Supporting Information

Comprehensive Insights into the Catalytic Mechanism of Middle East Respiratory Syndrome 3C-Like Protease and Severe Acute Respiratory Syndrome 3C-Like Protease

*Hao Wang[1,2], Shuai He[1,2], Weilong Deng[1,2], Ying Zhang[3], Guobang Li[1,2], Jixue Sun[1], Wei Zhao[1,2], Yu Guo[1,2], Zheng Yin[1,2,4], Dongmei Li[1] *, and Luqing Shang[1,2]**

1. State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and KLMDASR of Tianjin, Nankai University, No.38 Tongyan Road, Haihe Education Park, Tianjin 300350, People's Republic of China

2. Drug Discovery Center for Infectious Disease, Nankai University, 38 Tongyan Road, Haihe Education Park, Tianjin 300350, People's Republic of China

3. Laboratory of Structural Biological & Ministry of Education and Laboratory of Protein Science, School of Medicine and Life Sciences, Tsinghua University, Beijing 100084, People's Republic of China

4. Center of Basic Molecular Science, Department of Chemistry, Tsinghua University, Beijing 100084, People's Republic of China

* E-mail: dongmeili@nankai.edu.cn, shanglq@nankai.edu.cn

Contents

Supporting figures.....3

Supporting tables 16

Supporting chemical schemes.....23

Chemical synthesis.....24

NMR Figures.....34

Ab initio model scheme.....38

Supporting figures in MD simulation 39

References.....51

Supporting figures

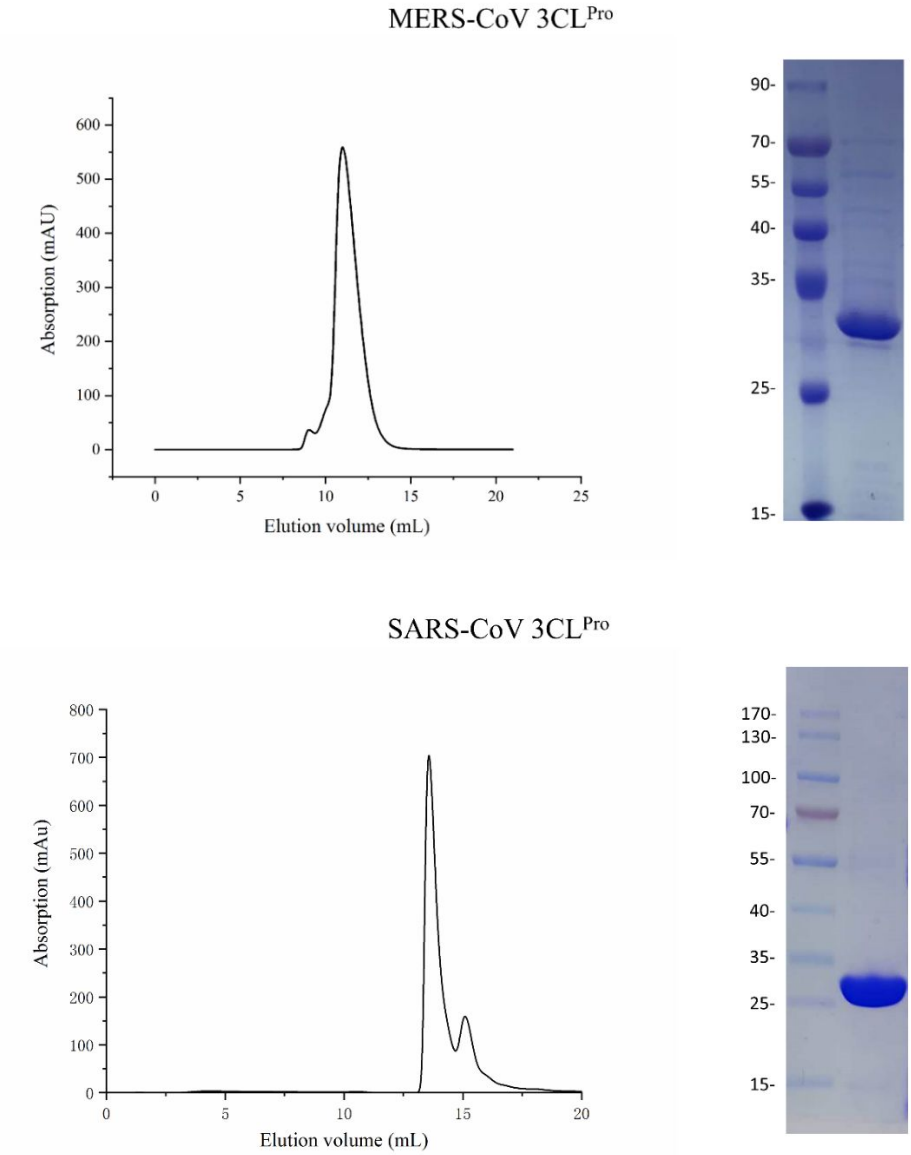


Figure S1. Size exclusion chromatography of MERS-CoV $3CL^{Pro}$ and SARS-CoV $3CL^{Pro}$. MERS-CoV $3CL^{Pro}$ is purified by SEC with retention volume of 11.5 mL by the superdex-75 gel Filtration columns, while SARS-CoV $3CL^{Pro}$ is purified by SEC

with retention volume of 13.7 mL by the superdex-200 gel Filtration columns. The sample obtained following purification are analyzed by SDS-PAGE and stained with Coomassie brilliant blue R-250.

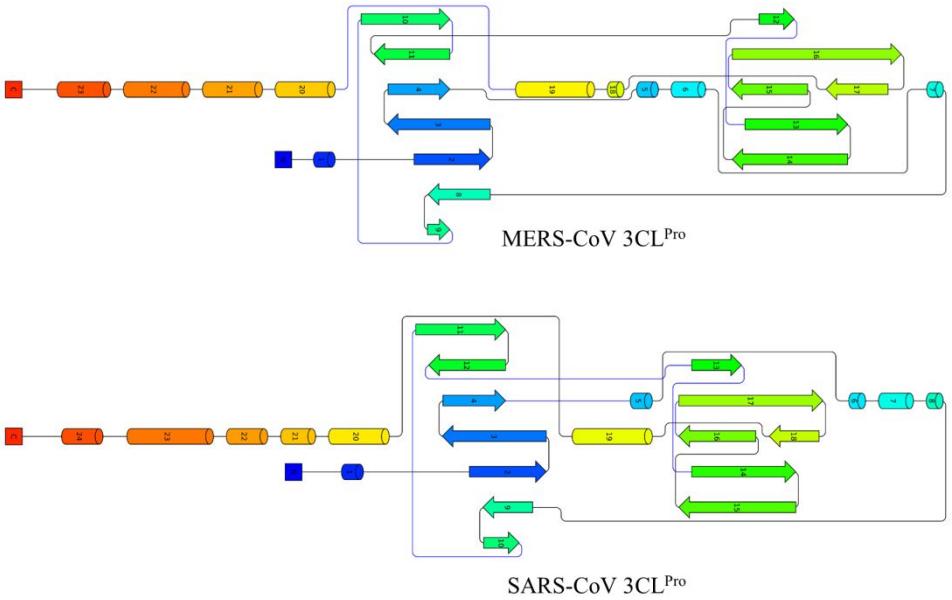


Figure S2. Topological research on MERS-CoV 3CL^{Pro} (left) and SARS-CoV 3CL^{Pro} (right). The figure is displayed via processed by Pro-origami website (http://munk.cis.unimelb.edu.au/pro-origami/porun.shtml). [1-4] The α-helix is shown as barrel and the β-sheet is exhibited as arrow. Meanwhile, α-helix and β-sheet are connected with loop, which is shown as string.

Protease	**MERS-CoV 3CL^{Pro}**			
Title	Total	Domain I	Domain II	Domain III
HCoV-HKU1 3CL^{Pro}	0.667	0.433	0.438	0.463
BCoV-HKU4 3CL^{Pro}	0.741	0.516	0.352	0.435
HCoV-229E 3CL^{Pro}	0.590	0.316	0.498	0.555
HCoV-NL63 3CL^{Pro}	0.556	0.333	0.443	0.542

Protease	**SARS-CoV 3CL^{Pro}**			
Title	Total	Domain I	Domain II	Domain III
HCoV-HKU1 3CL^{Pro}	0.836	0.452	0.458	0.898
BCoV-HKU4 3CL^{Pro}	0.848	0.444	0.364	0.841
HCoV-229E 3CL^{Pro}	0.774	0.342	0.403	0.988
HCoV-NL63 3CL^{Pro}	0.998	0.442	0.409	1.068

Figure S3. The alignment of four coronaviruses 3C like protease with MERS-CoV

3CL^{Pro} and SARS-CoV 3CL^{Pro}, respectively.

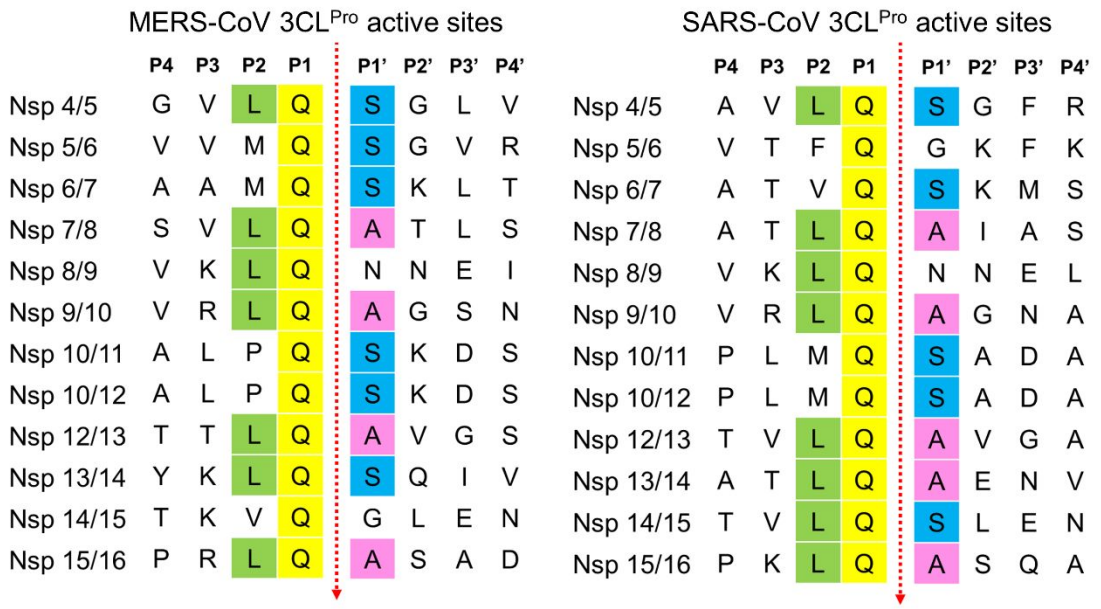


Figure S4. Bioinformatics analysis on the native substrate active sites of MERS-CoV 3CL^{Pro} (left) and SARS-CoV 3CL^{Pro} (right). Reported protease favoring residues of substrate are emphasized by color background.

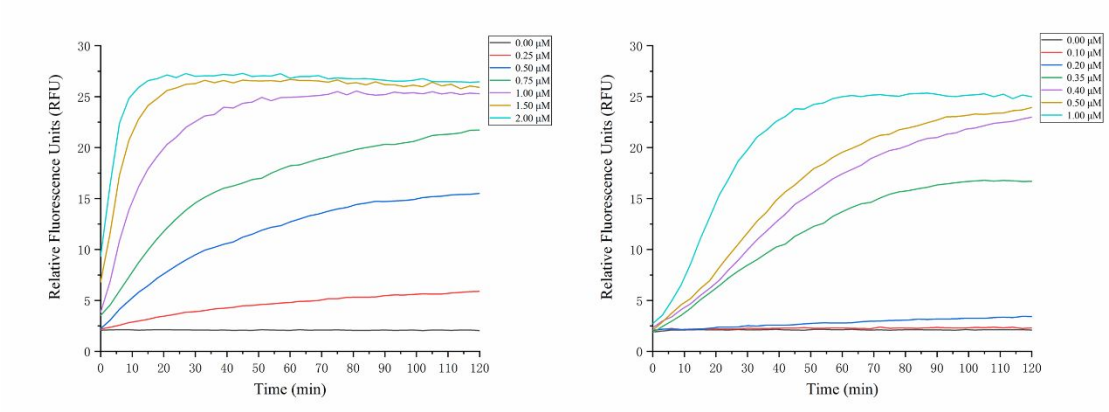


Figure S5. Substrate cleavage capability of MERS-CoV 3CL^{Pro} (left) and SARS-CoV 3CL^{Pro} (right) in dose dependent and time dependent manner.

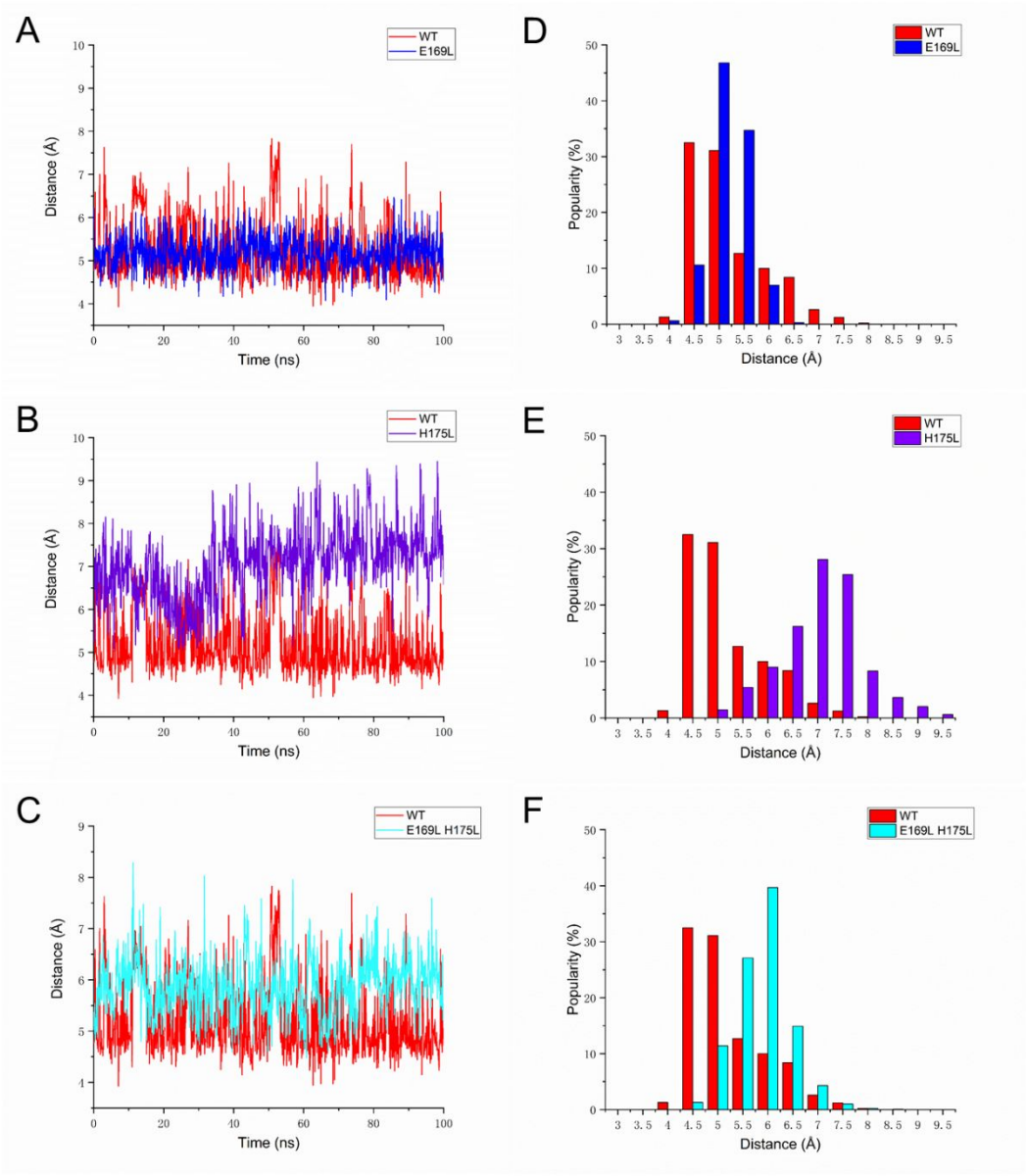


Figure S6. MD simulation calculation on the barycenter distance between 169th site residue side-chain terminal and 175th site residue side-chain terminal. (A) The time evolution of the distance in WT and mutation E169L. (B) The time evolution of the distance in WT and mutation H175L. (C) The time evolution of the distance in WT and mutation E169L H175L. (D) The distribution of the distance of (A). (E) The distribution of the distance of (B). (F) The distribution of the distance of (C).

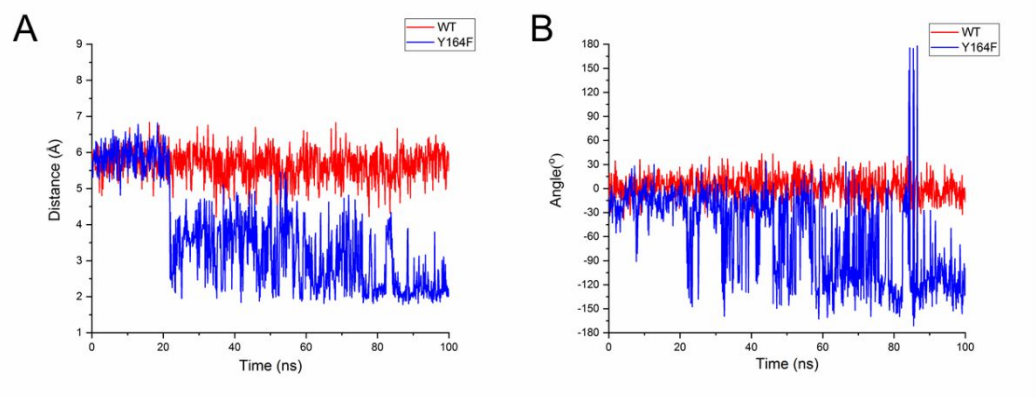


Figure S7. MD simulation calculation on the wild type MERS-CoV $3CL^{Pro}$ (in red) and mutation Y164F (in blue). (A) The time evolution of distance between H166 and substrate glutamine in WT (red) and mutation Y164F (blue). (B) The time evolution of dihedral of H166 (CD2-CG-CB-H) in WT (red) and mutation Y164F (blue).

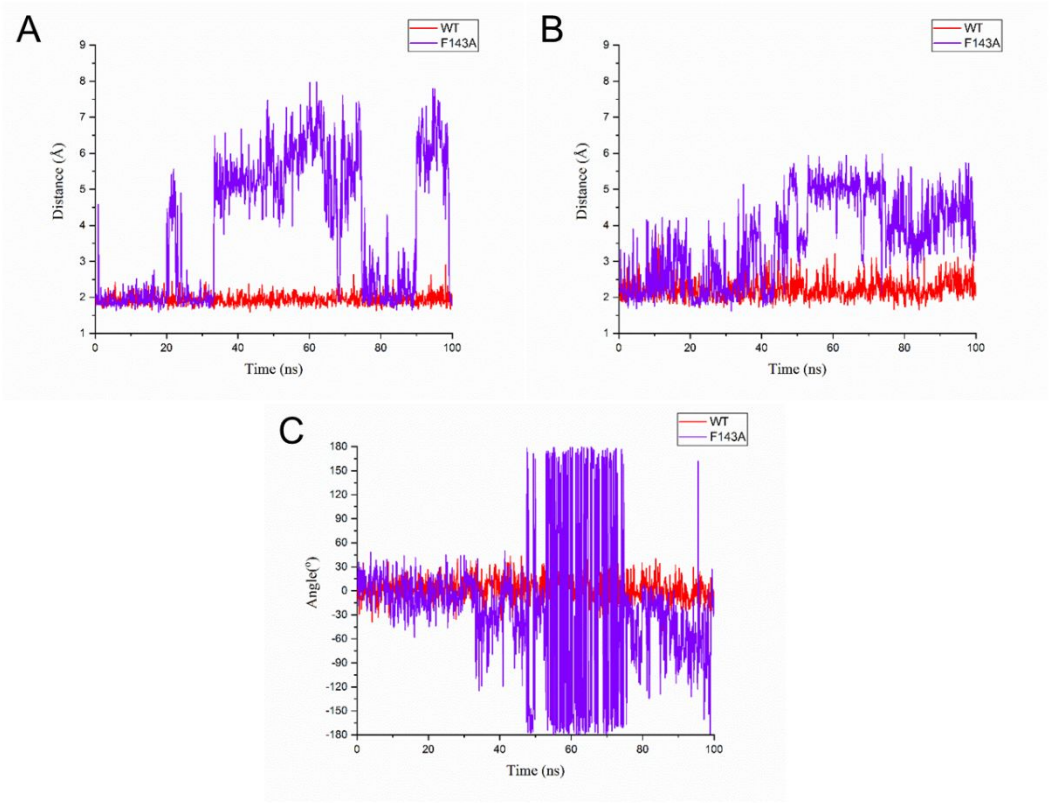


Figure S8. MD simulation calculation on the WT MERS-CoV $3CL^{Pro}$ (red) and mutation F143A (violet). (A) The time evolution of distance between H166 and substrate glutamine in WT (red) and mutation F143A (violet). (B) The time evolution of distance between H166 and Y164 in WT (red) and mutation F143A (violet). (C) The time evolution of dihedral of H166 (CD2-CG-CB-H) in WT (red) and mutation F143A (violet).

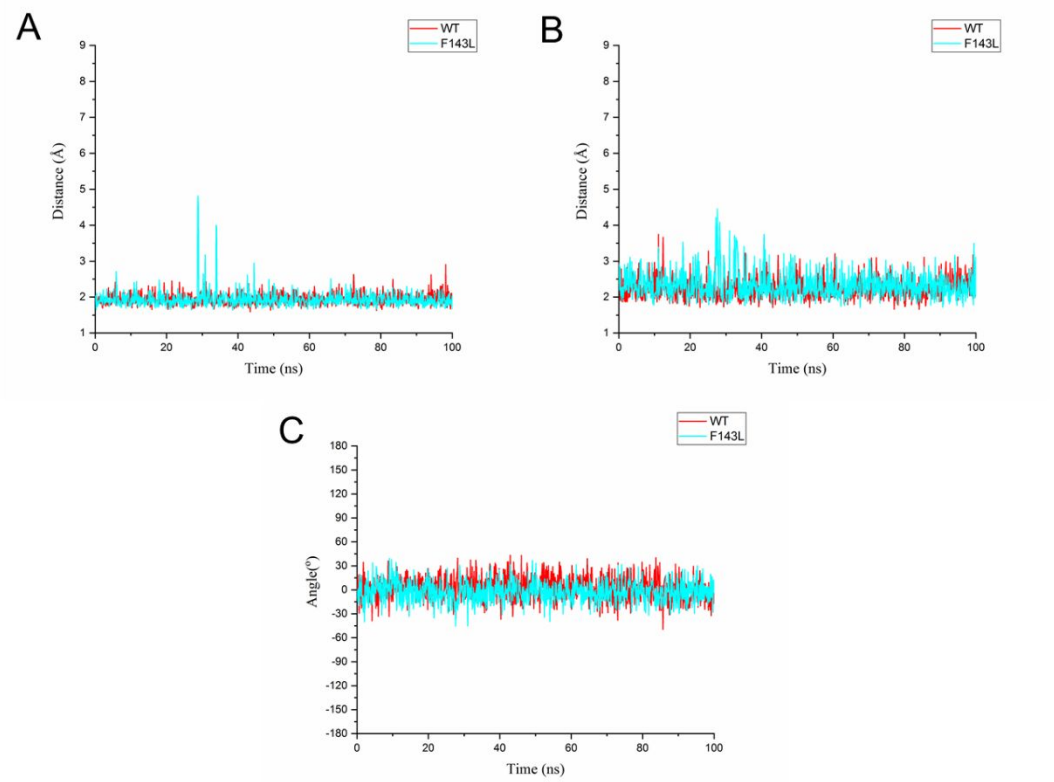


Figure S9. MD simulation calculation on the WT MERS-CoV $3CL^{Pro}$ (red) and mutation F143L (cyan). (A) The time evolution of distance between H166 and substrate glutamine in WT (red) and mutation F143L (cyan). (B) The time evolution of distance between H166 and Y164 in WT (red) and mutation F143L (cyan). (C) The time evolution of dihedral of H166 (CD2-CG-CB-H) in WT (red) and mutation F143L (cyan).

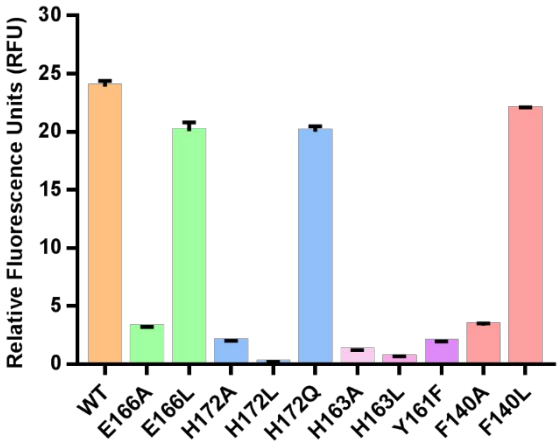


Figure S10. Activity evaluation of related representative mutations in SARS-CoV $3CL^{Pro}$. The data presented are mean values from experiments in triplicate and the error

bars indicate standard deviations.

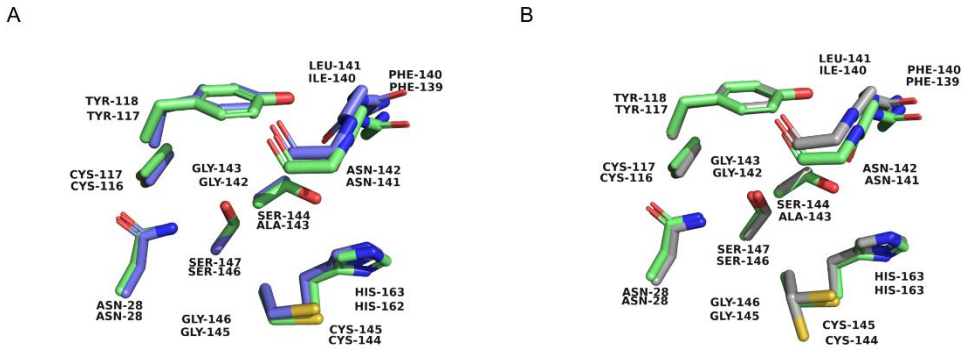


Figure S11. Superimposition of SARS-CoV $3CL^{Pro}$ (green) with HCoV-229E $3CL^{Pro}$ (purple) (A) and HCoV-NL63 $3CL^{Pro}$ (gray) (B). Owing to holding similar surrounding of the S144 in SARS-CoV $3CL^{Pro}$ with A143 in HCoV-229E $3CL^{Pro}$ and HCoV-NL63 $3CL^{Pro}$, SARS-CoV $3CL^{Pro}$ was utilized to verify the significant roles of S144.

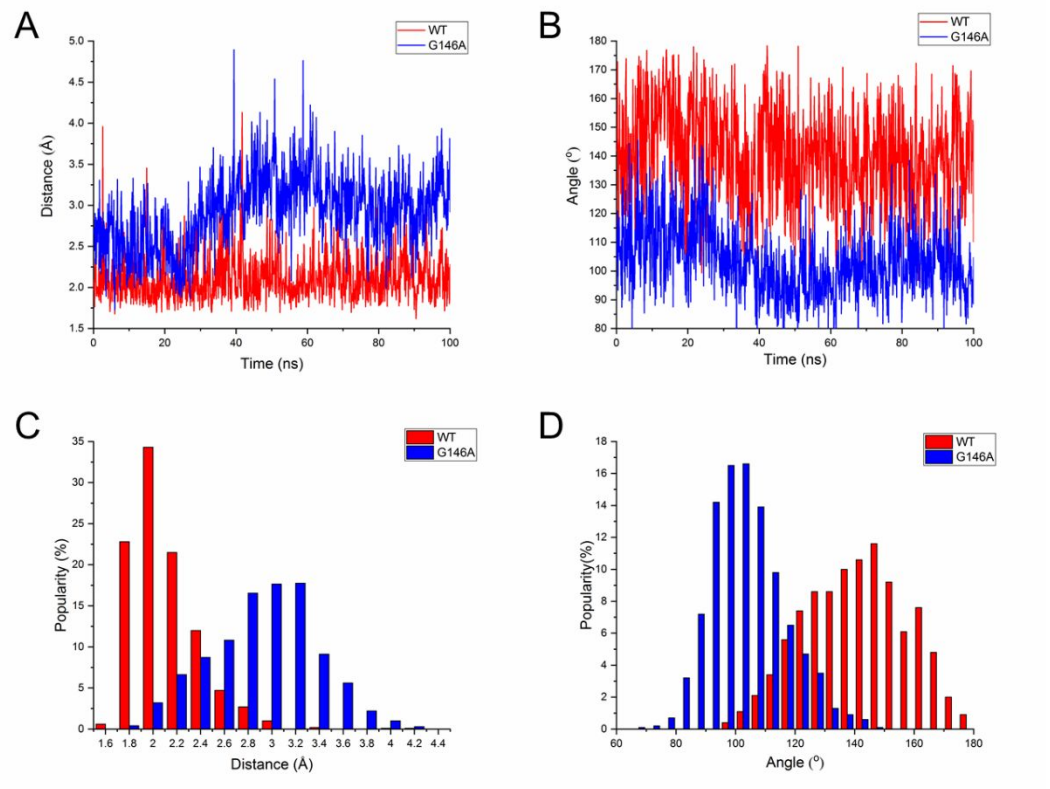


Figure S12. MD simulation calculation on the wild type MERS-CoV $3CL^{Pro}$ (red), mutation G146A (blue). (A) The time evolution of distance between 146 site residue and substrate glutamine. (B) The time evolution of angle between 146 site residue and substrate glutamine (N-H-O). (C) The distribution of the distance of (A). (D) The

distribution of the angle of (B).

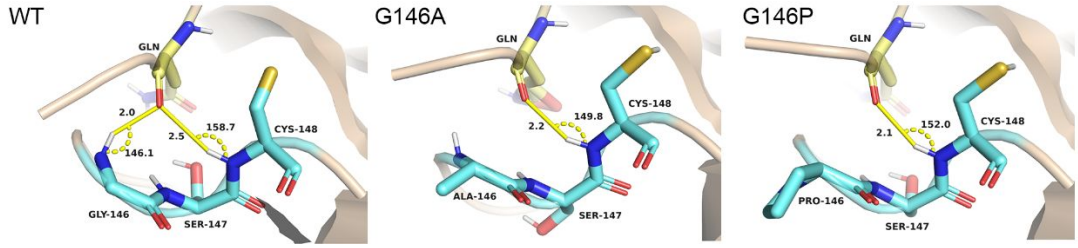


Figure S13. The average structure of the WT and mutants extracted from last 10 ns trajectory following MD simulation.

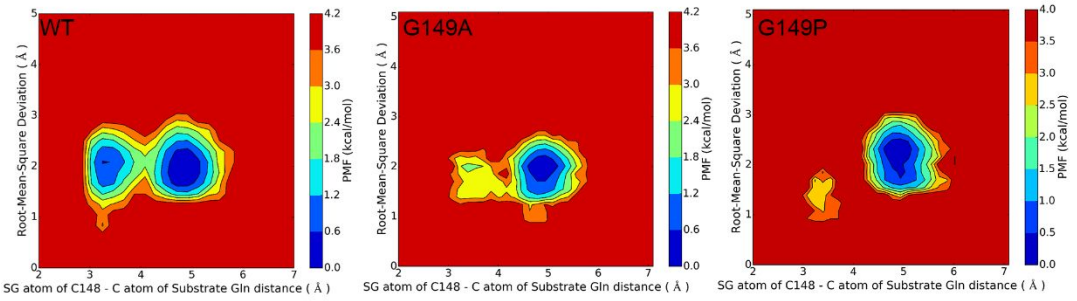


Figure S14. PMF calculation for the SG atom of C148 - substrate glutamine carbonyl group distance vs the RMSD of total system in WT MERS-CoV 3CL^{Pro} (left), G149A mutant (center), and G149P mutant (right).

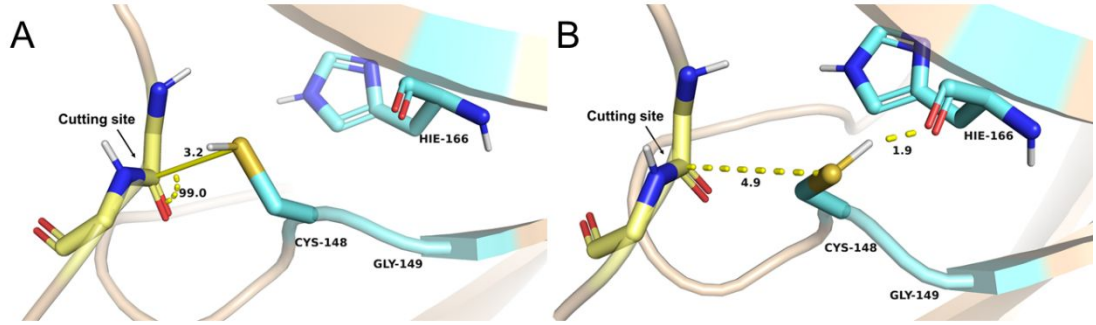


Figure S15. The average structure of attacking state (A) and resting-state (B) of MERS-CoV 3CL^{Pro} to manifest the detail features of the two-state.

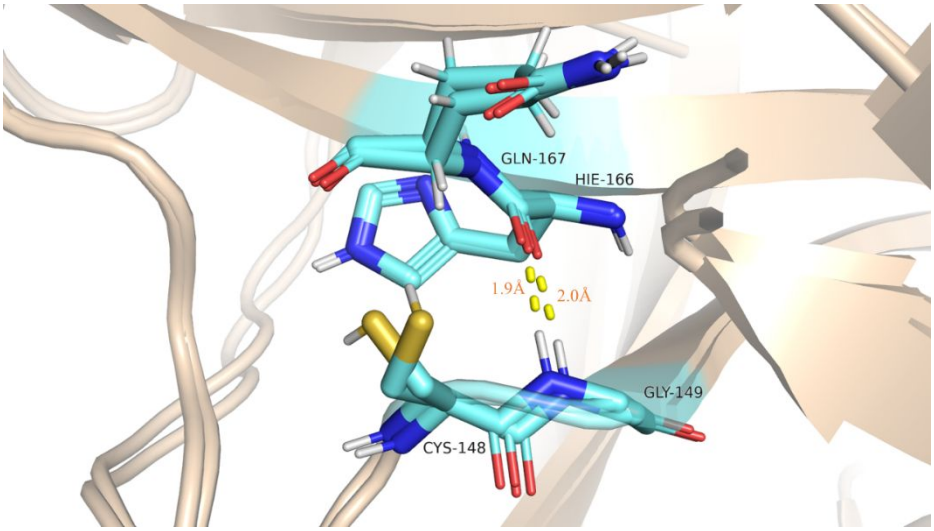


Figure S16. Superimposition of two typical conformation (attacking state and resting state) of MERS-CoV 3CL^{Pro} in MD simulation.

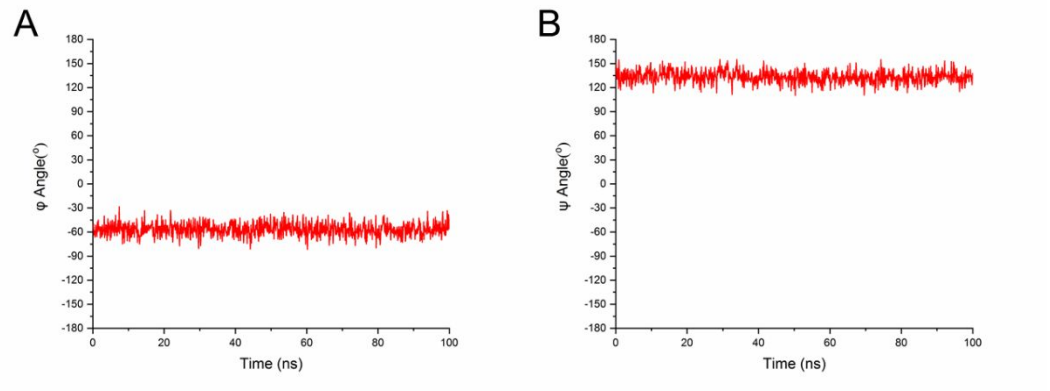


Figure S17. MD simulation calculation on the wild type MERS-CoV 3CL^{Pro}. (A) The time evolution of φ dihedral of C148. (B) The time evolution of ψ dihedral of C148.

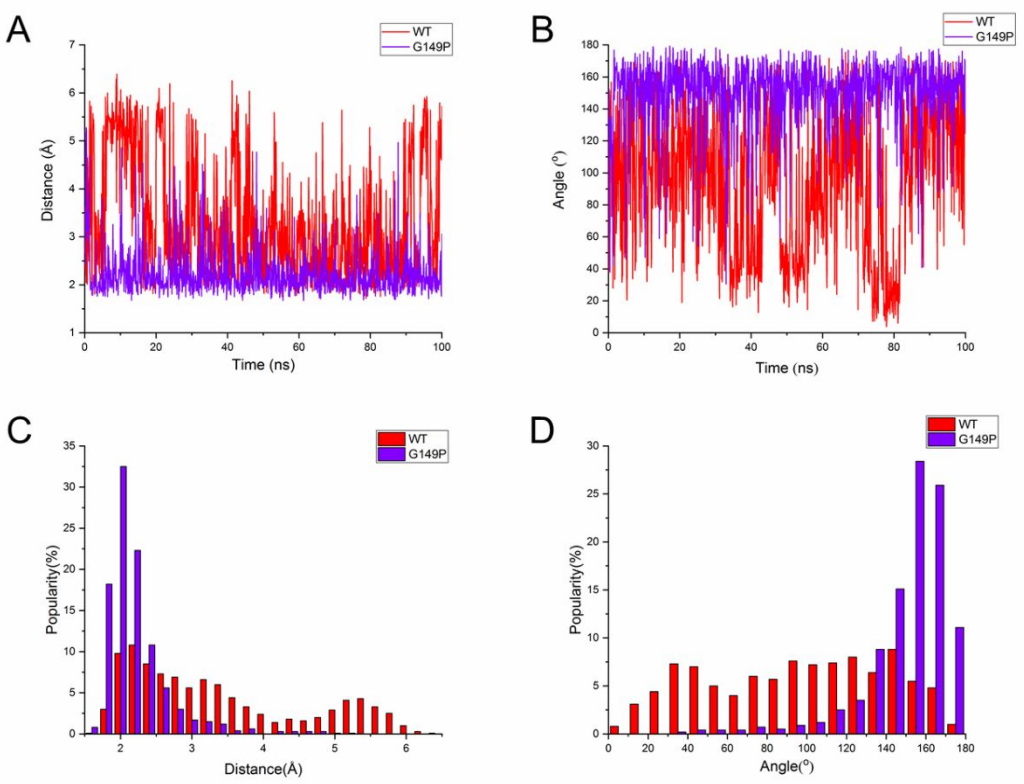


Figure S18. MD simulation calculation on the WT MERS-CoV 3CL^{Pro} (red), mutation G149P (purple). (A) The time evolution of distance between C148 and H166, which indicated C148 formed compact connection to H166 in mutation G149P. (B) The time evolution of angle between C148 and H166 (S-H-O), which indicated stable the hydrogen interaction between C148 and H166. (C) The distribution of the distance in (A). (D) The distribution of the angle in (B).

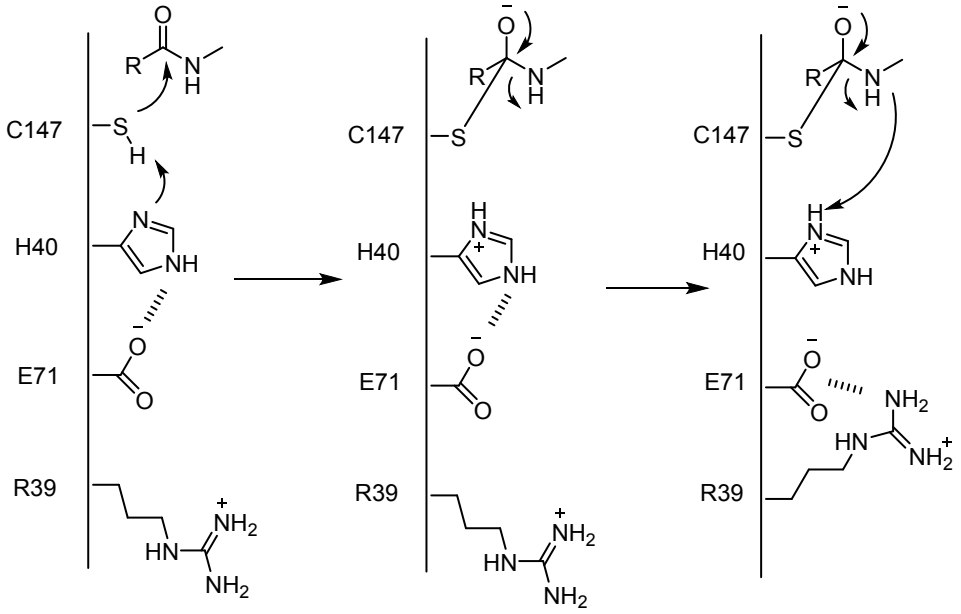


Figure S19. Molecular catalytic mechanism of EV71 $3C^{Pro}$.

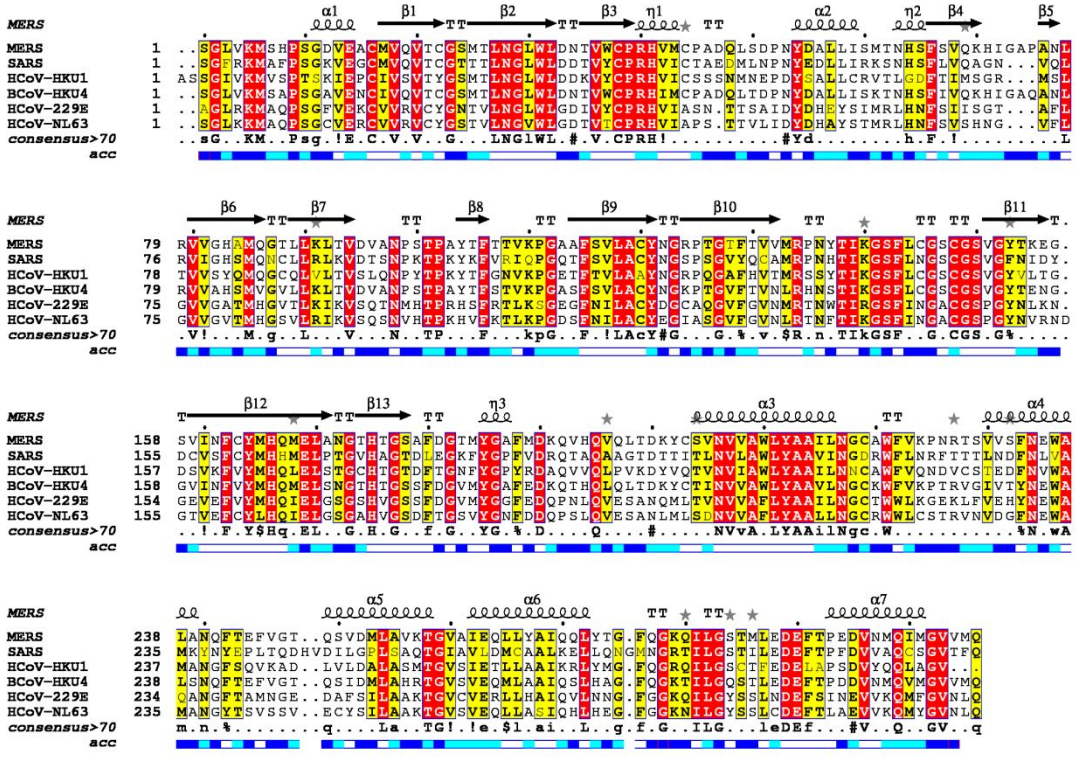


Figure S20. Sequence alignment of six common coronavirus $3CL^{Pro}$ (MERS-CoV $3CL^{Pro}$, SARS-CoV $3CL^{Pro}$, HCoV-HKU1 $3CL^{Pro}$, BCoV-HKU4 $3CL^{Pro}$, HCoV-229E $3CL^{Pro}$, HCoV-NL63 $3CL^{Pro}$). High conserved residues were shown in red background, α-helices in helix, β-sheets in arrow and turn in T symbol.

	Turn					Turn			
Position	**38**	**39**	**40**	**41**	**189**	**190**	**191**		**195**
MERS-CoV 3CL^{Pro}	C	P	R	H	M	D	K		Q
SARS-CoV 3CL^{Pro}	C	P	R	H	V	D	R		Q
HCoV-HKU1 3CL^{Pro}	C	P	R	H	R	D	A		Q
BCoV-HKU1 3CL^{Pro}	C	P	R	H	E	D	K		Q
HCoV-NL63 3CL^{Pro}	C	P	R	H	E	D	Q		Q
HCoV-229E 3CL^{Pro}	C	P	R	H	E	D	Q		Q

Figure S21. Conserved analysis on the sequence of six common coronavirus 3CL^{Pro}. The significant residues of MERS-CoV 3CL^{Pro} and SARS-CoV 3CL^{Pro} are highlighted.

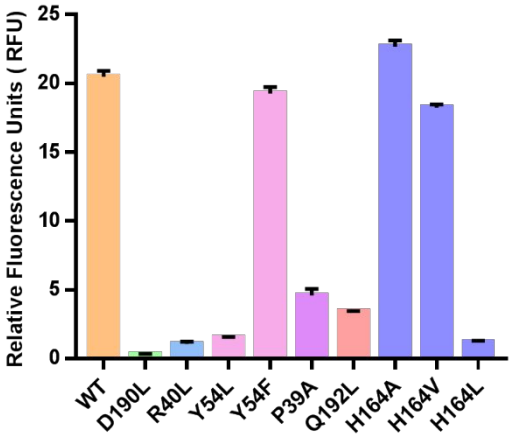


Figure S22. Enzymatic activity analysis on related representative mutations in SARS-CoV 3CL^{Pro}. The data presented are mean values from experiments in triplicate and the error bars indicate standard deviations.

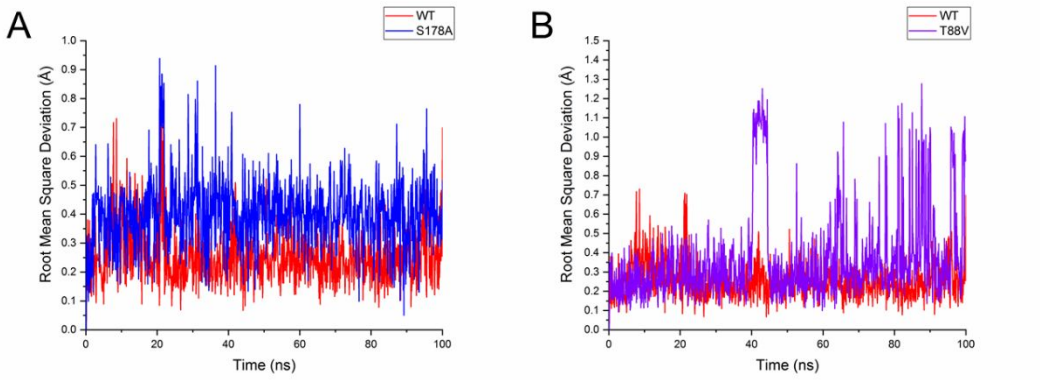


Figure S23. MD simulation calculation on the WT MERS-CoV 3CL^{Pro} (red), mutation

S178A (blue) and mutation T88A (violet). (A) The time evolution of RMSD of Q167. (B) The time evolution of RMSD of Q167.

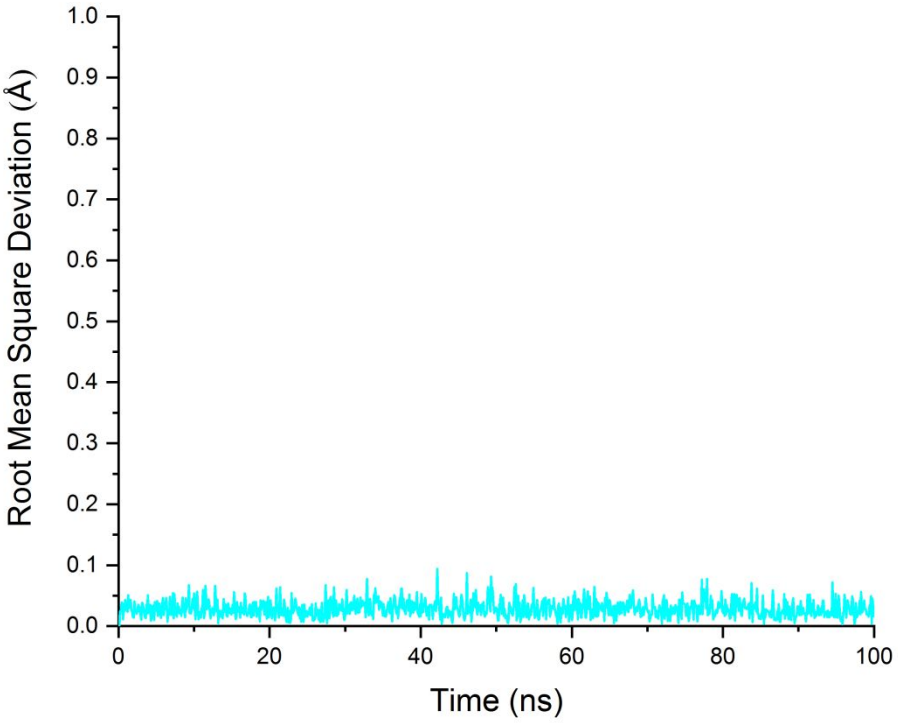


Figure S24. The time evolution of RMSD of Q167 in multipoints mutant T88C/S178T (cyan).

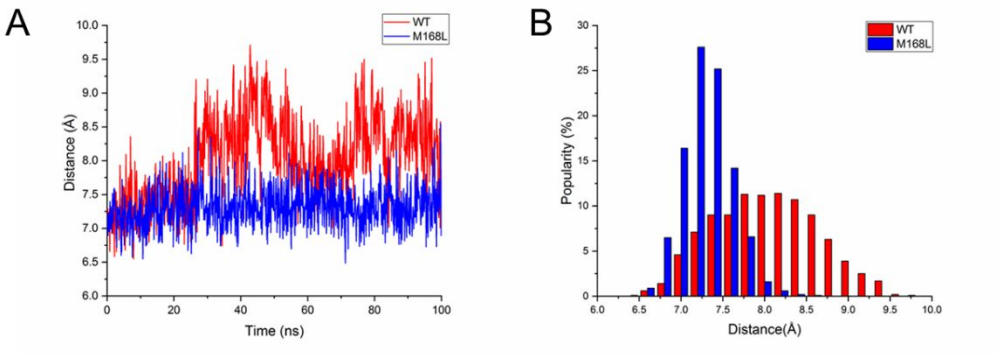


Figure S25. MD simulation calculation on the wild type MERS-CoV $3CL^{Pro}$ (red), mutation M168L (blue). (A) The time evolution of the distance between D190 and protonated histidine. (B) The relative distribution of the distance in (A).

	K_m (μM)	k_{cat}(min^{-1})	k_{cat}/K_m(mM^{-1}min^{-1})
SARS-CoV 3CL^{Pro}	6.518±1.048	0.275±0.016	42.25
M165L	7.186±1.038	0.498±0.025	69.30

Figure S26. Kinetic parameters of the WT and mutation M165L of SARS-CoV 3CL^{Pro}.

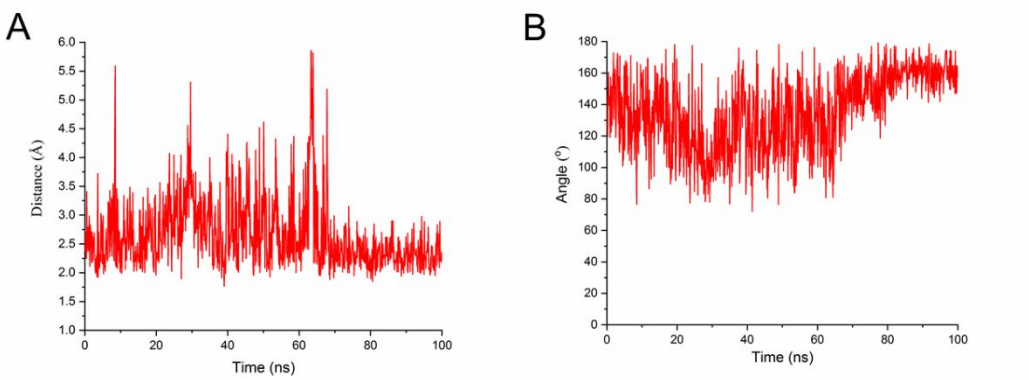


Figure S27. MD simulation calculation on the WT MERS-CoV 3CL^{Pro} binding to the compound **12b**. (A) The time evolution of the distance between Q195 and N atom of the pyridine ring. (B) The time evolution of the angle originated from N (main chain N atom in Q195)-H (main amide bond H atom in Q195)-N (N atom in the pyridine ring of the compound **12b**).

Supporting tables

Table S1. The designed primers for construction of plasmid.

		Sequence
MERS-CoV 3CL^{Pro}	Forward	5’-CG**GGATCC**AGCGGTTTGGTGAAAATGTC-3’
	Reverse	5’-CCG**CTCGAG**CTGCATAACCACACCCATAAT -3’
SARS-CoV 3CL^{Pro}	Forward	5’-CG**GGATCC**AGTGGTTTCAGGAAAATGGC-3’
	Reverse	5’- CCG**CTCGAG**TTGGAAGGTAACACCAGAGC-3’

Table S2. The designed primers for proteion mutagenesis.

Species	Mutation		Sequence (5'-3')
MERS-CoV	C148A	Forward	TTTCTGTGTGGTTCTGCTGGTAGTGTTG
		Reverse	GCAGAACCACACAGAAAGGAACCCTTA
	C148S	Forward	TCCTTTCTGTGTGGTTCTTCTGGTAGTGTTG
		Reverse	GAAGAACCACACAGAAAGGAACCCTTAATTG
	C145A	Forward	ATTAAGGGTTCCTTTCTGGCTGGTTCTTGTGG
		Reverse	GCCAGAAAGGAACCCTTAATTGTGTAGTTAGGGCG
	C145S	Forward	ATTAAGGGTTCCTTTCTGTCTGGTTCTTGTG
		Reverse	GACAGAAAGGAACCCTTAATTGTGTAGTTAGG
	H41A	Forward	ACAGTCTGGTGCCCACGAGCCGTAATGTGCC
		Reverse	GCTCGTGGGCACCAGACTGTGTTGTCAAGCC
	H41L	Forward	GTCTGGTGCCCACGACTTGTAATGTGCCC
		Reverse	AAGTCGTGGGCACCAGACTGTGTTGTCAAGC
	H166A	Forward	CAATTTCTGTTACATGGCTCAAATGGAAC
		Reverse	GCCATGTAACAGAAATTGATCACACTACC
	H166L	Forward	ATCAATTTCTGTTACATGCTTCAAATGGAAC
		Reverse	AGCATGTAACAGAAATTGATCACACTACCC
	H166Q	Forward	AATTTCTGTTACATGCAGCAAATGGAAC
		Reverse	CTGCATGTAACAGAAATTGATCACACTACCC
	H166N	Forward	ATCAATTTCTGTTACATGAATCAAATGGAAC
		Reverse	TCATGTAACAGAAATTGATCACACTACCCTC
	Y164A	Forward	AGTGTGATCAATTTCTGTGCCATGCATCAAATG
		Reverse	GCACAGAAATTGATCACACTACCCTCCTTG
	Y164F	Forward	GTGTGATCAATTTCTGTTTCATGCATCAAATGG
		Reverse	AAACAGAAATTGATCACACTACCCTCCTTG
	Y164R	Forward	GTGTGATCAATTTCTGTCGCATGCATCAAATGG

		Reverse	CGACAGAAATTGATCACACTACCCTCCTTG
	Y164K	Forward	GTGTGATCAATTTCTGTAAAATGCATCAAATGG
		Reverse	TTTACAGAAATTGATCACACTACCCTCCTTG
	Y164E	Forward	GTGTGATCAATTTCTGTGAGATGCATCAAATGG
		Reverse	CTCACAGAAATTGATCACACTACCCTCCTTGGTGTAA
	Y164D	Forward	AGTGTGATCAATTTCTGTGACATGCATCAAATG
		Reverse	CACAGAAATTGATCACACTACCCTCCTTGG
	F143Y	Forward	CACAATTAAGGGTTCCTATCTGTGTGGTTC
		Reverse	TAGGAACCCTTAATTGTGTAGTTAGGGCG
	F143W	Forward	CACAATTAAGGGTTCCTGGCTGTGTGGTTC
		Reverse	CCAGGAACCCTTAATTGTGTAGTTAGGGCG
	F143A	Forward	CACAATTAAGGGTTCCGCTCTGTGTGGTTC
		Reverse	GCGGAACCCTTAATTGTGTAGTTAGGGCG
	F143V	Forward	TACACAATTAAGGGTTCCGTTCTGTGTGG
		Reverse	CGGAACCCTTAATTGTGTAGTTAGGGCGCAT
	F143L	Forward	AATTAAGGGTTCCTTGCTGTGTGGTTC
		Reverse	CAAGGAACCCTTAATTGTGTAGTTAGGG
	F143M	Forward	ACAATTAAGGGTTCCATGCTGTGTGGTTCTTG
		Reverse	CATGGAACCCTTAATTGTGTAGTTAGGGCG
	E169A	Forward	TGTTACATGCATCAAATGGCACTTGCTAATGG
		Reverse	GCCATTTGATGCATGTAACAGAAATTGATCAC
	E169V	Forward	TGTTACATGCATCAAATGGTACTTGCTAATGGTACAC
		Reverse	ACCATTTGATGCATGTAACAGAAATTGATCAC
	E169L	Forward	TGTTACATGCATCAAATGCTACTTGCTAATG
		Reverse	AGCATTTGATGCATGTAACAGAAATTGATC
	E169M	Forward	TACATGCATCAAATGATGCTTGCTAATGGTACAC
		Reverse	CATCATTTGATGCATGTAACAGAAATTGATCAC
	E169D	Forward	ACATGCATCAAATGGACCTTGCTAATGGTACAC
		Reverse	GTCCATTTGATGCATGTAACAGAAATTGATCAC

	E169Q	Forward	GTTACATGCATCAAATGCAACTTGCTAATGG
		Reverse	GCATTTGATGCATGTAACAGAAATTGATCACAC
	H175A	Forward	AACTTGCTAATGGTACAGCTACCGGTTCAG
		Reverse	GCTGTACCATTAGCAAGTTCCATTTGATGC
	H175L	Forward	AACTTGCTAATGGTACACTTACCGGTTCAG
		Reverse	AGTGTACCATTAGCAAGTTCCATTTGATGC
	H175Q	Forward	AACTTGCTAATGGTACACAGACCGGTTCAG
		Reverse	CTGTGTACCATTAGCAAGTTCCATTTGATGC
	H175E	Forward	CTTGCTAATGGTACAGAGACCGGTTCAG
		Reverse	CTCTGTACCATTAGCAAGTTCCATTTGATGCATG
	H175N	Forward	GAACTTGCTAATGGTACAAATACCGGTTCAGC
		Reverse	TTGTACCATTAGCAAGTTCCATTTGATGCATGT
	T174V	Forward	GGAACTTGCTAATGGTGTACATACCGGTTCAG
		Reverse	ACACCATTAGCAAGTTCCATTTGATGCATGT
	G146A	Forward	AGGGTTCCTTTCTGTGTGCTTCTTGTGGTAG
		Reverse	GCACACAGAAAGGAACCCTTAATTGTGTAG
	G146P	Forward	AGGGTTCCTTTCTGTGTCCTTCTTGTGGTAG
		Reverse	GGACACAGAAAGGAACCCTTAATTGTGTAG
	S147A	Forward	GTTCCTTTCTGTGTGGTGCTTGTGGTAGTG
		Reverse	GCACCACACAGAAAGGAACCCTTAATTGTG
	G149A	Forward	CTGTGTGGTTCTTGTGCTAGTGTTGGTTA
		Reverse	GCACAAGAACCACACAGAAAGGAACCC
	G149P	Forward	CTGTGTGGTTCTTGTCCTAGTGTTGGTTA
		Reverse	GGACAAGAACCACACAGAAAGGAACCC
	S150A	Forward	TGTGGTTCTTGTGGTGCTGTTGGTTAC
		Reverse	GCACCACAAGAACCACACAGAAAGGAACC
	N28A	Forward	GCGGTAGCATGACTCTTGCTGGTCTTTGG
		Reverse	GCAAGAGTCATGCTACCGCAGGTAACCTGAAC
	N28L	Forward	CGGTAGCATGACTCTTCTTGGTCTTTGG

		Reverse	AGAAGAGTCATGCTACCGCAGGTAACCTGAAC
	N28D	Forward	GCGGTAGCATGACTCTTGATGGTCTTTGG
		Reverse	CAAGAGTCATGCTACCGCAGGTAACCTGAAC
	N28Q	Forward	GGTAGCATGACTCTTCAAGGTCTTTGGC
		Reverse	TTGAAGAGTCATGCTACCGCAGGTAACCTGAA
	N28H	Forward	GCGGTAGCATGACTCTTCATGGTCTTTGG
		Reverse	GAAGAGTCATGCTACCGCAGGTAACCTGAAC
	D190A	Forward	GTATGGTGCCTTTATGGCTAAACAAGTG
		Reverse	GCCATAAAGGCACCATACATAGTACCATC
	D190L	Forward	ATGTATGGTGCCTTTATGCTTAAACAAGTGC
		Reverse	AGCATAAAGGCACCATACATAGTACCATCAAAT
	D190E	Forward	TATGGTGCCTTTATGGAGAAACAAGTGCAC
		Reverse	CTCCATAAAGGCACCATACATAGTACCATC
	D190N	Forward	GTATGGTGCCTTTATGAATAAACAAGTG
		Reverse	TCATAAAGGCACCATACATAGTACCATC
	D190H	Forward	GTATGGTGCCTTTATGCATAAACAAGTG
		Reverse	GCATAAAGGCACCATACATAGTACCATC
	R40A	Forward	ACAGTCTGGTGCCCAGCACACGTAATGTGCCC
		Reverse	TGCTGGGCACCAGACTGTGTTGTCAAGCCA
	R40L	Forward	CAGTCTGGTGCCCACTTCACGTAATGTGCC
		Reverse	AAGTGGGCACCAGACTGTGTTGTCAAGCCA
	R40K	Forward	CAGTCTGGTGCCCAAAACACGTAATGTGCC
		Reverse	TTTTGGGCACCAGACTGTGTTGTCAAGCCA
	R40H	Forward	CAGTCTGGTGCCCACACCACGTAATGTGC
		Reverse	GTGTGGGCACCAGACTGTGTTGTCAAGC
	R40Y	Forward	CAGTCTGGTGCCCATACCACGTAATGTGCCC
		Reverse	GTATGGGCACCAGACTGTGTTGTCAAGCCA
	M85A	Forward	GTTGTTGGTCATGCCGCGCAAGGCACTC
		Reverse	GCGGCATGACCAACAACACGCAAGTTTGCTG

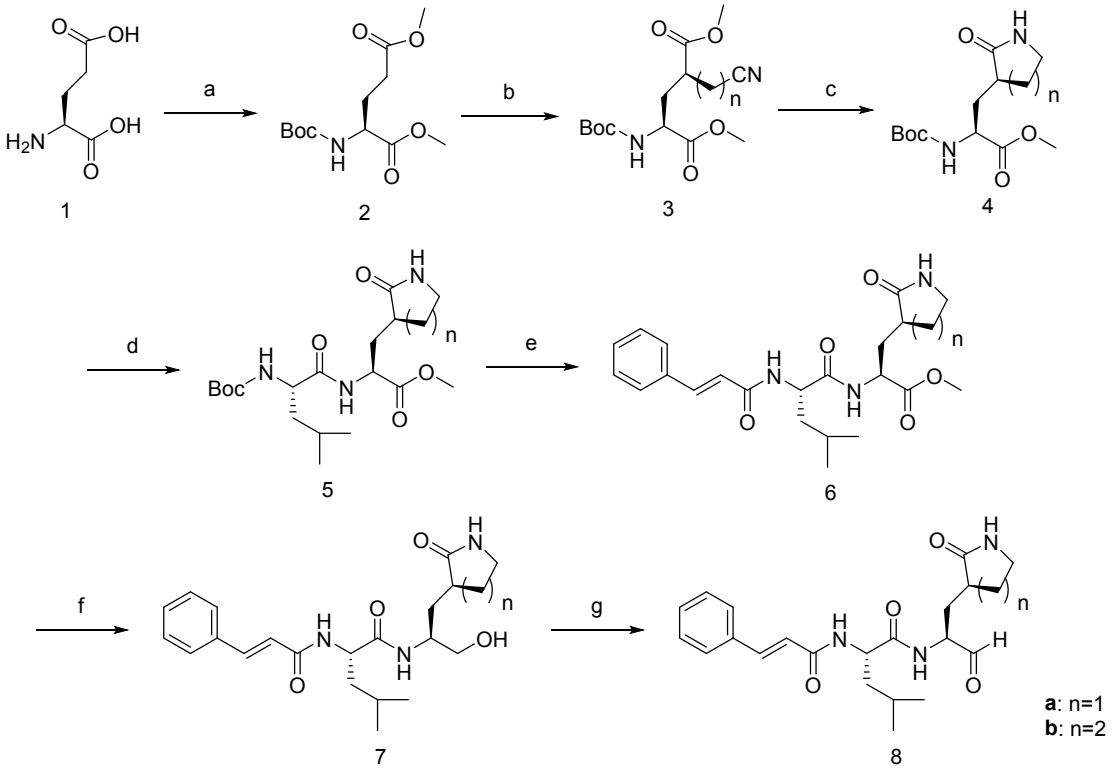
	M85L	Forward	CGTGTTGTTGGTCATGCCCTGCAAGGCACTC
		Reverse	GGGCATGACCAACAACACGCAAGTTTGCTG
	Y54A	Forward	TTGTCTGATCCTAATGCTGATGCCTTG
		Reverse	GCATTAGGATCAGACAACTGGTCAGCC
	Y54L	Forward	ATGCTTAATCCTAACCTTGAAGATCTG
		Reverse	AGGTTAGGATTAAGCATGTCTTCTGCTG
	Y54F	Forward	TGTCTGATCCTAATTTTGATGCCTTGTTG
		Reverse	AAATTAGGATCAGACAACTGGTCAGCCGG
	Y54W	Forward	TGTCTGATCCTAATTGGGATGCCTTG
		Reverse	CCAATTAGGATCAGACAACTGGTCAGC
	Y54R	Forward	TTGTCTGATCCTAATCGTGATGCCTTG
		Reverse	CGATTAGGATCAGACAACTGGTCAGCC
	C38A	Forward	TGACAACACAGTCTGGGCCCCACGACAC
		Reverse	GCCCAGACTGTGTTGTCAAGCCAAAGACC
	C38S	Forward	TGACAACACAGTCTGGTCCCCACGACAC
		Reverse	GACCAGACTGTGTTGTCAAGCCAAAGACC
	P39A	Forward	GACAACACAGTCTGGTGCGCACGACACGTAATG
		Reverse	CGCACCAGACTGTGTTGTCAAGCCAAAGACC
	P39G	Forward	AACACAGTCTGGTGCGGACGACACGTAAT
		Reverse	CCGCACCAGACTGTGTTGTCAAGCCAAAG
	Q195A	Forward	GATAAACAAGTGCACGCAGTTCAGTTAACAGAC
		Reverse	GCGTGCACTTGTTTATCCATAAAGGCA
	Q195L	Forward	GATAAACAAGTGCACCTCGTTCAGTTAACAG
		Reverse	GAGGTGCACTTGTTTATCCATAAAGGCA
	Q195H	Forward	ATAAACAAGTGCACCACGTTCAGTTAACAG
		Reverse	GTGGTGCACTTGTTTATCCATAAAGGCA
	Q195K	Forward	GGATAAACAAGTGCACAAAGTTCAGTTA
		Reverse	TGTGCACTTGTTTATCCATAAAGGCACCA
	Q167H	Forward	TTCTGTTACATGCATCACATGGAACTTG

		Reverse	GTGATGCATGTAACAGAAATTGATCACAC
	Q167E	Forward	TTTCTGTTACATGCATGAAATGGAACTTG
		Reverse	CATGCATGTAACAGAAATTGATCACACTAC
	Q167N	Forward	TTCTGTTACATGCATAACATGGAACTTG
		Reverse	GTTATGCATGTAACAGAAATTGATCACAC
	Q167A	Forward	TTCTGTTACATGCATGCAATGGAACTTG
		Reverse	GCATGCATGTAACAGAAATTGATCACAC
	Q167V	Forward	TTTCTGTTACATGCATGTGATGGAACTTG
		Reverse	CACATGCATGTAACAGAAATTGATCACACTA
	Q167L	Forward	TTCTGTTACATGCATCTAATGGAACTTGC
		Reverse	AGATGCATGTAACAGAAATTGATCACACTAC
	Q167S	Forward	TTCTGTTACATGCATTCAATGGAACTTG
		Reverse	GAATGCATGTAACAGAAATTGATCACAC
	T88C	Forward	GGTCATGCCATGCAAGGCTGCCTTTTGAAGTTGAC
		Reverse	GCAGCCTTGCATGGCATGACCAACAACACGCAAGTT
	T88A	Forward	GGTCATGCCATGCAAGGCGCTCTTTTGAAG
		Reverse	CGCCTTGCATGGCATGACCAACAACACGC
	T88V	Forward	GGTCATGCCATGCAAGGCGTCCTTTTGAAGTTGAC
		Reverse	GACGCCTTGCATGGCATGACCAACAACACGCAAGTT
	S178A	Forward	ATGGTACACATACCGGTGCAGCATTTGATG
		Reverse	CACCGGTATGTGTACCATTAGCAAGTTCC
	S178T	Forward	ATGGTACACATACCGGTACAGCATTTGATG
		Reverse	TACCGGTATGTGTACCATTAGCAAGTTCC
	M168L	Forward	TTCTGTTACATGCATCAACTGGAACTTGCTA
		Reverse	GTTGATGCATGTAACAGAAATTGATCACAC
SARS-CoV	H163A	Forward	TCTTTCTGCTATATGGCGCATATGGAGCTTC
		Reverse	CGCCATATAGCAGAAAGACACGCAATCAT
	H163L	Forward	TCTTTCTGCTATATGCTTCATATGGAGCTTC
		Reverse	AGCATATAGCAGAAAGACACGCAATCAT

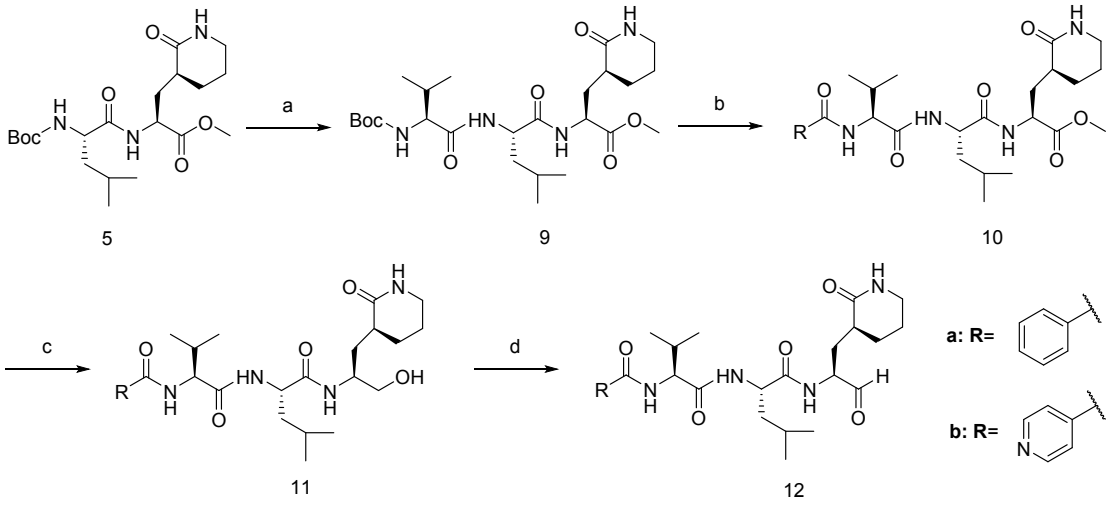
	Y161F	Forward	TGCGTGTCTTTCTGCTTTATGCATCATATG
		Reverse	AAGCAGAAAGACACGCAATCATAATCA
	F140A	Forward	ATACCATTAAAGGTTCTGCCCTTAATGGATC
		Reverse	GCAGAACCTTTAATGGTATGATTAGGTCTC
	F140L	Forward	ATACCATTAAAGGTTCTTTGCTTAATGGATC
		Reverse	CAAAGAACCTTTAATGGTATGATTAGGTCTC
	E166A	Forward	TATATGCATCATATGGCCCTTCCAACAGGAGT
		Reverse	GGCCATATGATGCATATAGCAGAAAGACAC
	E166L	Forward	TGCTATATGCATCATATGCTGCTTCCAACAGG
		Reverse	AGCATATGATGCATATAGCAGAAAGACACGC
	H172A	Forward	CTTCCAACAGGAGTAGCCGCTGGTACTG
		Reverse	GCTACTCCTGTTGGAAGCTCCATATGAT
	H172L	Forward	TTCCAACAGGAGTACTCGCTGGTACTGACC
		Reverse	AGTACTCCTGTTGGAAGCTCCATATGAT
	H172Q	Forward	GCTTCCAACAGGAGTACAGGCTGGTACTGAC
		Reverse	CTGTACTCCTGTTGGAAGCTCCATATGAT
	G143A	Forward	TCTTTCCTTAATGCCTCATGTGGTAGTGTTGG
		Reverse	GGCATTAAGGAAAGAACCTTTAATGGTATG
	S144A	Forward	GGTTCTTTCCTTAATGGAGCATGTGGTAGTG
		Reverse	CTCCATTAAGGAAAGAACCTTTAATGGTATG
	G146A	Forward	CTTAATGGATCATGTGCTAGTGTTGGTTTT
		Reverse	GCACATGATCCATTAAGGAAAGAACCTTTAA
	S147A	Forward	ATGGATCATGTGGTGCTGTTGGTTTTAACATT
		Reverse	GCACCACATGATCCATTAAGGAAAGAACC
	N28L	Forward	GGAACTACAACTCTTCTTGGATTGTGGTTGGAT
		Reverse	AGAAGAGTTGTAGTTCCACAGGTTACTTG
	N28D	Forward	GGAACTACAACTCTTGACGGATTGTGGTTGG
		Reverse	GTCAAGAGTTGTAGTTCCACAGGTTACTTGTAC
	N28H	Forward	GTGGAACTACAACTCTTCATGGATTGTG

		Reverse	GAAGAGTTGTAGTTCCACAGGTTACTTGTAC
	D187L	Forward	TATGGTCCATTTGTTCTGAGACAAACTGC
		Reverse	CAGAACAAATGGACCATAGAATTTACCTTC
	R40L	Forward	ACACAGTATACTGTCCACTTCATGTCATTTG
		Reverse	AAGTGGACAGTATACTGTGTCATCCAACCAC
	Y54F	Forward	ATGCTTAATCCTAACTTTGAAGATCTG
		Reverse	AAGTTAGGATTAAGCATGTCTTCTGCTG
	Y54L	Forward	ATGCTTAATCCTAACCTTGAAGATCTGCTCATTCGC
		Reverse	AGGTTAGGATTAAGCATGTCTTCTGCTGTGC
	C38A	Forward	GGATGACACAGTATACGCTCCAAGACATGTCATTTGC
		Reverse	GCGTATACTGTGTCATCCAACCACAATCCATT
	C38S	Forward	GGATGACACAGTATACTCTCCAAGACATGTCATTTGC
		Reverse	GAGTATACTGTGTCATCCAACCACAATCCATTAA
	P39A	Forward	GACACAGTATACTGTGCGAGACATGTCAT
		Reverse	CGCACAGTATACTGTGTCATCCAACCACA
	Q192L	Forward	GACAGACAAACTGCACTGGCTGCAGGTACA
		Reverse	AGTGCAGTTTGTCTGTCAACAAATGGAC
	Q192H	Forward	GACAGACAAACTGCACACGCTGCAGGTACAGAC
		Reverse	GTGTGCAGTTTGTCTGTCAACAAATGGAC
	H164A	Forward	TTCTGCTATATGCATGCCATGGAGCTTCCAAC
		Reverse	GGCATGCATATAGCAGAAAGACACGCAAT
	H164V	Forward	TTCTGCTATATGCATGTCATGGAGCTTCCAAC
		Reverse	GACATGCATATAGCAGAAAGACACGCAATC
	H164L	Forward	TTCTGCTATATGCATCTCATGGAGCTTCCAAC
		Reverse	GAGATGCATATAGCAGAAAGACACGCAATC
	M165L	Forward	TTCTGCTATATGCATCATCTGGAGCTTCCAAC
		Reverse	GATGATGCATATAGCAGAAAGACACGCAATC

Supporting chemical schemes



Scheme S1. Reagents and conditions: (a) (1) $SOCl_2$, MeOH, reflux, 2 h; (2) $(Boc)_2O$, TEA, anhydrous THF; 99% for two steps. (b) (1) LiHMDS, argon atmosphere, anhydrous THF; -78 ^oC; (2) $BrCH_2CN$ or $BrCH_2CH_2CN$, argon atmosphere, anhydrous THF, 3 h, -78 ^oC. 62% for two steps. (c) $CoCl_2 \cdot 6H_2O$, $NaBH_4$, MeOH, 0 ^oC, 2 days, 39.5%. (d) (1) TFA, DCM, RT, 3 h; (2) adjust pH value to 7, (S)-Boc-Leu-OH , EDCI, HOBt and TEA, DCM; 60.4% for two steps; (e) (1) TFA, DCM, RT, 3h; (2) adjust pH value to 7, cinnamic acid, EDCI, HOBt and TEA, DCM; 61% for two steps; (f) $NaBH_4$, MeOH, RT, 2 h, 81%; (g) Dess-Martin periodinane, DCM, RT, 1 h, 91%.



Scheme S2. Reagents and conditions: (a) (1) TFA, DCM, RT, 3 h; (2) adjust pH value to 7, (S)-Boc-Val-OH , EDCI, HOBt and TEA, DCM; 54% for two steps; (b) (1) TFA, DCM, RT, 3 h; (2) adjust pH value to 7, benzoic acid or isonicotinic acid, EDCI, HOBt and TEA, DCM; 52% for two steps; (c) $NaBH_4$, MeOH, RT, 2 h, 71%; (g) Dess-Martin periodinane, DCM, RT, 1 h, 83%.

Chemical synthesis

1. Reagents and instruments

All reagents were purchased from various commercial suppliers and used as received. NMR spectra data were recorded on a Bruker AVANCE-400 NMR spectrometer (400 MHz) (Bruker, Karlsruhe, Germany). Molecular mass was determined by using a Shimadzu LCMS-2020 ESI mass spectrometry (Shimadzu, Kyoto, Japan). All final compounds exhibited purities of > 95% as analyzed by HPLC (Dionex UltiMate 3000, Germany).

2. General procedure for the synthesis of compounds.

*2.1 Procedure for the synthesis of compounds **1-8***

*2.1.1 Procedure for the preparation of dimethyl (tert-butoxycarbonyl)-L-glutamate **2.***

To a suspension of L-glutamic acid (30 g, 203.9 mmol) in anhydrous MeOH (400 mL) was added $SOCl_2$ (11.83 mL, 203.9 mmol) in drop-wise at 0 ^oC. After 30 min of string, the reaction was heated to reflux for 3 h, and then cooled to the room temperature. After

evaporating the solvent, the residue was suspended in anhydrous THF (200 mL) and added others reagent (Ditertbutyl dicarbonate (66.75 g, 305.85 mmol) and triethylamine (30.95 g, 305.85mmol)) at ice-bath. Following the reaction mixture was stirred overnight at room temperature, the solvent of the mixture was evaporated and the residue was dissolved in DCM (400 mL). Following washed with H_2O (200 mL×2), saturated citric acid solution (200 mL×2), saturated $NaHCO_3$ solution (200 mL×2) and brine (200 mL×2), the organic phase was concentrated, and purified by column chromatography (EtOAc: Petroleum ether, 1: 5 v/v) to give the pure product as a colorless oil **2** (55.57 g, 201.86 mmol, 99.0 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.44 (d, *J* = 8.1 Hz, 1H), 4.33 (dd, *J* = 12.6, 7.5 Hz, 1H), 3.75 (s, 3H), 3.68 (s, 3H), 2.51 – 2.34 (m, 2H), 2.24 – 2.13 (m, 1H), 1.97 (td, *J* = 14.7, 8.2 Hz, 1H), 1.44 (s, 9H) . ^{13}CNMR (100 MHz, $CDCl_3$) δ: 173.02, 172.58, 155.32, 79.64, 52.73, 52.18, 51.56, 29.92, 28.11, 27.42. ESI-MS (m/z): 298.1$(M+Na)^+$.

2.1.2. Procedure for the preparation of compounds **3a-b**.

2.1.2.1. Procedure for the preparation of dimethyl (2S,4S)-2-((tert-butoxycarbonyl)amino)-4-(2-cyanoethyl)pentanedioate **3a.**

To a solution of **2** (20.0 g, 72.65 mmol) dissolved in anhydrous THF (500 mL), the solution of lithium hexamethyldisilazide/THF (159.83 mL, 1mol/L, 159.83 mmol) was added in drop-wise under argon atmosphere at -78°C. Following a further 2 h of stirring at -78°C, 3-Bromopropionitrile (10.71 g, 79.91 mmol) was diluted with anhydrous THF and added in drop-wise to the reaction mixture with the solution over a period of 2 h at -78°C under argon atmosphere. Following an additional 2 h at -78°C under the argon atmosphere, 20 mL pre-cooled methanol and 10 mL pre-cooled acetic acid were added to quench the reaction. After a further 10 min of stirring at -78 °C, the reaction was allowed to stir at room temperature overnight. Following filtered to removing the insoluble salt, the filtrate was evaporated and the obtained residue was further purified by column chromatography (EtOAc: Petroleum ether, 1: 5 v/v) to give the pure product as yellow oil **3a** (14.74 g, 44.89 mmol, 62.0%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.11 (d, *J* = 8.0 Hz, 1H), 4.48 – 4.31 (m, 1H), 3.74 (s, 3H), 3.72 (s, 3H), 2.70 – 2.58 (m, 1H),

2.48 – 2.35 (m, 2H), 2.13 – 1.92 (m, 4H), 1.45 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ: 174.39, 172.34, 155.38, 118.71, 80.28, 52.54, 52.16, 51.56, 40.79, 34.37, 28.24, 27.30, 15.12.

2.1.2.2.Procedure for the preparation of dimethyl (2S,4R)-2-((tert-butoxycarbonyl)amino)-4-(cyanomethyl)pentanedioate **3b.**

The similar procedure with the procedure to generate compound **3a** was executed to synthesize compound **3b** via the replacement of 3-bromopropionitrile into bromoacetonitrile in the process. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.34 (d, J = 8.8 Hz, 1H), 4.27 (dd, J = 14.2, 7.8 Hz, 1H), 3.70 – 3.58 (m, 6H), 2.83 – 2.72 (m, 1H), 2.70 – 2.57 (m, 2H), 2.11 – 1.99 (m, 2H), 1.33 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 172.49, 172.01, 155.55, 117.24, 80.19, 52.55, 52.52, 50.98, 38.16, 33.43, 28.11, 18.83. ESI-MS (m/z): 315.2 $(M+H)^+$. Yield 65%.

2.1.3. Procedure for the preparation of compounds **4a-b**.

2.1.3.1 Procedure for the preparation of methyl (S)-2-((tert-butoxycarbonyl)amino)-3-((S)-2-oxopiperidin-3-yl)propanoate **4a**

To a solution of **3a** (10 g, 30.45 mmol) dissolved into anhydrous MeOH (400 mL), $CoCl_2 \cdot 6H_2O$ (4.35 g, 18.27 mmol) was added at -10°C. Then, $NaBH_4$ (6.91 g, 182.7 mmol) was added portion-wise at 0 °C. Following maintained to stir at 0 °C for 48 h, the reaction mixture was added saturated ammonium chloride solution (50 mL) to quench the reaction and the mixture was filtered to remove insoluble substances. After evaporated the solvent, the residue was extracted with DCM (100 mL×3) and further purified by column chromatography (EtOAc: petroleum ether, 2.5:1 v/v) to give the pure yellow oil compound **4a** (3.61 g, 12.18 mmol, 39.45 %).^{1}H NMR (400MHz, $CDCl_3$) δ: 7.24 (s, 1H), 5.94 (d, J = 8.1 Hz, 1H), 4.29 (t, J = 7.8 Hz, 1H), 3.72 (s, 3H), 3.37 – 3.16 (m, 2H), 2.44 – 2.22 (m, 2H), 2.18 – 1.99 (m, 1H), 1.96 – 1.79 (m, 2H), 1.78 – 1.63 (m, 1H), 1.61 – 1.50 (m, 1H), 1.44 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ174.67, 173.23, 155.91, 79.51, 52.12, 51.68, 42.05, 37.85, 34.01, 28.21, 26.42, 21.43. ESI-MS (m/z): 301.3 $(M+H)^+$.

2.1.3.2 Procedure for the preparation of methyl (S)-2-((tert-butoxycarbonyl)amino)-3-((S)-2-oxopyrrolidin-3-yl)propanoate **4b**

The similar procedure with the procedure to generate compound **4a** was executed to synthesize compound **4b**.^{1}H NMR (400MHz, $CDCl_3$) δ: 7.56 (s, 1H), 5.99 (d, *J* = 8.1 Hz, 1H),4.37 – 4.24 (m, 1H), 3.74 (s, 3H), 3.42 – 3.25 (m, 2H), 2.56 – 2.38 (m, 2H), 2.22 – 2.07 (m, 1H), 1.93 – 1.75 (m, 2H), 1.44 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 180.04, 173.01, 155.79, 79.72, 52.32, 52.25, 40.46, 38.28, 33.91, 28.24, 27.96. ESI-MS (m/z): 287.2 $(M+H)^+$

2.1.4. Procedure for the preparation of compounds **5a-b**.

2.1.4.1.Procedure for the preparation of Methyl (S)-2-((S)-2-((tert-butoxycarbonyl)amino)-4-methylpentanamido)-3-((S)-2-oxopiperidin-3-yl)propanoate **5a.**

To a solution of **4a** (1.0 g, 3.33 mmol) dissolved in anhydrous DCM (50 mL), CF_3COOH (2.5 mL, 33.3 mmol) was added slowly at ice-bath. Subsequently, the reaction mixture was allowed to stir at room temperature for 3 h and concentrated to remove the redundant trifluoroacetic acid. Then, the triethylamine was added to the solution of the residue dissolved in DCM (60 mL) to adjust the pH value of the solution to 7.0. Subsequently, the Boc -Leu-OH (770 mg, 3.33 mmol), EDCI (765.9 mg, 4.00 mmol) and HOBt (540.0 mg, 4.00 mmol) were sequentially added. Following TEA (1.85 mL, 13.32 mmol) was added in drop-wise, the reaction mixture was stirred at ambient temperature overnight. Followed by washing with H_2O (50 mL×2), saturated citric acid solution (50 mL×2), saturated $NaHCO_3$ solution (50 mL×2) and saturated brine (50 mL×2), the organic phase was purified by column chromatography (DCM: MeOH, 100: 1 to 60: 1v/v) to afford the pure product as a light yellow foam **5a** (831.2 mg, 2.01mmol, 60.4%). ^{1}H NMR (400MHz, $CDCl_3$) δ: 7.90 (d, *J* = 7.2 Hz, 1H), 6.85 (s, 1H), 5.21 (d, *J* = 8.6 Hz, 1H), 4.54 (t, *J* = 7.3 Hz, 1H), 4.31 (dd, *J* = 14.2, 8.5 Hz, 1H), 3.71 (s, 3H), 3.32 – 3.22 (m, 2H), 2.47 – 2.26 (m, 2H), 2.11 – 2.01 (m, 1H), 1.90 – 1.82 (m, 2H), 1.79 – 1.68 (m, 2H), 1.67 – 1.60 (m, 1H), 1.57 – 1.48 (m, 2H), 1.43 (s, 9H), 0.95 (dd, *J* = 6.3, 3.9 Hz, 6H); ^{13}C NMR (101 MHz, $CDCl_3$) δ: 174.54, 173.38,

172.50, 155.57, 79.60, 52.82, 52.24, 50.20, 42.25, 42.11, 37.77, 33.19, 28.31, 26.33, 24.61, 22.87, 22.17, 21.59. ESI-MS (m/z):414.3 $(M+H)^+$.

2.1.4.2.Procedure for the preparation of Methyl (S)-2-((S)-2-((tert-butoxycarbonyl)amino)-4-methylpentanamido)-3-((S)-2-oxopyrrolidin-3-yl)propanoate **5b.**

^{1}H NMR (400MHz, $CDCl_3$) δ: 7.94 (d, *J* = 6.7 Hz, 1H), 7.30 (s, 1H), 5.41 (d, *J* = 8.3 Hz, 1H), 4.60 – 4.42 (m, 1H), 4.36 – 4.21 (m, 1H), 3.72 (s, 3H), 3.46 – 3.23 (m, 2H), 2.58 – 2.31 (m, 2H), 2.30 – 2.14 (m, 1H), 1.92 – 1.78 (m, 2H), 1.77 – 1.69 (m, 1H), 1.68 – 1.59 (m, 1H), 1.54 – 1.48 (m, 1H), 1.42 (s, 9H), 1.06 – 0.82 (m, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 179.91, 173.40, 172.25, 155.61, 79.54, 52.82, 52.24, 50.89, 42.06, 40.44, 38.26, 33.05, 28.23, 27.90, 24.55, 22.81, 22.07. ESI-MS (m/z): 400.6 $(M+H)^+$.

2.1.5. Procedure for the preparation of compounds **6a-b**.

2.1.5.1 Procedure for the preparation of Methyl (S)-2-((S)-2-cinnamamido-4-methylpentanamido)-3-((S)-2-oxopiperidin-3-yl)propanoate **6a**

The detailed equivalent and procedure of condensation reaction was referred to the procedure for the preparation of compound **5a**. ^{1}H NMR (400MHz, $CDCl_3$) δ: 8.46 (d, *J* = 6.8 Hz, 1H), 7.58 (d, *J* = 15.6 Hz, 1H), 7.51 – 7.43 (m, 2H), 7.39 – 7.30 (m, 3H), 7.19 (s, 1H), 7.10 (d, *J* = 8.8 Hz, 1H), 6.53 (d, *J* = 15.6 Hz, 1H), 5.01 (td, *J* = 8.8, 5.1 Hz, 1H), 4.53 – 4.44 (m, 1H), 3.69 (s, 3H), 3.34 – 3.25 (m, 2H), 2.60 – 2.48 (m, 1H), 2.38 – 2.25 (m, 1H), 2.05 – 1.96 (m, 1H), 1.89 – 1.79 (m, 2H), 1.79 – 1.71 (m, 2H), 1.71 – 1.57 (m, 2H), 1.53 – 1.39 (m, 1H), 0.97 (d, *J* = 3.7 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 174.18, 173.47, 172.48, 165.70, 141.13, 134.87, 129.62, 128.76, 127.84, 120.79, 52.25, 51.37, 50.49, 42.66, 42.15, 37.74, 32.94, 26.30, 24.71, 22.90, 22.26, 21.63. ESI-MS (m/z): 466.3 $(M+Na)^+$.

2.1.5.2. Procedure for the preparation of Methyl (S)-2-((S)-2-cinnamamido-4-methylpentanamido)-3-((S)-2-oxopiperidin-3-yl)propanoate **6b**

^{1}H NMR (400MHz, $CDCl_3$) δ: 8.43 (d, *J* = 6.8 Hz, 1H), 7.58 (d, *J* = 15.6 Hz, 1H), 7.46 (dd, *J* = 6.4, 2.6 Hz, 2H), 7.33 (dd, *J* = 8.9, 5.2 Hz, 4H), 7.19 (d, *J* = 9.1 Hz, 1H), 6.53

(d, *J* = 15.6 Hz, 1H), 5.02 (td, *J* = 8.9, 5.0 Hz, 1H), 4.46 – 4.37 (m, 1H), 3.70 (s, 3H), 3.43 – 3.24 (m, 2H), 2.45 – 2.25 (m, 3H), 1.89 – 1.70 (m, 4H), 1.70 – 1.58 (m, 1H), 0.98 (d, *J* = 4.0 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 179.47, 173.56, 172.15, 165.84, 141.26, 134.81, 129.69, 128.78, 127.85, 120.71, 52.31, 51.36, 51.34, 42.61, 40.55, 38.42, 32.74, 27.94, 24.72, 22.90, 22.20 . ESI-MS (m/z):452.4 $(M+H)^+$.

2.1.6. Procedure for the preparation of compounds **7a-b**.

To a solution of **6a-b** (0.90 mmol) dissolved into anhydrous MeOH (30.0 mL), $NaBH_4$ (0.51 g, 13.53 mmol) was added at 0 °C. Then, the reaction mixture was stirred at ambient temperature for 2 h. Saturated NH_4Cl solution (20 mL) was added to the mixture for quench the reaction. Following evaporating the solvent, EA (60 mL×2) was added to extract the aqueous components. The organic phase was washed with H_2O (30 mL×2), saturated brine (30 mL×2), and concentrated. Finally, the residue was purified by column chromatography (DCM: MeOH, 25:1 v/v) to afford the pure product as a white solid **7a-b** (yield 79% -81%).

2.1.6.1.(S)-2-cinnamamido-N-((S)-1-hydroxy-3-((S)-2-oxopiperidin-3-yl)propan-2-yl)-4-methylpentanamide **7a**

^{1}H NMR (400MHz, MeOD) δ: 8.01 (d, *J* = 8.7 Hz, 1H), 7.61 – 7.49 (m, 3H), 7.44 – 7.32 (m, 3H), 6.70 (d, *J* = 15.8 Hz, 1H), 4.49 (t, *J* = 7.4 Hz, 1H), 4.10 – 3.95 (m, 1H), 3.56 – 3.44 (m, 2H), 3.27 – 3.16 (m, 2H), 2.44 – 2.30 (m, 1H), 2.18 – 1.96 (m, 2H), 1.86 – 1.76 (m, 1H), 1.75 – 1.58 (m, 5H), 1.52 – 1.43 (m, 1H), 0.97 (dd, *J* = 11.7, 6.4 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ: 176.05, 173.83, 167.12, 140.72, 134.88, 129.51, 128.58, 127.49, 120.20, 64.23, 52.40, 48.45, 41.60, 40.75, 37.25, 32.43, 25.68, 24.64, 22.07, 20.67, 20.65. ESI-MS (m/z): 438.3 $(M+Na)^+$.

2.1.6.2.(S)-2-cinnamamido-N-((S)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)-4-methylpentanamide **7b**

^{1}H NMR (400MHz, $CDCl_3$) δ: 7.95 (d, *J* = 6.8 Hz, 1H), 7.49 (d, *J* = 15.6 Hz, 1H), 7.35 (s, 2H), 7.30 – 7.18 (m, 4H), 6.88 (s, 1H), 6.47 (d, *J* = 15.6 Hz, 1H), 4.75 – 4.57 (m, 1H), 4.23 – 4.11 (m, 1H), 3.96 (dd, *J* = 13.3, 12.9 Hz, 1H), 3.25 – 3.04 (m, 2H), 3.05 – 2.88 (m, 1H), 2.43 – 2.28 (m, 1H), 2.28 – 2.14 (m, 1H), 2.08 – 1.86 (m, 1H), 1.78 – 1.53 (m, 4H), 1.53 – 1.41 (m, 1H), 0.87 (s, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 181.03,

173.62, 166.17, 141.22, 134.74, 129.72, 128.80, 127.83, 120.67, 65.47, 52.26, 50.41, 42.16, 40.64, 38.31, 32.22, 28.32, 24.92, 23.06, 22.07. ESI-MS (m/z): 424.8 $(M+Na)^+$.

2.1.7. Procedure for the preparation of compounds **8a-b**.

To a solution of **7a-b** (0.48 mmol) in anhydrous DCM (30.0 mL), Dess-Martin reagent (306.2 mg, 0.72 mmol) was added at ice-bath. Then, the reaction mixture was allowed to stir at ambient temperature for 2 h. A solution of $NaHCO_3$ and solid $Na_2S_2O_3$ were added to quench the reaction. After 40 min of string, the reaction mixture was extracted by DCM (30.0 mL × 2). The organic phase was washed with brine (30 mL × 2), dried over Na_2SO_4, and concentrated, and the residue was purified by column chromatography (DCM: MeOH, 50:1 to 35:1 v/v) to afford the pure product as a white solid **8a-b** (yield 86%-91%). Owing the existence of MeOH, the aldehyde prefer to form diastereoisomers hemiacetals (Figure S28). For obtaining the purity aldehyde as far as possible, abundant CCl_4 and hexane were added into the eluent to form azeotropes. Then, the eluent was concentrated at 44 ^oC and give the residue which mainly contain aldehyde. Following the abundant hexane was added into the solution of residue dissolved into chloroform, the precipitation was filtered and give the purity aldehyde.

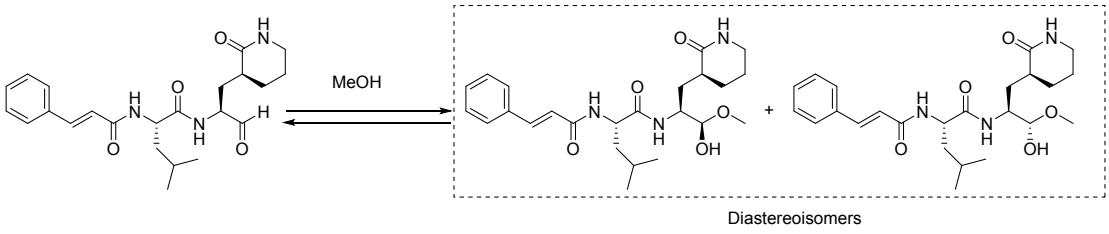


Figure S28. The forming process to generate hemiacetals from aldehyde.

2.1.7.1. (S)-2-cinnamamido-4-methyl-N-((S)-1-oxo-3-((S)-2-oxopiperidin-3-yl)propan-2-yl)pentanamide **8a**

^1H NMR (400MHz, $CDCl_3$) δ: 9.50 (s, 1H), 8.64 (d, *J* = 6.2 Hz, 1H), 7.58 (d, *J* = 15.6 Hz, 1H), 7.45 (d, *J* = 3.7 Hz, 2H), 7.32 (s, 3H), 7.06 (d, *J* = 8.8 Hz, 2H), 6.51 (d, *J* = 15.6 Hz, 1H), 5.04 – 4.86 (m, 1H), 4.44 – 4.28 (m, 1H), 3.37 – 3.15 (m, 2H), 2.44 – 2.19 (m, 2H), 2.05 – 1.90 (m, 1H), 1.90 – 1.78 (m, 2H), 1.77 – 1.60 (m, 4H), 1.52 – 1.38 (m, 1H), 0.98 (d, *J* = 5.2 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 199.96, 174.70, 173.88, 165.85, 141.29, 134.77, 129.71, 128.79, 127.84, 120.57, 57.01, 51.64, 42.48, 42.21, 37.17, 30.59, 27.01, 24.93, 22.94, 22.18, 21.43. ESI-MS (m/z): 414.2 $(M+H)^+$.

2.1.7.2. (S)-2-cinnamamido-4-methyl-N-((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)pentanamide **8b**

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.43 (s, 1H), 8.60 (d, *J* = 6.1 Hz, 1H), 7.51 (d, *J* = 15.6 Hz, 1H), 7.38 (dd, *J* = 6.2, 2.6 Hz, 2H), 7.32 – 7.21 (m, 3H), 7.03 (s, 1H), 6.97 (d, *J* = 8.8 Hz, 1H), 6.44 (d, *J* = 15.6 Hz, 1H), 4.95 – 4.82 (m, 1H), 4.28 – 4.17 (m, 1H), 3.35 – 3.14 (m, 2H), 2.49 – 2.32 (m, 1H), 2.32 – 2.17 (m, 2H), 2.12 – 1.97 (m, 1H), 1.81 – 1.71 (m, 1H), 1.70 – 1.64 (m, 2H), 1.63 – 1.54 (m, 1H), 0.91 (d, *J* = 5.5 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 199.73, 179.85, 174.02, 165.94, 141.46, 134.70, 129.79, 128.82, 127.85, 120.46, 57.65, 51.55, 42.57, 40.63, 38.14, 29.70, 28.30, 24.92, 22.94, 22.12. ESI-MS (m/z): 400.5 $(M+H)^+$.

2.1.8. Procedure for the preparation of compounds methyl (6S,9S,12S)-9-isobutyl-6-isopropyl-2,2-dimethyl-4,7,10-trioxo-12-(((S)-2-oxopiperidin-3-yl)methyl)-3-oxa-5,8,11-triazatridecan-13-oate **9**.

The detailed equivalent and procedure of condensation reaction was referred to the procedure for the preparation of compound **5a.** ^{1}H NMR (400MHz, $CDCl_3$) δ: 7.94 (d, *J* = 7.2 Hz, 1H), 7.22 (d, *J* = 7.8 Hz, 1H), 6.93 (s, 1H), 5.15 (d, *J* = 9.1 Hz, 1H), 4.66 (td, *J* = 8.9, 4.9 Hz, 1H), 4.61 – 4.50 (m, 1H), 3.88 (t, *J* = 7.9 Hz, 1H), 3.70 (s, 3H), 3.35 – 3.19 (m, 2H), 2.47 – 2.34 (m, 2H), 2.33 – 2.23 (m, 1H), 2.13 – 1.99 (m, 2H), 1.92 – 1.81 (m, 2H), 1.77 – 1.61 (m, 3H), 1.58 – 1.51 (m, 1H), 1.43 (s, 9H), 0.93 (td, *J* = 6.9, 3.2 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 174.57, 172.63, 172.29, 171.54, 155.96, 79.88, 60.18, 52.24, 51.64, 50.09, 42.18, 37.74, 33.50, 30.80, 28.31, 28.26, 26.31, 24.61, 22.89, 22.00, 21.44, 19.20, 18.03 . ESI-MS (m/z): 513.8 $(M+H)^+$.

2.1.9. Procedure for the preparation of compounds **10a-10b**.

The detailed equivalent and procedure of condensation reaction was referred to the procedure for the preparation of compound **5a.**

2.1.9.1. Methyl (S)-2-((S)-2-((S)-2-benzamido-3-methylbutanamido)-4-methylpentanamido)-3-((S)-2-oxopiperidin-3-yl)propanoate **10a**

^{1}H NMR (400MHz, $CDCl_3$) δ: 8.07 (d, *J* = 7.9 Hz, 1H), 7.81 (d, *J* = 7.3 Hz, 2H), 7.72 (d, *J* = 8.6 Hz, 1H), 7.49 (t, *J* = 7.3 Hz, 1H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.17 (d, *J* = 8.8

Hz, 1H), 7.04 (s, 1H), 4.71 – 4.58 (m, 2H), 4.54 (t, *J* = 8.1 Hz, 1H), 3.69 (s, 3H), 3.34 – 3.20 (m, 2H), 2.49 – 2.37 (m, 1H), 2.35 – 2.25 (m, 1H), 2.25 – 2.13 (m, 1H), 2.11 – 1.99 (m, 1H), 1.92 – 1.80 (m, 2H), 1.72 – 1.59 (m, 3H), 1.58 – 1.42 (m, 2H), 0.99 (dd, *J* = 6.6, 1.4 Hz, 6H), 0.86 (dd, *J* = 11.2, 6.1 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 174.46, 172.66, 172.35, 171.16, 167.49, 134.11, 131.67, 128.53, 127.20, 59.05, 52.26, 51.90, 49.91, 42.28, 41.92, 37.73, 33.58, 31.46, 26.20, 24.70, 22.76, 22.08, 21.48, 19.28, 18.56. ESI-MS (m/z): 517.5 $(M+H)^+$.

2.1.9.2. Methyl (S)-2-((S)-2-((S)-2-(isonicotinamido)-3-methylbutanamido)-4-methylpentanamido)-3-((S)-2-oxopiperidin-3-yl)propanoate **10b**

^{1}H NMR (400MHz, $CDCl_3$) δ: 8.66 (dd, *J* = 4.6, 1.2 Hz, 2H), 8.02 (d, *J* = 7.9 Hz, 1H), 7.88 (dd, *J* = 4.4, 1.8 Hz, 2H), 7.74 (d, *J* = 7.6 Hz, 1H), 7.68 (d, *J* = 8.2 Hz, 1H), 7.08 (s, 1H), 4.41 – 4.26 (m, 2H), 4.05 – 3.98 (m, 1H), 3.72 (s, 3H), 3.46 – 3.36 (m, 2H), 2.45 – 2.31 (m, 1H), 2.19 – 1.98 (m, 3H), 1.97 – 1.83 (m, 1H), 1.79 – 1.56 (m, 5H), 1.51 – 1.37 (m, 1H), 0.97 (d, *J* = 6.4 Hz, 6H), 0.91 (dd, *J* = 16.1, 6.1 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 175.47, 173.39, 172.14, 171.99, 165.42, 148.14, 142.66, 121.03, 58.20, 54.55, 53.81, 53.77, 42.72, 41.55, 36.61, 31.94, 30.44, 24.14, 23. 84, 21.65, 20.35, 19.31, 18.91. ESI-MS (m/z):540.8 $(M+Na)^+$.

2.1.10. Procedure for the preparation of compounds **11a-11b**.

The detailed equivalent and procedure of reduction reaction was referred to the procedure for the preparation of **7a.**

2.1.10.1. N-((S)-1-(((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopiperidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)amino)-3-methyl-1-oxobutan-2-yl)benzamide **11a.**

^{1}H NMR (400MHz, MeOD) δ: 7.89 – 7.79 (m, 2H), 7.58 – 7.50 (m, 1H), 7.46 (t, *J* = 7.5 Hz, 2H), 4.45 – 4.34 (m, 2H), 4.07 – 3.96 (m, 1H), 3.49 (ddd, *J* = 24.6, 11.0, 5.6 Hz, 2H), 3.28 – 3.16 (m, 2H), 2.41 – 2.29 (m, 1H), 2.24 – 1.99 (m, 3H), 1.85 – 1.76 (m, 1H), 1.75 – 1.53 (m, 5H), 1.51 – 1.40 (m, 1H), 1.02 (d, *J* = 6.7 Hz, 6H), 0.93 (dd, *J* = 18.2, 6.5 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ: 175.94, 173.44, 172.37, 168.93, 133.98, 131.46, 128.19, 127.12, 64.28, 59.57, 52.30, 48.28, 41.62, 40.55, 37.21, 32.65,

30.67, 25.72, 24.47, 21.97, 20.79, 20.75, 18.56, 17.94. ESI-MS (m/z):511.4 (M+ Na)$^+$.

2.1.10.2 N-((S)-1-(((S)-1-(((S)-1-hydroxy-3-((S)-2-oxopiperidin-3-yl)propan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)amino)-3-methyl-1-oxobutan-2-yl)isonicotinamide **11b.**

^{1}H NMR (400MHz, MeOD) δ: 8.69 (dd, *J* = 4.6, 1.5 Hz, 2H), 7.81 (dd, *J* = 4.6, 1.6 Hz, 2H), 4.39 (dd, *J* = 8.7, 5.5 Hz, 2H), 4.11 – 3.96 (m, 1H), 3.49 (ddd, *J* = 24.6, 10.9, 5.7 Hz, 2H), 3.28 – 3.16 (m, 2H), 2.41 – 2.28 (m, 1H), 2.24 – 1.99 (m, 3H), 1.87 – 1.76 (m, 1H), 1.76 – 1.54 (m, 5H), 1.53 – 1.39 (m, 1H), 1.02 (d, *J* = 6.7 Hz, 6H), 0.93 (dd, *J* = 18.1, 6.5 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ: 175.91, 173.31, 171.95, 166.62, 149.54, 142.23, 121.77, 64.28, 59.80, 52.27, 49.23, 48.27, 41.62, 40.55, 37.21, 32.64, 30.57, 25.73, 24.48, 21.97, 20.80, 20.78, 18.51, 17.98. ESI-MS (m/z):512.5 (M+ Na)$^+$.

2.1.11. Procedure for the preparation of compounds **12a-12b**.

The detailed equivalent and procedure of reduction reaction was referred to the procedure for the preparation of **8a.**

2.1.11.1 *N-((S)-3-methyl-1-(((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopiperidin-3-yl)propan-2-yl)amino)pentan-2-yl)amino)-1-oxobutan-2-yl)benzamide* **12a.**

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.50 (s, 1H), 8.34 (d, *J* = 6.5 Hz, 1H), 7.81 (d, *J* = 7.6 Hz, 3H), 7.49 (t, *J* = 7.3 Hz, 1H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.22 (d, *J* = 8.3 Hz, 1H), 6.99 (s, 1H), 4.80 – 4.56 (m, 2H), 4.54 – 4.38 (m, 1H), 3.39 – 3.11 (m, 2H), 2.43 – 2.28 (m, 1H), 2.27 – 2.11 (m, 2H), 2.04 – 1.93 (m, 1H), 1.93 – 1.78 (m, 2H), 1.75 – 1.56 (m, 4H), 1.51 – 1.35 (m, 1H), 0.98 (dd, *J* = 6.1, 4.1 Hz, 6H), 0.87 (dd, *J* = 10.0, 5.8 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 199.68, 174.81, 173.18, 171.48, 167.52, 134.10, 131.66, 128.52, 127.23, 58.91, 56.82, 52.02, 42.22, 41.54, 37.06, 31.49, 30.81, 26.91, 24.84, 22.73, 22.11, 21.33, 19.30, 18.53. ESI-MS (m/z): 509.7 (M+Na)$^+$.

2.1.11.1. *N-((S)-3-methyl-1-(((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopiperidin-3-yl)propan-2-yl)amino)pentan-2-yl)amino)-1-oxobutan-2-yl)isonicotinamide* **12b.**

^{1}H NMR (400MHz, $CDCl_3$) δ: 9.49 (s, 1H), 8.73 (s, 2H), 8.38 (d, *J* = 6.6 Hz, 1H), 7.68 (d, *J* = 4.5 Hz, 3H), 7.57 (d, *J* = 8.4 Hz, 1H), 6.77 (s, 1H), 4.73 – 4.54 (m, 2H), 4.51 – 4.36 (m, 1H), 3.34 – 3.20 (m, 2H), 2.41 – 2.27 (m, 1H), 2.24 – 2.11 (m, 2H), 2.05 –

1.94 (m, 1H), 1.93 – 1.79 (m, 2H), 1.77 – 1.63 (m, 3H), 1.61 – 1.46 (m, 2H), 0.97 (t, *J* = 5.2 Hz, 6H), 0.89 (dd, *J* = 10.2, 5.5 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ:199.49, 174.86, 173.03, 171.08, 165.66, 150.52, 141.17, 121.18, 59.08, 57.08, 52.04, 42.30, 41.64, 37.31, 31.52, 30.87, 27.12, 24.85, 22.74, 22.15, 21.28, 19.25, 18.54. ESI-MS (m/z): 510.8 $(M+Na)^+$.

NMR Figures

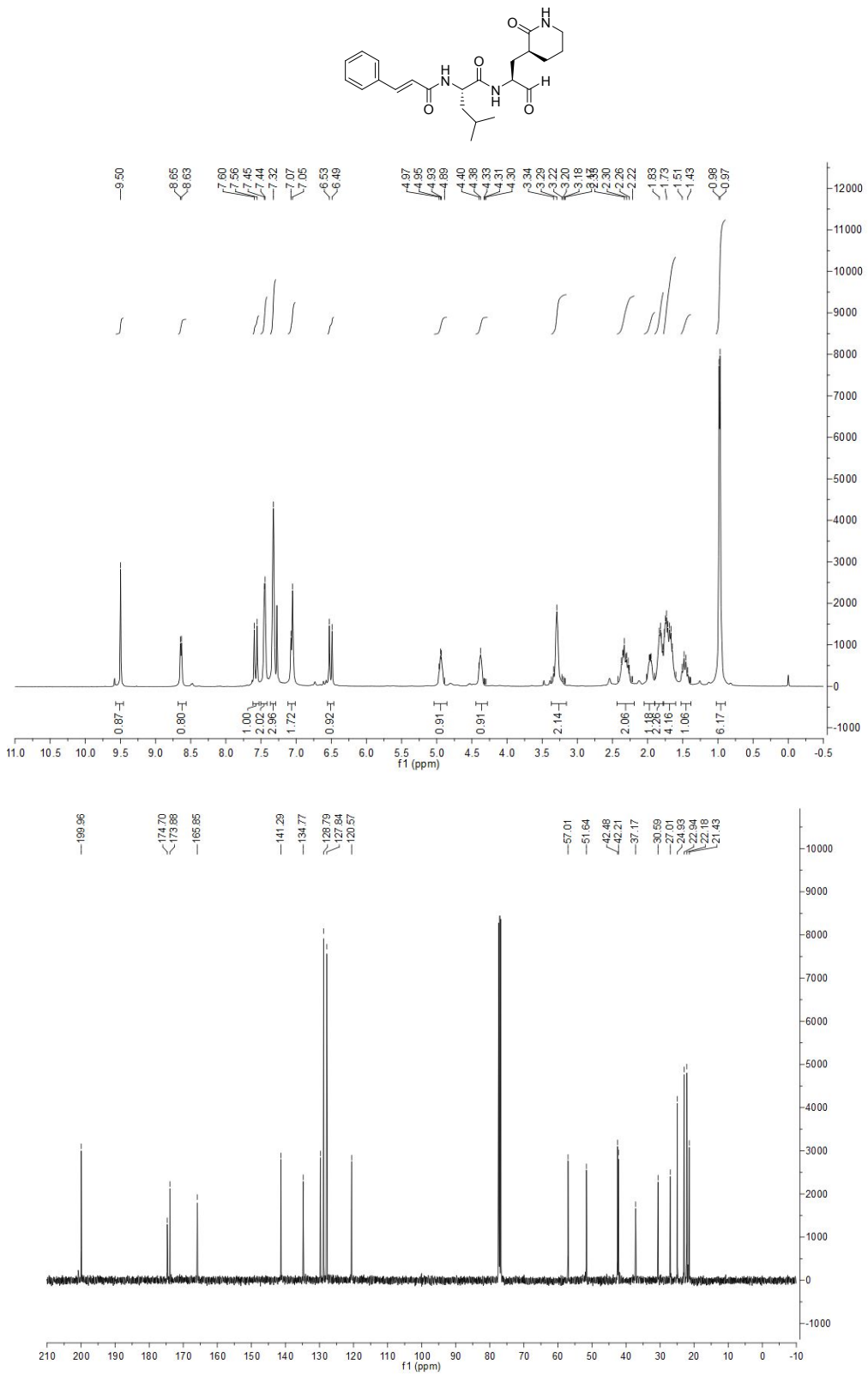


Figure S29. The NMR figures of compound **8a**.

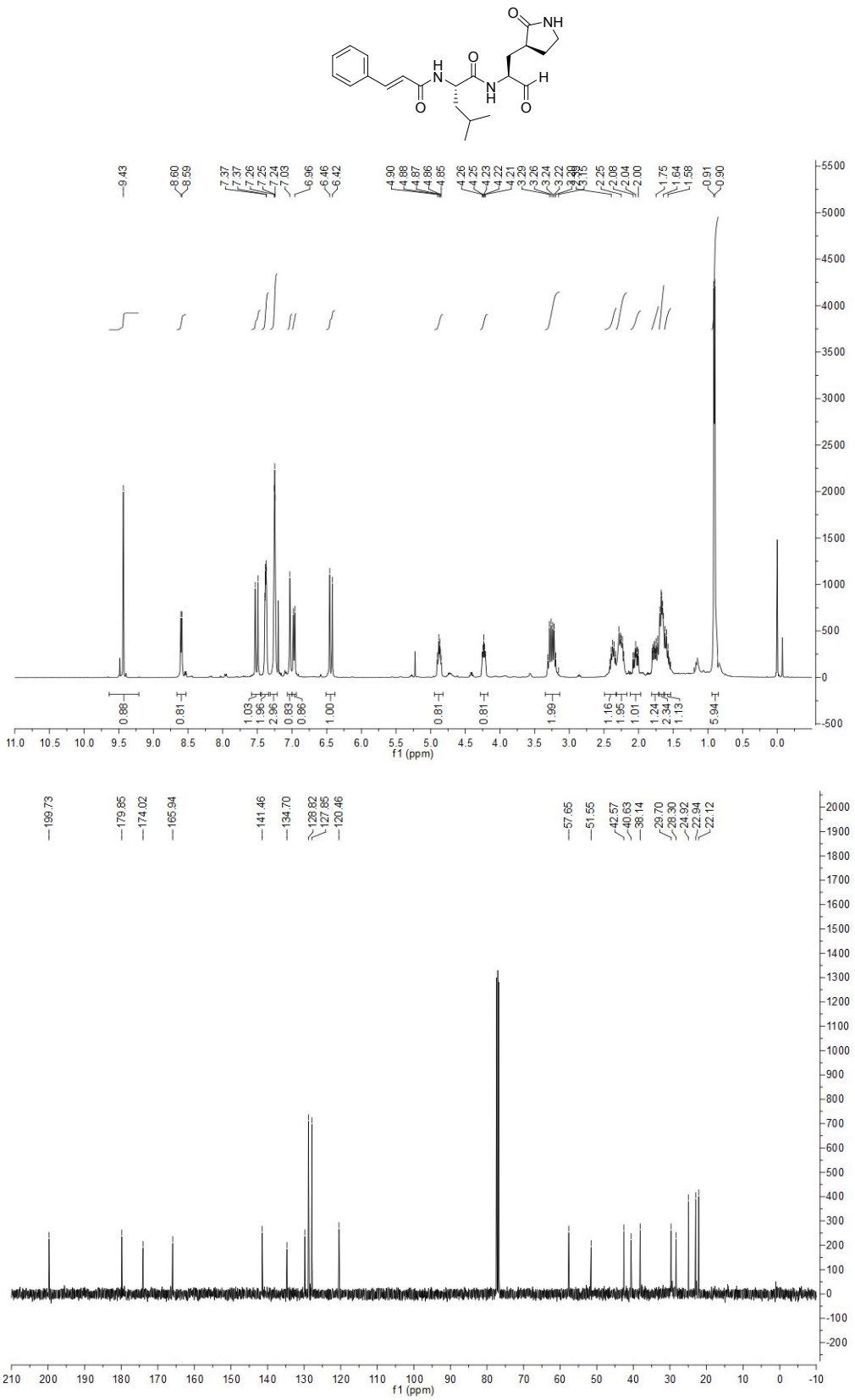


Figure S30. The NMR figures of compound **8b**.

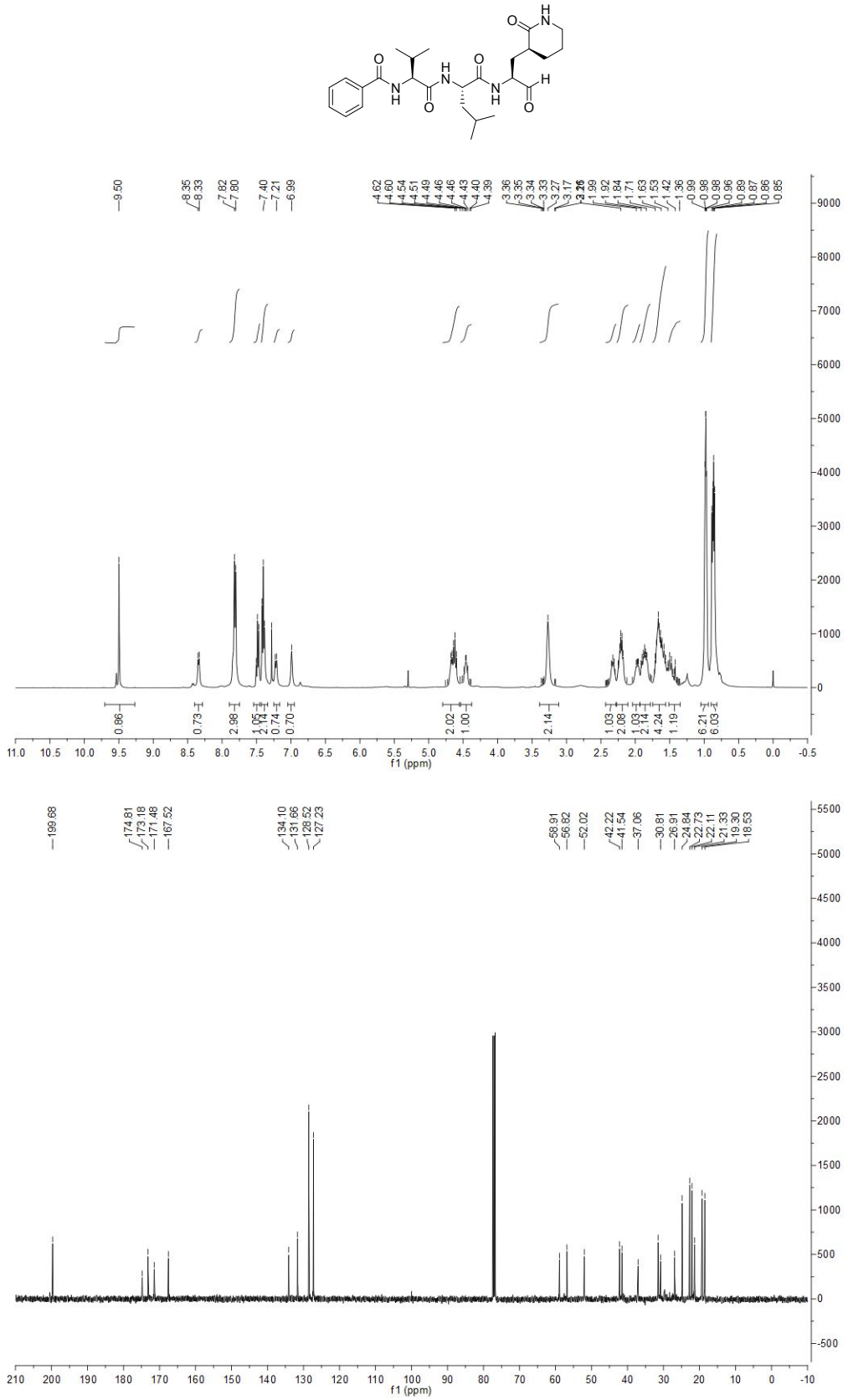


Figure S31. The NMR figures of compound **12a**.

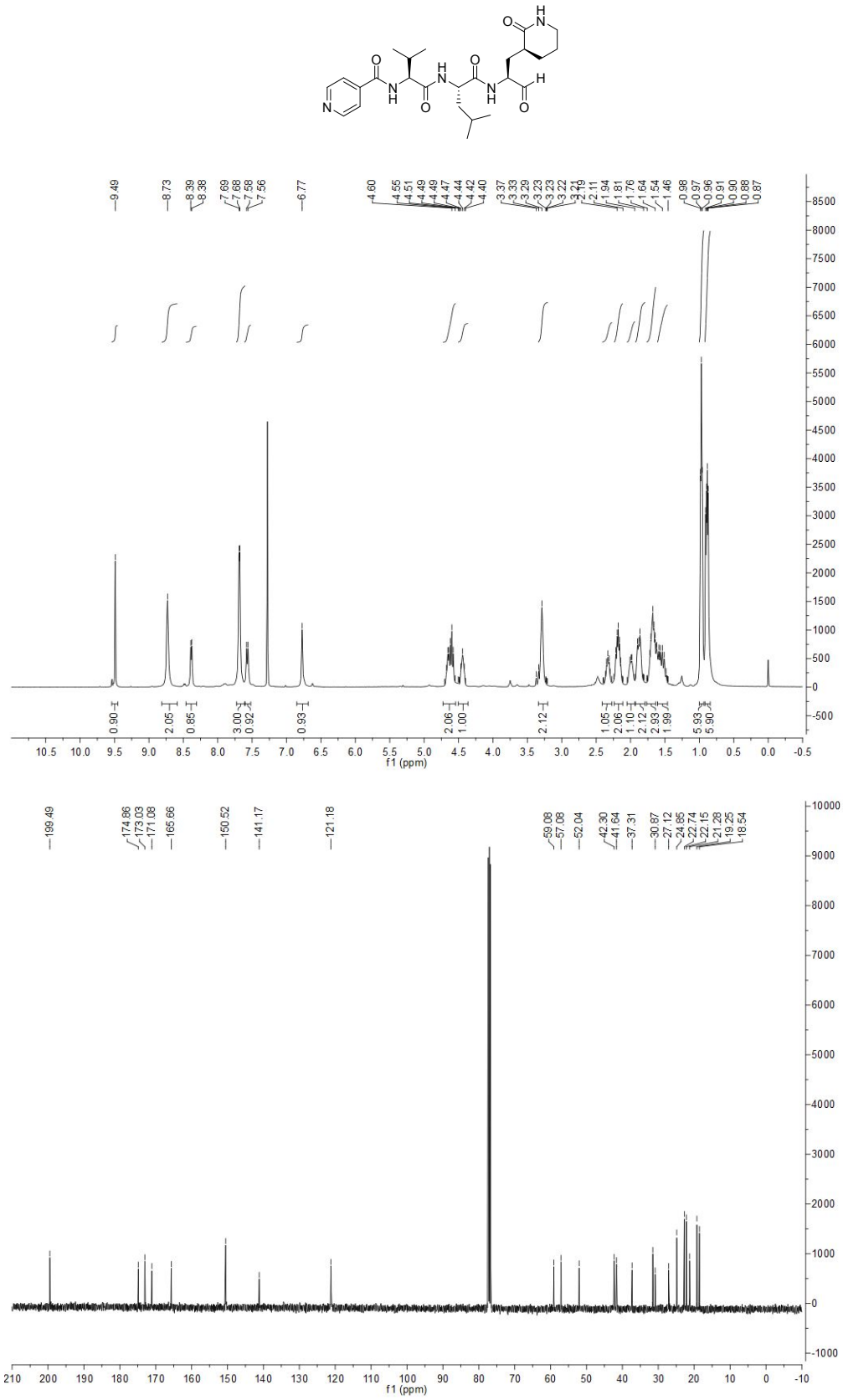
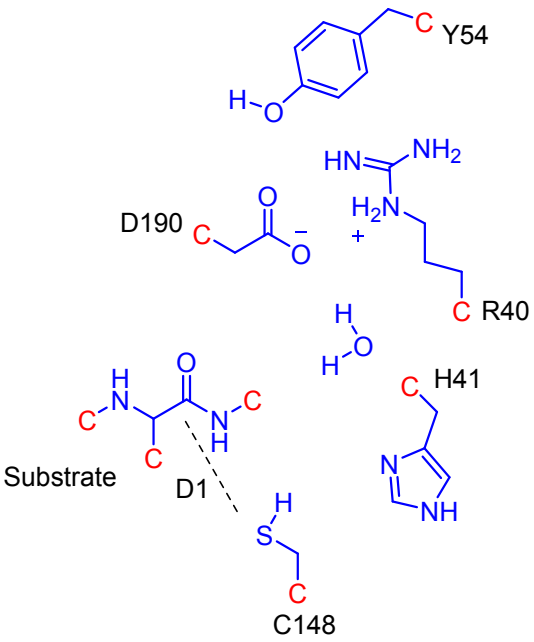


Figure S32. The NMR figures of compound **12b**.

Ab initio model scheme



Scheme S3. The model for the reaction system. During the geometry optimization, the boundary atoms are fixed at the position as they were in the protein environment.

Supporting figures in MD simulation

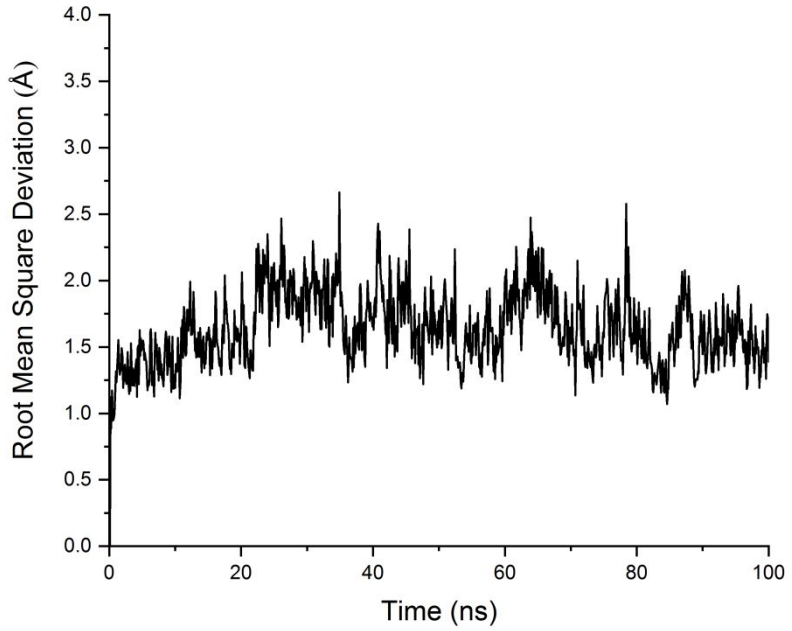


Figure S33. The time evolution of RMSD of WT MERS-CoV $3CL^{Pro}$ in MD simulation.

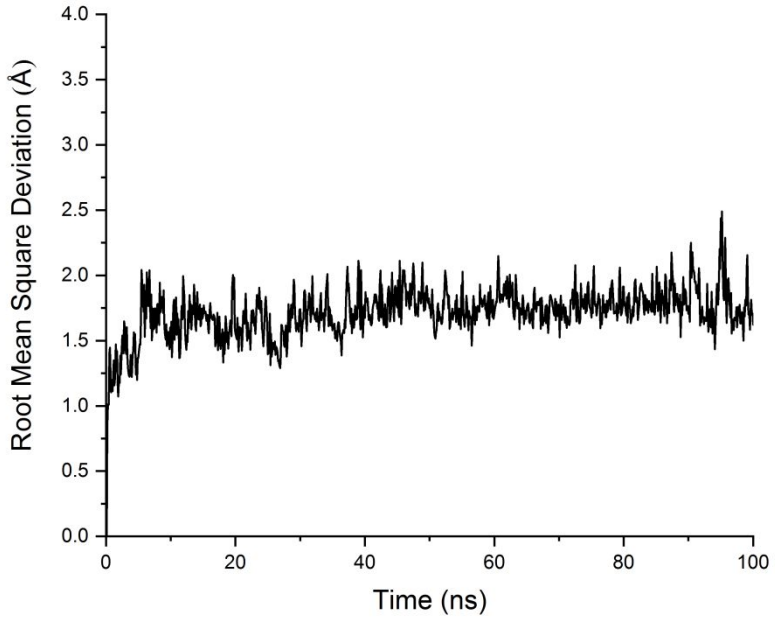


Figure S34.The time evolution of RMSD of mutant E169L in MD simulation.

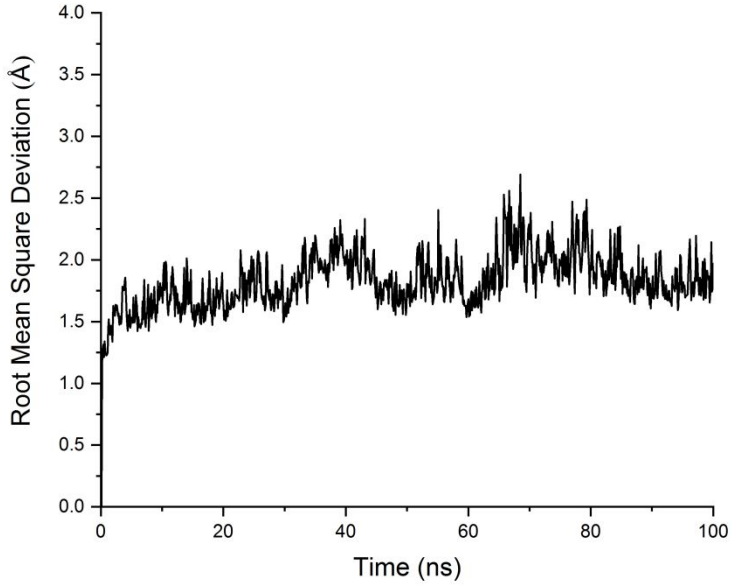


Figure S35. The time evolution of RMSD of mutant H175L in MD simulation.

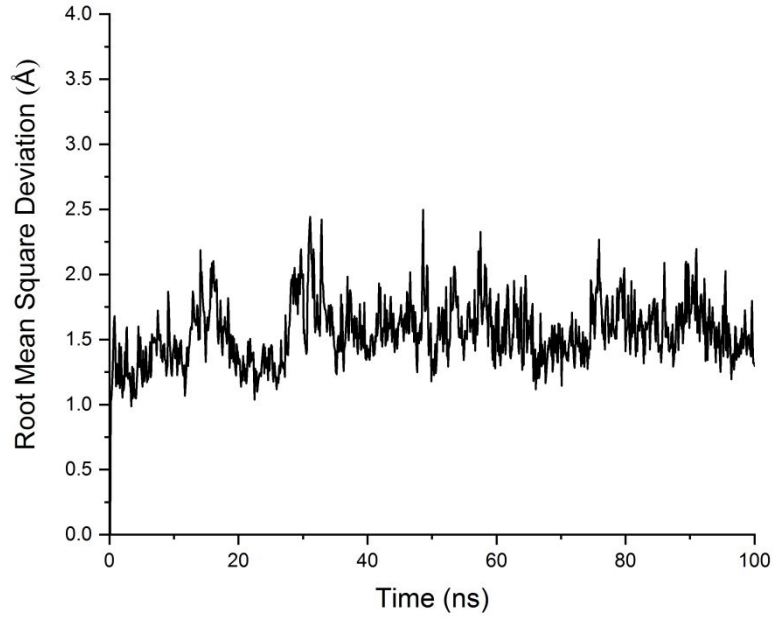


Figure S36. The time evolution of RMSD of mutant E169L/H175L in MD simulation.

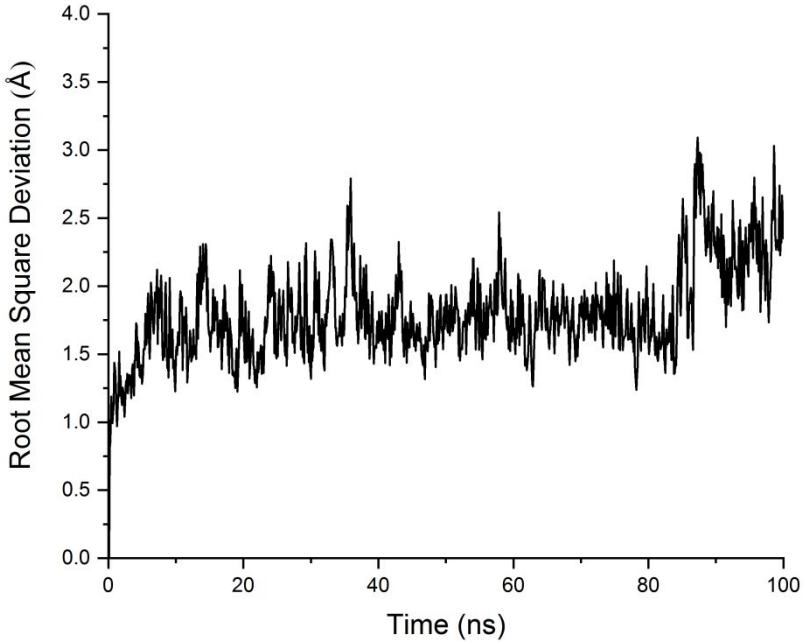


Figure S37. The time evolution of RMSD of mutant Y164F in MD simulation.

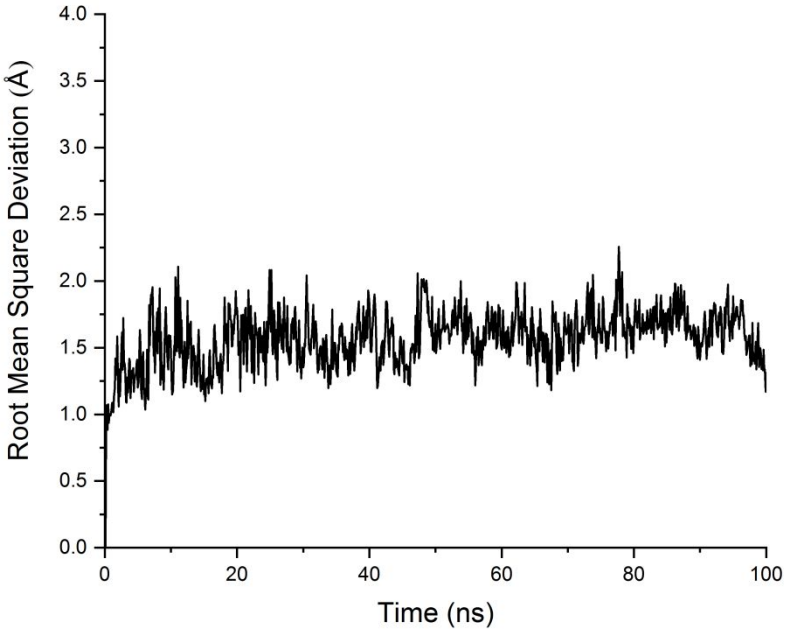


Figure S38. The time evolution of RMSD of mutant F143L in MD simulation.

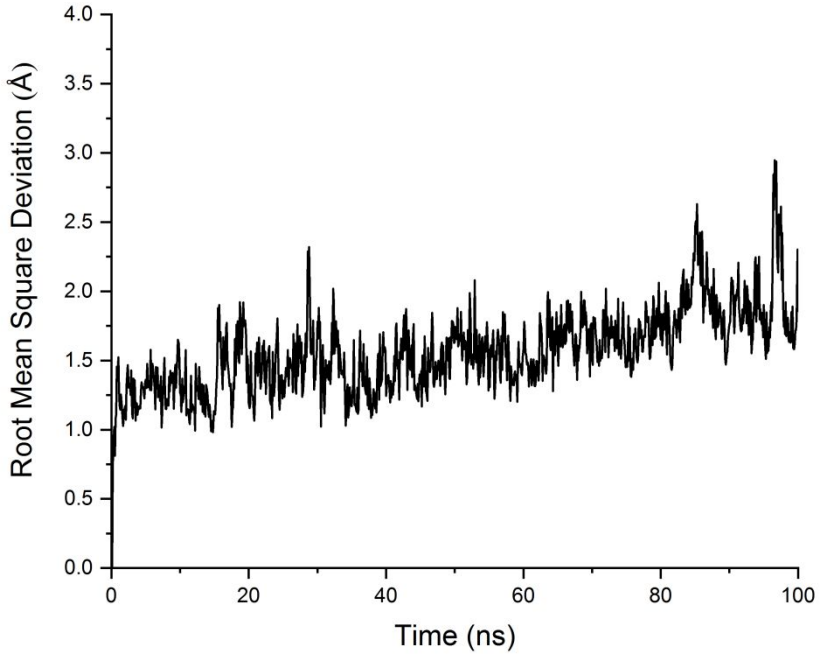


Figure S39. The time evolution of RMSD of mutant F143A in MD simulation.

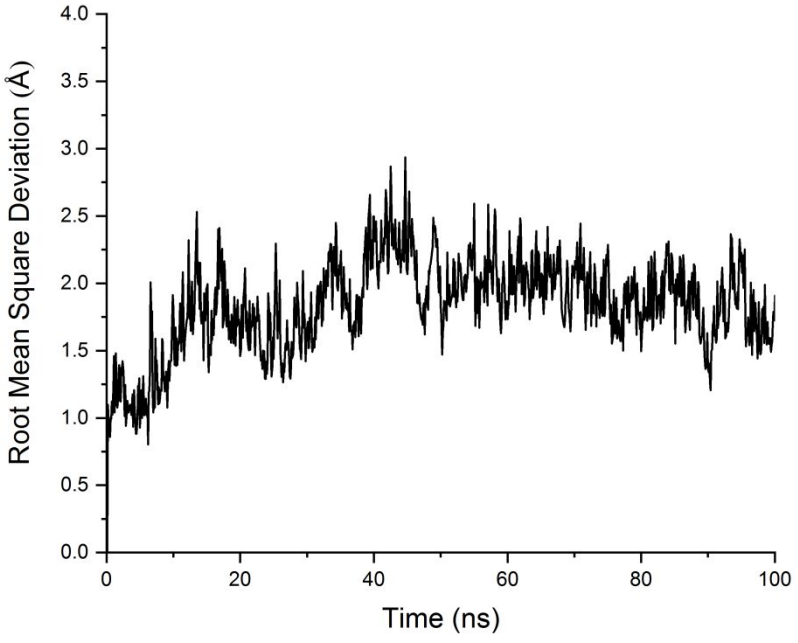


Figure S40. The time evolution of RMSD of mutant G146A in MD simulation.

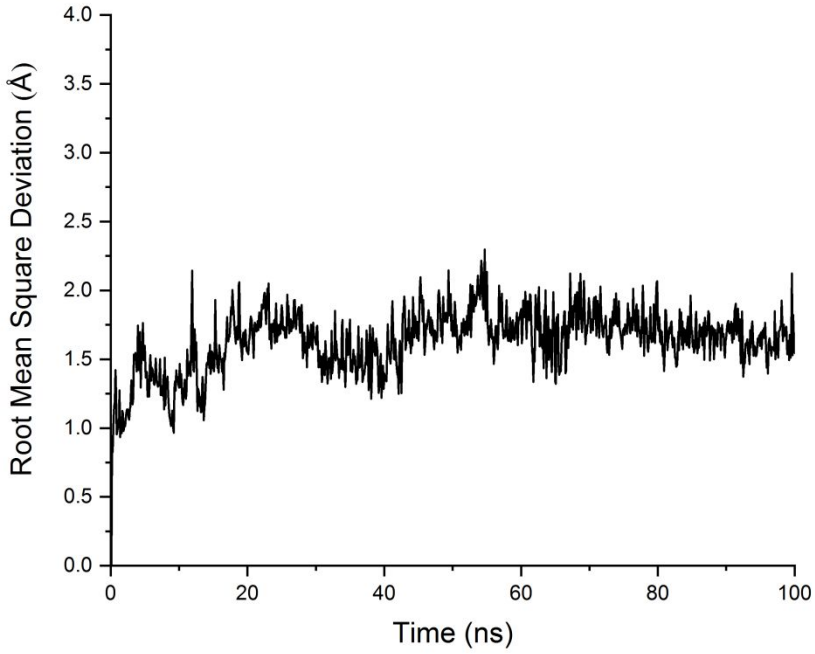


Figure S41. The time evolution of RMSD of mutant G146P in MD simulation.

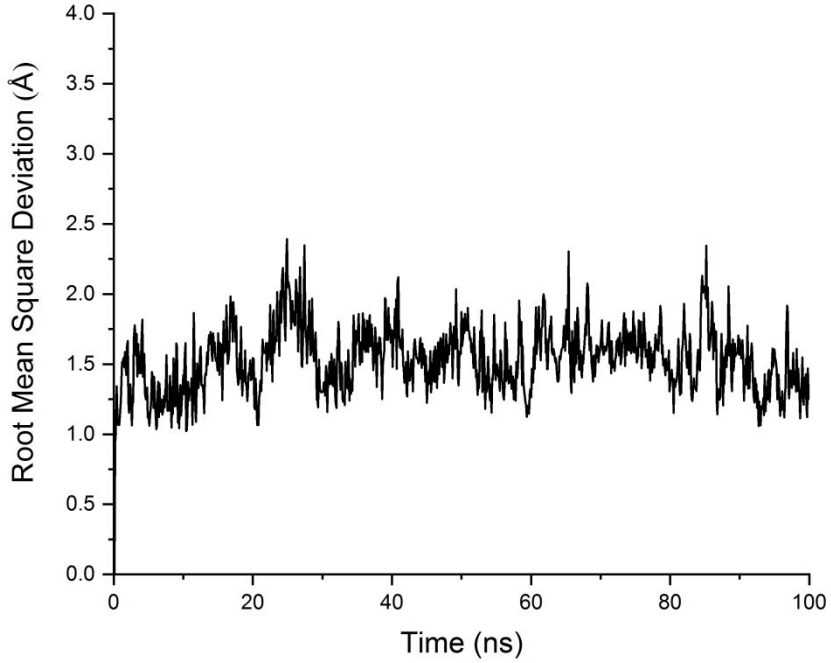


Figure S42. The time evolution of RMSD of mutant S147A in MD simulation.

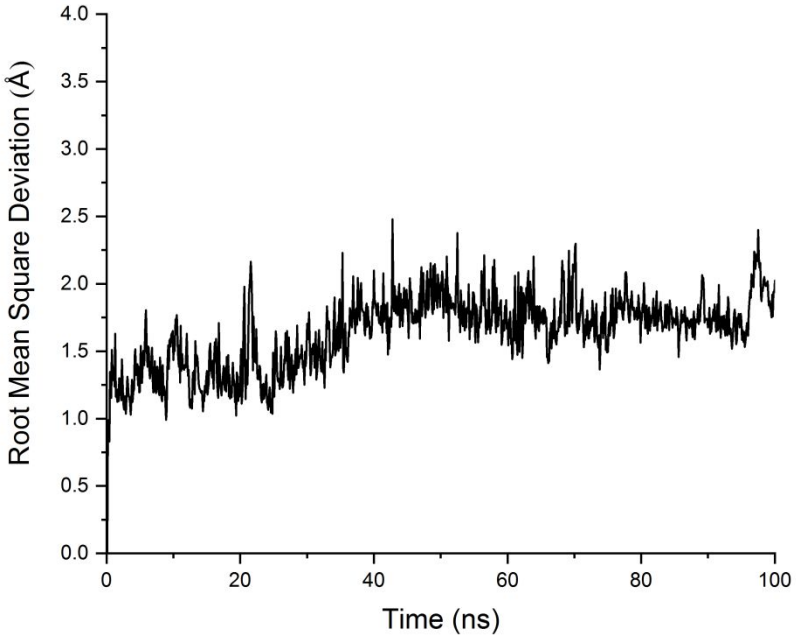


Figure S43. The time evolution of RMSD of mutant G149A in MD simulation.

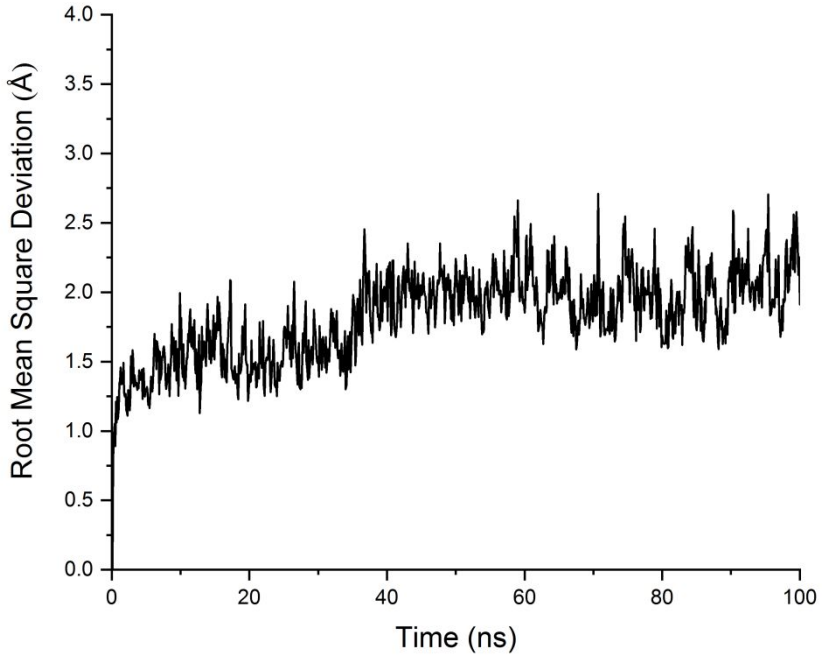


Figure S44. The time evolution of RMSD of mutant G149P in MD simulation.

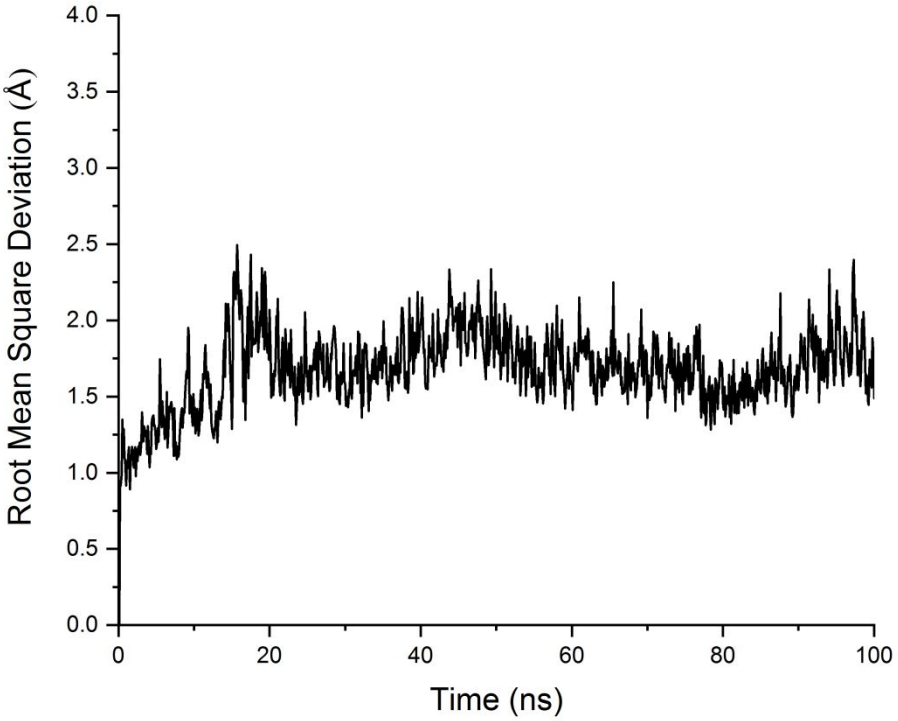


Figure S45. The time evolution of RMSD of mutant S150A in MD simulation.

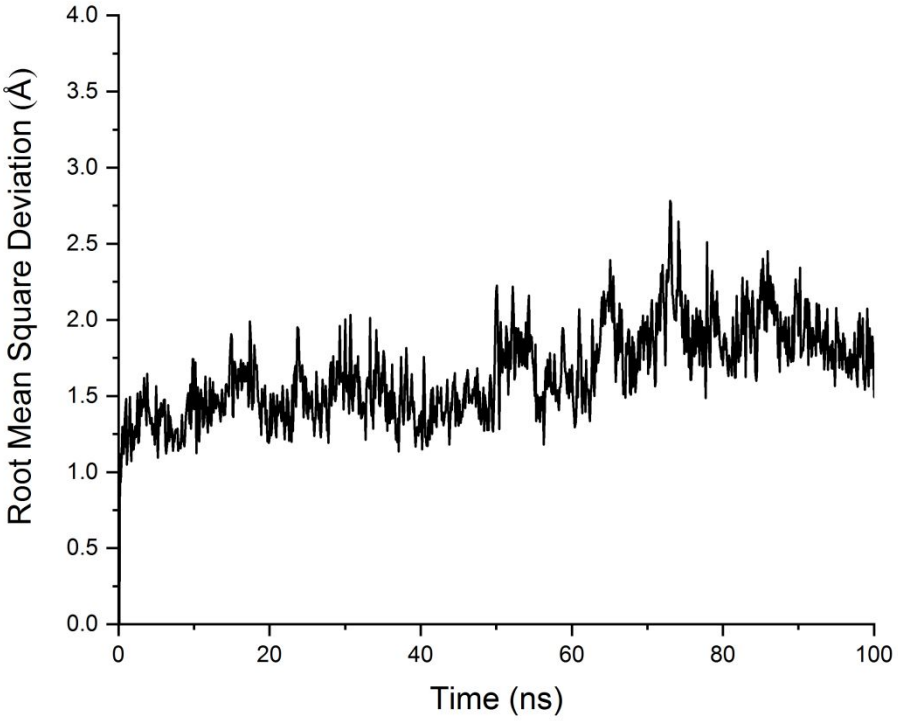


Figure S46. The time evolution of RMSD of mutant N28L in MD simulation:

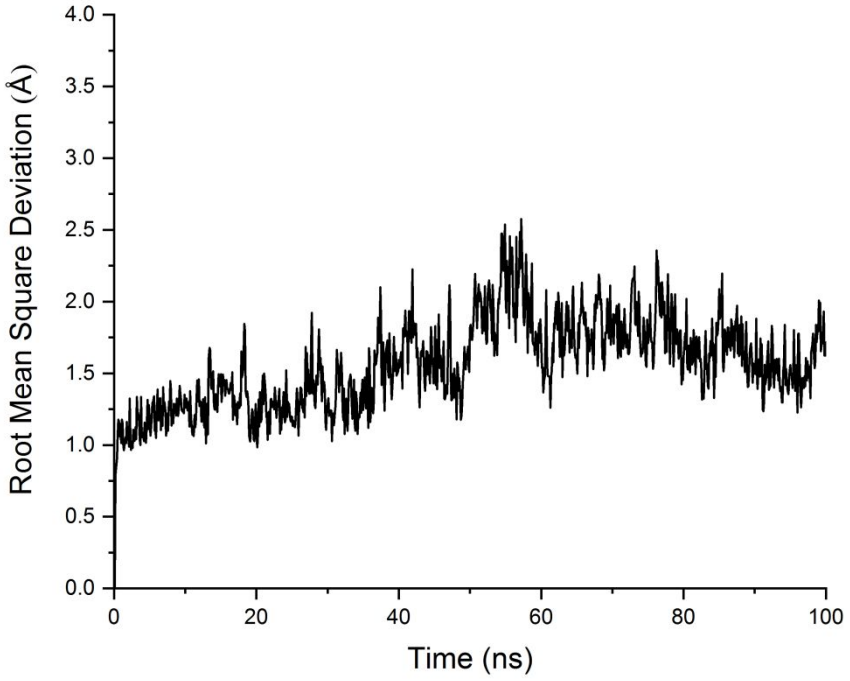


Figure S47. The time evolution of RMSD of wild types which was extracted water in MD simulation.

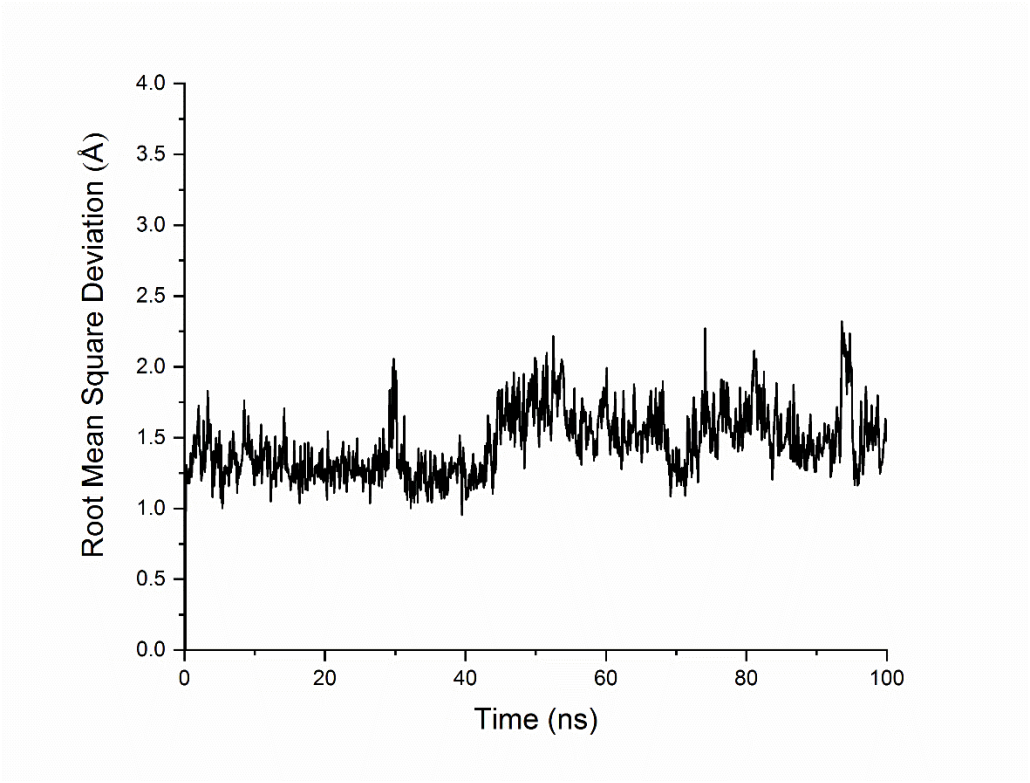


Figure S48.The time evolution of RMSD of mutant Q167A in MD simulation.

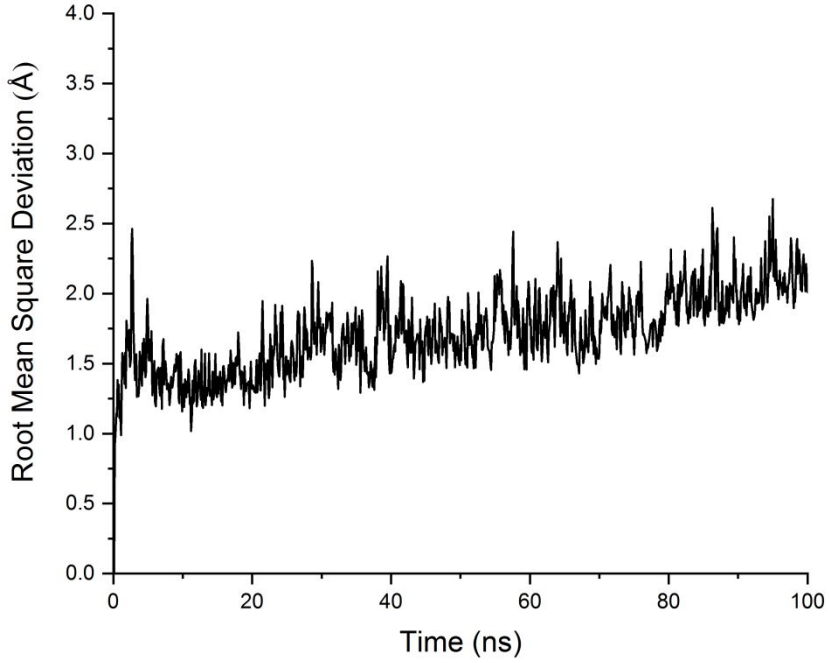


Figure S49.The time evolution of RMSD of mutant Q167V in MD simulation.

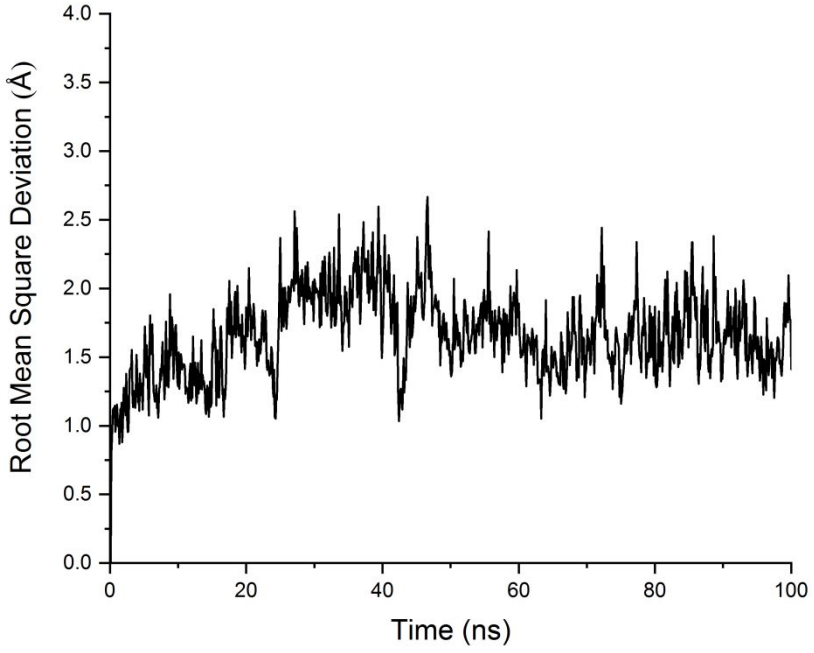


Figure S50.The time evolution of RMSD of mutant Q167L in MD simulation.

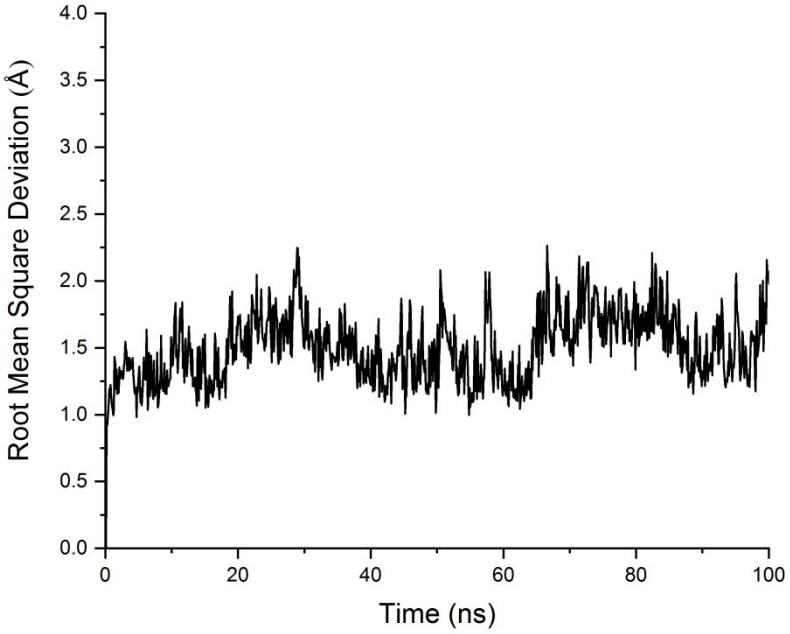


Figure S51.The time evolution of RMSD of mutant M168L in MD simulation.

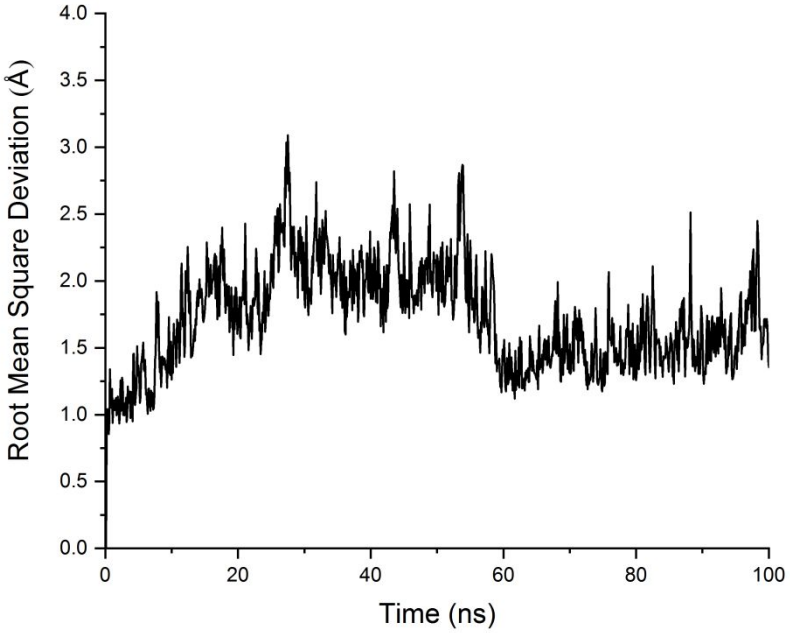


Figure S52.The time evolution of RMSD of mutant T88A in MD simulation.

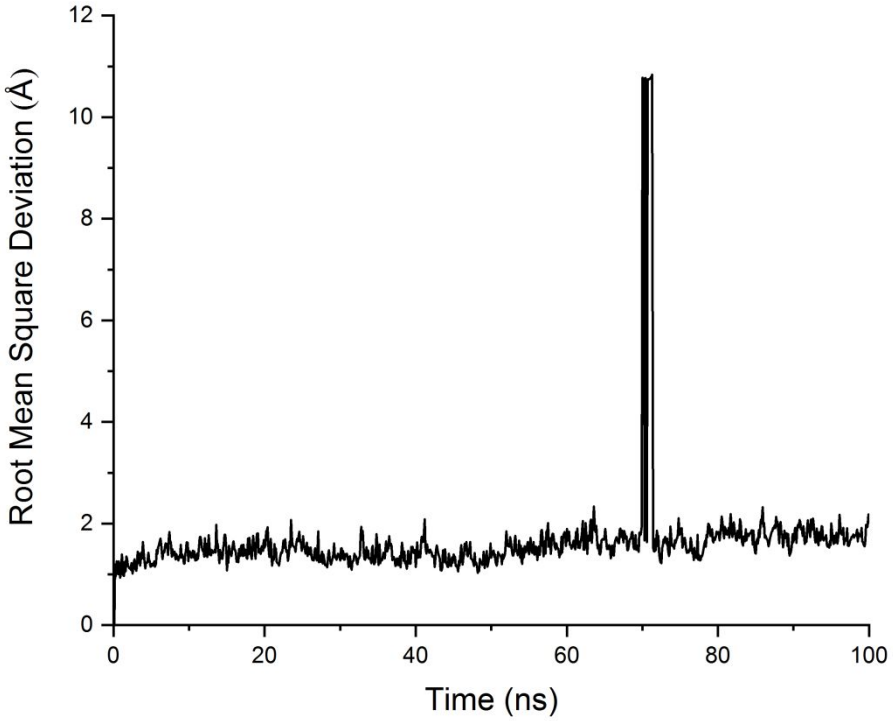


Figure S53.The time evolution of RMSD of mutant S178A in MD simulation.

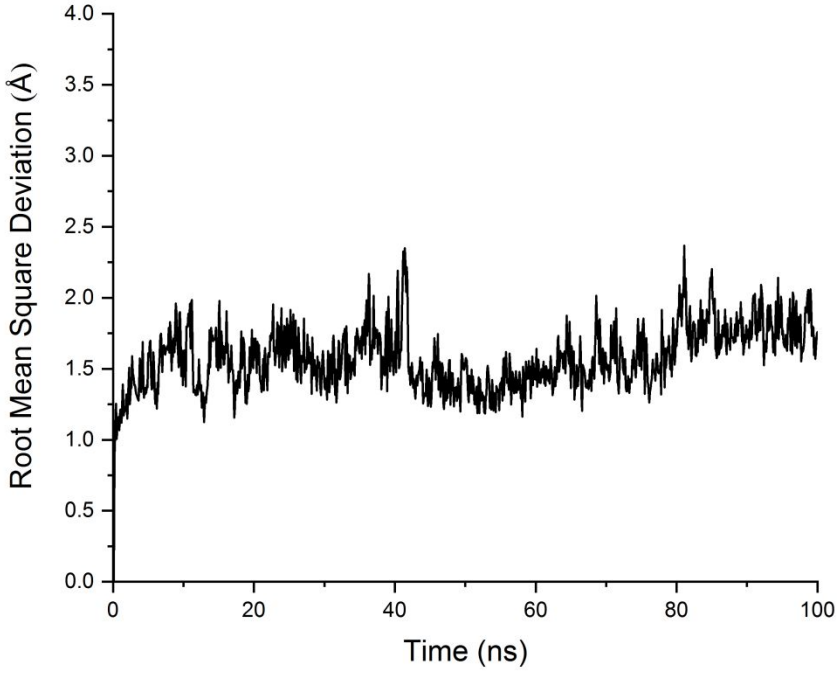


Figure S54. The time evolution of RMSD of mutant T88C/S178T in MD simulation.

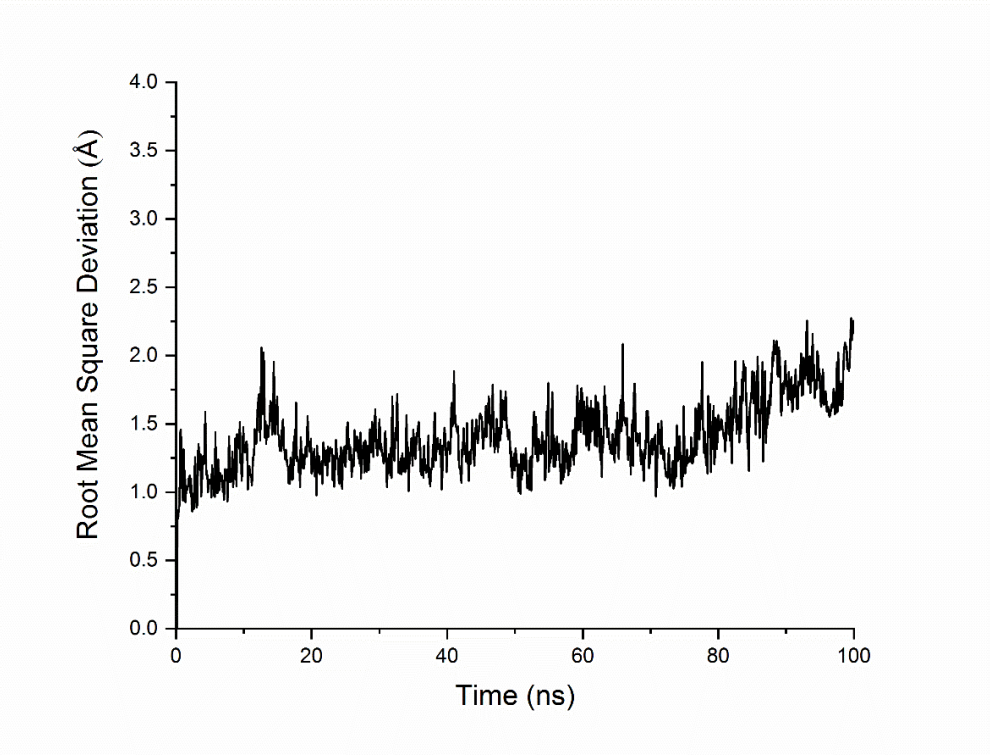


Figure S55. The time evolution of RMSD of MERS-CoV $3CL^{Pro}$ which binds to **12a** in MD simulation.

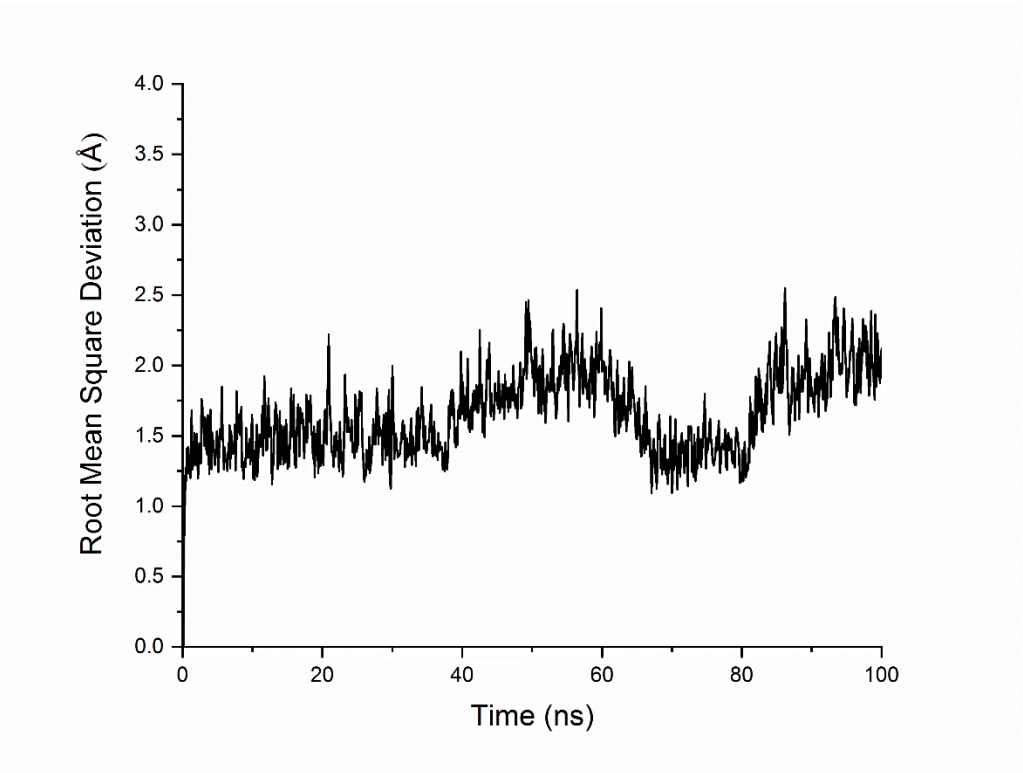


Figure S56. The time evolution of RMSD of MERS-CoV $3CL^{Pro}$ which binds to **12b** in MD simulation.

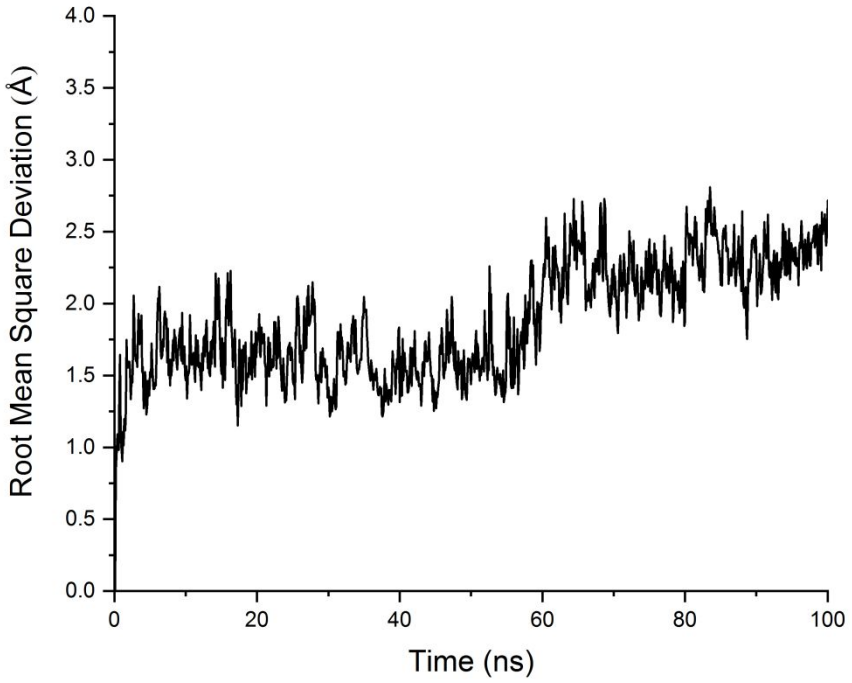


Figure S57.The time evolution of RMSD of wild type which was preconditioned before QM in MD simulation (first step of cleavage process).

References

1. Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, 22:2577-2637.
2. Frishman, D.; Argos, P., Knowledge-Based Protein Secondary Structure Assignment. *Proteins: Struct., Funct., Genet.* **1995**, 23:566-579.
3. Zhou, H.; Xue, B.; Zhou, Y., DDOMAIN: Dividing structures into domains using a normalized domain-domain interaction profile. *Protein Sci.* **2007**,16:947-955.
4. Pearl, F.; Todd, A.; Sillitoe, I.; Dibley, M.; Redfern, O.; Lewis, T.; Bennett, C.; Marsden, R.; Grant, A.; Lee, D.; Akpor, A.; Maibaum, M.; Harrison, A.; Dallman, T.; Reeves, G.; Diboun, I.; Addou, S.; Lise, S.; Johnston, C.; Sillero, A.; Thornton, J.; Orengo, C.; The CATH Domain Structure Database and related resources Gene3D and DHS provide comprehensive domain family information for genome analysis. *Nucleic Acids Res.* **2005**, 33: D247-D251.