf, P. A. *N. Engl. J. Med.* **2002**, *346*, 476–483.
4 Konishi, K.; Fujioka, M. *Biochemistry* **1987**, *26*, 8496–8502.
5 Takata, Y.; Huang, Y. F.; Komoto, J.; Yamada, T.; Konishi, K.; Ogawa, H.; Gomi, T.; Fujioka, M.; Takusagawa, F. *Biochemistry* **2003**, *42*, 8394–8402.
6 Luka, Z. *Vitam. Horm.* **2008**, *79*, 325–345.
7 Pauling, L. *Chem. Eng. News* **1946**, *24*, 1375–1377.
8 Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210–3235.
9 Nagel, Z. D.; Klinman, J. P. *Nat. Chem. Biol.* **2009**, *5*, 543–550.
10 Knapp, M. J.; Klinman, J. P. *Eur. J. Biochem.* **2002**, *269*, 3113–3121.
11 Antoniou, D.; Caratzoulas, S.; Kalyanaraman, C.; Mincer, J. S.; Schwartz, S. D. *Eur. J. Biochem.* **2002**, *269*, 3103–3112.
12 Basran, J.; Sutcliffe, M. J.; Scrutton, N. S. *Biochemistry* **1999**, *38*, 3218–3222.
13 Scrutton, N. S.; Basran, J.; Sutcliffe, M. J. *Eur. J. Biochem.* **1999**, *264*, 666–671.
14 Kohen, A. *Acc. Chem. Res.* **2015**, *48*, 466–473.
15 Olsen, J.; Wu, Y. S.; Borchardt, R. T.; Schowen, R. L. *Trans- methylation*. Usdin E, Borchardt RT, Creveling CR Eds, Elsevier, New York, 1979.
16 Moliner, V.; Williams, I. H. *J. Am. Chem. Soc.* **2000**, *122*, 10895–10902.
17 Ruggiero, G. D.; Williams, I. H.; Roca, M.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2004**, *126*, 8634–8635.
18 Kanaan, N.; Ruiz-Pernía, J. J.; Williams, I. H. *Chem. Commun.* **2008**, 6114–6116.
19 Lameira, J.; Bora, R. P.; Chu, Z. T.; Warshel, A. *Proteins: Struct. Funct. Genet.* **2015**, *83*, 318–330.
20 Dua, Q.; Wang, Z.; Schramm, V. L. *Proc. Nat. Acad. Sci. USA* **2016**, *113*, 2916–2921.
21 Boyd, R. J.; Kim, C. K.; Shi, Z.; Weinberg, N.; Wolfe, S. *J. Am. Chem. Soc.* **1993**, *115*, 10147–10152.
22 Zhang, J.; Klinman, J. P. *J. Am. Chem. Soc.* **2016**, *138*, 9158–9165.
23 Zhang, J.; Klinman, J. P. *J. Am. Chem. Soc.* **2011**, *133*, 17134– 17137.
24 Zhang, J.; Kulik, H. J.; Martinez, T. J.; Klinman, J. P. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7954–7959.
25 Glasstone, S.; Laidler, K. J.; Eyring, H. Diffusion and Electrochemical Phenomena, The Theory of Rate Processes: The Kinetics of Chemical Reactions, Viscosity, McGraw-Hill, New York, 1941.
26 Truhlar, D. G.; Hase, W. L.; Hynes, J. T. *J. Phys. Chem.* **1983**, *87*, 2664–2682; erratum: **1983**, *87*, 5523.
27 Truhlar, D. G. Arch. Biochem. Biophys. **2015**, *582*, 10–17.
28 Truhlar, D. G.; Garrett, B. C. Acc. Chem. Res. **1980**, *13*, 440–448.
29 Roca, M.; Martí, S.; Andrés, J.; Moliner, V.; Tuñón, M.; Bertran, J.; Williams, I. H. *J. Am. Chem. Soc.* **2003**, *125*, 7726–7737.
30 Roca, M.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertran, J. *J. Am. Chem. Soc.* **2005**, *127*, 10648–10655.
31 Roca, M.; Moliner, V.; Tuñón, I.; Hynes, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 6186–6193.
32 Soriano, A.; Castillo, R.; Christov, C.; Andrés, J.; Moliner, V.; Tuñón, I. *Biochemistry* **2006**, *45*, 14917–14925.

33 Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415–436.
34 Hui, L.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, **61**, 704–721.
35 Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. *J. Sci. Comp.* **1995**, *16*, 1190–1208.
36 Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
37 Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
38 Field, M. J. *A practical Introduction to the Simulation of Molecular Systems*. Cambridge University Press: Cambridge, *U.K.* 1999.
39 Field, M. J.; Albe, M.; Bret, C.; Proust-de Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088–1100.
40 Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
41 Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
42 Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
43 Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comp. Chem.* **2005**, *26*, 1781–1802.
44 Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graph. Mod.* **2006**, *25*, 247260.
45 Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comp. Chem.* **2004**, *25*, 1157–1174.
46 Hay, S.; Scrutton, N. S. Nat. Chem. 2012, 4, 161-168.
47 Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
48 Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
49 Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 5828–5833.
50 Alhambra, C. Corchado, J.; Sánchez, M. L.; Garcia-Viloca, M.; Gao, J.; Truhlar, D.G. *J. Phys. Chem. B* **2001**, *105*, 11326-11340.
51 Garcia-Viloca, M.; Alhambra, C.; Truhlar, D. G.; Gao, J. *J. Comput. Chem.* **2003**, *24*, 177–190.
52 Pu, J. Z.; Gao, J. L.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3140–3169.
53 Nguyen, K. A.; Rossi, I.; Truhlar, D. G. *J. Chem. Phys.* **1995**, *103*, 5522–5530.
54 Corchado, J. C.; Coitiño, E. L.; Chuang, Y.; Fast, P. L.; Truhlar, D. G. *J. Phys. Chem. A.* **1998**, *102*, 2424–2438.
55 Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A.* **1999**, *103*, 1140–1149.
56 Renka, R. J. *SIAM J. Stat. Comput.* **1987**, *8*, 393–415.
57 Renka, R. J. *ACM Trans. Math. Software* **1993**, *19*, 81–94.
58 Ruiz-Pernia, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. *J. Phys. Chem. B.* **2004**, *108*, 8427–8433.
59 Roca, M.; Moliner, V.; Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I. *J. Phys. Chem. A.* **2006**, *110*, 503–509.
60 Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S. *J. Phys. Chem. B.* **2006**, *110*, 17663–17670.
61 Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A.* **2003**, *107*, 1384–1388.
62 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.;

Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.,Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.;Cossi, M.;Rega, N.;Millam,N.J.;Klene,M.;Knox,J.E.;Cross,J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma,K.; Zakrzewski, V. G.; Voth, G. A., Salvadör, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.;Fox, D. J. Gaussian09, Gaussian, Inc.: Wallingford, CT, 2009.

63 Świderek, K.; Paneth, P. *Chem. Rev.* **2013**, *113*, 7851–7879.

64 Ruggiero, G. D.; Guy, S. J.; Martí, S.; Moliner, V.; Williams, I. H. *J. Phys. Org. Chem.* **2004**, *17*, 592–600.

65 Vardi-Kilshtain, A.; Major, D. T.; Kohen, A.; Engel, H.; Doron, D. *J. Chem. Theory Comput.* **2012**, *8*, 4786–4796.

66 Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.

67 Krzemińska, A.; Moliner, V.; Świderek, K.; *J. Am. Chem. Soc.* **2016**, *138*, 16283–16298.

68 Świderek, K.; Tuñón, I.; Moliner, V.; Bertran, J. *Chem. Eur. J.* **2017**, *23*, 7582–7589.

69 Vaissier, V.; Sharma, S. C.; Schaettle, K.; Zhang, T.; Head-Gordon, T. *J. Am. Chem. Soc.* **2017**, *139*, 18409−18427

Table of Contents (TOC)