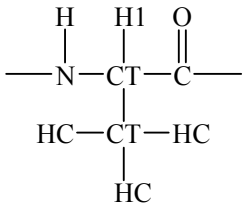
Supporting Information

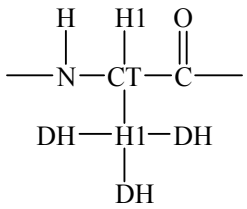
Free Energy Perturbation (FEP) Simulation on Transition States and Redesign of Butyrylcholinesterase

Wenchao Yang,[1,2,a] Yongmei Pan,[2,a] Fang Zheng,[2] Hoon Cho,[2] Hsin-Hsiung Tai,[2] and Chang-Guo Zhan[2,*]

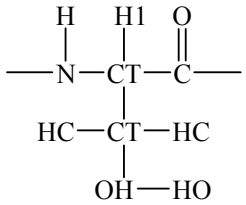
[1]Key Laboratory of Pesticide & Chemical Biology of Ministry of Education, College of Chemistry, Central China Normal University, Wuhan 430079, P. R. China

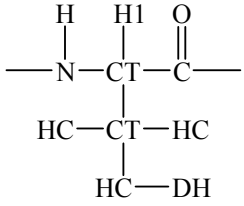
[2]Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 725 Rose Street, Lexington, KY 40536, USA



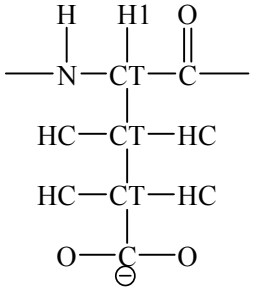


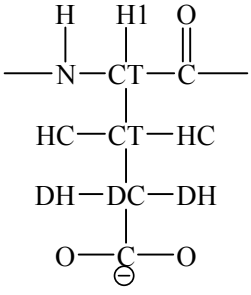
(A) Ala to Gly Before perturbation: Ala After perturbation: Gly





(B) Ser to Ala Before perturbation: Ser After perturbation: Ala

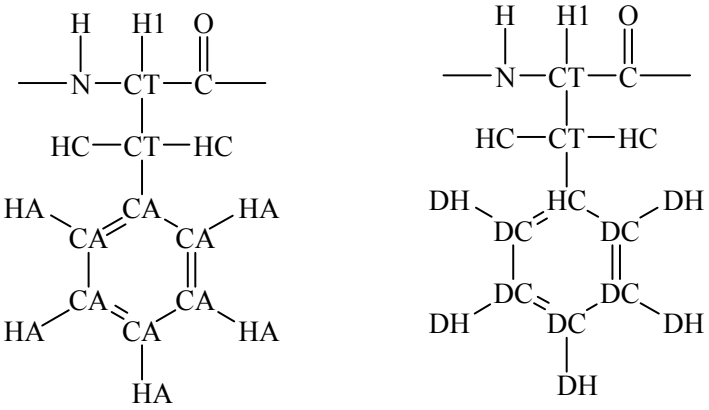




[a] These authors contributed equally to this work.

* Corresponding author. E-mail: zhan@uky.edu

(C) Glu to Asp Before perturbation: Glu After perturbation: Asp



(D) Phe to Ala Before perturbation: Phe After perturbation: Ala

Figure S1. Structures and atom types before and after perturbations associated with (A) Ala to Gly (B) Ser to Ala (C) Glu to Asp (D) Phe to Ala.

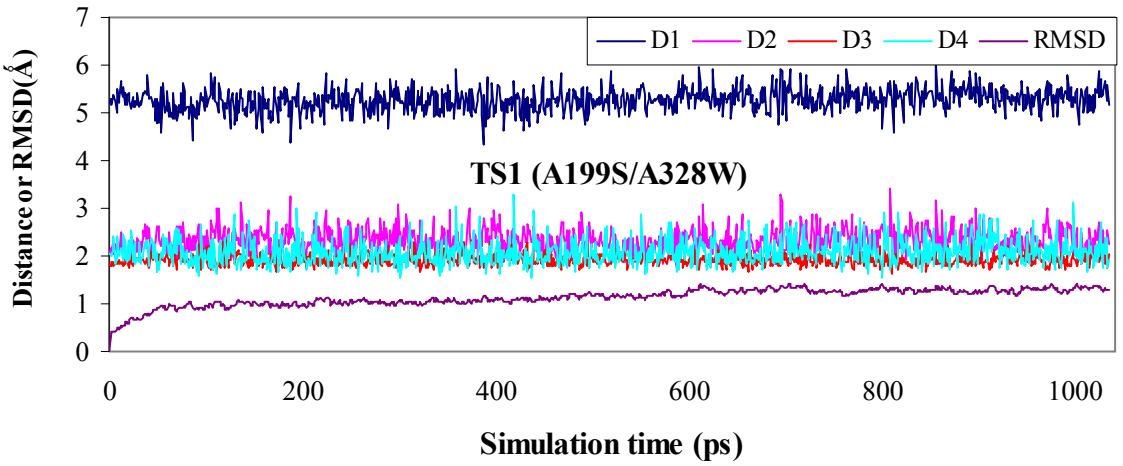


Figure S2. Plots of the key internuclear distances vs the time in the MD-simulated TS1 structure for (-)-cocaine hydrolysis catalyzed by A199S/A328W BChE. Traces D1, D2, and D3 refer to the distances between the carbonyl oxygen of (-)-cocaine and the hydrogen of G116, G117, and S199, respectively. Trace D4 is the internuclear distance between the carbonyl oxygen of (-)-cocaine benzoyl ester and the hydroxyl hydrogen of the S199 side chain. RMSD represents the root-mean-square deviation of the simulated positions of the protein backbone atoms from those in the initial structure.

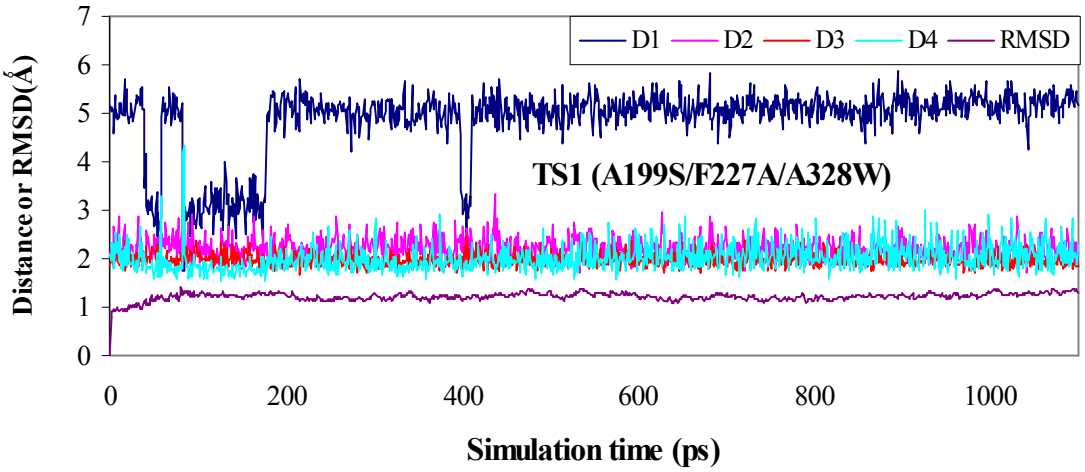


Figure S3. Plots of the key internuclear distances vs the time in the MD-simulated TS1 structure for (-)-cocaine hydrolysis catalyzed by A199S/F227A/A328W BChE. Traces D1, D2, and D3 refer to the distances between the carbonyl oxygen of (-)-cocaine and the hydrogen of G116, G117, and S199, respectively. Trace D4 is the internuclear distance between the carbonyl oxygen of (-)-cocaine benzoyl ester and the hydroxyl hydrogen of the S199 side chain. RMSD represents the root-mean-square deviation of the simulated positions of the protein backbone atoms from those in the initial structure.

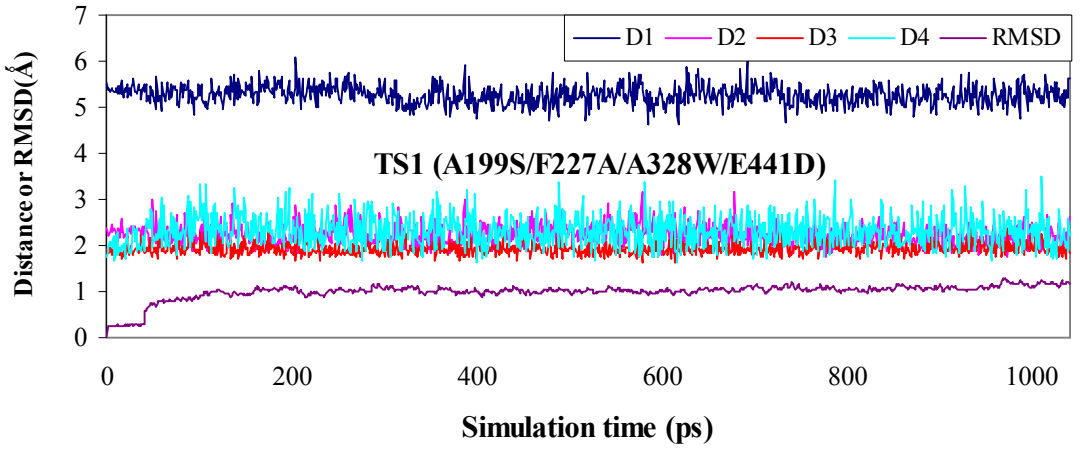


Figure S4. Plots of the key internuclear distances vs the time in the MD-simulated TS1 structure for (-)-cocaine hydrolysis catalyzed by A199S/F227A/A328W/E441D BChE. Traces D1, D2, and D3 refer to the distances between the carbonyl oxygen of (-)-cocaine and the hydrogen of G116, G117, and S199, respectively. Trace D4 is the internuclear distance between the carbonyl oxygen of (-)-cocaine benzoyl ester and the hydroxyl hydrogen of the S199 side chain. RMSD represents the root-mean-square deviation of the simulated positions of the protein backbone atoms from those in the initial structure.

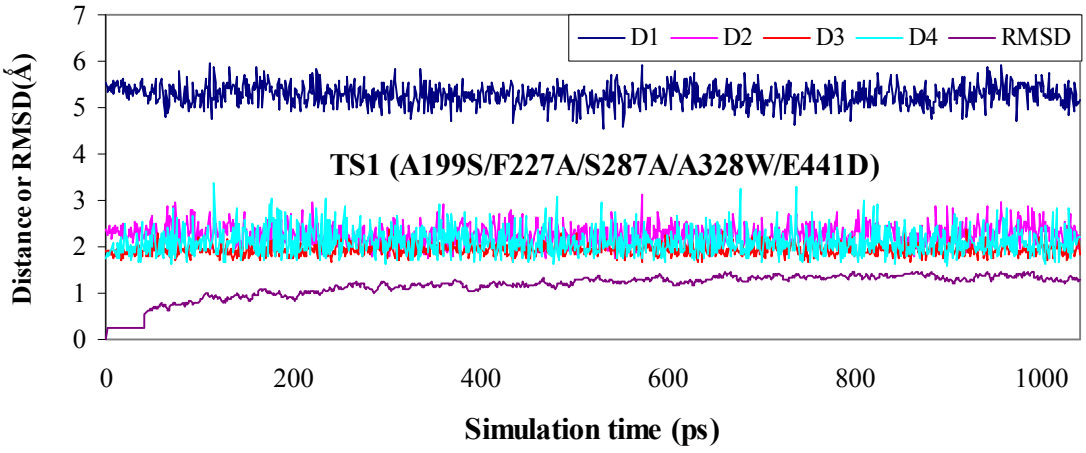


Figure S5. Plots of the key internuclear distances vs the time in the MD-simulated TS1 structure for (-)-cocaine hydrolysis catalyzed by A199S/F227A/S287A/A328W/E441D BChE. Traces D1, D2, and D3 refer to the distances between the carbonyl oxygen of (-)-cocaine and the hydrogen of G116, G117, and S199, respectively. Trace D4 is the internuclear distance between the carbonyl oxygen of (-)-cocaine benzoyl ester and the hydroxyl hydrogen of the S199 side chain. RMSD represents the root-mean-square deviation of the simulated positions of the protein backbone atoms from those in the initial structure.

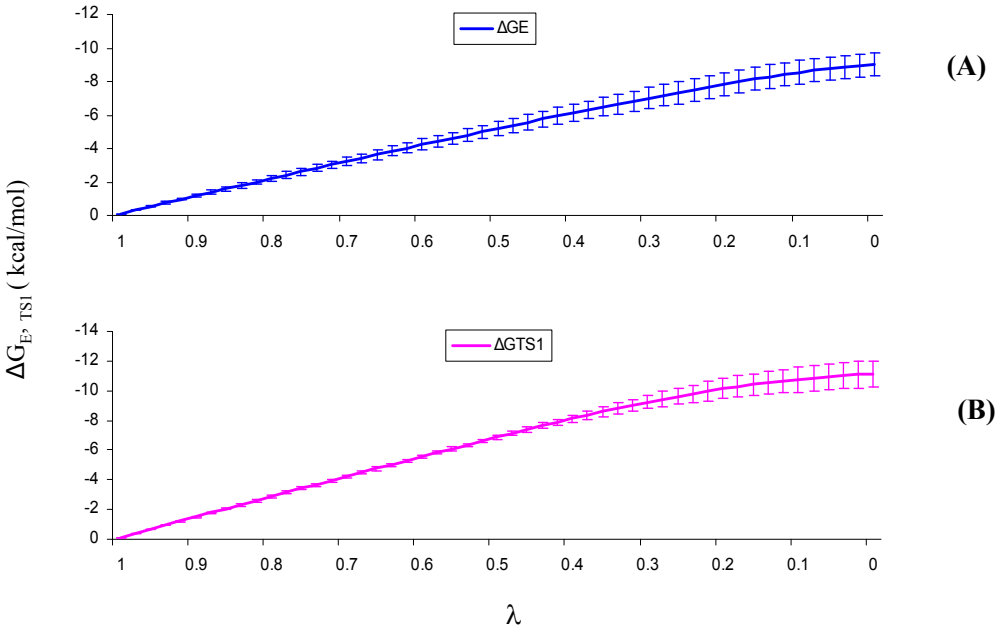
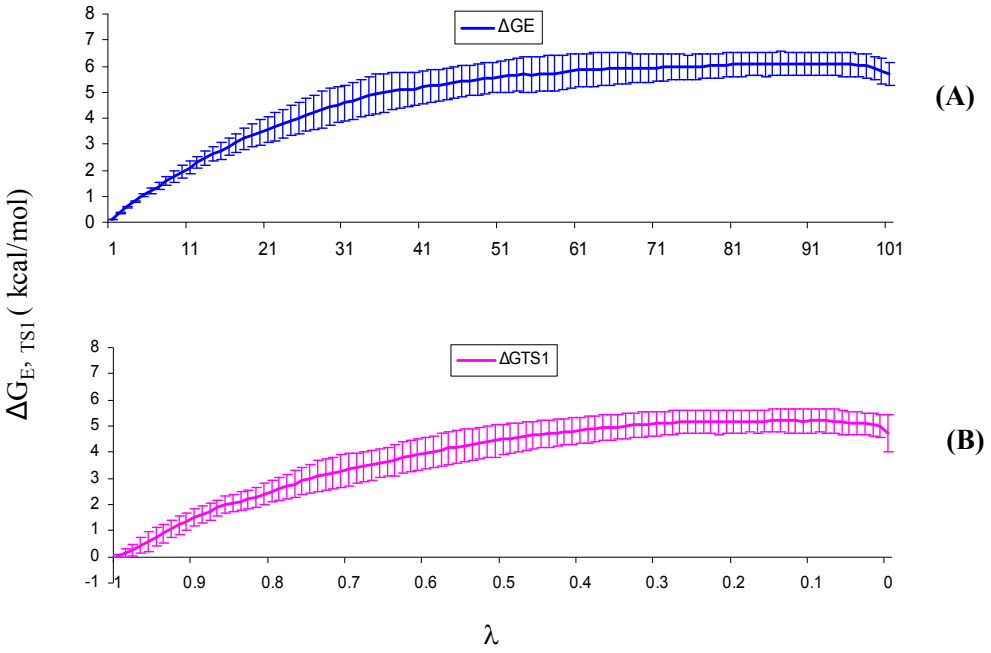


Figure S6. Incremental changes of ΔG_E **(A)** and ΔG_{TS1}**(B)** as function of λ for mutation from the free enzyme **(A)** or the TS1 structure **(B)** of A328W to A328W/A199S. The error bars indicate the standard deviation of each partial sum of the energy changes.



Gigure S7. Incremental changes of ΔG_E (A) and ΔG_{TS1}(B) as function of λ for mutation from the free enzyme (A) or the TS1 structure (B) of A328W/A199S to A328W/A199S/F227A. The error bars indicate the standard deviation of each partial sum of the energy changes.

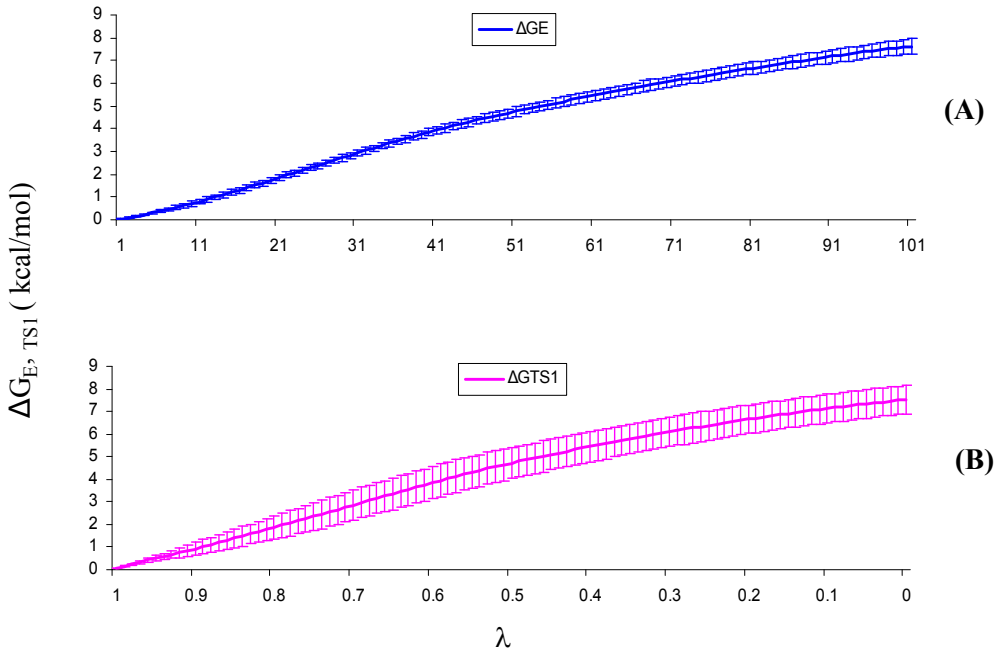


Figure S8. Incremental changes of ΔG_E (A) and ΔG_{TS1}(B) as function of λ for mutation from the free enzyme (A) or the TS1 structure (B) of A328W/A199S/F227A to A328W/A199S/F227/E441D. The error bars indicate the standard deviation of each partial sum of the energy changes.

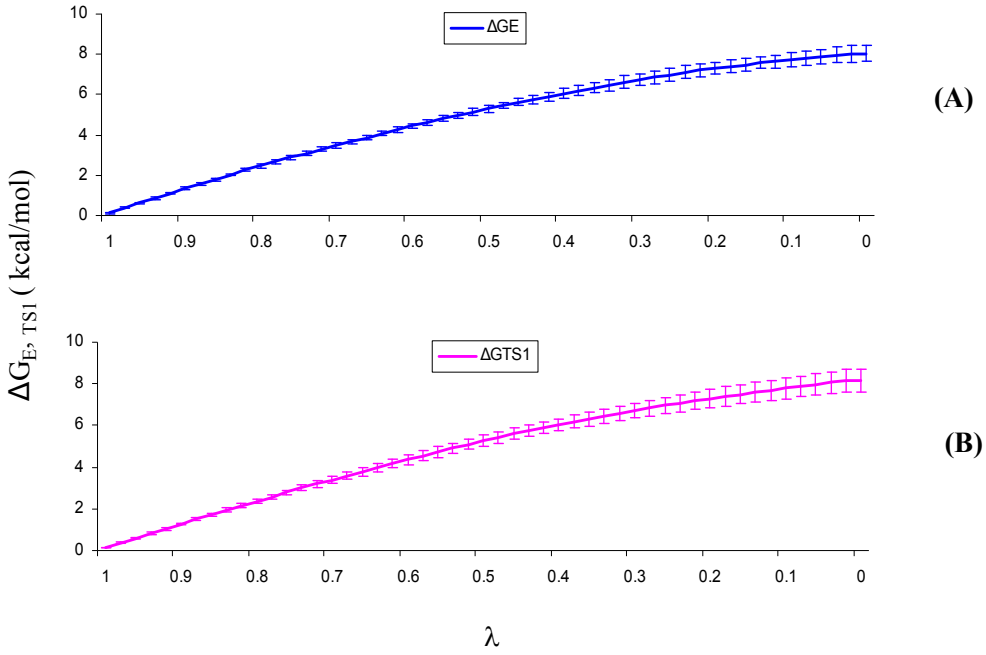


Figure S9. Incremental changes of ΔG_E (A) and ΔG_{TS1}(B) as function of λ for mutation from the free enzyme (A) or the TS1 structure (B) of A328W/A199S/F227A/E441D to A328W/A199S/F227A/S287A. The error bars indicate the standard deviation of each partial sum of the energy changes.

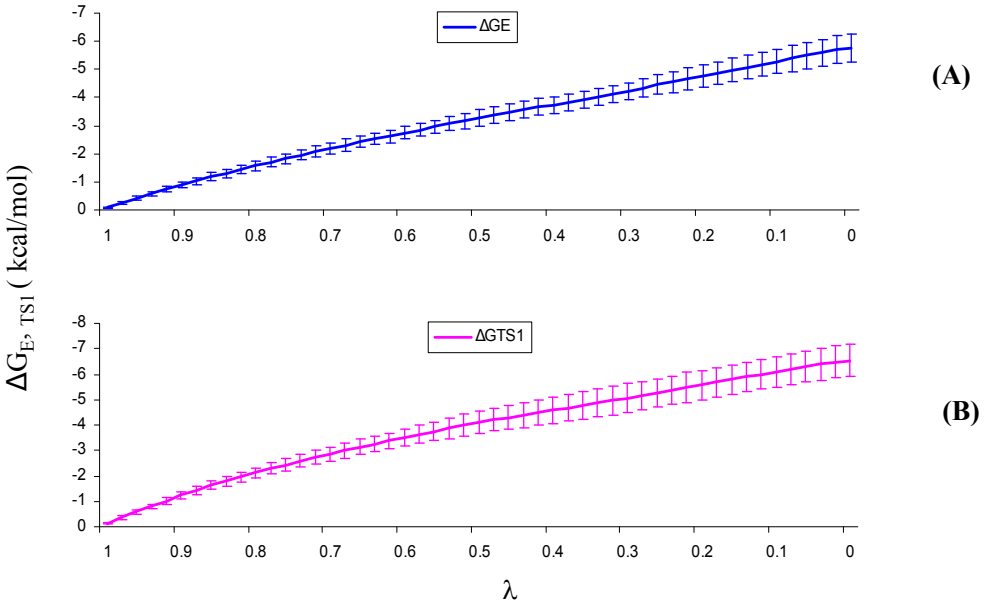


Figure S10. Incremental changes of ΔG_E (A) and ΔG_{TS1} (B) as function of λ for mutation from the free enzyme (A) or the TS1 structure (B) of A328W/A199S/F227A/E441D/S287A to A328W/A199S/F227A/E441D/S287G. The error bars indicate the standard deviation of each partial sum of the energy changes.

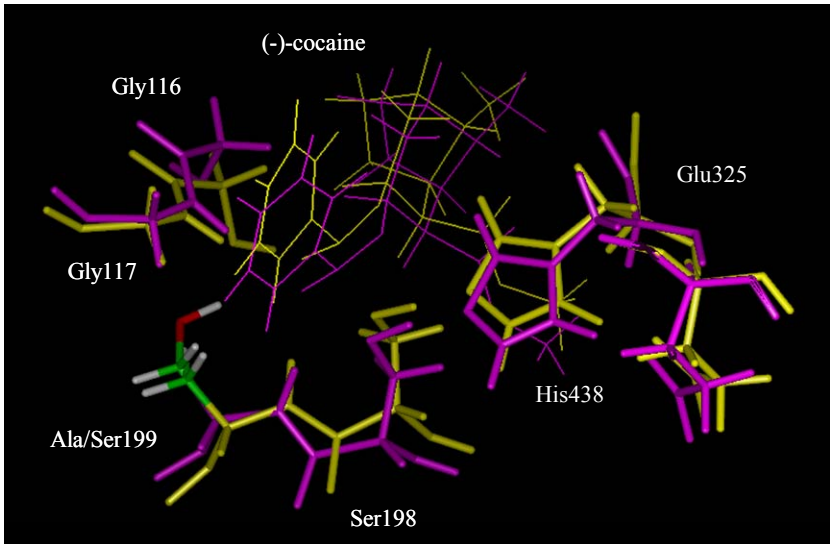


Figure S11. Superimposing of the TS1 structure corresponding to the A328W/A199S mutant (before the perturbation, magenta) with that corresponding to the mutant A328W (after the perturbation, yellow). The colors of the side chain atoms in residue Ser199 (before the perturbation) or Ala199 (after the perturbation) are shown according to their atom types

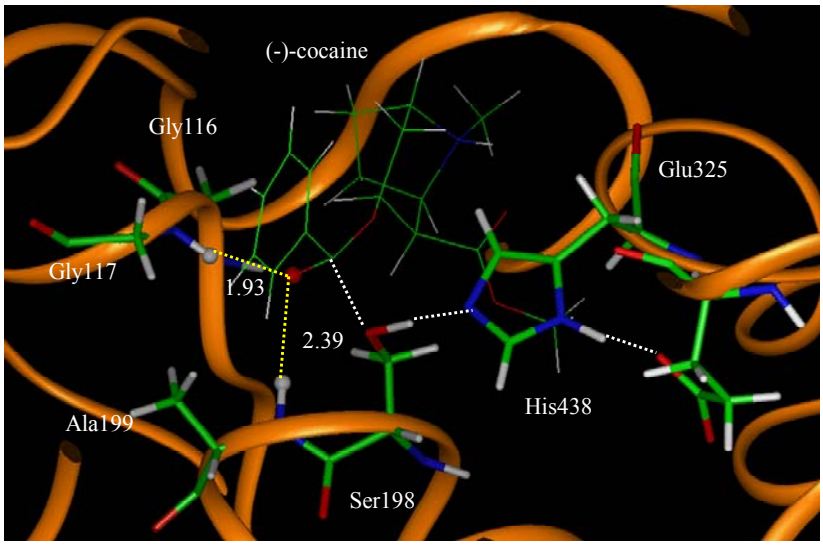


Figure S12. TS1 structure with the A328W mutant after the perturbation. The colors of key atoms are shown according to their atom types. The orange ribbon indicates the backbone of the molecule. The dashed lines in yellow indicate the hydrogen bonds between (-)-cocaine and the oxyanion hole of the mutant. The simulated average distances are given for the hydrogen bonds. The dashed lines in the white refer to the transition bonds. The atoms highlighted as balls include several key H atoms (grey balls) from the oxyanion hole of the protein and the O atom (red ball) of the (-)-cocaine carbonyl group.

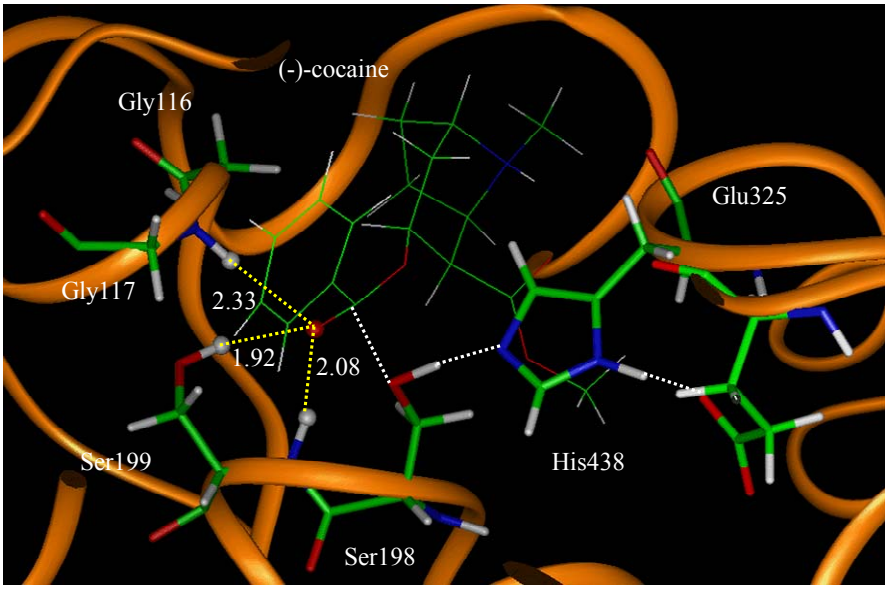


Figure S13. TS1 structure with the A199S/A328W mutant before the perturbation. The colors of key atoms are shown according to their atom types. The orange ribbon indicates the backbone of the molecule. The dashed lines in yellow indicate the hydrogen bonds between (-)-cocaine and the oxyanion hole of the mutant. The simulated average distances are given for the hydrogen bonds. The dashed lines in the white refer to the transition bonds. The atoms highlighted as balls include several key H atoms (grey balls) from the oxyanion hole of the protein and the O atom (red ball) of the (-)-cocaine carbonyl group.

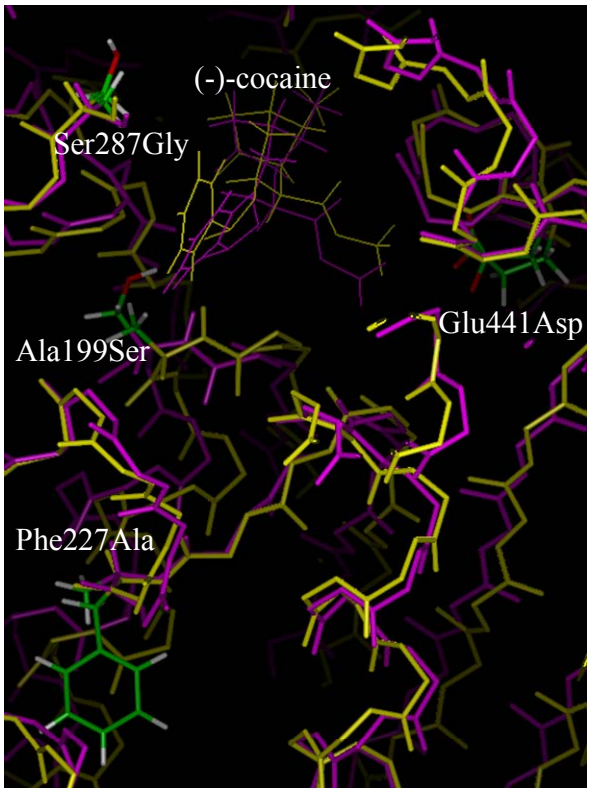


Figure S14. Superimposing of the two TS1 structures associated with the A328W mutant (yellow) and the A328W/A199S/F227A/E441D/S287G mutant (pink). The stick structures indicate the backbone of the TS1 structures. The four mutated residues are colored. This figure reveals that the active site cavity of the A328W/A199S/F227A/E441D/S287G mutant is slightly (not dramatically) larger than that of the A328W mutant.

Details of parameter development

The lengths of the transition bonds and the partial charges of the cocaine atoms in the TS1 structure were defined and calculated based on the geometries obtained from previous *ab initio* reaction coordinate calculations on the cocaine hydrolysis catalyzed by wild-type BChE.[20] The atoms specific to the TS1 structure are given new atom types and the involving bond and angle parameters are defined in a separate file. In detail, for the atoms involving the transition bonds, for example the carbonyl carbon of the benzoyl ester of cocaine, which has partially formed bond with the hydroxyl oxygen of Ser198 in the TS1 structure, a new atom type is defined and the corresponding parameters are set according to the geometry calculated from *ab initio* calculation. Another two new atom types are defined for the hydroxyl hydrogen of the attacking Ser198 and the delta hydrogen of His438. Other parameters concerning the substrate and the protein are those defined by Amber.

Parameter file of TS1 structure

```
Parameter file specific to transition state TS1
MASS
cx 12.01         0.616  !          carbonyl carbon of benzoyl ester of
cocaine
GN 1.008         0.161              Hydrogen attached to delta N in
His438
GO 1.008         0.135              Hydroxyl group H in Ser198

BOND
cx-os  401.8    1.350       COG in TS1
cx-o   540.4    1.260       COG in TS1
cx-OH  468.1    1.880       COG in TS1
cx-ca  346.5    1.490       COG in TS1
GN-NA  434.0    1.110       HD1 in HIS of TS1
GN-O2  434.0    1.430       Added bond for TS1
GO-OH  553.0    1.120       HG in SER of TS1
GO-NB  553.0    1.420       added bond

ANGLE
os-c -ca   67.662     110.765   Calculated with empirical approach
cx-os-c3    63.0      116.91    NEW_MP2/6-31G COG
o -cx-os    75.4      123.89    MP2/6-31G* COG
os-cx-ca   67.662     110.765   same as os-c -ca
ca-cx-o     67.1      126.40    NEW_MP2/6-31G COG
cx-ca-ca    62.0      122.13    NEW_MP2/6-31G COG
ct-OH-cx    50.5      109.50    new type for COG and ser197
CT-OH-cx    50.5      109.50    new type for COG and ser197
OH-cx-os    60.5      110.02    new type reduced force
OH-cx-o     60.5      110.02    only bond length important
ca-cx-OH    50.9      120.00
cx-OH-GO    50.9      109.06    above are for COG
cx-OH-HO    50.9      109.06    above are for COG
CC-NA-GN    50.0      125.00    GN is H in his in TS1
CR-NA-GN    50.0      125.00    GN is H, his in TS1
CT-OH-GO    55.0      108.50    GO is HG in SER for TS1
C -O2-GN    50.1      160.00
O2-GN-NA    50.1      170.00
OH-GO-NB    50.1      160.00
GO-NB-CR    50.1      125.00
GO-NB-CV    50.1      125.00

DIHE
X -cx-os-X    2    5.400       180.000          2.000
X -cx-ca-X    4   14.500       180.000          2.000
intrpol.bsd.on C6H6
X -cx-OH-X    2    0.000         0.000          3.000     3-fold, no
force
CT-C -O2-GN   2    0.000       180.000          2.000
C -O2-GN-NA   2    0.000       180.000          2.000
O2-C -O2-GN   2    0.000       180.000          2.000
O2-GN-NA-CC   2    0.000       180.000          2.000
O2-GN-NA-CR   2    0.000       180.000          2.000
CT-OH-GO-NB   2    0.000       180.000          2.000
OH-GO-NB-CR   2    0.000       180.000          2.000
```

```
OH-GO-NB-CV   2    0.000       180.000          2.000
cx-OH-GO-NB   2    0.000       180.000          2.000

IMPROPER
c3-o -c -os         1.1          180.0         2.0         Using default
value
ca-o -c -os         1.1          180.0         2.0         Using default
value
c -ca-ca-ca         1.1          180.0         2.0         Using default
value
ca-ca-ca-ha         1.1          180.0         2.0         Using default
value
c3-o -cx-os         1.1          180.0         2.0         Using default
value
ca-o -cx-os         1.1          180.0         2.0         Using default
value
cx-ca-ca-ca         1.1          180.0         2.0         Using default
value

NONBON
cx          1.9080  0.0860             OPLS (c type)
GN          0.6000  0.0157             Hydrogen on N
GO          0.0000  0.0000             Hydroxyl H
```

Partial charges on cocaine atoms in TS1 structure

Atom	atom type	charge
C1	c3	-0.082
H2	h1	0.122
H3	h1	0.122
H4	h1	0.122
O5	os	-0.335
C6	c	0.665
O7	o	-0.573
C8	c3	-0.021
H9	hc	0.073
C10	c3	-0.018
N20	n4	-0.020
C22	c3	-0.263
H23	hx	0.146
H24	hx	0.146
H25	hx	0.146
H21	hn	0.394
H11	hx	0.104
C12	c3	-0.067
H13	hc	0.069
H14	hc	0.076
C15	c3	-0.101
H16	hc	0.077
H17	hc	0.078
C18	c3	0.023
H19	hx	0.106
C26	c3	-0.144
H27	hc	0.079
H28	hc	0.102
C29	c3	0.111
H30	h1	0.152

```
O31    os                  -0.349
C32    cx                  0.580
O33    o                  -0.625
C34    ca                  -0.036
C35    ca                  -0.098
H36    ha                  0.154
C37    ca                  -0.147
H38    ha                  0.154
C39    ca                  -0.093
H40    ha                  0.157
C41    ca                  -0.150
H42    ha                  0.154
C43    ca                  -0.119
H44    ha                  0.129
```

Force field parameters for perturbed residues

Perturbed Gly from Ala

```
MASS
DH  0.00          0.0  !

BOND
H1-DH   340.0   1.09

ANGLE
DH-H1-DH   35.0       109.50
DH-H1-CT   50         109.50

DIHE
H1-CT-C-N  1    0.0     0.       4.
H1-CT-C-N  1    0.0     0.       2.
C-N-CT-H1  1    0.0     0.       1.
C-N-CT-H1  1    0.0     0.       3.
C-N-CT-H1  1    0.0     0.       4.
C-N-CT-H1  1    0.0     0.       2.
C-N-CT-CT  1    0.0     0.       2.
C-N-CT-CT  1    0.0     0.       1.
C-N-CT-CT  1    0.0     0.       3.
C-N-CT-CT  1    0.0     0.       4.
DH-H1-CT-N  1    0.0     0.       3.
DH-H1-CT-H1  1    0.0     0.       3.
O-C-CT-CT  1    0.0     0.       1.
O-C-CT-CT  1    0.0     0.       2.
O-C-CT-H1  1    0.0     0.       2.
O-C-CT-H1  1    0.0     0.       1.
O-C-CT-H1  1    0.0     0.       3
O-C-CT-CT  1    0.0     0.       3.
C-CT-H1-DH  1    0.0     0.       3.

IMPROPER
H1-DH-DH-DH         0.0            0.          2
CT-DH-HC-DH         0.0            0.0         2.

NONBON
DH          0.00  0.0
```

Perturbed Ala from Ser

```
MASS
DH  0.00          0.0  !

BOND
HC-DH   553.0   0.96

ANGLE
CT-HC-DH   55.0       108.50

DIHE
DH-HC-CT-HC   1    0.00          0.0             3.
CT-CT-HC-DH   1    0.00          0.0             3.0
CT-CT-HC-DH   1    0.00          0.0             2.0
CT-CT-HC-DH   1    0.0           0.              1
H1-CT-C-N     1    0.0           0.              4
CT-CT-OH-HO   1    0.0           0.              1
CT-CT-OH-HO   1    0.0           0.              2
CT-CT-OH-HO   1    0.0           0.              3
H1-CT-OH-HO   1    0.0           0.              2
H1-CT-OH-HO   1    0.0           0.              3
H1-CT-OH-HO   1    0.0           0.              2
H1-CT-CT-HC   1    0.0           0.              1
H1-CT-CT-HC   1    0.0           0.              3
H1-CT-CT-HC   1    0.0           0.              3
H1-CT-CT-OH   1    0.0           0.              3
H1-CT-CT-OH   1    0.0           0.              1
H1-CT-CT-H1   1    0.0           0.              1
H1-CT-CT-H1   1    0.0           0.              3
H1-CT-C-N     1    0.0           0.              2
HC-CT-HC-DH   1    0.0           0.              2.
C-CT-H1-DH    1    0.0           0.              3.
C-N-CT-H1     1    0.0           0.              1
C-N-CT-H1     1    0.0           0.              4
C-N-CT-H1     1    0.0           0.              3
C-N-CT-H1     1    0.0           0.              2
C-N-CT-CT     1    0.0           0.              1
C-N-CT-CT     1    0.0           0.              2
C-N-CT-CT     1    0.0           0.              3
C-N-CT-CT     1    0.0           0.              4
O-C-CT-CT     1    0.0           0.              1
O-C-CT-CT     1    0.0           0.              3
O-C-CT-CT     1    0.0           0.              2
O-C-CT-H1     1    0.0           0.              2
O-C-CT-H1     1    0.0           0.              1
O-C-CT-H1     1    0.0           0.              3
H1-CT-OH-HO   1    0.0           0.              1.
HC-CT-HC-DH   1    0.0           0.              1.
HC-CT-HC-DH   1    0.0           0.              2
HC-CT-HC-DH   1    0.0           0.              3.

IMPROPER
NA-CA-CN-CB         0.0            0.          2.

NONBON
DH          0.00  0.0
```

Perturbed Asp from Glu

```
MASS
DC  0.00          0.0  !            gamma C of Glu
DH  0.00          0.0

BOND
DC-DH   340.0    1.090
CT-DC   317.0    0.76
DC-C    310.0    0.76

ANGLE
CT-DC-C    63.0      180.00
DC-CT-CT   40.0      180.0
DC-C-O2    70.0      117.00
DH-DC-DH   35.0      109.50
DH-DC-C    50.0      60.0
DH-DC-CT    50.0      60.0
HC-CT-DC    50.0      109.50

DIHE
DH-DC-C-O2  1    0.0     0.       2.
DH-DC-C-O2  1    0.0     0.       3.
DC-CT-CT-C  1    0.0     0.       2.
DC-CT-CT-C  1    0.0     0.       3.
HC-CT-DC-DH  1    0.0     0.       2.
HC-CT-DC-DH  1    0.0     0.       3.
HC-CT-DC-C  1    0.0     0.       2.
HC-CT-DC-C  1    0.0     0.       3.
CT-DC-C-O2  1    0.0     0.       2.
CT-DC-C-O2  1    0.0     0.       3.
CT-CT-DC-DH  1    0.0     0.       2.
CT-CT-DC-DH  1    0.0     0.       3.
CT-CT-DC-C  1    0.0     0.       2.
CT-CT-DC-C  1    0.0     0.       3.
HC-CT-CT-HC  1    0.0     0.       2.
HC-CT-CT-HC  1    0.0     0.       3.
HC-CT-CT-C  1    0.0     0.       2.
HC-CT-CT-C  1    0.0     0.       3.
CT-CT-C-O2  1    0.0     0.       3.
CT-CT-C-O2  1    0.0     0.       2.
CT-CT-CT-HC  1    0.0     0.       2.
CT-CT-CT-HC  1    0.0     0.       3.
CT-CT-CT-C  1    0.0     0.       2.
CT-CT-CT-C  1    0.0     0.       3.
HC-CT-C-O2  1    0.0     0.       3.
CT-CT-CT-C  1    0.0     0.       2.
CT-CT-CT-C  1    0.0     0.       3.

IMPROPER
CT-DH-DC-C      0.0     0.       2.
CT-DH-DC-C      0.0     0.       3.
DH-DH-DC-C      0.0     0.       2.
CT-HC-CT-DC      0.0     0.       2.
CT-HC-CT-DC      0.0     0.       3.
HC-HC-CT-DC      0.0     0.       2.
CT-DH-DC-DH      0.0     0.       2.
CT-HC-CT-HC      0.0     0.       2.
```

```
CT-HC-CT-C      0.0     0.       2.
CT-HC-CT-C      0.0     0.       3.
CT-HC-CT-CT      0.0     0.       2.
CT-HC-CT-CT      0.0     0.       3.
HC-HC-CT-CT      0.0     0.       2.
HC-HC-CT-C      0.0     0.       2.

NONBON
DC          0.00  0.0
DH          0.00  0.0
```

Perturbed Ala from Phe

```
MASS
DH  0.00          0.0  !
DC  0.00          0.0                aromatic carbon of Phe

BOND
DC-DH   367.0   1.08
HC-DC   469.0   1.40
DC-DC   469.0   1.40

ANGLE
DC-DC-DH   50.0       120.00
HC-DC-DC   63.0       120.00
HC-DC-DH   50.0       120.0
CT-HC-DC   70.0       120.0
DC-DC-DC   63.0       120.0
DC-HC-DC   63.0       120.0

DIHE
DH-HC-CT-HC   1    0.0          0.0             3.
CB-HC-CT-CT   1    0.0          0.0             1
CT-CT-HC-DH   1    0.0          0.0             3.0
CT-CT-HC-DH   1    0.0          0.0             2.0
CT-CT-HC-DH   1    0.0          0.0             1
DC-HC-CT-HC   1    0.0          0.0             1
HC-CT-HC-DH   1    0.0          0.0             2.
CT-CT-HC-DC   6    0.0          0.0             2
CT-CT-HC-DC   6    0.0          0.0             1
CT-CT-HC-DC   6    0.0          0.0             3
HC-CT-HC-DC   6    0.0          0.0             2
DC-HC-CT-HC   6    0.0          0.0             3
HC-DC-DC-DH   4   14.5         180.0             2
CT-DC-HC-DC   4   14.5         180.0             2

IMPROPER
CT-DH-HC-DH         0.0            0.0         2.
HC-DH-DC-DC         0.0            0.0         2.
DC-DH-DC-DC         0.0            0.0         2.
DC-DH-DC-DC         0.0            0.0         2.
HC-DC-DC-DH         0.0            0.0         2.
CT-DC-HC-DC         0.0            0.0         2.

NONBON
DH          0.00  0.0
DC          0.00  0.0
```