Electronic Supplementary Material (ESI)

Significantly enhancing the stereoselectivity of a regioselective nitrilase for production of (*S*)-3-cyano-5-methylhexanoic acid using MM/PBSA method

Zhi Chen, Hualei Wang, Lin Yang, Shuiqing Jiang, Dongzhi Wei**

State Key Laboratory of Bioreactor Engineering, New World Institute of Biotechnology, East China University of Science and Technology, Shanghai200237, China. *Correspondence: Hualei Wang (hlwang@ecust.edu.cn), Dongzhi Wei (dzhwei@ecust.edu.cn).

Supplementary information

Table of contents

1. Materials.....3
2. Saturation mutagenesis and combinational mutagenesis.....3
3. Expression and purification of enzymes.....4
4. Biocatalytic reactions of *rac*-ISBN.....5
5. Analytical method.....5
6. Autodock analysis and computational details.....6
7. Supplementary Scheme, Tables and Figures.....7

1. Materials

The KOD-Plus-Mutagenesis Kit was purchased from TOYOBO Co., Ltd. (Japan) was utilized for all the mutant constructions. Substrate of racemic isobutylsuccinonitrile (ISBN) was synthesized by Aurora Phama Co., Ltd. (Shanghai, China). Isopropyl-β-D-thiogalactopyranoside (IPTG) and kanamycin were purchased from Sigma Co., Ltd. The recombinant strain harboring nitrilase gene *bll6402* from *Bradyrhizobium japonicum* USDA 110 was constructed by our group (named bll6402NIT) and stored in our lab. All the primers utilized in the present study were synthesized by Shanghai Generay Biotech Co., Ltd (Shanghai, China) and listed in Table S2-3. The other materials were all commercially available.

2. Saturation mutagenesis and combinational mutagenesis

All mutants of bll6402NIT were produced based on vector pET-28(a) with an N-terminal 6×His tag. The site-saturation mutagenesis (NNT) as well as site-directed mutagenesis was carried out utilizing the KOD-Plus-Mutagenesis Kit with a three-step reverse PCR strategy. A pair of primers for mutagenesis formation displayed in Table S1 was composed of one mutagenic forward primer P1 and one non-mutagenic reverse primer. Firstly, reverse PCR reaction mixtures (50 μl) consisted of ddH_2O (35μl), 10*PCR buffer (5 μl), 2 mM dNTPs (5 μl), forward and reverse primers (both 1.5 μl), template DNA (50-100 μg) and KOD-Plus- DNA polymerase (1μl) were performed utilizing the following procedure: 94 °C for 2 min, (98 °C for 10 s, 68 °C for 7 min) × 7 cycles, 4 °C hold. The resulting PCR mixtures were then digested at 37 °C for 1 h utilizing *Dpn* I in order to remove the template DNA. Finally, ligation reaction was conducted at 37 °C for 1 h utilizing the following reaction system (total 15 μl): PCR products digested by *Dpn* I (2 μl), ddH_2O (7 μl), Ligation high (7 μl) and T4 Polynucleotide Kinase (1 μl). The resulting ligation products were subsequently transformed into compenent cells and mutants with correctly sequencing were selected and subjected to enzyme expression.

3. Expression and purification of enzymes

For the expression of bll6402NIT mutants, the resulting recombinant *E. coli* cells were cultivated in Luria–Bertani liquid medium containing 50 mg/ml kanamycin at 37°C on a rotary shaker at 220 ×g. A final concentration of 0.1mM isopropyl-β-D-thiogalactoside was added for the induction when the optical density at 600 nm of the culture broth reached between 0.6 and 0.8. The cells were then further incubated at 37°C and 220 ×g for another 4-6 h. After centrifugation at 8,000 ×g for 10 min, the cells were harvested and preserved at -20 °C for 2h and then were lyophilized for 12 h. The cells lyophilization powder were stored at -20°C for subsequent experiments.

For the purification of bll6402NIT mutants, Nickle affinity chromatography (Histrap HP column, GE Healthcare UK Ltd) was applied to purify the recombinant nitrilase by exploiting the histidine tag. The obtained cells were suspended and washed twice with 10 ml of sodium phosphate buffer (20 mM, pH 7.4). Then the cells were resuspended in 20 ml of the same buffer and disrupted by sonication on ice at 200 W for 10 min. The soluble fractions of the sonicated solution were obtained by centrifugation at 8,000 ×g for 30 min to remove the cell debris. The resulting supernatant was passed through a 0.22 μm filter, and then loaded onto a Ni-NTA column previously equilibrated with a binding buffer (20 mM sodium phosphate buffer, 500 mM NaCl, 10 mM imidazole, pH 7.4). The column was subsequently washed with 10 ml of wash buffer (20 mM sodium phosphate buffer, 500 mM NaCl, 50, 100 and 250 mM imidazole, pH 7.4) sequentially to wipe out the non-target proteins and eluted with the elution buffer (20 mM sodium phosphate buffer, 500 mM NaCl, 500 mM imidazole, pH 7.4). The purified enzymes were further ultrafiltrated-concentrated and then analyzed by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Protein concentration was determined using the Bradford method with bovine serum albumin as the standard.[1] All purification steps were performed at 4 °C. Purified proteins were further analyzed by circular dichroism (Applied Photophysics, UK).

4. Biocatalytic reactions of *rac*-ISBN

The reactions were carried out in standard reaction mixtures (1 ml) consisting of sodium phosphate buffer (100 mmol/L, pH 7.5), 10 g/L (dry cell weight, DCW) and 50 mM substrate ISBN containing 5 % methol *(v/v*) as cosolvent. The reaction samples were conducted at 30 °C in a rotary shaker at 200 rpm for 12 h. The residual substrate was firstly extracted by ethyl acetate and upper organic lawyer was collected for chiral GC analysis after being dried by anhydrous sodium sulfate. Aqueous solution in lower layer was subsequently acidified by the addition of 50 μl of 6 M HCl. The organic layer in which the target product was located was dried again by anhydrous sodium sulfate and then for chiral GC analysis.

5. Analytical method

Chiral GC analysis was performed on an Agilent 7200 gas chromatography equipped with FID and chiral capillary column CP-Chirasil-DEXCB column (30 m×0.25 mm ×0.25 μm) with nitrogen as the carrier gas. The column flow rate was 1.5 mL/min. The temperatures of the injector and detector were set at 200 °C. The column temperature was programmed as follows: 120 °C (0 min)–10 °C/min–150 °C (20 min)–40 °C/min–180 °C (5 min). Enantiomeric excess (e.e) of product and substrate were calculated via the standard method.[2]

Mutants with the highest e.e value of (*S*)-CMHA were selected and product (*S*)-CMHA was further purified and conducted NMR and MS (ESI) analysis. NMR spectra were recorded on Bruker Avance III (Bruker, Switzerland) operating at 400 MHz for ^{1}HNMR and 101 MHz for ^{13}C NMR acquisitions, respectively with $CDCl_3$ as the solvent. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 3.12 – 2.94 (m, 1H), 2.84 – 2.67 (m, 1H), 2.64 – 2.35 (m, 2H), 1.72 – 1.56 (m, 1H), 1.43 – 1.16 (m, 2H), 1.04 – 0.78 (m, 6H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 77.23, δ 174.22, 121.26, 40.89, 37.00, 26.09 (d, J = 48.4 Hz), 23.10, 21.45. High-resolution MS spectra (ESI) were recorded on XEVO G2 TOF (Waters, USA). MS (ESI): m/z 154.0867 $[M-H]^-$.

6. Autodock and computational details

The homology model of bll6402NIT was obtained from our previous publication.[3] All the molecular docking experiments in the present study were performed utilizing Autodock4.2 program suite.[4] Structure of substrate was prepared and subsequently MM2 minimized by Chem3D sofeware in Chem office15 suite. The dock performance was carried out with default parameters except for the number of runs 100. The grid box was set as a cube with the number of points in x,y,z dimension 40 and utlizing the Cys x,y,z spatial coordinates as the center. The docking results were prepared by Open Babel and visualized by Pymol edu. Classical docking studies were firstly utilized for obtaining the enzyme-substrate complexes. Then molecular dynamics simulation was performed on the complexes and the more reliable complexes conformation was obtained. Based on the structures of complexes conformation, residues around the ligands with 5 Å were finally selected for the binding free energy calculation by using the MM-PBSA method. The potential hotspot sites were molecularly modified and the ideal results were obtained.

Molecular dynamics simulation was performed by Amber 12 with Amber ff99SB force field.[5, 6] The Generalized Amber Force Field (GAFF) was utilized to cope with the ligands.[7] The docked enzyme-ligand complexes were solvated in the system of a TIP3P water box. Sodium ions were then added to neutralize the system charge. A steepest descents minimization algorithm was utilized to perform energy minimization. The system was set to the equilibration phases for 100 ps using the NVT ensemble with a continuous temperature of 300 K and subsequently 100 ps using the NPT ensemble with the same temperature. Molecular dynamics simulation was run for 10 ns to conduct RMSD and Dc-s analysis. The binding free energy (ΔG_{bind}) were calculated by utilizing the molecular mechanics-poisson-boltzmann surface area (MM-PBSA) method with mm pbsa procedure in Amber 12, which is a classical and more suitable method to calculate the free energy distinctions between two states when compared with other methods.[8] The per-residue binding free energy decomposition was subsequently performed to calculate the binding free energies

($\Delta G_{ligand\text{-}residue}$) on the selected residues (5 Å of the ligand) to identify the residues responsible for potential effects on tendency of nitrilase-substrate binding.

7. Supplementary Scheme, Tables and Figures

TableS1 The calculated binding free energy (ΔG_{bind}) based on MM/PBSA method.

Enzyme	Substrate	ΔG_{bind} (kcal/mol)
WT	*R*-ISBN	-13.5861
	S-ISBN	-14.8072
W57F/V134M	*R*-ISBN	-12.3079
	S-ISBN	-18.4183
W57Y/V134M	*R*-ISBN	-12.4410
	S-ISBN	-16.2271

Table S2 Primers used for site-directed mutagenesis of wild type bll6402NIT.

Primers	Sequence (5'-3') *
52Y-F	GGCNNKCCTTGGTGGCTTTGGCTCG
52Y-R	CGGAATCCAGACCTCCGGGAACGC C
57W-F	CTTNNKCTCGGGACGCCGGCTTGGG
57W-R	CCACCAAGGATAGCCCGGAATCCA G
134V-F	CACNNKGAACGTACGCTCTATGGC G
134V-R	TGTCGGCTTGAGCTTTCGGCGCTTG
135E-F	GTCNNKCGTACGCTCTATGGCGAAG

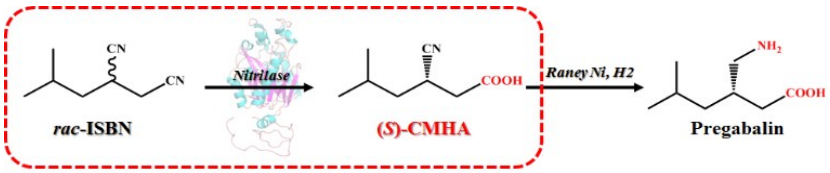
135E-R	GTGTGTCGGCTTGAGCTTTCGGCGC

*The underline parts represent the sites for saturated mutation.

Table S3 Primers used for combinational mutagenesis of mutants W57F and W57Y.

Primers	Sequence (5'-3') *
V134A-F	CACGCCGAACGTACGCTCTATGGC
	G
V134M-F	CACATGGAACGTACGCTCTATGGC
	G
V134Q-F	CACCAGGAACGTACGCTCTATGGC
	G
V134C-F	CACTGGGAACGTACGCTCTATGGC
	G
V134A/M/Q/C-R	TGTCGGCTTGAGCTTTCGGCGCTTG

*The underline parts represent the sites that 134Val mutated to Ala. Met, Gln and Cys using the recombinant plasmids of mutants W57F and W57Y as template.



Scheme S1 Nitrilase-mediated biocatalysis toward ***RAC*-ISBN** for preparation of pregabalin intermediate, (*S*)-3-cyano-5-methylhexanoic acid ((*S*)-CMHA).

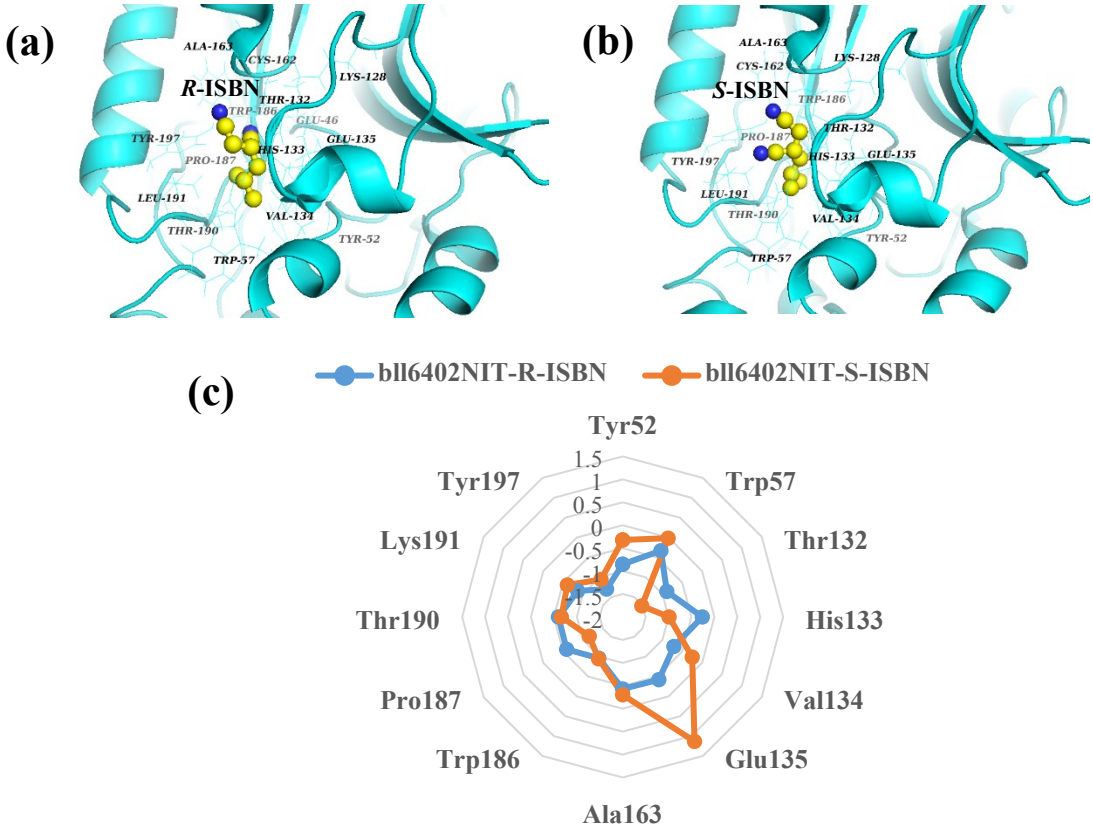


Figure S1 Residues around 5Å of 6402WT and *R*-ISBN complex **(a)** and 6402WT and *S*-ISBN complex **(b)** in the homology model of bll6402NIT were selected. Binding free energies ($\Delta G_{ligand\text{-}residue}$) **(c)** on the selected residues calculated by MM/PBSA method.

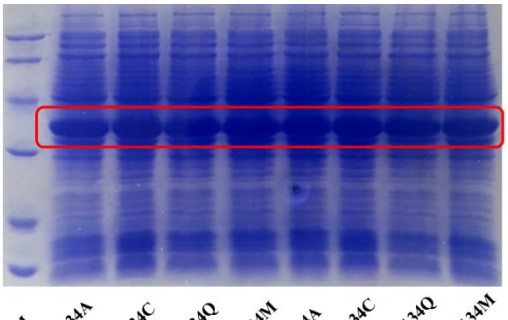


Figure S2 SDS-PAGE analysis of W57F, W57Y, V134A, V134C, V134Q and V134M

combinational mutants.

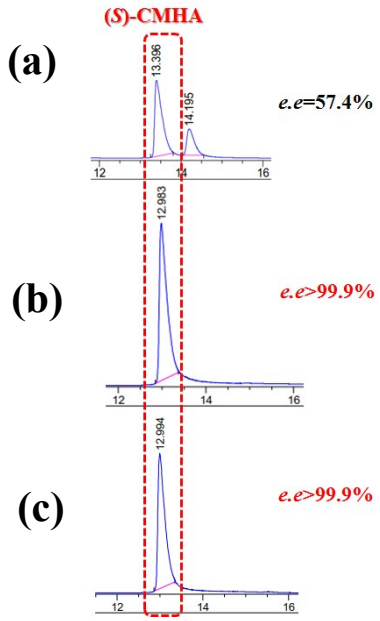
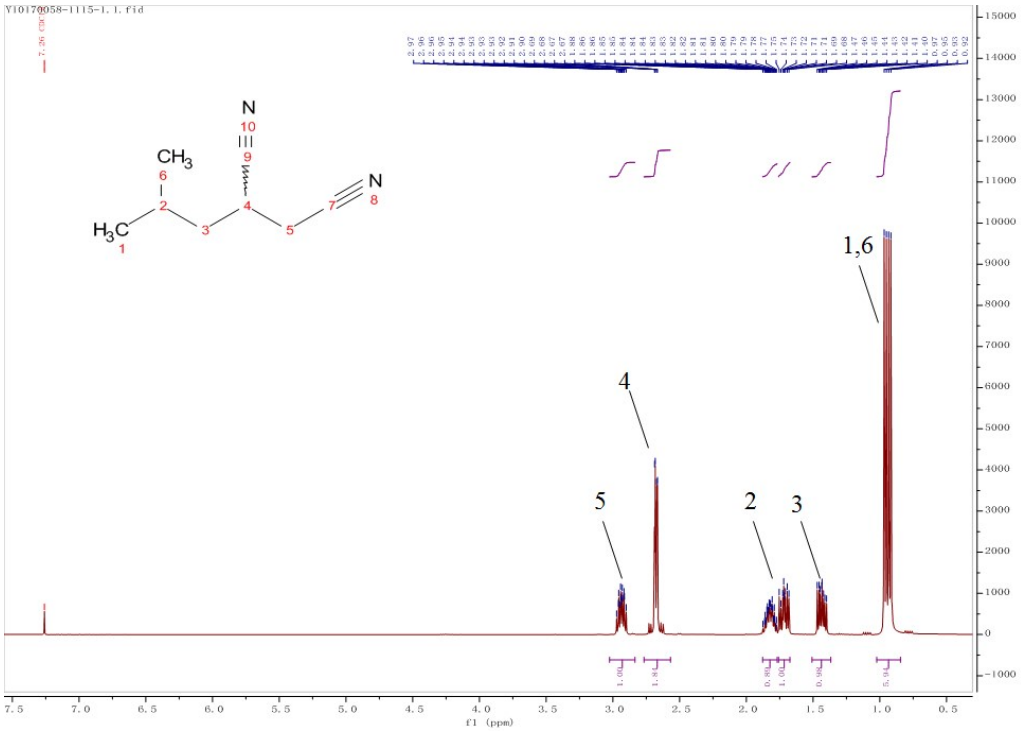


Figure S3 Biocatalytic products of wild type bll6402NIT **(a)**, W57F/V134M **(b)** and W57Y/V134M **(c)** detected by chiral GC.

(a)


Y10170058-1115-1.1.fid
N
10
CH3
6
9
N
8
2
4
7
H3C
1
3
5
1,6
4
5
2
3
f1 (ppm)

(b)

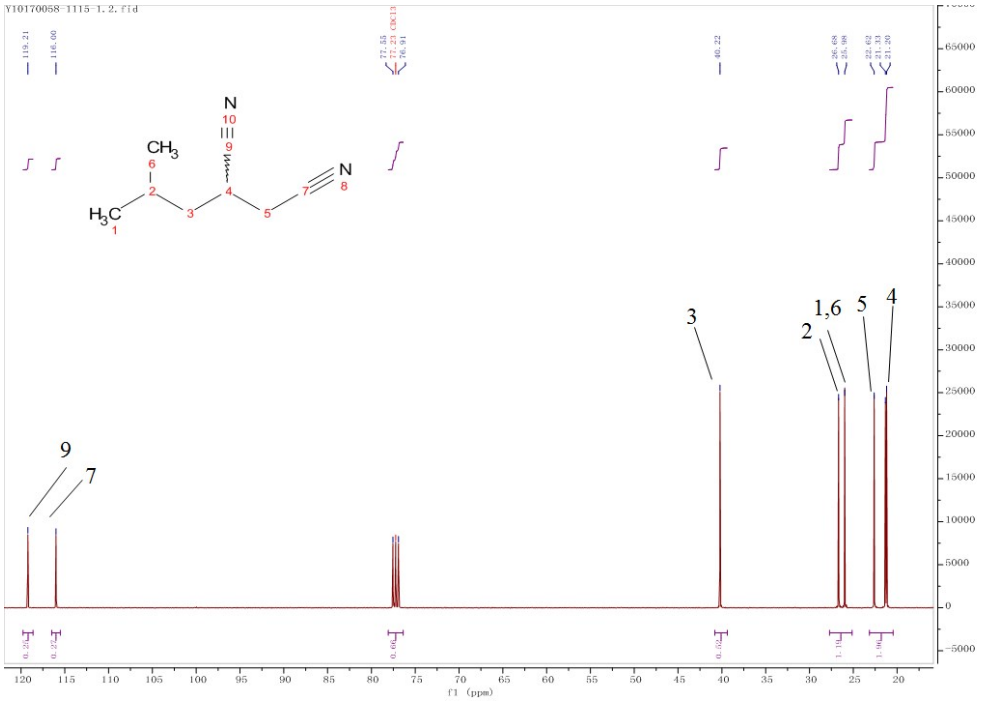
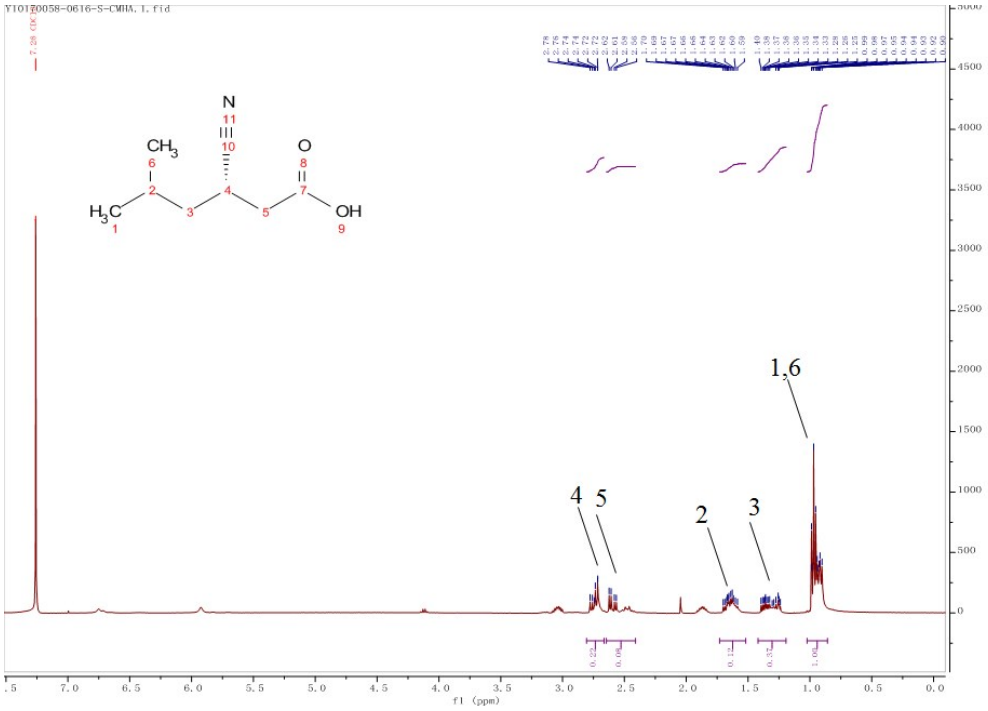
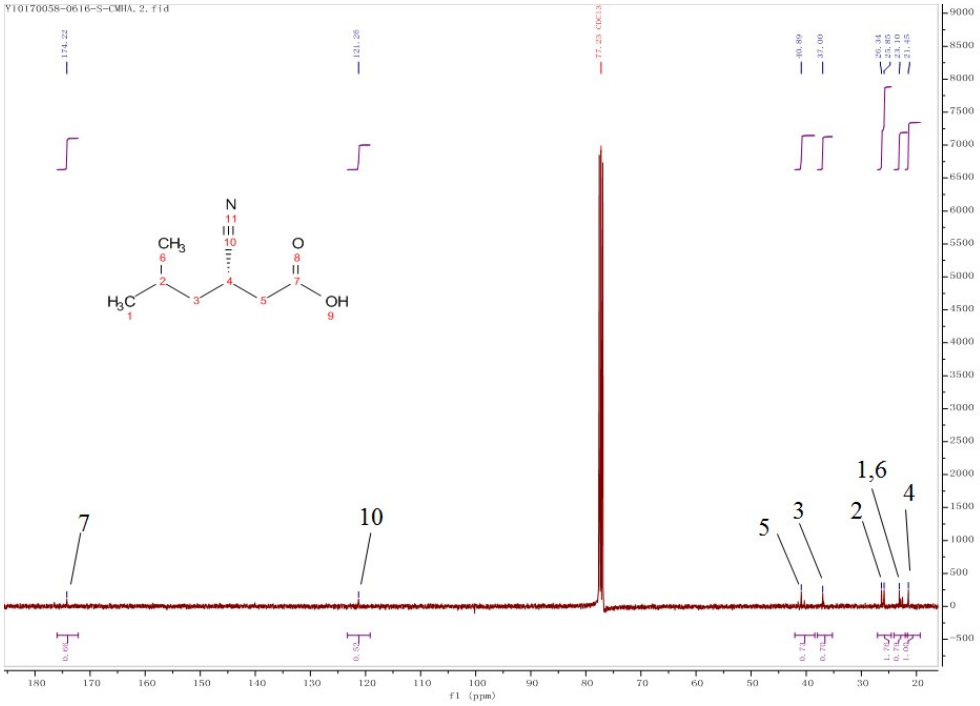

Y10170058-1115-1.2.fid
119.21
116.00
77.33
77.12
CDCl3
76.91
40.22
26.68
25.98
22.62
21.33
21.20
N
10
CH3
6
9
N
8
2
4
7
H3C
1
3
5
3
2
1,6
5
4
9
7
f1 (ppm)

Figure S4 ^{1}H NMR **(a)**, ^{13}C NMR **(b)** spectra of substrate isobutylsuccinonitrile (ISBN). ^{1}H NMR (400 MHz, Chloroform-*d*) 2.94 (dtd, *J* = 10.5, 6.4, 5.4 Hz, 1H), 2.74 – 2.60 (m, 1H), 1.89 – 1.77 (m, 1H), 1.72 (ddd, *J* = 13.5, 10.5, 5.1 Hz, 1H), 1.43 (ddd, *J* = 13.5, 9.2, 5.4 Hz, 2H), 0.94 (dd, *J* = 13.6, 6.6 Hz, 6H).

(a)



(b)



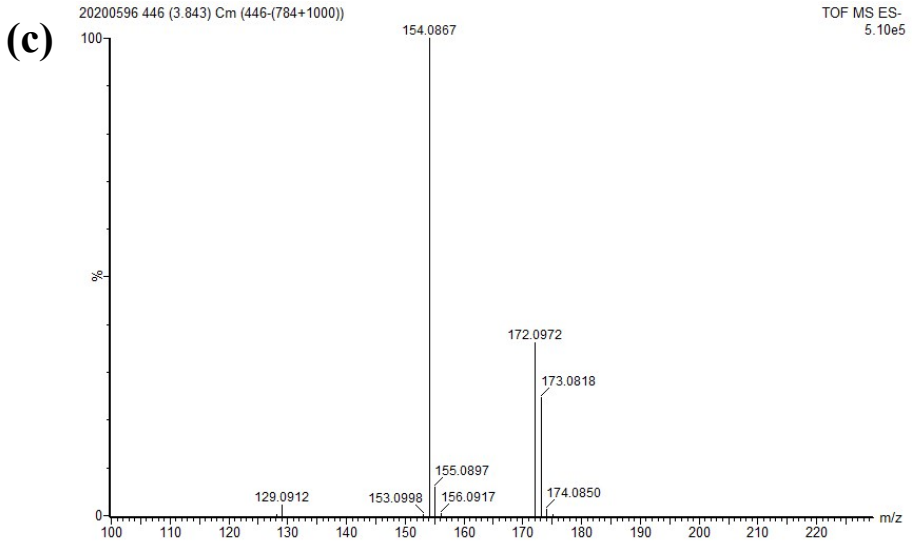


Figure S5 ^{1}H NMR (a), ^{13}C NMR (b) and MS (ESI) spectra (c) of product (*S*)-3-cyano-5-methylhexanoic acid ((*S*)-CMHA). MS (ESI): m/z=154.0867 $[M-H]^-$.

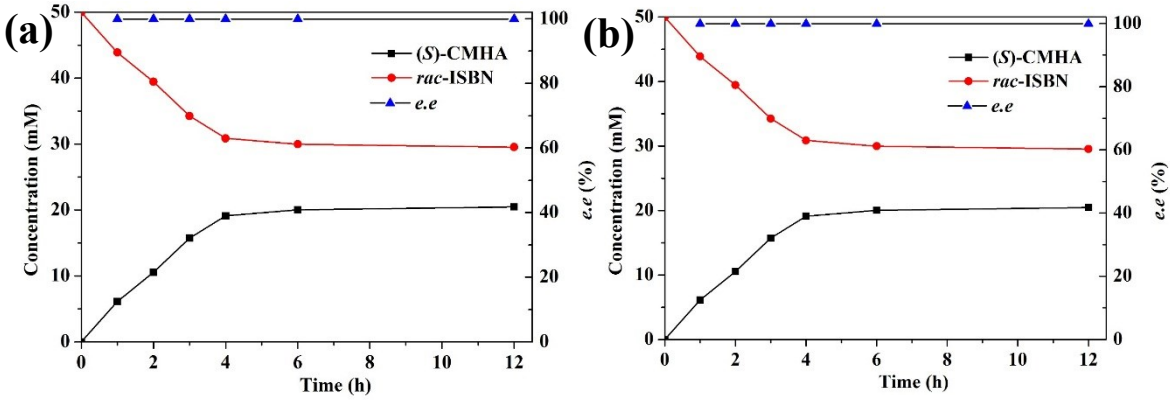


Figure S6 Time course hydrolysis of *rac*-ISBN by whole cells of the mutants W57F/V134M (a) and W57Y/V134M (b).

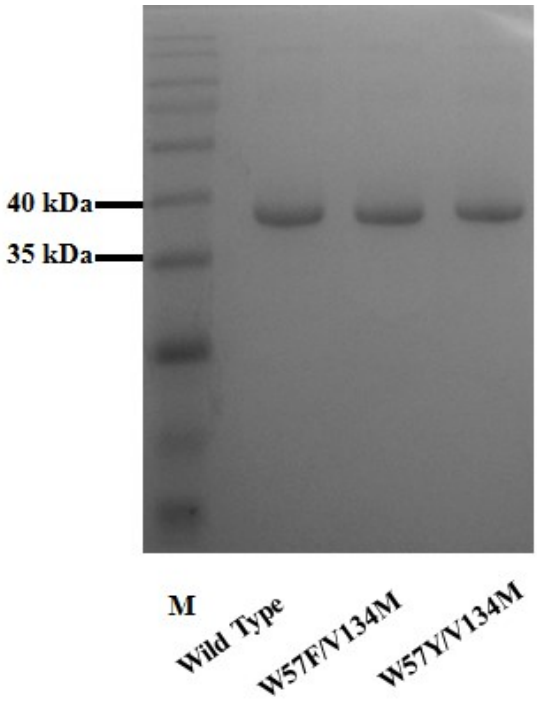


Figure S7 Expression and purification of wild type bll6402NIT, W57F/V134M and W57Y/V134M.

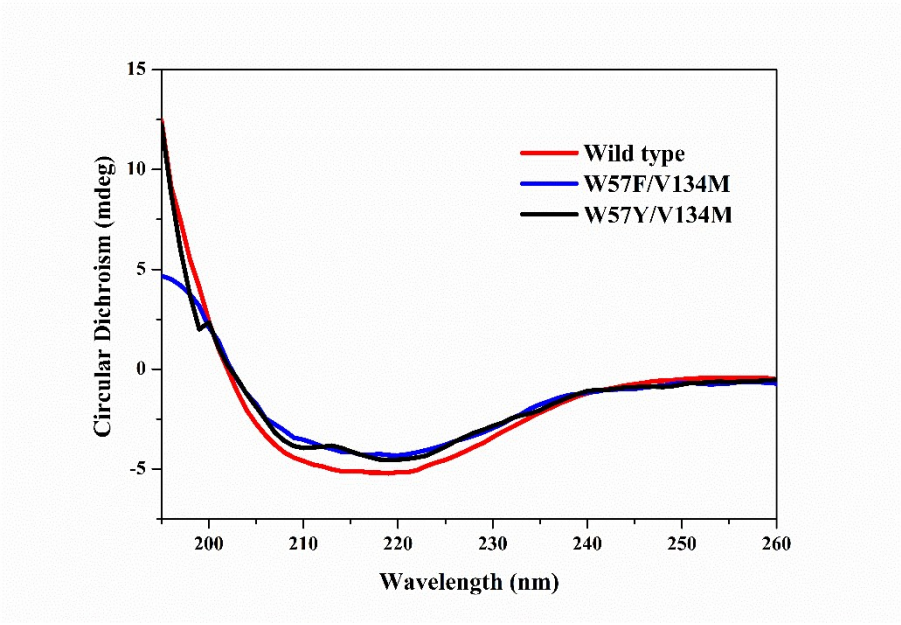


Figure S8 Spectroscopic structural analysis of wild type bll6402NIT (red), W57F/V134M (blue) and W57Y/V134M (black). Circular dichroism wavelength scans of enzyme at pH 7.4 at 30°C.

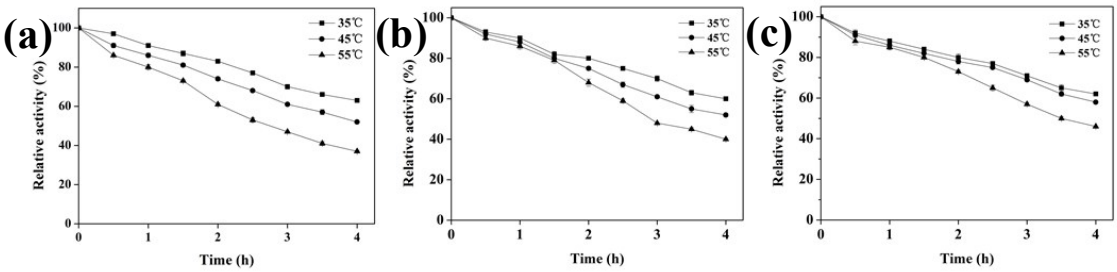


Figure S9 Thermostability analysis of wild type bll6402NIT (a), W57F/V134M (b) and W57Y/V134M (c).

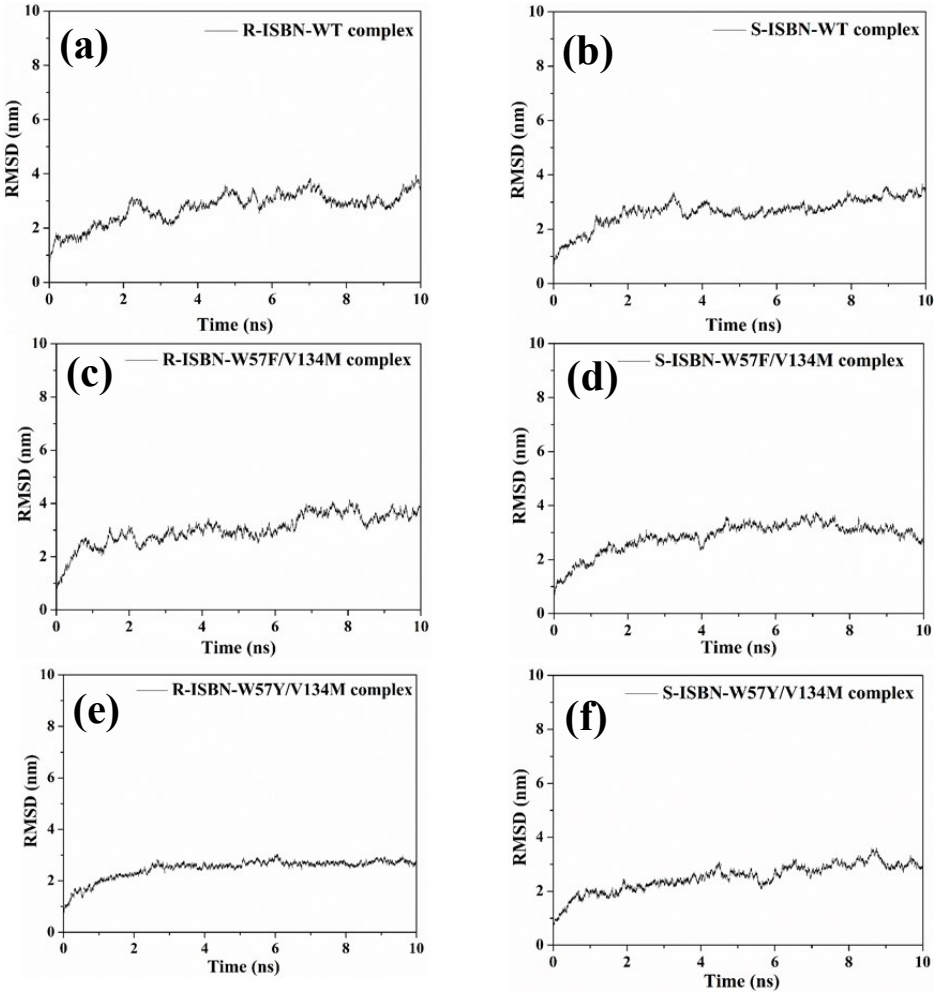


Figure S10 RMSD calculation of wild type bll6402NIT and mutants with substrate complexes during 10 ns molecular dynamic simulation.

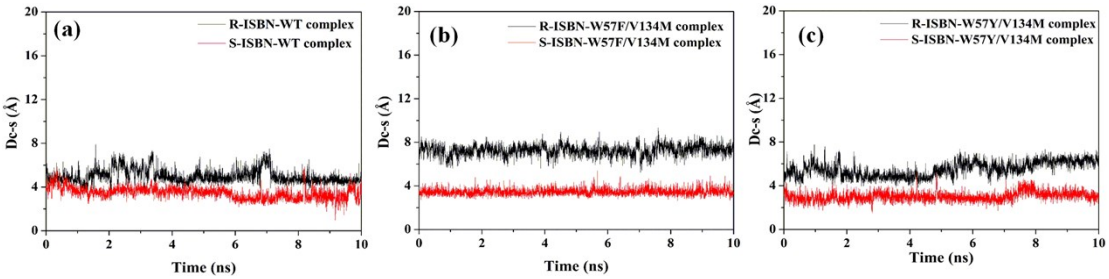


Figure S11 Dc-s Measurements of wild type bll6402NIT (A) and mutants W57F/V134M and W57Y/V134M during 10 ns molecular dynamic simulation.

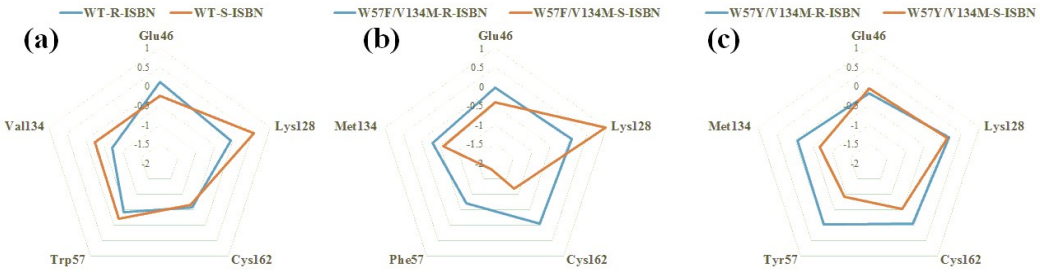


Figure S12 Binding free energies ($\Delta G_{ligand\text{-}residue}$) on hotspot residues and catalytic triad (EKC) of nitrilase for 6402WT-substrate complex **(a)**, W57Y/V134M-substarte complex **(b)** and W57Y/V134M complex **(c)**.

8. References

1. M. M. Bradford, *Anal Biochem*, 1976, **72**, 248-254.
2. C. S. Chen, Y. Fujimoto, G. Girdaukas and C. J. Sih, *J. Am. Chem. Soc.* , 1982, **104**, 7294-7299.
3. Z. Chen, S. Jiang, H. Wang, L. Wang and D. Wei, *Chem Commun (Camb)*, 2019, **55**, 2948-2951.
4. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J Comput Chem*, 2009, **30**, 2785-2791.
5. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J Am Chem Soc*, 1995.
6. D. A. Case, T. E. Cheatham, 3rd, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J Comput Chem*, 2005, **26**, 1668-1688.
7. J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J Comput Chem*, 2004,

25, 1157-1174.

8. P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, 3rd, *Acc Chem Res*, 2000, **33**, 889-897.