Week1 (7.1-7.7): Site 289 targeted saturation mutagenesis and screening

(i) 289 locus targeted mutation and function detection:

1.289 targeted mutations at the locus

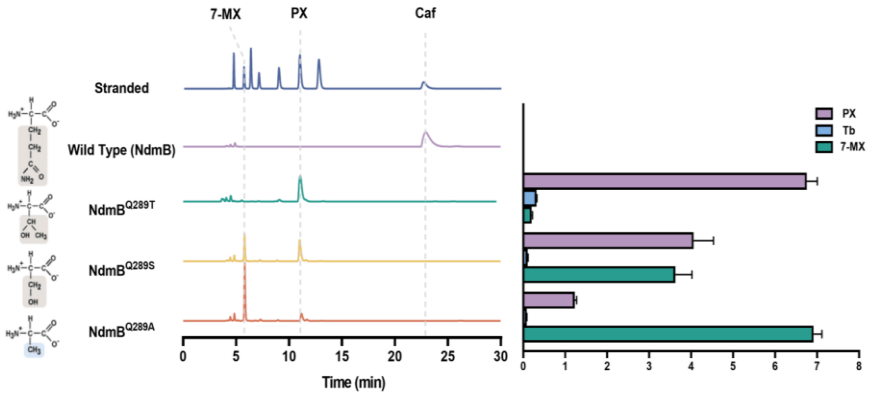
After identifying the 289th locus as the critical locus, we performed targeted mutagenesis of the 289th locus in the same way as before.

2. Functional testing of mutants

(1) Nearly 200 mutants were selected for validation. Those with high OD values were also selected for liquid-phase assay as well as sequencing. The results of liquid phase detection are as follows:

序号	名称	7-MX产量	Tb产量	PX产量	突变位点（Gln289）
1	1-A10	0.079687333	0.03175963	0.23418437	Gln289→Ser289
2	1-D2	0.034147903	0.01911288	0.16570454	Gln289→Thr289
3	1-C12	0.019101483	0.02087637	0.26901942	Gln289→Ala289
4	2-B12	0.64264753	0.01034581	0.03162134	Gln289→Ala289
5	2-B1	0.174261018	0.01098403	0.11833204	Gln289→Gly289
6	2-D6	0.153878988	0.01226046	0.18580015	Gln289→Gly289
7	2-G2	0.235060292	0.00764179	0.01307306	Gln289→Ser289
8	2-F2	0.013979298	0.02279102	0.30512408	Gln289→Ser289
9	1-C1	0.117490129	0.04067785	0.25289135	Gln289→Ala289
10	1-B5	0.034894888	0.03175963	0.30690948	Gln289→Ser289
11	1-A11	0.175034681	0.00476982	0.00805412	Gln289→Cys289
12	1-D5	0.128107993	0.03244823	0.13380547	Gln289→Ser289

(2) We selected the type of mutation with different amino acids at position 289 for the validation of the expanded system (100 OD, 8 mM).

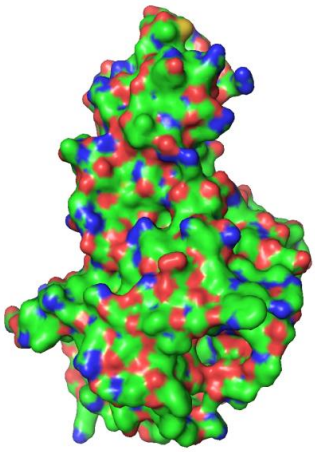


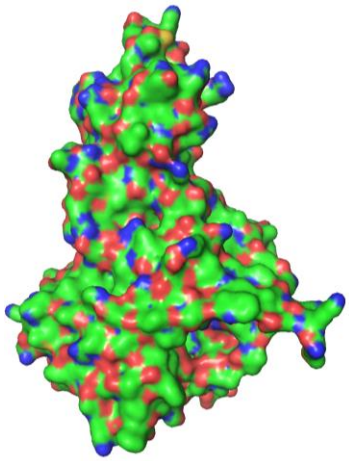
(3) Analysis of results

The experimental results revealed two key phenomena: first, when Gln289 was mutated to threonine, the yield of PX was significantly increased to 1.22 g/L (6.76 mM), which was equivalent to 7.3 times the highest yield available. Second, with the gradual decrease of the

R group (threonine, serine, and alanine), the types and proportions of the products showed a clear regular change: the smaller the R group, the yield of PX gradually decreased, while the yield of 7-MX increased accordingly. This phenomenon suggests that the volume of the R group has a significant effect on the substrate heteropantagonism of the enzyme: as the volume of the R group decreases, NdmB is able to catalyse not only the demethylation reaction at the N3 position, but also gradually exhibits the ability to catalyse the demethylation reaction at the N1 position.

Week2-4 (7.8-7.28): Interpretation of data analysis using molecular docking and molecular dynamics simulations





(i) Pre-testing:

Although AlphaFold has been reported to predict the 3D structure of proteins well, we used it to predict the structure of the original NdmB and found that it was far from the real structure (PDB data), so we used molecular dynamics simulation to obtain the mutant.

1. first download the standard NdmB.pdb format file from the database, and then transfer it to gromacs for kinetic simulation modelling, before we start due to the fact that there is no force field file for Fe3+ in the force field database of gromacs, we found the force field of metal ions such as Fe3+ in water by checking the literature and imported it into the force field file database of gromacs, and it applies to the spc water model used in our simulations.
2. We chose the AMBER99SB-ILDN force field for the simulation: (1) because the force field data of Fe3+ in this literature are applicable to the AMBER force field; (2) this is an improvement of the side chain of proteins based on AMBER99SB, which is usually considered more suitable for protein simulation.
3. However, we did not find the force field data for [2Fe-2S], so we deleted it when we did the simulation, which may have an effect on the final morphology of the protein, but since we mainly studied how the SPRBCC structural domain binds to and reacts with the substrate, the lack of the structural domain did not have a significant effect on our results.

	TIP3P		SPC/E		TIP4P$_{EW}$	
	$R_{min}/2$ (Å)	ε (kcal/mol)	$R_{min}/2$ (Å)	ε (kcal/mol)	$R_{min}/2$ (Å)	ε (kcal/mol)
Al^{3+}	1.297	0.00471279	1.296	0.00465074	1.285	0.00401101
Fe^{3+}	1.386	0.01357097	1.386	0.01357097	1.375	0.01205473
Cr^{3+}	1.344	0.00848000	1.343	0.00838052	1.333	0.00743559
In^{3+}	1.461	0.02808726	1.461	0.02808726	1.450	0.02545423

(ii) Optimisation of molecular dynamics simulation processes

1. Determination of reaction conditions

To better approximate the model to the real state in our experiments, the reaction conditions are to be determined first

Whole-cell catalytic conditions: pH = 9; T = 293.15 K; P = 1 bar

In whole-cell catalysis the intracellular temperature is close to the ambient temperature, so the temperature is still 293.15 K. The intracellular pressure ranges from about 0.8 to 1.2 bar, taking the common value of 1 bar, but the intracellular pH is usually between 7.4 and 7.8, and will not be exactly equivalent to the external pH.

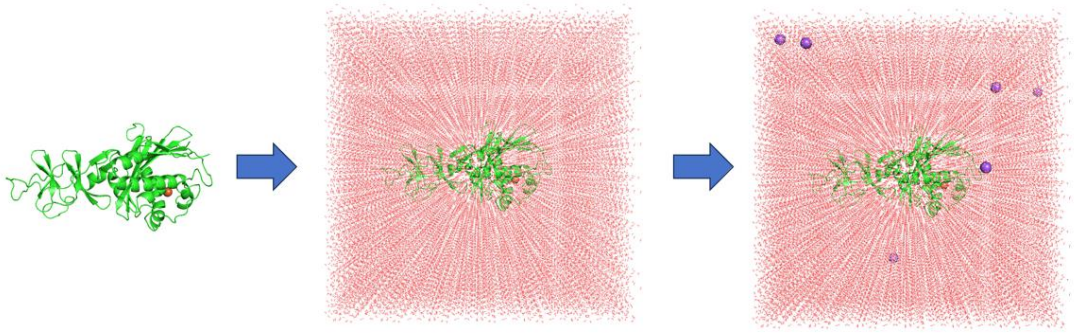
By reviewing the literature we found a model of the pH relationship between the inside and outside of the E. coli cell, and based on the model we calculated that the intracellular pH = 7.781 when the external pH = 9

After determining the pH of the protein's environment, we pre-processed the protein structure, i.e. protonated it at the corresponding pH using PDB2PQR.

(iii) Dynamics modelling

1. Start kinetic modelling after completion of pre-processing (protonation)

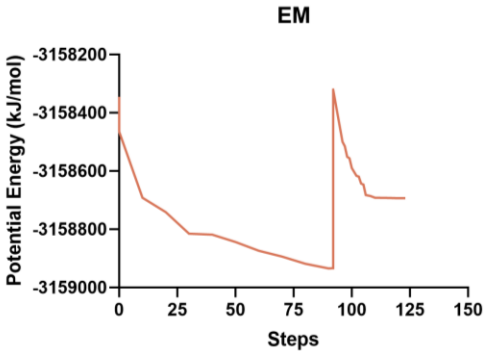
To better simulate the state of the protein in the cell, it was first placed in a water-filled environment with Na+ and Cl- to neutralise the charge.



Energy minimisation is carried out and optimised using the STEEP method and the CG method alternately,

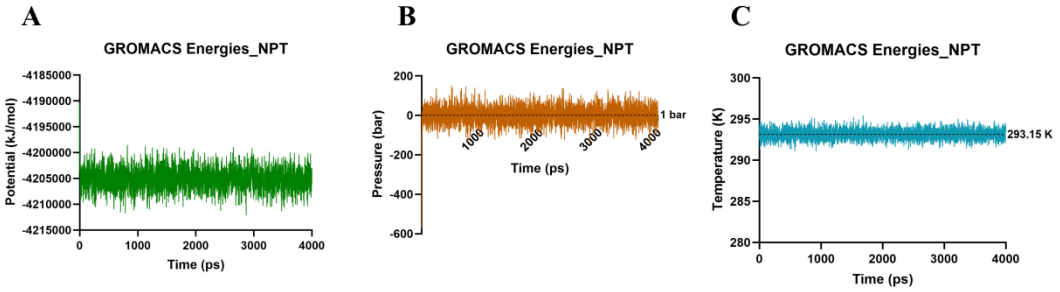
Ultimately, the maximum force Fmax < 50

As shown in the figure, the fifth algorithm is optimised alternatively and the maximum force obtained is less than the threshold value



NPT Balancing Results:

The results show that the potential energy is very low and remains stable, the pressure fluctuates but is stable near the target value, and the temperature is also stable near the target value, indicating that the protein equilibrium is more successful and the next simulation can be carried out.



After molecular dynamics simulation of the NdmB structure in the database according to our transformation conditions (pH=9, T=20°C, P=1bar) to get a more accurate protein state. The Q289 site of NdmB was changed to T289, S289, A289 by PyMOL, and the protein model was simulated by kinetic simulation under our whole-cell catalytic conditions.

(iv) Further refinement of the results of the dry run component

After obtaining a more accurate structure of the enzyme, we first used AutoDockTools (ADT) to dehydrate the protein receptor, add full hydrogen, and store it as a PDBQT format file. After that we obtained the planar conformation of caffeine from PubChem, performed operations such as hydrogenation and merging of non-polar hydrogen atoms, and the ligand was also stored as a PDBQT format file. The molecular docking analysis was done by AutoDock Vina 1.1.2, which is mainly used to assess the binding mode of substrates such as caffeine and PX at the active centre of NdmB enzyme. We ran the docking procedure multiple times and selected the conformation with the lowest binding energy for in-depth analysis. This procedure helped us to reveal the effect of different mutations on substrate binding capacity. Subsequently, we performed a detailed comparison of the structural dynamics and interactions between wild-type NdmB and NdmB^{Q289T} mutants in the presence of caffeine. The analyses focused on key aspects such as changes in distance relative to non-heme iron, radial

distribution function (RDF), principal component analysis (PCA) and covariance mapping. Special attention was paid to the formation of hydrogen bonds with the aim of further elucidating the mechanism of interaction between the mutant and caffeine during the binding process.