Supporting Information for:

The Role of Ligand-Driven Conformational Changes in Enzyme Catalysis: Modeling the Reactivity of the Catalytic Cage of Triosephosphate Isomerase

Yashraj S. Kulkarni[1], Qinghua Liao[1], Fabian Byléhn[1,2], Tina L. Amyes[3], John P. Richard[3,*] and Shina C. L. Kamerlin[1,*]

[1] Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, BMC Box 596, S-751 24 Uppsala, Sweden. [2] Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom. [3] Department of Chemistry, University at Buffalo, SUNY, Buffalo, New York 14260-3000, United States.

Corresponding authors e-mail addresses:

kamerlin@icm.uu.se

jrichard@buffalo.edu

Table of Contents

Extended Methodology

The empirical valence bond calculations (EVB)[1-2] in the present work were performed using an identical protocol to that used to model the reactivity of the full substrates, as described in detail in refs. [3-4]. These references describe all system setup, protonation patterns used for key ionizable residues, extended simulation details, and the EVB parameters used to describe the reactions of the full substrates DHAP and GAP as studied in our previous work.[3-4]

In brief, all simulations were performed using the 1.2 Å resolution structure of TIM in complex with DHAP (PDB ID: 1NEY[5-6]), which provides also a model based on which to generate the starting structures for simulations with the substrate pieces, by manual truncation of the $\mathbf{E_C}$•GAP and $\mathbf{E_C}$•DHAP complexes from our previous work. The systems were prepared for simulations as described in ref. [3-4], and each complex was solvated in a water sphere of a radius of 20 Å with TIP3P water molecules, centered on the C1 atom of the glycolaldehyde (GA) substrate (see **Figure S7** for atom numbering). All protein atoms outside of the water sphere were restrained to their crystallographic positions to avoid instabilities during the simulations. All the simulations in this work were performed using the OPLS-AA force field,[7-8] and OPLS-AA compatible parameters for the deprotonation of the substrate pieces were obtained using Schrödinger's Macromodel 9.1[9] and the standard RESP procedure,[10] respectively. The parameters used to describe the reaction of the substrate pieces are all included in **Tables S4**-**S15**, and the corresponding parameters used to describe the reaction of the full substrates are presented in the **Supporting Information** of our previous work.[3-4]

All EVB simulations were performed with the Q simulation package[11] with a time step of 1 fs. Each system was gradually heated up from 0.01 to 300 K over 140 ps of simulation time, while positional restraints of 200 kcal mol^{-1} $Å^{-2}$ on all heavy atoms were gradually dropped until only

weak restraints of 0.5 kcal mol^{-1} $Å^{-2}$ were maintained on only the reacting atoms. These weak restraints were maintained in the subsequent molecular dynamics and EVB simulations. Once the system had been heated to 300 K, a further 40 ns MD simulation was performed to equilibrate the system in preparation for subsequent EVB simulations. The endpoint of the equilibration was then used to generate ten independent starting structures for EVB simulations by reseeding for 110 ps with ten new random seeds. The entire procedure was repeated three times to produce thirty equilibrated starting structures in total. The subsequent EVB simulations were performed using the standard EVB free energy perturbation/umbrella sampling (EVB-FEP/US) procedure, the details of which have been described elsewhere,[1-2] and the successful application of EVB to TIM has been demonstrated both by us[3-4] and by others.[12] The EVB simulations were performed in 51 mapping windows of 200 ps length each, resulting in a total of 120 ns equilibration time and 306 ns EVB sampling time per system, and a total of 720 ns equilibration and 1.836 μs simulation time over all systems (taking into account that two different GA positions were tested for each of the simulations with the substrate pieces). As in our previous work,[3-4] the calculations were equilibrated at the EVB transition state, and the subsequent EVB trajectories were propagated from the transition state in order to both save simulation time and improve convergence. Finally, all energy analysis of the resulting EVB trajectories was performed using the QCalc module of Q,[11] the structural analysis was performed using VMD 1.9.1,[13] and the root mean square fluctuation (RMSF) and root mean square deviations (RMSD) analysis of TIM during the simulations (**Figures S3**-**S6**) was performed using GROMACS 2016.4.[14-15]

Finally, in order to determine the conformational space sampled for the substrate pieces GA and GA•HP_i, as well as the full substrate GAP (**Figure 2** of the main text), the two distances between the donor carbon atoms of the substrates and the C_α atoms of D111 in Chain B and I19 in Chain A,

respectively, were calculated at the Michaelis complexes and transition states of 30 independent EVB trajectories for each system (corresponding to 6000 individual snapshots per system). These distances were selected as both residues I19 (Chain A) and D111 (Chain B) fall into the "excluded" region of the simulation sphere, which is constrained to crystallographic coordinates during the simulations, as described above. This gave us a fixed reference point against which to extract the conformational space of the different ligands during our simulations. The two calculated distances were then binned in two dimensions, from which the number of conformations that fell into each grid were determined. The color scale used on this figure is yellow -> green -> purple (from most to least populated), with white indicating regions of conformational space that were not sampled, as shown on the 2D landscapes in **Figure 2** of the main text.

Supplementary Figures

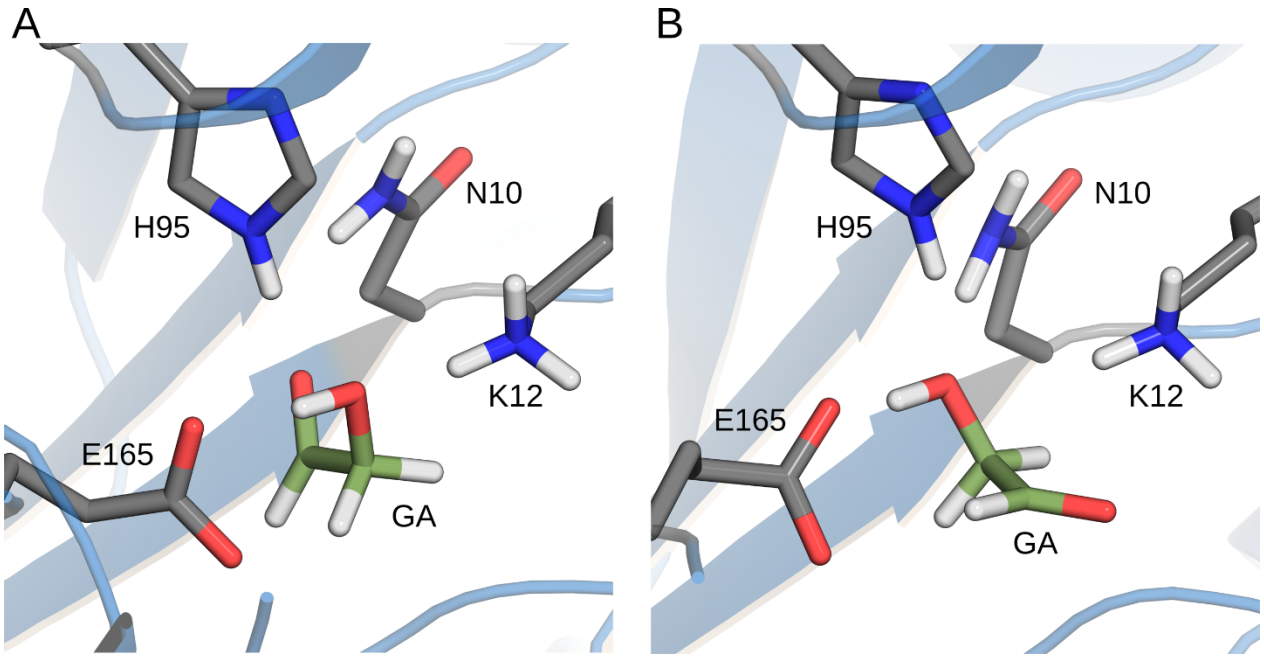

Figure S1: Structural representations of the Michaelis complexes between TIM and GA in the (**A**) GAP-like, and (**B**) DHAP-like conformations at the active site of TIM, as described in the main text. The structures shown here are endpoints of EVB trajectories propagated from the transition state for the deprotonation reaction, as described in the **Extended Methodology** section. The positions of catalytic side chains – E165, H95, N10 and K12 – are also shown here. For averages of the interacting distances between GA and the amino acid side chains illustrated here, see **Table S1** (these distances have not been annotated on the figure for clarity).

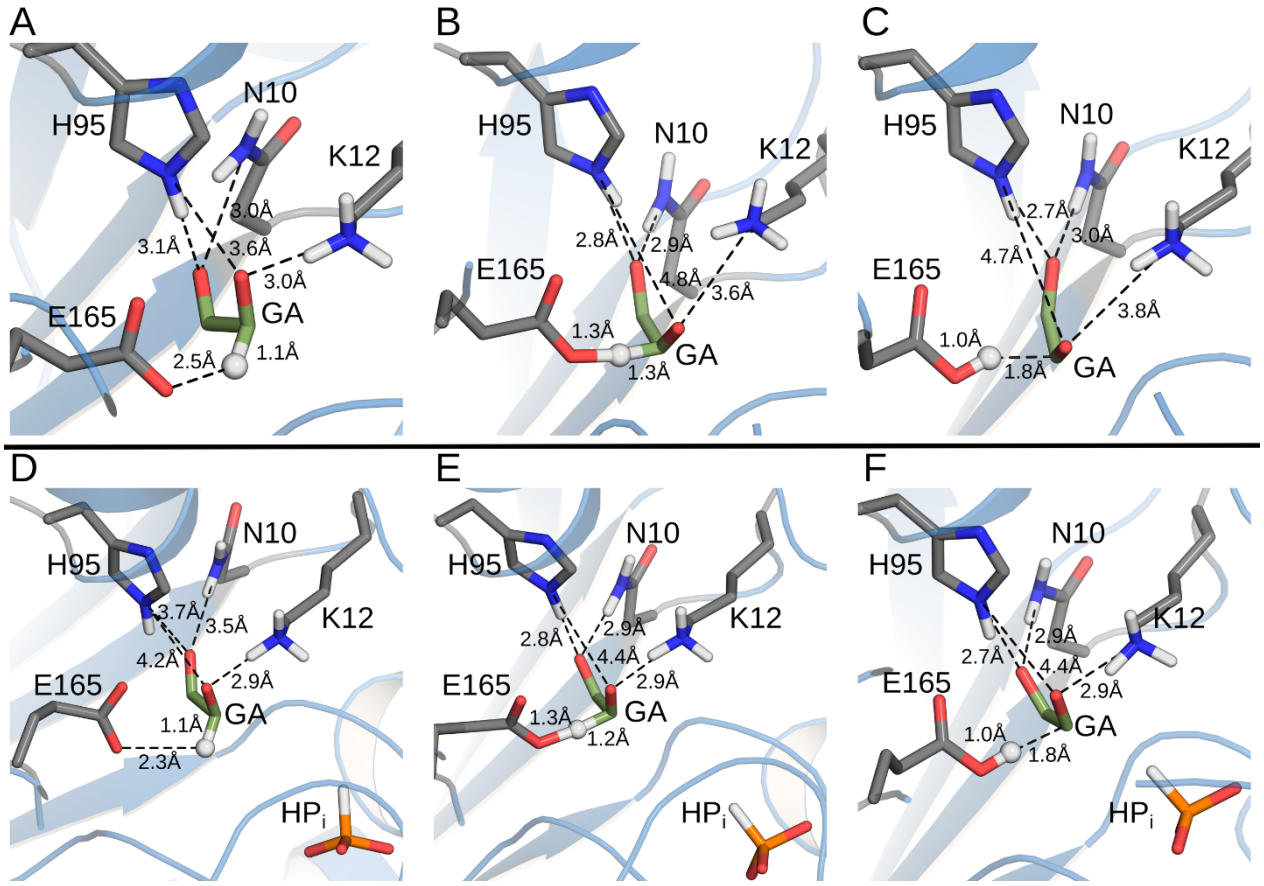

Figure S2: Representative structures of the stationary points for the deprotonation of GA (**A-C**) in the absence of and (**D-F**) in the presence of the phosphite dianion. The Michaelis complexes (**A**,**D**), transition states (**B**,**E**) and intermediate complexes (**C**,**F**) are shown here. The distances annotated on this figure are average distances over the entire simulation trajectory. These structures were obtained from equilibrated structures of the transition state for the deprotonation, as well as the endpoints (Michaelis complex and intermediate) of corresponding EVB trajectories propagated from this transition state as described in the **Methodology** section. The donor-hydrogen and acceptor-hydrogen distances involved in the proton transfer reactions are presented in **Table S1**, and the corresponding average key enzyme-substrate distances are presented in **Table S2**.

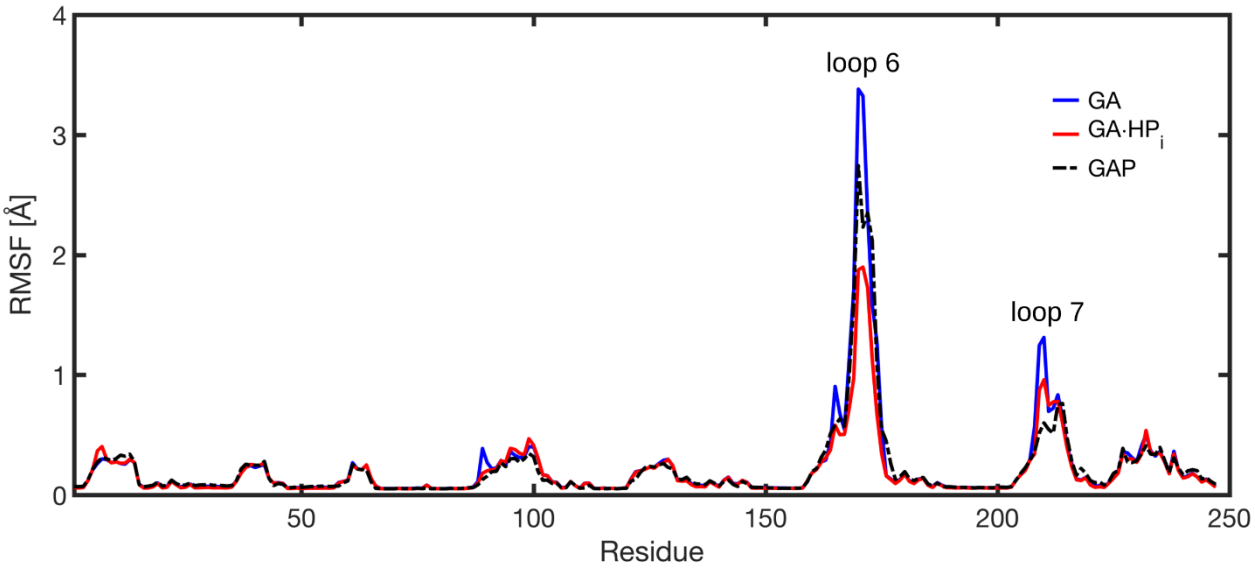

Figure S3: Root mean square fluctuations (RMSF) of the C_α–atoms at the Michaelis complexes for the TIM-catalyzed deprotonation of GA (GAP-like conformation) – both in the absence and presence of HP_i – as well as the full substrate GAP. The data was determined for a single subunit of the enzyme, focused on the flexible region of the simulation sphere (all residues within 20 Å of the simulation center, as described in the **Extended Methodology** section above). Here, the total fluctuations observed for the wild-type enzyme are shown. Data was obtained from running 30 independent EVB trajectories from the corresponding EVB transition state for each system, and averaged over 6 ns of total simulation time per system (corresponding to 6000 individual snapshots per system). As can be seen from this figure, both loop 6 and loop 7 of TIM are most flexible when only GA is bound (for a discussion of the structural and catalytic importance of these loops, see *e.g.* ref. [16]). Interestingly, loop 6 is *least flexible* in the presence of GA and HP_i, but this is likely due to the fact that unlike the full substrate GAP, the two fragments can move independently of each other and the phosphite ion can bind in such a way as to keep loop 6 more closed. The greater flexibility of the protein in the presence of only GA is likely linked to the weaker electrostatic interactions illustrated in **Figure 1** of the main text.

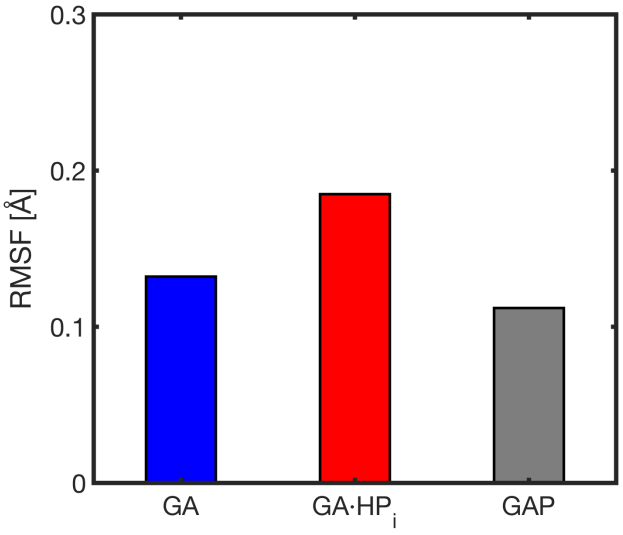

Figure S4: Root mean square fluctuations (RMSF) of the donor carbon atom of the substrate piece GA (GAP-like conformation) – both in the absence and presence of HP_i – as well as the full substrate GAP, at the respective Michaelis complexes for each system. Data was obtained from running 30 independent EVB trajectories from the corresponding EVB transition state for each system, and averaged over 6 ns of total simulation time per system (corresponding to 6000 individual snapshots per system). As can be seen from this figure, the substrate piece is slightly more flexible in the active site than the full substrate GAP.

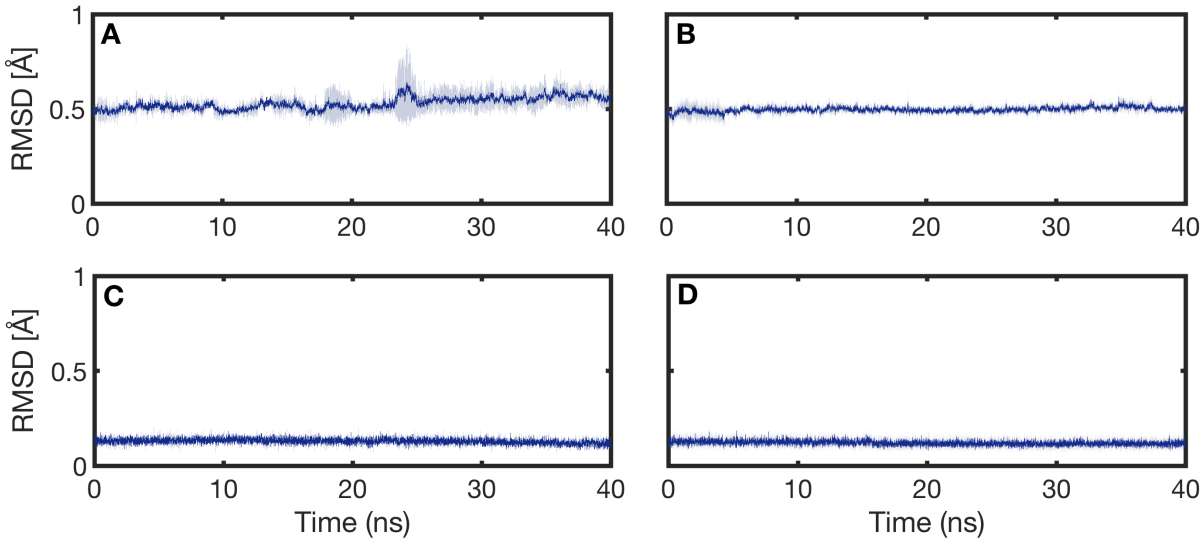

Figure S5: Root mean square deviations (RMSD) of (A, B) the backbone atoms at the transition state for the TIM-catalyzed deprotonation of GA (GAP-like conformation) – both in the absence and presence of HP_i – determined for a single subunit of the enzyme, focused on the flexible region of the simulation sphere (all residues within 20 Å of the simulation center, as described in the **Extended Methodology** section above). Here, the total fluctuations observed for the wild-type enzyme are shown. Shown here are also (C, D) the corresponding RMSD values for just the substrate piece GA. Data was collected every 5 ps from the initial equilibration runs. The data shown here are averages over three individual 40 ns MD simulations per system, *i.e.* a total of 120 ns of simulation time per system. The shaded regions indicate the standard deviation per point over all three trajectories.

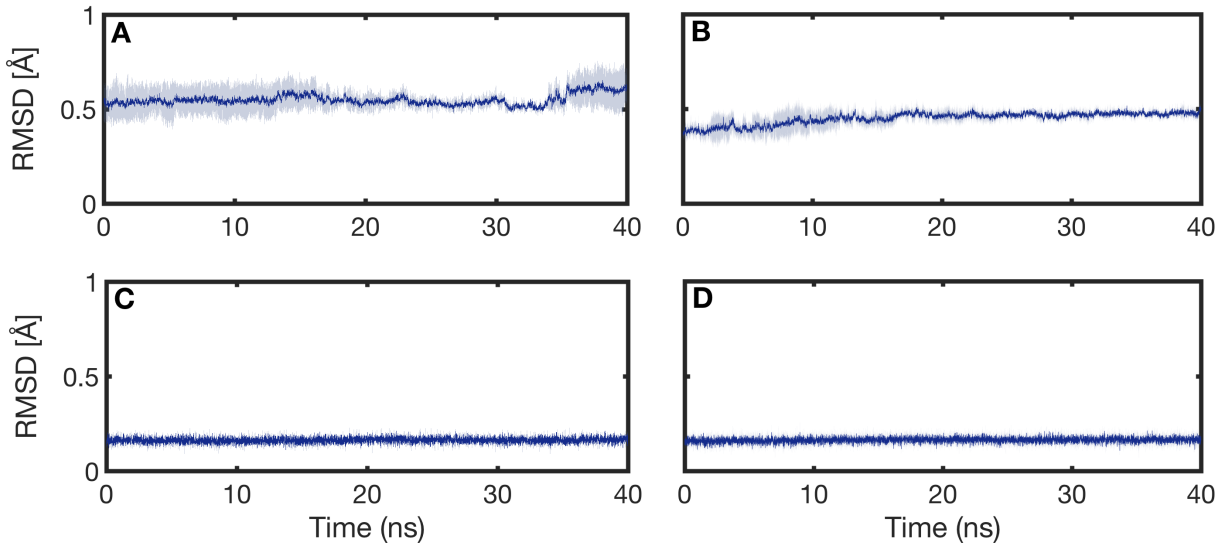

Figure S6: Root mean square deviations (RMSD) of (A, B) the backbone atoms at the transition state for the TIM-catalyzed deprotonation of GA (DHAP-like conformation) – both in the absence and presence of HP_i – determined for a single subunit of the enzyme, focused on the flexible region of the simulation sphere (all residues within 20 Å of the simulation center, as described in the **Extended Methodology** section above). Here, the total fluctuations observed for the wild-type enzyme are shown. Shown here are also (C, D) the corresponding RMSD values for just the substrate piece GA. Data was collected every 5 ps from the initial equilibration runs. The data shown here are averages over three individual 40 ns MD simulations per system, *i.e.* a total of 120 ns of simulation time per system. The shaded regions indicate the standard deviation per point over all three trajectories.

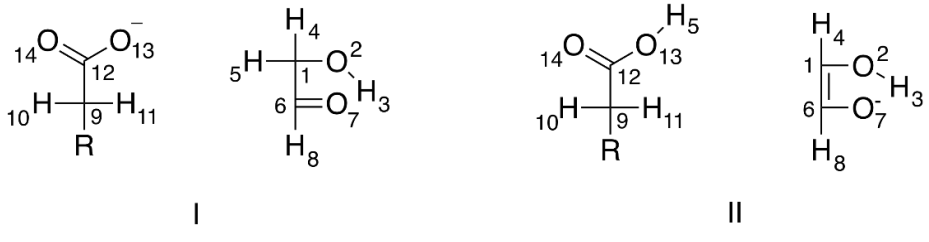

Figure S7: Numbering of atoms constituting the reacting part of the system for the valence bond states describing the deprotonation of GA. State I corresponds to the Michaelis complex whereas State II corresponds to the intermediate resulting after the proton transfer step.

Supplementary Tables

Table S1: Average values of key distances and angles between key reacting atoms at the Michaelis complex (MC), transition state (TS) and the intermediate state (IS) for the deprotonation of GA (in the preferred GAP-like conformation) in the absence and presence of phosphite dianion (HP_i), as well as the full substrate GAP, by wild type *y*TIM.[a]

Reacting State	Distance/Angle	GA	GA•HP_i	GAP
		Water Reaction[b]		
MC	D-H	1.09 ± 0.03		1.09 ± 0.03
	A-H	3.36 ± 0.89		2.33 ± 0.34
	D-A	3.41 ± 0.36		3.23 ± 0.28
	D-H..A	83.9 ± 42.1		141.3 ± 16.7
TS	D-H	1.31 ± 0.06		1.39 ± 0.06
	A-H	1.22 ± 0.04		1.18 ± 0.04
	D-A	2.51 ± 0.05		2.56 ± 0.05
	D-H..A	170.5 ± 4.9		168.7 ± 5.5
IS	D-H	1.82 ± 0.13		2.00 ± 0.15
	A-H	1.00 ± 0.03		0.99 ± 0.03
	D-A	2.77 ± 0.11		2.91 ± 0.13
	D-H..A	161.1 ± 8.6		152.9 ± 11.2
		TIM-catalyzed Reaction		
MC	D-H	1.09 ± 0.03	1.09 ± 0.03	1.09 ± 0.03
	A-H	2.46 ± 0.46	2.34 ± 0.44	2.04 ± 0.17
	D-A	3.08 ± 0.18	3.12 ± 0.31	3.03 ± 0.17
	D-H..A	117.6 ± 22.7	131.6 ± 21.4	151.9 ± 13.3
TS	D-H	1.27 ± 0.06	1.24 ± 0.05	1.32 ± 0.06
	A-H	1.25 ± 0.05	1.29 ± 0.05	1.23 ± 0.04
	D-A	2.50 ± 0.04	2.51 ± 0.04	2.53 ± 0.04

	D-H..A	171.3 ± 4.6	169.6 ± 5.1	171.0 ± 4.7
IS	D-H	1.79 ± 0.13	1.79 ± 0.12	1.98 ± 0.18
	A-H	1.00 ± 0.03	1.00 ± 0.03	0.99 ± 0.02
	D-A	2.75 ± 0.11	2.76 ± 0.11	2.88 ± 0.13
	D-H..A	163.5 ± 9.1	164.4 ± 7.8	154.1 ± 13.1

[a] D-H, A-H and D-A are the donor-hydrogen, acceptor-hydrogen and donor-acceptor distances, respectively, in Å. D-H…A denotes the corresponding angle between the donor atom, the H-atom being transferred, and the acceptor atom, in °. All values are averages and standard deviations over 30 independent EVB trajectories, as described in the **Extended Methodology** section. [b] In the case of the non-enzymatic reaction, the reacting distances and reacting angle are the same with and without phosphite, and therefore only one set of values was shown here (only the reaction of GA itself was studied in the absence of the enzyme, for calibration of the corresponding EVB parameters).

Table S2: Average values of key distances between atom pairs of the substrate and the side-chains of H95, N10 and K12 at the Michaelis complex (MC), transition state (TS) and the intermediate state (IS) for the deprotonation of GA (in the preferred GAP-like conformation) in the absence and presence of the phosphite dianion (HP_i), as well as the full substrate GAP, by wild type *y*TIM.[a]

Substrate	Residue	Atom Pairs	MC	TS	IS
GA (GAP-like)	H95	NE2—O2	3.60 ± 0.67	4.84 ± 0.33	4.71 ± 0.62
		NE2—O7	3.05 ± 0.44	2.79 ± 0.11	2.73 ± 0.09
	K12	NZ—O2	3.01 ± 0.45	3.58 ± 0.41	3.84 ± 0.68
	N10	NZ—O7	2.98 ± 0.38	2.90 ± 0.15	2.95 ± 0.19
GA•HP_i (GAP-like)	H95	NE2—O2	4.24 ± 0.88	4.38 ± 0.49	4.38 ± 0.68
		NE2—O7	3.71 ± 0.94	2.81 ± 0.13	2.73 ± 0.09
	K12	NZ—O2	2.87 ± 0.18	2.91 ± 0.19	2.91 ± 0.35
	N10	NZ—O7	3.45 ± 0.84	2.92 ± 0.21	2.90 ± 0.21
GAP	H95	NE2—O2	3.53 ± 0.44	3.49 ± 0.25	3.53 ± 0.21
		NE2—O7	2.87 ± 0.19	2.73 ± 0.11	2.72 ± 0.09
	K12	NZ—O2	3.37 ± 0.49	3.32 ± 0.24	3.36 ± 0.23
	N10	NZ—O7	3.02 ± 0.20	2.88 ± 0.13	2.89 ± 0.15

[a] NE2, ND2 and NZ atoms denote the N atom at the epsilon position in H95, the N atom of the carboxamide group of N10 and the N atom of the amino group of K12 respectively. [b] The data for GAP and DHAP were originally presented in refs. [3-4]. All distances are in Å and are averages and standard deviations over 30 individual EVB trajectories, as described in the **Extended Methodology** section.

Table S3: Electrostatic contribution of individual amino acids to the calculated activation free energies ($\Delta G^{\ddagger}$) for the TIM-catalyzed deprotonation of GA (preferred GAP-like conformation) in the absence and presence of the phosphite dianion (HP_i), as well as the full substrate GAP, in the wild type variant of TIM.[a]

Residue	GA	GA•HP_i	GAP[b]
N10	-1.2 ± 0.1	-0.9 ± 0.3	-0.9 ± 0.1
K12	-1.9 ± 0.3	-2.2 ± 0.3	-2.7 ± 0.1
H95	-0.9 ± 0.3	-0.6 ± 0.4	-0.9 ± 0.1
S96	0.4 ± 0.1	0.7 ± 0.3	0.3 ± 0.1
E97	0.7 ± 0.1	0.6 ± 0.2	0.8 ± 0.1
R99	0.4 ± 0.1	0.6 ± 0.2	0.6 ± 0.1
E129	-0.6 ± 0.0	-0.6 ± 0.1	-0.7 ± 0.1
G210	0.4 ± 0.3	0.1 ± 0.4	-0.3 ± 0.1
G232	-0.1 ± 0.0	-0.1 ± 0.1	-0.3 ± 0.1
K237	-0.2 ± 0.0	-0.3 ± 0.1	-0.5 ± 0.1

[a] All data is shown as averages over 30 independent EVB trajectories, in kcal·mol^{-1}. All values were obtained using the linear response approximation,[5-7] scaling assuming a dielectric constant of 4 for the active site, as described in the main text. [b] Data for the full substrate GAP was originally presented in refs. [3-4]. Shown here are also the total contributions from all residues for each reaction.

Empirical Valence Bond (EVB) Parameters Used to Describe the Deprotonation of GA

Table S4: EVB off-diagonal element (H_{ij}) and gas phase shift (α_i) parameters. These parameters were fitted to reproduce an activation free energy of $\Delta G^{\ddagger}$ = 24.0 kcal·mol^{-1} and a reaction free energy of ΔG_0 = 16.0 kcal·mol^{-1}, based on the energetics of the uncatalyzed deprotonation of the full substrate GAP.[3-4, 12, 17]

Substrate	H_{ij} (kcal·mol^{-1})	α_i (kcal·mol^{-1})
GA	58.6	191.6

Table S5: List of the atom types and van der Waals parameters used to describe atoms constituting the reacting part of the system.

Type	A_i (kcal$^{1/2}$·mol$^{1/2}$·Å^6)	B_i (kcal$^{1/2}$·mol$^{1/2}$·Å^3)	C_i (kcal·mol^{-1})	α_i (Å^2)	$A_{1\text{-}4}$ (kcal$^{1/2}$·mol$^{1/2}$Å^3)	$B_{1\text{-}4}$ (kcal$^{1/2}$·mol$^{1/2}$·Å^3)	Mass (a.u.)
C2	1802.24	34.18	1	2.5	1274.38	24.17	12.01
CDH	1103.59	24.67	180	2.6	780.35	17.44	12.01
CT	944.52	22.03	91	2.5	667.88	15.58	12.01
HA	69.58	4.91	1	2.5	49.20	3.47	1.01
HC	84.57	5.41	5	2.5	59.80	3.83	1.01
HDH	109.18	6.99	1	2.5	77.20	4.94	1.01
HO	0.00	0.00	5	2.5	0.00	0.00	1.01
O	616.44	23.77	1	2.5	435.89	16.81	16.00
O1	616.44	23.77	1	2.5	435.89	16.81	16.00
O2	616.44	23.77	250	1.5	435.89	16.81	16.00
ODE	601.15	22.27	250	1.5	425.08	15.74	16.00
ODH	976.93	31.26	1	2.5	690.79	22.10	16.00
OH	760.65	25.05	1	2.5	537.86	17.71	16.00

[a] For all atoms except the reacting atoms, a standard 6-12 Lennard-Jones potential was used. In the case of the reacting atoms, which change bonding patterns between atoms *i* and *j*, an alternate function of the form $V_{react} = C_i C_j exp(-\alpha_i \alpha_j r_{ij})$ was used to prevent artificial repulsion between these atoms as bonding patterns change. r_{ij} denotes the distance (Å) between atoms *i* and *j*. For atom type assignment see **Table S6**.

Table S6: Atom types in the different VB states used to describe the deprotonation of GA.[a]

Atom number	State I	State II
1	CT	CDH
2	OH	OH
3	HO	HO
4	HC	HA
5	HC	HO
6	C2	CDH
7	O	ODH
8	HC	HDH
9	CT	CT
10	HC	HC
11	HC	HC
12	C2	C2
13	O2	ODE
14	O2	O1

[a] See **Figure S7** for the atom numbering, **Table S5** for the corresponding van der Waals parameters and **Table S7** for the corresponding partial charges.

Table S7: Atomic partial charges in the different VB states used to describe the deprotonation of GA.[a]

Atom number	State I	State II
1	0.2981	-0.3830
2	-0.6880	-0.5843
3	0.4414	0.4002
4	-0.0240	0.1278
5	-0.0240	0.4500
6	0.5176	0.2468
7	-0.4723	-0.7567
8	-0.0488	-0.0508
9	-0.2200	-0.1200
10	0.0600	0.0600
11	0.0600	0.0600
12	0.7000	0.5200
13	-0.8000	-0.5300
14	-0.8000	-0.4400

[a] For the corresponding atom numbering, see **Figure S7**, and for details of how these charges were derived, see the main text.

Table S8: Bond types and corresponding parameters for covalent bonds of the reacting part of the system used to describe the deprotonation of GA.[a]

Bond type	D_e (kcal·mol^{-1})	α (Å^{-2})	r_0 (Å)	k_b (kcal·mol^{-1}·Å^{-2})	r_0 (Å)
0			Not Set		
1				1312	1.250
2				900	1.364
3				1140	1.229
4	245.8	1.5	0.945		
5				640	1.410
6				900	1.370
7				680	1.090
8				680	1.080
9	85.0	2.0	1.090		
10				634	1.522
11				1098	1.340
12				1140	1.229
13				900	1.370
14				1106	0.945
15				536	1.529

[a] The bonds between non-reacting atoms are described using harmonic potentials $V_{\text{Harmonic}} = 0.5k_b\,(r_{ij} - r_0)^2$, while bonds between reacting atoms are described using Morse potentials $V_{\text{Morse}} = D_e\,[1 - exp[-\alpha(r_{ij} - r_0)]]^2$. For the bond type assignments, see **Table S9**.

Table S9: Bond types used to describe the covalent bonds of the reacting part of the system in the different VB states, for the deprotonation of GA.

Atom number		Bond type	
#1	#2	State I	State II
1	2	5	6
1	4	7	8
1	5	9	0
1	6	10	11
2	3	20	20
6	7	12	13
6	8	7	8
9	10	7	7
9	11	7	7
9	12	10	10
12	13	1	2
12	14	1	3
13	5	0	4
R[a]	9	15	15

[a] R corresponds to C_β in E165, which has no assigned atom number as it is not part of the reacting system (the corresponding PDB atom number is used in the setup of the simulations). For the corresponding atom numbering, see **Figure S7**.

Table S10: Angle types and the corresponding parameters used for bending adjacent bonds in the reacting part of the system used to describe the deprotonation of GA.[a]

Angle type	k_a (kcal·mol^{-1}·rad^{-2})	Θ (°)	Angle type	k_a (kcal·mol^{-1}·rad^{-2})	Θ (°)
0	Not Set		8	70.0	109.50
1	140.0	117.00	9	70.0	114.50
2	160.0	120.40	10	100.0	109.50
3	70.0	113.00	11	66.0	107.80
4	110.0	108.50	12	70.0	123.00
5	140.0	123.00	13	160.0	126.00
6	70.0	115.00	14	75.0	110.70
7	70.0	120.00	15	126.0	111.10

[a] The angle potential is described using the potential $V_{angle} = 0.5\ \Sigma k_a(\Theta-\Theta_0)^2$. For the assignment of angle types, see **Table S11**.

Table S11: Angle type assignment in the different VB states used to describe the deprotonation of GA.[a]

Atom number			Angle type	
#1	#2	#3	State I	State II
1	2	3	4	3
1	6	7	2	5
1	6	8	6	7
2	1	4	8	9
2	1	5	8	0
2	1	6	10	5
4	1	5	11	0
4	1	6	8	7
5	1	6	8	0
7	6	8	12	9
9	12	13	1	1
9	12	14	1	2
10	9	11	11	11
10	9	12	8	8
11	9	12	8	8
12	13	5	0	3
13	12	14	13	13
R[a)]	9	10	14	14
R[a)]	9	11	14	14
R[a)]	9	12	15	15

[a]See **Figure S7** for the atom numbering. R corresponds to C_β in E165, which has no assigned atom number as it is not a part of the reacting system (the corresponding PDB atom number is used in the setup of the simulations).

Table S12: Torsion types and the corresponding parameters for rotation of dihedrals in the reacting part of the system used to describe the deprotonation of GA.[a]

Torsion type	V_1	V_2	V_3	Torsion type	V_1	V_2	V_3
	0.5·barrier height (kcal·mol^{-1})				0.5·barrier height (kcal·mol^{-1})		
0		Not Set		7	0.0000	7.0000	0.0000
1	0.0000	0.2730	0.0000	8	0.0000	0.0000	0.2340
2	0.0000	0.5830	0.0000	9	0.0000	0.0000	0.2250
3	1.5000	2.4500	0.0000	10	0.0000	2.7500	0.0000
4	0.0000	0.0000	0.0000	11	-0.1780	-0.0870	0.2460
5	0.0000	2.4500	0.0000	12	0.7500	0.8410	0.0000
6	2.1590	0.0000	0.0000	13	0.0000	0.0000	0.1800

[a] The torsion angle potential is described using the potential $V_{torsion} = V_1(1+\cos(n\varphi-\delta)) + V_2(1+\cos2(n\varphi-\delta)) + V_3(1+\cos3(n\varphi-\delta))$. Here n is the periodicity (number of maxima per turn) and δ is the phase shift. For assignment of torsion types, see **Table S13**.

Table S13: Torsion type assignment in the different VB states used to describe the deprotonation of GA.[a]

Atom number				Torsion type	
#1	#2	#3	#4	State I	State II
2	1	6	7	6	7
2	1	6	8	8	7
3	2	1	4	9	10
3	2	1	5	9	0
3	2	1	6	11	12
4	1	6	7	4	7
4	1	6	8	13	7
5	1	6	8	13	0
5	1	6	7	4	0
9	12	13	5	0	3
10	9	12	13	4	4
10	9	12	14	4	4
11	9	12	13	4	4
11	9	12	14	4	4
14	12	13	5	0	5
R [a)]	9	12	13	1	2
R[a)]	9	12	14	1	1

[a] See **Figure S7** for the atom numbering. R corresponds to C_β in E165, which has no assigned atom number as it is not a part of the reacting system (the corresponding PDB atom number is used in the setup of the simulations).

Table S14: Improper torsion types and the corresponding parameters in the reacting part of the system used to describe the deprotonation of GA.[a]

Improper torsion type	k_i (kcal·mol^{-1}·rad^{-2})	τ_0 (°)
0	Not set	
1	10.5	180
2	15.0	180

[a] The improper torsion potential is described using the potential $V_{improper} = k_i(\tau - \tau_0)^2$, where k_i is the force constant and τ is the equilibrium angle (in degrees). For assignment of improper torsion types, see **Table S15**.

Table S15: Improper torsion type assignment in the different VB states used to describe the deprotonation of GA.[a]

Atom number				Improper torsion type	
#1	#2	#3	#4	State I	State II
1	6	8	7	1	2
6	1	4	2	0	2
9	12	13	14	1	1

[a] See **Figure S7** for the corresponding atom numbering.

References

1. Warshel, A.; Weiss, R. M., *J. Am. Chem. Soc.* **1980,** *102*, 6218-6226.

2. Kamerlin, S. C. L.; Warshel, A., *WIREs Comput. Mol. Sci.* **2011,** *1*, 30-45.

3. Amrein, B. A.; Steffen-Munsberg, F.; Szeler, I.; Purg, M.; Kulkarni, Y.; Kamerlin, S. C. L., *IUCrJ* **2017,** *4*, 50-64.

4. Kulkarni, Y. S.; Liao, Q.; Petrović, D.; Krüger, D. M.; Strodel, B.; Amyes, T. L.; Richard, J. P.; Kamerlin, S. C. L., *J. Am. Chem. Soc.* **2017,** *139*, 10514-10525.

5. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., *Nucleic Acids Res.* **2000,** *28*, 235-242.

6. Jogl, G.; Rozovsky, S.; McDermott, A. E.; Tong, L., *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100*, 50-55.

7. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., *J. Am. Chem. Soc.* **1996,** *118*, 11225-11236.

8. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *J. Phys. Chem. B* **2001,** *105*, 6474-6487.

9. *Schrödinger Release 2013-1: MacroModel*, 9.1; Schrödinger, LCC: New York, NY, 2013.

10. Wang, J.; Cieplak, P.; Kollman, P. A., *J. Comput. Chem.* **2000,** *21*, 1049-1074.

11. Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J., *J. Mol. Graph. Model.* **1998,** *16*, 213-225.

12. Åqvist, J.; Fothergill, M., *J. Biol. Chem.* **1996,** *271*, 10010-10016.

13. Humphrey, W.; Dalke, A.; Schulten, K., *J. Mol. Graph.* **1996,** *14*, 33-38.

14. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., *Bioinformatics* **2013,** *29*, 845-854.

15. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., *J. Chem. Theory Comput.* **2008,** *4*, 435-437.

16. Wierenga, R. K., *Cell. Mol. Life Sci.* **2010,** *67*, 3961-3982.

17. Richard, J. P., *J. Am. Chem. Soc* **1984,** *106*, 4926-4936.