Supplementary information

Enzymatic hydrolysis on L-azetidine-2-carboxylate ring opening

Xuexia Xu[b*§], Qin Yang[a,b*], Lanteng Wang[b*], Jie Zheng [b], Yang Gu [b], Xiwen Xing [a§], Jiahai Zhou [b§]

*same contribution

§ Correspondence to xuxx@ siat.ac.cn, xingxiwen@jnu.edu.cn and jiahai@siat.ac.cn

a. Guangdong Provincial Key Laboratory of Bioengineering Medicine, Department of Cell Biology, College of Life Science and Technology, Jinan University, Guangzhou 510632, China.

b. Shenzhen Key Laboratory for the Intelligent Microbial Manufacturing of Medicine, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.

Figures and tables

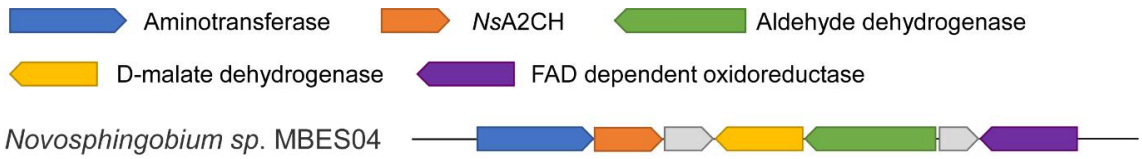

Figure S1. Gene cluster containing *Ns*A2CH encoded gene in *Novosphingobium sp*. MBES04.

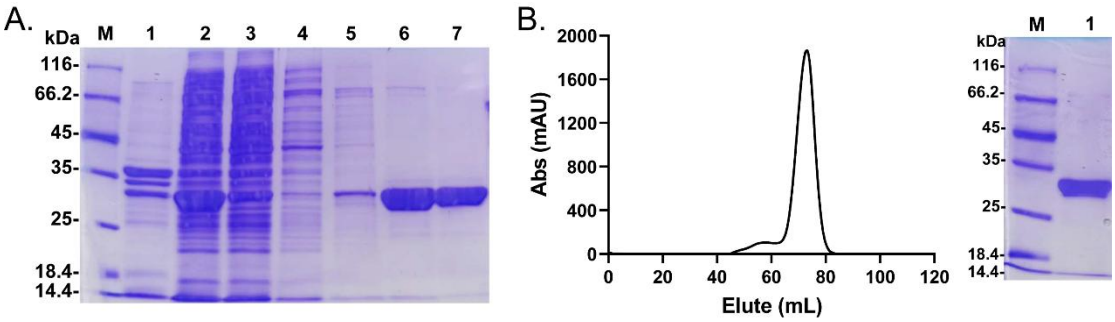

Figure S2. Protein purification and SDS-PAGE analysis of *Ns*A2CH. (A) SDS-PAGE after affinity chromatography. Lane M, protein marker; Lane 1, precipitation; Lane 2, cell Lysates; Lane 3, flow through; lanes 4-7, protein was eluted with 10, 50, 200, 200 mM imidazole, respectively. (B) Gel exclusion chromatography spectrogram and SDS-PAGE analysis. Lane M, protein marker; Lane 1, proteins collected in the peak tip (RT: 73.268 min) by size-exclusion chromatography.

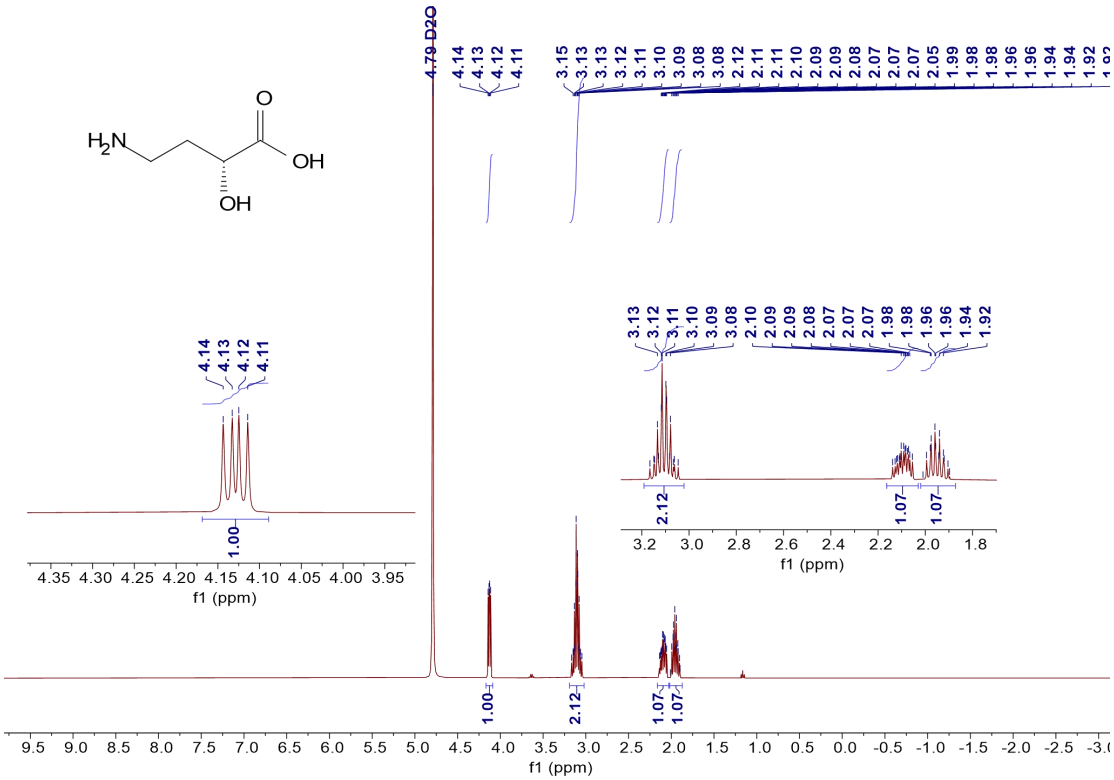

Figure S3. ^{1}H-NMR spectrum of (2*R*)-4-amino-2-hydroxybutanoic acid (**11b**).

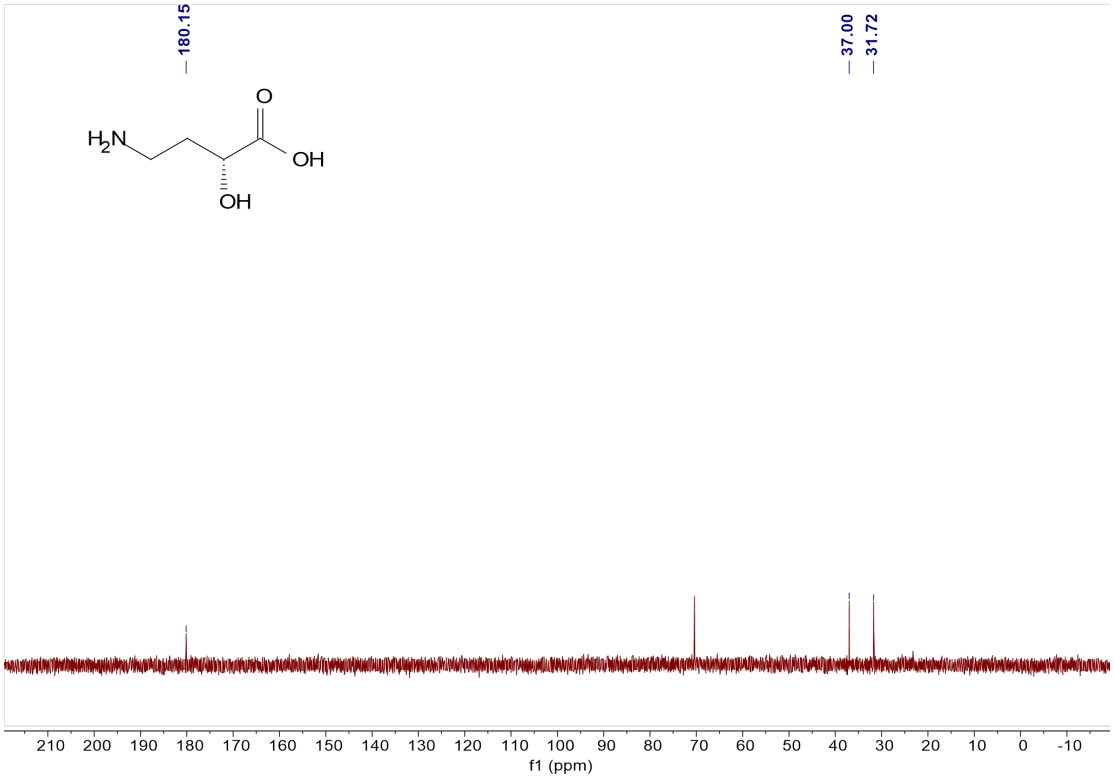

Figure S4. ^{13}C-NMR spectrum of (2*R*)-4-amino-2-hydroxybutanoic acid (**11b**).

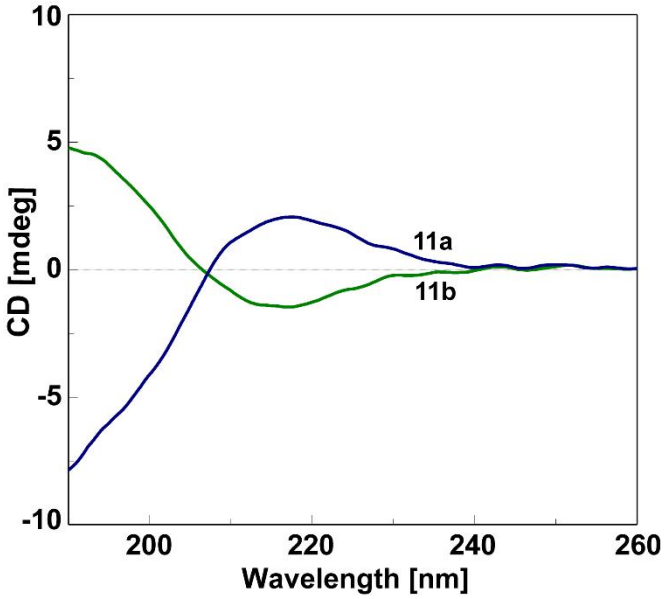

Figure S5. Circular dichroism (CD) spectrum of the reaction product 4-amino-2-hydroxybutanoic acid (**11b**). Blue line represents (*S*)-4-amino-2-hydroxybutanoic acid (**11a**) standard; green line represents product (**11b**).

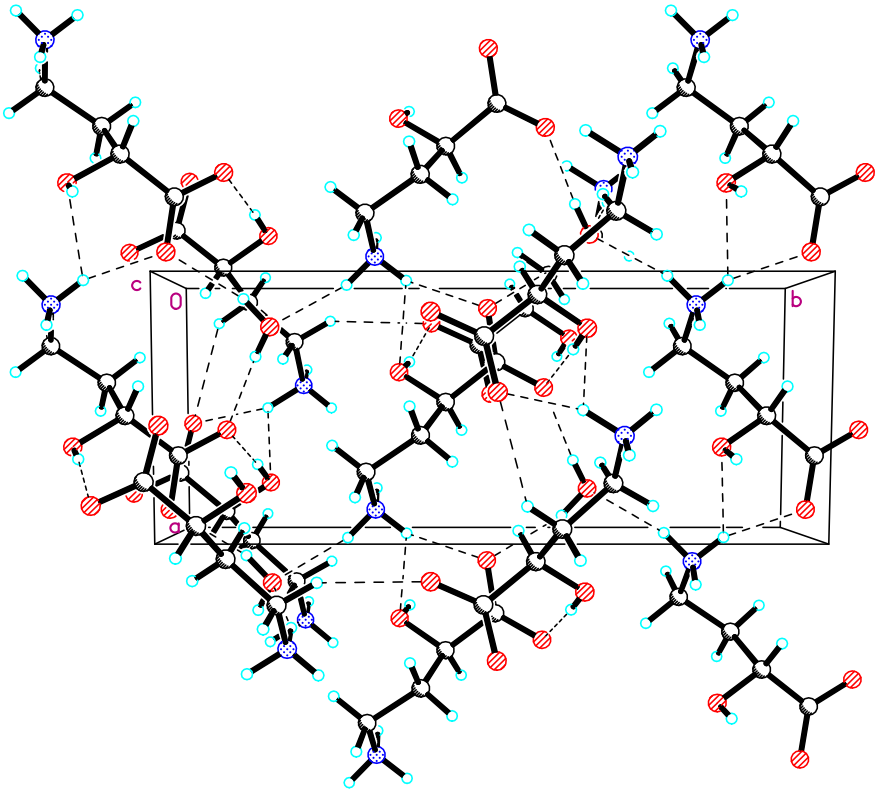

Figure S6. X-Ray Structure of the reaction product (2*R*)-4-amino-2-hydroxybutanoic acid. Projection of cell accumulation along the C axis.

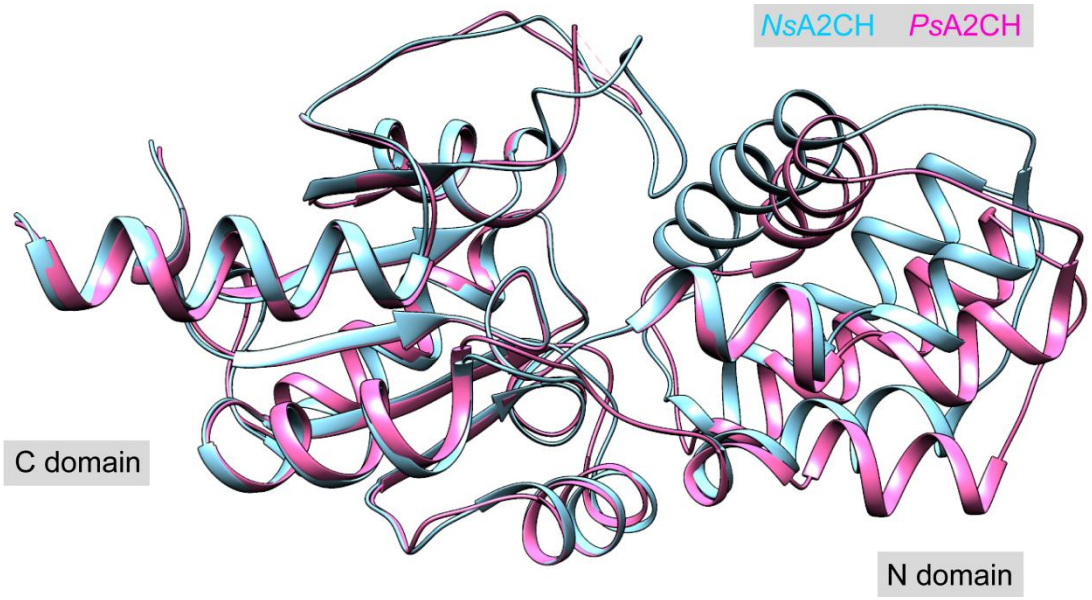

Figure S7. Structure alignment of WT *Ns*A2CH and *Ps*A2CH (PDB accession: 3SMV).

Both the N-terminal and the C-terminal domain of *Ns*A2CH can be well superimposed by the corresponding part of *Ps*A2CH (RMSD of 1.83 Å and 0.98 Å, respectively).

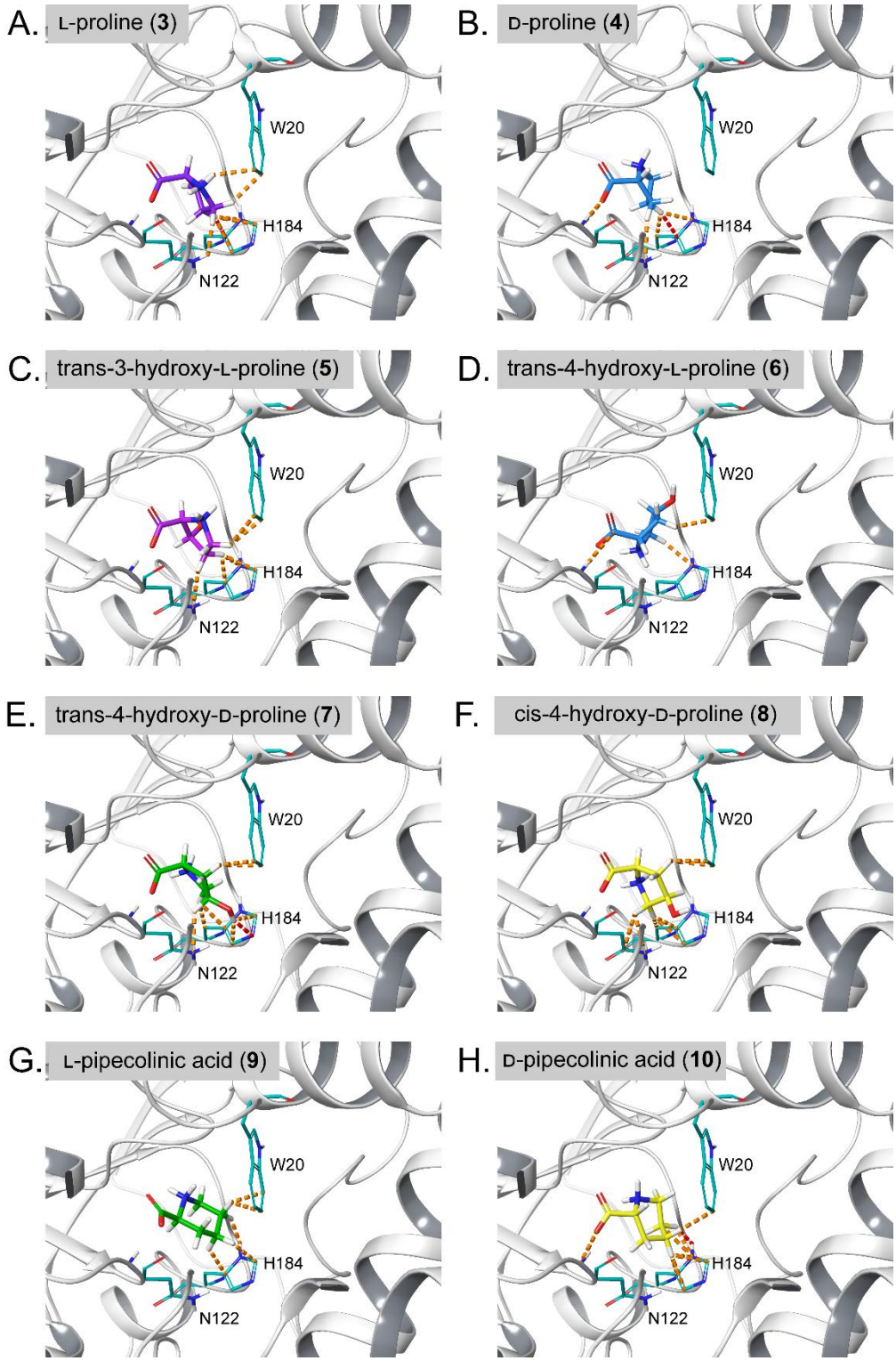

Figure S8. The docking results of *Ns*A2CH with test compounds. Compounds including L-proline (**3**), D-proline (**4**), trans-3-hydroxy-L-proline (**5**), trans-4-hydroxy-L-proline (**6**), trans-4-hydroxy-D-proline (**7**), cis-4-hydroxy-D-proline (**8**), L-pipecolinic acid (**9**), D-pipecolinic acid (**10**).

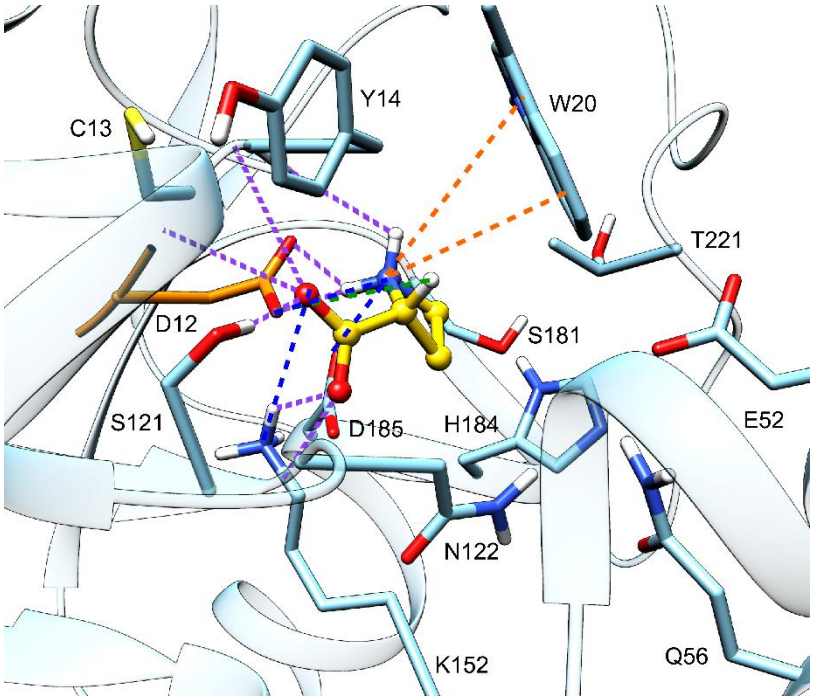

Figure S9. Molecular interactions in active sites of D-AZC docked *Ns*A2CH structure. Hydrogen bond, salt bridge and π-π stacking are labeled in purple, blue and green dot line, respectively.

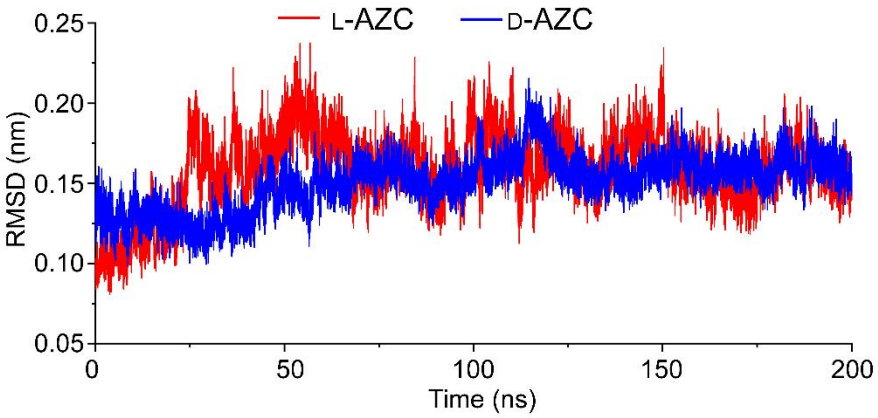

Figure S10. Plots of the Root Mean Square Deviation (RMSD) of all Cα atom over the 200 ns MD simulations of L-AZC docked and D-AZC docked *Ns*A2CH complex.

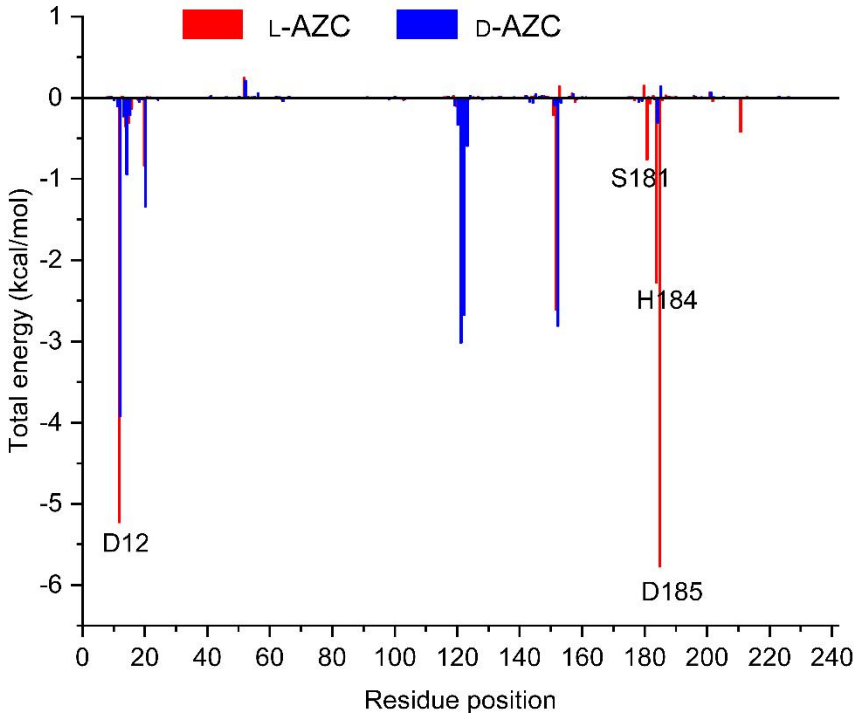

Figure S11. The total energy contribution of per amino acid residue in MMGBSA analysis for $_{L}$-AZC docked and $_{D}$-AZC docked *Ns*A2CH complex.

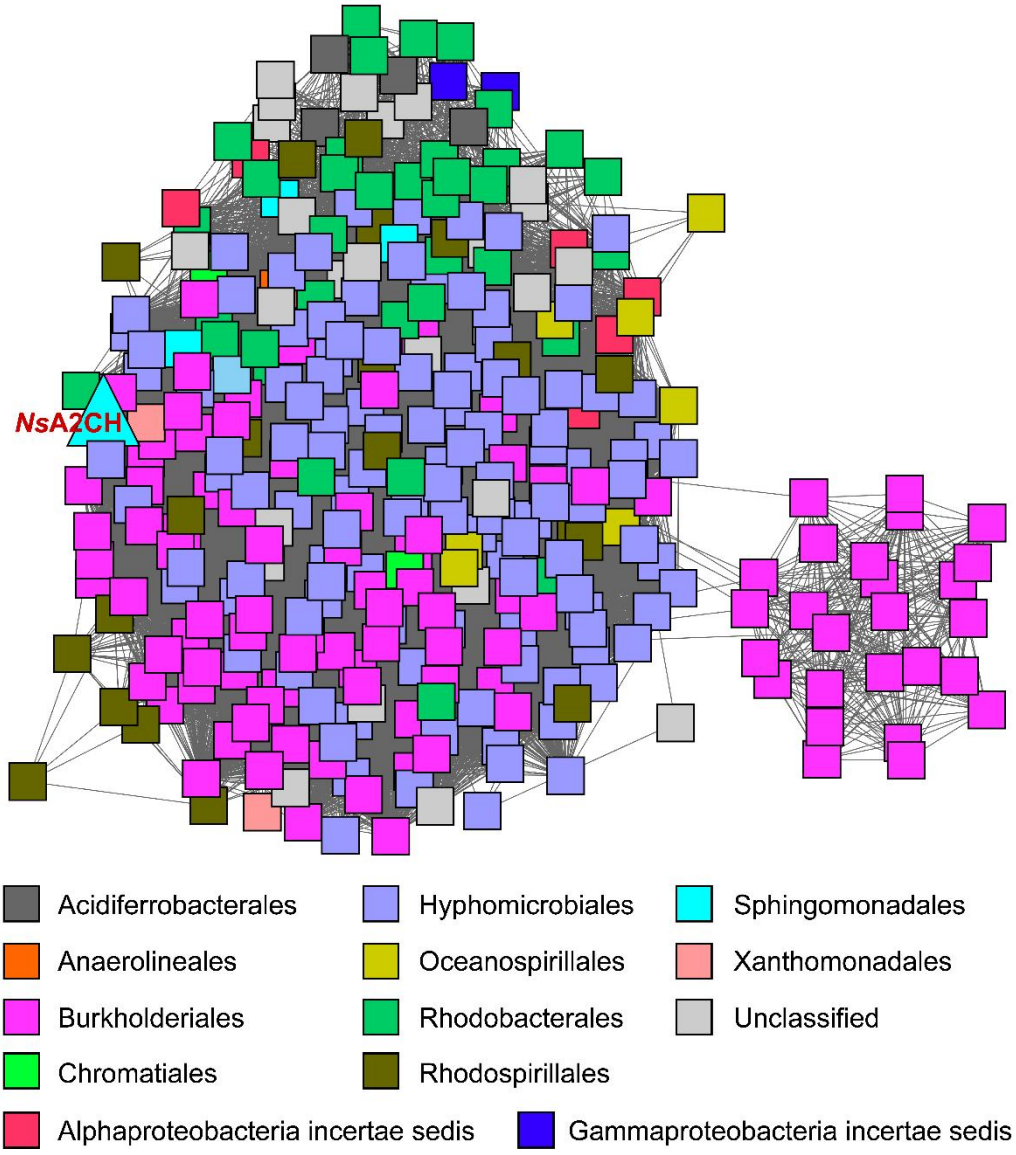

Figure S12. Sequence similarity network of the homologous L-AZC hydrolase (A2CHs) of *Ns*A2CH. The network is displayed at Evalue of 10^{-100}, where each edge represents sequences > 63% sequence identity. Sequences were colored by different Order.

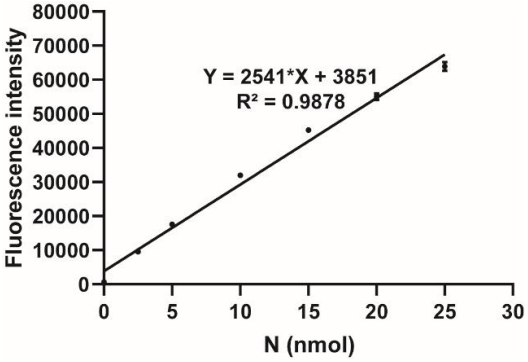

Figure S13. Standard curve of fluorescence intensity measured with different equivalents of 4-amino-2-hydroxybutanoic acid at λ_{ex}= 340 nm and λ_{em}= 455 nm. All determinations were performed in triplicate, and error bars represent the standard deviation.

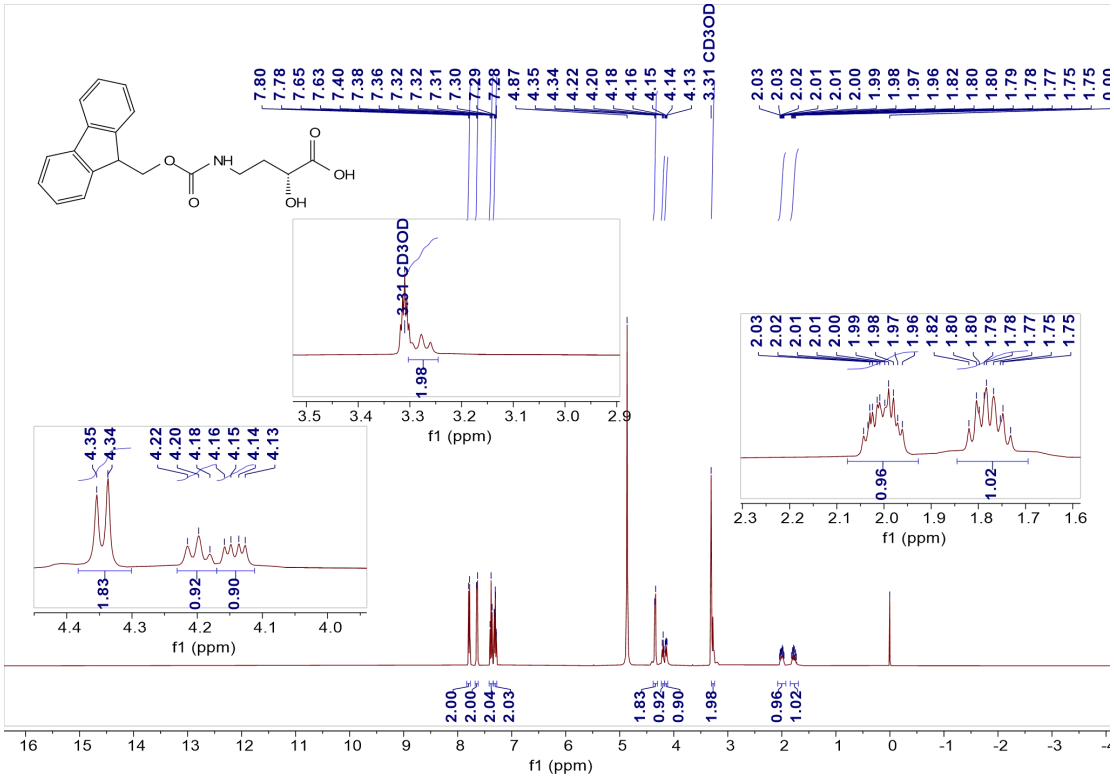

Figure S14. ^{1}H-NMR spectrum of (2*S*)-4-(9H-Fluoren-9-ylmethoxycarbonylamino)-2-hydroxybutyric acid (**12a**).

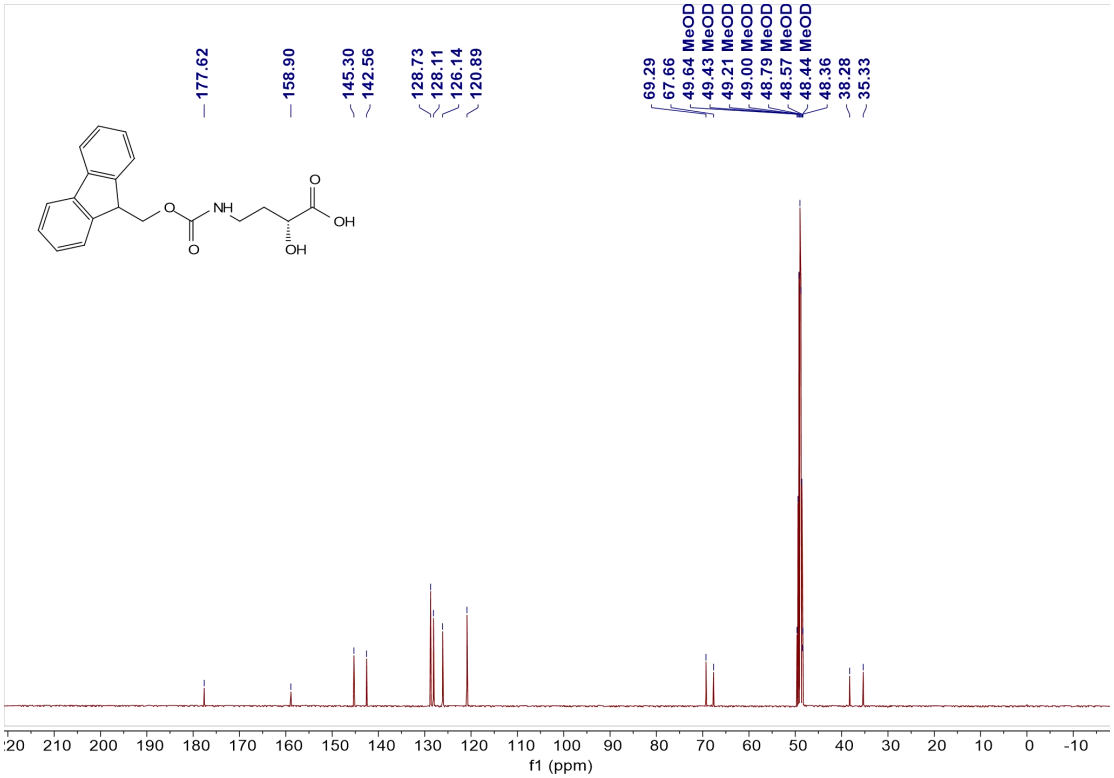

Figure S15. ^{13}C-NMR spectrum of (2*S*)-4-(9H-Fluoren-9-ylmethoxycarbonylamino)-2-hydroxybutyric acid (**12a**).

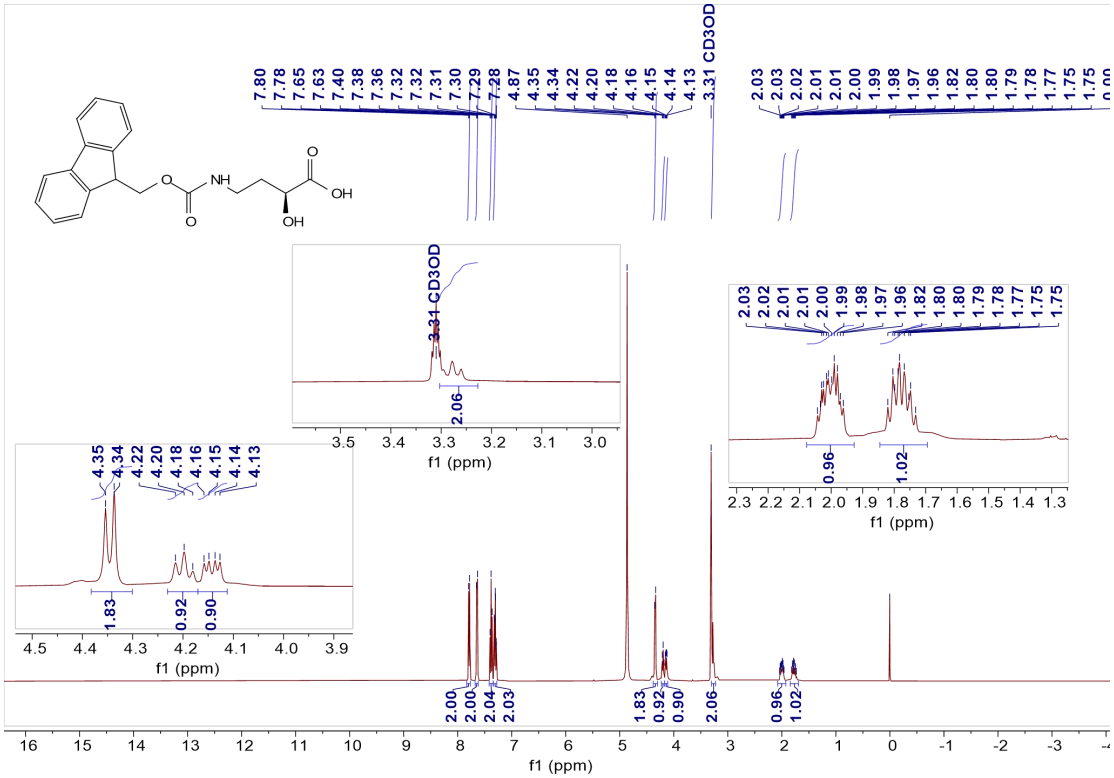

Figure S16. ^{1}H-NMR spectrum of (2*R*)-4-(9H-Fluoren-9-ylmethoxycarbonylamino)-2-hydroxybutyric acid (**12b**).

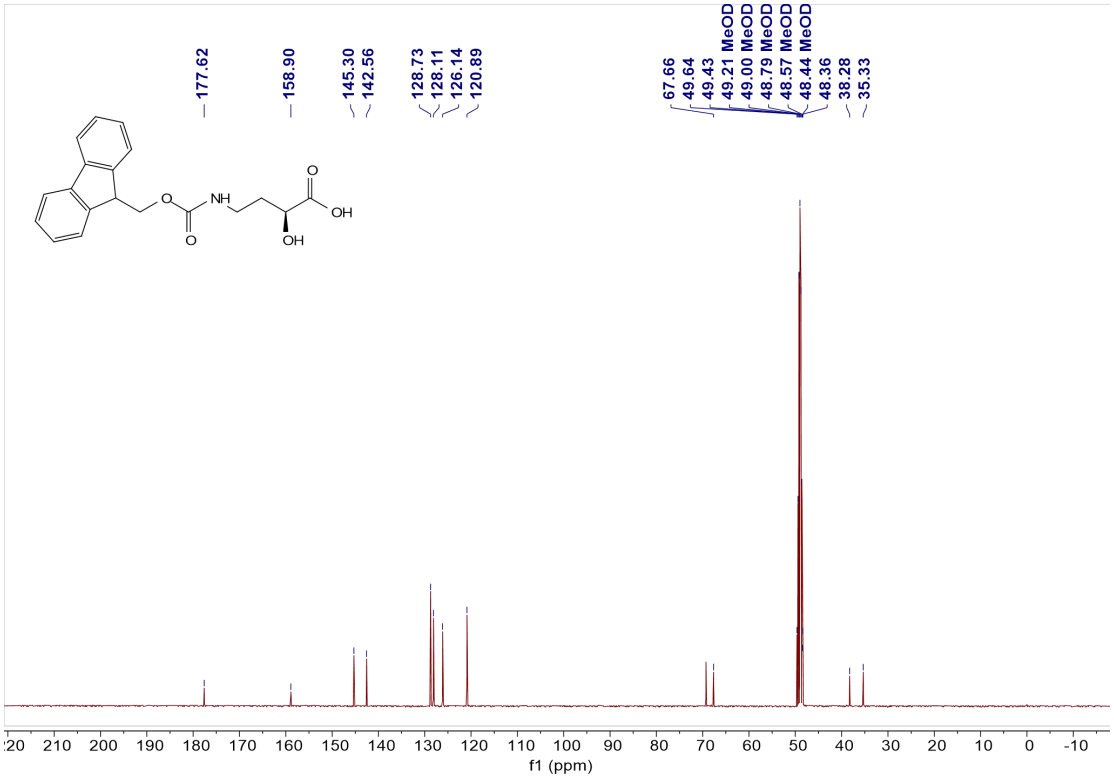

Figure S17. ^{13}C-NMR spectrum of (2*R*)-4-(9H-Fluoren-9-ylmethoxycarbonylamino)-2-hydroxybutyric acid (**12b**).

Table S1. Crystal data and structure refinement for (2*R*)-4-amino-2-hydroxybutyric acid.

Crystal Parameters	(2*R*)-4-amino-2-hydroxybutyric acid
Identification code	CCDC: 2225395
Empirical formula	C4 H11 N O4
Formula weight	137.14
Temperature	103(2) K
Wavelength	1.54178 Å
Crystal system, space group	Monoclinic, P2(1)
Unit cell dimensions	a = 5.8966(5) Å α = 90°
	b = 14.2652(11) Å β = 98.338(2)°
	c = 7.4117(6) Å γ = 90°
Volume	616.85(9) $Å^3$
Z, Calculated density	4, 1.477 Mg/m^3
Absorption coefficient	1.139 mm^{-1}
F(000)	296
Crystal size	0.30 x 0.05 x 0.05 mm^3
Theta range for data collection	6.03 to 67.90°
Limiting indices	-7<=h<=7, -17<=k<=17, -8<=l<=8
Reflections collected / unique	9343 / 2175 [R(int) = 0.0496]
Completeness to theta = 67.90°	98.3 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.7542 and 0.4550
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	2175 / 1 / 187
Goodness-of-fit on F^2	1.116
Final R indices [I>2sigma(I)]	R1 = 0.0337, wR2 = 0.0841
R indices (all data)	R1 = 0.0338, wR2 = 0.0842
Absolute structure parameter	0.07(15)
Largest diff. peak and hole	0.237 and -0.237 e.$Å^{-3}$

Table S2. Summary of the data collection and refinement statistics

Crystal Parameters	WT Apo	K152A complex
PDB ID	8HP5	8HP7
Space group	P 2_1 3	C 2
Cell dimensions		
a,b,c (Å)	151.8, 151.8, 151.8	146.0, 84.2, 102.4
α,β,γ (˚)	90, 90, 90	90, 133.9, 90
Data Collection		
Wavelength (Å)	0.979	0.979
Resolution (Å)	47.97-2.50 (2.60-2.50)[a]	41.02-1.43 (1.45-1.43)
No. of measured reflections	225499 (25797)	1069882 (52108)
No. of unique reflections	40320 (4529)	153210 (7519)
Completeness (%)	99.7 (100)	93.4 (93.1)
Multiplicity	5.6 (5.7)	7.0 (6.9)
Mean I/σ(I)	5.5 (2.6)	8.4 (3.5)
CC1/2	0.969 (0.788)	0.989 (0.919)
R_{merge} (%)[b]	19.5 (54.5)	17.3 (56.4)
Refinement		
Resolution range (Å)	42.09-2.50 (2.59-2.50)	33.26-1.43 (1.48-1.43)
Reflections used in refinement	40258 (4005)	152719 (15216)
R_{work}	0.166 (0.210)	0.193 (0.234)
R_{free} [c]	0.216 (0.282)	0.211 (0.257)
RMS bonds (Å)	0.008	0.006
RMS angles (°)	1.23	0.87
Protein residues	720	720
No. of non-hydrogen atom/ B-factor (Å^2)		
All/average	6324/32.06	6860/12.95
Macromolecules	5880/32.02	5944/11.80
Ligands	12/32.92	53/13.03
Solvent	432/32.61	863/20.90
Ramachandran plot		
Favored (%)	97.90	99.01
Allowed (%)	2.10	0.99
Outliers (%)	0	0

[a] Values in parentheses are for highest-resolution shell.

[b] $R_{merge} = \sum_{hkl}\sum_i|I_i - \langle I\rangle|/\sum_{hkl}\sum_i|\langle I\rangle|$, where I_i is the intensity for the *i*th measurement of an equivalent reflection with indices h, k, and l.

[c] R_{free} was calculated with the 5% of reflections set aside randomly throughout the refinement

Table S3. Primers used in alanine-scanning mutagenesis.

Name	Sequence (5' to 3')
D12A-F	GCACTGAGCTTTGCTTGCTATGGCACCC
D12A-R	CAGGGTGCCATAGCAAGCAAAGCTCAGTGC
W20A-F	GCACCCTGATTGATGCGGAAAGTGGTATGATTG
W20A-R	CAATCATACCACTTTCCGCATCAATCAGGGTGC
E52A-F	GCACATGCACGTCATGCAAGCCGTCAGCAGGCCC
E52A-R	GGGCCTGCTGACGGCTTGCATGACGTGCATGTGC
Q56A-F	CATGAAAGCCGTCAGGCGGCCCAGAC
Q56A-R	GTCTGGGCCGCCTGACGGCTTTCATG
S121A-F	CTGATCATCCTGGCCAACGTGGATAATAAGACC
S121A-R	GGTCTTATTATCCACGTTGGCCAGGATGATCAG
N122A-F	CTGATCATCCTGAGCGCCGTGGATAATAAGACC
N122A-R	GGTCTTATTATCCACGGCGCTCAGGATGATCAG
K152A-F	GCGCCGAAGATGTGGGCGCATATGCACCGAGTGATCG
K152A-R	CGATCACTCGGTGCATATGCGCCCACATCTTCGGCGC
S181A-F	CTGCATACCGCAGAAGCCCTGTTTCATGATCATG
S181A-R	CATGATCATGAAACAGGGCTTCTGCGGTATGCAG
H184A-F	GCAGAAAGCCTGTTTGCTGATCATGTGCCGGC
H184A-R	GCCGGCACATGATCAGCAAACAGGCTTTCTGC
D185A-F	GCCTGTTTCATGCTCATGTGCCGGCACG
D185A-R	CGTGCCGGCACATGAGCATGAAACAGGC
T211A-F	GGTTTTGGTGCAGCCATGACCCCGAGCC
T211A-R	GGCTCGGGGTCATGGCTGCACCAAAACC

Table S4. Chiral HPLC conditions and retention times for (2*S*)/(2*R*)-4-amino-2-hydroxybutyric acid.

Sample	Flow rate (mL/min)	T (ºC)	Column	Eluent[a]	RT (min)
12a	0.5	30	Chiral MJ (2)	n-hexane- /ethanol (0.1%TFA) 80:20	21.313 (*S*)
12b					14.710 (*R*)
12a+**12b**					14.787 (*R*); 21.278 (*S*)

[a] Experiments were performed with isocratic eluent.