Supporting Information for

EnzyHTP Computational Directed Evolution with Adaptive Resource Allocation

Qianzhen Shao[1], Yaoyukun Jiang[1] and Zhongyue J. Yang[1-5,*]

[1]*Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37235, United States*
[2]*Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States*
[3]*Vanderbilt Institute of Chemical Biology, Vanderbilt University, Nashville, Tennessee 37235, United States*
[4]*Data Science Institute, Vanderbilt University, Nashville, Tennessee 37235, United State*
[5]*Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, Tennessee 37235, United States*
**Corresponding author's email: zhongyue.yang@vanderbilt.edu*

Contents

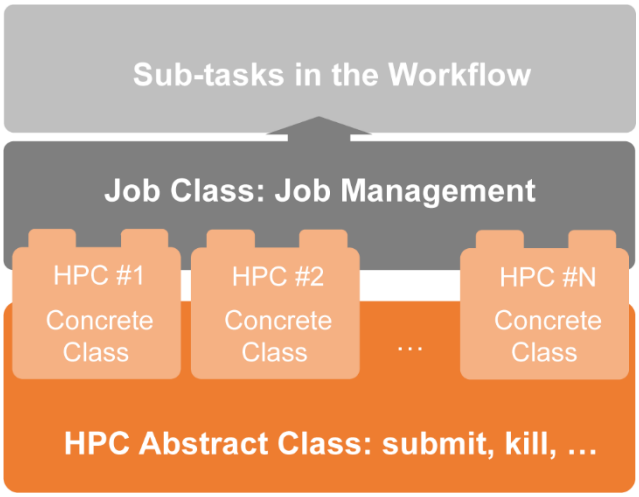

Figure S1. Architecture of the ARMer Python library

Text S1. Details about the architecture of the ARMer Python library

The Job class: Defined in the Job class, a job object is instantiated using the constructor *config_job()*, where the arguments *commands*, *cluster*, *env_settings*, and *res_keywords* are provided by the user (Supporting Information, Figure S2). Specifically, *commands* refers to the target shell commands for running external software for a specific enzyme modeling sub-task (e.g., *commands* = "g16 < filename.gjf > filename.out"); *cluster* refers to an HPC class object that contains miscellaneous details about user's local HPC (e.g., *cluster* = Accre(), in which ACCRE refers to our local HPC at Vanderbilt University. Note that it is easy to support other HPCs by defining other HPC classes base on the "plug-in" interface); *env_settings* states environment settings of external software (e.g., *env_settings* = ‘’’module load Gaussian/16.B.01’’’). Even for the same external software (e.g., Gaussian 16[1]), the environment settings likely differ in different HPC clusters. Finally, *res_keywords* configures computing resources for the job (e.g., *res_keywords* = ‘core_type’ : ‘cpu’, ‘nodes’:’1’, ‘node_cores’ : ‘8’, ‘job_name’ : ‘EnzyHTP_QMCluster’, ‘partition’ : ‘production’, ‘mem_per_core’ : ‘3G’, ‘walltime’ : ‘24:00:00’), in which the parameters depend on the computational task (e.g., 8-24 CPUs for QM calculation, 1 GPU for MD calculation, etc.).

The HPC classes: The HPC class files are stored in a folder named "cluster". In the folder, *_interface.py* defines an abstract HPC class as the code interface and *accre.py* defines an example concrete HPC class we made for our local HPC at Vanderbilt. Users can create new files under this folder defining new concrete HPC classes to easily modify ARMer Python library to be compatible with their local HPC cluster. The instances of the HPC class are used as input for generating the Job instance. The methods of the HPC class are used by the Job instance through the HPC instance to interface with the corresponding local HPC cluster. The new HPC class user defines is required to fulfill the code interfaced defined by the abstract HPC class in *_interface.py* to make sure they are compatible with the Job class. It is enforced by requiring (by the Job class) all HPC classes to inherit the abstract HPC class so that the new HPC class has to define some required methods (otherwise python will raise an error). Here is an example explaining how a new HPC is supported by a new HPC class:
To execute the *submit()* method in the Job class for job submission, only the *parser_resource_str()* method and the *submit_job()* method from *cluster* (a concrete HPC class instance) are called. Both *parser_resource_str()* and *submit_job()* are required by the code interface in the abstract HPC class to "parser general resource keywords to an HPC specified paragraph (Figure S3, upper part)" and "run shell commands to submit the job submission script in the submission directory (e.g.: sbatch xxx.sh)", respectively. A user can define a new HPC class inheriting the abstract HPC class and defines the two methods mentioned above easily with several lines of code (e.g.: running the corresponding shell command of the job scheduler on the HPC, similar to the corresponding method in *accre.py*) to have the *submit()* method in Job class support a new HPC. It is the same for all the other methods in the Job class, as long as the new HPC class fulfills the code interface, which is enforced, the method in the Job class will function normally on the corresponding new HPC.

Notably, ARMer achieves job management through interacting with the job scheduling system of HPCs. As such, ARMer cannot function properly with the HPCs where there is no implementation of the job scheduling system. Neither does it work with the job scheduling system that does not fully support the task of 1) submitting a job, 2) killing/holding/releasing a job with the job ID, or 3) obtaining job status (for job monitoring) with the job ID.

```
from armer import ClusterJob
import clusters

cluster = clusters.accre.Accre()

# case 1 single submission
job = ClusterJob.config_job(
    commands = "g16 < xxx.gjf > xxx.out",
    cluster = cluster,
    env_settings = cluster.G16_ENV['CPU'],
    res_keywords = {'core_type' : 'cpu',
                    'nodes':'1',
                    'node_cores' : '8',
                    'job_name' : 'TEST',
                    'partition' : 'production',
                    'mem_per_core' : '3G',
                    'walltime' : '24:00:00',
                    'account' : 'xxx'}
)
job.submit(sub_dir="./")
job.wait_to_end(period=10)

# case 2 array submission
jobs = []
for i in range(10):
    jobs.append(ClusterJob.config_job(
        commands = f"g16 < xxx_{i}.gjf > xxx_{i}.out",
        cluster = cluster,
        env_settings = cluster.G16_ENV['CPU'],
        res_keywords = {'core_type' : 'cpu',
                        'nodes':'1',
                        'node_cores' : '8',
                        'job_name' : 'TEST',
                        'partition' : 'production',
                        'mem_per_core' : '3G',
                        'walltime' : '24:00:00',
                        'account' : 'xxx'},
        sub_dir="./"                )
                )
ClusterJob.wait_to_array_end(jobs, period=30, array_size=5)
```

Figure S2. Example code of the ARMer API

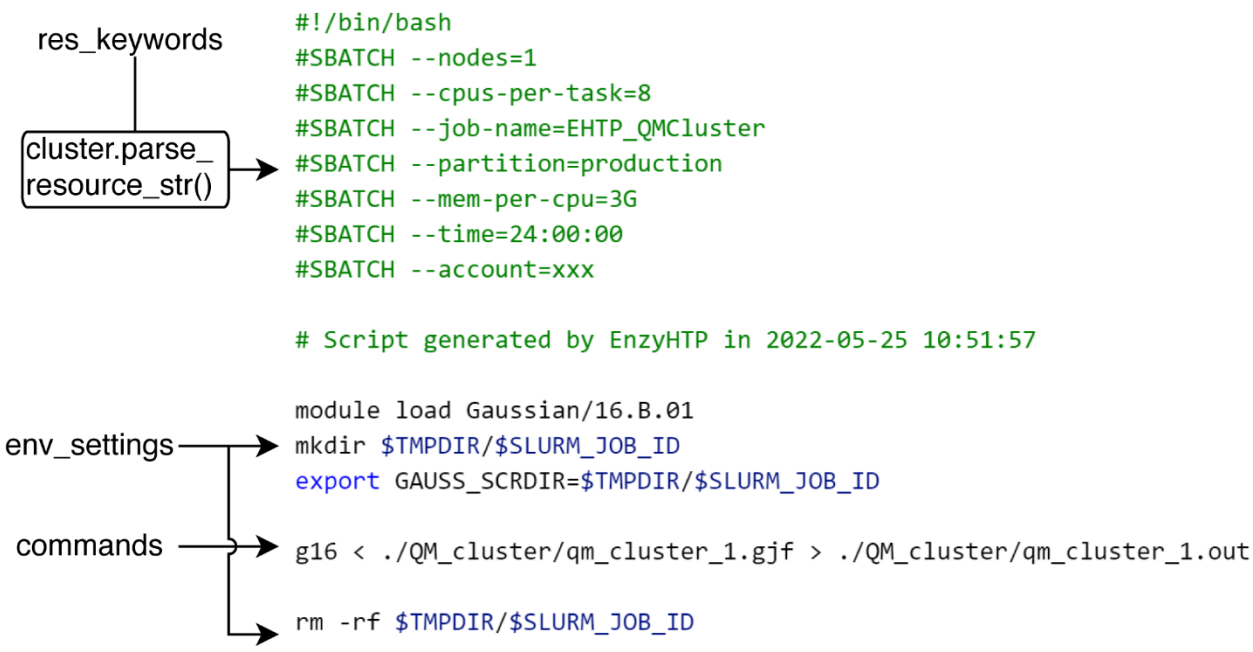

Figure S3. A sample job script that performs QM calculation using Gaussian 16. The job script is generated by the "workflow script" using commands implemented in ARMer.

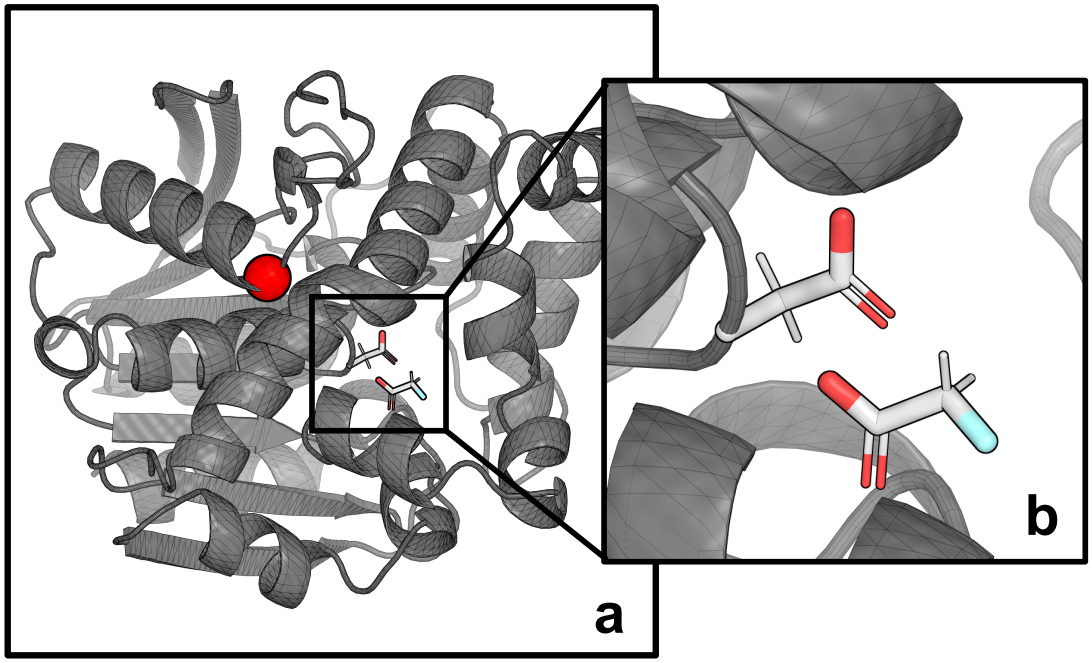

Figure S4. (a) Structure of FAcD K83D mutant and (b) the active site cluster.

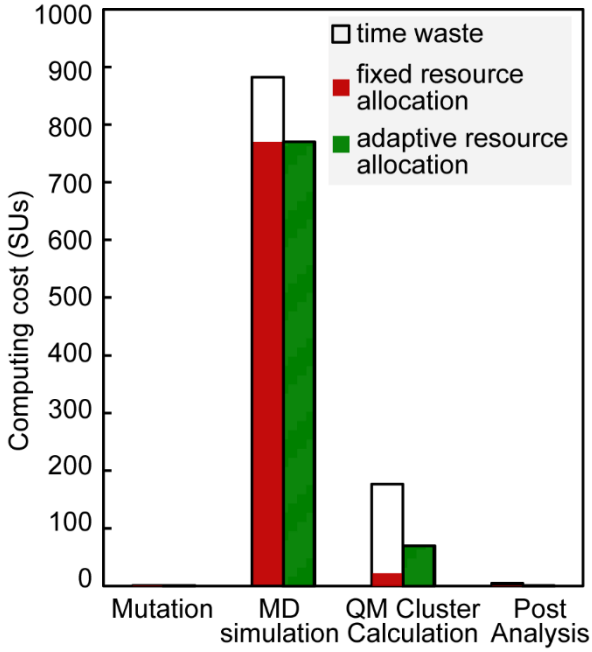

Figure S5. The comparison of computing resource cost between fixed resource allocation (in red) and adaptive resource allocation (in green) for each sub-task in the enzyme FAcD modeling workflow using EnzyHTP. The resource time-waste is colored in white. The computational cost is represented by core-hours or SUs. To normalize the computational cost across CPU and GPU, a weighting factor (i.e., 1 SU of GPU = 54 SUs of CPU) is applied based on the high-performance linpack benchmark.

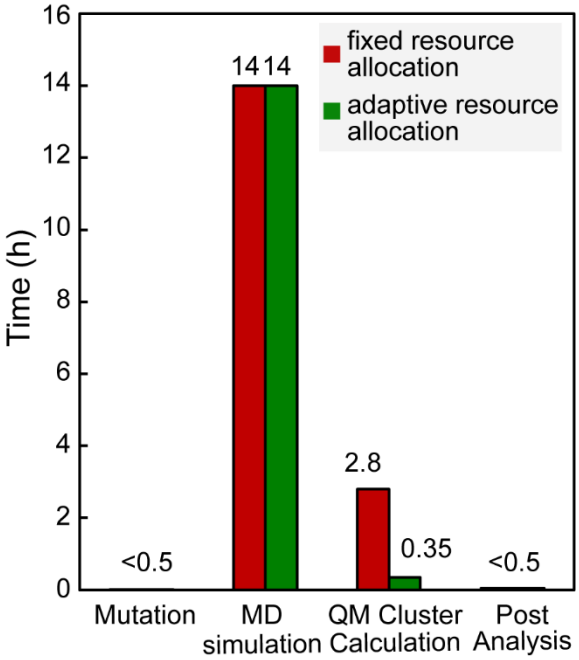

Figure S6. The comparison of wall-clock time cost between fixed resource allocation (in red) and adaptive resource allocation (in green) for each sub-task in the enzyme FAcD modeling workflow using EnzyHTP. Noticeably, the magnitude of efficiency acceleration (8-fold) is less than the ideal condition with parallel job submission (25-fold). The discrepancy is mainly caused by difference of CPU performance in ACCRE and the queueing time (Supporting Information Text S2).

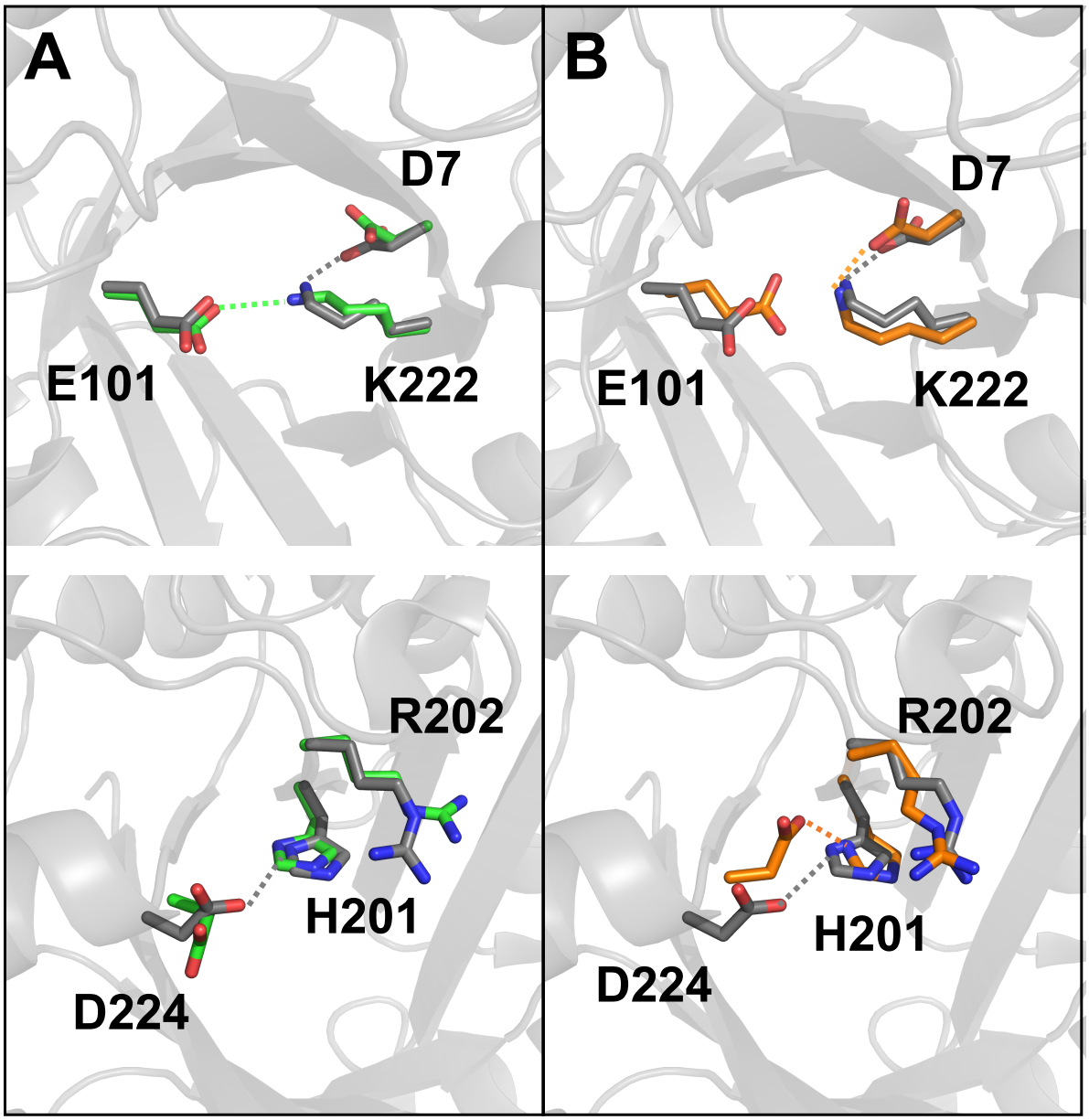

Figure S7. The comparison between EnzyHTP-modeled sidechain conformation (A: generated by the mutation engine; B: the last frame of the MD equilibrium run) and the experimentally determined sidechain conformation upon the mutation of I7D (shown in the upper two panes) and N224D (shown in the lower two panes). The structures colored in grey are obtained from the crystal structure of an apo-R4 mutant (I7D K146E G202R N224D, PDB code: 3IIO). The structures colored in green are obtained from the structure of R5_11-substrate complex generated by the mutation engine of EnzyHTP. The structures colored in orange are obtained from the structure of R5_11-substrate complex after the MD equilibrium run. All hydrogens are hidden. The key H-bond interaction are noted as dotted line. The structures are oriented to be comparable with the analysis from Khersonsky *et al.* [2]

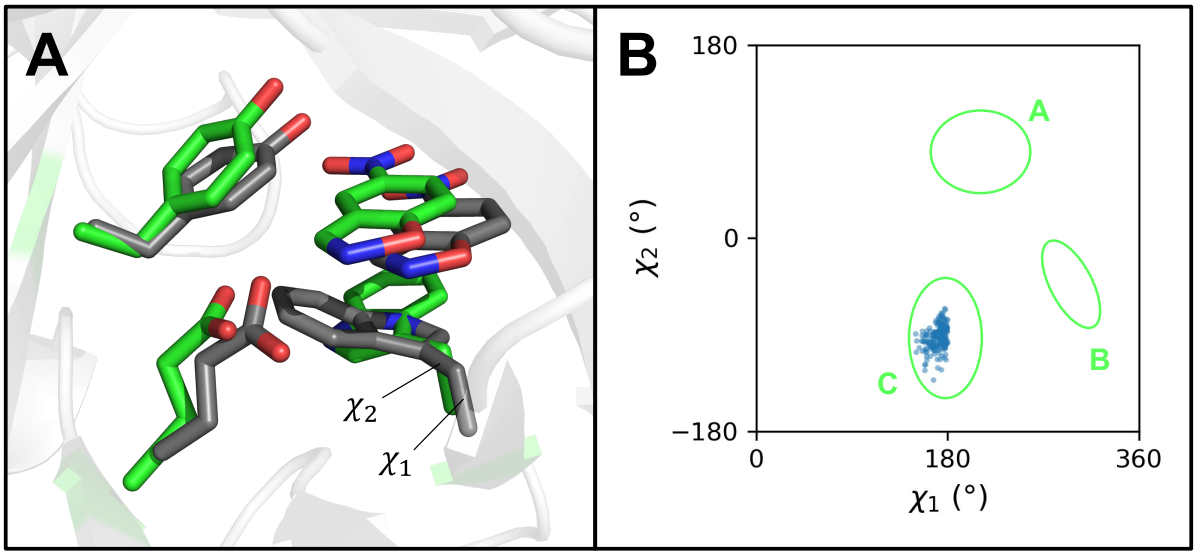

Figure S8. The comparison between modelled substrate binding configurations and experimentally determined binding configurations. (A) The comparison between EnzyHTP modelled substrate binding configuration and experimentally determined binding configuration C. Grey structure is representative structure of the MD trajectory generated by EnzyHTP. The structure of R5_11 is used as an example here. In total 6 representative structures are collected, all of them are in configuration A, only the 1st structure is shown here for clarity. (Find all representative structures for all top-ranked mutants in SI.zip) All hydrogens are hidden here. χ_1 and χ_2 are defined here by the corresponding rotating bond of labelling. (B) The distribution of W50 conformations of all representative structures of all top-ranking mutants. The figure is made that it is comparable to the figure from Hong *et al.*.[3] The region corresponding to the 3 different binding configurations A, B, and C are noted in green (defined by the work of Hong *et al.*)

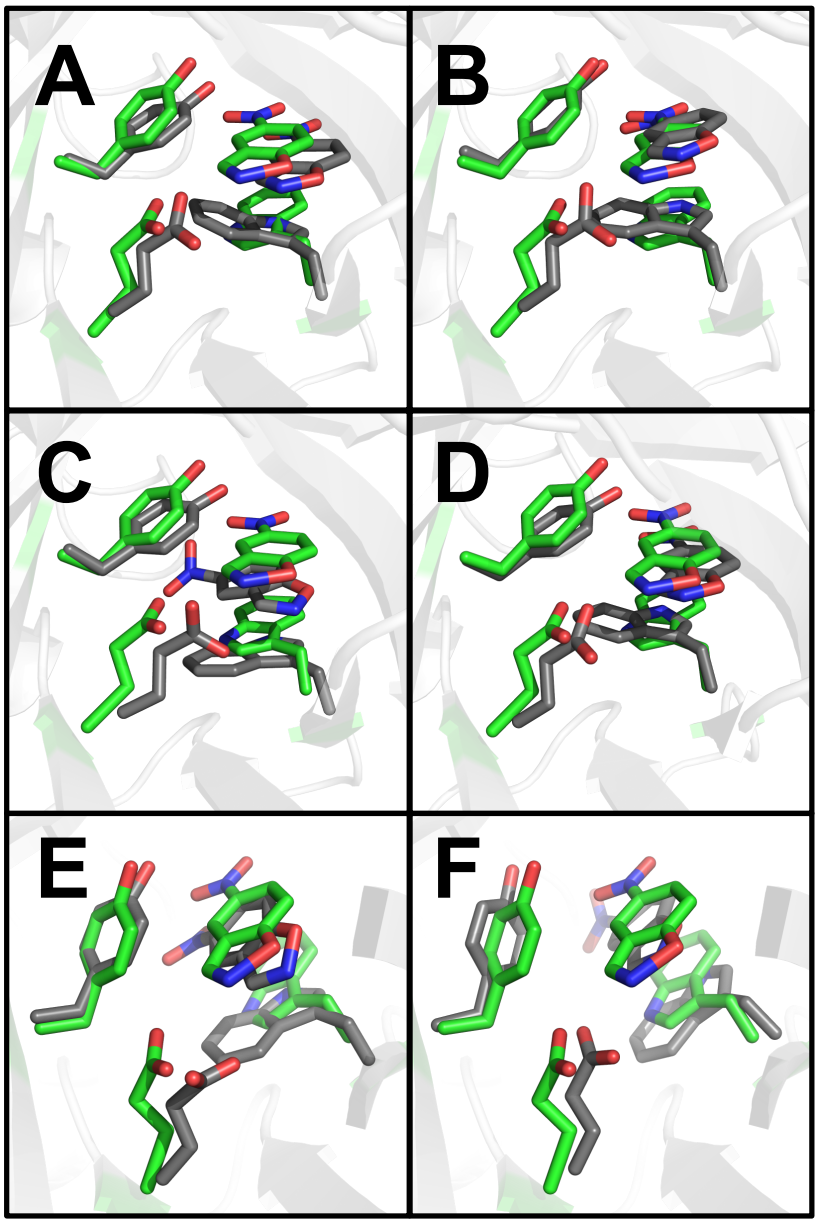

Figure S9. The binding configuration C (grey) from crystal structure (PDB code: 6CT3) comparing with 1st representative structure (green) of MD trajectories generated by EnzyHTP for the four experimentally-observed mutants (A) R5_11 (B) R5_12 (C) R6_71 (D) R6_72 and the No. 1 ranking mutant of each round (E) R5_1 - I7D K146E G202R N224D I102L (F) R6_85 - I7S K19E G202R N224D S144H R175W H201K G145W

Text S2. Details of the benchmark workflow

The first sub-task of the workflow uses the mutant generation module in EnzyHTP to generate a mutant structure (i.e, K83D, Supporting Information Figure S3a) and then optimizes the structure with molecular mechanics using the PMEMD module in AMBER. The time cost for mutant structure construction is nearly negligible and that for minimization takes ~30 seconds on 1 GPU. The second sub-task samples 100 conformers from a 100 ns MD trajectory simulated by the PMEMD module. The time cost of this step is ~14 hours on 1 GPU. Similarly, the computational cost for automatic preparation of MD input files can be omitted. Based on each of the sampled snapshot, the third sub-task conducts QM calculations of single point energy on the active site cluster (i.e., substrate + Asp110, Supporting Information Figure S3b) at the level of PBE0/def2TZVP using Gaussian 16. In total, 100 independent QM calculations should be conducted. With 8 CPUs, each calculation takes ~2 minutes to complete. The last sub-task involves post electronic structure analysis to obtain electrostatic stabilization energy, which is derived from the projection of protein electric field (computed by the point charges of enzyme residues with built-in EnzyHTP function) to the dipole of reacting bond (computed by Multiwfn[4]). The whole analysis take ~0.1 hours using 1 CPU. Overall, the entire workflow involves different types of modeling sub-tasks with diverse requirement of computing resources.

Text S3. Choice of parallel strategy

Parallelization of QM calculation (QM cluster single point energy for 100 snapshots) can be done by requesting either 1) 800 CPUs in one job and handle the parallelization of 100 Gaussian calculation in Python, or 2) 8 CPUs in each job but creating 100 jobs for each Gaussian calculation. We did not choose the first strategy since the resource request goes beyond the capacity in our HPC cluster. HPC are designed to share resources with users across campus. Requesting such a large amount of resources in one job will have a very long queue time and is disturbing the ecosystem of the HPC. For a situation the involves the operation of 100 mutations, there will be 80000 CPU expected to run at the same time. The resource demand is more than the total amount of CPU cores on ACCRE. In contrast, the second solution suits the ecosystem of HPC. It submits jobs to the job queue in a unit of 8 CPU instead of 800 CPU. This will make the queue time less significant and ARM's array submission is set perfectly for its implementation. There will be a maximum number of jobs that can exist on the job queue so that other users won't be influenced by the unexpected long queue time.

Text S4. Discrepancy between ideal and actual speed up from ARMer parallelization

Two reasons contribute to the discrepancy. First, the CPU performances are different. On ACCRE, the CPU installed on the GPU node is tested to be ~1.8 times faster than a regular CPU used in the resource allocation test. Second, queue time become significant when 100 calculations are submitted as 100 individual jobs. There are ~10 min queue time in total while the total running time cost is merely 20 min. However, with larger QM region and higher level

of QM theory, the total running time is expected to be much longer. As such, the fraction of queueing time will become less significant.

Text S5. Details of the thermostability score calculation using cartesian_ddg.

We interfaced the Rosetta cartesian_ddg application in EnzyHTP to perform the thermostability calculation in a high-throughput workflow. The interface was implemented as *get_rosetta_ddg()* function in EnzyHTP. While traditional cartesian_ddg calculation is limited to running on a single CPU process, ARMer library provides a simple solution to parallelize the cartesian_ddg calculation by submitting the calculation of each mutant as a single-core job in a job array using *wait_to_array_end()* function of the ARMer library.

Below are detailed configurations of the cartesian_ddg calculation in this work:
The reference structure, i.e., one of the starting variants of R5, I7D K146E G202R N224D, was used as the starting structure for the thermostability calculation. We used the same structure as the reference for all mutants to reduce the errors involved in the evaluation of mutation effects. The reference structure was first relaxed using the Rosetta *relax* application[5] (*relax_with_rosetta()* in EnzyHTP) as recommend in the cartesian_ddg tutorial by RosettaCommons. 20 decoys were made per input structure and the one with the lowest total energy is returned by the function as the input structure for the subsequent carteisn_ddg calculation. Cartesian_ddg's calculation is based on the reference and mutant conformations sampled in Rosetta using a Monte Carlo method. The conformational sampling involves only all contacting residues within 5.0 Å away from the mutation site. We conducted 10 iterations of such conformational sampling for both the reference and the mutants. The folding free energy calculation of these structures was performed using a Cartesian coordinate representation of the protein structure, and the pairwise interactions were evaluated by the full-atom energy function within 9.0 Å from the mutation site. The free energy difference between the average of the 10 iterations of the mutant and the reference is defined as folding free energy change upon mutation and is used as the thermostability score with a Rosetta energy unit (REU).

We also tested the protocol in which the thermostability score was computed directly from the crystal structure of the mutant. We used the crystal structure of R6_72, which corresponds to R6 3/7F in the original experimental study, as the input structure (PDB code: 5D32[3]) and calculated the Rosetta cartesian_ddg score relative to the reference structure by mutating it back. The result is 3.29 REU, which is close to the value calculated using the reference structure (i.e., 1.45 REU).

Text S6. Details of structure preparation and parameterization.

The crystal structure of KE07 was curated from the Protein Data Bank (PDB ID: 2RKX[6]). All crystallizing water molecules were removed for model construction. We used PyMol[7] to align the substrate-bound KE07 design model[6] with the crystal structure. The aligned design model was used to add the missing residues in the crystal structure (Glu251 to Leu253) and to construct the KE07-substrate complex, which was then prepared for MD simulations with the

AMBER 18 tLEaP tool.[8] The protein was modeled using AMBER ff14SB force field.[9] The generalized AMBER force field[10, 11] was used to parametrize the substrate whose structure was obtained from the PDB structure with the annotation of H5J (5-nitro-1,2-benzoxazole). AM1-BCC model was used to determine the atomic charges of the substrate.[12] The coordinate of hydrogens and protonation states of titratable residues in the enzyme were determined by the *get_protonation()* method in EnzyHTP. (powered by PDB2PQR[13] and PROPKA3[14])

Text S7. Details of the molecular dynamics simulation using AMBER18.

We performed MD simulations of the mutant-substrate complexes using the AMBER 18 software package.[8] The SHAKE algorithm[15] was used to constrain all hydrogen-containing bonds. To sample near transition state conformations throughout the simulations, we applied geometric restraints on the forming O-H distance and the O-H-C orbital interacting angle between the substrate and Glu101 from minimization to production runs (Figure S10). The enzyme complexes were solvated in a periodic truncated octahedron box with a 10 Å buffer of TIP3P water and were neutralized with Na^+ counterions. For each mutant complex, we first relaxed the whole solvent box using the steepest descent method for 10000 steps followed by the conjugate gradient method for another 10000 steps. After minimization, we heated each box from 0 to 300 K within 36 ps with constant volume, equilibrated it for 4 ps under constant volume at 300 K, and further equilibrated at 300 K and 1 atm for 1 ns. In addition, we restrained the backbone C_α and amide C and N atoms with a weight of 2 $kcal \cdot mol^{-1} \cdot Å^{-2}$ from minimization to equilibration. After equilibration, we carried out production runs for 110 ns and output the trajectories every 100 ps. The snapshots from the last 100 ns of the production run were used for analyses, resulting in a total of 1000 snapshots for each production run. All simulations were performed with a time step of 2 fs. The Langevin thermostat[16] and Berendsen barostat[17] were used to control the temperature and pressure, respectively.

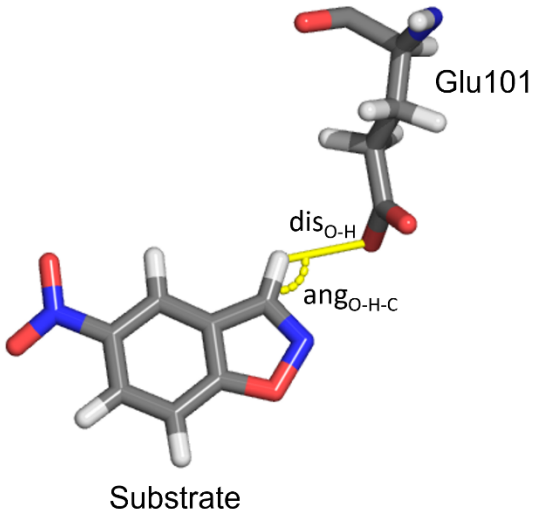

Figure S10. Geometric restraints applied in the MD simulation. The upper distance limit between Glu101 carbonyl O and substrate leaving H, dis_{O-H}, is 2.50 Å. The allowed angle range of Glu101 carbonyl O, substrate leaving H, and substrate deprotonated C, ang_{O-H-C}, is between

150.0 and 180.0°. If the distance or angle is out of the defined range, a 100 kcal mol^{-1} $Å^{-2}$ or 100 kcal mol^{-1} rad^{-2} weight is applied, respectively.

Text S8. Details of the electrostatic stabilization energy calculation.

The *get_field_strength()* method in EnzyHTP is used to calculate the field strength of the enzyme electric field on the center of the reacting bond (the breaking C–H bond) projecting along the bond direction. MM charges generated by tLEaP is used in the calculation. The method returns the result as a list of field strength for all sampled frames from MD. The *get_bond_dipole()* method in EnzyHTP is used to calculate the bond dipole of the same reacting bond. The method returns the result as a list of bond dipole for all sampled frames from MD in the same order as the field strength list. To obtain the electrostatic stabilization energy G_{elec} for a specific MD snapshot, one needs to calculate the negative dot product between the field strength and bond dipole moment. Notably, the G_{elec} is only calculated for the pre-reaction complex (the conformation that is assumed to readily form transition state with minimal geometrical change), but not the reactant or the transition state structure. Although G_{elec} values have been used to identify the beneficial mutants, a more physically meaningful index would be ΔG_{elec}, which is the difference of electrostatic stabilization energy between the transition state and the reactant. Additionally, the electric field calculated in this work is based on MM charges. We are working on implementing a Tinker interface in EnzyHTP2 to support electric field calculation using polarizable force field. Polarizable force field has been suggested by Bradshaw *et al.* to be more accurate in modeling the electric field.[18]

Text S9. Specifications of hardware used in this work.

The simulations enabled by the EnzyHTP & ARMer workflow was preform on the Advanced Computing Center for Research and Education (ACCRE) cluster at Vanderbilt University. For CPU jobs, the "production" partition on the cluster was used. Typical compute nodes in the partition each have 128 or 256 GB of memory and equipped with Intel(R) Xeon(R) CPU E5-2420 or Intel(R) Xeon(R) CPU E5-2630 v3. They all run a 64-bit Linux OS and have a 250 GB – 1 TB hard drive and dual copper gigabit Ethernet ports.
For GPU jobs, the "Pascal" and "Turing" partition on the cluster was used. Typical compute nodes in the partition are each equipped with Nvidia Titan X or GeForce RTX 2080 Ti GPU cards, and are also interconnected with a low-latency 25 or 40/56 Gb/s RoCE network. All compute nodes are monitored via Nagios. Resource management, scheduling of jobs, and usage tracking are handled by an integrated scheduling system by SLURM.

Text S10. Statistical analysis for the chance of randomly identifying the two target mutants in the top-ranked mutants.

For R5, the total number of mutants is 62, the number of experimentally-observed mutants (target) is 2.

Chance for randomly picking a group of 12 mutants and finding the 2 targets in the group $= \frac{C_{60}^{10}}{C_{62}^{12}} = 3.5\%$

For R6, the total number of mutants is 122, the number of experimentally-observed mutants (target) is 2.

Chance for randomly picking a group of 20 mutants and finding the 2 targets in the group $= \frac{C_{120}^{18}}{C_{122}^{20}} = 2.6\%$

For the overall process, the chance for randomly picking the 2 targets from R5 and also 2 targets from R6 is $3.5\% \times 2.6\% = 0.091\%$.

Table S1. List of mutants in different stages of the computational directed evolution workflow. (Mutation notations are relative to KE-07, thermostabilities are calculated relative to one of the starting variants of R5, I7D K146E G202R N224D)

6 Starting variants (R5)

I7D K146E G202R N224D
I7T K146T I199Q F86L I173V L176D F227L
I7S K19E G202R N224D
I7V K146E I199Q N224D L162P I173A L176I F229S
I7Q K146E I199V N224D F227L
I7Q K146E G202R N224D F229S

Mutant library of 62 mutants (single amino acid substitution) (R5)

Index	Mutant	Relative Thermostability (REU)	G_{elec} (kcal/mol)
R5_1	I7D K146E G202R N224D I102L	-2.4294	0.462064
R5_2	I7D K146E G202R N224D G80I	31.2548	1.734618
R5_3	I7D K146E G202R N224D I151T	2.1779	1.177025
R5_4	I7D K146E G202R N224D I83A	10.4976	2.115255
R5_5	I7D K146E G202R N224D L241K	19.8741	1.442738
R5_6	I7D K146E G202R N224D S55P	18.3155	1.423922
R5_7	I7D K146E G202R N224D S122C	-1.8241	2.108269
R5_8	I7D K146E G202R N224D D28H	6.1621	0.568817
R5_9	I7D K146E G202R N224D L10E	12.2746	1.706855
R5_10	I7D K146E G202R N224D I83E	13.2609	1.404584
R5_11	I7D G202R N224D V12M	1.126	0.873856
R5_12	I7D K146E G202R N224D L47I	2.776	2.000201
R5_13	I7T K146T I199Q F86L I173V L176D F227L V125N	13.683	1.394202
R5_14	I7T K146T I199Q F86L I173V L176D F227L I75T	16.2216	2.124284
R5_15	I7T K146T I199Q F86L I173V L176D F227L G136A	7.5485	0.549972
R5_16	I7T K146T I199Q F86L I173V L176D F227L L215C	5.204	1.584209
R5_17	I7T K146T I199Q F86L I173V L176D F227L T53Y	4.9345	2.400544
R5_18	I7T K146T I199Q F86L I173V L176D F227L E231S	1.576	1.307885
R5_19	I7T K146T I199Q F86L I173V L176D F227L D85W	3.768	2.09752
R5_20	I7T K146T I199Q F86L I173V L176D F227L L94R	5.6372	1.202644
R5_21	I7T K146T I199Q F86L I173V L176D F227L E64W	-4.8982	2.001677
R5_22	I7T K146T I199Q F86L I173V L176D F227L F77V	7.3068	1.682689
R5_23	I7S K19E G202R N224D L26A	13.3037	2.419008
R5_24	I7S K19E G202R N224D V190M	3.312	1.81938
R5_25	I7S K19E G202R N224D E87I	-2.0295	2.176079
R5_26	I7S K19E G202R N224D S144Y	1.0709	2.341676

R5_27	I7S K19E G202R N224D S144H	-0.3474	1.931238
R5_28	I7S K19E G202R N224D E71Q	-2.5053	2.061092
R5_29	I7S K19E G202R N224D E239M	-3.3136	1.782684
R5_30	I7S K19E G202R N224D E64M	-3.6519	2.481016
R5_31	I7S K19E G202R N224D G164Q	-0.5316	2.071931
R5_32	I7S K19E G202R N224D I168W	16.0304	2.259422
R5_33	I7V K146E I199Q N224D L162P I173A L176I F229S R5K	39.5956	1.553637
R5_34	I7V K146E I199Q N224D L162P I173A L176I F229S H209E	50.3112	3.37698
R5_35	I7V K146E I199Q N224D L162P I173A L176I F229S N109F	29.9434	2.632079
R5_36	I7V K146E I199Q N224D L162P I173A L176I F229S G121R	38.4373	2.36312
R5_37	I7V K146E I199Q N224D L162P I173A L176I F229S T195Q	43.1603	1.745177
R5_38	I7V K146E I199Q N224D L162P I173A L176I F229S L211N	52.1307	2.53346
R5_39	I7V K146E I199Q N224D L162P I173A L176I F229S V79N	53.1849	3.001068
R5_40	I7V K146E I199Q N224D L162P I173A L176I F229S L237E	56.044	1.309283
R5_41	I7V K146E I199Q N224D L162P I173A L176I F229S L196Q	47.4261	2.370785
R5_42	I7V K146E I199Q N224D L162P I173A L176I F229S L35E	56.2644	2.962428
R5_43	I7Q K146E I199V N224D F227L F38G	12.5548	2.560619
R5_44	I7Q K146E I199V N224D F227L V127M	6.0127	1.298944
R5_45	I7Q K146E I199V N224D F227L E231I	-6.3458	2.875185
R5_46	I7Q K146E I199V N224D F227L L176S	-0.342	2.815547
R5_47	I7Q K146E I199V N224D F227L W50E	6.5927	-0.31737
R5_48	I7Q K146E I199V N224D F227L Q123C	-4.9201	2.41841
R5_49	I7Q K146E I199V N224D F227L E212T	-3.5428	3.11804
R5_50	I7Q K146E I199V N224D F227L I232G	13.5628	1.11532
R5_51	I7Q K146E I199V N224D F227L V33K	10.0982	1.250266
R5_52	I7Q K146E I199V N224D F227L K4E	11.844	1.888153
R5_53	I7Q K146E G202R N224D F229S G96L	-0.93	2.03562
R5_54	I7Q K146E G202R N224D F229S N103Q	-2.5568	1.278448
R5_55	I7Q K146E G202R N224D F229S G177P	16.0494	2.216208
R5_56	I7Q K146E G202R N224D F229S A223R	15.2292	2.154399
R5_57	I7Q K146E G202R N224D F229S L237R	17.543	2.364616
R5_58	I7Q K146E G202R N224D F229S A220F	4.7706	2.259814
R5_59	I7Q K146E G202R N224D F229S P197W	-0.0449	1.872679

R5_60	I7Q K146E G202R N224D F229S K132S	1.5713	1.797461
R5_61	I7Q K146E G202R N224D F229S G36W	41.3776	2.335375
R5_62	I7Q K146E G202R N224D F229S D51P	9.2742	2.060396

12 Starting Variants (R6)

I7D K146E G202R N224D I102L
I7D G202R N224D V12M
I7D K146E G202R N224D I151T
I7Q K146E G202R N224D F229S N103Q
I7T K146T I199Q F86L I173V L176D F227L E231S
I7S K19E G202R N224D E239M
I7Q K146E G202R N224D F229S K132S
I7S K19E G202R N224D V190M
I7Q K146E G202R N224D F229S P197W
I7S K19E G202R N224D S144H
I7D K146E G202R N224D L47I
I7T K146T I199Q F86L I173V L176D F227L E64W

Mutant library of 122 mutants (triple amino acid substitution) (R6)

Index	Mutant	Relative Thermostability (REU)	G_{elec} (kcal/mol)
R6_1	I7D K146E G202R N224D I151T A200E T104P T61V	21.2327	0.614574
R6_2	I7D K146E G202R N224D I151T E34Y Q118Y E212T	6.5956	1.628815
R6_3	I7D K146E G202R N224D I151T V134F E64Q S111M	0.3563	1.018563
R6_4	I7D K146E G202R N224A I151T N247E Y143K	9.1533	0.55926
R6_5	I7D K146E G202R N224D I151T E34T E159L E161S	9.2208	0.865038
R6_6	I7D K146E G202R N224D I151T E87R M139G I102W	18.7026	0.888853
R6_7	I7D K146E G202R N224D I151T H228G L241V L65V	20.2239	1.776604
R6_8	I7D K146E G202R N224D I151T K19A A124H R59H	8.1041	2.028205
R6_9	I7D K146E G202R N224D I151T A9I V140Y K147R	32.8882	1.550033
R6_10	I7D K146E G202R N224D I151T G43V N25M G80D	39.5216	1.313835
R6_11	I7Q K146E G202R N224D F229S N103Q F227K G164E V169T	24.7841	0.979464
R6_12	I7Q K146E G202R N224D F229S N103Q E239L K60R K243H	-8.3477	1.985088
R6_13	I7Q K146E G202R N224D F229S N103Q P192M R230S W50P	16.52	0.607212
R6_14	I7Q K146E G202R N224D F229S N103Q D219G L241S V246D	33.8276	0.518982
R6_15	I7Q K146E G202R N224D F229S N103Q A218W P76N E159C	47.2909	1.65398
R6_16	I7Q K146E G202R N224D F229S N103Q L47V V18K K206F	4.3307	1.826064
R6_17	I7Q K146E G202R N224D F229S N103Q Q72K E251Y E71T	-1.4923	1.806021
R6_18	I7Q K146E G202R N224D F229S N103Q F120I D183S E108Y	7.5634	1.486437
R6_19	I7Q K146E G202R N224D F229S N103Q R175Q F77T R163C	6.8093	0.953207
R6_20	I7Q K146E G202R N224D F229S N103Q D183P F86K N22E	5.8305	1.335689
R6_21	I7S K19E G202R N224D V190M G15N I52T I102Q	15.0419	0.830404

R6_22	I7S K19E G202R N224D V190M S225G F138D K60W	16.7671	2.002205
R6_23	I7S K19E G202R N224D V190M D174N A8P A106Q	35.6279	2.212158
R6_24	I7S K19E G202R N224D V190M T104S F38Q G80A	23.6187	2.535019
R6_25	I7S K19E G202R N224D V190M H228F E236Y Y39E	7.5988	1.900614
R6_26	I7S K19E G202R N224D V190M I83T E57W E24M	12.2654	1.643638
R6_27	I7S K19E G202R N224D V190M F227D G96Y E91A	25.4014	2.091616
R6_28	I7S K19E G202R N224D V190M Y182L S29V G181M	6.6493	1.619869
R6_29	I7S K19E G202R N224D V190M S180E I168K I83Q	34.5987	2.077871
R6_30	I7S K19E G202R N224D V190M G96Y G80H G121H	57.3862	2.681765
R6_31	I7S K19E G202R N224D E239M D85W E159L V18S	5.3568	2.290612
R6_32	I7S K19E G202R N224D E239M Q118A K68A G171Y	3.5626	2.112947
R6_33	I7S K19E G202R N224D E239M S180T I151T I93F	0.4268	2.044038
R6_34	I7S K19E G202R N224D E239M L250E N25T L153A	10.9356	1.630455
R6_35	I7S K19E G202R N224D E239M R154E D51F S111I	1.4912	2.038115
R6_36	I7S K19E G202R N224D E239M G81A E34P V248I	9.2252	1.785763
R6_37	I7S K19E G202R N224D E239M G36W A218N T114W	49.1664	1.883796
R6_38	I7S K19E G202R N224D E239M A131H V66T V248F	25.9938	2.131537
R6_39	I7S K19E G202R N224D E239M K162W P192M V160G	8.2398	2.293581
R6_40	I7S K19E G202R N224D E239M G203K N22D A130W	22.6763	1.906366
R6_41	I7T K146T I199Q F86L I173C L176D F227L E64W G166M S180P	1.0461	2.366244
R6_42	I7T K146T I199Q F86L I173V L176D F227L E64W F210C H209V L215T	17.397	2.135462
R6_43	I7T K146T I199Q F86L I173V L176D F227L E64W L10D S111H V158E	12.6641	1.383808
R6_44	I7T K146T I199Q F86L I173V L176D F227L E64W T114Y K222M S90R	-12.5303	0.769945
R6_45	I7T K146T I199Q F86L I173V L176D F227L E64W K13N N109A A54E	1.5054	2.386343
R6_46	I7T K146T I199Q F86L I173V L176D F227L E64W A204G A200C L112N	10.1206	0.77211
R6_47	I7T K146T I199Q F86L I173V L176D F227L E64W K179M F210M Q115R	10.0342	1.989846
R6_48	I7T K146T I199Q F86L I173V L176D F227L E64W Q123K V126T R133V	-0.082	2.371224
R6_49	I7T K146T I199Q F86L I173V L176D F227L E64W L63V E161T V157H	10.0044	2.357851
R6_50	I7T K146T I199Q F86L I173V L176D F227L E64W V158W R163I L112R	11.4778	2.119153
R6_51	I7T K146T I199Q F86L I173V L176D F227L E231S L65A V234Q E57G	19.895	1.576543
R6_52	I7T K146T I199Q F86L I173V L176D F227L E231S A54I A70G L170D	20.9452	1.174878
R6_53	I7T K146T I199Q F86L I173V L176D F227L E231S N224L V234H E161V	15.8692	2.017679

R6_54	I7T K146T I199Q F86L I173V L176D F227L E231S K238S E185H V160D	29.0821	2.125947
R6_55	I7T K146T I199Q F86L I173V L176D F227L E231S M1T M207E	12.5544	1.032211
R6_56	I7T K146T I199Q F86L I173V L176D F227L E231S D45K K222A E236D	15.1899	0.163357
R6_57	I7T K146T I199Q F86L I173V L176D F227L E231S W156Y E167F T114M	12.4308	1.896938
R6_58	I7T K146T I199Q F86L I173V L176D F227L E231S D98F F189I A130G	-8.3726	1.809391
R6_59	I7T K146T I199Q F86L I173V L176D F227L E231S T119W Q118W A200M	-8.8877	2.049167
R6_60	I7T K146T I199Q F86L I173V L176D F227L E231S D85R G181S T119P	22.7848	1.032753
R6_61	I7D G202R N224D V12M R5D E251I G164W	10.3214	0.691923
R6_62	I7D G202R N224D V12M T149F I129Y W50C	38.1102	0.427876
R6_63	I7D G202R N224D V12M V157F D31I A200Q	22.2338	1.709476
R6_64	I7D G202R N224D V12M I113P N25T E87L	9.7947	-0.23825
R6_65	I7D G202R N224D V12R Q115C S48C	9.6933	1.054112
R6_66	I7D G202R N224D V12M A200D H228T G82N	18.9655	1.687562
R6_67	I7D G202R N224D V12M L211D T53D L215I	17.6574	1.345295
R6_68	I7D G202R N224D V12M M207F G80H L253D	53.2226	0.605108
R6_69	I7D G202S N224D V12M G150P E91L	7.7947	1.799092
R6_70	I7D G202R N224D V12M A3C L152A G177R	7.6975	0.447504
R6_71	I7D G202R N224D V12M A54D G171A	-3.1845	1.306708
R6_72	I7D G202R N224D K19E K146T	1.453	1.523727
R6_73	I7Q K146E G202R N224K F229S K132S F141H T142L	11.9457	1.099562
R6_74	I7Q K146E G202R N224D F229S K132S E46M G80L L10K	41.0735	1.050319
R6_75	I7Q K146E G202R N224D F229Y K132S K60F E185H	-1.763	1.644657
R6_76	I7Q K146E G202R N224D F229S K132S I6E V190L D135A	27.5338	1.711406
R6_77	I7Q K146E G202R N224D F229S K132S V56H E251Y Y182N	21.5343	1.796814
R6_78	I7Q K146E G202Q N224D F229S K132S D74S R163W	8.1405	2.483228
R6_79	I7Q K146E G202R N224D F229S K132S V140Q F227A Y39V	36.7006	1.235729
R6_80	I7Q K146E G202R N224D F229S K132S Y143M K4V K242W	-3.0529	1.259893
R6_81	I7Q K146E G202R N224D F229S K132S V125R D174N I173H	21.2476	1.069762
R6_82	I7Q K146E G202R N224D F229S K132S A216V T104V D183L	6.8825	1.606626
R6_83	I7S K19E G202R N224G S144H L35A F214K	27.2149	1.587074
R6_84	I7S K19E G202R N224D S144H G217Q T178N S122C	4.3257	2.118727
R6_85	I7S K19E G202R N224D S144H R175W H201K G145W	2.1054	-0.42473
R6_86	I7S K19E G202R N224D S144H I93T V169S A204M	26.0211	2.25983
R6_87	I7S K19E G202R N224D S144H K13Y Y240S G166C	8.6285	1.523624
R6_88	I7S K19E G202R N224D S144H V140F G43E G217Y	34.3372	1.755721
R6_89	I7S K19E G202R N224D S144H T178S A105C A8Y	15.9739	1.673982
R6_90	I7S K19E G202R N224D S144H V234M Y143F W156G	23.7117	1.729843

R6_91	I7S K19E G202R N224D S144H S90W D155L W156F	-7.6782	1.896569
R6_92	I7S K19E G202R N224D S144H A9I Q115V G150C	6.5945	1.903842
R6_93	I7D K146E G202R N224D I102L E108P K206R A8P	24.7477	2.032606
R6_94	I7D K146E G202R N224D I102L F38Y G80I D183V	39.9394	0.887196
R6_95	I7D K146E G202R N224D I102L L253M D135A Y143D	3.2468	1.638967
R6_96	I7D K146E G202R N224D I102L A130H I168L E91I	9.69	1.407671
R6_97	I7D K146E G202R N224D I102L K99M P76V I232Y	10.4951	1.5209
R6_98	I7D K146E G202R N224D I102L V12C H84W V79M	5.7513	0.768532
R6_99	I7D K146E G202R N224D I102L L170A F38A T53R	17.8256	1.554059
R6_100	I7D K146E G202R N224D I102L V248K F120E D174M	28.1362	1.807758
R6_101	I7D K146E G202R N224D I102L A97C I42L F49M	4.8473	0.973602
R6_102	I7D K146E G202R N224D I102L V18E I42G I187V	13.0953	0.086593
R6_103	I7Q K146E G202R N224D F229S P197W R188Q H84C V69N	9.9712	1.337716
R6_104	I7Q K146E G202W N224D F229S P197W Q72V Y143L	-0.9068	2.546014
R6_105	I7Q K146E G202R N224D F229S P197W K13I D135C M62R	10.1141	2.42903
R6_106	I7Q K146E G202R N224D F229S P197W L241F S111T R154W	13.412	1.339951
R6_107	I7Q K146E G202R N224D F229S P197W T61L L94H M207A	7.0261	2.135224
R6_108	I7Q K146E G202R N224D F229S P197W W156A K37L V69M	16.94	1.886372
R6_109	I7Q K146E G202R N224D F229S P197W K147S I93V T78G	10.0105	1.348725
R6_110	I7Q K146E G202R N224D F229S P197W L35A A105N F120Q	25.592	1.356338
R6_111	I7Q K146E G202R N224D F229S P197W D219T A9I L250S	24.1969	1.302424
R6_112	I7Q K146E G202R N224D F229S P197W V157G A221C V33F	15.2964	1.457007
R6_113	I7D K146E G202R N224D L47I S111G I232P S90H	19.9695	1.878916
R6_114	I7D K146E G202R N224D L47I A105H M1K E251V	15.4443	1.787848
R6_115	I7D K146E G202R N224P L47I A70Q V125L	10.5668	0.68505
R6_116	I7D K146E G202R N224D L47I L176F A218Q G82S	24.5314	1.945509
R6_117	I7D K146E G202R N224D L47I I168E Q72H D135V	19.7095	0.296108
R6_118	I7D K146E G202R N224D L47I P110N I42T L237V	21.9016	1.450983
R6_119	I7D K146E G202R N224D L47I S144K S111C I42Q	3.7206	1.727913
R6_120	I7D K146E G202R N224D L47I T194P A105T L152G	30.4214	1.459889
R6_121	I7D K146E G202R N224D L47I G164P I187V A117R	68.3046	2.121582
R6_122	I7D K146E G202R N224D L47I K147V T195P A9F	8.6732	0.830204

Table S2. R5 and R6 mutants with G_{elec} values calculated based on the first 50 ns out of each 100 ns that used in the calculation of Table S1.

R5 Mutants

Index	Mutant	G_{elec} (kcal/mol)
R5_1	I7D K146E G202R N224D I102L	1.05767
R5_2	I7D K146E G202R N224D G80I	1.400463
R5_3	I7D K146E G202R N224D I151T	1.366507
R5_4	I7D K146E G202R N224D I83A	2.144801
R5_5	I7D K146E G202R N224D L241K	1.495664
R5_6	I7D K146E G202R N224D S55P	1.775713
R5_7	I7D K146E G202R N224D S122C	2.418011
R5_8	I7D K146E G202R N224D D28H	0.932835
R5_9	I7D K146E G202R N224D L10E	1.727324
R5_10	I7D K146E G202R N224D I83E	1.338194
R5_11	I7D G202R N224D V12M	1.285187
R5_12	I7D K146E G202R N224D L47I	2.413852
R5_13	I7T K146T I199Q F86L I173V L176D F227L V125N	1.553837
R5_14	I7T K146T I199Q F86L I173V L176D F227L I75T	2.134146
R5_15	I7T K146T I199Q F86L I173V L176D F227L G136A	0.386611
R5_16	I7T K146T I199Q F86L I173V L176D F227L L215C	2.267702
R5_17	I7T K146T I199Q F86L I173V L176D F227L T53Y	2.219869
R5_18	I7T K146T I199Q F86L I173V L176D F227L E231S	1.361821
R5_19	I7T K146T I199Q F86L I173V L176D F227L D85W	2.232541
R5_20	I7T K146T I199Q F86L I173V L176D F227L L94R	2.0791
R5_21	I7T K146T I199Q F86L I173V L176D F227L E64W	2.078755
R5_22	I7T K146T I199Q F86L I173V L176D F227L F77V	1.750233
R5_23	I7S K19E G202R N224D L26A	2.509062
R5_24	I7S K19E G202R N224D V190M	1.910481
R5_25	I7S K19E G202R N224D E87I	2.298244
R5_26	I7S K19E G202R N224D S144Y	2.31771
R5_27	I7S K19E G202R N224D S144H	2.01021
R5_28	I7S K19E G202R N224D E71Q	2.011669
R5_29	I7S K19E G202R N224D E239M	1.771831
R5_30	I7S K19E G202R N224D E64M	2.659577
R5_31	I7S K19E G202R N224D G164Q	2.064943
R5_32	I7S K19E G202R N224D I168W	2.332201
R5_33	I7V K146E I199Q N224D L162P I173A L176I F229S R5K	2.413945
R5_34	I7V K146E I199Q N224D L162P I173A L176I F229S H209E	3.405914
R5_35	I7V K146E I199Q N224D L162P I173A L176I F229S N109F	2.787596

R5_36	I7V K146E I199Q N224D L162P I173A L176I F229S G121R	2.751543
R5_37	I7V K146E I199Q N224D L162P I173A L176I F229S T195Q	1.548805
R5_38	I7V K146E I199Q N224D L162P I173A L176I F229S L211N	2.753345
R5_39	I7V K146E I199Q N224D L162P I173A L176I F229S V79N	3.075296
R5_40	I7V K146E I199Q N224D L162P I173A L176I F229S L237E	1.393274
R5_41	I7V K146E I199Q N224D L162P I173A L176I F229S L196Q	2.567829
R5_42	I7V K146E I199Q N224D L162P I173A L176I F229S L35E	3.024855
R5_43	I7Q K146E I199V N224D F227L F38G	2.405695
R5_44	I7Q K146E I199V N224D F227L V127M	1.414554
R5_45	I7Q K146E I199V N224D F227L E231I	2.866997
R5_46	I7Q K146E I199V N224D F227L L176S	2.825156
R5_47	I7Q K146E I199V N224D F227L W50E	-0.67384
R5_48	I7Q K146E I199V N224D F227L Q123C	2.742804
R5_49	I7Q K146E I199V N224D F227L E212T	2.98337
R5_50	I7Q K146E I199V N224D F227L I232G	1.484956
R5_51	I7Q K146E I199V N224D F227L V33K	1.299182
R5_52	I7Q K146E I199V N224D F227L K4E	2.35174
R5_53	I7Q K146E G202R N224D F229S G96L	1.921055
R5_54	I7Q K146E G202R N224D F229S N103Q	1.801213
R5_55	I7Q K146E G202R N224D F229S G177P	2.42626
R5_56	I7Q K146E G202R N224D F229S A223R	2.062878
R5_57	I7Q K146E G202R N224D F229S L237R	2.201105
R5_58	I7Q K146E G202R N224D F229S A220F	2.214115
R5_59	I7Q K146E G202R N224D F229S P197W	1.610854
R5_60	I7Q K146E G202R N224D F229S K132S	1.78619
R5_61	I7Q K146E G202R N224D F229S G36W	2.257044
R5_62	I7Q K146E G202R N224D F229S D51P	2.482978

R6 Mutants

Index	Mutant	G_{elec} (kcal/mol)
R6_1	I7D K146E G202R N224D I151T A200E T104P T61V	0.668798
R6_2	I7D K146E G202R N224D I151T E34Y Q118Y E212T	1.661571
R6_3	I7D K146E G202R N224D I151T V134F E64Q S111M	1.063145
R6_4	I7D K146E G202R N224A I151T N247E Y143K	0.61898
R6_5	I7D K146E G202R N224D I151T E34T E159L E161S	0.955816
R6_6	I7D K146E G202R N224D I151T E87R M139G I102W	1.461106
R6_7	I7D K146E G202R N224D I151T H228G L241V L65V	1.881447
R6_8	I7D K146E G202R N224D I151T K19A A124H R59H	1.770084
R6_9	I7D K146E G202R N224D I151T A9I V140Y K147R	1.479986
R6_10	I7D K146E G202R N224D I151T G43V N25M G80D	1.906273

R6_11	I7Q K146E G202R N224D F229S N103Q F227K G164E V169T	0.990702
R6_12	I7Q K146E G202R N224D F229S N103Q E239L K60R K243H	1.804799
R6_13	I7Q K146E G202R N224D F229S N103Q P192M R230S W50P	0.628775
R6_14	I7Q K146E G202R N224D F229S N103Q D219G L241S V246D	0.521394
R6_15	I7Q K146E G202R N224D F229S N103Q A218W P76N E159C	1.388519
R6_16	I7Q K146E G202R N224D F229S N103Q L47V V18K K206F	1.609724
R6_17	I7Q K146E G202R N224D F229S N103Q Q72K E251Y E71T	1.714088
R6_18	I7Q K146E G202R N224D F229S N103Q F120I D183S E108Y	1.348386
R6_19	I7Q K146E G202R N224D F229S N103Q R175Q F77T R163C	1.830235
R6_20	I7Q K146E G202R N224D F229S N103Q D183P F86K N22E	1.594762
R6_21	I7S K19E G202R N224D V190M G15N I52T I102Q	0.401359
R6_22	I7S K19E G202R N224D V190M S225G F138D K60W	1.89525
R6_23	I7S K19E G202R N224D V190M D174N A8P A106Q	2.373056
R6_24	I7S K19E G202R N224D V190M T104S F38Q G80A	2.555123
R6_25	I7S K19E G202R N224D V190M H228F E236Y Y39E	2.148678
R6_26	I7S K19E G202R N224D V190M I83T E57W E24M	1.790345
R6_27	I7S K19E G202R N224D V190M F227D G96Y E91A	2.217455
R6_28	I7S K19E G202R N224D V190M Y182L S29V G181M	1.681322
R6_29	I7S K19E G202R N224D V190M S180E I168K I83Q	2.007352
R6_30	I7S K19E G202R N224D V190M G96Y G80H G121H	2.689698
R6_31	I7S K19E G202R N224D E239M D85W E159L V18S	2.460356
R6_32	I7S K19E G202R N224D E239M Q118A K68A G171Y	2.127553
R6_33	I7S K19E G202R N224D E239M S180T I151T I93F	1.950752
R6_34	I7S K19E G202R N224D E239M L250E N25T L153A	1.823458
R6_35	I7S K19E G202R N224D E239M R154E D51F S111I	2.043237
R6_36	I7S K19E G202R N224D E239M G81A E34P V248I	2.018056
R6_37	I7S K19E G202R N224D E239M G36W A218N T114W	1.934048
R6_38	I7S K19E G202R N224D E239M A131H V66T V248F	2.1142
R6_39	I7S K19E G202R N224D E239M K162W P192M V160G	2.290127
R6_40	I7S K19E G202R N224D E239M G203K N22D A130W	1.203576
R6_41	I7T K146T I199Q F86L I173C L176D F227L E64W G166M S180P	2.483159
R6_42	I7T K146T I199Q F86L I173V L176D F227L E64W F210C H209V L215T	2.248835
R6_43	I7T K146T I199Q F86L I173V L176D F227L E64W L10D S111H V158E	1.555916
R6_44	I7T K146T I199Q F86L I173V L176D F227L E64W T114Y K222M S90R	0.748774
R6_45	I7T K146T I199Q F86L I173V L176D F227L E64W K13N N109A A54E	2.424472
R6_46	I7T K146T I199Q F86L I173V L176D F227L E64W A204G A200C L112N	1.052807
R6_47	I7T K146T I199Q F86L I173V L176D F227L E64W K179M F210M Q115R	1.823341
R6_48	I7T K146T I199Q F86L I173V L176D F227L E64W Q123K V126T R133V	2.61254
R6_49	I7T K146T I199Q F86L I173V L176D F227L E64W L63V E161T V157H	2.606737
R6_50	I7T K146T I199Q F86L I173V L176D F227L E64W V158W R163I L112R	2.246874
R6_51	I7T K146T I199Q F86L I173V L176D F227L E231S L65A V234Q E57G	1.883324
R6_52	I7T K146T I199Q F86L I173V L176D F227L E231S A54I A70G L170D	1.25915

R6_53	I7T K146T I199Q F86L I173V L176D F227L E231S N224L V234H E161V	2.095579
R6_54	I7T K146T I199Q F86L I173V L176D F227L E231S K238S E185H V160D	2.062308
R6_55	I7T K146T I199Q F86L I173V L176D F227L E231S M1T M207E	1.852295
R6_56	I7T K146T I199Q F86L I173V L176D F227L E231S D45K K222A E236D	0.185846
R6_57	I7T K146T I199Q F86L I173V L176D F227L E231S W156Y E167F T114M	2.47455
R6_58	I7T K146T I199Q F86L I173V L176D F227L E231S D98F F189I A130G	2.118176
R6_59	I7T K146T I199Q F86L I173V L176D F227L E231S T119W Q118W A200M	2.128687
R6_60	I7T K146T I199Q F86L I173V L176D F227L E231S D85R G181S T119P	1.163173
R6_61	I7D G202R N224D V12M R5D E251I G164W	0.587225
R6_62	I7D G202R N224D V12M T149F I129Y W50C	0.497627
R6_63	I7D G202R N224D V12M V157F D31I A200Q	1.149388
R6_64	I7D G202R N224D V12M I113P N25T E87L	-0.10702
R6_65	I7D G202R N224D V12R Q115C S48C	1.085609
R6_66	I7D G202R N224D V12M A200D H228T G82N	1.286209
R6_67	I7D G202R N224D V12M L211D T53D L215I	1.718081
R6_68	I7D G202R N224D V12M M207F G80H L253D	0.757418
R6_69	I7D G202S N224D V12M G150P E91L	2.002389
R6_70	I7D G202R N224D V12M A3C L152A G177R	0.885939
R6_71	I7D G202R N224D V12M A54D G171A	1.120153
R6_72	I7D G202R N224D K19E K146T	1.596297
R6_73	I7Q K146E G202R N224K F229S K132S F141H T142L	1.169497
R6_74	I7Q K146E G202R N224D F229S K132S E46M G80L L10K	1.189239
R6_75	I7Q K146E G202R N224D F229Y K132S K60F E185H	1.587524
R6_76	I7Q K146E G202R N224D F229S K132S I6E V190L D135A	1.671097
R6_77	I7Q K146E G202R N224D F229S K132S V56H E251Y Y182N	1.597737
R6_78	I7Q K146E G202Q N224D F229S K132S D74S R163W	2.771105
R6_79	I7Q K146E G202R N224D F229S K132S V140Q F227A Y39V	1.315407
R6_80	I7Q K146E G202R N224D F229S K132S Y143M K4V K242W	1.229661
R6_81	I7Q K146E G202R N224D F229S K132S V125R D174N I173H	1.684372
R6_82	I7Q K146E G202R N224D F229S K132S A216V T104V D183L	1.485294
R6_83	I7S K19E G202R N224G S144H L35A F214K	1.297993
R6_84	I7S K19E G202R N224D S144H G217Q T178N S122C	2.228668
R6_85	I7S K19E G202R N224D S144H R175W H201K G145W	-0.55657
R6_86	I7S K19E G202R N224D S144H I93T V169S A204M	2.596152
R6_87	I7S K19E G202R N224D S144H K13Y Y240S G166C	1.512919
R6_88	I7S K19E G202R N224D S144H V140F G43E G217Y	1.576891
R6_89	I7S K19E G202R N224D S144H T178S A105C A8Y	1.724893
R6_90	I7S K19E G202R N224D S144H V234M Y143F W156G	1.766006
R6_91	I7S K19E G202R N224D S144H S90W D155L W156F	1.946453
R6_92	I7S K19E G202R N224D S144H A9I Q115V G150C	1.811797
R6_93	I7D K146E G202R N224D I102L E108P K206R A8P	1.845815

R6_94	I7D K146E G202R N224D I102L F38Y G80I D183V	1.790019
R6_95	I7D K146E G202R N224D I102L L253M D135A Y143D	1.579808
R6_96	I7D K146E G202R N224D I102L A130H I168L E91I	1.557445
R6_97	I7D K146E G202R N224D I102L K99M P76V I232Y	1.510566
R6_98	I7D K146E G202R N224D I102L V12C H84W V79M	1.20597
R6_99	I7D K146E G202R N224D I102L L170A F38A T53R	1.599371
R6_100	I7D K146E G202R N224D I102L V248K F120E D174M	1.764707
R6_101	I7D K146E G202R N224D I102L A97C I42L F49M	1.114438
R6_102	I7D K146E G202R N224D I102L V18E I42G I187V	0.122115
R6_103	I7Q K146E G202R N224D F229S P197W R188Q H84C V69N	1.290974
R6_104	I7Q K146E G202W N224D F229S P197W Q72V Y143L	2.543158
R6_105	I7Q K146E G202R N224D F229S P197W K13I D135C M62R	1.98117
R6_106	I7Q K146E G202R N224D F229S P197W L241F S111T R154W	1.556844
R6_107	I7Q K146E G202R N224D F229S P197W T61L L94H M207A	2.033685
R6_108	I7Q K146E G202R N224D F229S P197W W156A K37L V69M	1.90771
R6_109	I7Q K146E G202R N224D F229S P197W K147S I93V T78G	1.266248
R6_110	I7Q K146E G202R N224D F229S P197W L35A A105N F120Q	1.566663
R6_111	I7Q K146E G202R N224D F229S P197W D219T A9I L250S	1.185164
R6_112	I7Q K146E G202R N224D F229S P197W V157G A221C V33F	1.617854
R6_113	I7D K146E G202R N224D L47I S111G I232P S90H	1.689164
R6_114	I7D K146E G202R N224D L47I A105H M1K E251V	1.740965
R6_115	I7D K146E G202R N224P L47I A70Q V125L	0.432665
R6_116	I7D K146E G202R N224D L47I L176F A218Q G82S	2.0323
R6_117	I7D K146E G202R N224D L47I I168E Q72H D135V	0.979069
R6_118	I7D K146E G202R N224D L47I P110N I42T L237V	1.561373
R6_119	I7D K146E G202R N224D L47I S144K S111C I42Q	1.662601
R6_120	I7D K146E G202R N224D L47I T194P A105T L152G	1.636615
R6_121	I7D K146E G202R N224D L47I G164P I187V A117R	2.212662
R6_122	I7D K146E G202R N224D L47I K147V T195P A9F	0.756254

References

1. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

2. Khersonsky, O.; Röthlisberger, D.; Dym, O.; Albeck, S.; Jackson, C. J.; Baker, D.; Tawfik, D. S., Evolutionary Optimization of Computationally Designed Enzymes: Kemp Eliminases of the KE07 Series. *J. Mol.*

Biol. **2010**, *396*, 1025-1042.
3. Hong, N.-S.; Petrović, D.; Lee, R.; Gryn'Ova, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C.-Y.; Mabbitt, P. D.; Zhang, W.; Altamore, T.; Easton, C.; Coote, M. L.; Kamerlin, S. C. L.; Jackson, C. J., The evolution of multiple active site configurations in a designed enzyme. *Nat. Commun.* **2018**, *9*.
4. Lu, T.; Chen, F., Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580-592.
5. Nivón, L. G.; Moretti, R.; Baker, D., A Pareto-Optimal Refinement Method for Protein Design Scaffolds. *PLoS ONE* **2013**, *8*, e59004.
6. Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; Dechancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D., Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190-195.
7. Schrodinger, LLC, In; 2015.
8. D.A. Case, H. M. A., K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman Amber 2021.
9. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.
10. Wang, N. X.; Wilson, A. K., The behavior of density functionals with respect to basis set. I. The correlation consistent basis sets. *The Journal of Chemical Physics* **2004**, *121*, 7632-7646.
11. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling* **2006**, *25*, 247-260.
12. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623-1641.
13. Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A., PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **2007**, *35*, W522-W525.
14. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525-537.
15. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23*, 327-341.
16. Loncharich, R. J.; Brooks, B. R.; Pastor, R. W., Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N′-methylamide. *Biopolymers* **1992**, *32*, 523-535.
17. Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*, 3684-3690.
18. Bradshaw, R. T.; Dziedzic, J.; Skylaris, C.-K.; Essex, J. W., The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields? *J. Chem. Inf. Model.* **2020**, *60*, 3131-3144.