Protein engineering of nirobenzene dioxygenase for enantioselective synthesis of chiral sulfoxides

Janna Shainsky[1], Kalia Bernath-Levin[1], Sivan Isaschar-Ovdat[1], Fabian Glaser[2] and Ayelet Fishman[1,3]

[1]Department of Biotechnology and Food Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel and [2]Bioinformatics Knowledge Unit, Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel

[3]To whom correspondence should be addressed.
E-mail: afishman@tx.technion.ac.il

Nitrobenzene dioxygenase (NBDO) from *Comamonas* sp. is shown here to perform enantioselective oxidation of aromatic sulfides. Several *para*-substituted alkyl aryl sulfides were examined and it was found that the activity of the enzyme is dependent on the size of the substrate. Saturation mutagenesis was performed on different residues in the active site in order to improve activity and selectivity. Mutagenesis at position 258 in the α-hydroxylase subunit of NBDO improved both activity and enantioselectivity. Substitutions in position 293 improved the activity on all substrates and had diverse influence on enantioselectivity. Mutagenesis in position 207 provided two interesting variants, V207I and V207A, with opposite enantioselectivities. Furthermore, combining two favorable mutations, N258A and F293H, provided an improved variant with both higher activity (5.20 ± 0.01, 2.12 ± 0.21, 2.64 ± 0.14 and 4.01 ± 0.34 nmol min^{-1} mg protein^{-1} on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively, which is 1.7, 4.6, 7.1 and 26.7-fold compared with wild type) and improved enantioselectivity (e.g. 67% enantiomeric excess for Cl-thioanisole vs. 5% for wild type). Molecular docking and active site volume calculations were used to correlate between the structure of the substrates and the function of the enzymes. The results from this work suggest that the location of pro-chiral sulfides in the active site is coordinated by hydrophobic interactions and by steric considerations, which in turn influences the activity and enantioselectivity of NBDO.

Keywords: nitrobenzene dioxygenase/sulfoxidation/biocatalysis/saturation mutagenesis/*in silico* docking

Introduction

Utilization of enzymes as biocatalysts for production of enantiopure compounds has become an established manufacturing process in the pharmaceutical industry (Nestl *et al.*, 2011; Zheng and Xu, 2011). The greater availability of biocatalysts combined with mutagenesis and computational tools contributed to the development of new application areas (Nestl *et al.*, 2011). The advantage of biocatalysis over chemical synthesis is that enzyme-catalyzed reactions are often highly enantioselective and regioselective (Beloqui *et al.*, 2008). Naturally occurring chiral sulfoxides possess a wide range of biological activities from flavor and aroma precursors to antimicrobial properties. Chiral sulfoxides are also effective chiral auxiliaries that facilitate asymmetric transformations (Fernandez and Khiar, 2003; Ronald, 2005). Furthermore, one of the most significant applications of chiral sulfoxides is in the pharmaceutical industry (Bentley, 2005). The third best-selling drug in the USA, esomeprazole, that is marketed under the name Nexium® by AstraZeneca and prescribed for ulcer disorders, is a chiral sulfoxide (Bartholow, 2011). In addition, armodafinil, that is marketed under the name Provigil® by Cephalon and is prescribed for the treatment of excessive daytime sleepiness, contains a chiral sulfoxide center (Bartholow, 2011). Chiral sulfoxides can be synthesized *via* biological or chemical means; however, enantioselective oxidation of prochiral sulfides is considered the most efficient one (de Gonzalo *et al.*, 2006).

Several biocatalysts were reported to oxidize prochiral sulfides selectively, among them naphthalene dioxygenase (NDO) from *Pseudomonas* sp. NCIB 9816-4 and toluene dioxygenase (TDO) from *Pseudomonas putida* F1 and from *Pseudomonas putida* UV4 (Lee *et al.*, 1995; Boyd *et al.*, 1998; Kerridge *et al.*, 1999). Purified, non-purified and cloned forms of these enzymes exhibited moderate to high yields with high enantioselectivity (Lee *et al.*, 1995; Boyd *et al.*, 1998; Kerridge *et al.*, 1999). In addition, toluene monooxygenases were evaluated in chiral sulfoxide synthesis. It was demonstrated that alteration in residues that are close to the active site can improve activity and selectivity on various pro-chiral sulfides (Feingersch *et al.*, 2008). A third family of biocatalysts used for obtaining chiral sulfoxides are chloroperoxidase from *Caldariomyces fumago* and horseradish peroxidase. These isolated peroxidases were able to oxidize a wide spectrum of prochiral sulfides and epoxides (Colonna *et al.*, 1999; Vargas *et al.*, 1999; Fernandez and Khiar, 2003). Nevertheless, the usage of H_2O_2 as an oxidizing agent is a drawback of these reactions. Furthermore, flavin-containing Baeyer–Villiger monooxygenases have been implemented as biocatalysts for the production of valuable drugs such as omeprazole which comprises a chiral sulfoxide (Bong *et al.*, 2010; Bornscheuer *et al.*, 2012). Finally, various yeasts and fungi are also able to produce chiral sulfoxides to different extents of selectivity (Holland, 2001; Ricci *et al.*, 2005).

Nitrobenzene dioxygenase (NBDO) from *Comamonas* sp. strain JS765, which was isolated from an industrial waste treatment plant in New Jersey (Nishino and Spain, 1995), is a member of the Rieske non-heme dioxygenase family, which includes, among others, NDO and 2-nitrotoluene dioxygenase (2NTDO), and has a hetero-hexameric structure (Friemann *et al.*, 2005). Genes encoding NBDO are clustered in a catabolic operon (*nbzAaAbAcAd*) and include: an iron–sulfur flavoprotein reductase (*nbzAa*), a Rieske [2Fe-2S]

ferredoxin (*nbzAb*) and a Rieske [2Fe-2S] non-heme iron dioxygenase (composed of α-oxygenase (*nbzAc*) and β-oxygenase (*nbzAd*) heterodimer) (Lessner *et al.*, 2002). Electrons from NADH are extracted by the reductase and then transferred to the ferredoxin. The ferredoxin transfers the electrons to the Rieske cluster in the α-oxygenase. The α-oxygenase binds molecular oxygen and a substrate in the active site. Electrons are transferred from the oxygenase Rieske cluster to the mononuclear iron, and an oxidized product is formed. The role of the β-oxygenase is suggested to be structural only (Friemann *et al.*, 2005). The crystal structure of NBDO was solved by Parales and co-workers including a structure with bound nitrobenzene (NB) (Friemann *et al.*, 2005; Parales *et al.*, 2005). NBDO has a relatively relaxed substrate preference and is able to oxidize a wide range of substrates including nitrotoluene, nitrochlorobenzene, nitrochlorotoluene, dinitrobenzene, naphthalene and dinitrotoluene (Lessner *et al.*, 2002; Ju and Parales, 2006, 2009). NBDO has been investigated for its hydroxylation capability using site-specific mutagenesis (Ju and Parales, 2006, 2009).

To date, there is very limited knowledge on the enantioselectivity of NBDO and there are no reports on its utilization for the synthesis of chiral sulfoxides. It was our objective to investigate the ability of NBDO to selectively oxidize alkyl aryl sulfides and to improve the activity and selectivity using saturation mutagenesis. An attempt to correlate structure and function of the obtained variants and examined substrates was performed using molecular docking.

Materials and methods

Chemicals

Thioanisole (99%), methyl *p*-tolyl sulfide (*p*tolyl) (99%), 4-bromophenyl methyl sulfide (Br-thioanisole) (97%), 4-chlorophenyl methyl sulfide (Cl-thioanisole) (98%), methyl phenyl sulfoxide (97%), (*R*)- and (*S*)-methyl *p*-tolyl sulfoxide (98%), sodium *meta*-periodate, nitrobenzene (99%) (NB), sulfanilamide (99%), *N*-(1-naphthyl)ethylenediamine (98%) and *tert*-butyl benzene (TBZ) were purchased from Sigma-Aldrich (Rehovot, Israel). 1-Bromo-4-(methyl sulfinyl)benzene (*p*-bromophenyl methyl sulfoxide) was purchased from Combi Blocks Inc. (San Diego, CA, USA), *p*-chlorophenyl methyl sulfoxide was purchased from Absolute Standards Inc. (Hamden, CT, USA), ethyl acetate was obtained from Gadot (Haifa, Israel). Ethanol absolute high-performance liquid chromatography (HPLC) grade was purchased from Biolab (Jerusalem, Israel), *n*-hexane, isopropyl alcohol (IPA) and sodium nitrite from Merck (Darmstadt, Germany). Phosphate buffer (PB) was prepared using KH_2PO_4 and K_2HPO_4 by Carlo Erba (Barcelona, Spain). Hydrochloric acid was purchased from J.T. Baker (Mallinckrodt Baker, Inc., Phillipsburg, NJ, USA). All materials used were of the highest purity available and were used without further purification.

Bacterial strains and plasmids

The *Nbz* operon encoding the NBDO (GenBank: AF379638.1), cloned in Bluescript II KS plasmid (Stratagene, La Jolla, CA, USA) and expressed in *Escherichia coli* TG1 (*supE hsd*Δ5 *thi* Δ(*lac-proAB*) F′

[*traD*36 *proAB*+ *lacI*q *lacZ*ΔM15]) (Keenan and Wood, 2006) was kindly provided by Prof. Thomas Wood from Pennsylvania State University.

The four-gene operon was re-cloned into pET Duet 1 plasmid (Novagen®, Merck KgaA, Darmstadt, Germany), and expressed in *E.coli* BL21 (DE3) (F⁻ *ompT gal dcm lon hsdS_B*(r_B^- m_B^-) λ *(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]*) (Novagen®, Merck KgaA). The pET Duet 1 contains two multiple cloning sites, each of which is preceded by a T7 promoter and a ribosome binding site. The Duet-Nbz vector was designed to have the α-oxygenase and β-oxygenase under one promoter and the reductase and ferredoxin under the second promoter (Supplementary Fig. S1). The *Nbz* operon was amplified in three segments using expand high fidelity DNA polymerase (Roche Diagnostics Corporation, Indianapolis, IN, USA) according to a standard amplification protocol supplied by the manufacturer. The first segment contained the reductase (*nbzAa*) and ferredoxin (*nbzAb*) genes (red-fer segment), the second segment contained the α-oxygenase gene (*nbzAc*) and the third segment contained the β-oxygenase gene (*nbzAd*). The three segments were cloned into the pET Duet 1 vector in three sequential steps. In each step a single segment was cloned into the vector, transformed into *E.coli* DH5α (*fhuA2* Δ(*argF-lacZ*)U169 *phoA glnV44* Φ80 Δ(*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) (Novagen®) competent cells *via* electroporation using MicroPulser electroporator (Biorad, Hercules, CA, USA) and purified using Qiagen plasmid mini kit (Qiagen, Valencia, CA, USA). The purified plasmid was used as a vector for the cloning of the next segment. Before cloning of the red-fer segment, it was modified using two primers (Reduct GCG Bc3: 5′GCGCTCC GGCGATGGTCGAGG3′ and Reduct GCG Fo3: 5′CCTCG ACCATCGCCGGAGCGC3′) that changed a C nucleotide into a G nucleotide within the reductase sequence in order to abolish an NcoI restriction site within the reductase gene. Each segment was cloned into the pET-Duet 1 vector using NcoI and EcoRI, NotI and EcoRI and NdeI and XhoI restriction enzymes (New England Biolabs, Ipswich, MA, USA) for α-oxygenase, β-oxygenase and reductase + ferredoxin segments, respectively, and using T4 DNA ligase (Takara, Clontech Laboratories Inc., Mountain View, CA, USA).

In order to simplify the mutagenesis process of the α-oxygenase, it was cloned separately into pET28a plasmid using the same primers as in pET-Duet 1 (DuetNcoBc and Duet EcorFo, Supplementary Table S1) and using the same restriction enzymes (NcoI and EcoRI; New England Biolabs). For protein amplification and biotransformations pET-Duet 1 was transformed into *E.coli* BL21 (DE3) (NovaGen) competent cells *via* electroporation. Each transformation step was verified by DNA sequencing using the dideoxy chain termination technique (Hylabs Ltd, Rehovot, Israel).

Media and growth conditions

Bacterial glycerol stock (−80°C) was streaked on Luria–Bertani (LB) plate solidified with 1.5% Bacto agar (Acumedia, Neogen Corporation, Lansing, MI, USA) (Sambrook *et al.*, 1989) supplemented with ampicillin at 100 µg ml⁻¹ (LB-amp100) or with kanamycin at 25 µg ml⁻¹ (LB-kan25), to maintain the plasmids. A single colony was used to inoculate fresh LB-amp100 or LB-kan25 media, the

culture was grown overnight (o/n) at 37°C with shaking at 250 rpm on a TU-400 incubator shaker (Orbital shaker incubator, MRC, Holon, Israel) and used as a starter to inoculate fresh LB-kan25 media the following day at 1 : 50 starter : medium ratio. The culture was grown at 37°C to $OD_{600} \approx$ 0.4–0.6, upon which 0.5 mM (using pET-Duet 1) or 1 mM (using pET28a) of isopropyl β-D-1-thiogalactopyranoside (IPTG) were added. The culture was further grown at 30°C until it reached $OD_{600} = 1.3$ and was then harvested.

Saturation mutagenesis (NNK libraries)

Gene libraries encoding all possible amino acids at positions 207, 222, 258, 293 and 295 of the α-oxygenase *nbzAc* in pET28a were constructed by replacing the target codon with NNK (N stands for A, T, C or G and K stands for G or T) *via* overlap extension polymerase chain reaction (PCR). For each alteration two inner primers containing NNK in the desired position were designed (Supplementary Table S1). In addition, two external primers, containing restriction sites for NcoI and EcoRI were used (Supplementary Table S1). Expand high fidelity DNA polymerase (Roche) was used for the amplifications. At the first step one inner primer containing NNK and one external primer were used (for example: Nbz-V207-fo + Duet EcorFo and Nbz-V207-rev + DuetNcoBc), which provided two short degenerate segments. Degenerate segments were incubated for 1 h with Dpn1 (New England Biolabs) in order to eliminate the plasmid template, followed by purification with QIAquick PCR purification kit (Qiagen). The amplified segments were used as templates for the next assembly PCR. The segments were combined in 1 : 1 molar ratio and amplified using DuetNcoBc and Duet EcorFo primers to obtain a full degenerate segment (1300 bp). After additional purification and digestion with NcoI and EcoRI the segment was cloned into pET Duet 1 plasmid that contained three genes, *nbzAaAbAd,* using T4 DNA ligase (Takara).The resulting plasmid was transformed into *E.coli* BL21(DE3) (Novagen®) competent cells *via* electroporation.

Site-directed mutagenesis for obtaining double-mutant N258A/F293H

Site-directed mutagenesis for obtaining variant N258A/F293H was performed *via* overlap extension PCR, as described above. One pair of primers was designed to obtain mutation F293H (Supplementary Table S1) while N258A served as a template.

Screening of libraries

Activity on nitrobenzene. (A) Screening for mutants that retained the native activity on nitrobenzene (NB) was performed by modification of a screening method previously developed by Fishman *et al.* (2004). At neutral pH, the catechol derivatives formed from NB auto-oxidize to quinones and semiquinones that readily polymerize to form a brown color. The formation of brown color was accelerated by addition of sodium periodate ($NaIO_4$) and measured at 405 nm. The assay was started by picking colonies from glycerol stocks stored in 96-well plates using a library copier VP 381 (V&P scientific, Inc., San Diego, CA, USA) and transferring into 96-deep-well plates with a plastic lid (ABgene, Thermo Fisher Scientific, Epson, UK) containing 1.2 ml of LB-Amp100. The cells were grown as described in the 'Media and

and growth conditions' section until they reached $OD_{600} = 1.3$. The culture was centrifuged at $3000 \times g$ for 10 min, at 25°C using 11144/13145 rotor, of Sigma-4K15 centrifuge (Sigma, Osterode, Germany). The cell pellets were re-suspended in 200 μl of potassium PB 100 mM, pH 7, containing 1 mM NB (prepared in ethanol). The deep well plates were shaken at 30°C for 2 h in a TU-400 incubator shaker and spun-down ($3000 \times g$ for 10 min). Hundred microliters of the resulting supernatant were transferred into 96-well plates (Nunc, Rochester, NY, USA) that contained in each well 100 μl of 2 mM sodium periodate (prepared in 100 mM potassium PB pH 7). After 10 min, color formation was measured at 405 nm using a multi-plate reader (Molecular Devices, Sunny Vale, CA, USA). All liquid handling steps were conducted using an *epMotion 5070* robotic system (Eppendorf AG, Hamburg, Germany). In order to ensure the probability of 95% that all 32 possible outcomes from the single site random mutagenesis have been sampled, 94 colonies needed to be screened (Reetz *et al.*, 2008). In practice, at least 190 colonies were screened for each library.

(B) In order to compare the native activity on NB between wild type (WT) and variants, a colorimetric nitrite release assay was used (Ju and Parales, 2006). The growth of the variants was performed in 96-deep-well plates as described in section (A). After resuspension of the cell pellet in PB containing 1 mM NB, the plates were shaken at 30°C for 30 min in a TU-400 incubator shaker. The reaction was stopped by centrifugation ($3000 \times g$ for 10 min). Fifty microliters of the resulting supernatant were transferred into new 96-well plates and mixed with 50 μl of 1% (w/v) sulfanilamide in 1.5 M HCl. After 3 min, 50 μl of 0.02% (w/v) *N*-(1-naphthyl)ethylenediamine in 1.5 M HCl was added. After an additional 3 min incubation, formation of a pink colored azo dye complex was quantified with a multi-plate reader by measuring the absorbance at 550 nm in comparison with a sodium nitrite standard curve.

Activity on four pro-chiral sulfides. Variants that retained native activity on NB were further screened for improved activity on thioanisole, *p*tolyl, Br-thioanisole and Cl-thioanisole (Fig. 1). The cultures were grown overnight in 96-deep-well plates as described in section (i) and were used as starter cultures for inoculation into test tubes filled with 5 ml LB-Amp100 (one test tube for each variant and each substrate). The cells were grown according to the 'Media and growth conditions' section and were harvested at $7000 \times g$, for 5 min, at 25°C using 12256 rotor of Sigma-4K15 centrifuge (Sigma, Osterode, Germany). The cell pellets were re-suspended in 1.2 ml PB pH 7, 100 mM to a final OD_{600} of ~4. Each one of the four substrates was added to 1 ml of cell suspension to a final concentration of 1 mM (from 100 mM
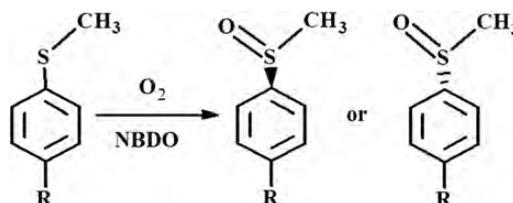

Fig. 1. The model reaction used in this research. $R=CH_3$ for methyl *p*-tolyl sulfide, $R=Br$ for 4-bromophenyl methyl sulfide, $R=Cl$ for 4-chlorophenyl methyl sulfide and $R=H$ for thioanisole.

stock in ethanol). The biotransformation was carried out for 3 h at 30°C in 16 ml glass vials for thioanisole and *p*tolyl, and in 22 ml glass vials equipped with teflon-sealed caps for Br-thioanisole and Cl-thioanisole (in order to prevent substrate evaporation during the biotransformation period) using a TU-400 incubator shaker. The reaction was stopped by addition of ethyl acetate (1 : 1 volume) followed by vigorous vortex. Phase separation was facilitated by centrifugation (13 $400 \times g$ for 2 min) using a Minispin centrifuge (Eppendorph, Hamburg, Germany). The organic supernatant analyzed by a gas chromatograph equipped with mass spectra detector (GC/MS) for activity determination and by HPLC for enantiomeric excess (%ee) determination. The criteria for better performance were higher amount of product and/or higher %ee values compared with WT.

Whole-cell biotransformations

The growth was performed as described in the 'Media and growth conditions' section. Shortly, exponential phase cultures were used in all experiments by diluting overnight culture (1 : 50 volume), growing them until reaching $OD_{600} = 0.4-0.6$, adding IPTG and growing until $OD_{600} \sim$ 1.3. The exponentially grown *E.coli* BL21 (DE3) cells harboring pET-Duet 1 vector that included NBDO or its variants were centrifuged at $7000 \times g$ for 10 min at 25°C in a Sigma-4K15 centrifuge using the 12256 rotor, and resuspended in PB to an OD_{600} of 11. The biotransformation was carried out at 30°C in a TU-400 incubator shaker with 2 ml cell suspension and 1 mM substrate in 16 ml glass vials for thioanisole and *p*tolyl and in 22 ml glass vials equipped with teflon-sealed caps for Br- and Cl-thioanisole. The reaction was stopped periodically (a vial was sacrificed) using 2 ml of ethyl acetate (1 : 1 volume) with vigorous vortex. The product and the remaining substrate were extracted into the organic phase and measured using the GC/MS and/or HPLC. The initial product formation rates were determined by sampling at 10 min intervals during the first 90 min ($\sim$20% substrate consumption). The specific activity (nmol min^{-1} mg protein^{-1}) was calculated as the ratio of the initial transformation rate and the total protein content. Enantioselectivity was determined after 3 h of reaction using HPLC. *Escherichia coli* BL21 (DE3) cells with and without pET-Duet 1 plasmid (that did not contain NBZ genes) were used as negative controls for the biotransformations and did not provide any products.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Cell pellets from bacterial suspensions at $OD_{600} = 1.6$ were resuspended in Tris–EDTA (TE) buffer. Protein concentration was calculated using the Bradford method in comparison with a calibration curve of bovine serum albumin (Bradford, 1976). For each variant a total amount of 30 μg protein was used. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer concentrated 4-fold was added. The sample was boiled for 10 min, and separated by SDS-PAGE (12% polyacrylamide) (Laemmli, 1970; Ausubel *et al.*, 1993). The electrophoresis was performed using mini-gel device (Bio-Rad, Richmond, CA, USA) at 150 V current voltage. Molecular weight of the proteins was estimated according to a commercial marker, BLUeye prestained protein ladder, GeneDireX (Hylabs Ltd,).

Molecular docking

AutoDock Vina was used for docking experiments (Scripps Research Institute, La Jolla, CA, USA) (Trott and Olson, 2010). At first, a comparison of three known crystal structures of NBDO was carried out (PDB codes: 2BMO (native structure without substrate); 2BMQ (NBDO with NB); 2BMR (NBDO with 3-nitro toluene). It was noticed that substrate binding causes slight conformational changes in the active site, in which the Fe and its coordination layer (His206, His211 and Asp360) are slightly displaced (Friemann *et al.*, 2005). Therefore, it was decided to use PDB 2BMQ, since the difference between 2BMQ and 2BMR conformations was not significant, and NB is more similar to aryl sulfides since it does not have a substituent at the *meta* position. In AutoDock Vina, all four pro-chiral sulfides were treated as flexible ligands by modifying their rotatable torsion, while NBDO (PDB code 2BMQ; Friemann *et al.*, 2005) was considered as a rigid receptor. All variants of NBDO were prepared using Pymol (DeLano, 2002) and verified by UCSF Chimera (Resource for Biocomputing, Visualization, and Informatics, University of California, CA, USA). Energy minimization of the obtained structure was done with assistance of UCSF Chimera. The Fe atom and three water molecules located at a distance of <4 Å from the metal are considered part of the receptor (H_2O numbers in 2BMQ: 2334, 2401 and 2227). The receptor was prepared using AutoDock tools package (Scripps Research Institute). All residues were checked to guarantee that the program assigned the correct charges to each ionizable side chain. The structures of the substrates were built using Chem3D Ultra10 (CambridgeSoft Corporation, Cambridge, MA, USA) and their energy was minimized using molecular mechanics method (MM2). Following minimization, all hydrogen atoms and Gasteiger atomic charges were added in UCSF Chimera (Resource for Biocomputing, Visualization, and Informatics, University of California, CA, USA). The addition of torsions and the rest of substrates preparations were made using AutoDock tools package (Scripps Research Institute). To define the grid size and location, AutoGrid function from Autodock was implemented. A grid box with sufficient space to cover the active site of NBDO (Lawrence P, 2002) was centered at the location of the C6 atom in NB, with dimensions of $10 \times 12 \times 12$ Å and a spacing of 1 Å in each dimension. After examination of a wide range of exhaustiveness values between 8 and 500, the value was set at 200. Beyond this value the approximation was not improved; however, the run time significantly increased. Exhaustiveness indicates the number of parallel runs performed by AutoDock Vina. Each run can produce several results with promising intermediate results being stored and remembered. All intermediate results are finally merged, clustered and sorted automatically to produce the final result. Increasing the exhaustiveness increases the run time linearly, and decreases the probability of not finding the global minimum exponentially (Trott and Olson, 2010). Thus, exhaustiveness of 200 reflects 200 parallel runs on the ligand molecule.

Under these parameter settings, NB was re-docked into 2BMQ providing a model with a root mean standard deviation (RMSD) of <1.12 Å (RMSD of <2 Å is considered successful; Trott and Olson, 2010). These docking parameters for NB were used for all docking experiments

involving NBDO and its variants on all four sulfides and on NB. Each run provided ~9 possible substrate conformations. The selected conformation had to fulfill two conditions: lowest affinity energy (kcal mol^{-1}) and to comply with a distance of <5.5 Å between the sulfur in the sulfides or the C6-C1 atoms in NB and the Fe atom. This distance was chosen after analyzing seven different crystal structures of dioxygenases co-crystalized with their substrates (PDB codes: 1ULJ, 2BMR, 2HMK, 2HMM, 2HMO, 3GZX, 2BMQ) (Moore *et al.*, 2010). The distance between the oxygenation site and the catalytic metal in those structures ranged from 4 to 5.5 Å.

More than 10 individual runs were made for each substrate and each variant in order to receive reliable and comparable results.

Analytical methods

The identification of the products and the conversion rates was determined by GC/MS using a GC 6890N (Agilent Technologies, Santa Clara, CA, USA) instrument equipped with a capillary column (30 m × 0.32 mm × 0.25 μm), filled with HP-5 (5%-phenyl-methylsiloxane) (Agilent Technologies), and an HP-5975 mass spectra detector (Agilent Technologies). As the HP-5 capillary column is not chiral, the (R) and (S) sulfoxides were seen as one peak. TBZ was used as an internal standard. The temperature was programmed as follows: $T_1 = 90°C$; $dT/dt = 10°C$ min^{-1}, $T_2 = 190°C$; 13 min, split ratio 1 : 10. Under these conditions, the retention times were: $R_T = 3.12$ min for TBZ, $R_T = 3.96$ min for thioanisole, $R_T = 6.73$ min for methyl phenyl sulfoxide; $R_T = 7.21$ min for Br-thioanisole, $R_T = 10.03$ min for *p*-bromophenyl methyl sulfoxide; $R_T = 6.03$ min for Cl-thioanisole and $R_T = 8.85$ min for *p*-chlorophenyl methyl sulfoxide; $R_T = 5.06$ min for me-*p*-tolyl sulfide and $R_T = 8.18$ min for me-*p*-tolyl sulfoxide. The identity of the substances was determined by comparison with commercial standards and with a NIST library and the amount of the substances was quantified according to calibration curves.

Following the extraction of the substrates and products with ethyl acetate, the solvent was evaporated with nitrogen and then resuspended (1 : 1 volume) in HPLC-grade ethanol due to the sensitivity of the chiral column to organic solvents. The %ee of the products was determined using HPLC 1100-series instrument (Agilent Technologies) equipped with LUX 5 μM Cellulose-2 (Phenomenex, Torrance, CA, USA) chiral column (250 × 4.6 mm). Detection was carried out with a UV-VIS detector at 225 nm. The temperature was ambient. The mobile phase was hexane/IPA 80 : 20 and the flow was 2.0 ml min^{-1} for *p*-bromo- and *p*-chlorophenyl methyl sulfoxides or 1.0 ml min^{-1} for methyl phenyl sulfoxide. The mobile phase for me-*p*-tolyl sulfoxide separation was hexane/EtOH 85 : 15 and the flow was 2.0 ml min^{-1}. Under these conditions, the retention times were: $R_T = 16.97$ min for (R)-methyl phenyl sulfoxide and $R_T = 18.55$ min for (S)-methyl phenyl sulfoxide; $R_T = 6.93$ min for (R)-*p*-bromophenyl methyl sulfoxide and $R_T = 7.91$ min for (S)-*p*-bromophenyl methyl sulfoxide; $R_T = 6.39$ min for (R)-*p*-chlorophenyl methyl sulfoxide and $R_T = 7.38$ min for (S)-*p*-chlorophenyl methyl sulfoxide, and $R_T = 8.61$ min for (R)-methyl *p*-tolyl sulfoxide and $R_T = 7.92$ min for (S)-methyl *p*-tolyl sulfoxide. The order of R and S sulfoxides was determined using purified R and S enantiomers (for me-*p*-tolyl sulfoxide) or by injecting the obtained products to a 30 m × 0.32 mm × 0.25 μm GC capillary column packed with γ-cyclodextrin tri-fluoroacetyl (CHIRALDEX^{TM} G-TA, ASTEC, Bellefonte, PA, USA) and comparing the retention order with published data by Anderson *et al.* (2002).

Results

Construction of pET-Duet-NBZ plasmid

NBDO is a large multicomponent enzyme complex that is encoded by four genes (*nbzAaAbAcAd*) (Parales *et al.*, 2005). When all the genes are clustered under one operon it may lead to low expression levels of the proteins that are translated from the genes that are located farther away from the promoter and from the ribosome binding site (Tolia and Joshua-Tor, 2006). pET-Duet 1 that contains two multiple cloning sites, each of which is preceded by a T7 promoter and a ribosome binding site, should be more suitable for multi-gene systems. The Duet-Nbz plasmid was designed to have the α-oxygenase and β-oxygenase under one promoter and the reductase and ferredoxin under the second promoter. This provided two shorter clusters containing ~2000 bp each, while the larger genes (α-oxygenase and reductase) are located nearby the promoter (Supplementary Fig. S1). All gene manipulations were confirmed by sequencing, and expression levels were verified by SDS-PAGE and analyzed using ImageJ software (ImageJ, http://rsb.info.nih.gov/ij/index.html), indicating better expression (4.7-fold) with the Duet system compared with Bluescript II KS. Additional proof for better expression using pET-Duet1 was obtained by the determination of native activity on 1 mM NB. The initial rate measured was 8.55 ± 0.14 nmol min^{-1} mg protein^{-1} compared with 0.392 ± 0.005 nmol min^{-1} mg protein^{-1} in pBS(kan), which is 22-fold higher.

Saturation mutagenesis (NNK libraries) and screening for improved variants

Due to the large plasmid size (~10 kb) it was impossible to use a direct mutagenesis method such as QuickChange (Stratagene) for saturation mutagenesis. Therefore in order to simplify the DNA manipulations, the α-oxygenase gene was cloned separately to pET28a. After performing saturation mutagenesis *via* overlap extension PCR, the mutated gene was cut and ligated into the Duet plasmid harboring the β-oxygenase, ferredoxin and reductase genes to form a complete and active operon. Positions for mutagenesis, V207, F222, N258, F293 and N295, were chosen based on different considerations. These include prior work on NBDO with other substrates (N258, F293) (Friemann *et al.*, 2005; Ju and Parales, 2006), alignment with similar enzymes such as aniline dioxygenase (V207) (Ang *et al.*, 2009), or the HotSpot Wizard developed by Damborsky and co-workers (F222, N295) (Pavelka *et al.*, 2009). This program gives predictions to 'hot spots' in enzymes based on structural, functional and evolutionary information obtained from different databases. In the output, HotSpot Wizard lists annotated residues ordered by estimated mutability. F222 received mutability rank of 9 (high), while N295 received a value of 3 (low). However, both positions are close to the active site and were not previously investigated. Therefore they were chosen for mutagenesis and for further analysis.

Saturation mutagenesis using degenerate primers NNK requires that 94 variants should be screened to ensure the coverage of all possible codon combinations with a probability of 95% (Reetz et al., 2008). In practice, ~200 colonies (two 96-well plates) were screened for each randomized position according to a two-step approach. Initially, libraries were screened on NB, and those variants which retained native activity were further screened on thioanisole, ptolyl, Br-thioanisole and Cl-thioanisole. Active variants on NB showed formation of brown color with intensity similar to WT, indicative of catechol formation (Fishman et al., 2004). Variants that did not show color formation were considered inactive and were not further investigated, under the assumption that if the native activity is lost, most probably a new activity would not be obtained (Soskine and Tawfik, 2010). The percentage of active variants ranged from 25 to 56% in every plate. The second step was evaluation of the active variants on all four pro-chiral sulfides (Fig. 1).

Although a few rapid screening assays for determination of sulfoxide formation were developed (Shainsky et al., 2008), each of them detects indirectly only one substrate, thus if we decided to use one of the assays, the determined variants would be improved only on the examined substrate (Feingersch et al., 2008). Therefore, to broaden the scope of the work, it was decided to perform one point measurements on all four substrates using GC/MS and HPLC. The screening results are graphically summarized in Fig. 2. Each point represents one experiment on one substrate and two separate analyses (activity and %ee). Gray boxes show the threshold applied for choosing variants for further work. Improved variants were those showing higher enantioselectivity than WT on one of the substrates and/or at least 1.5-fold activity. Based on these criteria, 19 variants were selected: N258A, N258F, N258A, F222M, F222V, F222G, F222P, N295G, N295D, N295A, N295H, N295L, N295S, V207I, V207A, F293H, F293Q, F293I, F293T.
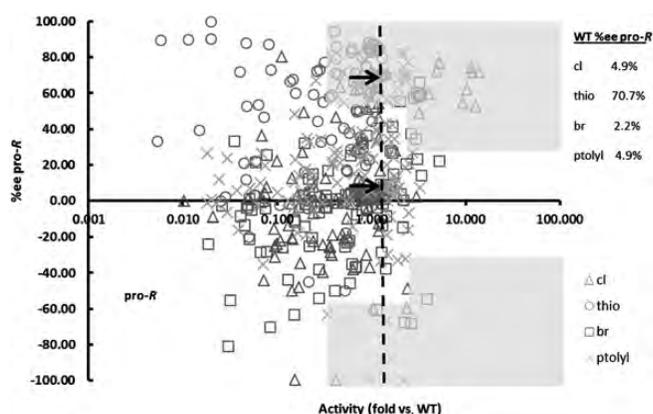

Fig. 2. Schematic representation of all screening results of NBDO variants (libraries 207, 222, 258, 293, 295) on four pro-chiral sulfides. X-axis is presented in logarithmic scale and shows the activity of each variant compared with WT that was fixed as 1. Enantioselectivity of WT on all four substrates is described in the upper right corner of the graph and marked with black arrows on the plot. Y-axis shows the enantioselectivity of each variant. Every point on the plot is a result of an independent experiment measuring activity (GC/MS) and selectivity (HPLC) after 3 h of biotransformation. Gray boxes denote the threshold for choosing the variants for further investigation (better activity and/or selectivity compared with WT).

Whole-cell biotransformations

Initially, WT's activity and selectivity on thioanisole, ptolyl, Cl-thioanisole and Br-thioanisole were evaluated (Fig. 3). The results clearly indicate that the highest activity and selectivity were on thioanisole, 3.11 ± 0.01 nmol min^{-1} mg protein^{-1} and $70.7 \pm 0.2\%$ee (proR) compared with 0.46 ± 0.02 nmol min^{-1} mg protein^{-1} and $4.9 \pm 0.2\%$, 0.37 ± 0.02 nmol min^{-1} mg protein^{-1} and $4.9 \pm 0.2\%$ and 0.150 ± 0.003 nmol min^{-1} mg protein^{-1} and $2.21 \pm 0.03\%$, on ptolyl, Cl-thioanisole and Br-thioanisole, respectively. Subsequently, all of the 19 variants selected following the first stage were evaluated and compared with WT, and the best performing ones are presented in Fig. 3. Generally, the activity on all four substrates was according to the order thioanisole > ptolyl ≥ Cl-thioanisole > Br-thioanisole.

All of the variants from library N258 improved dramatically the enantioselectivity to a maximum of 87, 77, 73 and 64% ee (proR) on thioanisole, ptolyl, Cl-thioanisole and Br-thioanisole compared with 70.7, 4.9, 4.9 and 2.2% in NBDO WT, respectively, (Fig. 3B). However, the activity was improved only by N258A: 4.9 ± 0.3, 0.89 ± 0.09, 1.03 ± 0.04 and 0.53 ± 0.01 nmol min^{-1} mg protein^{-1} on thioanisole, ptolyl, Cl-thioanisole and Br-thioanisole, respectively, which is 1.6-, 2.0-, 2.8- and 3.5-fold compared with WT, respectively (Fig. 3A). Regarding residue F293, the most active variant on thioanisole was F293T with an
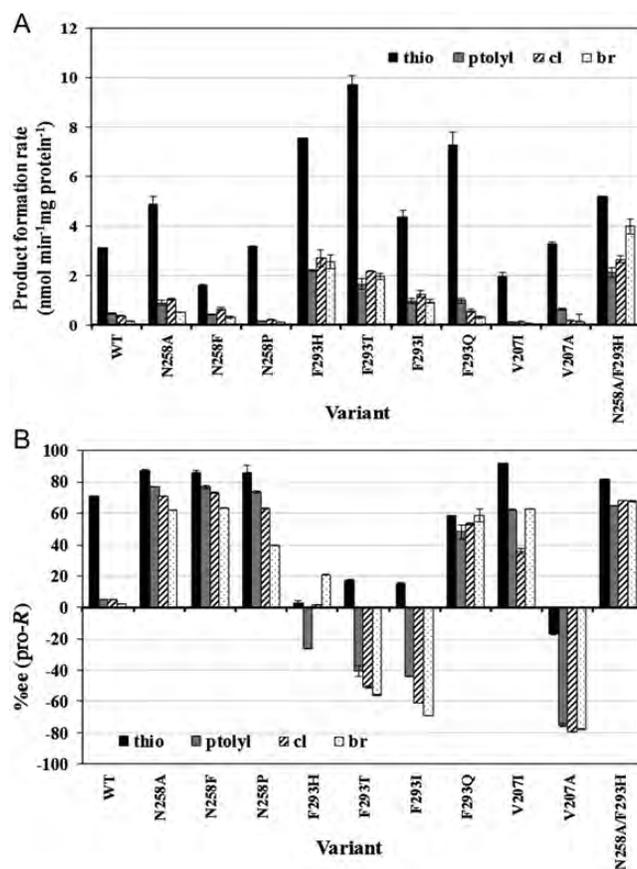

Fig. 3. Product formation rate (**A**) and % enantioselectivity (proR) (**B**) of E.coli BL21(DE3) cells expressing WT NBDO and NbzAc mutants. Initial substrate concentration was 1 mM. The results represent an average of at least three independent experiments. Activity measurements were obtained by GC/MS analyses and enantioselectivity by HPLC using a chiral column.

activity 3-fold higher than WT. However, overall, the best variant from this library was F293H that caused a marked increase in activity on all four substrates, 2.4-, 4.8-, 7.4- and 17-fold improvement on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively. In contrast, the influence of F293H on enantioselectivity was in general negative: the selectivity on thioanisole was diminished, the selectivity on Br-thioanisole improved a little and on *p*tolyl was slightly reversed. It is interesting to note that while variants F293T and F293I reversed the selectivity of three substrates (*p*tolyl, Cl-thioanisole and Br-thioanisole) to favor the (*S*)-sulfoxide to a pronounced degree, mutant F293Q improved the pro-*R* selectivity on the same substrates. Thus, position 293 is shown to influence the orientation of larger substrates in the active site. Similar results were obtained by two variants from library V207—a decrease in activity accompanied by a substantial change in enantioselectivity. Variant V207I improved the original pro-*R* enantioselectivity while V207A reversed it on all four substrates: 92.46 ± 1.28 vs. $-16.68 \pm 0.67\%$, 62.32 ± 0.32 vs. $-75.25 \pm 1.14\%$, 35.92 ± 1.99 vs. $-79.22 \pm 0.09\%$ and 62.73 ± 0.19 vs. $77.70 \pm 0.02\%$ ee (proR) on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively.

The best performing variants in library N295 were N295D on thioanisole (1.01 ± 0.3 nmol min^{-1} mg protein^{-1} and $83.9 \pm 0.1\%$ ee (proR)) and N295G on *p*tolyl, Cl-thioanisole and Br-thioanisole exhibiting 0.52 ± 0.02 nmol min^{-1} mg protein^{-1} and $64.6 \pm 0.2\%$, 1.11 ± 0.09 nmol min^{-1} mg protein^{-1} and $48.2 \pm 0.5\%$, and 0.20 ± 0.06 nmol min^{-1} mg protein^{-1} and $48.0 \pm 0.3\%$ ee pro R, respectively.

In library F222, the best variant was F222M with 0.44 ± 0.02 nmol min^{-1} mg protein^{-1} and $87.1 \pm 1.8\%$, 0.03 ± 0.001 nmol min^{-1} mg protein^{-1} and $37.1 \pm 4.5\%$ and 0.02 ± 0.001 nmol min^{-1} mg protein^{-1} and $-9.0 \pm 2.3\%$ ee pro-*R*, on thioanisole, *p*tolyl and Cl-thioanisole, respectively and negligible activity on Br-thioanisole. Due to the general low performance exhibited by variants from libraries N295 and F222, the results were not included in Fig. 3.

Next, two beneficial mutations were combined to investigate whether both activity and selectivity could be improved jointly. Using site-directed mutagenesis the F293H substitution was introduced into the plasmid of N258A. While variant N258A showed a fair activity improvement (1.6-, 2.0-, 2.8- and 3.5-fold better on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively) accompanied by a dramatic improvement in enantioselectivity, variant F293H had remarkable activity improvement (2.4-, 4.8-, 7.4-, and 17-fold compared with WT, on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively) and almost no influence on selectivity (Fig. 3). The double-mutant N258A/F293H showed improved activity of 5.20 ± 0.01, 2.12 ± 0.21, 2.64 ± 0.14 and 4.01 ± 0.34 nmol min^{-1} mg protein^{-1} on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively, which is 1.7-, 4.6-, 7.14- and 26.7-fold compared with WT, and improved enantioselectivity similar to N258A variant (82, 65, 68, and 68 ee (proR) on thioanisole, *p*tolyl, Cl-thioanisole and Br-thioanisole, respectively). The most dramatic effect was with Br-thioanisole as a substrate, 26.7-fold activity improvement, and the least significant effect was using thioanisole. However, it is clear that this combination was favorable for the overall performance.

It should be noted that in prolonged experiments (>3 h of biotransformation) a peak of sulfone was detected by the GC/MS, indicating further oxidation of the sulfoxides by the enzymes. However, within the short time limits of the experiments described in this work, oxidation of sulfoxides to sulfones was not detectable. In addition, no other oxidation by-products were observed during the entire biotransformation period.

In order to gain further understanding on the influence of each mutation, the activity on the natural substrate NB, was examined (Supplementary Fig. S2). It was discovered that the native activity of all variants, except V207I, was decreased by 2–20-fold. To verify that the increase in activity of the mutants in the sulfoxidation reactions was derived from the amino acid substitutions rather than expression level changes, SDS-PAGE was used to visualize three of the four subunits: *NbzAc* (50 kDa), *NbzAa* (35 kDa) and *NbzAd* (23 kDa) (Parales *et al.*, 2005). The gel included the pET-Duet 1 plasmid without an insert as a control. Densitometric analyses indicated that differences in mutant and WT bands were <10%. As the cell growth, the biotransformation conditions and the amount of protein loaded were identical for WT and mutants, the changes in activity appear to arise from the mutations and not from different expression levels.

Molecular docking

For investigating structure-function correlations, knowledge on the orientation of the substrate in the active site is essential. Ideally, the investigation involves comprehensive studies that include quantum mechanics and molecular mechanics methods (QM/MM) (Hu and Yang, 2009). However, such analysis was beyond the scope of our research. Therefore simple molecular docking software was implemented in order to provide visualization of the substrate conformation before the enzymatic reaction has actually occurred (Trott and Olson, 2010). Measures were taken in order to establish that the docking results may be indicative of the correct substrate conformation in the active site; however, only crystallization of the mutants can provide unambiguous evidence. Initially, NB was re-docked into the enzyme structure (PDB 2BMQ) that was maintained either rigid or partially flexible (residues Val207, Phe293, Asn258, His211, His206 and Asp360). The best model correlating well with the actual crystal structure that contains NB was achieved using a rigid receptor, providing an RMSD (root mean standard deviation) of <1.12 Å (RMSD of 2 Å is considered successful; Trott and Olson, 2010). These docking parameters were subsequently used for all docking experiments involving NBDO and its variants. Six representative models are shown in Fig. 4: (A) WT with thioanisole, presenting the smallest substrate that is the most similar to NB; (B) WT with *p*tolyl as a representative of the large substrates (*p*tolyl, Cl-thioanisole, Br-thioanisole) that have similar size (6.2, 6.4 and 6.5 Å, respectively), despite the differences in their chemical properties; (C) and (D) present Br-thioanisole docked in WT and in F293H, respectively. The latter showed a 17-fold improvement in activity (the largest activity improvement using single amino acid substitution) (2.56 ± 0.28 nmol min^{-1} mg protein^{-1} vs. 0.150 ± 0.003 nmol min^{-1} mg protein^{-1} in WT) and minor pro-*R* enantioselectivity ($20.91 \pm 0.13\%$ ee (pro-*R*) vs. $2.21 \pm 0.03\%$ in WT). In Fig. 4C and D
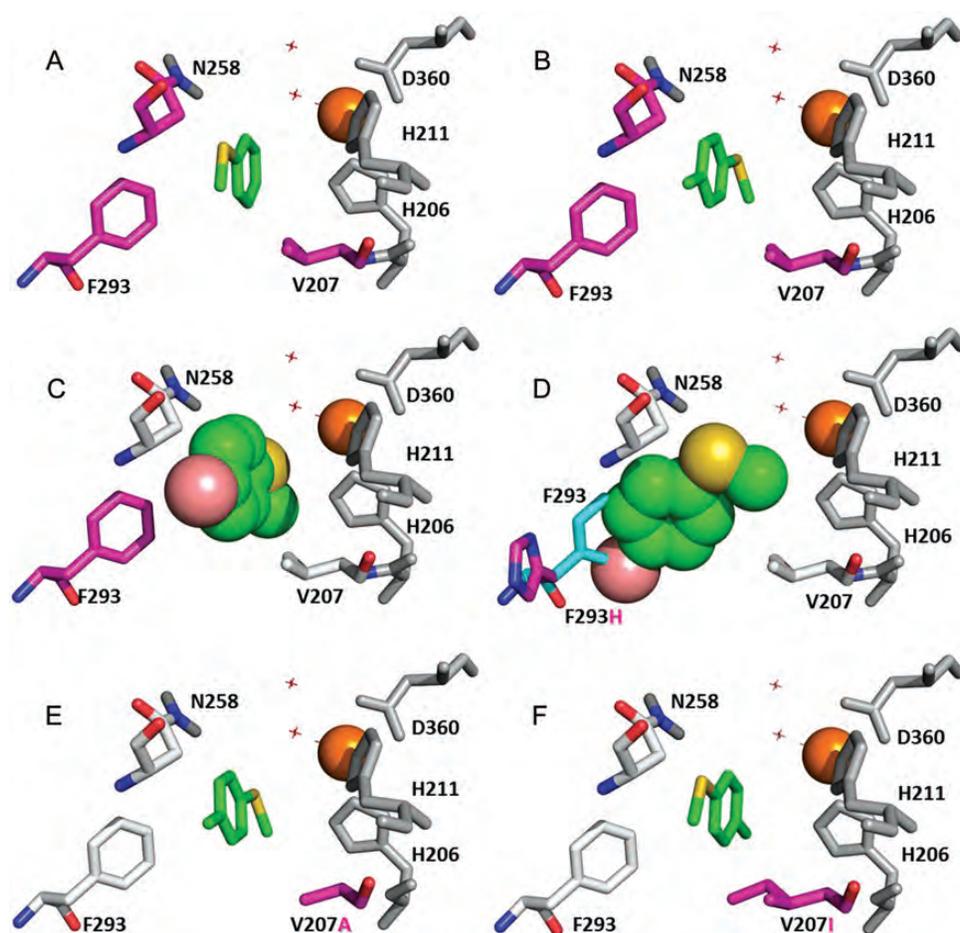
Fig. 4. Molecular models of thioanisole, Br-thioanisole and *p*tolyl in the active site predicted by AutoDock Vina. The most reasonable conformation was visualized by PyMOL. The residues that were subjected to mutagenesis (V207, F293 and N258) are presented as magenta or grey sticks. His–His–Asp triad is presented as grey sticks. Fe atom is presented as an orange sphere and water molecules as red crosses. (**A**) Thioanisole in WT active site; (**B**) *p*tolyl in WT active site; (**C**) Br-thioanisole in WT active site; (**D**) Br-thioanisole in F293H. Residue 293 in WT is colored in cyan and in the mutant in magenta. Br-thioanisole in (C) and (D) is presented as spheres in order to emphasize the volume the substrate occupies in the active site of the enzyme; (**E**) *p*tolyl in V207A; (**F**) *p*tolyl in V207I. Color coded sticks: nitrogen = blue, oxygen = red, sulfur = yellow; methyl = green.

Br-thioanisole is presented in spheres in order to emphasize the volume it occupies in the active site. Fig. 4E and F present variants V207I and V207A with *p*tolyl docking in which the enantioselectivity was reversed by changing Ile to Ala ($62.32 \pm 0.32\%$ in V207I vs. $-75.25 \pm 1.14\%$ ee (pro-*R*) in V207A). These docking models (Fig. 4) suggest that alterations in active site residues lead to different positioning of the substrate in the active site. In Fig. 4A thioanisole was located according to the docking program in the same orientation as NB, the native substrate, which may explain the highest activity of the enzyme on this substrate compared with *p*tolyl that was positioned in an opposite manner (Fig. 4B and Br- and Cl-thioanisole, which are not shown in the figure). The activity on *p*tolyl was 15% of the activity on thioanisole. In F293H (Fig. 4D) the substrate is suggested to be located in the diagonal between residue 293 and the Fe atom unlike in WT (Fig. 4C) where it was located between residues V207 and N258. The change from the bulky hydrophobic Phe to the smaller His residue (F293H) enlarged the active site, contributing to the increase in activity. However, it is assumed that the substrate is not stabilized in one specific orientation, resulting in poor enantioselectivity of this variant. The docking models propose that the orientation of *p*tolyl in V207A and V207I (Fig. 4E and F) is reversed. The drastic change in orientation of the substrate (close to 180°) in V207A vs. V207I, may explain the reversion in enantioselectivity.

Calculation of active site volume using CASTp

In order to gain additional insight into the influence of each amino acid substitution, the volume of the active site pocket of WT NBDO (PDB code 2BMQ) and its variants was measured using CASTp web server (Dundas *et al.*, 2006). The calculation is based on the alpha-shape theory and discrete-flow method (Liang *et al.*, 1998). The output includes all the pockets and cavities identified by the server (based on water radii of 1.4 Å); however, we analyzed only the active site (Supplementary Table S2). The volume of WT's active site pocket was calculated to be 322 Å^3. Most of the alterations resulted in an increase of the volume of the active site. For example, double-mutant N258A/F293H had an active site pocket larger by 37.3% compared with WT. The exceptions were V207I and N258F for which the active site volume was reduced by 8.44 and 22.08%, respectively. Interestingly, this reduction was generally correlated with the decrease in activity.

Discussion

Chiral sulfoxides are of great importance to the pharmaceutical industry and in organic synthesis as functionalized substances and as chiral auxiliaries (Fernandez and Khiar, 2003; Bentley, 2005; de Gonzalo *et al.*, 2006). The focus of the current research was the study of structure–function correlations between NBDO and prochiral sulfides. Saturation mutagenesis was applied to improve the enantioselectivity and oxidation rate of four model substrates, namely, thioanisole, *p*tolyl, Br-thioanisole and Cl-thioanisole. Furthermore, *in-silico* docking was used to analyze the experimental results.

WT NBDO exhibited relatively high activity and enantioselectivity on thioanisole, and a dramatic decrease in activity (7–20-fold) and selectivity (<5% ee pro-*R*) on the other three substrates (Fig. 3), suggesting that the size of the substrate influences the positioning in the active site and as a consequence its oxidation rate (NB and thioanisole have a length of ~4.5 Å, while *p*tolyl, Cl-thioanisole and Br-thioanisole – ~6.4 Å). Docking experiments also proposed that thioanisole is oriented similarly to NB, while *p*tolyl, Cl-thioanisole and Br-thioanisole are rotated by ~180° (RMSD >3.2 Å) (Fig. 4A and B). Since the active site of NBDO is relatively small and tight ~6 Å × 8 Å × 10 Å (Lawrence P, 2002) the larger substrates are sterically restricted and thus may not be oriented favorably for oxygen attack, resulting in reduced activity and selectivity. The calculated volumes of the substrates (thioanisole to Br-thioanisole) according to Chimera are in the range of 107–157 Å^3.

In order to improve the activity and selectivity of NBDO, saturation mutagenesis was implemented at five positions (N258, V207, F222, F293 and N295). Screening was performed on all four substrates in order to prevent bias (Fig. 1). Libraries N295 and F222 that were chosen based on HotSpot Wizard did not provide improved variants. The reason may be the relatively distant location from the active site, 10.7 and 10.1 Å, respectively. In contrast, residues V207, N258 and F293 are part of the active site and are located 5.9, 6.5 and 8.9 Å, respectively, from the mononuclear iron. V207 is a conserved residue (rank 7 in ConSurf server; Ashkenazy *et al.*, 2010) while N258 and F293 are not (ranks 1 and 3, respectively). N258 creates a hydrogen bond with the oxygen in NB which stabilizes the substrate in the active site and enables the oxidation at C1 and C6 of the aromatic ring (Friemann *et al.*, 2005). When Asn was replaced by Val (the respective residue in NDO), regioselectivity towards various nitroarenes was changed since the hydrogen bond could not be formed (Ju and Parales, 2006). For example, with 4-nitrotoluene as a substrate, WT NBDO formed 4-methyl catechol with 95% preference, while N258V produced 4-nitrobenzyl alcohol with 99% preference (Ju and Parales, 2006). In the present work, variants N258A, N258F and N258P improved the enantiomeric excess of all four substrates towards the *R*-isomer with the most active and selective variant being N258A (Fig. 3). Ala is smaller and more hydrophobic than Asn, perhaps these qualities facilitated better positioning of the substrate in the active site. Docking results suggest that thioanisole, Cl-thioanisole and *p*tolyl are situated in the same manner as in WT, which may explain the moderate increase in activity (1.5–2.5-fold),

and dramatic increase in selectivity. Br-thioanisole is suggested to be positioned in a completely different orientation (the RMSD between its location in WT and in N258A is 3.7 Å) which may explain the higher improvement in activity (3.5-fold). As expected the activity on NB for all N258 mutants was drastically reduced for all variants (3–10-fold) since the hydrogen bond to the substrate could not be formed. Furthermore, docking experiments showed that NB is possibly positioned differently in all three variants in comparison with WT (RMSD > 3.37 Å, results not shown) substantiating the experimental results.

Residue F293 is part of the hydrophobic cleft providing an appropriate environment for the binding of aromatic substrates (Friemann *et al.*, 2005). Parallel residues in 2NTDO, NDO and di-nitrotoluene dioxygenase (DNTDO) are Ile, His and Gln, respectively. These differences are believed to be the basis for substrate specificity (Friemann *et al.*, 2005). Previously, when F293 was altered to one of these residues differences in activity were substrate dependent. For example, with 3-nitrotoluene as the substrate there was almost no effect on the enzyme's regioselectivity, but with 4-nitrotoluene the selectivity was changed from 95% toward 4-methyl catechol to 70% toward 4-nitrobenzyl alcohol using F293H. Using 2-nitrotoluene, the regeoselectivity of F293H and F293Q was not affected, but the activity of F293I was completely lost (Ju and Parales, 2006). Our screening resulted in four favorable mutations: F293I, F293H, F293Q, F293T (Fig. 3). The first three alterations are similar to amino acids at parallel positions in 2NTDO, NDO and DNTDO, respectively, while the substitution to Thr is unique. Generally, all the substitutions improved the activity of the enzyme toward all four examined substrates, but the effect on enantioselectivity was diverse. F293H provided the largest improvement in activity (except on thioanisole where F293T brought the largest impact), accompanied with a marked decrease in selectivity of thioanisole. F293I and F293T had the same tendency changing the enantioselectivity on *p*tolyl, Cl-thioanisole and Br-thioanisole towards the S-enantiomer and reducing the selectivity on thioanisole (Fig. 3B). F293Q presented a selectivity of 50–60% towards the *R*-enantiomer on all substrates. Volume change calculations revealed an increase in the active site dimensions in all F293 variants at a range of 6.2–35.4% (Supplementary Table S2), which probably facilitated better positioning of the larger substrates such as Br-thioanisole. Therefore, it is reasonable that the most dramatic change with the bulkiest substrate was in F293H (17-fold improvement for Br-thioanisole). Docking results of Br-thioanisole in variant F293H suggest positioning of the substrate between the His and the mononuclear iron with the sulfur oriented towards the iron (Fig. 4D). Results of all the docking experiments of F293 variants show that in most cases (except F293T) the larger substrates (*p*tolyl, Cl-thioanisole and Br-thioanisole) are positioned in a similar orientation while the smaller substrates (NB and thioanisole) are in a different one. Thus, this implies that the improvement in activity for the large substrates is a result of the increase in active site volume by decreasing the amino acid at position 293. It is suggested that the substrate is stabilized in the new orientation by π–π interactions with three aromatic residues (F251, F222 and W356) that are located <5.5 Å from the aromatic ring of the substrate (Hunter and Sanders, 1990). As expected the

343

activity of the variants on NB was decreased (2.5–6-fold, Supplementary Fig. S2) supporting the conclusion of Ju and Parales that the reduced size of the hydrophobic residue at 293 position is not favorable for the appropriate positioning of a small substrate for ring oxidation (Ju and Parales, 2006).

Position V207, not investigated previously with NBDO, is part of the hydrophobic cavity of the active site (Friemann *et al.*, 2005) and is located 5.9 Å from the mononuclear iron. It is highly conserved among the dioxygenases. Alteration in a parallel position in aniline dioxygenase (V205), which shares 19% similarity with NBDO, enabled the enzyme to accept a new substrate, 2-isopropylaniline. The new activity came at the expense of the native activity (Ang *et al.*, 2009). Screening of the 207 NNK library did not provide improved variants however, replacing Val to Ile improved the pro-*R* enantiosectivity, while substitution to Ala changed it to pro-*S*. The substitution to Ile reduced the active site volume by 8.4%. This may explain the decreased activity, on the one hand, and the increased enantioselectivity on the other. Probably, the larger residue does not enable free access of the substrate molecules to the active site; however, it stabilizes it in one orientation which provides higher selectivity. Substitution to Ala provided opposite enantioselectivity and almost no improvement in activity. Docking experiments suggest that the substrate binds similarly in V207I and in WT which may explain the same enantioselectivity, while in V207A it has a different orientation. In addition, the orientation of *p*tolyl and Cl-thioanisole is opposite in both variants which may also hint on the opposite enantioselectivity (partial results Fig. 4E and F). The native activity on NB decreased 10-fold in V207A, while was not affected much in V207I. As NB is a small molecule, substitution to Ile did not impair its positioning in the active site, while substitution to Ala increased the active site volume by 9.9% and affected the orientation. Docking experiments also support this assumption by placing NB in V207I in the same orientation as in WT, and in an opposite manner in V207A.

A further step in the investigation of the influence of critical residues in the active site was the combination of two favorable mutations in positions 258 and 293 resulting in variant N258A/F293H. The new variant exhibited the high enantioselectivity of N258A and high activity of F293H accompanied by an increase of 37% in the active site volume. Not surprising, the most pronounced effect was in the activity on Br-thioanisole (27-fold improvement). Using molecular docking, the suggested position of the substrates is similar to F293H, while for NB there is no appropriate position that enables efficient oxidation (none of nine proposed positions was close enough to the mononuclear iron to enable oxidation). This correlates well with the nearly complete loss of activity of N258A/F293H on NB (26-fold, Supplementary Fig. S2).

Taking into account the limitations of the docking program and the fact that pro-chiral sulfides cannot create hydrogen bonds with neighboring residues, it is suggested that their location in the active site is coordinated by hydrophobic interactions and by steric considerations. In summary, without usage of complex computational design that involves QM/MM methods we can only investigate the structure of the ligand and the receptor before the enzymatic reaction has occurred. As a consequence, we conclude that the main influence on the activity and selectivity is the size of the substrate and the active site pocket volume.

Chiral sulfoxides have been previously produced by several mono- and dioxygenases. The activity of toluene *ortho*-monooxygenase (TOM) and its TomA3 V106 variants on thioanisole and methyl *p*-tolyl sulfide was lower compared with WT NBDO and its variants (Feingersch *et al.*, 2008). For example, WT TOM activity on thioanisole was 1.6 ± 0.2 nmol min^{-1} mg protein^{-1} (Feingersch *et al.*, 2008) while WT NBDO activity was 3.11 ± 0.02 nmol min^{-1} mg protein^{-1}. However, the alteration in position 106 of TomA3, and in the parallel position 100 of TmoA in toluene 4-monooxygenase (T4MO), brought to pronounced improvement in enantioselectivity toward the *S*-enantiomer. Naphthalene and toluene dioxygenases were previously evaluated toward a vast range of prochiral sulfides exhibiting a wide variety of conversion rates (2–100%) and high enantioselectivity (>90%) (Lee *et al.*, 1995; Kerridge *et al.*, 1999; Boyd *et al.*, 2004). Evolved NBDO variants need to be further improved by successive rounds of mutagenesis in order to reach potential efficiency required by industrial processes. Nevertheless, some advantages are the complete regioselectivity (no ring oxidations) and pro-*R* enantioselectivity that is competent with several chiral drugs, such as Armodafinil.

Supplementary data

Supplementary data are available at *PEDS* online.

Acknowledgements

References

Anderson,J.L., Ding,J., McCulla,R.D., Jenks,W.S. and Armstrong,D.W. (2002) *J. Chromatogr. A*, **946**, 197–208.

Ang,E., Obbard,J. and Zhao,H. (2009) *Appl. Microbiol. Biotechnol.*, **81**, 1063–1070.

Ashkenazy,H., Erez,E., Martz,E., Pupko,T. and Ben-Tal,N. (2010) *Nucleic Acids Res.*, **38**, 529–533.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1993) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. New York, NY.

Bartholow,M. (2011). *Top 200 Drugs of 2011*.

Beloqui,A., de Maria,P.D., Golyshin,P.N. and Ferrer,M. (2008) *Curr. Opin. Microbiol.*, **11**, 240–248.

Bentley,R. (2005) *Chem. Soc. Rev.*, **34**, 609–624.

Bong,Y.K., Clay,M.D., Collier,S.J., Mijts,B., Vogel,M., Zhang,X., Zhu,J., Nazor,J., Smith,D. and Song,S. (2010) *Synthesis of prazole compounds*, In Codexis,I. (ed). Codexis Inc., WO2011071982.

Bornscheuer,U.T., Huisman,G.W., Kazlauskas,R.J., Lutz,S., Moore,J.C. and Robins,K. (2012) *Nature*, **485**, 185–194.

Boyd,D.R., Sharma,N.D., Haughey,S.A., Kennedy,M.A., McMurray,B.T., Sheldrake,G.N., Allen,C.R., Dalton,H. and Sproule,K. (1998) *J. Chem. Soc., Perkin Trans. 1*, 1929–1934.

Boyd,D.R., Sharma,N.D., Byrne,B.E., Haughey,S.A., Kennedy,M.A. and Allen,C.C.R. (2004) *Org. Biomol. Chem.*, **2**, 2530–2537.

Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.

Colonna,S., Gaggero,N., Richelmi,C. and Pasta,P. (1999) *Trends Biotechnol.*, **17**, 163–168.

de Gonzalo,G., Torres Pazmino,D.E., Ottolina,G., Fraaije,M.W. and Carrea,G. (2006) *Tetrahedron: Asymmetry*, **17**, 130–135.

DeLano,W.L. (2002), http://www.pymol.org.

Dundas,J., Ouyang,Z., Tseng,J., Binkowski,A., Turpaz,Y. and Liang,J. (2006) *Nucleic Acids Res.*, **34**, 116–118.

Feingersch,R., Shainsky,J., Wood,T.K. and Fishman,A. (2008) *Appl. Environ. Microbiol.*, **74**, 1555–1566.

Fernandez,I. and Khiar,N. (2003) *Chem. Rev.*, **103**, 3651–3705.

Fishman,A., Tao,Y., Bentley,W.E. and Thomas,K.W. (2004) *Biotechnol. Bioeng.*, **87**, 779–790.

Friemann,R., Ivkovic-Jensen,M.M., Lessner,D.J., Yu,C.-L., Gibson,D.T., Parales,R.E., Eklund,H. and Ramaswamy,S. (2005) *J. Mol. Biol.*, **348**, 1139–1151.

Holland,H.L. (2001) *Nat. Prod. Rep.*, **18**, 171–181.

Hu,H. and Yang,W. (2009) *J. Mol. Struct. Theochem*, **898**, 17–30.

Hunter,C.A. and Sanders,J.K.M. (1990) *J. Am. Chem. Soc.*, **112**, 5525–5534.

ImageJ (http://rsb.info.nih.gov/ij/index.html).

Ju,K.-S. and Parales,R.E. (2006) *Appl. Environ. Microbiol.*, **72**, 1817–1824.

Ju,K.S. and Parales,R.E. (2009) *Microb. Biotechnol.*, **2**, 241–252.

Keenan,B. and Wood,T. (2006) *Appl. Microbiol. Biotechnol.*, **73**, 827–838.

Kerridge,A., Willetts,A. and Holland,H. (1999) *J. Mol. Catal., B Enzymol.*, **6**, 59–65.

Laemmli,U.K. (1970) *Nature*, **227**, 680–685.

Lawrence P,W. (2002) *Enzyme Microbiol. Technol.*, **31**, 577–587.

Lee,K., Brand,J.M. and Gibson,D.T. (1995) *Biochem. Biophys. Res. Commun.*, **212**, 9–15.

Lessner,D.J., Johnson,G.R., Parales,R.E., Spain,J.C. and Gibson,D.T. (2002) *Appl. Environ. Microbiol.*, **68**, 634–641.

Liang,J., Edelsbrunner,H. and Woodward,C. (1998) *Protein Sci.*, **7**, 1884–1897.

Moore,C.D., Shahrokh,K., Sontum,S.F., Cheatham,T.E. and Yost,G.S. (2010) *Biochemistry*, **49**, 9011–9019.

Nestl,B.M., Nebel,B.A. and Hauer,B. (2011) *Curr. Opin. Chem. Biol.*, **15**, 187–193.

Nishino,S.F. and Spain,J.C. (1995) *Appl. Environ. Microbiol.*, **61**, 2308–2313.

Parales,R.E., Huang,R., Yu,C.-L., *et al.* (2005) *Appl. Environ. Microbiol.*, **71**, 3806–3814.

Pavelka,A., Chovancova,E. and Damborsky,J. (2009) *Nucl. Acids Res.*, **37**, 376–383.

Reetz,M.T., Kahakeaw,D. and Lohmer,R. (2008) *ChemBioChem*, **9**, 1797–1804.

Ricci,L.C., Comasseto,J.V., Andrade,L.H., Capelari,M., Cass,Q.B. and Porto,A.L.M. (2005) *Enzyme Microbiol. Technol.*, **36**, 937–946.

Ronald,B. (2005) *Chem. Soc. Rev.*, **34**, 609–624.

Sambrook,J., Fritsch,E. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn, Cold Spring Harbor, NY.

Shainsky,J., Derry,N.-L., Leichtmann,Y., Wood,T.K. and Fishman,A. (2008) *Appl. Environ. Microbiol.*, **75**, 4711–4719.

Soskine,M. and Tawfik,D.S. (2010) *Nat. Rev. Genet.*, **11**, 572–582.

Tolia,N.H. and Joshua-Tor,L. (2006) *Nat. Methods*, **3**, 55–64.

Trott,O. and Olson,A.J. (2010) *J. Comput. Chem.*, **31**, 455–461.

Vargas,R.R., Bechara,E.J.H., Marzorati,L. and Wladislaw,B. (1999) *Tetrahedron: Asymmetry*, **10**, 3219–3227.

Zheng,G.W. and Xu,J.H. (2011) *Curr. Opin. Biotechnol.*, **22**, 784–792.