Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2018

Electronic Supplementary Information (ESI):

Exposing Catalytic Versatility of GTPases:

Taking Reaction Detours in Mutants of hGBP1 Enzyme without Additional Energetic Cost

Ravi Tripathi, Jan Noetzel and Dominik Marx

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

1 RMSD analysis

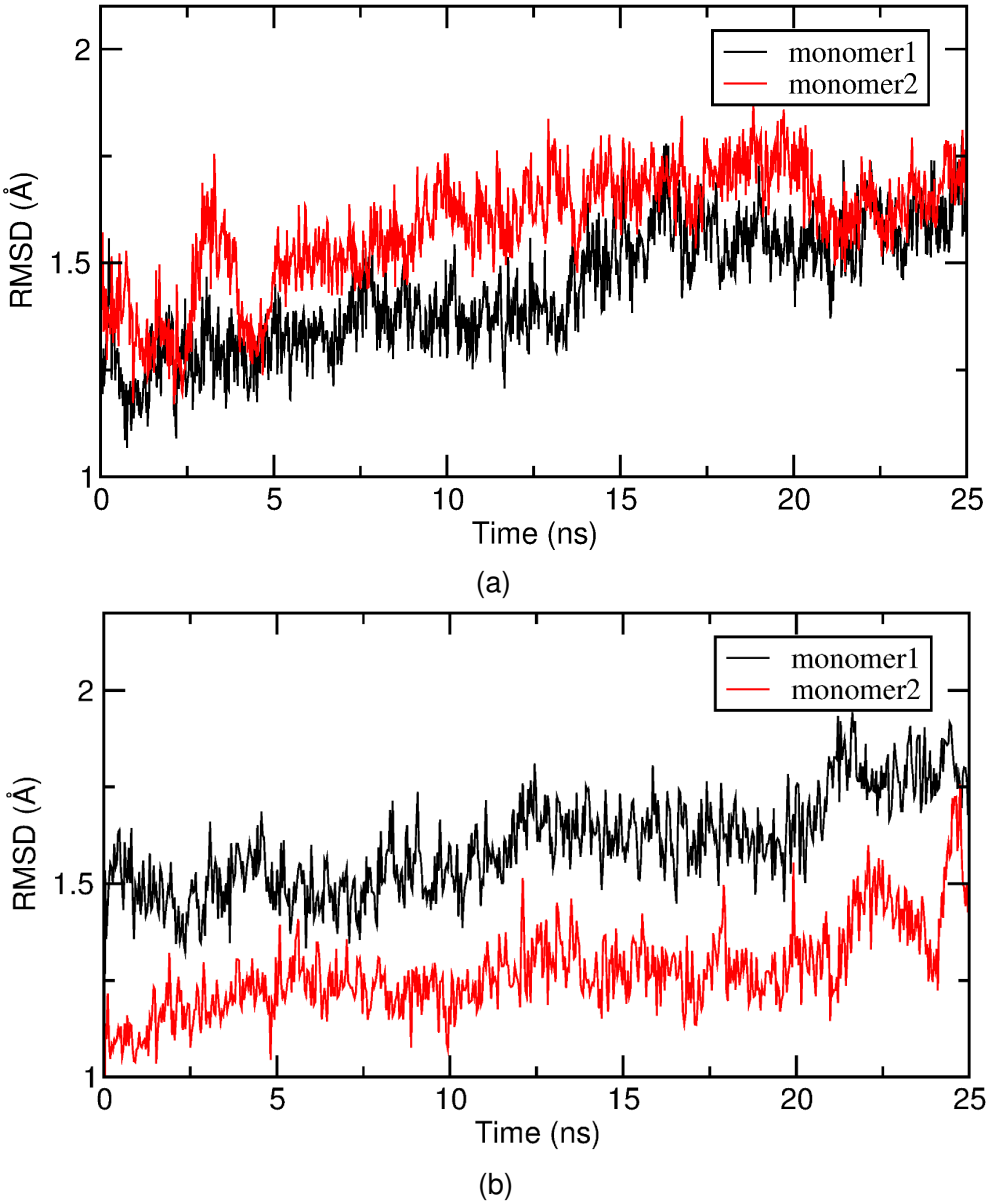

Fig. ESI 1: Backbone RMSD of protein backbone with respect to the starting structure during the NVT simulations of (a) E99A and (b) S73A. The residues 1-7, 157-173 and 305-314 are excluded in calculating the RMSD due to either their non availability in the X-ray structure or their high flexibility as being the terminal residues of the protein chains.

2 QM/MM simulations

The full details of our QM/MM simulations have already been provided in Ref. 1, however, we reproduce them here as a courtesy to the reader. The QM part of the system was dealt with using fully self-consistent Kohn-Sham density functional theory employing the BLYP functional [2,3]. The TZV2P-GTH Gaussian basis sets were utilized together with GTH norm-conserving pseudopotentials [4]. The GEEP (Gaussian Expansion of Electrostatic Potential) method was used to handle the Coulomb interaction between the QM and MM atoms [5,6]. The boundaries of the QM/MM interface were dealt with using the IMOMM link scheme [7], in which the cut bond is saturated with a hydrogen in order to maintain neutrality. The cut positions were always chosen between carbon-carbon single bonds for all QM/MM links in order to minimize the polarization artifacts near the boundary. The temperature of the system was kept fixed at $T = 300\,\mathrm{K}$ using Nosé-Hoover chain thermostats [8] and a timestep of 0.5 fs was employed.

3 Validating the electronic structure method

One of expected sources of error in any QM/MM-based free energy calculation is the employed electronic structure method that is used to treat the electronic structure within the QM part, which is the BLYP density functional in the present case. In our previous work[1] we have already shown for the wild type enzyme that the BLYP results are in good agreement with the ones obtained using a well-established correlated wavefunction-based method, namely SCS-(RI-)MP2 [9–11]. In order to reconfirm the validity and reliability of BLYP functional, we now compare the performance of BLYP with even more accurate coupled cluster data obtained by using the DLPNO-CCSD(T) method[12] together with the recommended def2-TZVPP [13] basis set, all as implemented in the ORCA program package[14]. To this end, we follow the same protocol as introduced previously[1] for the purpose of benchmarking: We have sampled representative configurations of the reactant (**1**), transition (**2**) and product (**3**) states using now the E99A mutant for GTP hydrolysis and have computed the electronic energy differences between these configurations as reported in Fig. ESI 2. Note that due to the efficiently of the ORCA implementation of the CCSD(T) method, we are not

forced to reduce the system size but can take into account all atoms in the full QM domain as treated by BLYP in the dynamical QM/MM simulations!

As repeatedly demonstrated[1, 15–17] for different classes of (bio)chemical reactions, the BLYP energies are found to be reliable without any need to adjust parameters as often required when using semi-empirical or even empirical electronic structure methods within QM/MM molecu-lar dynamics. In the present case, the BLYP data are shown be in good agreement with those obtained using the current "gold standard in quantum chemistry", CCSD(T), as our benchmark method, featuring a maximum difference of ≈ 2 kcal/mol on the relevant energy scale of about 20 kcal/mol. Therefore, the present validation demonstrates – again[1, 15–17] – that the BLYP density functional is reliable in providing trustworthy results for GTP hydrolysis, being very comparable to those obtained from state-of-the-art correlated wavefunction-based methods, such as SCS-MP2 and even CCSD(T), which clearly cannot be used within multi-dimensional QM/MM-based free energy sampling simulations.

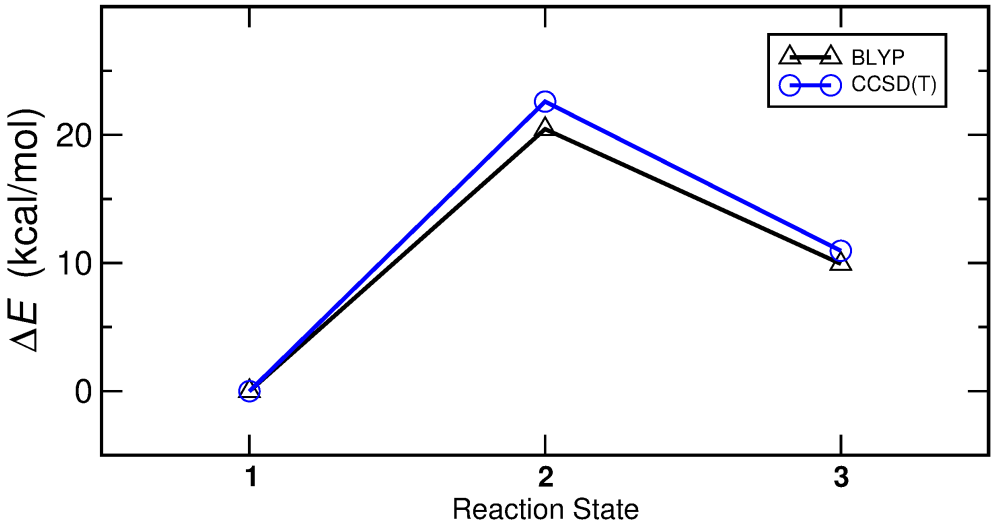

Fig. ESI 2: Comparison of electronic energy differences obtained from sampled configurations representing reactant (**1**), transition (**2**) and product (**3**) states for GTP hydrolysis in the E99A mutant of hGBP1 according to BLYP (black triangles) versus CCSD(T) (blue circles) calculations.

4 Metadynamics parameters and coordination number

The Gaussian potentials of heights varying between ≈ 1 to $\approx 2\,k_{\mathrm{B}}T$ were used to sample the free energy landscapes whereas all remaining parameters were the same as used in our study of the wild type[1].

The coordination number (CN) between two sets of atoms A and B, denoted as $C[\mathrm{A}-\mathrm{B}]$, is defined as

$$C[\mathrm{A}-\mathrm{B}] = \sum_{I\in\mathrm{A}}\sum_{J\in\mathrm{B}} \frac{1-\left(\frac{d[I-J]}{d^0_{\mathrm{AB}}}\right)^p}{1-\left(\frac{d[I-J]}{d^0_{\mathrm{AB}}}\right)^{p+q}} \quad . \tag{1}$$

Here, $d\,[I-J]$ is the distance between the two nuclei I and J belonging to the set of atoms A and B, respectively, d^0_{AB} is a fixed parameter which is defined based on the nature of the involved atoms, whereas p and q (being both set equal to 6) are constants that determine the steepness of the coordination number function. The value of d^0_{AB} is chosen to be 2.0 Å for the CN defining the P_γ-O_{LG} bond, 2.5 Å for the O_{Nu}-P_γ bond, and 1.5 Å for all other CN functions.

5 Proton transfer from Asp112 to HPO_4^{2-}

The proton transfer event from Asp112 to HPO_4^{2-}, leading to the final product $H_2PO_4^-$, was sampled by employing two CVs in the QM/MM metadynamics simulation. To study this reaction step, one CV (called CV3 herein) was defined as C[O2-H] + C[O1-H2] − C[$O_{\gamma,\mathrm{Asp112}}$ -H] − C[O2-H2], whereas the other one (CV4) is C[O_γ - $H_{\gamma,\mathrm{Ser73}}$] + C[$O_{\gamma,\mathrm{Ser73}}$ - H1] − C[$O_{\gamma,\mathrm{Ser73}}$ - $H_{\gamma,\mathrm{Ser73}}$] − C[O1-H1]. Importantly, these two CVs, which are based on a total of 8 CN different functions, allow for the breaking and making of multiple bonds during QM/MM metadynamics sampling that ultimately govern long-distance proton transfer from Asp112 to HPO_4^{2-} via a hydrogen-bonded network. This setup thus allowed us to successfully simulate multi-step proton transfer from Asp112 to the far-distant HPO_4^{2-} species.

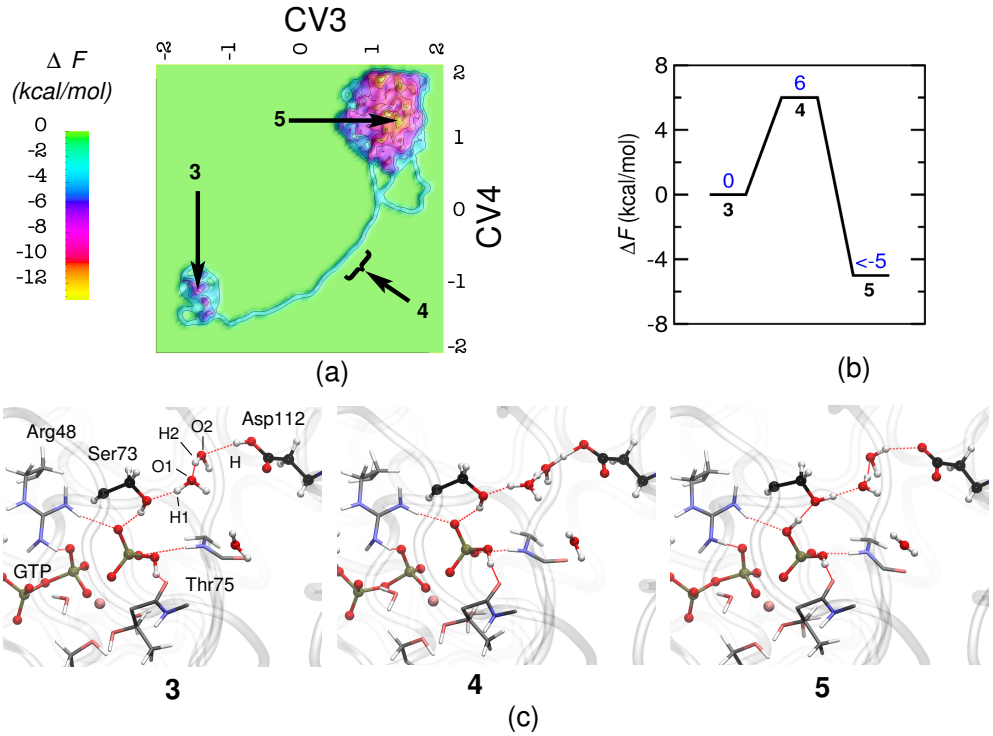

Fig. ESI 3: (a) Free energy surface and (b) corresponding free energy profile obtained for the proton transfer event from the protonated Asp112 residue to HPO_4^{2-}, thus leading to the final product $H_2PO_4^-$. Representative snapshots of active site residues corresponding to the minima and TS structure are depicted in (c).

6 Mechanistic analyses of reaction pathways

An an effort to understand the nature of the observed reaction pathways, we have performed More O'Ferrall-Jencks (MOFJ) analysis[18] of both mutants. This is carried out in terms of two crucial bond distances, namely, P_γ-O_{LG} and P_γ-O_{Nu}, that describe the evolution of bond cleavage and bond formation, respectively, during the course of the reaction. The correlation obtained from the produced trajectories manifests that GTP hydrolysis in both mutants follows essentially a concerted-like pathway according to Fig. ESI 4. In both cases, however, the TS is found to be a bit shifted towards the region corresponding to associative mechanisms. These analyses, hence, suggest that GTP hydrolysis in the E99A and S73A mutants of hGBP1 follow "concerted-associative" pathways in their respective rate-determining step.

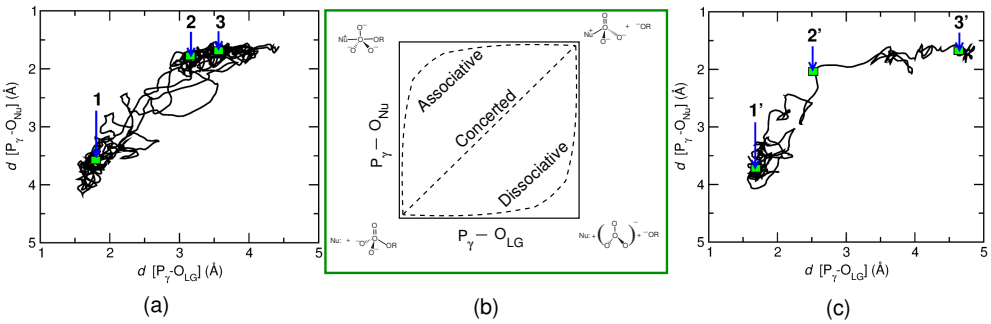

Fig. ESI 4: More O'Ferrall-Jencks analysis of GTP hydrolysis in the (a) E99A and (c) S73A mutants of hGBP1. The locations of reactant (**1** and **1'**), transition (**2** and **2'**) and product (**3** and **3'**) states are marked in the diagram. A generic MOFJ diagram illustrating the idealized associative, concerted, and dissociative reaction pathways is shown in the central panel (b).

In order to get confidence on the results obtained from this MOFJ analysis, we furthermore examined the amount of charge transferred on three key atoms, namely P_γ, O_{LG} and O_{Nu}, while proceding from the reactant state to the TS (see Table ESI 1). The increase of positive charge on P_γ is indicative of a dissociative pathway, whereas the charge on P_γ is more negative in the associative mechanism. The charge on P_γ, however, remains unchanged for the concerted pathway, as the positive charge developed on it (due to P_γ-O_{LG} bond breaking) will be compensated

Table ESI. 1: Mulliken charges (in |e|) and their standard deviations for structures corresponding to the reactant state (**1** and **1'**) and to the TS (**2** and **2'**) obtained from GTP hydrolysis in the E99A and S73A mutants of hGBP1. In order to perform this analysis, all the structures that fall within a certain cutoff from points representing reactant and TS configurations in the CV space are taken into the account.

Structure	P_γ Charge	O_{LG} Charge	O_{Nu} Charge
1	1.55 ± 0.08	-0.96 ± 0.04	-0.86 ± 0.04
2	1.36 ± 0.004	-0.90 ± 0.001	-0.87 ± 0.002
1'	1.47 ± 0.04	-0.95 ± 0.03	-0.92 ± 0.04
2'	1.44 ± 0.01	-0.84 ± 0.01	-0.93 ± 0.02

by the resulting negative charge due to the P_γ-O_{Nu} bond formation. On the other hand, the charge on O_{LG} changes to more negative in the dissociative and concerted pathway, whereas it remains unchanged (or changes only marginally) in the associative pathway. We observed a slight decrease in positive charge on P_γ. The negative charge on O_{LG} was also slightly decreased, whereas the charge on O_{Nu} remains unchanged. Therefore, also Mulliken population analysis supports a pathway that shares aspects of concerted and associative character, which is in line with the observation from the MOFJ analysis of the two mutants.

Finally, the shift of the reaction pathways from "concerted-dissociative" in wild type[1] hGBP1 to "concerted-associative" in the two mutant enzymes might be tentatively attributed to changes of the local electrostatic potential in the TS due to replacing the '$–CH_2OH$' and '$–(CH_2)_2COO^-$' side chains of serine and glutamate, respectively, with the charge-neutral '$–CH_3$' group of alanine.

7 SAC pathway for GTP hydrolysis in E99A

In order to verify if GTP hydrolysis can follow a SAC pathway in E99A, we performed a QM/MM metadynamics simulation utilizing the following two CVs: The CN difference $C[P_\gamma\text{-}O_{LG}] - C[O_{Nu}\text{-}P_\gamma]$ and the CN of H_{Nu} to all three oxygen atoms of the γ-phosphate. The importance of the first CV is already discussed in the main text, whereas the second CV is crucial to simulate the direct proton transfer from the nucleophilic water molecule, Nu, to one of the γ-oxygen atoms of GTP. This CV subspace is clearly sufficient to describe the direct SAC pathway, but no reaction was observed even though the free energy basin corresponding to the reactant minimum has been sampled via metadynamics up to a very high free energy level of 34 kcal/mol (see Fig. ESI 5). This result, hence, indicates that the SAC pathway in E99A either does not exist at all, or if exists is can only be accessed at the expense of unacceptably high free energy costs for enzymatic catalysis.

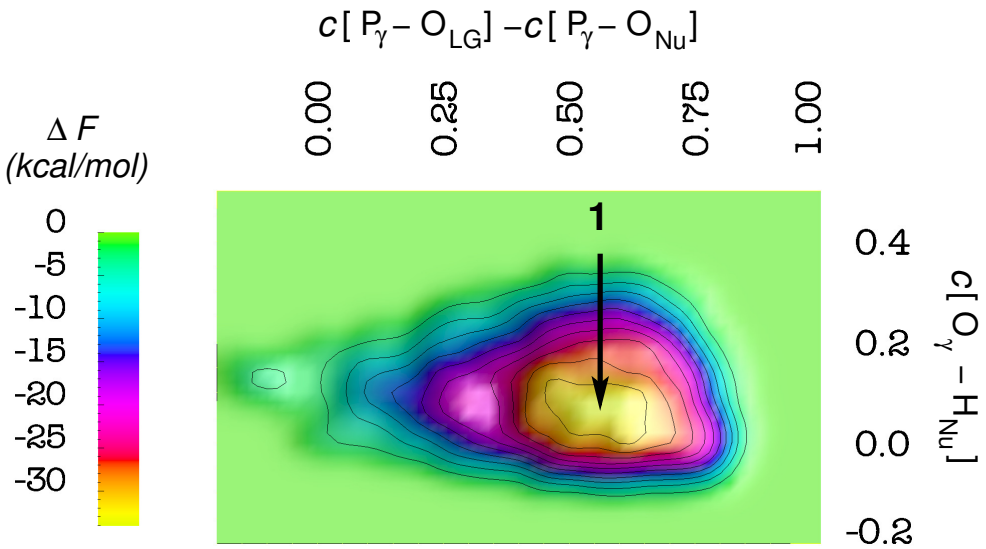

Fig. ESI 5: Free energy surface obtained from metadynamics simulation where the CVs are explicitly chosen to simulate the SAC pathway as explained in the accompanying text. It is clear from the reconstructed free energy surface that the reaction did not occur even when the free energy minimum corresponding to the reactant state has been mapped up to 34 kcal/mol.

8 Asp103 as a base for GTP hydrolysis in S73A

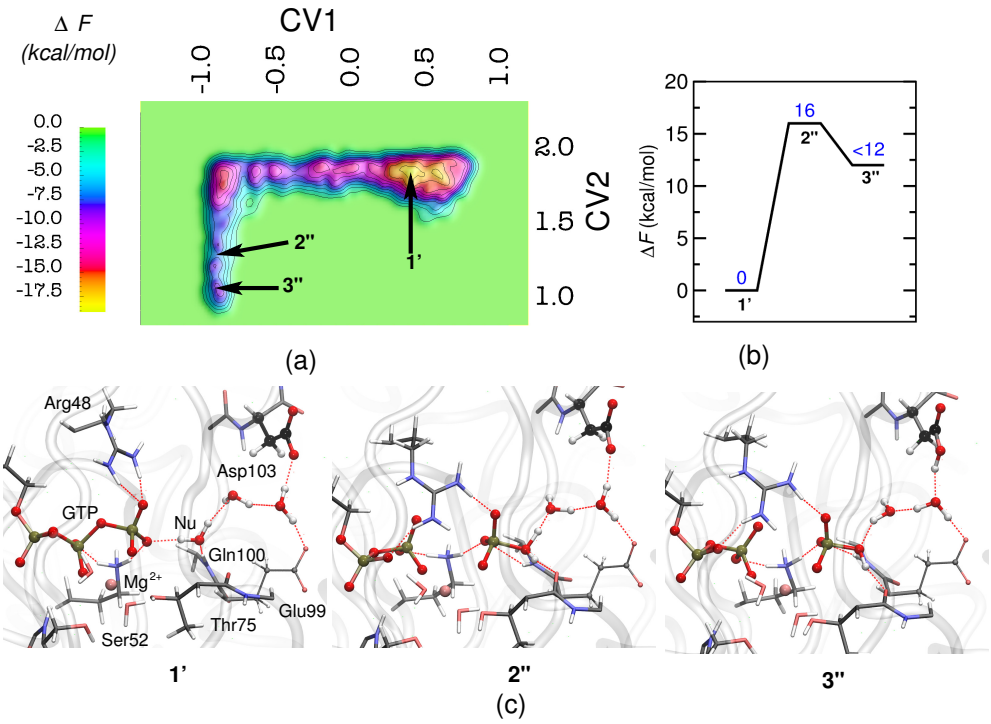

Fig. ESI 6: (a) Free energy surface and (b) corresponding free energy profile obtained for GTP hydrolysis in S73A where Glu99 was excluded from the QM subsystem and treated in its deprotonated form using a force field in order to block its protonation during QM/MM metadynamics sampling as described in the main text. The reactant and product minima **1**' and **3**" as well as the interconnecting transition state structure **2**" are depicted using representative configuration snapshots in (c).

References

[1] Tripathi, R., Glaves, R., and Marx, D. (2017) The GTPase hGBP1 converts GTP to GMP in two steps via proton shuttle mechanisms. *Chem. Sci. 8*, 371–380.

[2] Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A 38*, 3098–3100.

[3] Lee, C. T., Yang, W. T., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation energy formula into a functional of the electron-density. *Phys. Rev. B 37*, 785–789.

[4] Goedecker, S., Teter, M., and Hutter, J. (1996) Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B 54*, 1703–1710.

[5] Laino, T., Mohamed, F., Laio, A., and Parrinello, M. (2005) An efficient real space multigrid QM/MM electrostatic coupling. *J. Chem. Theory Comput. 1*, 1176–1184.

[6] Laino, T., Mohamed, F., Laio, A., and Parrinello, M. (2006) An efficient linear-scaling electrostatic coupling for treating periodic boundary conditions in QM/MM simulations. *J. Chem. Theory Comput. 2*, 1370–1378.

[7] Maseras, F. and Morokura, K. (1995) IMOMM: A new integrated *ab initio* + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem. 16*, 1170–1179.

[8] Martyna, G. J., Klein, M. L., and Tuckerman, M. E. (1992) Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys. 97*, 2635–2643.

[9] Weigend, F. and Häser, M. (1997) RI-MP2: First derivatives and global consistency. *Theor. Chem. Acc. 97*, 331–340.

[10] Weigend, F., Häser, M., Patzelt, H., and Ahlrichs., R. (1998) RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett. 294*, 143–152.

[11] Grimme, S. (2003) Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys. 118*, 9095–9102.

[12] Riplinger, C. and Neese, F. (2013) An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys. 138*, 034106.

[13] Hellweg, A., Hättig, C., Höfener, S., and Klopper, W. (2007) Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn. *Theoret. Chem. Acc. 117*, 587–597.

[14] Neese, F. (2012) The orca program system. *Wiley Interdiscip Rev Comput Mol Sci 2*, 73–78.

[15] Schreiner, E., Nair, N. N., and Marx, D. (2009) Peptide Synthesis in Aqueous Environments: The Role of Extreme Conditions on Peptide Bond Formation and Peptide Hydrolysis. *J. Am. Chem. Soc. 131*, 13668–13675.

[16] Glaves, R., Mathias, G., and Marx, D. (2012) Mechanistic insights into the hydrolysis of a nucleoside triphosphate model in neutral and acidic solution. *J. Am. Chem. Soc. 134*, 6995–7000.

[17] Zelleke, T. and Marx, D. (2017) Free-Energy Landscape and Proton Transfer Pathways in Oxidative Deamination by Methylamine Dehydrogenase. *ChemPhysChem 18*, 208–222.

[18] Anslyn, E. V. and Dougherty, D. A. (2006) *Modern Physical Organic Chemistry*. (University Science Books).