Semi-Rational Design of Nitroarenes Dioxygenase for Catalytic Ability Towards 2,4-Dichloronitrobenzene

Jia Xu[a#], Tao Li[a#], Wei E. Huang[b]*, Ning-Yi Zhou[a]*

[a]State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 200240 Shanghai, China

[b]Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK

# Jia Xu and Tao Li contributed equally to this work. Author order was determined by drawing straws.

*Correspondence to Ning-Yi Zhou, Email: ningyi.zhou@sjtu.edu.cn; Wei E. Huang, Email: wei.huang@eng.ox.ac.uk

Abstract

Rieske non-heme dioxygenase family enzymes play an important role in the aerobic biodegradation of nitroaromatic pollutants, but no active dioxygenases are available in nature for initial reactions in the degradation of many refractory pollutants like 2,4-dichloronitrobenzene (24DCNB). Here, we report the engineering of hotspots in 2,3-dichloronitrobenzene dioxygenase from *Diaphorobacter* sp. strain JS3051, achieved through molecular dynamic simulation analysis and site-directed mutagenesis, with the aim of enhancing its catalytic activity towards 24DCNB. The computationally predicted activity scores were largely consistent with the detected activities in wet experiments. Among them, the two most beneficial mutations (E204M and M248I) were obtained, and the combined mutant reached up to a 62-fold

increase in activity towards 24DCNB, generating a single product 3,5-dichlorocatechol which is a naturally occurring compound. *In silico* analysis confirmed that residue 204 affected the substrate preference for *meta*-substituted nitroarenes, while residue 248 may influence substrate preference by interaction with residue 295. Overall, this study provides a framework for manipulating nitroarene dioxygenases by computational methods to address various nitroarene contamination problems.

Importance

As a result of human activities, various nitroaromatic pollutants continue to enter the biosphere with poor degradability, and dioxygenation is an important kickoff step to remove toxic nitro-groups and convert them into degradable products. The biodegradation of many nitroarenes has been reported over the decades, however, many others still lack corresponding enzymes to initiate their degradation. Although rieske non-heme dioxygenase family enzymes play extraordinarily important roles in the aerobic biodegradation of various nitroaromatic pollutants, prediction of their substrate specificity is difficult. This work greatly improved catalytic activity of dioxygenase against 24DCNB by computer-aided semi-rational design, paving a new way for evolution strategy of nitroarene dioxygenase. This study highlights the potential for using the enzyme structure-function information with computational pre-screening methods to rapidly tailor the catalytic functions of enzymes towards poorly biodegradable contaminants.

Keywords: biodegradation; molecular dynamics simulation; nitroarene dioxygenase; semi-rational design; site-directed mutagenesis.

Introduction

Hundreds of synthetic nitroarenes are globally produced and used as chemical feedstocks for the manufacturing of drugs, dyes, pesticides, and explosives (1). However, contaminants of various nitroarenes have been released into biosphere from industrial wastes or improper handling of chemical products (2). These compounds are more toxic than their parent aromatic compounds and are resistant to biodegradation due to the presence of electron-withdrawing nitro groups (sometimes halogen-groups) (3, 4). Exposure to nitroarenes may cause the formation of DNA adducts, which further lead to carcinogenesis, mutagenesis, and teratogenesis (2). Therefore, a number of nitroarenes have been listed as priority pollutants by United State Environmental Protection Agency (EPA) (5).

The prior treatment for nitroarenes in environment would be an effective biodegradation that eliminates its poisonous nitro-group. The potentials of microorganisms to degrade recalcitrant nitroarenes has been invoked by the characterization of bacteria capable of growing on nitroaromatic substrates, for example, 2-nitrotoluene (2NT) (6), nitrobenzene (NB) (7), 3-nitrotoluene (3NT) (8), 2,4-dinitrotoluene (24DNT) (9, 10), 2-chloronitrobenzene (2CNB) (11), 2,3- and 3,4-dichloronitrobenzene (DCNB) (12). Rieske non-heme iron dioxygenases are versatile enzymes which play an important role in xenobiotic degradation, notably, they kick-start the nitroarenes degradation pathway with the formation of the biodegradable catechols and the release of nitrite (1). This type of enzymes contains three components, namely ferredoxin reductase, Rieske ferredoxin, and $\alpha_3\beta_3$ terminal oxygenase. Each α subunit of oxygenase contains a Rieske [2Fe−2S] cluster and a mononuclear iron (II) catalytic center, which is related to electron transfer and

78 substrate oxidation, respectively (13, 14). Generally, nitroarene dioxygenases share ≥
79 80% amino acid sequence identity (**Fig. 1A**), and some key residues from the catalytic
80 domain of the α subunit contribute to the difference in their substrate specificity (3,
81 15-18).

82 Extensive research has been conducted on several nitroarene dioxygenases,
83 including crystal structure data (19, 20), computational chemical studies (21-25),
84 peroxide shunt (26), and substrate turnover (27, 28) experiments. Therefore, the
85 catalytic site and substrate binding site at the active center in α subunits of oxygenase
86 and the substrate oxidation mechanism are gradually uncovered and become clear.
87 Nitroarene dioxygenases in general share a similar substrate binding mode, with the
88 right exposure of their susceptible sites for the electrophilic attack to the iron center.
89 The entry of substrate into the active site leads to the rearrangement of water
90 molecules and ligands to form Fe^{III}-(hydro)peroxo species ($[Fe_{III}–OOH]^{2+}$), and a
91 dioxygenation attack is initiated through the generation of a peroxo-bridged substrate
92 radical between the substrate and Fe-oxygen species (24-28). In this process, a good
93 substrate fit in the active site is a prerequisite for O_2 activation. Some nitroarene
94 dioxygenases possess an asparagine at position 258, which will form a hydrogen bond
95 (H-bond) between its amino group and the nitro group of the substrate (14-16, 20).
96 This interaction is believed to play an extremely important role in the proper
97 positioning and stable binding of nitroarene substrates.

98 However, in spite of decades of efforts made in isolating nitroarene degraders and
99 identifying enzymes involved, many nitroarene pollutants have remained poorly
100 biodegradable. Take 24DCNB as an example, which is a High Production Volume
101 (HPV) chemical (greater than one million pounds per year) (29) used for the synthesis
102 of 2,4-difluoroaniline, and detected in multifarious industrial wastewater (30-32).

Compared to its isomers 23DCNB or 34DCNB, 24DCNB is not only genotoxic but also considered to be a possible human carcinogen (33-35). The large quantity of usage, refractoriness, genotoxicity, carcinogenicity and toxicity of 24DCNB have prompted researchers to search for efficient and economical environmental remediation approaches (30-32, 36, 37). Various bio-electrochemical methods have been proposed to remove 24DCNB from wastewater (30-32, 36, 37), however, no strains or enzymes have been reported to be capable of degrading 24DCNB yet.

Computer-aided directed evolution of proteins has been proven to have the superior capability of modifying protein for desired functions (38, 39). Especially for those enzymes that lack a high-throughput screening method, rational design based on information on protein structure and mutagenesis hotspots plus the support of computational tools offer an effective solution, which creates a significantly smaller variant library and provides *in silico* prescreening (40). Recently developed tools, including computer-based prediction of protein structures (41, 42), molecular docking (43, 44) and molecular dynamic (MD) simulation (45, 46), would dramatically bridge the gap between limited enzyme resources and explosive growth of artificial chemical compounds.

Herein, we applied an *in silico* method to expedite the improvement of the catalytic activity of nitroarene dioxygenase towards 24DCNB. Through computationally predicting and experimentally verifying catalytic activity, two advantageous mutations (E204M and M248I) were selected from a small mutant library containing 14 single-site mutants of 23DCNB dioxygenase; and after combining these two advantageous mutations together, E204M-M248I showed a 62-fold increase in the activity towards 24DCNB. The computer-assisted methodology employed in this study provides a framework for enhancing the activity

of nitroarene dioxygenases toward recalcitrant nitroarenes.

Materials and methods

Chemicals, plasmids, bacterial strains, and culture conditions.

2-nitrotoluene, 4-nitrotoluene, 1-nitronaphthalene, 2,5-dichloronitrobenzene, 2,4-dinitrotoluene, 2,6-dinitrotoluene, 3-methylcatechol, 4-methylcatechol, 3-chlorocatechol, 4-chlorocatechol, 3,4-dichlorocatechol, 3,5-dichlorocatechol, 4,5-dichlorocatechol and 4-methyl-5-nitrocatechol were purchased from Sigma-Aldrich (USA) and their purity are >99%. 2-chloronitrobenzene (>99%), 3-chloronitrobenzene (>99%), 4-chloronitrobenzene (>99.5%) and 2-chloro-3-nitrotoluene (>98%) were purchased from Aladdin (China). 3-nitrotoluene (>99%) and 2,4-dichloronitrobenzene (>99%) were purchased from TCL (China). 4-chloro-3-nitrotoluene (99.36%) and 3,5-dichloronitrobenzene (98%) were purchased from bidepharm (China). 6-chloro-2-nitrotoluene (99%) and 2,3-dichloronitrobenzene (99%) were purchased from Macklin (China). In addition, 3,4-dichloronitrobenzene (Alfa Aesar, China, 99%); 2,4-dinitrochlorobenzene (Adamas-beta, China, 99%). The plasmid pETDuet-DCB harboring the genes for 23DCNB dioxygenase was constructed previously (47) by inserting the *dcbAaAb* fragment between *Nco*I and *Sac*I, and the *dcbAcAd* fragment between *Nde*I and *Kpn*I. *Escherichia coli* DH5α and *Escherichia coli* BL21(DE3) were used for cloning and expressing mutant proteins, respectively. *E. coli* strains were cultured at 37°C in lysogeny broth (LB) or LB agar with appropriate antibiotics.

Site-directed mutagenesis.

Site-directed mutagenesis of *dcbAc* was performed as the method described previously (47). Briefly, plasmid pETDuet-DCB which was used as the template was

amplified by Phanta Max Super-Fidelity DNA Polymerase (Vazyme Biotech Co., Ltd, China) with complementary mutagenic primers described in **Table 1**. Then the PCR products were digested using *Dpn*I, and the resulting products were transformed in the host bacteria into *E. coli* DH5α and screened on LB agar with 100 μg/ml ampicillin.

Whole-cell biotransformation assays to determine the specific activity of dioxygenase mutants.

E. coli BL21 (DE3) (pETDuet-DCB) cells were grown in LB medium at 37°C until the optical density at 600 nm ($OD_{600\ nm}$) reached 0.8, which was measured in a BioTek EPOCH 2 microplate reader. Then 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added before incubating overnight at 16°C to induce dioxygenase expression. The cells were harvested, washed twice with phosphate-buffered saline (PBS) and suspended with PBS containing 80-200 μM 24DCNB. Specific activities were determined by measuring the rate of nitrite formed at appropriate intervals (depending on the activity of each mutant) during exposure to 200 μM 24DCNB or 23DCNB with shaking (220 rpm, 30°C). Nitrite was detected by the Griess method as described previously (48). To measure the protein concentration, cell pellets were recollected by centrifugation, suspended in equal volumes of 0.1 M NaOH and boiled for 10 min. After that, protein concentration was measured by the Bradford method (49) with bovine serum albumin as the standard. Concentrations of 24DCNB and 35DCC were quantified by high-performance liquid chromatography (HPLC). To determine protein expression pattern, 1 ml of cell cultures were harvested and resuspended with equal volumes of sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) running buffer. Samples (20 μl each) were analyzed by 12% SDS-PAGE.

Analytical methods.

Reverse-phase HPLC with a Waters e2695 separation module equipped with a Waters 2998 photodiode array detector and a C_{18} reversed-phase column (5 μm, 4.6 × 250 mm) at 30°C was used to quantify the compound concentrations of 24DCNB and 35DCC. The mobile phase consisted of water containing 0.1% (vol/vol) acetic acid (A) and methanol (B), eluted with 20% of solvent B for 5 min and linearly increased to 90% B after 30 minutes.

The products of biotransformation were identified by gas chromatography-mass spectrometry (GC-MS), which was performed with a TSQ™ 8000 Evo Triple Quadrupole GC-MS/MS (Thermo Fisher Scientific Inc., MA, USA) equipped with a capillary column HP-5MS (0.25 mm × 30 m, Agilent technologies., CA, USA). For GC-MS analysis, biotransformation samples were extracted with an equal volume of ethyl acetate, which was then evaporated to dryness and dissolved in 0.2 ml ethyl acetate. Samples were further derivatized by adding equal volumes of *N*, *O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) at 70°C for 30 min. GC/MS program setting: the inlet temperature was set at 280°C, and the initial temperature was 70°C for 2 min, raised to 130°C at 5°C/min, increased to 180°C at 10°C/min, increased to 285°C at 5°C/min, and held for 5 min. Mass spectrometer conditions: 33–750 *m/z* mass range at the electron energy of 70 eV, EI energy source.

Simulation system preparation.

The initial structure of mutated α subunits of dioxygenases was generated by AlphaFold2 (42), and all models were obtained with a high confidence level (an average plDDT (confidence) of 97). The protonation states of protein residues were verified and their hydrogen atoms were added by using H++ web server (50). The enzyme-ligand complex was assumed to be in its transition state, with Fe^{III}-(hydro)peroxo species formed in the active site. Therefore, a hydro(peroxo)

moiety was placed next to the central iron by following the method reported previously (14), and the ligand (24DCNB) was docked to the predicted models with AutoDock Vina (43). Then the program MCPB (51, 52) was used to generate force field parameters for $[Fe_{III}–OOH]^{2+}$ and Rieske [2Fe−2S] cluster using the ff99SB force field, while the Gaussian 16 (53, 54) was used to calculate the optimized geometries, force constants and ESP charges. The resultant amber topology was further transformed into a GROMACS topology using ACPYPE (55).

Molecular dynamic simulation.

The molecular dynamic (MD) simulations were performed using the GROMACS 2020 (56) with the ff99SB force field. TIP3P-type water molecules counter ions were filled into an extended 1 nm cubic box to generate a neutralized system and the temperature and pressure of the simulation system were set at 300 K and 1 bar separately. The prepared system first went through 50,000 steps of steepest descent minimization until the maximum force < 5.0 kJ/mol, followed by 100 ps of equilibration using NVT and NPT simulations respectively. During these two simulation phases, the protein and ligand were held fixed by using position restraints with a force constant of 1000 kJ mol^{-1} nm^{-2}. The V-rescale method was used for the maintenance of constant temperature and the Parrinello-Rahman method for constant pressure. LINCS algorithm (57) was used to constrain all covalent bonds involving hydrogen atoms. The particle mesh Ewald method (58) with a grid length of 0.16 Å was used to calculate the electrostatic interactions. For each system, a 20 ns of MD simulation was carried out under the same conditions and repeated for three times with a different random number, and some selected ones were extended to 200 ns.

After obtaining the trajectory of the MD simulation, root mean square deviation (RMSD), root-mean-square fluctuation (RMSF), the distance and angle between

atoms, and conformational clustering analysis were calculated using the *gmx rms, gmx rmsf, gmx distance, gmx gangle,* and *gmx cluster* commands. The resulting structures were visualized by PyMOL (59). POVME 3.0 (60) was used to measure the substrate pocket volume distribution of conformations captured every 5 ns beginning from 100 ns trajectory. A dynamical cross-correlation matrix (DCCM) was obtained by using a python script to calculate correlation coefficients of the wild-type MD simulation every 50 ps in the last-10 ns trajectory.

Results

Analysis of key residues and construction of the mutant library of 23DCNB dioxygenase. Considering the facts that the α subunits of Nag-like nitroarene dioxygenases exhibit high sequence conservation and they determine the substrate specificity (3, 15-18), the differential residues of the α subunits are thus considered as potential hotspots for protein engineering (**Fig. 1B**). In terms of the structural resemblance of 23DCNB to our intended substrate 24DCNB, we chose the 23DCNB dioxygenase from *Diaphorobacter* sp. strain JS3051 (12, 47) as the starting enzyme to carry out further modifications. We focused on the residues around the active center that were loosely conserved, or those conserved but different sites (61) (**Fig. 1B and C**), especially the sites that may have an effect on substrate specificity mentioned in previous reports (3, 15-18, 47).

To enhance the dioxygenase's activity towards 24DCNB, a *para*-substituted nitroarene. Residues VAL207, SER242, and MET248, predicting to affect the accommodation of the enzyme toward *para*-substituted group, were mutated into the corresponding residues in 24DNT dioxygenase or 34DCNB dioxygenase. Residue 204, serving as a "gatekeeper," occupies a pivotal position, bridging the enzyme

tunnel and binding pocket, and controlling access to the active center (47, 62). While most Nag-like dioxygenases feature a hydrophobic amino acid at this position, the 23DCNB dioxygenase has a glutamic acid, which is believed to form a halogen bond with the C3 chlorine atom of 23DCNB (47). However, for our target substrate, 24DCNB, GLU at position 204 might disrupt proper substrate positioning. Hence, we substituted GLU204 with other hydrophobic amino acids featuring bulky side chains (ILE, LEU, MET, PHE, TRP, TYR) to better shape the active center or enzyme tunnel. The remaining two selected hotspots were modified with amino acids from other characterized dioxygenases. In the end, a total of 14 mutations (E204F, E204I, E204L, E204M, E204Y, E204W, V207I, S242T, M248I, L293H, L293I, L293Q, I350T, and I350V) were generated. Three-dimensional structures of these variants, along with the wild type 23DCNB dioxygenase, were predicted using AlphaFold2 (42) for further computational analysis.

***In silico* prediction and *in vivo* assessment of enzyme activity on 24DCNB.** To roughly investigate the effect of six different residues on dioxygenase-24DCNB interactions and pre-screen the mutant library by catalytic potential, we carried out 3 times of 20 ns MD simulations for 14 mutants and wild type, all with 24DCNB docked at the active center of each α subunit. The RMSD values of all systems reached a plateau after the first 5 ns simulations (**Fig. S1**), indicating that most systems reached their equilibrium. Therefore, all data were observed and collected from 5 ns to 20 ns.

According to the catalytic process of nitroarene dioxygenases (14-16, 20, 24-28), the nitroarene substrate can be anchored through an H-bond interaction, facilitating subsequent electrophilic attack on its benzene ring and denitration. A stable H-bond between the amino group of Asn258 and the nitro group of the substrate ensures a

stable conformation and correct position of 24DCNB with respect to center [Fe_{III}–OOH]$^{2+}$, which is an essential precondition for reaching the pre-reaction state. The probability of H-bond formation was assessed using two key parameters: 1) the Oa-Nd (d) distance, which represents the distance between the oxygen atom (acceptor) of the nitro group in 24DCNB and the nitrogen atom (donor) of Asn258, and 2) the Nd-H···Oa angle (θ), which signifies the donor-hydrogen···acceptor angle (**Fig. 2**). In general, a hydrogen bond is considered proper when the acceptor-donor distance is around 3.0 Å and the donor-hydrogen···acceptor angle approaches 180° (64). The closer the hydrogen bond approaches these ideal geometric values, the stronger the bond becomes. Consequently, we established the geometric hydrogen bond criterion as d(Oa-Nd) ≤ 3.5 Å and θ (Nd-H···Oa) ≥ 150°. We then calculated the proportion of the frequency of the correct conformations to the total frequency within each 20-ns simulation system that meets this established criterion, and recorded it as the probability of H-bond formation (HB-Probability) (**Fig. 3A**).

On the other hand, a productive conformation should guarantee that C1 and C6 carbon of 24DCNB are the dihydroxylated positions, yielding the only possible product, 3,5-dichlorocatechol (35DCC). Two parameters were used to evaluate productive and nonproductive conformations, characterized by O1-C1 (D_1) and O2-C6 (D_2) distances between the O1 and O2 oxygen of [Fe_{III}–OOH]$^{2+}$ and the C1 and C6 carbon of 24DCNB respectively (**Fig. 2**). The average oxidation attack distances D_1 and D_2 were calculated and presented in **Fig. 3B**, all with a standard deviation between 0.03-0.07 nm. To allow a subsequent dioxygenated attack on the substrates, both distances must be appropriate.

In fact, the transiently equilibrated conformations and coordinates of 24DCNB in each 20 ns MD simulation system may not always accurately reflect their true state.

The reliability of a docking conformation is directly proportional to its frequency, as the most frequently occurring one is more likely to be deemed trustworthy upon repetition. The majority of mutant dioxygenases and wild-type dioxygenase exhibited a low probability or no possibility of H-bond formation in most repetitions, with only five mutants generally showing high probability when repeated (**Fig. 3A**). Hence, these five mutants, namely E204I, E204L, E204M, E204Y and M248I, were identified as being able to effectively establish H-bonds with the substrate 24DCNB.

Averagely, there were more than half populations successfully formed the H-bond in these five mutants' simulation systems. Specifically, E204M performed the best of all, with the highest HB-probability of 85.1% and an average probability of 62.7%.

The distribution of oxidation attack distance (**Fig. 3B**) exhibited a certain correlation with the distribution of HB-probability (**Fig. 3A**), owing to the relatively fixed positions of the iron center and Asn258. Specifically, when an H-bond was well-formed, 24DCNB tended to be anchored, thereby exposing the C1 and C6 carbon sites to the iron-oxygen complex; in other words, a high HB-probability often coincided with short oxidation attack distances. As demonstrated by the instances of E204I, E204L, E204M and M248I, the mean distances of both D_1 and D_2 were roughly less than 0.4 nm. The distances between the oxygen atoms of iron-oxygen complex and the reacting carbon atoms in the crystal structure of naphthalene dioxygenase (PDB ID: 1O7N) (19) are 0.32 nm and 0.28 nm, which is slightly smaller than the corresponding experimental values presented here. The conformation clustering analysis has confirmed that 24DCNB in the active site of these four mutants adopted a stable docking conformation throughout the 20 ns simulation (**Fig. S2C**), which is considered to have reached the pre-reaction state. However, E204Y is an exception with a high probability of forming hydrogen bonds but relatively long

distances for oxidation attacks. Different docking conformations were adopted during the 20 ns simulation of E204Y, with the C1 and C2 adjacent carbon atoms of 24DCNB being closest to the oxygen atom pair of $[Fe_{III}–OOH]^{2+}$ (**Fig. S2AB**). Despite this proximity, the presence of nitro- and chlorine-group at these two sites hinders productive dioxygenation, as the simultaneous denitration and dichlorination has not been observed thus far (3).

Considering both aspects that may reflect the catalytic potential of mutant dioxygenases towards 24DCNB, four mutants E204I, E204L, E204M, and M248I were supposed to be candidates for 24DCNB dioxygenase.

Meanwhile, to verify the actual dioxygenase activity, we conducted whole-cell biotransformation assays to obtain the specific enzyme activities of wild type and 14 mutant dioxygenases towards 23DCNB and 24DCNB. Based on SDS-PAGE (**Fig. S3**), all of the mutations produced comparable amounts of dioxygenase proteins. As shown in **Fig. 3C**, except for L293I, the specific enzyme activity on the natural substrate 23DCNB was reduced to varying degrees in all mutants. More importantly, E204I, E204L, E204M, and M248I developed noticeable catalytical activity towards 24DCNB from nearly nothing, which is in agreement with the results of *in silico* prediction. It's interesting to note that three of them are mutations at residue 204, suggesting its importance in controlling substrate specificity and catalytic activity. And when the glutamic acid at position 204 was mutated to methionine, the dioxygenase showed the highest activity towards both 23DCNB and 24DCNB among these three mutants.

Combination of beneficial mutations at sites 204 and 248. Epistasis describes a genetic phenomenon in which the combined effect of multiple mutations is not a simple addition of their individual effects, but a more complex result, which also

occurs in protein evolution (65). As mutations at residues 204 and 248 conferred the dioxygenase with a unique catalytic capacity, it is intriguing to explore the potential synergy of combining these two beneficial mutations and their impact on the enzyme's properties. Hence, a double mutant E204M-M248I was obtained. The *in vivo* specific activities of single-site and double-site mutants were determined with various nitroarene substrates (**Fig. 4**).

Compared to wild-type dioxygenase, E204M mutation modulated the substrate preference of dioxygenase towards favoring *ortho*-substituted substrates more, significantly increasing the specific activity on 2NT and 2CNB by 5-to 9-fold, while having reduced activities on substrates with a *meta*-substituted group including 3NT, 3CNB, 6C2NT, 2C3NT, 23DCNB, 34DCNB, 35DCNB, 24DNCB, and 26DNT. The specific activity of the M248I mutant towards the majority of substrates tested were comparable to that of wild-type dioxygenase. But 34DCNB and 24DNCB, both have 3-and 4-substituted groups, saw a 2-to 4-fold increase in specific activity of M248I mutant compared with the wild-type. After combining these two beneficial substitutions, the substrate specificity of dioxygenase was broadened notably. In particular, the E204M-M248I mutant greatly enhanced the activity towards substrates that have 2-or/and 4-substituted groups (2NT, 4NT, 2CNB, 4CNB, and 24DCNB). And it is worth noting that E204M-M248I showed a 62-fold increase in specific activity on 24DCNB over the wild-type. Such results indicate that E204M and M248I mutations exhibit positive epistasis, in which greater improvements in specificity and activity have been generated than expected.

Identification of biotransformation products by E204M-M248I. The course illustrating the biotransformation of 24DCNB to 3,5-dichlorocatechol (35DCC) catalyzed by E204M-M248I is presented in **Fig. 5**. Remarkably, E204M-M248I

exhibited the ability to transform 24DCNB into 35DCC and nitrite in a stoichiometric 1:1:1 ratio. These results provide compelling evidence for a dioxygenation reaction targeting the C1 and C6 carbons of 24DCNB, thus validating the hypothesis previously proposed in our computer-aided prediction method.

Additionally, we performed characterization of the reaction products catalyzed by wild-type, E204M, M248I, and E204M-M248I mutants toward 18 distinct nitroaromatic substrates, as depicted in **Fig. S5**. Despite variable relative specific activities exhibited by different mutant dioxygenases towards various substrates (**Fig. 4**), the post-oxidation product composition of most substrates was largely consistent. All substrates except 4C3NT demonstrated susceptibility to oxidation at nitro-substituted positions on the aromatic ring, leading to the formation of catechol derivatives. Additionally, oxidation of certain substrates could also occur at methyl substituents, producing alcohol compounds, albeit without detectable dechlorination products.

In the case of substrate 2NT, substitution of glutamate with methionine at position 204 resulted in an enhanced specificity for dioxygenation at the nitro group, yielding predominantly 3-methylcatechol as the primary product (66.3% and 77.2% for E204M and E204M-M248I, respectively). The M248I mutation impacts substrate specificity towards 34DCNB evidently, which shifts the dioxygenation site from positions 1 and 2 to positions 1 and 6, resulting in 90.7% of the generated products being 4,5-diochlorocatechol. The relative specific activity of M248I mutant towards 34DCNB is also the highest compared to other mutants (**Fig. 4**). A significant alteration in substrate regiospecificity is observed with substrate 24DNT. E204M-M248I resulted in nearly all 24DNT being converted to 4-methyl-5-nitrocatechol, whereas the wild-type or single mutants convert less than

half of the substrate to 4-methyl-5-nitrocatechol.

Further Molecular dynamics analysis of E204M-M248I and two single-site mutants. 3×20 ns MD simulations of E204M-M248I and an additional 200 ns MD simulations separately for wild-type, E204M, M248I, and E204M-M248I based on their 20 ns MD simulations were performed (**Fig. S4**), trying to shed light on the role of E204M and M248I mutations in influencing the enzyme properties. The H-bond formation and oxidation attack distances were also analyzed during the 5-200 ns simulation and presented in **Fig. S6**.

The substitution of glutamic acid at position 204 with methionine contributes to the hydrophobic environment in the binding site while maintaining a similar-sized sidechain. Given that residue 204 lies at the junction of the enzyme tunnel and binding pocket, a properly sized and hydrophobic sidechain can effectively shield the binding site from accessing the solvent and shape the binding pocket by steric hindrance. It's postulated that the mutation E204M could add compactness to the binding pocket so that the 24DCNB's spatial movement was much more constrained inside the pocket. It's reflected by decreased substrate pocket volume (**Fig. 6**) and decreased root-mean-square fluctuations (RMSFs) of 24DCNB (**Fig. 6**). Furthermore, relative to the productive conformation of 24DCNB, the side chain of residue 204 is close to its C3 atom. A more compact active site chamber with narrowed space around the C3 atom of 24DCNB contributes to the improvement of catalytical activity of E204M mutant towards smaller *ortho*-substituted substrates like 2NT and 2CNB while leading to the weakening of activity towards *meta*-substituted substrates.

As for residue 248, its position has a certain distance away from the center iron at about 14 Å and the orientation of its sidechain is in the opposite direction. Therefore, we applied dynamical cross-correlation matrix (DCCM) to determine pairwise

cross-correlation coefficients (C_{ij}) indicating the potential allosteric sites (66, 67). According to the computed DCCM of wild-type (**Fig. S7**), the most obvious positive dynamic correlation was between M248 and N295 with a C_{ij} higher than 0.45, suggesting a potential epistatic effect between these two structural neighbors. Residue 295, as a well-conserved hydrophobic residue in Nag-like family dioxygenases (**Fig. 1B**), lies around the C4 atom of a properly-docked 24DCNB and plays a critical part in maintaining and stabilizing the binding pocket. Mutation M248I thereby may make an impact on N295 and in turn affect the interactions between substrate and the binding pocket, as evidenced by the improved specific activity of M248I mutant towards substrates with 4-substituted groups. Besides, the M248I mutant has the largest substrate pocket volume among the wild-type, E204I, M248I and E204I-M248I variants, providing larger binding space for bulky substrates like 1-NN, which is also in agreement with the experimental results (**Fig. 4**).

By combing two beneficial mutations, E204I-M248I has a higher possibility of H-bond formation, appropriate oxidation attack distances and compact binding pocket (**Fig. S6 and 6**). But it is unfortunate that such inconspicuous improvements obtained from the MD simulation failed to explain the obvious synergistic epistasis between E204I and M248I observed in *in vivo* specific activities.

Discussion

The nitroarene dioxygenases from Gram-negative strains to date have all evolved from a common Nag-like naphthalene dioxygenase ancestor (10, 68-70). These dioxygenases possess a broad substrate spectrum but exhibit a greater preference for various nitro-aromatic compounds. The divergence in substrate specificity and regiospecificity of nitroarene dioxygenases is primarily determined by the matchup

between the substrate and binding pocket (27, 28). Numerous studies, typically involving sequences or structures alignment and site-directed mutagenesis, have demonstrated that such divergence is often caused by substitutions of several key residues, particularly those surrounding the active site (3, 15, 16, 18, 47, 62). An asparagine introduced at position 258 offers hydrogen bonding with nitroarene substrates, capturing and pinning the substrate in the active center and providing the preconditions for efficient denitration (3, 14-16, 20). Substitution of residues at position 204, 293 and 350 would impact the activity and regiospecificity of nitrobenzene dioxygenase (3, 16, 18) and the well-studied naphthalene dioxygenase from *Pseudomonas* sp. NCIB 9816-4 (correspond to positions 206, 295, and 352 in $NDO_{9816\text{-}4}$) (71-73). Ju *et al.* (17) and Mahan *et al.* (18) conducted long-term laboratory evolution experiments to generate mutants of 2-nitrotoluene dioxygenase from *Acidovorax* sp. strain JS42 that exhibited the ability to utilize 3- or 4-nitrotoluene. They found residues outside the active site (238, 242, and 248 as reported by Ju *et al.*, and 405 according to Mahan *et al.*) could also modulate the catalytic activity, thereby offering a novel perspective for dioxygenase modification. The application of random mutagenesis (74) in the directed evolution of nitroarene dioxygenases has been limited in previous studies, primarily due to its potential to generate a vast library of mutants without efficient high-throughput screening methods.

This study introduces an *in silico* prediction approach for canonical site-directed mutagenesis to engineer 23DCNB dioxygenase with improved catalytic activity towards 24DCNB. The *in silico* results are basic anastomotic with the experimental findings, while the amount of time and labor involved in simulation is dramatically less than in the experiment. It’s expected that this approach could be extended to other

nitroarene dioxygenases for desired activity towards various nitroarenes. However, due to limitation in our computing resources, we were only able to construct and evaluate a small mutant library. It is possible that other mutations, which we have not yet identified, also play a role in the transformation of this 23DCNB dioxygenase into a 24DCNB dioxygenase. Expanding the mutant library tested during the pre-screening phase may unveil additional advantageous mutations, thereby augmenting the catalytic activity of final mutant dioxygenase towards 24DCNB.

Our results bear out that the substituted positions of the substrate, rather than the types of its substituted groups, greatly influence the fitting process of nitroarene substrates in the active pocket of dioxygenase (27, 47). This feature further gives rise to the catalytic promiscuity of dioxygenases: as shown in the mutant dioxygenase substrate spectrum above (**Fig. 4**), the activity of a dioxygenase towards various nitroarene substrates is diverse, and mutations often result in the transfer of this activity to a class of substrates with similar substituted patterns.

The E204M mutation brought about a shift in preference of dioxygenase from *meta*- to *ortho*-substituted substrates, resulting from a more compact substrate pocket around the C3 atom of substrate and thus a better fit of it. In general, the access through the enzyme channel and the orientation inside the active site of substrate both account for the catalytic efficiency, selectivity and specificity (75-77). In view of the gatekeeper role of residue 204, it would be arbitrary to ignore any of the possible impact of residue 204 on substrate as it passes through the enzyme tunnel or positioning in the active site. However, unlike $NDO_{9816\text{-}4}$, where an opening and obvious tunnel could be observed from the protein surface (78), our wild-type or mutant dioxygenases have occluded tunnels which could not be observed nor be calculated. Our results highlighted residue 204 as a critical residue lining the catalytic

cavity and supported a potential interaction between residue 204 and the *meta*-substituted group of nitroarenes (when it is properly bound to the active site) (18, 47).

The residue 248, although located outside the active site, has been identified as a crucial residue that potentially influences substrate specificity through its interaction with Asn295. Through this indirect effect, M248I mutation may exert some allosteric control on the active binding cavity around the C4 atom of substrate, making it easier to accommodate *para*-substituted nitroarenes. Our findings corroborate the research carried out by Ju *et al.* (17), in which long-term evolution provided 2-nitrotoluene dioxygenase with improved activity towards 4-nitrotoluene, and all three evolved mutants shared a common mutation—M248I.

Collectively, the coordinated interplay between E204M and M248I mutations retained the improved activity towards *ortho*- and *para*-substituted substrates, enhanced the catalytic activity towards 24DCNB, and thus the substrate spectrum was greatly broadened. E204M-M248I dioxygenase could be a versatile biocatalyst, supplying various nitroarene pollutants with hydroxyl groups for further transformation via ring cleavage and providing a new option for bioremediation in nitroarene-contaminated areas. The products transformed from 24DCNB was identified to be 35DCC, which can be completely mineralized through the degradation pathways discovered in *Cupriavidus necator* JMP134 (79-82), *Burkholderia cepacia* 2a (83), *Cupriavidus gilardii* T-1 (84).

Acknowledgements

This work was funded by the National Key R&D Program of China (2018YFA0901200 and 2021YFA0909500).

References

1. Booth G. 2000. Nitro compounds, aromatic. In Ullmann's Encyclopedia of Industrial Chemistry, (Ed.).
2. Ju K-S, Parales RE. 2010. Nitroaromatic compounds, from synthesis to biodegradation. Microbiol Mol Biol Rev 74:250–272.
3. Ju K-S, Parales RE. 2009. Application of nitroarene dioxygenases in the design of novel strains that degrade chloronitrobenzenes. Microb Biotechnol 2:241–252.
4. Isayev O, Rasulev B, Gorb L, Leszczynski J. 2006. Structure-toxicity relationships of nitroaromatic compounds. Mol Diversity 10:233-245.
5. Agency USEP. 2014. Priority Pollutant List. https://www.epa.gov/sites/default/files/2015-09/documents/priority-pollutant-list-epa.pdf. Accessed 8th July 2023.
6. Haigler BE, Wallace WH, Spain JC. 1994. Biodegradation of 2-nitrotoluene by *Pseudomonas* sp. strain JS42. Appl Environ Microbiol 60:3466-3469.
7. Nishino SF, Spain JC. 1995. Oxidative pathway for the biodegradation of nitrobenzene by *Comamonas* sp. Strain JS765. Appl Environ Microbiol 61:2308-2313.
8. Singh D, Ramanathan G. 2013. Biomineralization of 3-nitrotoluene by *Diaphorobacter* species. Biodegradation 24:645-655.
9. Johnson GR, Jain RK, Spain JC. 2000. Properties of the trihydroxytoluene oxygenase from *Burkholderia cepacia* R34: an extradiol dioxygenase from the 2,4-dinitrotoluene pathway. Arch Microbiol 173:86-90.
10. Johnson GR, Jain RK, Spain JC. 2002. Origins of the 2,4-dinitrotoluene pathway. J Bacteriol 184:4219-4232.
11. Liu H, Wang S-J, Zhou N-Y. 2005. A new isolate of *Pseudomonas stutzerithat* degrades 2-chloronitrobenzene. Biotechnol Lett 27:275-278.
12. Palatucci ML, Waidner LA, Mack EE, Spain JC. 2019. Aerobic biodegradation of 2,3- and 3,4-dichloronitrobenzene. J Hazard Mater 378:120717.
13. Parales RE. 2003. The role of active-site residues in naphthalene dioxygenase. J Ind Microbiol Biotechnol 30:271-278.
14. Pabis A, Geronimo I, York DM, Paneth P. 2014. Molecular dynamics simulation of nitrobenzene dioxygenase using AMBER force field. J Chem Theory Comput 10:2246-2254.
15. Lee K-S, Parales JV, Friemann R, Parales RE. 2005. Active site residues controlling substrate specificity in 2-nitrotoluene dioxygenase from *Acidovorax* sp. strain JS42. J Ind Microbiol Biotechnol 32:465-473.
16. Ju K-S, Parales RE. 2006. Control of substrate specificity by active-site residues in nitrobenzene dioxygenase. Appl Environ Microbiol 72:1817-1824.
17. Ju K-S, Parales RE. 2011. Evolution of a new bacterial pathway for 4-nitrotoluene degradation. Mol Microbiol 82:355-364.
18. Mahan KM, Penrod JT, Ju K-S, Al Kass N, Tan WA, Truong R, Parales JV, Parales RE. 2015. Selection for growth on 3-nitrotoluene by 2-nitrotoluene-utilizing *Acidovorax* sp. strain JS42 identifies nitroarene dioxygenases with altered specificities. Appl Environ Microbiol 81:309-319.

19. Karlsson A. 2003. Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. Science 299:1039-1042.
20. Friemann R, Ivkovic-Jensen MM, Lessner DJ, Yu C-L, Gibson DT, Parales RE, Eklund H, Ramaswamy S. 2005. Structural insight into the dioxygenation of nitroarene compounds: the crystal structure of nitrobenzene dioxygenase. J Mol Biol 348:1139-1151.
21. Csizi KS, Eckert L, Brunken C, Hofstetter TB, Reiher M. 2022. The apparently unreactive substrate facilitates the electron transfer for dioxygen activation in rieske dioxygenases. Chem Eur J 28:80.
22. Wolfe MD, Parales JV, Gibson DT, Lipscomb JD. 2001. Single turnover chemistry and regulation of O_2 activation by the oxygenase component of naphthalene 1,2-dioxygenase. J Biol Chem 276:1945-1953.
23. Wolfe MD, Lipscomb JD. 2003. Hydrogen peroxide-coupled *cis*-diol formation catalyzed by naphthalene 1,2-dioxygenase. J Biol Chem 278:829-835.
24. Bassan A, Blomberg MRA, Siegbahn PEM. 2004. A theoretical study of the *cis*-dihydroxylation mechanism in naphthalene 1,2-dioxygenase. JBIC, J Biol Inorg Chem 9:439-452.
25. Bassan A, Borowski T, Siegbahn PEM. 2004. Quantum chemical studies of dioxygen activation by mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad. Dalton Trans 20:3153-3162.
26. Sutherlin KD, Rivard BS, Böttger LH, Liu LV, Rogers MS, Srnec M, Park K, Yoda Y, Kitao S, Kobayashi Y, Saito M, Seto M, Hu M, Zhao J, Lipscomb JD, Solomon EI. 2018. NRVS studies of the peroxide shunt intermediate in a Rieske dioxygenase and its relation to the native $Fe^{II}O_2$ reaction. J Am Chem Soc 140:5544-5559.
27. Bopp CE, Bernet NM, Kohler H-PE, Hofstetter TB. 2022. Elucidating the role of O_2 uncoupling in the oxidative biodegradation of organic contaminants by Rieske non-heme iron dioxygenases. ACS Environ Au 2:428-440.
28. Pati SG, Bopp CE, Kohler H-PE, Hofstetter TB. 2022. Substrate-specific coupling of O_2 activation to hydroxylations of aromatic compounds by Rieske non-heme iron dioxygenases. ACS Catal 12:6444-6456.
29. Agency. USEP. The United States High Production Volume (USHPV) database. https://comptox.epa.gov/dashboard/chemical-lists/EPAHPV&search=DTXSID3024998. Accessed 8th July 2023.
30. Chen H, Gao X, Wang C, Shao J, Xu X, Zhu L. 2017. Efficient 2,4-dichloronitrobenzene removal in the coupled BES-UASB reactor: Effect of external voltage mode. Bioresour Technol 241:879-886.
31. Chen L, Shao J, Chen H, Wang C, Gao X, Xu X, Zhu L. 2018. Cathode potential regulation in a coupled bioelectrode-anaerobic sludge system for effective dechlorination of 2,4-dichloronitrobenzene. Bioresour Technol 254:180-186.
32. Liu Y, Wang C, Zhang K, Zhou Y, Xu Y, Xu X, Zhu L. 2020. Rapid degradation of 2,4-dichloronitrobenzene in single-chamber microbial electrolysis cell with pre-acclimated bioanode: A comprehensive assessment. Sci Total Environ 724:138053.
33. Kano H, Suzuki M, Senoh H, Yamazaki K, Aiso S, Matsumoto M, Nagano K, Fukushima S. 2012. 2,4-Dichloro-1-nitrobenzene exerts carcinogenicities in both rats and mice by two years feeding. Arch Toxicol 86:1763-1772.

34. Aleksic M, Pease CK, Basketter DA, Panico M, Morris HR, Dell A. 2008. Mass spectrometric identification of covalent adducts of the skin allergen 2,4-dinitro-1-chlorobenzene and model skin proteins. Toxicol In Vitro 22:1169-1176.
35. Agency USEP. International Agency for Research on Cancer (IARC): Group 2B: Possibly carcinogenic to humans. https://comptox.epa.gov/dashboard/chemical-lists/IARC2B&search=DTXSID3024998. Accessed 8th July 2023.
36. Jiang X, Shen J, Han Y, Lou S, Han W, Sun X, Li J, Mu Y, Wang L. 2016. Efficient nitro reduction and dechlorination of 2,4-dinitrochlorobenzene through the integration of bioelectrochemical system into upflow anaerobic sludge blanket: A comprehensive study. Water Res 88:257-265.
37. Chen H, Lu D, Chen L, Wang C, Xu X, Zhu L. 2019. A study of the coupled bioelectrochemical system-upflow anaerobic sludge blanket for efficient transformation of 2,4-dichloronitrobenzene. Environ Sci Pollut Res 26:13002-13013.
38. Amrein BA, Steffen-Munsberg F, Szeler I, Purg M, Kulkarni Y, Kamerlin SCL. 2017. CADEE: Computer-Aided Directed Evolution of Enzymes. IUCrJ 4:50-64.
39. Wu L, Qin L, Nie Y, Xu Y, Zhao Y-L. 2022. Computer-aided understanding and engineering of enzymatic selectivity. Biotechnol Adv 54:107793.
40. Go MK, Zhao LN, Xue B, Supekar S, Robinson RC, Fan H, Yew WS. 2020. Directed computational evolution of quorum-quenching lactonases from the amidohydrolase superfamily. Structure 28:635-642.e3.
41. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296-W303.
42. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583-589.
43. Trott O, Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455-461.
44. Vavra O, Filipovic J, Plhak J, Bednar D, Marques SM, Brezovsky J, Stourac J, Matyska L, Damborsky J. 2019. CaverDock: a molecular docking-based tool to analyse ligand transport through protein tunnels and channels. Bioinformatics 35:4986-4993.
45. Franz F, Daday C, Gräter F. 2020. Advances in molecular simulations of protein mechanical properties and function. Curr Opin Struct Biol 61:132-138.
46. Surpeta B, Sequeiros-Borja C, Brezovsky J. 2020. Dynamics, a powerful component of current and future *in silico* approaches for protein design and engineering. Int J Mol Sci 21:2713.
47. Li T, Gao Y-Z, Xu J, Zhang S-T, Guo Y, Spain JC, Zhou N-Y. 2021. A recently assembled degradation pathway for 2,3-dichloronitrobenzene in *Diaphorobacter* sp. strain JS3051. mBio 12:e0223121.
48. An D, Gibson DT, Spain JC. 1994. Oxidative release of nitrite from 2-nitrotoluene by a

three-component enzyme system from *Pseudomonas* sp. strain JS42. J Bacteriol 176:7462-7467.
49. Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254.
50. Anandakrishnan R, Aguilar B, Onufriev AV. 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res 40:W537-W541.
51. Li P, Merz KM. 2016. MCPB.py: A python based metal center parameter builder. J Chem Inf Model 56:599-604.
52. Peters MB, Yang Y, Wang B, Füsti-Molnár L, Weaver MN, Merz KM. 2010. Structural survey of zinc-containing proteins and development of the zinc AMBER force field (ZAFF). J Chem Theory Comput 6:2935-2947.
53. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery Jr. JA, Peralta JE, Ogliaro F, et al. 2016. Gaussian 16 Rev. C.01, Wallingford, CT.
54. Gaussian. Inc., 340 Quinnipiac Street, Building 40, Wallingford, CT 06392.
55. Sousa da Silva AW, Vranken WF. 2012. ACPYPE - AnteChamber PYthon Parser interfacE. BMC Res Notes 5:367.
56. Lindahl A, Hess, & van der Spoel. 2020. GROMACS 2020 Source code (Version 2020). https://doi.org/10.5281/zenodo.3562495.
57. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. 1997. LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18:1463-1472.
58. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. 1995. A smooth particle mesh Ewald method. J Chem Phys 103:8577-8593.
59. Schrodinger, LLC. 2015. The JyMOL molecular graphics development component, version 1.8.
60. Wagner JR, Sørensen J, Hensley N, Wong C, Zhu C, Perison T, Amaro RE. 2017. POVME 3.0: Software for mapping binding pocket flexibility. J Chem Theory Comput 13:4584-4592.
61. Yu H, Ma S, Li Y, Dalby PA. 2022. Hot spots-making directed evolution easier. Biotechnol Adv 56:107926.
62. Kumari A, Singh D, Ramaswamy S, Ramanathan G. 2017. Structural and functional studies of ferredoxin and oxygenase components of 3-nitrotoluene dioxygenase from *Diaphorobacter* sp. strain DS2. PLoS One 12:e0176398.
63. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189-1191.
64. Dannenberg JJ. 1998. An introduction to hydrogen bonding by George A. Jeffrey (University of Pittsburgh). J Am Chem Soc 120:5604-5604.
65. Starr TN, Thornton JW. 2016. Epistasis in protein evolution. Protein Science 25:1204-1218.
66. Wang Z, Zhou H, Yu H, Pu Z, Xu J, Zhang H, Wu J, Yang L. 2022. Computational redesign of

the substrate binding pocket of glutamate dehydrogenase for efficient synthesis of noncanonical L-amino acids. ACS Catal 12:13619-13629.

67. Yu H, Dalby PA. 2018. Coupled molecular dynamics mediate long- and short-range epistasis between mutations that affect stability and aggregation kinetics. Proc Natl Acad Sci U S A 115:E11043-E11052.

68. Liu H, Wang S-J, Zhang J-J, Dai H, Tang H, Zhou N-Y. 2011. Patchwork assembly of nag-like nitroarene dioxygenase genes and the 3-chlorocatechol degradation cluster for evolution of the 2-chloronitrobenzene catabolism pathway in *Pseudomonas stutzeri* ZWLR2-1. Appl Environ Microbiol 77:4547-4552.

69. Gao Y-Z, Liu X-Y, Liu H, Guo Y, Zhou N-Y. 2020. A Bph-like nitroarene dioxygenase catalyzes the conversion of 3-nitrotoluene to 3-methylcatechol by *Rhodococcus* sp. strain ZWL3NT. Appl Environ Microbiol 86: e02517-19.

70. Li T, Xu J, Brower AL, Xu Z-J, Xu Y, Spain JC, Zhou N-Y. 2023. Molecular basis and evolutionary origin of 1-nitronaphthalene catabolism in *Sphingobium* sp. strain JS3065. Appl Environ Microbiol 89:e0172822.

71. Yu CL, Parales RE, Gibson DT. 2001. Multiple mutations at the active site of naphthalene dioxygenase affect regioselectivity and enantioselectivity. J Ind Microbiol Biotechnol 27:94-103.

72. Parales RE, Resnick SM, Yu CL, Boyd DR, Sharma ND, Gibson DT. 2000. Regioselectivity and enantioselectivity of naphthalene dioxygenase during arene cis-dihydroxylation: control by phenylalanine 352 in the alpha subunit. J Bacteriol 182:5495-504.

73. Seo J, Ryu JY, Han J, Ahn JH, Sadowsky MJ, Hur HG, Chong Y. 2013. Amino acid substitutions in naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816-4 result in regio- and stereo-specific hydroxylation of flavanone and isoflavanone. Appl Microbiol Biotechnol 97:693-704.

74. Bernath-Levin K, Shainsky J, Sigawi L, Fishman A. 2014. Directed evolution of nitrobenzene dioxygenase for the synthesis of the antioxidant hydroxytyrosol. Appl Microbiol Biotechnol 98:4975-4985.

75. Kreß N, Halder JM, Rapp LR, Hauer B. 2018. Unlocked potential of dynamic elements in protein structures: channels and loops. Curr Opin Chem Biol 47:109-116.

76. Kokkonen P, Bednar D, Pinto G, Prokop Z, Damborsky J. 2019. Engineering enzyme access tunnels. Biotechnol Adv 37:107386.

77. Liu J, Tian J, Perry C, Lukowski AL, Doukov TI, Narayan ARH, Bridwell-Rabb J. 2022. Design principles for site-selective hydroxylation by a Rieske oxygenase. Nat Commun. 13:255.

78. Escalante DE, Aukema KG, Wackett LP, Aksan A. 2017. Simulation of the bottleneck controlling access into a Rieske active site: predicting substrates of naphthalene 1,2-dioxygenase. J Chem Inf Model 57:550-561.

79. Kumar A, Trefault N, Olaniran AO. 2014. Microbial degradation of 2,4-dichlorophenoxyacetic acid: Insight into the enzymes and catabolic genes involved, their regulation and biotechnological implications. Crit Rev Microbiol 42:1-15.

80. Pérez-Pantoja D, Ledger T, Pieper DH, González B. 2003. Efficient turnover of chlorocatechols is essential for growth of *Ralstonia eutropha* JMP134(pJP4) in 3-chlorobenzoic acid. J Bacteriol 185:1534-1542.

81. Plumeier I, PéRez-Pantoja D, Heim S, GonzáLez B, Pieper DH. 2002. Importance of different *tfd* genes for degradation of chloroaromatics by *Ralstonia eutropha* JMP134. J Bacteriol 184:4054-4064.
82. Pérez-Pantoja D, Guzmán L, Manzano M, Pieper DH, González B. 2000. Role of *tfd*C(I)D(I)E(I)F(I) and *tfd*D(II)C(II)E(II)F(II) gene modules in catabolism of 3-chlorobenzoate by *Ralstonia eutropha* JMP134(pJP4). Appl Environ Microbiol 66:1602-1608.
83. Smith ARW, Beadle CA. 2008. Induction of enzymes of 2,4-dichlorophenoxyacetate degradation in *Burkholderia cepacia* 2a and toxicity of metabolic intermediates. Biodegradation 19:669-681.
84. Wu X, Wang W, Liu J, Pan D, Tu X, Lv P, Wang Y, Cao H, Wang Y, Hua R. 2017. Rapid biodegradation of the herbicide 2,4-dichlorophenoxyacetic acid by *Cupriavidus gilardii* T-1. J Agric Food Chem 65:3711-3720.

Figure legends

FIG 1 23DCNB dioxygenase mutant library design. (A) Comparison of various Nag-like dioxygenase gene clusters, arranged in order of pairwise identity. The corresponding substrates are presented on the right, with their attack sites of dioxygenation marked in red shades. **(B)** Sequence alignment of the α subunit of various Nag-like dioxygenases. The degree of conservation was calculated by Jalview (63), shown from low (white) to high (blue) as well. Six chosen sites are highlighted by frames in red. Chosen reasons are given above the sites in form of small circles in corresponding colors. **(C)** Chosen residues for mutation marked in the homologous model of α subunit of 23DCNB dioxygenase.

FIG 2 Parameters for the evaluation of H-bond formation probability and dioxygenation attack distance.

FIG 3 Predicted catalytic potential evaluated by H-bond formation probability (A) and dioxygenation attack distance (B) and actual specific activity (C) of wild-type and mutant dioxygenases on 24DCNB. Each mutant has nine sets of data in (A) and (B) because there were three active sites in each $\alpha_3\beta_3$ system which was repeated three times. The specific activity was obtained through whole-cell biotransformation assays with 23DCNB and 24DCNB as substrates. Values are data averages from at least three parallel experiments, and error bars are standard deviations.

FIG 4 Substrate specificity of wild-type dioxygenase, E204M, M248I, and E204M-M248I mutants towards different nitroarenes. The specific activity was obtained through whole-cell biotransformation assays. Values are data averages from at least three parallel experiments, and error bars are standard deviations. 2NT, 2-nitrotoluene; 3NT, 3-nitrotoluene; 4NT, 4-nitrotoluene; 2CNB, 2-chloronitrobenzene; 3CNB, 3-chloronitrobenzene; 4CNB, 4-chloronitrobenzene; 4C3NT, 4-chloro-3-nitrotoluene; 6C2NT, 6-chloro-2-nitrotoluene; 2C3NT, 2-chloro-3-nitrotoluene; 23DCNB, 2,3-dichloronitrobenzene; 24DCNB, 2,4-dichloronitrobenzene; 25DCNB, 2,5-dichloronitrobenzene; 34DCNB, 3,4-dichloronitrobenzene; 35DCNB, 3,5-dichloronitrobenzene; 24DNCB,

2,4-dinitrochlorobenzene; 24DNT, 2,4-dinitrotoluene; 26DNT, 2,6-dinitrotoluene; 1NN, 1-nitronaphthalene.

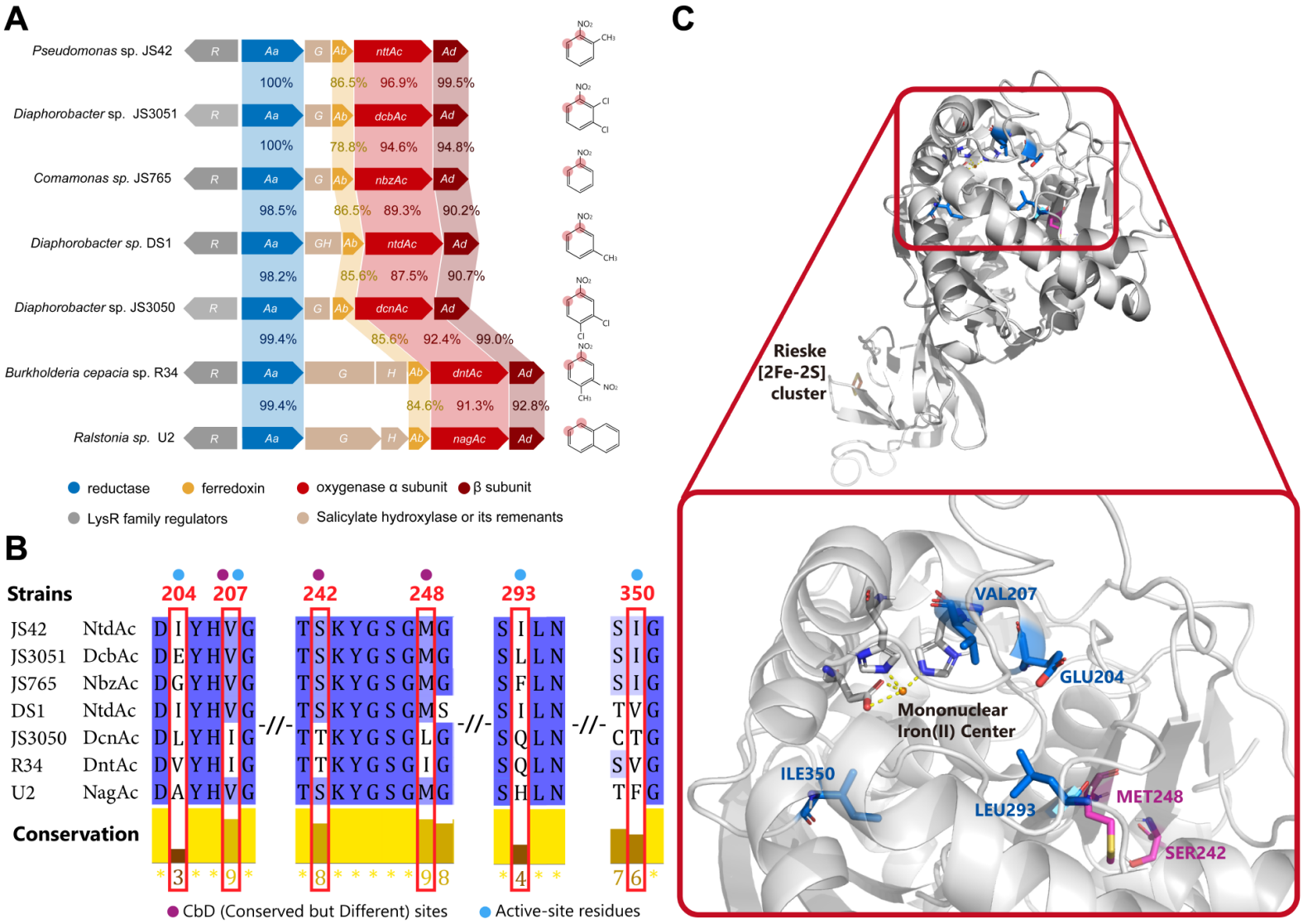
FIG 5 Whole-cell biotransformation of 24DCNB. (A) Reaction scheme of 24DCNB biotransformation. (B) Time course of 24DCNB conversion to 35DCC with nitrite release by IPTG-induced *E. coli* BL21(DE3) expressing E204M-M248I dioxygenase. The data shown are the averages of three technical duplicates of a typical experiment, and the outcomes of all separate trials were broadly consistent. Standard deviations are shown by error bars. (C, D and E) GC/MS total ion chromatogram. The product of 24DCNB biotransformation catalyzed by wild-type (C) and E204M-M248I (D) was detected and compared with the authentic 35DCC (E). The mass spectra refer to 35DCC after derivatization.

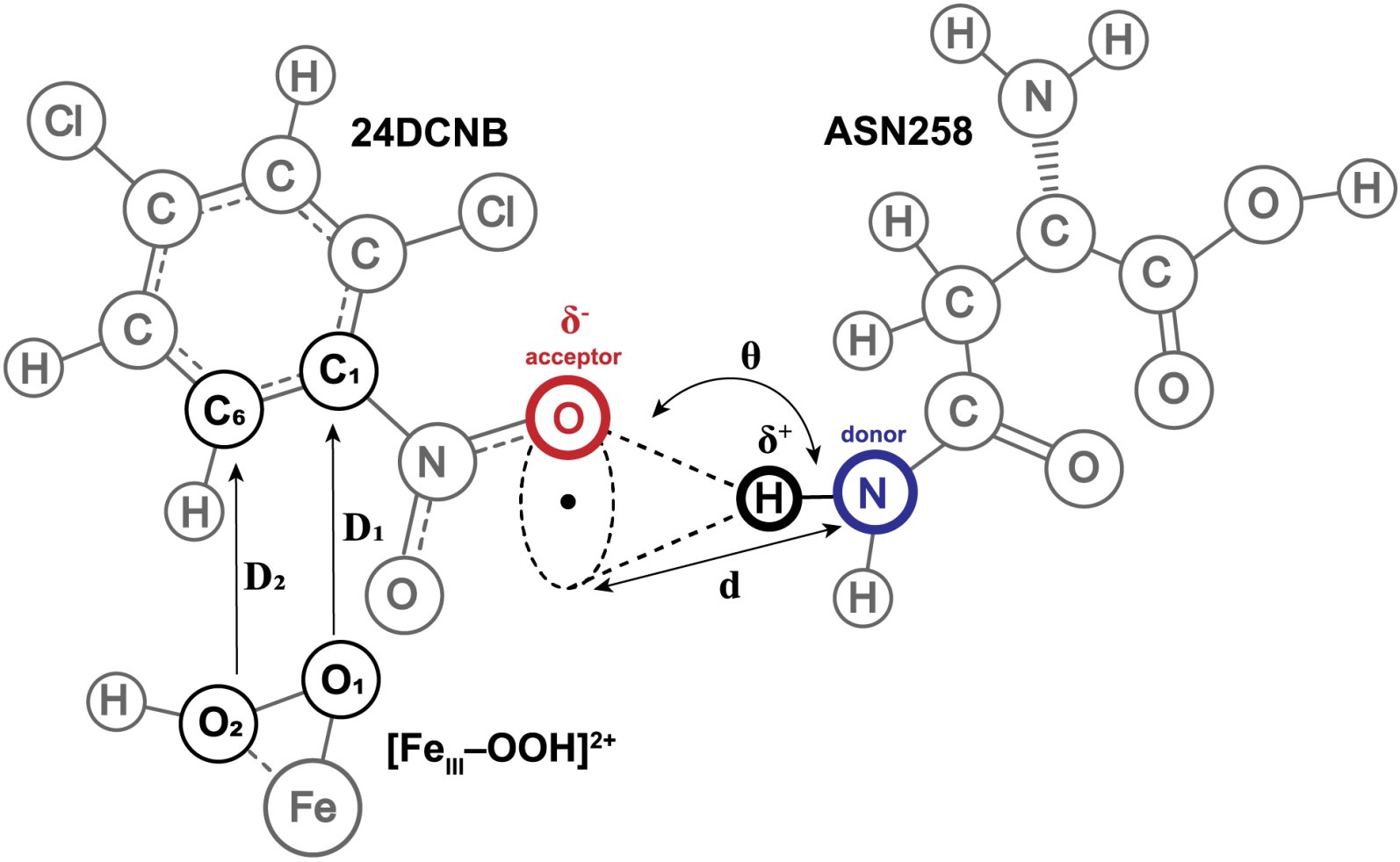
FIG 6 Pocket volume analysis and stability analysis of substrate docking. (A) The superposition of binding pockets of wild-type (grey) and E204M mutant (yellow) captured every 5 ns during 100-200 ns MD simulation trajectory (take α1 as an example). (B) The average pocket volume of each α subunit. (C) RMSF values of the 24DCNB in each α subunit.

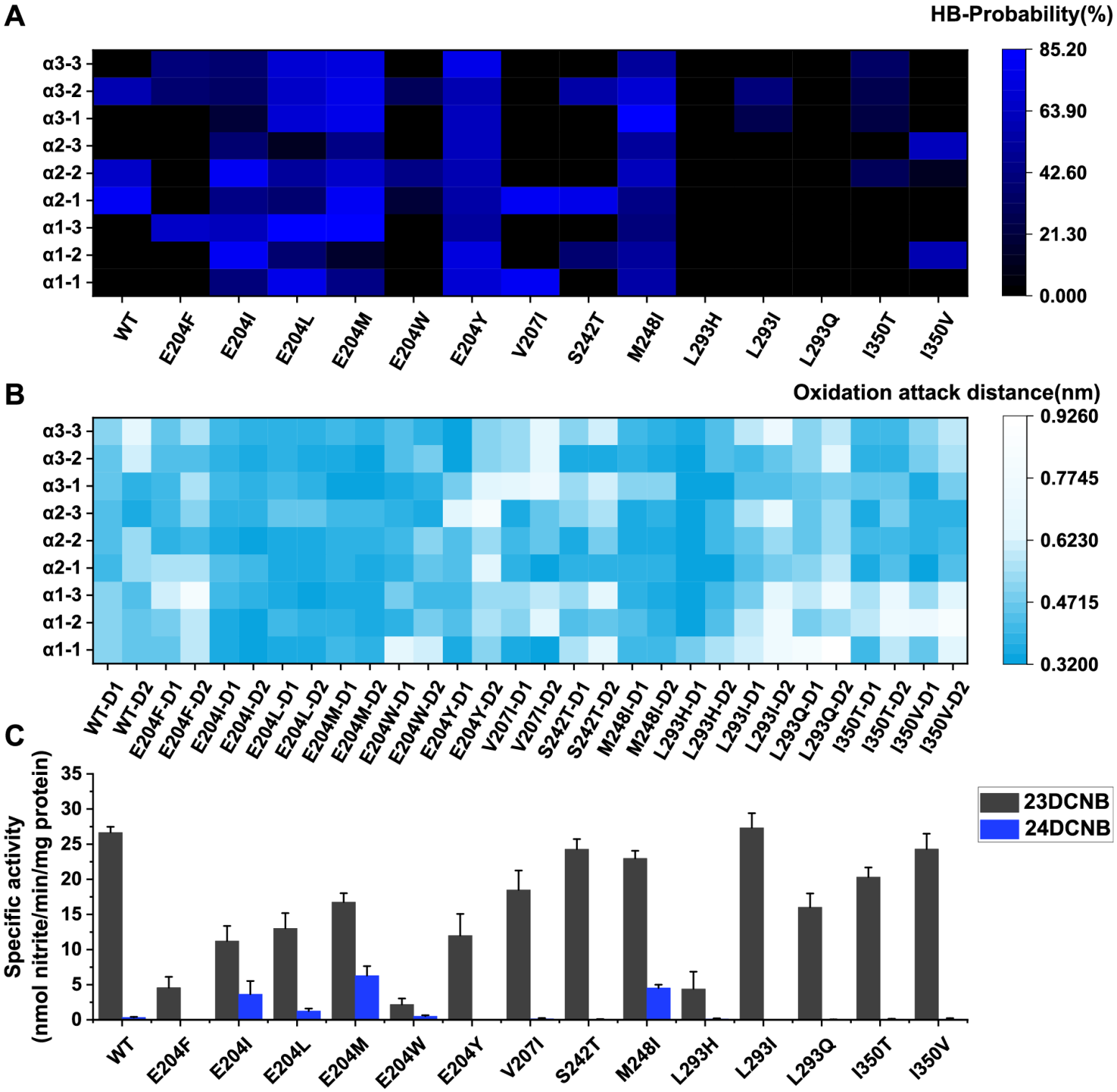
Table

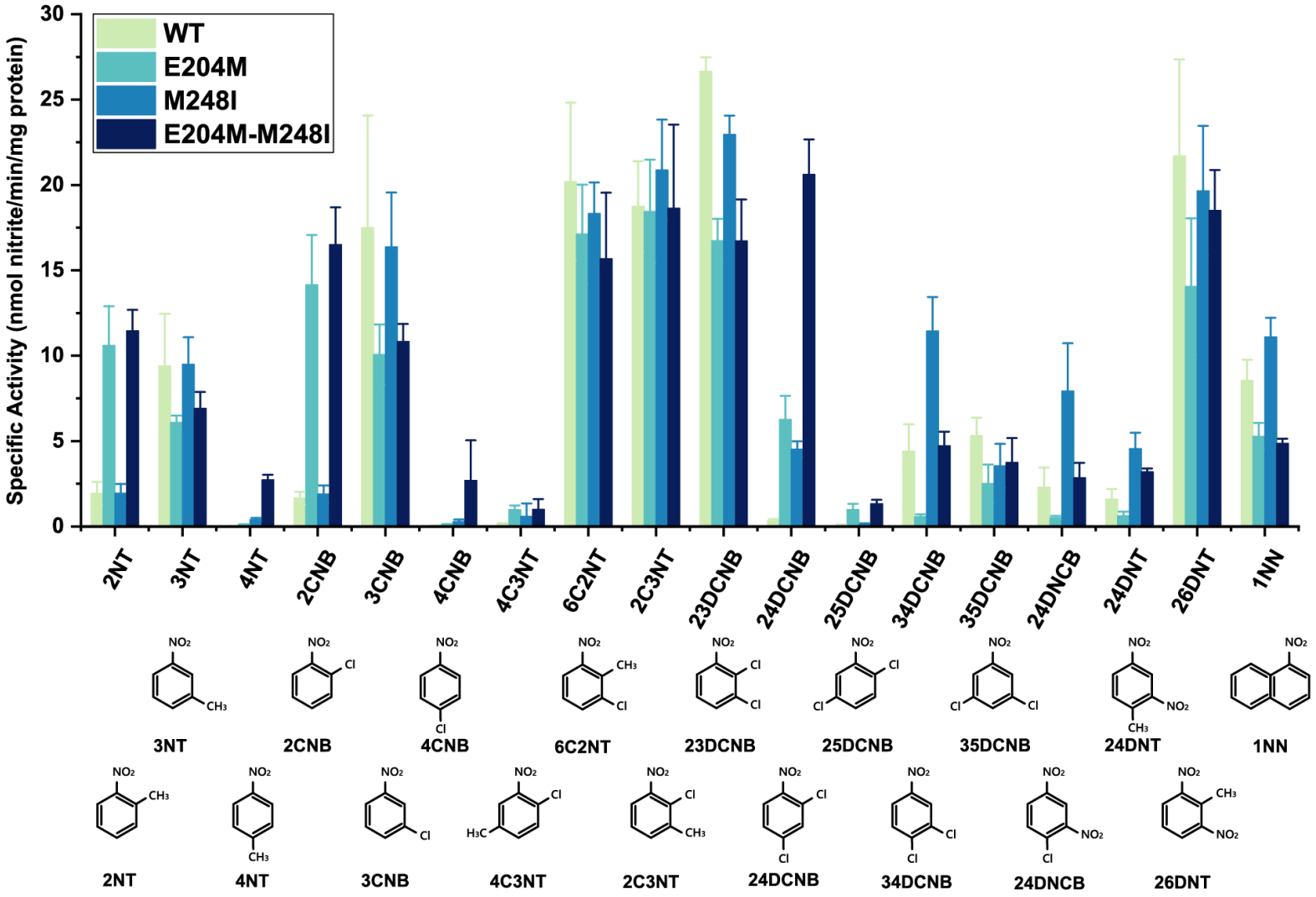
TABLE 1 mutagenic primers used in this study.

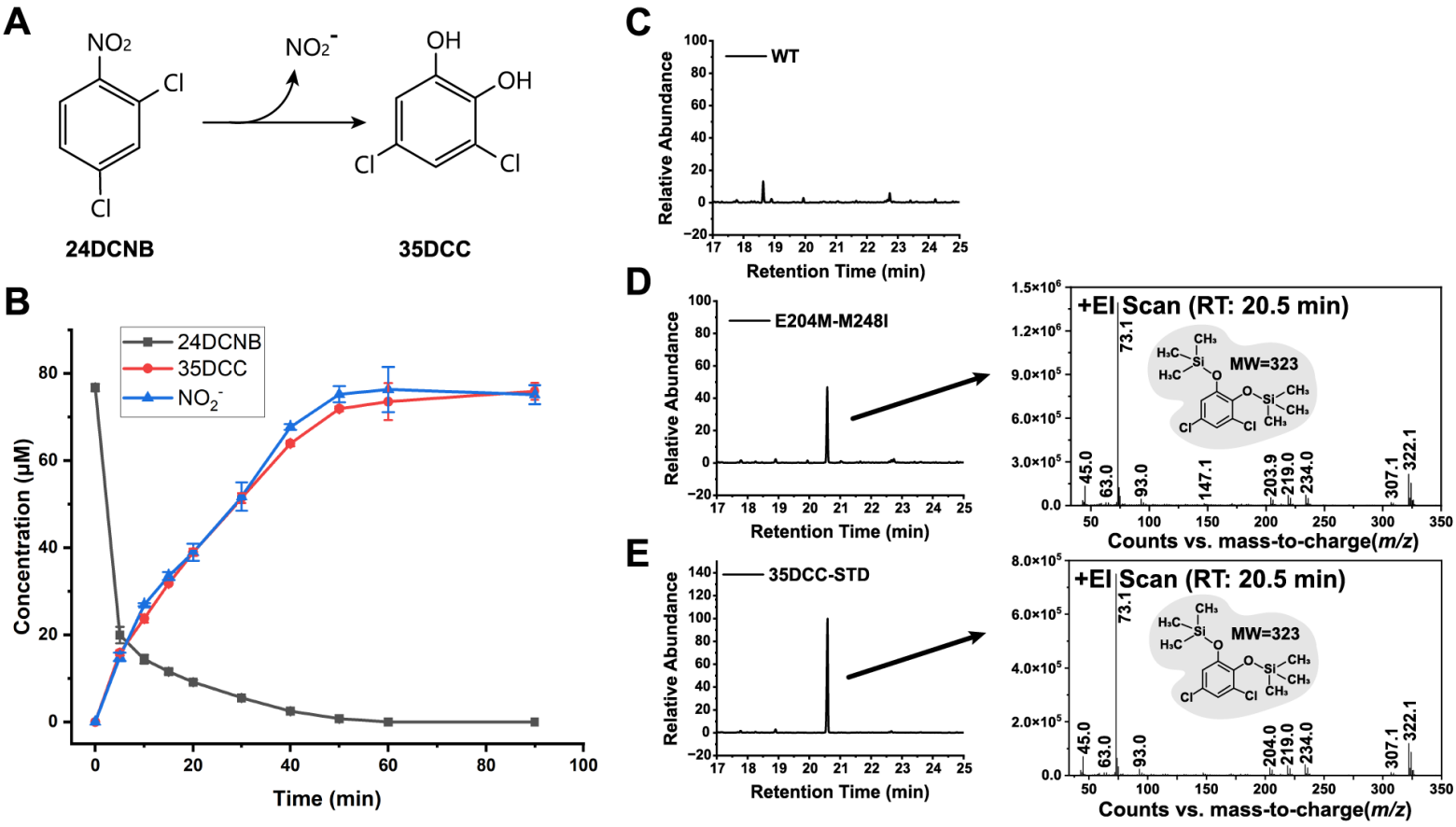
Primers	Sequence
E204F-F	5'-GCTGAAAACTTCGTTGGTGACTTCTACCACGTTGGTTGGACCCAC-3'
E204F-R	5'-GTGGGTCCAACCAACGTGGTAGAAGTCACCAACGAAGTTTTCAGC-3'
E204I-F	5'-TGAAAACTTCGTTGGTGACATATACCACGTTGGTTGGACC-3'
E204I-R	5'-GGTCCAACCAACGTGGTATATGTCACCAACGAAGTTTTCA-3'
E204L-F	5'-GGTCCAACCAACGTGGTATATGTCACCAACGAAGTTTTCA-3'
E204L-R	5'-GGTCCAACCAACGTGGTATAAGTCACCAACGAAGTTTTCA-3'
E204M-F	5'-GCTGAAAACTTCGTTGGTGACATGTACCACGTTGGTTGGACCCAC-3'
E204M-R	5'-GTGGGTCCAACCAACGTGGTACATGTCACCAACGAAGTTTTCAGC-3'
E204W-F	5'-GCTGAAAACTTCGTTGGTGACTGGTACCACGTTGGTTGGACCCAC-3'
E204W-R	5'-GTGGGTCCAACCAACGTGGTACCAGTCACCAACGAAGTTTTCAGC-3'
E204Y-F	5'-GTGGGTCCAACCAACGTGGTACCAGTCACCAACGAAGTTTTCAGC-3'
E204Y-R	5'-GTGGGTCCAACCAACGTGGTACCAGTCACCAACGAAGTTTTCAGC-3'
V207I-F	5'-GGTGACGAATACCACATTGGTTGGACCCACG-3'
V207I-R	5'-CGTGGGTCCAACCAATGTGGTATTCGTCACC-3'
S242T-F	5'-CTGGTCTGCAGATGACCACTAAATACGGTTCTGGT-3'
S242T-R	5'-ACCAGAACCGTATTTAGTGGTCATCTGCAGACCAG-3'
M248I-F	5'-ACCTCTAAATACGGTTCTGGTATAGGTCTGACCTGG-3'
M248I-R	5'-CCAGGTCAGACCTATACCAGAACCGTATTTAGAGGT-3'
L293H-F	5'-GCTCGTATCTACCGTTCTCATCTGAACGGTACTGTTTTCC-3'
L293H-R	5'-GGAAAACAGTACCGTTCAGATGAGAACGGTAGATACGAGC-3'
L293I-F	5'-TGCTCGTATCTACCGTTCTATACTGAACGGTACTGTTTTCC-3'
L293I-R	5'-GGAAAACAGTACCGTTCAGTATAGAACGGTAGATACGAGCA-3'
L293Q-F	5'-TCGTATCTACCGTTCTCAGCTGAACGGTACTGTTT-3'
L293Q-R	5'-AAACAGTACCGTTCAGCTGAGAACGGTAGATACGA-3'
I350T-F	5'-CGCTGTTCAGCGTTCTACCGGTCCGGC-3'
I350T-R	5'-GCCGGACCGGTAGAACGCTGAACAGCG-3'
I350V-F	5'-TGACGCTGTTCAGCGTTCTGTCGGTCCGGC-3'
I350V-R	5'-GCCGGACCGACAGAACGCTGAACAGCGTCA-3'

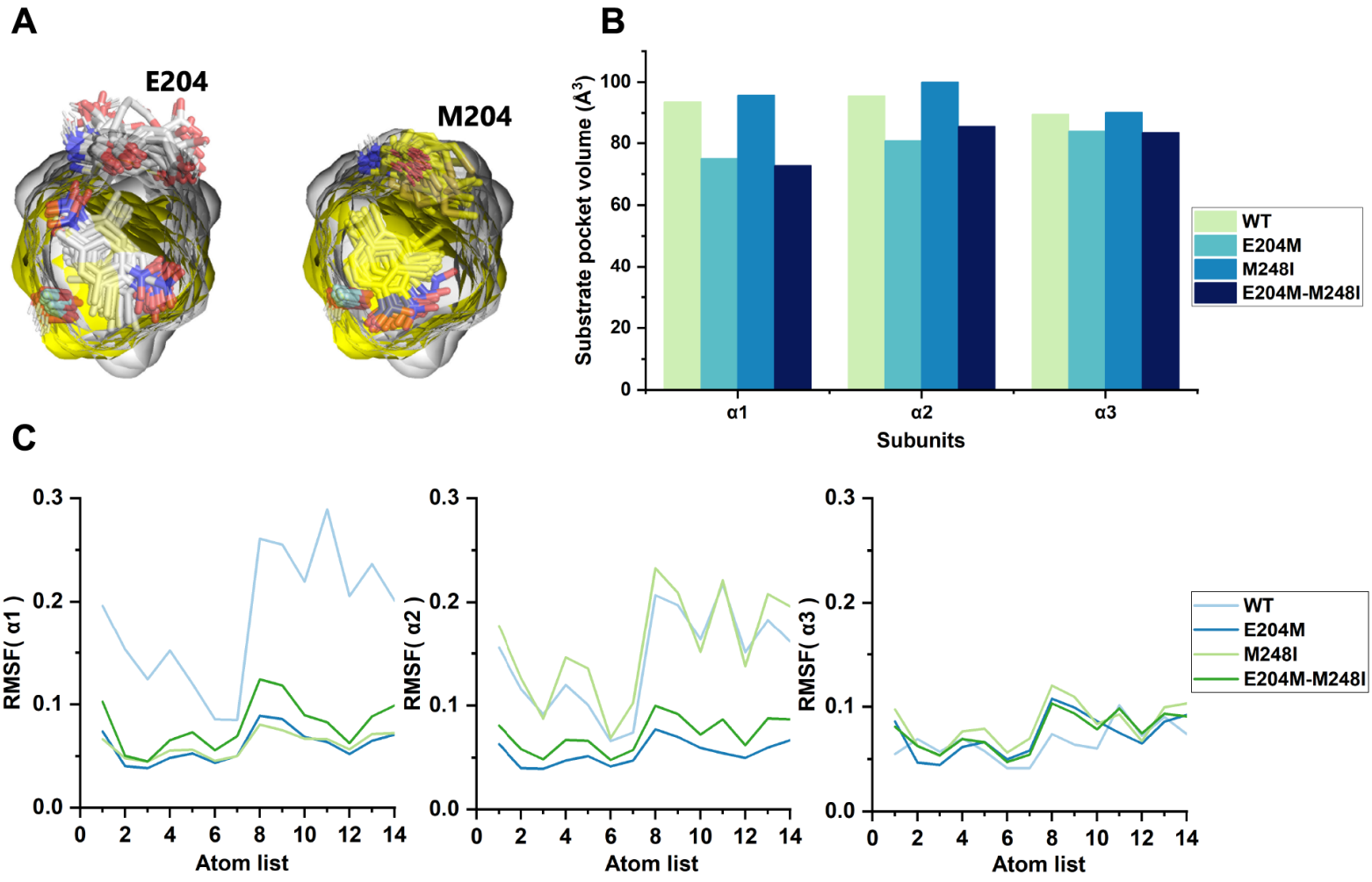

A
Pseudomonas sp. JS42
Diaphorobacter sp. JS3051
Comamonas sp. JS765
Diaphorobacter sp. DS1
Diaphorobacter sp. JS3050
Burkholderia cepacia sp. R34
Ralstonia sp. U2
R
Aa
G
Ab
nttAc
Ad
dcbAc
nbzAc
GH
ntdAc
dcnAc
H
dntAc
nagAc
100%
86.5%
96.9%
99.5%
78.8%
94.6%
94.8%
98.5%
86.5%
89.3%
90.2%
98.2%
85.6%
87.5%
90.7%
99.4%
85.6%
92.4%
99.0%
99.4%
84.6%
91.3%
92.8%
reductase
ferredoxin
oxygenase α subunit
β subunit
LysR family regulators
Salicylate hydroxylase or its remenants
B
Strains
204
207
242
248
293
350
JS42 NtdAc D I Y H V G T S K Y G S G M G S I L N S I G
JS3051 DcbAc D E Y H V G T S K Y G S G M G S L L N S I G
JS765 NbzAc D G Y H V G T S K Y G S G M G S F L N S I G
DS1 NtdAc D I Y H V G T S K Y G S G M S S I L N T V G
JS3050 DcnAc D L Y H I G T T K Y G S G L G S Q L N C T G
R34 DntAc D V Y H I G T T K Y G S G I G S Q L N S V G
U2 NagAc D A Y H V G T S K Y G S G M G S H L N T F G
-//-
Conservation
* 3 * * 9 *
* 8 * * * * * 9 8
* 4 * *
7 6 *
CbD (Conserved but Different) sites
Active-site residues
C
Rieske
[2Fe-2S]
cluster
VAL207
GLU204
Mononuclear
Iron(II) Center
ILE350
LEU293
MET248
SER242


24DCNB
ASN258
δ-
acceptor
θ
δ+
donor
D₁
D₂
d
[Fe$_{III}$–OOH]$^{2+}$


A
HB-Probability(%)
85.20
63.90
42.60
21.30
0.000
α3-3
α3-2
α3-1
α2-3
α2-2
α2-1
α1-3
α1-2
α1-1
WT
E204F
E204I
E204L
E204M
E204W
E204Y
V207I
S242T
M248I
L293H
L293I
L293Q
I350T
I350V
B
Oxidation attack distance(nm)
0.9260
0.7745
0.6230
0.4715
0.3200
WT-D1
WT-D2
E204F-D1
E204F-D2
E204I-D1
E204I-D2
E204L-D1
E204L-D2
E204M-D1
E204M-D2
E204W-D1
E204W-D2
E204Y-D1
E204Y-D2
V207I-D1
V207I-D2
S242T-D1
S242T-D2
M248I-D1
M248I-D2
L293H-D1
L293H-D2
L293I-D1
L293I-D2
L293Q-D1
L293Q-D2
I350T-D1
I350T-D2
I350V-D1
I350V-D2
C
Specific activity
(nmol nitrite/min/mg protein)
0
5
10
15
20
25
30
35
23DCNB
24DCNB
WT
E204F
E204I
E204L
E204M
E204W
E204Y
V207I
S242T
M248I
L293H
L293I
L293Q
I350T
I350V


WT
E204M
M248I
E204M-M248I
Specific Activity (nmol nitrite/min/mg protein)
0
5
10
15
20
25
30
2NT
3NT
4NT
2CNB
3CNB
4CNB
4C3NT
6C2NT
2C3NT
23DCNB
24DCNB
25DCNB
34DCNB
35DCNB
24DNCB
24DNT
26DNT
1NN


A
NO2
Cl
Cl
NO2-
OH
OH
Cl
Cl
24DCNB
35DCC
B
24DCNB
35DCC
NO2-
Concentration (μM)
Time (min)
C
WT
Relative Abundance
Retention Time (min)
D
E204M-M248I
Relative Abundance
Retention Time (min)
+EI Scan (RT: 20.5 min)
MW=323
Counts vs. mass-to-charge(m/z)
E
35DCC-STD
Relative Abundance
Retention Time (min)
+EI Scan (RT: 20.5 min)
MW=323
Counts vs. mass-to-charge(m/z)


A
E204
M204
B
Substrate pocket volume (Å³)
0
20
40
60
80
100
α1
α2
α3
Subunits
WT
E204M
M248I
E204M-M248I
C
RMSF(α1)
RMSF(α2)
RMSF(α3)
0.0
0.1
0.2
0.3
0
2
4
6
8
10
12
14
Atom list
WT
E204M
M248I
E204M-M248I