Supporting Information

Molecular dynamics study of taxadiene synthase catalysis

Andrés M. Escorcia, Jeaphianne P. M. van Rijn, Gui-Juan Cheng, Patrick Schrepfer, Thomas B. Brück, and Walter Thiel*

[*]W.T.: Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim, Germany. Email: thiel@kofo.mpg.de

Content:

1. Additional details on the MD setup

The residues kept fixed during the MD simulations correspond to residues of the N-terminal domain that are distant from the reaction core and to residues at the surface of the C-terminal domain (Figure S1). These residues are listed below:

➢ N-terminal: residues 101-130 and 537-552.
➢ C-terminal: residues 553-554, 592-594, 634, 636, 638-639, 641-642, 666-672, 675-676, 697, 699, 702, 704, 728-734, 738, 742, 778-789, 792-793, 796, 799, 803, 807, 809-818, 821 and 857-862.
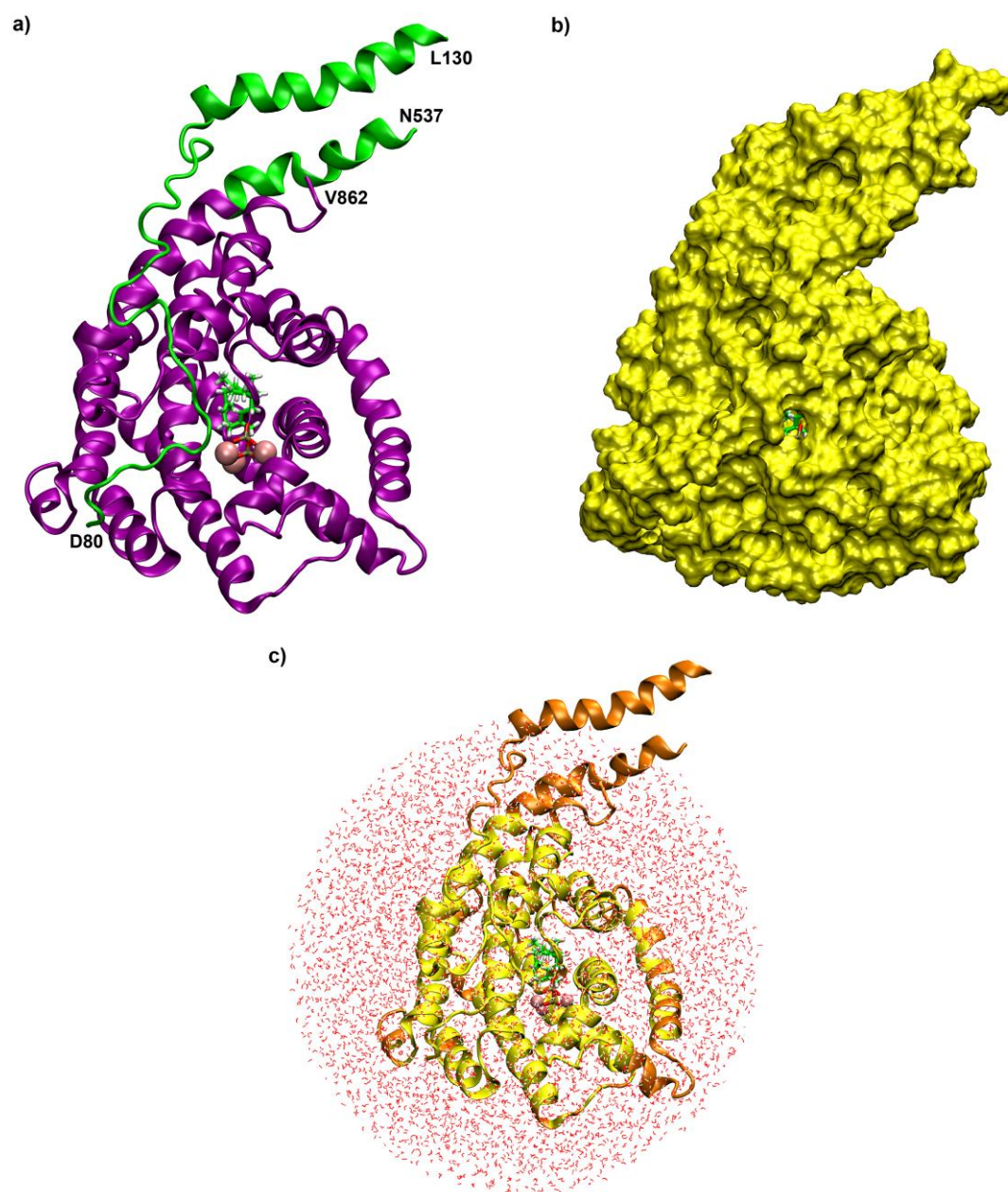
Figure S1. (a) Representative structure of TXS used in the MD simulations. It comprises the C-terminal domain (S553-V862, purple) and part of the N-terminal domain (D80-L130 and N537-Q552, green). (b) Surface representation of TXS showing more clearly that the active site is enclosed. (c) Solvated system with the protein residues kept fixed during the MD simulations depicted in orange.

2. Structural stability of the protein

Table S1

Average RMSD (with respect to the initial structure) for all heavy atoms of the protein backbone in the MD simulations of the **TXS·cation** complexes[a]

COMPLEX	RMSD (Å)		
	MD1	MD2	MD3
TXS·C	1.2 (0.2)	1.04 (0.07)	1.07 (0.08)
TXS·D1	1.1 (0.1)	0.97 (0.05)	1.2 (0.2)
TXS·F	1.1 (0.1)	0.97 (0.08)	1.07 (0.04)
TXS·E	1.1 (0.1)	1.04 (0.05)	1.03 (0.05)

[a] Standard deviations are given in parentheses. MD1-MD3 correspond to MD simulations with different initial velocity distributions.

Table S2

Average RMSD (with respect to the initial structure) for all atoms of PPi in the MD simulations of the **TXS·cation** complexes[a]

COMPLEX	RMSD (Å)		
	MD1	MD2	MD3
TXS·C	0.8 (0.2)	0.7 (0.1)	0.8 (0.2)
TXS·D1	0.5 (0.2)	0.5 (0.2)	0.4 (0.1)
TXS·F	0.6 (0.2)	0.5 (0.2)	0.6 (0.2)
TXS·E	0.5 (0.1)	0.5 (0.2)	0.5 (0.1)

[a] Standard deviations are given in parentheses. MD1-MD3 correspond to MD simulations with different initial velocity distributions. The RMSD values for PPi are lower than those for the protein, which is indicative of its high positional and conformational stability in the active site of TXS.

Table S3

Interactions identified to play a key role for the stability of the closed conformation of TXS (indicated by CLO) and for the rigidity of the cavity of the active site harboring the carbocation (indicated by RIG) in the MD simulations of the **TXS·cation** complexes

Interactions[a]		Role		Lifetime[d]			
		RIG	**CLO**	**C**	**D1**	**F**	**E**
hydrogen bonds[b]	PPi:O1,O3,O4-N757:*ND2*	X		97	99	97	99
	PPi:O5,O7-R754:*NH1,NH2*	X		100	100	100	100
	D614:OD1,OD2-R580:*NH1,NH2*	X		100	62	48	65
	D614:OD1,OD2-Y89:*OH*		X	95	97	84	67
	Y89:OH-R580:*NH1,NH2*	X	X	83	72	56	68
	S587:OG-Y841:*OH*	X	X	99	94	95	98
	N845:OD1-H90:*N,ND1*		X	85	70	64	100
	A844:O-H90:*ND1*		X	43	66	21	47
	D92:OD1,OD2-H579:*N*		X	100	96	100	100
	G715:O-Q609:*NE2*	X		92	100	93	69
	Y684:<u>OH</u>-Y688:<u>OH</u>	X		99	100	100	89
	Y835:OH-N757:*ND2*	X		92	68	91	68
water-mediated hydrogen bonds[b,c]	PPi:O1-W$_{(1,2)}$-Y835:*OH*	X		100	100	100	100
	PPi:O1-W$_{(1)}$-Y89:<u>OH</u>		X	99	65	54	70
	PPi:O6,O7-W$_{(2-3)}$-Q770:OE1,*NE2*		X	100	100	100	100
	PPi:O1-W$_{(1-3)}$-R580:*NH2*	X		77	100	100	91
	PPi:O1,O2,O3-W$_{(1-3)}$-R768:*NH1,NH2*		X	100	100	100	100
	Y841:O-W$_{(1)}$-R580:*NH2*	X	X	66	37	31	62
	S713:OG-W$_{(1-2)}$-Mg$^{2+}_C$	X		88	60	31	85
metal-ligand coordinations	N757:OD1-Mg$^{2+}_B$	X		100	100	100	100
	T761:OG1-Mg$^{2+}_B$		X	100	100	100	100
	E765:OE1,OE2-Mg$^{2+}_B$		X	100	100	100	100
	D613:OD2-Mg$^{2+}_A$	X		100	100	100	100
	D613:OD1-Mg$^{2+}_C$	X		100	100	100	100
π-stacking	Y841-F834	X	X	100	100	100	100

[a] There are pairs of residues that form different hydrogen bonds during the MD simulations (e.g. D614 and R580), some of which occur simultaneously during certain intervals of time. In this case the respective residues are specified once, and all acceptor and/or donor atoms are listed separated by commas. The same notation is used for residues coordinating to a Mg^{2+} ion through more than one atom.

[b] Atoms acting as hydrogen bond acceptors and donors are indicated in regular and italic fonts, respectively, whereas atoms changing from donor to acceptor character and vice versa during the MD simulations are underlined. A maximum distance of 3 Å between acceptor and hydrogen was chosen as criterion for the presence of a hydrogen bond.

[c] W$_{(n)}$ refers to water molecules, "n" being their number. "n" is given as a range when the number of water molecules fluctuates during and/or between the MD simulations.

[d] The entry under Lifetime is the percentage of the simulation time that the type of interaction between the respective pair of residues (independent of the atoms involved) occurs during the MD simulations. The values reported here correspond to average values obtained from the three MD runs with different seed velocities.
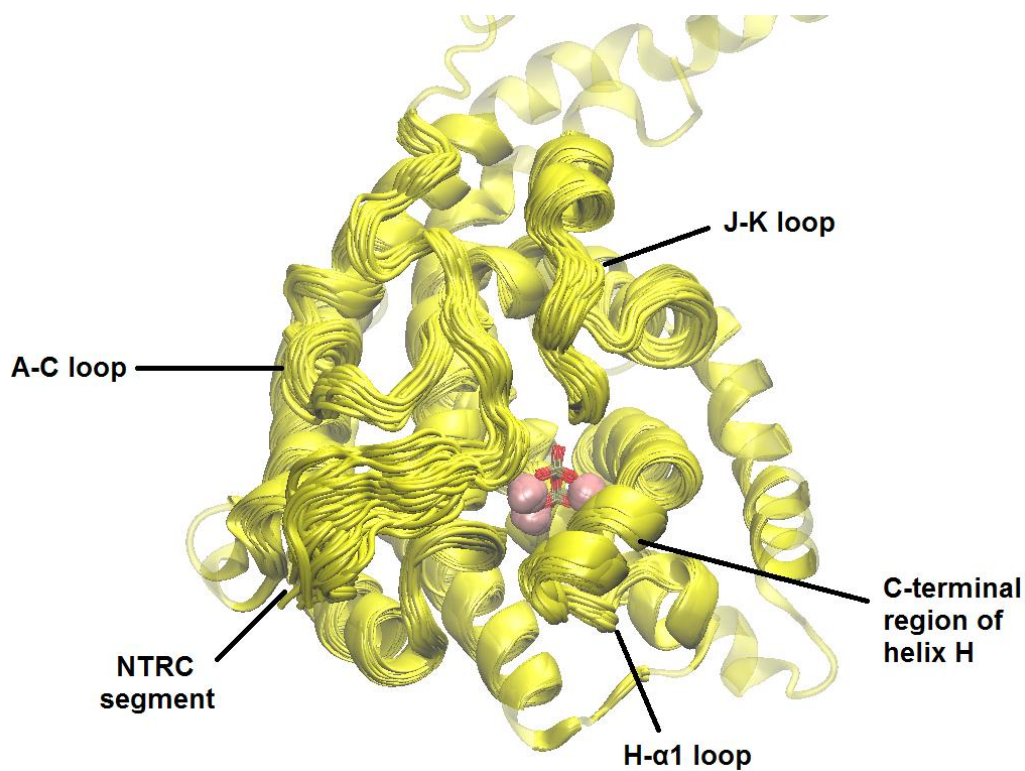
Figure S2. Superimposed structures of PPi (oxygens in red), catalytic Mg^{2+} ions (pink) and TXS (yellow), at intervals of 10 ps during one of the three MD runs of the **TXS·C** complex. All other simulations show similar superimposition profiles. Protein residues which were kept fixed during the MD simulations are shown as transparent loops (on top, see also Figure S1c).
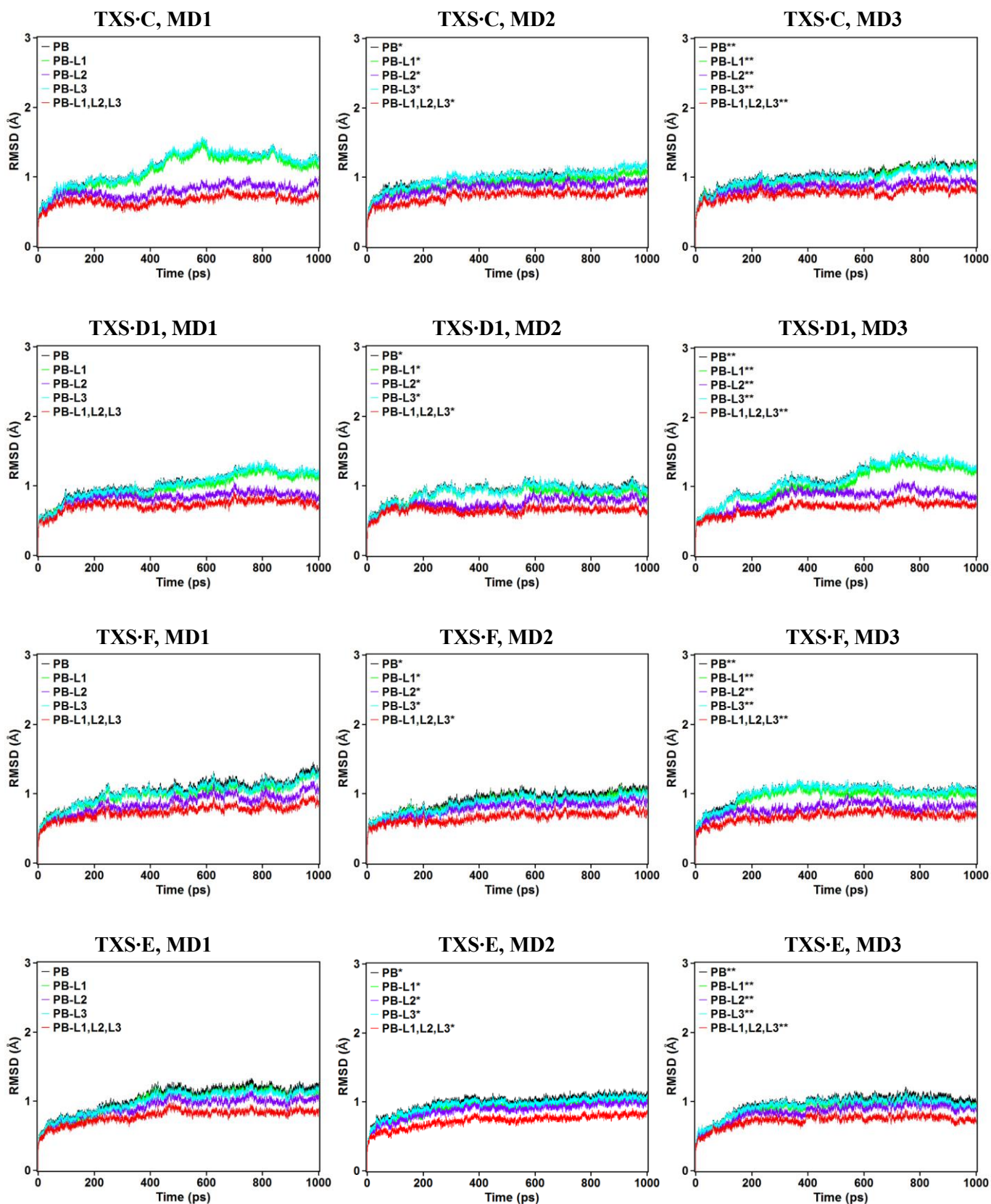
Figure S3. Time evolution of the RMSD for the protein structure in the MD simulations of the **TXS·cation** complexes. PB stands for all heavy atoms of the protein backbone, while "PB-L1, L2 and/or L3" refers to PB without the A-C loop (L1), the NTRC segment (L2) and/or the J-K loop (L3). MD1-MD3 correspond to MD simulations with different seed velocities. Superscripts * are added to distinguish between these MD simulations for more clarity. In all MD simulations the RMSD values become smaller overall when removing the contribution of these protein regions, especially that from the NTRC segment, which is indicative of its high flexibility.
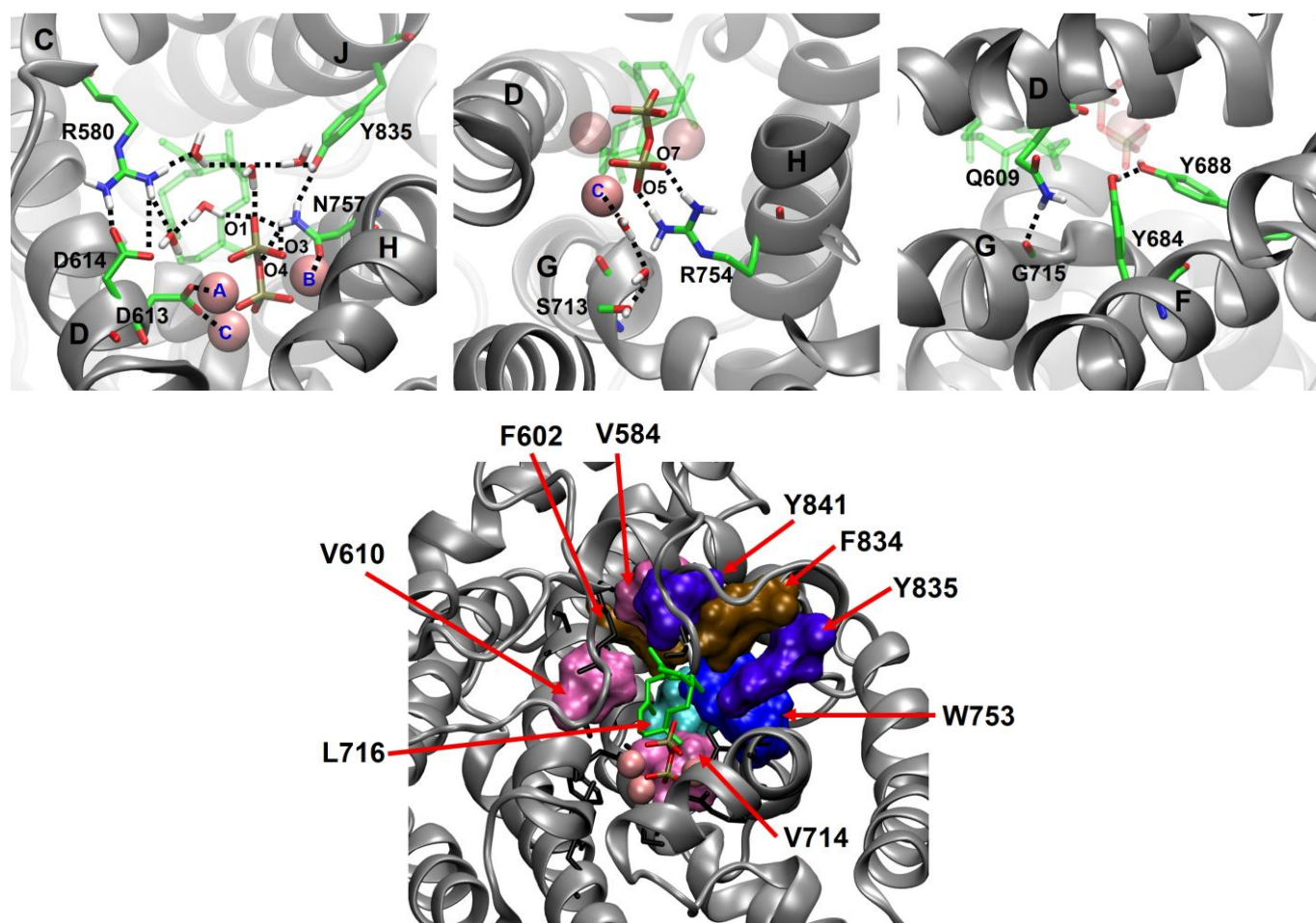
Figure S4. Representative snapshots of the **TXS·C** complex to illustrate the interactions (top) and the residues with bulky side chains (bottom) which contribute to the rigidity of the active-site cavity of TXS harboring the carbocations. Interactions that are important for both the rigidity of the cavity and the stability of the closed conformation of TXS (see Table S3) are omitted, as they are shown in Figure 4 of the main text.

3. Cation C

Table S4

Average RMSD (with respect to the initial structure) for all carbon atoms of cation **C** in the MD simulations of the **TXS·C** complex[a]

MD	RMSD (Å)
MD1	0.8 (0.2)
MD2	0.9 (0.4)
MD3	0.8 (0.2)

[a] Standard deviations are given in parentheses. MD1-MD3 correspond to MD simulations with different initial velocity distributions. The low RMSD values and standard deviations in all simulations are indicative of the high positional and conformational stability of cation **C** in the active site of TXS.
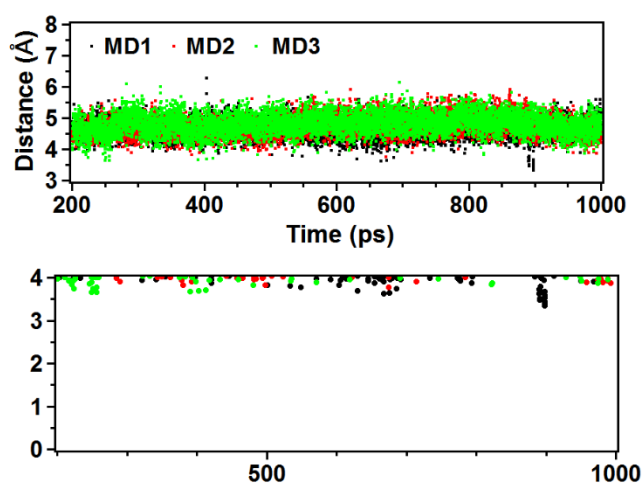
Figure S5. Time evolution of the PPi:O1-**C**:H12a distance in the MD simulations of the **TXS·C** complex. The periods of time when PPi:O1-**C**:H12a ≤ 4.0 Å are highlighted in the panel at the bottom. MD1-MD3 correspond to MD simulations with different initial velocity distributions. See Figure 6 of the main text for conventions on the atom labels.

Table S5

Average values of the relevant interatomic distances and angles involved in the possible water-assisted deprotonation paths of cation **C** (DPC) identified in the MD simulations of the **TXS·C** complex[a]

DPC (lifetime[b], %)	Distance ; Angle Definitions	Distance (Å) ; Angle (°) Values		
		MD1	**MD2**	**MD3**
DPC1 (86 - 61 - 9)	**C**:H12a-W1$_C$:O ; **C**:(C12-H12a)-W1$_C$:O	2.7 (0.2) ; 142 (12)	2.7 (0.2) ; 132 (12)	2.8 (0.2) ; 150 (11)
	W1$_C$:H-PPi:O1 ; W1$_C$:(O-H)-PPi:O1	1.7 (0.1) ; 165 (8)	1.7 (0.1) ; 165 (8)	1.7 (0.1) ; 162 (9)
DPC2 (65 – 49 – 0)	**C**:H12a-W1$_C$:O ; **C**:(C12-H12a)-W1$_C$:O	2.7 (0.2) ; 142 (11)	2.7 (0.2) ; 132 (13)	NI[c]
	W1$_C$:H-W2$_C$:O ; W1$_C$:(O-H)-W2$_C$:O	1.8 (0.1) ; 165 (8)	1.8 (0.1) ; 164 (8)	
	W2$_C$:H-D614:O$_{D1}$; W2$_C$:(O-H)-D614:O$_{D1}$	1.8 (0.1) ; 161 (10)	1.8 (0.1) ; 161 (10)	
DPC3 (0 – 0 – 74)	**C**:H12a-W1$_C$:O ; **C**:(C12-H12a)-W1$_C$:O	NI[c]	NI[c]	2.7 (0.2) ; 127 (12)
	W1$_C$:H-W2$_C$:O ; W1$_C$:(O-H)-W2$_C$:O			1.8 (0.1) ; 162 (9)
	W2$_C$:H-PPi:O1 ; W2$_C$:(O-H)-PPi:O1			1.7 (0.1) ; 162 (9)

[a] Average values were calculated considering only the frames relevant to the proposed DPCs, *i.e.* those in which the following conditions are met simultaneously: (i) the distance between the **C**:H12a and W1$_C$:O (deprotonating base) atoms is ≤ 3.0 Å, and (ii) the distance in the hydrogen bonds that may facilitate the deprotonation of cation **C** via multiple proton transfer reactions is ≤ 2.0 Å. Standard deviations are given in parentheses. See Figure 6 of the main text for conventions on the atom labels.

[b] The entries under DPC(lifetime) are the percentages of the simulation time, during which the conditions for DPCs are met (according to the geometrical criteria listed above). They are given in parentheses in the order MD1-MD2-MD3. MD1-MD3 correspond to MD simulations with different initial velocity distributions.

[c] Not identified.

* Note: one may use different cutoff values for the relevant interatomic distances, for which it is still reasonable to expect that a multiple proton transfer reaction can occur (e.g. 2.5 Å instead of 2.0 Å for the W1$_C$:H-W2$_C$:O hydrogen bond). This leads to different numerical values for the average distances and DPC lifetimes; however, the conclusions on the deprotonation of cation **C** (as described in the main text) remain the same.
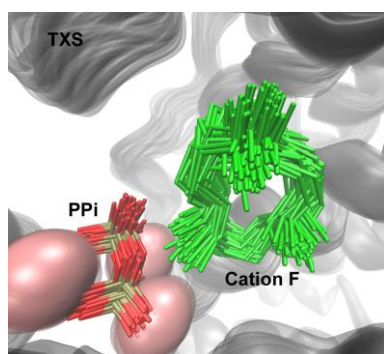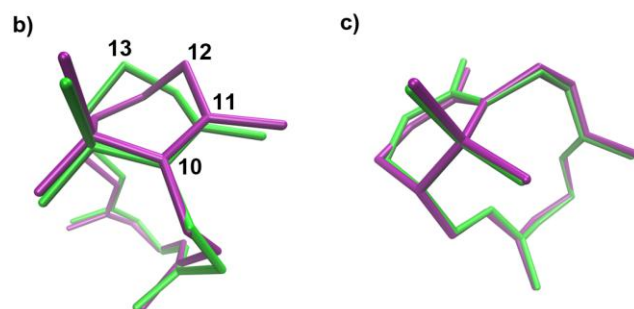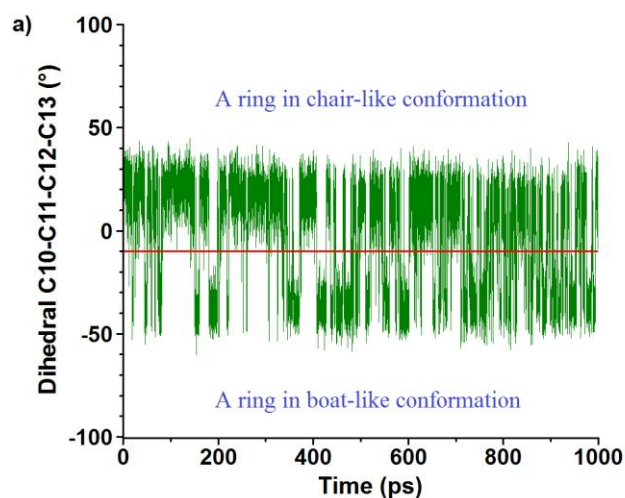
4. Cation F

Table S6
Average RMSD (with respect to the initial structure) for all carbon atoms of cation **F** in the MD simulations of the **TXS·F** complex[a]

MD	RMSD (Å)
MD1	1.2 (0.2)
MD2	0.8 (0.2)
MD3	1.1 (0.2)

[a] Standard deviations are given in parentheses. MD1-MD3 correspond to MD simulations with different initial velocity distributions.

Figure S6. Superimposed structures of the **TXS·F** complex at intervals of 10 ps during MD1. MD2 and MD3 show similar superimposition profiles.

Figure S7. (a) Time evolution of the C10-C11-C12-C13 dihedral angle of cation **F** during the MD1 simulation of the **TXS·F** complex showing the flipping of the A ring between a chair-like (F_{chair}, dihedral > -10°) and boat-like (F_{boat}, dihedral < -10°) conformation. MD2 and MD3 have similar profiles. (b) F_{chair} (green) and F_{boat} (purple) superimposed. (c) HT-QM structure of cation **F** (purple) [1] superimposed with F_{chair} (green, taken from the MD simulations of the **TXS·F** complex).

Table S7

Average values of the relevant interatomic distances and angles involved in the possible deprotonation paths of cation **F** (DPF) identified in the MD simulations of the **TXS·F** complex[a]

DPF (lifetime[b], %)	Distance ; Angle Definitions	Distance (Å) ; Angle (°) Values		
		MD1	MD2	MD3
DPF1$_{C20}$ (70 - 48 - 30)	**F**:H20-PPi:O1 ; **F**:(C20-H20)-PPi:O1	2.6 (0.2) ; 153 (13)	2.6 (0.2) ; 145 (15)	2.7 (0.2) ; 140 (12)
DPF2$_{C20}$ (27 - 9 - 3)	**F**:H20-W1$_F$:O ; **F**:(C20-H20)-W1$_F$:O	2.8 (0.2) ; 135 (15)	2.7 (0.2) ; 139 (16)	2.8 (0.2) ; 126 (11)
	W1$_F$:H-PPi:O1 ; W1$_F$:(O-H)-PPi:O1	1.7 (0.1) ; 165 (8)	1.7 (0.1) ; 165 (8)	1.7 (0.1) ; 166 (7)
DPF1$_{C2}$ (1 - 5 - 6)	**F**:H2-PPi:O1 ; **F**:(C2-H2)-PPi:O1	2.8 (0.2) ; 124 (12)	2.7 (0.2) ; 124 (14)	2.8 (0.2) ; 112 (11)
DPF2$_{C2}$ (71 - 61 - 84)	**F**:H2-W2$_F$:O ; **F**:(C2-H2)-W2$_F$:O	2.7 (0.2) ; 118 (13)	2.7 (0.2) ; 129 (12)	2.6 (0.2) ; 129 (13)
	W2$_F$:H-PPi:O1 ; W2$_F$:(O-H)-PPi:O1	1.7 (0.1) ; 166 (7)	1.7 (0.1) ; 165 (8)	1.7 (0.1) ; 166 (8)

[a] Average values were calculated considering only the frames in which the distance between the deprotonating base (PPi:O1, W1$_F$:O or W2$_F$:O) and the hydrogen atom to be abstracted from cation **F** (H20 or H2) is ≤ 3.0 Å. Regarding DPF2$_{C20/C2}$, the frames were selected by also requiring a distance ≤ 2.0 Å for the water-PPi hydrogen bond that may facilitate deprotonation of **F** via a double proton transfer reaction. Standard deviations are given in parentheses. See Figure 7 in the main text for conventions on the atom labels.

[b] The entries under DPF(lifetime) are the percentages of the simulation time, during which the conditions for DPFs are met (according to the geometrical criteria listed above). They are given in parentheses in the order MD1-MD2-MD3. MD1-MD3 correspond to MD simulations with different initial velocity distributions.

* Note: one may use different cutoff values for the relevant interatomic distances, for which it is still reasonable to expect that the reactions in question can occur (e.g. 3.3 Å instead of 3.0 Å for **F**:H20-PPi:O1 and **F**:H20-W1$_F$:O). This leads to different numerical values for the average distances and DPC lifetimes; however, the conclusions on the deprotonation of cation **F** (as described in the main text) remain the same.
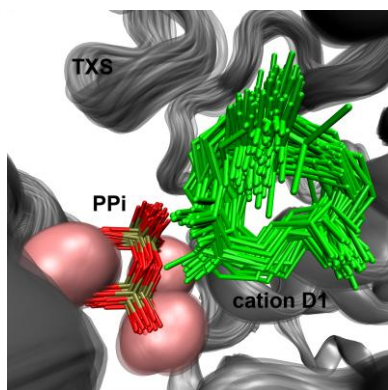
11

5. Cation D1

Table S8

Average RMSD (with respect to the initial structure) for all carbon atoms of cation **D1** in the MD simulations of the **TXS·D1** complex[a]
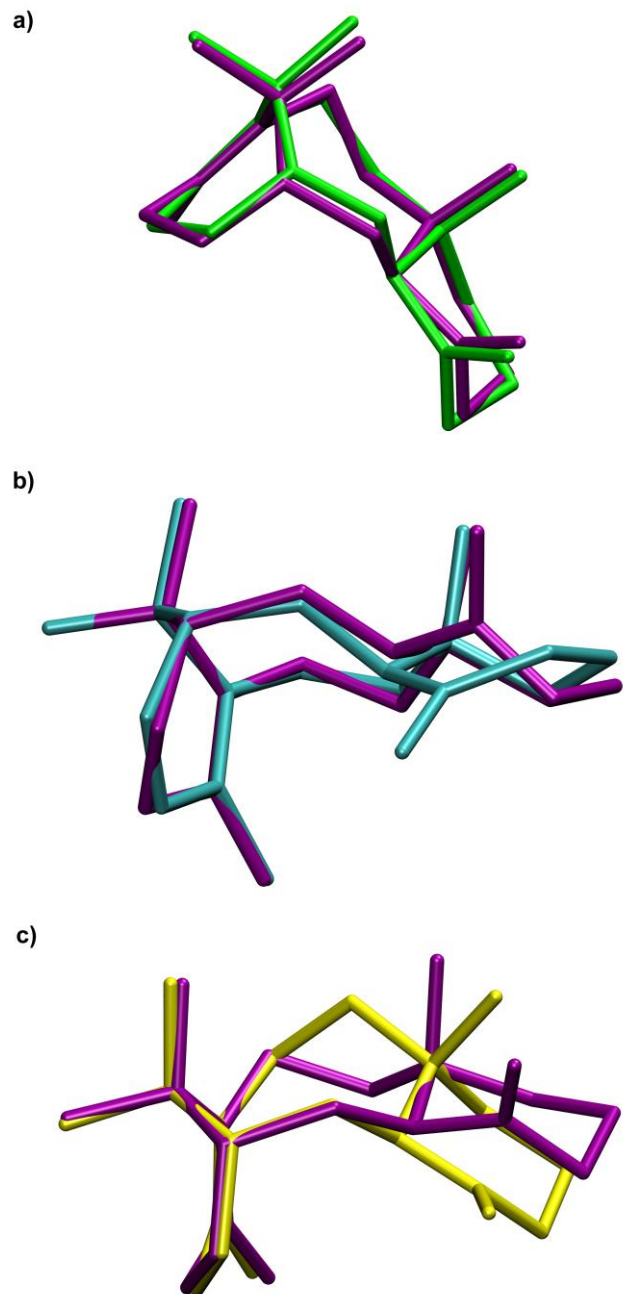
MD	RMSD (Å)
MD1	1.5 (0.4)
MD2	0.7 (0.2)
MD3	0.8 (0.2)

[a] Standard deviations are given in parentheses. MD1-MD3 correspond to MD simulations with different initial velocity distributions. The higher occurrence of **D1$_{boat}$** during MD1 contributes to the higher RMSD value obtained for this simulation (see Table 1 of the main text).

Figure S8. Superimposed structures of the **TXS·D1** complex at intervals of 10 ps during MD1. MD2 and MD3 show similar superimposition profiles.
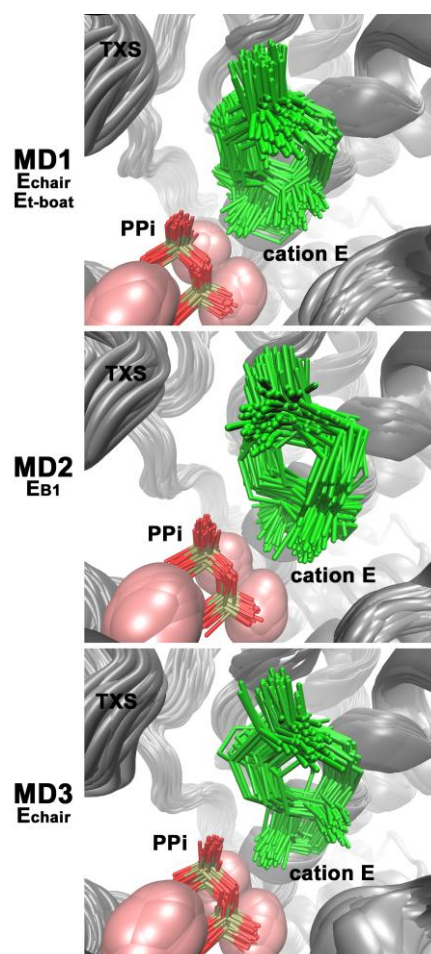
6. Cation E

a)

b)

c)

Figure S9. E_{chair} (purple) superimposed with: (a) the HT-QM structure of cation **E** (green) [1], (b) $E_{t\text{-}boat}$ (cyan), and (c) E_{B1} (yellow).

Table S9

Average RMSD (with respect to the initial structure) for all carbon atoms of cation **E** in the MD simulations of the **TXS·E** complex[a]

MD	RMSD (Å)
MD1	0.9 (0.3)
MD2	1.6 (0.2)
MD3	1.0 (0.2)

[a] Standard deviations are given in parentheses. MD1-MD3 correspond to MD simulations with different initial velocity distributions. The formation of E_{B1} during MD2 contributes to the higher RMSD value obtained for this simulation (see Figure S9c).

Figure S10. Superimposed structures of the **TXS·E** complex at intervals of 10 ps during MD1-MD3. The conformers of cation **E** identified in each simulation are indicated on the left.

13

7. References

1. Hong YJ, Tantillo DJ (2011) The Taxadiene-Forming Carbocation Cascade. J Am Chem Soc 133:18249–18256.