Electronic Supplementary Information (ESI):

The GTPase hGBP1 converts GTP to GMP in two steps via proton shuttle mechanisms

Ravi Tripathi*, Rachel Glaves* † and Dominik Marx*

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

† Present address : Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany

1 Methods and models

1.1 Setup and equilibration of the hGBP1 homodimer models

The crystal structures with PDB IDs 2BC9 and 2B8W [1] were taken as the starting structures for our calculations. The PDB ID 2BC9 corresponds to the isolated large GTPase domain of hGBP1 in the presence of the non-hydrolyzable GTP analogue GppNHp, whereas PDB ID 2B8W corresponds to the large GTPase domain of hGBP1 in complex with GMP/AlF_4^-, which mimics the transition state analogue of GDP hydrolysis. It has been shown that this domain alone proved sufficient to hydrolyze GTP to GMP [1], making hGBP1 a manageable model to handle computationally. The missing amino acid residues from the X-ray structures were introduced using the MODELLER [2] software suite and a symmetry partner was also added to 2BC9 in order to form the structure of the LG homodimer which is catalytically active. For the system constructed from 2BC9, which will be identified as GTP-P hereafter, the GppNHp residue was replaced by the GTP substrate and the dimer was solvated with 66 249 water molecules. The system was then neutralized by adding 148 Na^+ and 136 Cl^- ions, in order to ensure a physiological salt concentration, in a periodic cubic box of dimensions 127.6 $Å^3$. The smooth particle-mesh Ewald method [3] was employed to calculate the long-range Coulomb interactions. The protein molecule together with the substrate was described using the OPLS all-atom force field [4], whereas the water molecules were treated using TIP3P [5] model. The Lincs constraint [6] was applied to restrict the vibrational motion of all hydrogen atoms.

The equilibration of the system was first performed making use of position restraints to keep the positions of the protein backbone, the GTP substrate, the crystal water molecules, and the Mg^{2+} cation fixed. The entire system was then equilibrated using the NPT ensemble for about 34 ns. The well-equilibrated structure from the NPT simulation was then further subjected to 20 ns of NVT simulation in readiness for the forthcoming *ab initio* QM/MM molecular dynamics NVT simulations. The RMSD of the protein backbone with respect to the X-ray structure [1] (PDB ID 2BC9) was measured as an indication of the protein's stability. The average RMSD from the NVT simulation was found to be within 2.16 Å, which is reasonable, and gives us confidence in the force field employed.

In order to study the second enzymatic step, i.e. GDP hydrolysis in hGBP1, an intact GDP molecule was created out of the structural coordinates of GMP/AlF_4^- (available in PDB ID 2B8W) in the active site of hGBP1. The system was then neutralized by adding 61 Na^+ and 51 Cl^- ions to ensure a physiological salt concentration. A total of 52 170 water molecules were added to solvate the system inside a periodic cubic box of dimensions 118.9 $Å^3$. Other technical details remained the same as described before for GTP-P. The system (called GDP-P hereafter) was first minimized keeping the enzyme and GDP substrate fixed and then subjected to an equilibration, where the position restraints were gradually released from the backbone and side chains in a stepwise fashion. The whole system was then equilibrated using the NPT ensemble for about 5 ns until the density of the system was converged. This was followed by a further 70 ns of NVT simulation using the converged cell volume obtained from the NPT simulation. The RMSD of the protein backbone from the NVT run with respect to the X-ray structure was measured to be 1.67 Å, which provides adequate confidence in the force field parameters used to define the system. Note that this X-ray structure does not correspond to the actual GDP-bound reactant state, as it is in complex with GMP/AlF_4^-, mimicking the transition state (TS) of the GDP→GMP reaction, so the comparison of the RMSDs should be considered with care.

1.2 *Ab initio* QM/MM imulations

The system setups and subsequent force field MD simulations were then followed by *ab initio* QM/MM simulations [7] which were carried out using the `CP2k` code [8,9] using iterative Born-Oppenheimer propagation [7]. The QM part of the system was dealt with using fully self-consistent Kohn-Sham density functional theory employing the BLYP functional [10,11]. The TZV2P-GTH Gaussian basis sets were utilized together with GTH norm-conserving pseudopotentials [12]. A cubic box of 26 $Å^3$ was employed for the QM subsystem and the QM electrostatics were calculated using the Martyna-Tuckerman Poisson solver [13] for finite systems. The GEEP (Gaussian Expansion of Electrostatic Potential) method was used to handle the Coulomb interaction between the QM and MM atoms [14,15]. The boundaries of the QM/MM interface were dealt with using the IMOMM link scheme [16], in which the cut bond is saturated with a hydrogen in order to main-

tain 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 [17] and a timestep of 0.5 fs was employed.

The QM part chosen for the GTP-P system comprises the triphosphate tail of the GTP molecule cut at the carbon bond to the sugar, the Mg^{2+} ion, the side-chain/backbone atoms of Ser73, Tyr47, Arg48, Thr49, Gly50, Lys51, Thr98, Glu99, and Gly100. Also, the hydroxyl groups of Ser52 and Thr75, which coordinate the Mg^{2+} ion, were included in the QM subsystem. Additionally, the waters chosen for the QM system were two waters coordinating the Mg^{2+} ion, the catalytic water (Nu), the water molecule bridging Ser73 and Glu99 (WAT), and the water forming the hydrogen bond with the catalytic water. A total of 145 atoms were included in the QM part, making the total charge of the system equal to -1 that was properly taken into account by applying finite cluster boundary conditions [13]. The equilibrated structure from the NVT simulation of GTP-P was taken as the starting structure for the QM/MM simulation, where it was further equilibrated for about 5 ps. The average RMSD of the protein backbone with respect to the X-ray structure was obtained to be 2.09 Å here and no substantial structural changes were observed compared to the force field simulations.

On the other hand, the QM system selected for GDP-P included the diphosphate tail of the GDP molecule cut at the carbon bond to the sugar. Also, the water molecules chosen here comprise the two water molecules coordinating the Mg^{2+} ion, the catalytic water (Nu), and the two water molecules bridging the catalytic water and Glu99. The rest of the QM system, stemming from protein residues, is the same as was chosen for the GTP-P system. Overall, the total charge of the QM system was zero and the total number of atoms making up the QM system was 141. The equilibrated structure from the MM MD simulation of GDP-P, which correlates well with the X-ray structure, is taken as the starting structure for the subsequent QM/MM MD simulation. 5 ps of MD simulation were then performed within the QM/MM framework in order to substantiate the stability of the interactions of the active site residues observed from the MM MD simulation. All of the interactions from the force field MD simulation were found to remain intact in the reactive QM/MM MD simulation.

1.3 Sampling the free energy landscapes

Extended Lagrangian metadynamics simulations [7,18–22] were then performed in order to simulate the GTP and GDP hydrolysis, which are not directly accessible on the timescale of *ab initio* molecular dynamics simulations. This non-Markovian accelerated rare-event method works by enhancing the sampling along some predefined generalized variables, termed as collective variables (CVs), that are constructed to provide a coarse grain description of the system in terms of a suitable reaction subspaces in which the multi-dimensional free energy surfaces are generated in practice. A slowly growing Gaussian biasing potential is then added to enhance the sampling in the multi-dimensional reaction subspace. The multi-dimensional free energy landscape is then reconstructed by accumulating the added potentials, thus enabling the reaction mechanism along the minimum energy pathway to be determined including free energy differences such as activation free energies. Gaussian potentials of heights varying between $\approx 0.5\,k_{\mathrm{B}}T$ to $\approx 2\,k_{\mathrm{B}}T$ were used to fill the minima. The mass of the auxiliary variables was set to 50 amu and a coupling parameter, k_α, of 2.0 au was selected. Also, a width scaling factor of $\delta s = 0.03$ was defined for each CV.

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

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

where p and q are constants that determine the steepness of the coordination number function and d^0_{AB} is a fixed cutoff parameter chosen based on the van der Waals parameters of the atoms involved. The fraction inside the sum is nearly unity when $d[I-J] > d^0_{\mathrm{AB}}$, and close to zero otherwise. The values of both p and q were taken to be 6 for all coordination number type CVs, however, the value of d^0_{AB} was chosen to be 2.0 Å for all CVs involving Pγ and to be 1.48 Å otherwise.

In total, about 0.2 ns of *ab initio* QM/MM metadynamics simulations have been carried out in order to arrive at the data that underly the results of this investigation.

1.4 Convergence and sampling error in free energy barriers

The standard protocol to check the convergence of the free energy barrier is to continue the simulation until multiple (re)crossings involving all free energy minima within the reaction subspace spanned by the respective CVs. However, simulating many such recrossings is computationally impracticable for most large-scale simulation studies, in particular when computing the forces and energies from electronic structure calculations as required when studying covalent chemical reactions as done in the present case. Besides, different sets of CVs are often required to properly sample the forward and backward reactions. This is particularly true if the product state is "unbound" as a result of a cleavage reaction where the leaving group (cleavage product) might diffuse very far away so that its return (i.e. recrossing toward the reactant state) becomes increasingly improbable to sample.

A well-established practical solution is to estimate the error in metadynamics simulations by repeating it for a distinct reaction pathway (e.g. either the forward **or** the backward reaction) upon systematically decreasing the height of the Gaussian functions that are used to fill the free energy minimum corresponding to the reactant state (when assessing the forward reaction) until it reaches the product state for the first time (and thus escapes the reactant state once). Upon decreasing the Gaussian height W systematically allows one to explicitly estimate the error bar of the activation free energy barrier relative to the reactant state [23]. According to this "refinement protocol" [23], relatively large Gaussians (usually of height 2 to $3k_{\mathrm{B}}T$) are initially used to simulate the reaction by efficiently filling the reactant minimum. Once the first crossing event from reactant to product state is observed, the obtained activation free energy estimate is refined subsequently by restarting the simulation in the reactant minimum sometime *well before* the first crossing toward the product occurred – but using a smaller Gaussian height this time. Moreover, the Gaussian height can be adapted on demand during metadynamics sampling by using two or more W parameters during an otherwise continuous run. The convergence in the free energy barrier with decreasing height of the Gaussians allows one to explicitly compute an error bar estimate that is independent from sampling recrossings events.

We have performed such convergence calculations for both the GTP $\rightarrow$ GDP and the GDP $\rightarrow$

GMP transformations, i.e. the two forward steps in hGBP1 hydrolysis, in order to compute the activation free energy barriers as reported in Table 1. They are both found to be converged within an error of about 2 kcal/mol upon decreasing the Gaussian height parameter W. Importantly, the reaction mechanisms of both GTP and GDP hydrolysis steps were found to be identical upon variation of the Gaussian height. Hence, we conclude that the error of the activation free energy for hydrolysis due the metadynamics sampling procedure is approximately 2 kcal/mol. Clearly, the free energy difference of the product state w.r.t. the corresponding reactant state cannot be reliably computed when using this refinement approach. Yet, upper bounds for these relative free energies can be provided which are indicated in the free energy profiles in Figs. 1 to 3 of the main text using the $<$ signs for the free energies of the **3P**, **5P**, **8P** and **3R** product states relative to the **1P**, **3P**, **6P** and **1R** reactant states, respectively.

Table 1: Convergence of the computed activation free energy barriers $\triangle F^{\ddagger}$ for both hydrolysis steps as indicated (which is the free energy difference of **2P** w.r.t. **1P** and of **7P** w.r.t. **6P** for GTP and GDP hydrolysis, respectively) with respect to the Gaussian height parameters W that have been used for refinement (see text).

W (kcal/mol)	$\triangle F^{\ddagger}$ (kcal/mol)	
	1P$\rightarrow$**3P**	**6P**$\rightarrow$**8P**
1.2	20.0	22.0
1.2/0.6	18.0	21.0
1.2/0.6/0.3	18.0	–
1.2/0.3	–	22.0

1.5 Validating the electronic structure method

Systematic errors in the estimation of free energies and, thus, possibly even in the reaction mechanism can also arise due to the electronic structure method, which is the BLYP density functional in our case, that is used to carry out all *ab initio* molecular dynamics or *ab initio* QM/MM metadynamics simulations reported herein. In order to validate this approach, the energetics of various configurations were compared between BLYP and a wave function-based correlation method,

namely SCS-(RI-)MP2 [24–26] together with the triple-zeta valence plus two polarization function (TZVPP) basis set [25,27] as implemented in Turbomole [28]. This is done in terms of several instantaneous snapshot configurations from the GTP hydrolysis reaction in hGBP1 that have been sampled to represent the full course of the reaction from reactant **1P** to product **5P** via the TS **2P**, see Fig. ESI 1. The triphosphate tail of the GTP molecule, cut at the carbon bond to the sugar, the side chains of Arg48, Lys51, Ser73, and Glu99, and the Mg^{2+} ion plus the two water molecules coordinating it were included in this study. In addition, the hydroxyl groups of Ser52 and Thr74, saturated with hydrogen atoms, were also included to make up the hexacoordinated shell of the Mg^{2+} ion. The obtained potential energy differences from the BLYP functional along the minimum free energy pathway (see Fig. ESI 1) were found to be in good agreement with the ones obtained from the SCS-MP2/TZVPP reference; the maximum deviation was obtained to be circa 2.0 kcal/mol. Moreover, in order to explicitly assess the influence of dispersion interactions in our simulations with direct reference to BPYP, we have computed the potential energy differences along the same configurations by including the corresponding D3 correction [29]. We found that including dispersion does not improve the quality of result.

2 Non-enzymatic reference scenario: MeTP hydrolysis in bulk water

A molecular level understanding of the mechanism of triphosphate hydrolysis in a bulk water environment is necessary to achieve a one-to-one comparison between the enzymatic and non-enzymatic reactions (see main text). To this end, we have investigated a reference system comprising a fully deprotonated methylated triphosphate molecule (MeTP) plus a magnesium dication solvated in 113 water molecules subject to periodic boundary conditions in order to represent bulk solvation. The electronic structure details as well as the *ab initio* MD parameters chosen for the system remained the same as for the GTP-P and GDP-P systems and can be obtained from the previous section. The whole system was represented completely within the BLYP framework in view of the QM subsystem underlying the QM/MM simulations of the enzymatic case. The total

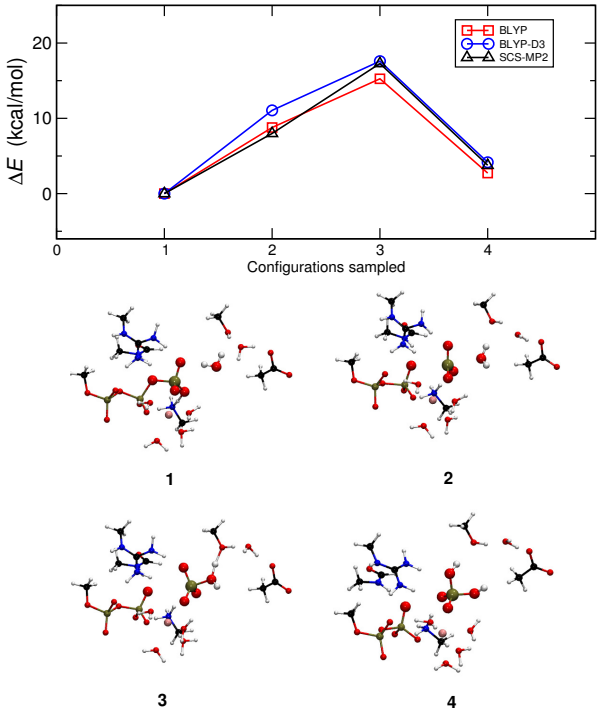

Fig. ESI 1: Comparison of the computed potential energy differences of sampled representative configurations (see text) according to BLYP, and BLYP-D3, and SCS-MP2 electronic structure while moving from reactant to product via the TS during GTP hydrolysis by hGBP1.

charge of the fully periodic system was -2 and a cubic box of volume $V = 15.374\,\text{Å}^3$ was employed. An equilibrated structure obtained from the circa 32 ps of simulation from our previous *ab initio* MD study [30] was taken to be the starting structure, and it was further equilibrated for about 6 ps using NVT *ab initio* MD simulations.

We want to emphasize that the CVs used here are chosen to be virtually identical to the ones used to simulate GTP hydrolysis in hGBP1, which is required to enable a one-to-one comparison between these two reactions in the same two-dimensional reactive subspace that hosts the free energy landscape. In contrast, in the investigation devoted to studying hydrolysis in a bulk water environment [30], where it is *a priori* unknown which of the many water molecules in the vicinity of the terminus of the triphosphate tail will eventually be involved in the chemical reaction step as opposed to distinct water in the enzyme, a more general set of CVs spanning a three-dimensional

reactive subspace was considered in our previous study [30]. As expected, the previous, more flexible approach resulted in a slightly lower free energy barrier ($\approx$ 29 kcal/mol) in comparison to the one obtained within the current study ($\approx$ 33 kcal/mol). In addition, the very shallow intermediate on the free energy surface reported earlier lies within the limitations of the sampling accuracy as already stressed in Ref. 30: "The initial barrier of ca. 29 kcal/mol (which corresponds to a thermal energy of roughly 50 $k_{\mathrm{B}}T$ at ambient conditions) represents the breakage of the γ-phosphate to bridging oxygen bond and the positioning of the attacking water molecule (structure 2N) leading to the formation of a very shallow intermediate on the FES at around 27 kcal/mol, which lies just within the limitations of our sampling accuracy"

3 Hydrogen bond interactions in the transition state for GTP hydrolysis in hGBP1

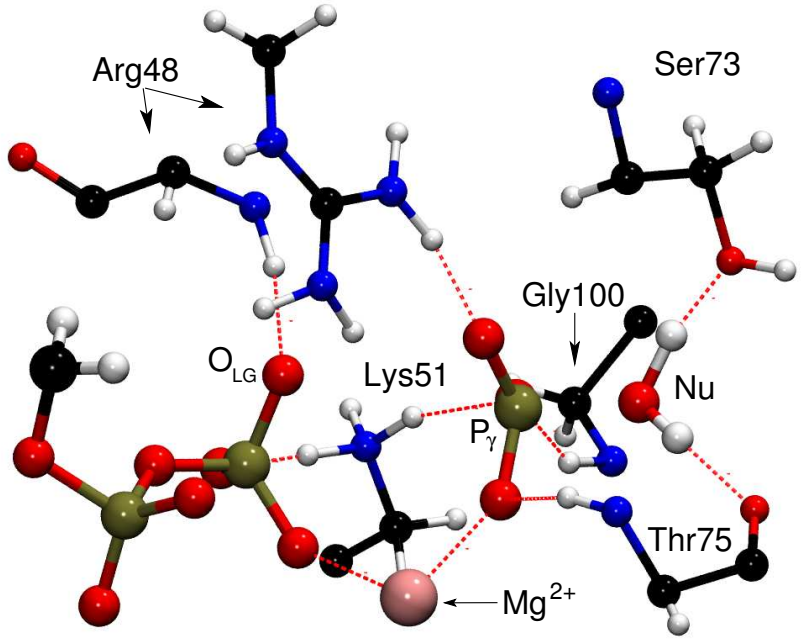

Fig. ESI 2: The key stabilizing hydrogen bond contacts in the TS structure for **1P** $\rightarrow$ **3P** conversion are highlighted using dotted lines. They involve the positively charged Lys51 and Arg48, the backbone NH group of Thr75 and Gly100, as well as the substrate-bound Mg^{2+}. The nucleophilic water Nu is found to be stabilized/oriented suitably by interacting with Ser73 and with the backbone carbonyl group of Thr75.

4 Evolution of the CVs trajectories generated by metadynamics

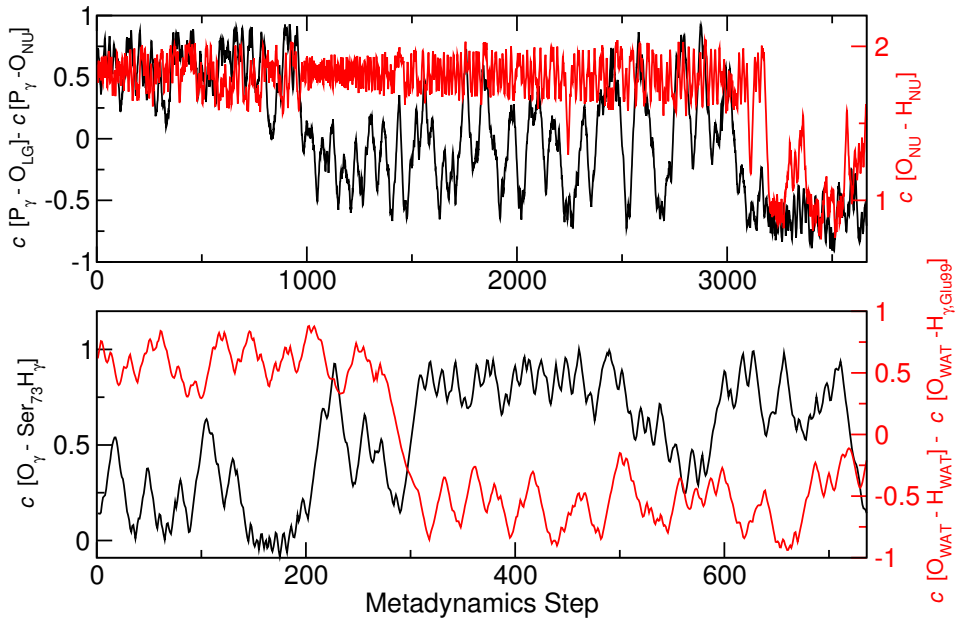

Fig. ESI 3: Evolution of the CVs as obtained for (a) **1P** $\rightarrow$ **3P** conversion (upper figure), and (b) **3P** $\rightarrow$ **5P** conversion (lower figure).

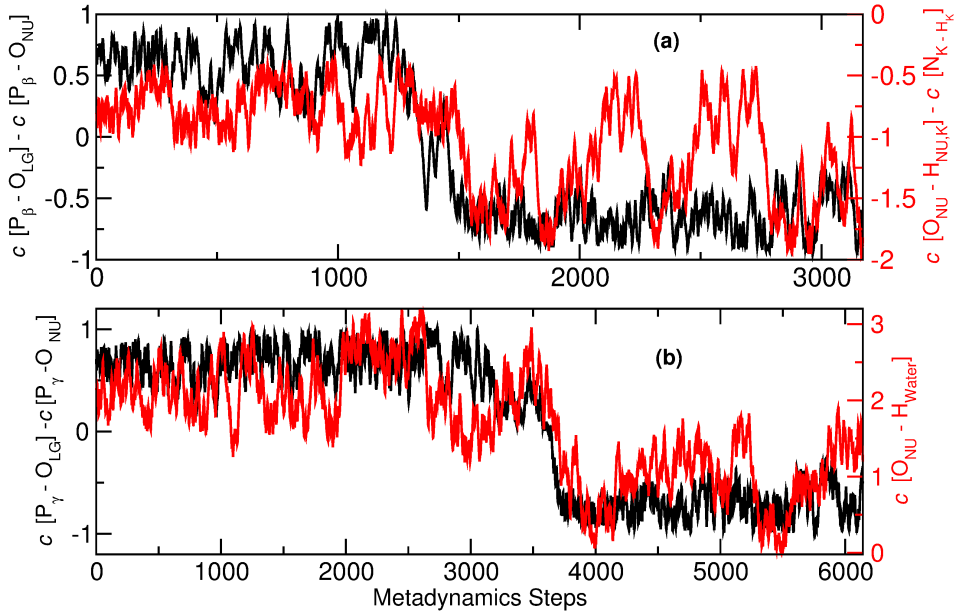

Fig. ESI 4: Evolution of CVs for (a) **6P** → **8P** transformation (upper figure), and (b) **1R** → **3R** transformation (lower figure)

5 GTP→GDP conversion via the concerted one–step mechanism

In order to verify the substrate assisted catalysis pathway in hGBP1, a metadynamics simulation was performed where the following coordinates were utilized: The coordination number difference C [P_γ-O_{LG}] − C [O_{Nu}-P_γ] and the coordination number of H_{Nu} to all the three oxygens of the γ-phosphate. The importance of the first CV has already been discussed before, and the second CV was utilized to probe the direct proton transfer from the Nu to one of the γ-oxygens of GTP. The reaction took place with the simultaneous breaking and formation of the P_γ-O_{LG} and P_γ-O_{Nu} bonds, respectively, and was accompanied by the deprotonation of the water Nu thus resulting in a one-step mechanism (see Fig. ESI 5(b)). Interestingly, although the reaction subspace that spans this free energy landscape was optimal to simulate the direct proton transfer from Nu to GTP, the proton transfer to the γ-phosphate oxygen took place, once again, indirectly via a sequence of proton relays involving Ser73, Glu99, and WAT, which are all strongly engaged in hydrogen

bonding. Careful analysis of representative configurations along the respective minimum free energy pathway discloses that this indirect one-step mechanism is structurewise very similar to the earlier one since **2'P**, **3'P**, and **4'P** in Fig. ESI 5 are very similar to their counterparts **2P**, **3P**, and **4P** in Fig. 1. Moreover, they are located at topologically similar points along the reaction path on the free energy landscape as unravelled by the analysis in Fig. ESI 5(c). Yet, the one-step process from **1P** directly to **5P** is found to possess a higher free energy barrier (26 kcal/mol) in comparison to the two-step process according to Fig. 1(b).

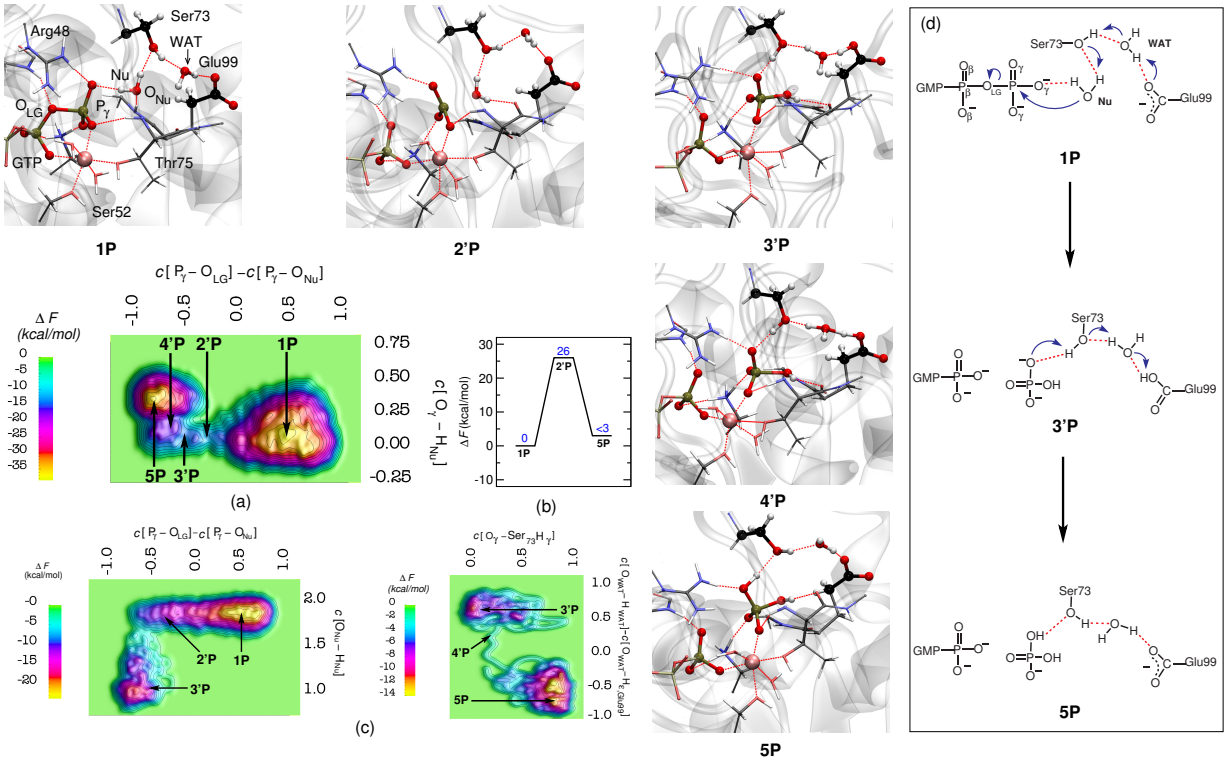

Fig. ESI 5: Free energy surface, (a), and corresponding free energy profile, (b), obtained for the concerted one-step hydrolysis mechanism of GTP to GDP by hGBP1 (see main text). The reactant (GTP) and product (GDP) minima **1P** and **5P**, the corresponding TS structure **2'P**, and two more intervening (non-stationary state) structures along the minimum free energy pathway, **3'P** and **4'P**, are depicted using representative configuration snapshots. The corresponding locations of these structures on the free energy surfaces obtained for the energetically preferred two-step "indirect SAC" mechanism (see Fig.1 in the main text) are indicated in panel (c). This demonstrates, together with the depicted respresentative configuration snapshots, that these structures are very similar to the respective states **2P**, **3P**, and **4P** according to the two-step mechanism. The mechanism of concerted one-step hydrolysis of GTP by hGBP1 according to the free energies in panels (a) and (b) is summarized schematically in panel (d). The color code used here is as follows: O (red); C (black); N (blue); H (white); Mg (pink).

6 Metadynamics study to simulate GTP→GDP conversion by restricting Ser73 participation in the proton transfer relay

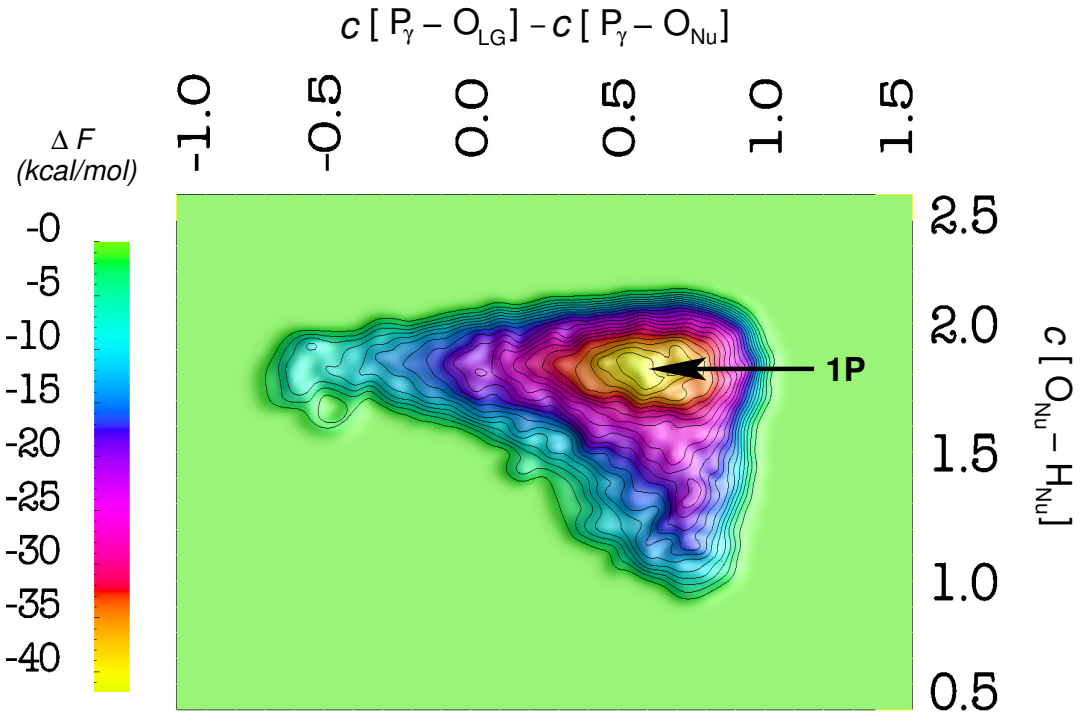

Fig. ESI 6: Free energy landscape after introducing a soft repulsive wall that acts on the distance between $O_{\gamma,Ser73}$ and H_{Nu} in order to *prevent* Ser73 to be part of the proton relay that connects via hydrogen bonding Glu99 via WAT, Ser73 and Nu to the γ-phosphate oxygen of GTP; see text and compare to the free energy landscape in Fig. 1(a). This *ab initio* QM/MM metadynamics simulation was continued until the depth of reactant free energy minimum exceeded about 40 kcal/mol. Despite the unusually high free energy this simulation did not lead to GTP hydrolysis (for instance via **2P** to **3P**) and only the minimum corresponding to the GTP reactant state **1P** was explored.

Table 2: Average bond lengths (in Å) and Mulliken charges (in |e|) as well as their standard deviations for structures corresponding to the reactant (**1P**) and to the TS (**2P**), obtained from the GTP hydrolysis in hGBP1. In order to calculate the averages, we have selected an ensemble of snapshots that fall within a certain cutoff from a point, representing reactant and TS, in CV space. A minimum of 10 random configurations, representing reactant and TS, were chosen in order to perform the charge calculations.

Structure	P_γ-O_{LG}	P_γ-O_{Nu}	P_γ Charge	O_{LG} Charge	O_{Nu} Charge
1P	1.87 ± 0.04	3.44 ± 0.21	1.50 ± 0.05	-0.79 ± 0.07	-0.75 ± 0.02
2P	3.50 ± 0.20	1.84 ± 0.04	1.40 ± 0.07	-0.96 ± 0.03	-0.63 ± 0.07

7 Charge transfer from population analysis

The amount of charge transfer on the three main atoms i.e. P_γ, O_{LG}, and O_{Nu} was also examined (based on Mulliken population analysis of the electronic structure underlying our *ab initio* QM/MM metadynamics simulations) with the aim to gain a better understanding of the nature of the TS structure **2P** gauged with respect to the reactant state **1P**. For a dissociative pathway, the charge on the P_γ becomes more positive, whereas it is either more negative or does not change at all in the associative mechanism [31]. Also, the charge on the P_γ remains unchanged for the concerted pathways as the increase in the positive charge on the P_γ due to P_γ–O_{LG} bond dissociation is compensated by charge transfer to the P_γ from O_{Nu} as the bond between them forms. On the other hand, there is an increased localization of negative charge on the O_{LG} for the dissociative and the concerted pathways. We found that the charge on the P_γ changed only marginally on going from reactant to TS, moreover, a sharp increase in the negative charge on the O_{LG} was noticed along the GTP hydrolysis pathway (see Table 2). These observations again favor a concerted mechanism of the rate-determining hydrolysis step which is in line with the conclusion drawn from the MOFJ plots in the main text. Also, the charge on the O_{Nu} is decreases slightly when moving from reactant to TS, which further supports the concerted pathway.

8 Selected interactions within the active site of GTP-P

Table 3: Average bond distances (in Å) and their standard deviations for structures corresponding to reactant and TS obtained during GTP hydrolysis in hGBP1. All structures that fall within a certain cutoff from a point, representing reactant and TS, in CV space, were selected to measure the average bond distances.

	1P	**2P**
$H_{Arg48}\cdots O_{LG}$	2.21 ± 0.24	1.80 ± 0.13
$H_{\eta 12,Arg48}\cdots O_{\gamma}$	1.66 ± 0.13	1.69 ± 0.09
$H_{\eta 22,Arg48}\cdots O_{LG}$	2.71 ± 0.27	2.77 ± 0.26
$H_{\zeta 3,Lys51}\cdots O_{\gamma}$	1.62 ± 0.13	1.94 ± 0.12
$H_{\zeta 2,Lys51}\cdots O_{\beta}$	2.00 ± 0.31	1.71 ± 0.18
$Mg^{2+}\cdots O\gamma$	2.00 ± 0.07	2.19 ± 0.10
$Mg^{2+}\cdots O\beta$	2.10 ± 0.10	2.00 ± 0.06
$H_{Thr75}\cdots O_{\gamma}$	2.42 ± 0.23	2.14 ± 0.10
$O_{Thr75}\cdots O_{Nu}$	3.74 ± 0.92	2.72 ± 0.15
$N_{Gly100}\cdots O_{Nu}$	3.17 ± 0.33	3.78 ± 0.22
$H_{His74}\cdots O_{\gamma}$	2.21 ± 0.26	2.19 ± 0.16

9 Second step: GDP→GMP conversion

As commonly done in order to restrict the sampling space to the interesting regions on the free energy landscape with the aim to increase sampling efficiency, we have applied a soft repulsive wall potential at a value of -0.6 along the second CV while simulating the GDP hydrolysis step. This gentle restraint restricts the possibility of Nu to abstract a proton from Lys51 while providing a proton to the γ-oxygen of GTP.

10 Ser73 mutation in the GDP-bound state of hGBP1

Ghosh *et al.* have shown that mutation of Ser73 (S73A) in hGBP1 results in low intrinsic activity and loss of cooperativity for both the GTP and GDP hydrolysis reactions. This led them to

propose an identical mechanism for both the GTPase and the GDPase activity [1]. In contrast with this proposal, our study shows no direct role of Ser73 during GDP hydrolysis. In order to gain a deeper understanding of this experimental observation, we performed an additional force field simulation utilizing the S73A mutation in the GDP-bound structure of hGBP1. It was observed in this MD simulation of the wild type enzyme that Ser73 always forms a hydrogen bond with Lys76 (see Fig. ESI 7). This interaction was lost upon S73A mutation, allowing Lys76 to rotate towards Glu99, and resulting in the formation of a salt bridge between the two. This obstructs the interaction of Glu99 with the chain of water molecules interacting with WAT, thereby blocking the proton relay mechanism (see Fig. ESI 7), and explaining the low intrinsic activity seen in the experimental study of the S73A mutation.

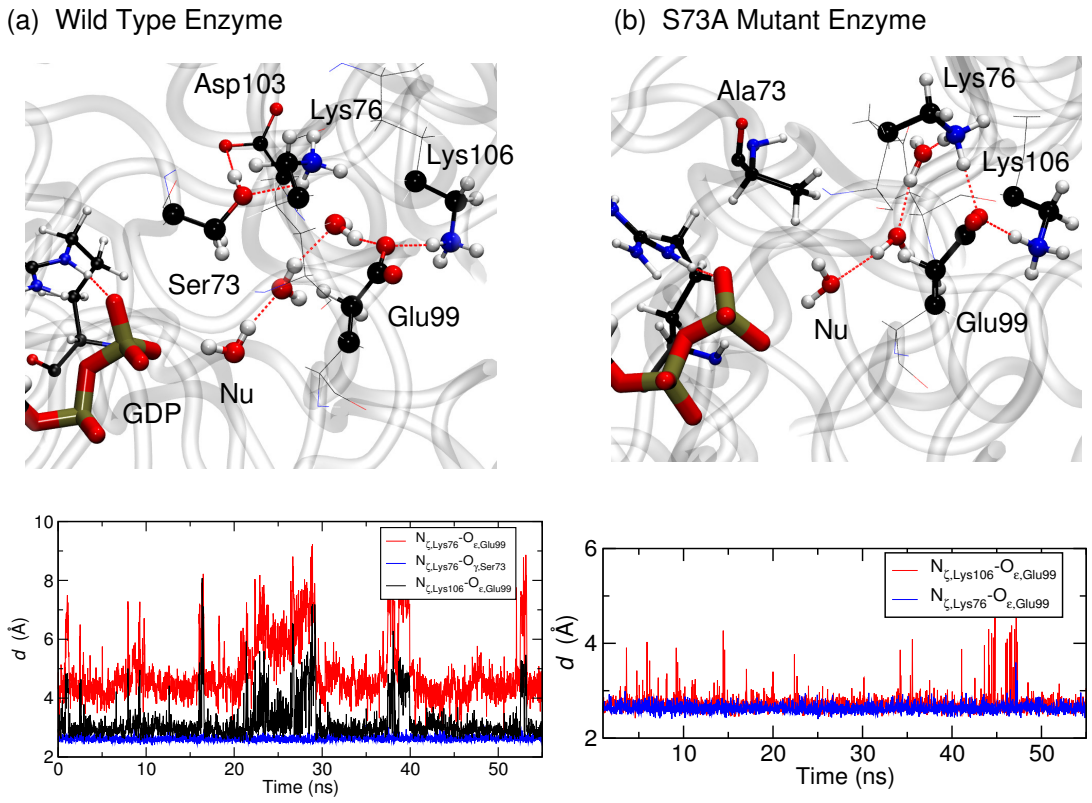

Fig. ESI 7: Average active site structures of (a) wild type GDP-P and (b) GDP-P with the S73A mutation. Distances between crucial atoms are shown in the lower graphs. The first 15 ns of the trajectory (from a total of 70 ns of simulation time) was ignored in the analysis to allow for equilibration.

11 Structural superimposition of 7P with the X-ray structure

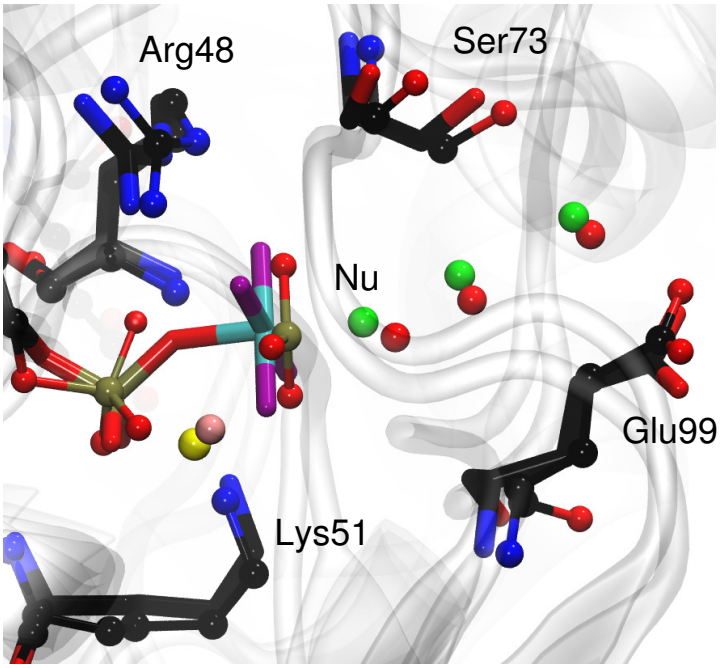

Fig. ESI 8: Structural superimposition of the X-ray structure PDB ID 2B8W (shown in the stick representation) with the TS structure of GDP hydrolysis obtained from our calculation (shown in the CPK representation). Magnesium ions from X-ray and simulation are shown in yellow and pink, respectively. The crystallographic water molecules are shown as green spheres, whereas the usual color convention is followed for the rest of the atoms (O: red, C: black, N: blue, Al: cyan, F: purple).

12 More O'Ferrall–Jencks analysis in terms of bond orders

In order to complement the usual More O'Ferrall–Jencks (MOFJ) analysis [32] carried out in terms of bond distances (as shown in Fig. 4) by the corresponding bond order representation, we depict this graph in Fig. ESI 9. It is obtained upon applying the simple bond–order / bond–distance relationship introduced by Pauling [33] in its usual parameterization $n(r) = \exp[-(r - r_e)/0.6]$ (according to Eq. (5) in Ref. 34) to the trajectory for GTP hydrolysis by hGBP1. The bond order analysis is found to highlight the dissociative character of this reaction even more nicely when

compared to using bond distances (see Fig. 4), which supports the concerted-dissociative nature of the process as worked out in detail in the main text.

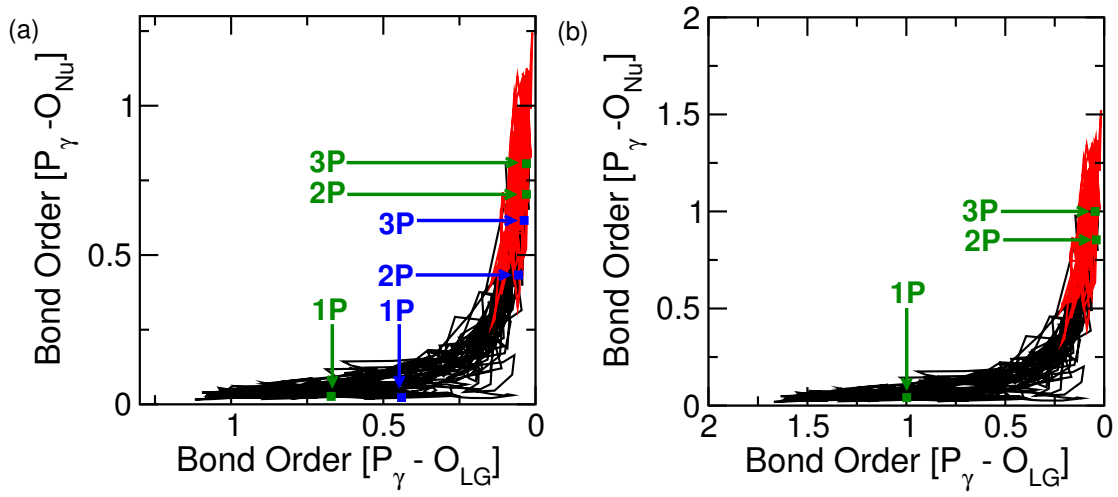

Fig. ESI 9: (a) More O'Ferrall-Jencks analysis of the GTP hydrolysis step in hGBP1 as carried out in terms of Pauling's bond order equation (see text) using $r_e = 1.63$ Å is represented using a solid line; the trajectory piece after the deprotonation of Nu is highlighted in red. The corresponding Pauling bond orders of the transition state, **2P**, along with those in the reactant and product states, **1P** and **3P**, respectively, are marked in green. The average values of the quantum–mechanical bond orders calculated from sets of representative configurations of the **1P**, **2P** and **3P** states using NBO analysis (see text) are displayed in blue. (b) More O'Ferrall-Jencks analysis of the GTP hydrolysis step in hGBP1 based on Pauling's bond order equation (see text) but using $r_e = 1.87$ Å and 1.75 Å for the Pγ–O$_{LG}$ and Pγ–O$_{NU}$ bonds, respectively, as computed from AIMD simulations (see text) is represented using a solid line; the trajectory piece after the deprotonation of Nu is highlighted in red.

However, although conceptually appealing, using Pauling's simple bond order transformation of P–O distances might introduce complications into the analysis of phosphate hydrolysis as lucidly reviewed in Section 2.3.2 of Ref. 35. In an effort to explicitly check if the empirical Pauling relation, that has been suggested to describe carbon–carbon bonds, also holds for the vastly different P–O bonds relevant to phosphate hydrolysis, we have numerically computed their bond

orders for configurations that are representative for the reactant, transition and product states of GTP hydrolysis corresponding to **1P**, **2P** and **3P**, respectively. For this purpose, subsets of those structures that were sampled to perform Mulliken population analysis in SI Section 7 were picked, for which Natural Bond Orbital (NBO) analysis [36] has been carried out in order to compute the bond orders directly from electronic structure using the `Gaussian09` package [37]. For this purpose, the triphosphate tail of the GTP molecule, cut at the carbon bond to the sugar, the Mg^{2+} ion and the side-chain/backbone atoms of Tyr47, Arg48, Thr49, Gly50, Lys51, Ser52, Ser73, His74, Thr75, Lys76, Glu99 as well as Gly100 were included to perform these calculation. Moreover, the four water molecules among those included in the QM subsystem during the full QM/MM metadynamics simulation underlying the GTP hydrolysis trajectory were also taken into account when performing this NBO analysis. The underlying single–point electronic structure calculations were carried out using again the BLYP functional in conjunction with the 6–311G(d) basis set as implemented in `Gaussian09`.

The average quantum-mechanical NBO-based bond orders (marked using blue squares and arrows) obtained for the **1P**, **2P** and **3P** structures (marked using green squares and arrows) are depicted in panel (a) of Fig. ESI 9 along with the MOFJ analysis based on Pauling's bond order equation (lines) of the GTP hydrolysis step in hGBP1 according to metadynamics. It can be seen that the bond orders obtained using the NBO calculations follow a qualitatively similar pattern as that obtained from Pauling's equation (using the usual reference bond distance of $r_e = 1.63$ Å for both the breaking and forming P–O bonds). This confirms in passing that Pauling's bond order equation is not strictly limited to reactions involving C–C bonds, but apparently also provides results that are consistent with numerical quantum chemistry when applied to phosphate hydrolysis in enzymes. In addition, we demonstrate via panel (b) that using the average bond lengths as numerically computed from the simulated trajectory specifically for the $P\gamma$–O_{LG} and $P\gamma$–O_{NU} bonds (being 1.87 and 1.75 Å, respectively, as reported in Fig. 1) in defining the reference value r_e in Pauling's equation instead of 1.63 Å does not lead to any qualitative difference in the resulting MOFJ pathways compared to the simple relation used in panel (a).

Based on this combination of MOFJ analysis in terms of bond distances (in the main text) and bond orders (herein) we can safely conclude that GTP hydrolysis in hGBP1 proceeds via a

concerted–dissociative reaction pathway.

References

[1] Ghosh, A.; Praefcke, G. J. K.; Renault, L.; Wittinghofer, A.; Herrmann, C. (2006). How guanylate-binding proteins achieve assembly-stimulated processive cleavage of GTP to GMP. *Nature*, *440*, 101–104.

[2] Sali, A.; Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, *234*, 779–815. `http://salilab.org/modeller`.

[3] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.*, *103*, 8577–8593.

[4] Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B*, *105*, 6474–6487.

[5] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, *79*, 926–935.

[6] Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.*, *18*, 1463–1472.

[7] Marx, D.; Hutter, J. *Ab initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press, 2009.

[8] Lippert, G.; Hutter, J.; Parrinello, M. (1997). A hybrid Gaussian and plane wave density functional scheme. *Mol. Phys.*, *92*, 477–487.

[9] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. (2005). QUICKSTEP: fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.*, *167*, 103–128.

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

[11] Lee, C.; Yang, W.; 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.

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

[13] Martyna, G. J.; Tuckerman, M. E. (1999). A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.*, *110*, 2810–2821.

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

[15] Laino, T.; Mohamed, F.; Laio, A.; 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.

[16] Maseras, F.; Morokuma, 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.

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

[18] Laio, A.; Parrinello, M. (2002). Escaping free-energy minima. *Proc. Natl. Acad. Sci.*, *99*, 12562–12566.

[19] Iannuzzi, M.; Laio, A.; Parrinello, M. (2003). Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.*, *90*, 238302–1–4.

[20] Laio, A.; Parrinello, M. (2006). Computing Free Energies and Accelerating Rare Events with Metadynamics. *Lect. Notes Phys.*, *703*, 315–347.

[21] Laio, A.; Gervasio, F. L. (2008). Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.*, *71*, 126601–126623.

[22] Barducci, A.; Bonomi, M.; Parrinello, M. (2011). Metadynamics. *WIREs Comput. Mol. Sci.*, *1*, 826–843.

[23] Nair, N. N.; Schreiner, E.; Marx, D. (2008). Peptide synthesis in aqueous environments: The role of extreme conditions on amino acid activation. *J. Am. Chem. Soc.*, *130*, 14148–14160.

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

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

[26] 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*(20), 9095–9102.

[27] Weigend, F.; Ahlrichs., R. (2005). Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, *7*, 3297–3305.

[28] Ahlrichs, R.; Bär, M.; Häser, M.; Horn., H.; Kölmel, C. (1989). Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.*, *162*, 165–169. `http://www.turbomole.com`.

[29] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction DFT-D for the 94 elements H-Pu. *J. Chem. Phys.*, *132*, 154104–19.

[30] Glaves, R.; Mathias, G.; 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.

[31] Carvalho, A. T. P.; Szeler, K.; Vavitsas, K.; Åqvist, J.; Kamerlin, S. C. L. (2015). Modeling the mechanisms of biological GTP hydrolysis. *Arch. Biochem. Biophys.*, *582*, 80–90.

[32] Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*. University Science Books, 2006.

[33] Pauling, L. (1947). Atomic Radii and Interatomic Distances in Metals. *J. Am. Chem. Soc.*, *69*, 542–553.

[34] Houk, K. N.; Gustafson, S. M.; Black, K. A. (1992). Theoretical secondary kinetic isotope effects and the interpretation of transition state geometries. 1. The Cope rearrangement. *J. Am. Chem. Soc.*, *114*, 8565–8572.

[35] Kamerlin, S. C. L.; Sharma, P. K.; Prasad, R. B.; Warshel, A. (2013). Why nature really chose phosphate. *Q. Rev. Biophys.*, *46*, 1–132.

[36] Weinhold, F.; Landis, C. R. (2001). Natural bond orbitals and extensions of localized bonding concepts. *Chem. Educ. Res. Pract.*, *2*, 91–104.

[37] 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.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox., D. J. *Gaussian 09, Revision B.01*. Gaussian, Inc., Wallingford CT, 2010.

see also `http://www.gaussian.com/` .