Oxidative biotransformations of phenol substrates catalysed by toluene dioxygenase: a molecular docking study

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:

Link to publication record in Queen's University Belfast Research Portal

Accepted Manuscript

Title: Oxidative biotransformations of phenol substrates catalysed by toluene dioxygenase: a molecular docking study

Author: Patrick Hoering Kyle Rothschild-Mancinelli Narain D. Sharma Derek R.Boyd Christopher C.R.Allen

Please cite this article as: Patrick Hoering, Kyle Rothschild-Mancinelli, Narain D.Sharma, Derek R.Boyd, Christopher C.R.Allen, Oxidative biotransformations of phenol substrates catalysed by toluene dioxygenase: a molecular docking study, Journal of Molecular Catalysis B: Enzymatic http://dx.doi.org/10.1016/j.molcatb.2016.10.013

Oxidative biotransformations of phenol substrates catalysed by toluene dioxygenase: a molecular docking study

Patrick Hoering[a], Kyle Rothschild-Mancinelli[a], Narain D. Sharma[b], Derek R.Boyd[b], Christopher C.R.Allen[a]*

[a] *School of Biological Sciences and Institute for Global Food Security, Queen's University Belfast, Belfast, Northern Ireland*

[b] *School of Chemistry and Chemical Engineering, Queen's University Belfast, Belfast, Northern Ireland*

*corresponding author, e-mail: *c.allen@qub.ac.uk; address: School of Biological Sciences Queen's University Belfast, Medical biology Centre, 97 Lisburn Rd., Belfast BT9 7BL, Northern Ireland*

GRAPHICAL ABSTRACT
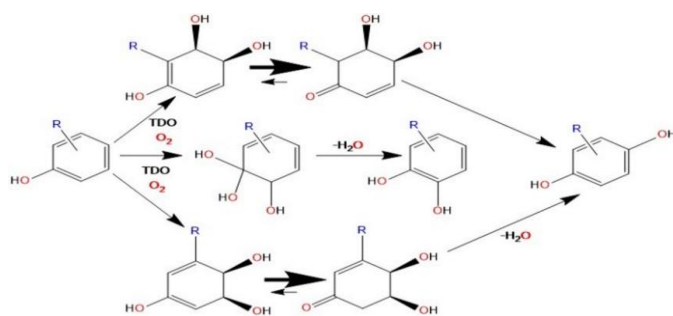
Toluene dioxygenase-catalysed dihydroxylation of phenols   Docking of 3-
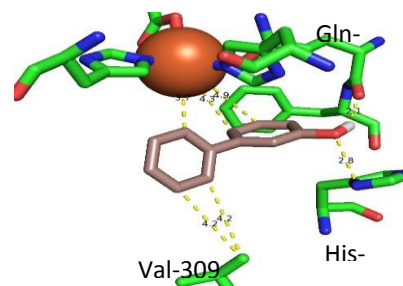
Highlights

- Comparison of docking of toluene and its X-ray crystal structure with TDO
- Binding interactions of substituted phenol substrates at the TDO active site.
- Evidence for *cis*-dihydrodiol tautomers derived from phenols and TDO.
- Support for phenol oxidation to catechols *via* triol intermediates using TDO.

ABSTRACT

Toluene dioxygenase-catalysed (TDO) oxidation converts substituted phenol substrates into catechols, hydroquinones, and chiral cyclohexenone *cis*-diol products. The ratio between the isolated products varied widely even between similar substrates, e.g. *o*-cresol, *m*-cresol and *p*-cresol. These differences are caused by different binding interactions within the active site of TDO. This study provides insight into the binding interactions by molecular docking using AutoDock tools. The nature of binding of phenolic substrates was of major interest, in order to explain the observed regio- and stereo-selectiviy of product formation. The ellipse-shaped binding pocket of TDO consists of a polar and a hydrophobic region, limiting the possible substrate orientations. The phenolic hydroxyl group was preferentially hydrogen bonded with Gln-215 and His-311 in the active site. In some cases, a hydrogen bond was formed with other amino acids, *e.g.* Asp-219and Met-220, instead. The position and type of the substituent on the phenol ring influences the formation of transient intermediates, and thus the nature and stability of the major isolated product.

Keywords: Toluene dioxygenase

Phenols

Catechols

Hydroquinones

Cyclohexenone cis-diols

1. Introduction

The biotransformations of phenols **A** (Scheme 1) to yield catechols **D**, as major metabolites, have been widely investigated using a range of bacteria expressing ring hydroxylating dioxygenase enzymes (RHD). [1-14]. Thus, toluene dioxygenase (TDO) [2, 3, 7-11], naphthalene dioxygenase (NDO) [9, 12], biphenyl dioxygenase (BPDO) [9, 12, 13] and *o*-xylene dioxygenase (OXDO) [14] have all been used to catalyse the formation of catechols **B**, as major metabolites of phenols **A**. To date, the binding interactions of phenolic substrates with amino acid residues at the active site of RHDs are relatively unexplored. In an earlier publication [9], a phenol-binding role was postulated for the imidazole ring of His-311 in the active site of TDO. The current investigation provides further insight into phenolic substrate interactions with the active site.

Recent reports showed that, in many cases, catechols **B** were accompanied by a new family of cyclohex-2-en-1-one *cis*-diols, as major metabolites (**F** and **G**) [7-11, 13]. Biotransformations of phenols, with the TDO-expressing constitutive mutant *Pseudomonas putida* UV4 and recombinant strain *Escherichia coli* p-CL-4t, resulted in more than twenty cyclohex-2-en-1-one *cis*-diols (**F** and **G**) being added to this new family of metabolites [7-11]. Hydroquinones **E** were also detected as minor phenol derivatives using bacteria expressing TDO [7-11], NDO [12], BPDO [12]and OXDO [14].

TDO-catalysed hydroxylation mechanisms have been postulated to account for: (a) the formation of benzene-1,2-diols (catechols, **D**) *via* triol intermediates (**C**) [3], (b) benzene-1,4-diols (hydroquinones, **E**) [11, 14] and (c) cyclohex-2-en-1-one *cis*-diols (**F** and **G**) *via cis*-dihydrodiol metabolites (**B** and **H**) [7-11,13] (Scheme 1). While the dehydration of cyclohex-2-en-1-one*cis*-diols **F** and **G**, to yield hydroquinones **E**, was observed in the presence of acid and occurred during biotransformations, neither triol **C** nor *cis*-dihydrodiol intermediates **B** and **H** were detected. An X-ray crystal structure of TDO from *P. putida* F1 was published, showing the position of the toluene substrate relative to the iron atom [15]. A figure was produced earlier, showing postulated binding interactions between toluene and proximate amino acids at the TDO active site, with calculated bond lengths below 3.7Å [9]. As the TDO binding site has an ellipse-shaped pocket, formed by seventeen amino acid residues, proposed binding modes included Van der Waals interactions and/or edge-to-face bonding to nearby amino acids, having non-polar alkyl or aryl groups (Phe-366, Phe-216, Ile-324) [9]. It was noted that, within the TDO active site, the *meta*-position of bound toluene was very close to the imidazole ring of His-311. This observation led to the proposal, that an attractive interaction between the acidic OH group of *meta*-phenol substrates and the basic imidazole ring, could explain the observed regio- and stereo-selective dihydroxylation of phenolic substrates, to yield *cis*-dihydrodiols **B** and **H** and the preferred ketotautomers **F** and **G** [7–11]. Conversely, the TDO-catalysed *cis*-dihydroxylation was inhibited by the presence of an alkyl group at the *meta*-position of non-phenolic disubstituted benzene substrates, *e.g. m*-chlorotoluene [16].

The major objectives of this study of TDO were to examine:

(i)   The use of molecular docking methods, to investigate the postulated formation of transient *cis*-diol (**B**, **H**) and triol (**C**) intermediates.

(ii)   The possible binding interactions between active site amino acids and phenol substrates, and proximity of phenol C=C bonds to the iron-dioxygen complex at the active site [9].

3

(iii)   The docking orientations and experimental results from phenol substrate biotransformations with TDO, in order to determine which arene bond (regiochemistry) and face (stereoselectivity) is preferred.

Docking of the substrates was conducted using two models based on the TDO crystal structure. The first model in this study used the native form without bound dioxygen (PDB code: 3EN1) [15], herein called 3EN1M. The second model was a derivative of 3EN1 with dioxygen incorporated (called 3EN1M-O_2). Due to the absence of a crystal structure of TDO with bound dioxygen, the O_2 molecule was placed manually [17]. While theFe(III) prosthetic group is fully coordinated with bound dioxygen, an Fe(II) group in the absence of dioxygen would be coordinated by two histidines and an aspartate [18], leaving no possibility of acting as a Lewis acid. The iron in both models was thus considered to be without charge during the docking. It also has to be noted that the crystal sructure of TDO showed no water molecules in the active site, which could influence substrate binding.

2. Computational modeling

Molecular modeling. Docking studies were performed with AutoDock suite 4.2 (autodock4 and autogrid4 scripts) utilising the user interface and python scripts of AutoDock Tools 1.5.6. The AutoDock suite and AutoDock tools (ADT) are available for download from the Scripps Research Institute (http://autodock.scripps.edu/) [19]. The crystal structure of TDO was obtained through ADT from the Protein Data Bank (PDB code 3EN1, resolution of 3.2 Å) as a .pdb-file. The toluene structure, contained inside the downloaded protein structure, was removed with UCSF Chimera prior to receptor preparation. This new structure was minimized, then validated using the SAVES server, which includes the ERRANT program used to check the quality factor of the model and Ramachandran plot to certify that the residues were in the correct positions.

Water molecules were removed and the protein was protonated with ADT. Receptor preparation was then performed using the 'prepare_receptor4.py' python script of ADT in order to repair missing atoms, add hydrogens, detect aromatic carbons, add Gasteiger charges and merge non-polar hydrogens. The resulting .pdbqt file for TDO was used for all docking experiments. The.pdb files for chemical compounds were created with UCSF Chimera 1.10.2 (https://www.egl.ucsf.edu/chimera/). The imported .pdb files of the chemical compounds into ADT, were automatically transformed into the required .pdbqt format.

The dimensions for the docking grid were adjusted to include all amino acids within 5 Å of toluene in the crystal structure. The grid resolution was set to 0.247 Å, and specific grid parameters were: Size(X, Y, Z): 40, 60, 72, and offset (X, Y, Z): 12.889, -4.528, 3.806. The resulting grid parameter file (.gpf) was used to create the grid map files (.glg) using autogrid4. The docking search protocol used the internal default docking parameters of AutoDock 4.2, starting the ligand at a random location. The docking was set to be performed as 100 runs of the Lamarckian genetic algorithm with a population size of 150 each, terminating after 2 500 000 energy evaluations or 27 000 generations (whichever came first). The docking results were analysed with ADT to obtain docking coordinates, calculated binding energy (in kJ) and calculated dissociation constant K_D (in μM). All docking studies were performed as rigid docking, keeping positions, orientations and protonation states of all amino acids locked in place. The 100 docking orientations were automatically analysed by ADT and divided into conformationally and energetically similar groups. The orientation with the highest binding energy of each of these groups is shown.

4

Addition of dioxygen to 3EN1M. Dioxygen was placed in the 3EN1M by overlaying the iron prosthetic group (Fe, dioxygen, His-208, His-213, and Asp-362) of the NDO crystal structure (PDB code: 1O7M) with the prosthetic group of TDO (Fe. His-222, His-228, Asp-376) using PyMol. This was performed to create a similar complex as previously published [17].

Visualization. The amino acids of the TDO active site were visualized with PyMol 1.7.4.5 from the Protein Data Bank (PDB) 3EN1 file coordinates. The docking results were imported into PyMol from the respective .pdb files created with AutoDock4 and AutoDock Tools. Measurements between atoms were calculated with PyMol's incorporated measurement tool.

3.

4. Results and Discussion

Biotransformations using *P. putida* UV4 and *E. coli* (p-CL-4t) whole cells, with toluene **1a**, phenol **2a**, 2-methylphenol (*o*-cresol) **3a**, 3-methylphenol (*m*-cresol) **4a**, *m*-methoxyphenol **5a**, 3-hydroxybiphenyl (3-phenylphenol) **6a** and 4-methylphenol (*p*- cresol) **7a** substrates, yielded a range of metabolites including *cis*-dihydrodiols, cyclohexenone *cis*-diols, catechols and hydroquinones [7-11]. The observed relative and isolated yields of phenol metabolites were dependent on the substrate type, TDO source (*i.e.* clone or mutant), the possibility of further metabolism and metabolite stability. It was anticipated that these molecular docking studies could, in principle, allow both a comparison with experimental results and a prediction of the regio- and stereo-selectivity of TDO for *cis*-dihydroxylation of the substrate.

The binding pocket of TDO is formed by ten hydrophobic amino acids: Phe-216, Met-220, Ala-223, Leu-272, Ile-276, Val-309, Leu-321, Ile-324, Phe-366 and Phe-372, and three polar amino acids: Gln-215, Asp-219 and His-311 [15]. The polar amino acids are more likely to facilitate strong interactions with the OH group of phenol substrates from a preliminary inspection of the active site. For clarity, only Gln-215, Phe-216, Val-309 and His-311are depicted in the figures, since the other non-polar amino acids are only responsible for non-substrate specific interactions with the docked ligands.

3.1. TDO-derived metabolites of toluene1a

Biotransformations of toluene **1a** using *P. putida* UV4 and *E. coli* (p-CL-4t) (TDO sources), followed by extraction and NMR analysis of the products, confirmed that *cis*-dihydroxylation had occurred, exclusively on one face of the 2,3-bond, resulting in (1*S,2R*)-*cis*-dihydrodiol **1b** with >95% relative yield and >98% enantiomeric excess (*ee*) (Scheme 2). A very small proportion of TDO-catalysed benzylic monohydroxylation, followed by *cis*-dihydroxylation also occurred, to yield the relatively stable triol **1c** (<5% relative yield) [20]. This exclusive preference for TDO-catalysed *cis*-dihydroxlation at the 2,3-bond, rather than the 1,2- or 3,4 bonds, was found in virtually all other monosubstituted benzene substrates, as indicated in a recent review and references therein [21].

The published crystal structure of TDO [15] is consistent with the observed regio- and stereo-selectivity of product formation, based on the orientation and proximity of toluene substrate **1a** towards the iron atom prosthetic group. The crystal structure shows how close the Me group of toluene **1a** is located to Ala-223, Val-309, and Ile-324, which are the determining interactions for the observed orientation (see Fig.S1). Additionally, His-222 stabilises the observed conformation by an edge-to-face (T-bonded) interaction with the aromatic ring (Fig.S2A). The *in silico* docked orientation of toluene **1a** (100% iterations

in both models, see Fig.S2B, Fig.S2C and Table 1) was comparable to the position of toluene in the crystal structure (Fig.S2A) with only slight differences (Figs.1A+B, Table 1). This confirmed the viability of the docking procedure and the parameters used. The biggest divergence (1.4 Å) was observed at the Me group of bound toluene. For this study, a variation in interatomic distances of less than 1.6 Å was considered to be within the limits imposed by the low crystal structure resolution of 3EN1 (3.2Å) [15]. Some differences between the crystal structure and computational model (Fig.1A+B) were expected, but it was important to show that a similar orientation was obtainable with the applied software. The *in silico* docking of toluene **1a** revealed no additional orientations, which could explain the presence of the very minor metabolite **1c**of the TDO-catalysed oxidation of toluene **1a** [20]. These docking results also suggest that bound dioxygen is unlikely to affect the toluene binding position significantly.

3.2. TDO-catalysed oxidation pathways of phenol 2a

Early studies of the biotransformation of phenol **2a**, using wild type *P. putida* F1 and the derived mutant *P. putida* 39D strains, showed evidence of catechol formation **2d**, *via* TDO-catalysed oxidation (Scheme 3) [2, 3]. Both *P. putida* 39D and UV4 mutant strains appeared to express an identical TDO enzyme [21]. It was postulated earlier [3] that triols of similar structure to **2c** would be undetected intermediates during TDO-catalysed oxidation of phenols to yield catechols. Support for this *ipso*-dihydroxylation pathway, was thus sought during the molecular docking study. Using the *P. putida* UV4 strain, catechol **2d** was only detected as a very minor metabolite (<5%), with hydroquinone **2e** being the major metabolite (>95%) [7]. This result was consistent with a TDO-catalysed *cis*-dihydroxylation at the 3,4-bond of phenol **2a**, to form the undetected *cis*-dihydrodiol **2b** and cyclohexenone *cis*-diol **2f**, followed by dehydration.

The *in silico* docking studies, with 3EN1M and phenol **2a**, suggested that formation of the unstable triol intermediate **2c** would only occur as a very minor pathway. Only 2% of iterations (Table 1) produced the required orientation for this pathway, leading to dehydration and catechol formation**2d** (Fig.S3A), as observed experimentally when using *P. putida* UV4. The remaining docking iterations showed two quite different orientations (Fig. 2A+B) for *cis*-dihydroxylation at the C-3 and C-4 atoms of the proximate C=C bond of phenol **2a**, which would produce an enantiomeric mixture of the *cis*-dihydrodiol intermediate **2b**. Dihydroxylation of the major orientation (60% iterations, Fig.2A) would result in a (1*S*,2*R*) configuration, while the minor orientation (38% iterations, Fig.2B) would lead to a (1*R*,2*S*) configuration with a theoretical *ee* value of *ca*. 20%. As keto *cis*-diol **2f** was not detected, it was assumed that the initially formed enol intermediate **2b** spontaneously eliminated water to form the main product, hydroquinone **2e**, or was further metabolised [7].

When using 3EN1M-O_2 (Figs. 2C+D), the orientation in Figure 2B was not obtained. However, while the major orientation (55% iterations, Fig.2C) was similar to Figure 2A, the previously minor orientation (2%, Fig.S3A), was much more abundant (45% iterations, Fig.2D). In the context of the published biotransformation results [3,7,21], when dioxygen was present in the docking study, a marked increase in the proportion of catechol **2d** relative to hydroquinone **2e** was predicted.

TDO-catalysed *cis*-dihydroxylation of monosubstituted benzene substrates at the 3,4-bond, to yield the corresponding *cis*-dihydrodiols, *e.g.* **2a** ⟶ **2b**, is unusual [21]. A precedent was found using

6

fluorobenzene as substrate with *P. putida* UV4, when *cis*-dihydroxylation also occurred at the 3,4-bond. This yielded the corresponding *cis*-dihydrodiol, as a very minor isomer (< 5% relative yield), with an *ee* value of *ca*. 20% favouring the (1*S*,2*R*) configuration [22]. The OH group and F atom are both relatively small substituents of similar size, capable of H-bonding at the active site. This suggests that TDO can only partially distinguish between these substituents. The orientations (Fig.2A+B) would produce opposite enantiomers of *cis*-dihydrodiol product **2b**, due to their different binding interactions within the active site of TDO. In Figure 2A the phenol OH group not only interacts with His-311 (distance between the O atom in phenol **1a** and nearest N atom of His-311, 4.0 Å), but also shows a possible interaction with Asp-219 (distance: 2.1 Å, not shown) and Met-220 (distance: 2.9 Å, not shown). The orientation in Figure 2B shows interaction with His-311 (distance: 3.0 Å) and a shorter hydrogen bond with another amino acid, namely Gln-215 (distance: 1.8 Å).

Distances between the Fe atom and C-3 and C-4 atoms of the nearest C=C bond in phenol **2a** at the TDO active site were: (i) 4.9 Å and 5.4 Å, thus yielding the (1*S*,2*R*) enantiomer of *cis*-dihydrodiol **2b** (Fig.2A) and (ii) 4.6 Å and 4.4 Å yielding the (1*R*,2*S*) enantiomer **2b** (Table 1). X-ray crystallography of NDO, showed the dioxygen molecule bound side-on to the Fe atom with O-Fe distances of 2.3 Å and 2.1Å [23]. It is anticipated that similar distances between dioxygen and the Fe atom will be found during TDO-catalysed *cis*-dihydroxylations of phenols **2a-7a** in the current study. The results obtained support this observation, as the space between docked phenol substrates and the Fe-atom is clearly sufficient to accommodate a dioxygen molecule complexed in a similar side-on manner. Indeed, as noted earlier, the docked toluene **1a** is further from the iron atom than in the crystal structure. All docked substrates in our study were slightly further away from the iron atom than the toluene observed in this crystal structure.

The stronger interaction for two of the orientations of the phenol molecule **2a** appears to be a hydrogen bond between the phenol OH group and Gln-215, either with the amide C=O group in the backbone (Fig.2B), or that of the side chain amide C=O group (Fig.S3A). The major orientation (Fig.2A) probably has a similar interaction with the backbone C=O group of Asp-219 (distance: 2.1 Å, not shown) and Met-220 (distance: 2.9 Å, not shown). This interaction was only observed with the parent phenol **2a** and *o*-cresol **3a** (Fig.3B) as docking ligands.

3.3. TDO-catalysed oxidation pathways of *o*-cresol 3a.

o-Cresol **3a** was a poor substrate for *P. putida* UV4 and *E. coli* p-CL-4t, compared to *m*-phenols **4a**, **5a** and **6a**. Extraction of metabolites, and chromatographic separation, produced (4*S*,5*R*,6*S*)-cyclohex-2-en-1-one *cis*-diol **3f** (>98% *ee*), as the preferred tautomer of *cis*-dihydrodiol **3b**, but in very low isolated yield (1%) [7]. LC-TOFMS analysis confirmed that *cis*-dihydroxylation at the 3,4-bond had occurred to give metabolite **3f**. GC-MS analysis, after silylation of the dried crude extract, indicated that hydroquinone **3e** (90% relative yield) and catechol **3d** (10% relative yield) were also present. The formation of aromatic products **3d** and **3e** could occur *via* dehydration of the undetected triol **3c** and enol *cis*-diol **3h** intermediates respectively (Scheme 4).

Three orientations (Fig.3A+B, Fig.S4A; Table 1) were predicted from 3N1M docking of *o*-cresol **3a**. The most abundant orientation (50% iterations, Fig.3A), would lead to the formation of the undetected *cis*-dihydrodiol **3h**. However, neither compound **3h** nor its keto tautomer **3g** were detected. In common

7

with the preferred orientation of phenol **2a** (Fig.2A), the *cis*-dihydroxylation was again found to occur at an unusual 4.5-bond position, when using *o*-cresol **3a** as substrate. Dehydration of enol *cis*-diol **3h** would then yield hydroquinone **3e** (Scheme 4). The orientation in Figure 3A shows an interaction with His-311 and a strong hydrogen bond to Gln-215, but weak interaction with Val-309.

Formation of the (4*S*,5*R*,6*S*)-cyclohex-2-en-1-one*cis*-diol **3f**, *via* the initially formed (1*S*,2*R*)-*cis*-dihydrodiol **3b**, could be predicted from the orientation shown in Figure 3B (24% docking iterations). The OH group shows interaction with His-222, Asp-219, Met-220 and His-311; potentially the Me group could have a weak interaction with Val-309. However, both intermediates **3b** and **3h** can readily dehydrate to yield hydroquinone **3e**.

The third orientation of *o*-cresol (26% docking iterations, Fig.S4AC) does not present any interactive site of the aromatic ring to the Fe atom, which could readily account for the formation of the undetected triol **3c** as a precursor of the minor metabolite **3d**. The only strong amino acid interaction with this orientation was with Gln-215, with a calculated hydrogen bond length of 2.1 Å.

The dioxygen-coordinated docking study with 3EN1M-O$_2$ando-cresol **3a**, did not show an orientation leading to the formation of **3h**. The predominant orientation (93% iterations, Fig.3C) would lead to the formation of the *cis*-diol tautomer **3b** and the experimentally observed *cis*-diol **3f** and hydroquinone **3e**. This orientation is superimposable on Figure 3B (overlay in Fig.S4B). On the other hand, Figure 3D (7% docking iterations) is a new orientation and does not account for any identified metabolites.

3.4. TDO-catalysed oxidation pathways of *m*-cresol 4a, *m*-methoxyphenol 5a, 3-phenylphenol 6a

Starting with LC-TOFMS analysis, cyclohex-2-en-1-one *cis*-diol metabolites **4f–6f**, were detected from biotransformation of the corresponding phenols **4a–6a** with *P. putida* UV4 and *E. coli* p-CL-4t whole cells (Scheme 5). *P. putida* UV4 biotransformations of these substrates gave enantiopure cyclohex-2-en-1-one *cis*-diol metabolites (isolated yields): (4*R*,5*S*)-**4f** (13%) [9], (4*S*,5*S*)-**5f** (38%) [11], and (4*R*,5*S*)-**6f** (28%) [9]. The intermediate *cis*-dihydrodiol tautomers **4b–6b** were not detected. A stable *cis*-dihydrodiol **6b'** was isolated as a minor metabolite (10% yield) from 3-phenylphenol **6a** [9]. GC-MS analysis of the silylated metabolites from phenols **4a** and **5a**, consistently showed the hydroquinones **4e** and **5e**, in much higher relative yields (>95% and 85%), compared with the corresponding catechols **4d** and **5d** (<5% and 15%). Dehydration of the *cis*-dihydrodiols **4b** and **5b** could account for the presence of hydroquinones **4e** and **5e**. LC-TOFMS analysis also showed the presence of the ring-opened metabolites **5i** and **6i** and the triol metabolite **6j** [9].

Molecular docking of *m*-cresol **4a**, using 3EN1M, produced two orientations (Figs.4A+B, Table 1). The major orientation (66% of docking iterations; Fig.4A) would result in TDO-catalysed *cis*-dihydroxylation to form enol *cis*-dihydrodiol **4b**. The phenol OH group hydrogen bonded toa C=O group (Gln-215) and an imidazole ring (His-311). The second orientation (34% iterations; Fig.4B) would result in *cis*-dihydroxylation of the 1,2-bond, to form the postulated triol intermediate **4c**, which dehydrates to yield catechol **4d**. When using 3EN1M-O$_2$ the only obtained orientation would favour formation of triol **4c**, and thus catechol **4d**, over *cis*-diol **4f** (100% iterations, Fig. 4C). Figure 4B and 4C are superimposable (see overlay in Fig.S5).

8

The major orientation of *m*-methoxyphenol **5a** (54% iterations, Fig.5A) would yield the unstable triol intermediate **5c**, which would spontaneously dehydrate to yield catechol **5d**. The substrate orientation shown in Figure 5B (12% docking iterations) would result in TDO-catalysed *cis*-dihydroxylation, to yield *cis*-dihydrodiol **5b**. This metabolite could either dehydrate to yield hydroquinone **5e** or tautomerise to yield the isolated cyclohex-2-en-1-one *cis*-diol tautomer **5f** (Scheme 5). None of the other two orientations (22% iterations, Fig.S6A; and 12% iterations, Fig.S6B) are consistent with known metabolites of *m*-methoxyphenol **4a**. When using 3EN1M-O_2, a similar situation was observed to the docking of *m*-cresol **4a**. where orientations would favour the formation of a catechol (Figs.5C+D, Table 1).

All orientations of *m*-methoxyphenol **5a** in 3EN1M showed close proximity to residue Gln-215 (1.7 – 2.2 Å, Fig.5 and Fig.S6; Table 1), forming hydrogen bonds to either the C=O group in the side chain or the polypeptide backbone. The methoxy group also appeared to have hydrophobic interactions with Val309 (3.1–3.5 Å, Fig.5 and Fig.S6A; Table 1).

The molecular docking of 3-phenylphenol **6a** using 3EN1M produced a single orientation (100% docking iterations; Fig.6A, Table 1). This orientation would allow the formation of the major *cis*-dihydroxylated product **6b** and its preferred keto tautomeric form **6f**. It showed a possible phenyl interaction with Val-309 (4.2 Å), and hydrogen bond interactions of the phenol OH group with His-311 (2.8 Å) and Gln-215 (2.1 Å). No orientation was observed which would account for the formation of the minor *cis*-dihydrodiol **6b'**. With 3EN1M-O_2, the major orientation observed (97% iterations, Fig.6B, Table 1) differs only slightly to Figure 6A, presenting a similar orientation to the iron atom as before. The major difference is observed at the phenyl ring, which shifts away from the iron atom to make space for the dioxygen molecule (see overlay Fig.S7). Although the major product **6f** of this biotransformation would still be the same, a very minor orientation (3% iterations, Fig.6C) was observed that would be dihydroxylated to form a *cis*-diol metabolite **6b'**.

3.5. TDO-catalysed oxidationpathway of *p*-cresol7a

Analysis of the biotransformation of *p*-cresol **7a** by *P. putida* UV4 or *E. coli* p-CL-4t, showed the presence of catechol **7d** by comparison with an authentic sample (Scheme 6) [7]. No evidence was found of *cis*-dihydrodiol **7b** or its cyclohexenone *cis*-diol tautomer **7f**. Molecular docking of *p*-cresol **7a** using 3EN1M provided a single orientation (100% iterations, Fig.7A, Table 1), which is consistent with *ipso*-dihydroxylation at the 1,2-bond, to form the undetected triol intermediate **7c**, which spontaneously dehydrated to form catechol **7d**. The docked substrate showed an attractive interaction with the Val-309 residue (3.4 Å) and hydrogen bond formation with Gln-215 (2.0 Å).

While cyclohexenone *cis*-diol metabolites were isolated from 2- and 3-substituted and 1,4-disubstituted phenols [7-11], none were obtained from 4-substituted phenols. Based on earlier studies of non-phenolic 1,4-disubstituted benzene substrates [16], 4-substituted phenols, *e.g.* **7a**, might be expected to undergo TDO-catalysed *cis*-dihydroxylation, to yield conjugated *cis*-dihydrodiols, *e.g.* **7b**. It is also possible that intermediate **7b** could dehydrate forming catechol **7d**, rather than the non-conjugated cyclohexenone **7f**. These docking observations support formation of the catechol product **7d** from triol intermediate **7c**, as the preferred pathway rather than dehydration of *cis*-dihydrodiol **7b**.

The docking studies with 3EN1M (Fig.7A) suggest that *p*-cresol **7a** would have a reduced degree of freedom within the TDO binding pocket, yielding only triol **7c**, ultimately dehydrating to form **7d**. The

OH group of phenol **7a** forms a hydrogen bond with the side chain C=O group of Gln-215 and the Me group interacts with Val-309. A similar conclusion can be drawn from docking with 3EN1M-O_2, as no orientational changes were observed (100%, Fig.7B, Table 1).

3.6. Docking of phenolate anions formed from phenols 2a-7a

The docking of the phenolate anion variant for each phenol substrate was also conducted, as the change in protonation state could influence substrate-amino acid interactions. These studies showed clearly that docking of the phenolate anion of substrates **2a-7a** could not account for any of the observed metabolites. This was expected because of the high pK_a-value of the phenols (*ca.*10), which, at intracellular pH, favours the protonated form.

3.7. Binding energies

AutoDock 4.2 calculates the dissociation constant (K_D) directly from the binding energy [19], and therefore any discussion of binding energies directly relates to the dissociation constant. The calculated binding energies of the phenol substrates **2a-5a** and **7a** are between -4.3 and -5.0 kJ/mol, while substrate **6a** had an exceptionally high value of -7.25 kJ/mol (Table 1). The narrow range of binding energies is a result of similar binding interactions of phenols with TDO: The phenolic OH group forms a hydrogen bond to an amido carbonyl group (predominantly of Gln-215), while the alkyl or phenyl substituent interacts with Val-309 (and other non-polar amino acid residues). While the phenol ring of 3-phenylphenol **6a** binds in a similar manner to the other substituted phenols (**1a-5a** and **7a**), the phenyl-substituent is located in close proximity to Phe-366, Phe-372 and Ile-324, thus lowering the binding energy significantly (-7.25 kJ/mol). An edge-to-face interaction between the phenyl-ring of substrate **6a** and Phe-366 also contributes to the stability of the preferred orientation (Fig.8). The variation in binding energies between the orientations of each substrate differ only slightly, and can therefore be considered energetically equivalent for the purposes of substrate binding and product formation.

Dockings with 3EN1M-O_2 show a small, but consistent, increase in calculated binding energies of 0.3 – 0.4 kJ/mol compared with 3EN1M dockings. For all substrates, except **6a**, this difference is caused by simulated interactions with the non-polar dioxygen molecule. The incorporation of dioxygen had no effect on the binding energy of compound **6a**, probably due to its relatively large size compared with the othersubstrates (**2a - 5a** and **7a**). The position change of the phenyl ring (see Fig. S7) altered interactions with the hydrophobic region, *i.e.* the phenyl ring moved away from Phe-366 and Phe-372, but got closer to Val-309, Leu-321 and Ile-324. This probably lowered the contribution to the binding energy, whereas presence of dioxygen increased the binding energy, with both effects negating each other.

The binding energies given in this study only indicate how well a substrate would bind theoretically, and do not indicate if a reaction can be catalysed, as inhibitors would also exhibit high binding energies without being converted to a product. The relative rates of reaction of specific orientations are unrelated to the binding energy.

4. Conclusions

Any comparison of the relative proportions of predicted phenol metabolites using models 3EN1M and 3EN1M-O$_2$, with experimental results, should take account of the following factors: (i) spontaneous decomposition of metabolites can occur ,*e.g.* dehydration of 2,3-*cis*-dihydrodiols to yield hydroquinones and of 1,2-*cis*-diols to give catechols., (ii) further enzyme-catalysed reactions of initial metabolites, *e.g.* catechol-2,3-dioxygenase catalysed oxidative ring-opening of catechols, (iii) the qualitative nature of docking results might not quantitatively reflect experimental results.

The placement of the O$_2$ molecule in model 3EN1M to give 3EN1M-O$_2$, is an approximation from an X-ray crystal structure of NDO (pdb code: 1O7M) [23] and is similar to a model used earlier [17]. Although either model predicted nearly identical orientations for toluene **1a**, other dockings showed slight differences between 3EN1M and 3EN1M-O$_2$, *i.e.* **3a-6a**. While 3-phenylphenol **6a** docking with 3EN1M showed only one orientation, leading to *cis*-diol **6f**, docking with 3EN1M-O$_2$ also revealed an orientation that would explain the formation of *cis*-diol **6b'**. On the other hand, docking of *m*-cresol **4a** with 3EN1M-O$_2$ only showed the catechol **4d** forming orientation, whereas 3EN1M docking also showed the cis-diol **4f** forming orientation. It is therefore evident that both models need to be considered to better link the docking to the experimental results. The regio- and stereo-selectivity of the isolated *cis*-diol metabolites **1b, 3f-6f, 6b'** were in full accord with *in silico* predictions.

The docking results provide support for earlier postulates [3], suggesting that triol intermediates precede catechol formation during phenol biotransformation (**2c-5c, 7c**). Formation of isolated cyclohexenone *cis*-diols **2f-6f** would occur in a similar manner *via* enol-diol intermediates **2b-6b**.

The active site of TDO is divided into hydrophobic and polar regions, allowing for the binding of phenol substrates in specific orientations. The polar region consists of amino acids Asp-219, Gln-215 and His-311. The results with both docking models, 3EN1M and 3EN1M-O$_2$, suggest that hydrogen bonds, between the phenol OH-group with His-311 and Gln-215, play a significant role in substrate binding. This observation is consistent with the previous postulate [9]. The results show that there are some differences between docking with 3EN1M and 3EN1M-O$_2$, and it is concluded that the presence of the O$_2$ molecule in the active site does not adversely affect the hydrogen bonding associated between His-311 and Gln-215. We propose that the hydrogen bond interactions highlighted in this context, are as significant as the attractive interaction between the Me group of toluene and the hydrophobic region (exemplified by Val-309) within the TDO crystal structure [15]. It is also reasonable to presume that the polar part of the TDO binding pocket, involving His-311 and Gln-215, is responsible for the reported inhibitory effect of an alkyl-group in the *meta*-position (termed the '*meta*-effect') of non-phenolic substrates (*e.g. m*-halotoluenes) leading to benzylic monohydroxylation rather than *cis*-dihydroxylation [16]. The predicted binding interactions of phenols with His-311 and Gln-215 may also apply to other types of substrates capable of hydrogen bonding, *e.g.* substituted anilines. These amino acids are excellent targets for site-directed mutagenesis studies, which could confirm some of the observations of this study.

Earlier studies suggested that TDO was generally only able to catalyse *cis*-dihydroxylation of monocyclic arenes and other relatively small aromatic substrates [21]. The docking studies and biotransformation of 3-phenylphenol **6a**, demonstrate an ability of TDO to bind and dihydroxylate larger substrates.

The docking results with AutoDock show sufficient promise for use as a predictive tool for TDO-catalysed biotransformations. However, as differences and unexpected docking orientations, yielding no or as yet undetected products, were obtained with both models, further analysis of these binding interactions, using whole protein dynamic simulation studies, would now be warranted.

5. Acknowledgements

References

[1] R.C. Bayly, S. Dagley, D. T. Gibson, Biochem. J.101 (1966), 293-301.

[2] J. C. Spain, D. T. Gibson, Appl. Environ. Microbiol. 54 (1988) 1399-1404.

[3] J. C. Spain, G. J. Zylstyra, C. K. Blake, D. T. Gibson, Appl. Environ.Microbiol.55 (1989) 2648-2652.

[4] F. K. Higson, D. D. Focht, Appl. Environ. Microbiol.55 (1989) 946-952.

[5] C. Hinterregger, R. Leitner, M. Loidl, A. Ferschl, F. Streichsbier, Appl. Microbiol. Biotechnol. 37 (1992) 252-259.

[6] G. Bestetti, E. Galli, B. Leoni, F. Pelizzoni, G. Sello, Appl. Microbiol. Biotechnol. 37 (1992) 260-263.

[7] D. R. Boyd, N. D. Sharma, J. F. Malone, C. C. R. Allen, Chem. Commun. (2009), 3633-3635.

[8] M. Kwit, J. Gawronski, D. R. Boyd, N. D. Sharma, M. Kaik, Org. Biomol. Chem. 8 (2010) 5635-5645.

[9] D. R. Boyd, N. D. Sharma, P. Stevenson, M. Blain, C.McRoberts, J. T. G. Hamilton, J. M. Argudo, H. Mundi, L. A. Kulakov, C. C. R. Allen, Org. Biomol.Chem. 9 (2011) 1479-1490.

[10] D. R. Boyd, N. D. Sharma, J. F. Malone, P. B. A. McIntyre, P. J. Stevenson, C. C. R. Allen, M. Kwit, J. Gawronski, Org. Biomol. Chem. 10 (2012) 6217-6229.

[11] D. R. Boyd, N. D. Sharma, J. F. Malone, P. B. A. McIntyre, C. McRoberts, S. Floyd, C. C. R. Allen, A. Gohil, S. J. Coles, P. N. Horton, P. J. Stevenson, J. Org. Chem. 80(2015) 3429-3439.

[12] K. Lee, FEMS Microbiol. Lett. 255 (2006) 316-320.

[13] M. Sondossi, D. Barriault, M. Sylvestre, Appl. Environ.Microbiol.70 (2004) 174-181.

[14] D. Kim, J. S. Lee, K. Y. Choi, Y-S. Kim, J. N. Choi, S-K. Kim, J-C. Chae, G. J. Zlystra, C. H. Lee, E. Kim, Enzyme. Microb. Technol. 41 (2007) 221-225.

[15] R. Friemann, K. Lee, E. N. Brown, D. T. Gibson, H. Eklund, S. Ramaswamy, Acta. Cryst. D65 (2009) 24-33.

[16] D. R. Boyd, N. D. Sharma, N. I. Bowers, H. Dalton, M. D. Garrett, J. S. Harrison, G. N. Sheldrake, Org. Biomol. Chem.4 (2006) 3343-3349.

[17] M.A. Vila, M. Pazos, C. Iglesias, N. Vega, G. Seoane, I. Carrera, Chem Bio Chem, 17(2016) 291-295.

[18] S. M. Barry, G. L. Challis, ACS Catal. 3 (2013), 2362 - 2370

[19] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsel, A. Olsen, J. Comput. Chem.16 (2009) 2758-2791.

[20] D. R. Boyd, N. D. Sharma, N. I. Bowers, J. Duffy, J. S. Harrison, H. Dalton, (2000) J. Chem. Soc. Perkin Trans.1, 1345-1350.

[21] C.C.R. Allen, Science of Synthesis: Biocatalysis in Organic Synthesis, 3.1 (2015), 1-20.

[22] D. R. Boyd, N. D. Sharma, G. P. Coen, F. Hempenstall, V. Ljubez, J. F.Malone, C. C. R. Allen, J. T. G. Hamilton, Org. Biomol.Chem. 6 (2008) 3957-3966.

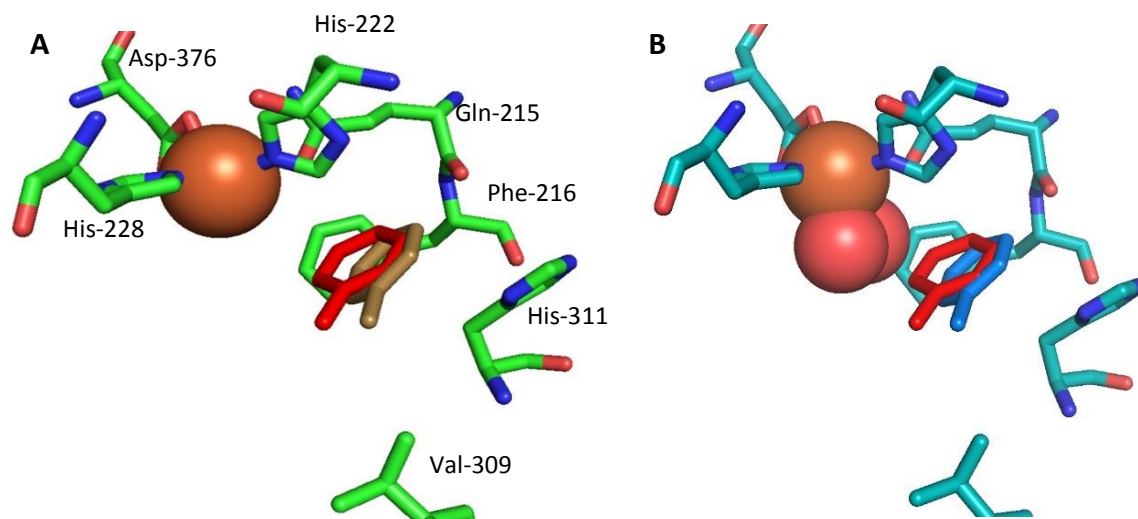[23] A. Karlsson, J. V. Parales, R. E. Parales, D. T. Gibson, H. Eklund, S. Ramaswamy, Science, 299, (2003) 1039-1042

Fig. 1. Comparison of docked and crystal structure orientations of toluene without (A) and with dioxygen (B).
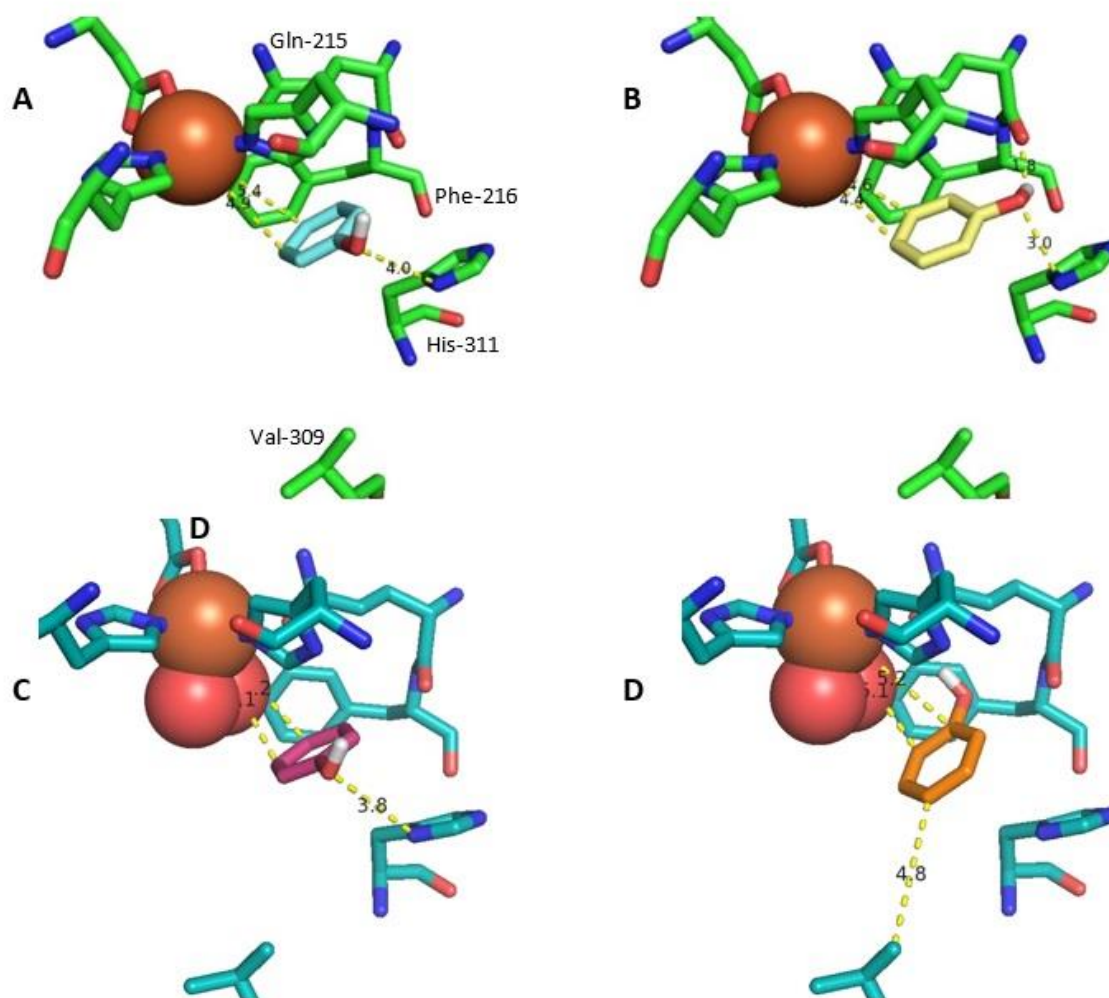
Fig. 2 Docking orientations of phenol **2a** without (A and B) and with (C and D) dioxygen.
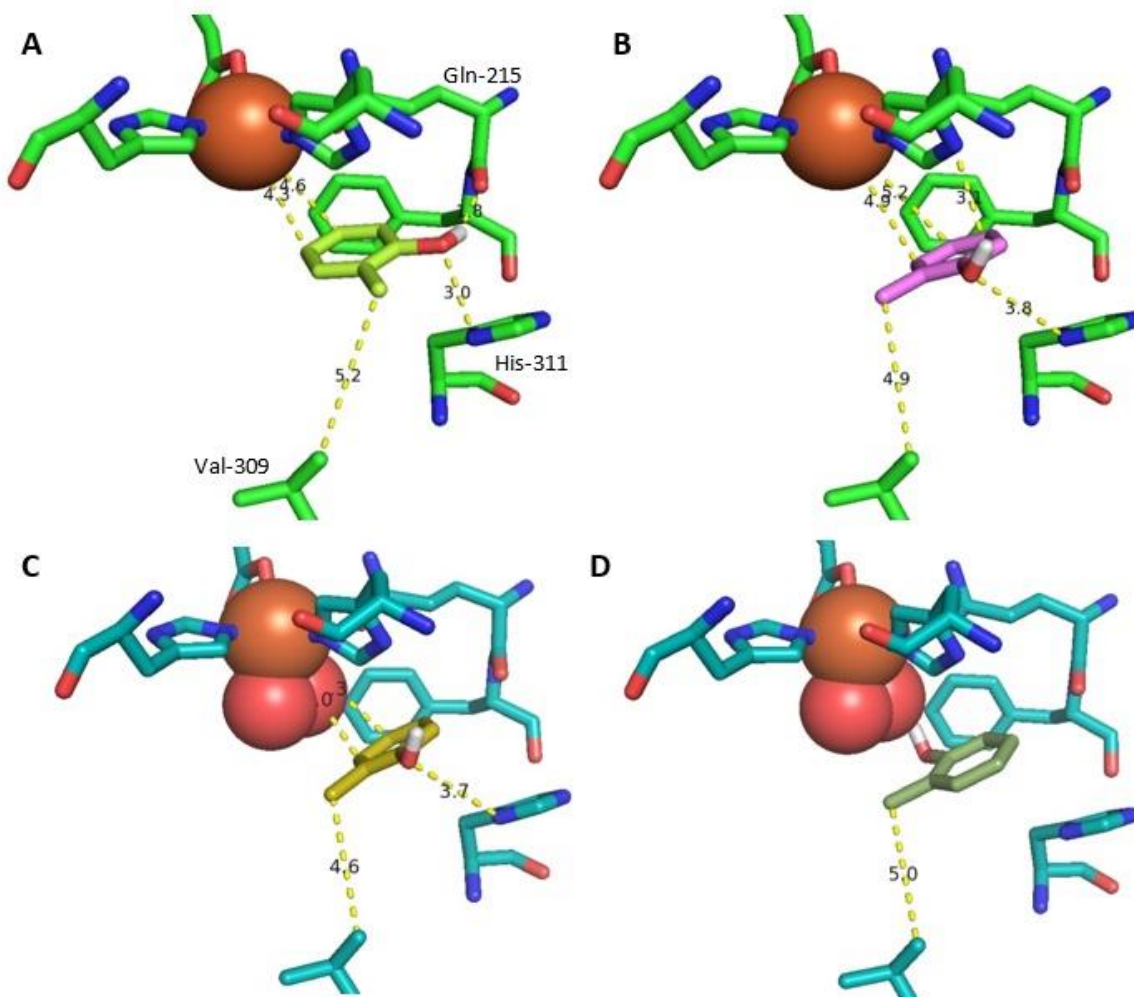
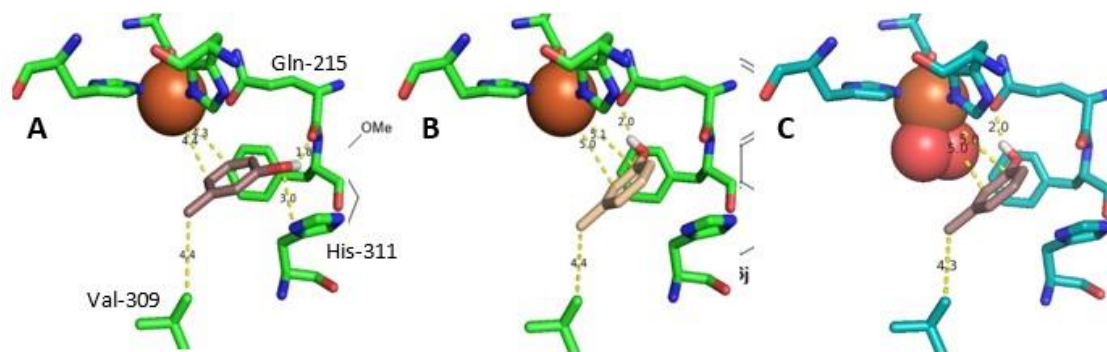**Fig. 3**Molecular docking orientations of *o*-cresol **3a**without (A + B) and with (C + D) dioxygen.

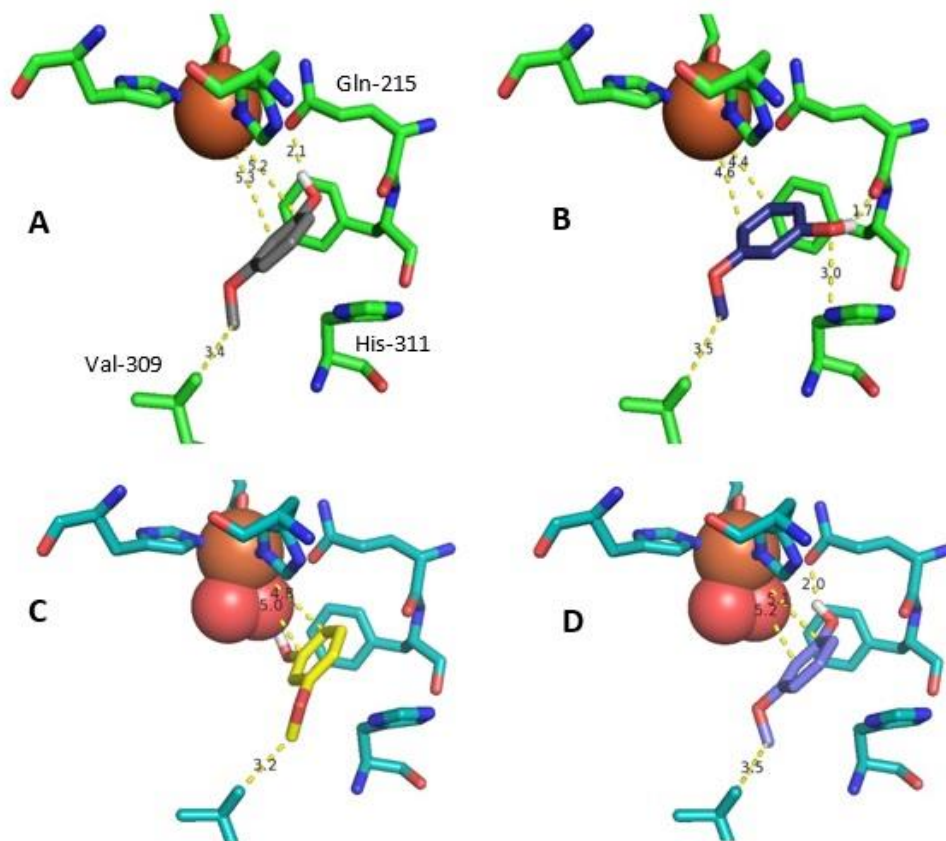**Fig. 4** Molecular docking of *m*-cresol **4a** without (A and B) and with (C) dioxygen.

Fig. 5 Molecular docking of *m*-methoxyphenol **5a** with (C + D) and without dioxygen (A + B) at the TDO active site.
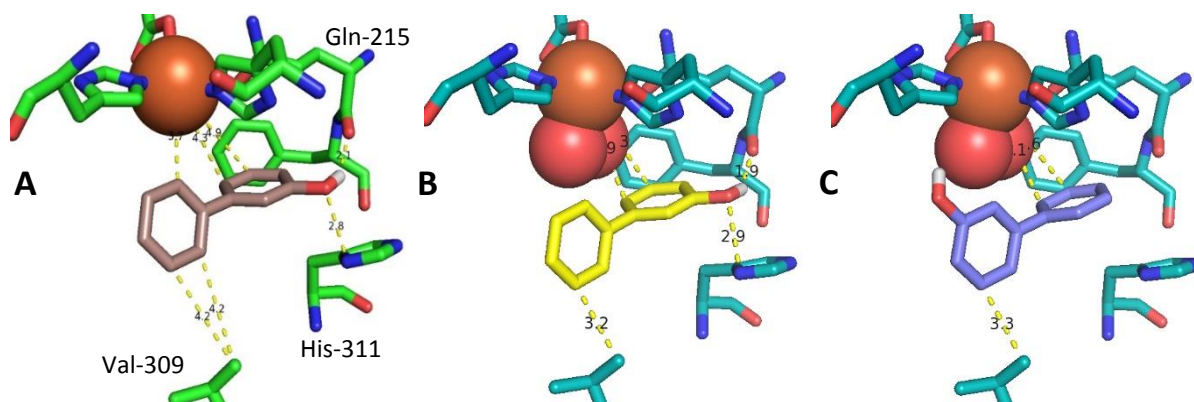
Fig. 6 Molecular docking of 3-hydroxybiphenyl **6a** with (B + C) and without (A) dioxygen at the TDO active site. Fig. 6C shows a possible orientation for the formation of **6b'**.
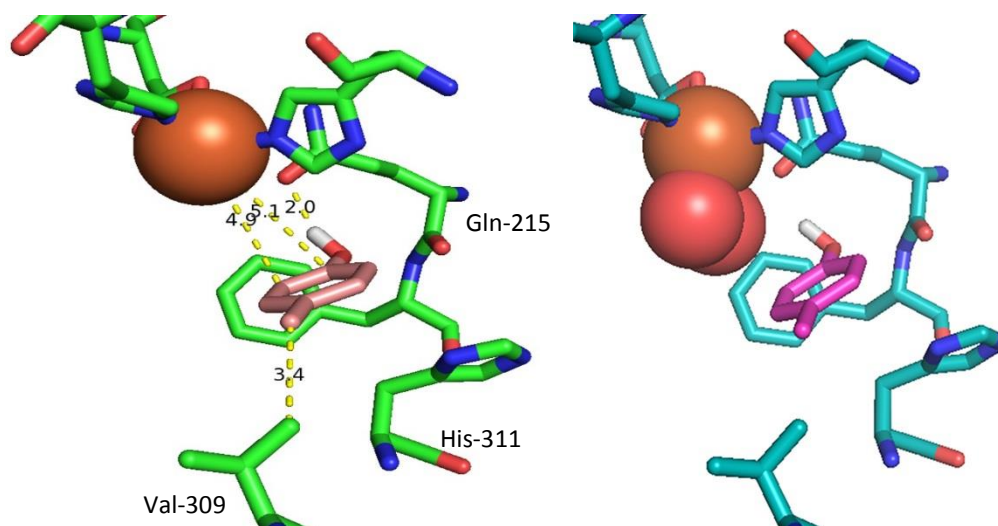
Fig. 7 Molecular docking of *p*-cresol without (A) and with (B) dioxygen.
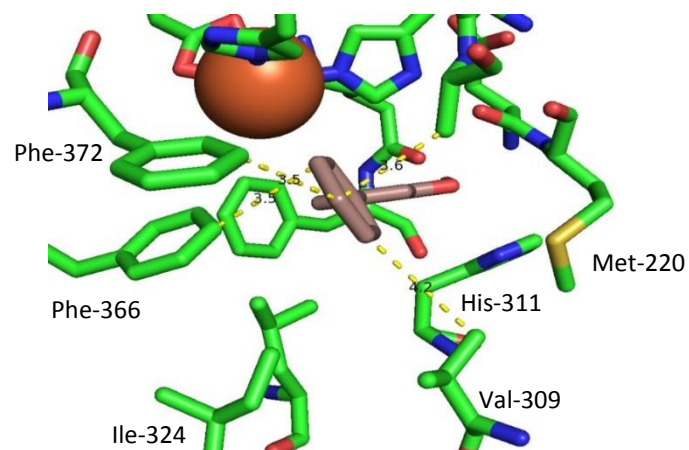
Fig.8 Orientation of the phenyl group of 3-phenylphenol 6a with a 60° dihedral angle relative to the phenol group
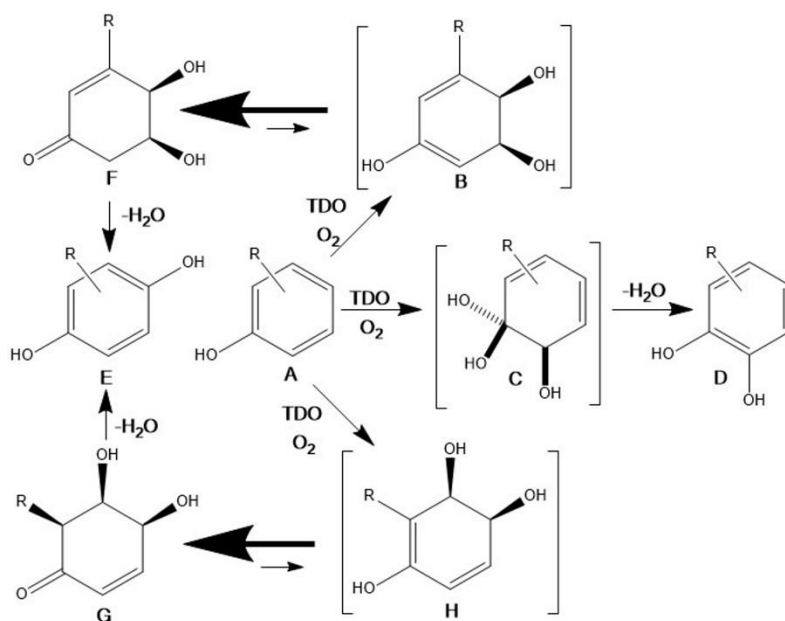
Scheme 1 Isolated catechols (**D**),hydroquinones (**E**), cyclohexenone*cis*-diols (**F** and **G**),postulatedtriols (**C**) and *cis*-dihydrodiol intermediate metabolites (**B** and **H**) derived from phenols (**A**).
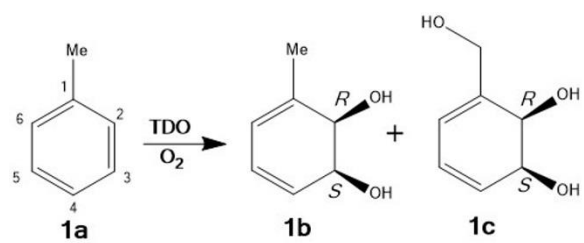
Scheme 2 *cis*-Dihydrodiol metabolites **1b** and **1c** derived from toluene**1a**.
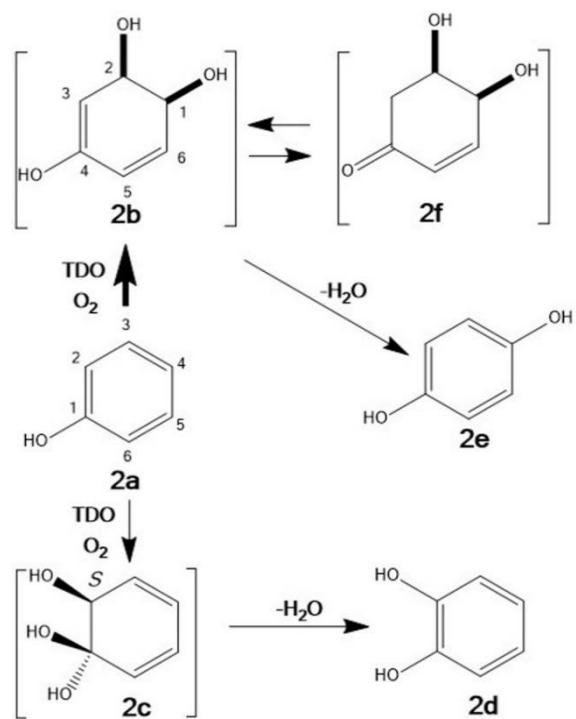
Scheme 3 Hydroquinone **2e** and catechol **2d** metabolites of phenol **2a**.
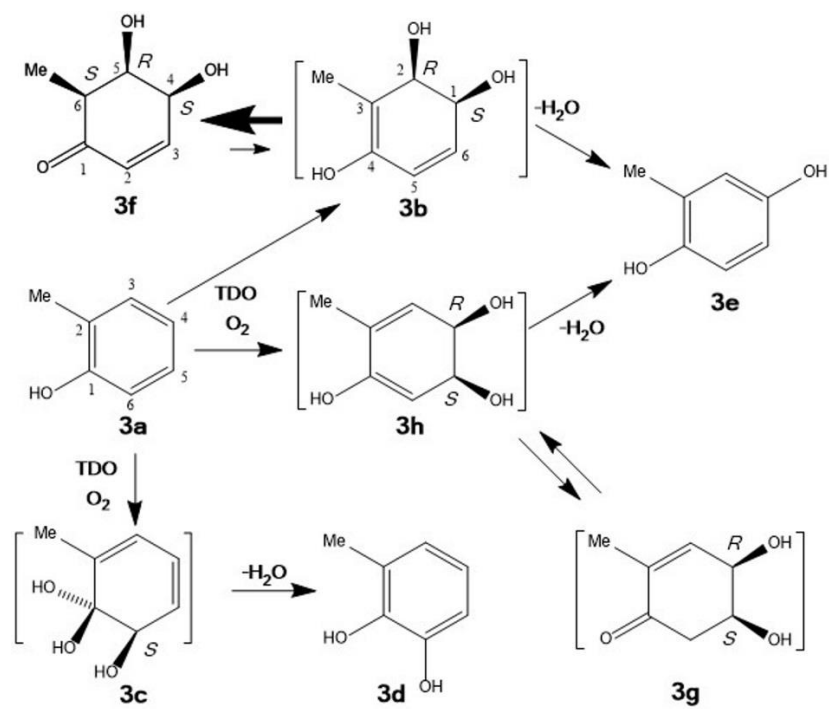
Scheme 4 Isolated metabolites (cyclohex-2-en-1-one *cis*-diol **3f**, catechol **3d**, hydroquinone **3e**) and intermediates (**3b** and **3c**) of *o*-cresol **3a.**
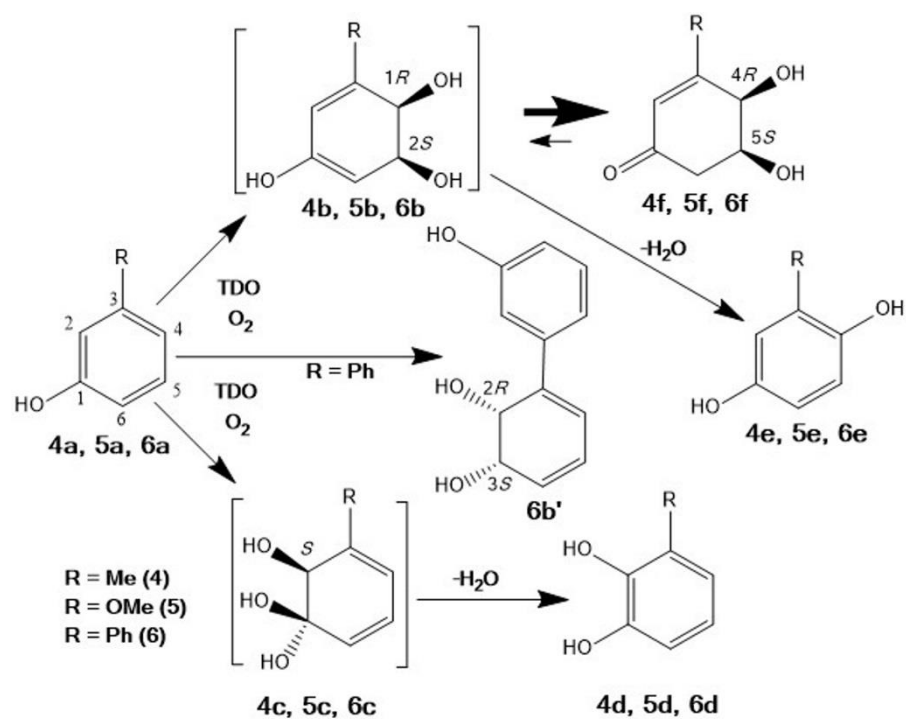
Scheme 5 *cis*-Dihydrodiol **4b**, **5b**, **6b**, **6b'**, cyclohex-2-en-1-one *cis*-diol **4f**, **5f**, **6f**, hydroquinone **4e**, **5e**, **6e,** triol **4c**, **5c**, **6c** and derived catechol **4d**, **5d**, **6d** metabolites of *m*-phenols **4a**, **5a** and **6a**.
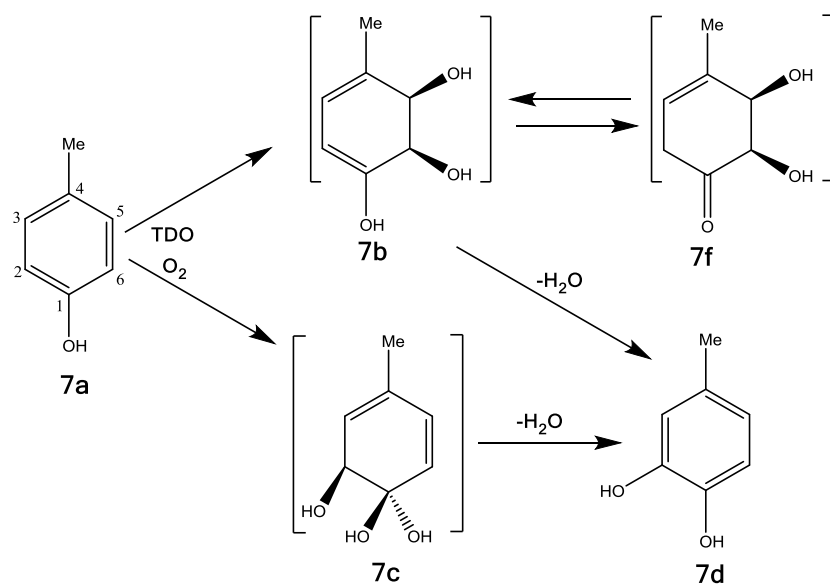
Scheme 6 Formation of catechol metabolite **7d** derived from phenol **7a**

28

Table 1 Summary of *in silico* substrate docking results with models 3EN1M and 3EN1M-O_2 (O_2).

	Substrate	Fig.[a]	IM[b]	DM[c]	Iter.[d]	ΔG[e]	K_D[f]	His[g]	Val[h]	Gln[i]	1st C[j]	2nd C[k]
1a	**Toluene**	S2A		**1b**	crystal.				4.5		4.3	4.3
		S2B	**1b**	**1b**	100	**-4.61**	420.73		3.8		5.3	5.1
	Toluene (O_2)	S2C	**1b**	**1b**	100	**-4.94**	237.85		3.8		5.3	5.1
2a	**Phenol**	2A	**2b**	**2e**	60	**-4.42**	579.04	4.0	5.7		4.9	5.4
		2B	**2b**	**2e**	38	**-4.39**	601.70	3.0	5.6	1.8	4.6	4.4
		S3A	**2c**	**2d**	2	**-4.37**	624.15		4.8	3.7	5.1	5.2
	Phenol (O_2)	2C	**2b**	**2e**	55	**-4.73**	341.26	3.8	5.5		5.1	5.2
		2D	**2c**	**2d**	45	**-4.70**	359.55		4.8	2.1	5.1	5.2
3a	**o-Cresol**	3A	**3h**	**3e**	50	**-4.98**	223.06	3.0	5.2	1.8	4.3	4.6
		3B	**3b**	**3f**	24	**-4.94**	237.83	3.8	4.9		4.9	5.3
		S4A	-	-	26	**-4.97**	228.25		5.1	2.1	4.5	5.0
	o-Cresol (O_2)	3C	**3b**	**3f**	93	**-5.33**	124.68	3.7	4.6		5.0	5.3
		3D	-	-	7	**-5.22**	150.42	3.1	5.0		5.0	5.3
4a	**m-Cresol**	4A	**4b**	**4f**	66	**-4.89**	259.18	3.0	4.4	1.8	4.3	4.4
		4B	**4c**	**4d**	34	**-4.87**	271.32		4.4	2.0	5.0	5.1
	m-Cresol (O_2)	4C	**4c**	**4d**	100	**-5.23**	145.45		4.3	2.0	5.0	5.0
5a	**3-methoxyphenol**	S6A	-	-	22	**-4.79**	307.79		3.1	2.2	4.3	4.0
		5A	**5c**	**5d**	54	**-4.76**	323.34		3.4	2.1	5.2	5.3
		5B	**5b**	**5f**	12	**-4.66**	382.58	3.4	3.5	1.7	4.4	4.6
		S6B	-	-	12	**-4.71**	352.15	3.3	5.2	1.8	4.7	4.5
	3-methoxyphenol (O_2)	5C	-	-	46	**-5.20**	154.55		3.2		4.8	5.0
		5D	**5c**	**5d**	54	**-5.10**	182.80		3.5	2.0	5.1	5.2
6a	**3-phenylphenol**	6A	**6b**	**6f**	100	**-7.25**	4.88	3.1	4.2	2.1	4.3	4.9
	3-phenylphenol (O_2)	6B	**6b**	**6f**	97	**-7.25**	4.89	2.9	3.2	1.9	4.9	5.3
		6C	**6b'**	**6b'**	3	**7.01**	4.90		3.3		5.1	5.6
7a	**p-cresol**	7A	**7c**	**7d**	100	**-4.80**	302.87		3.4	2.0	4.9	5.1
	p-cresol (O_2)	7B	**7c**	**7d**	100	**-5.13**	172.59		3.4	2.0	4.9	5.1

[a] Figure

[b] Initial metabolite

[c] Derived metabolie

[d] .Orientation occurrence (iterations)

[e] Binding energy [kJ/mol]

[f] Dissociation constant KD [μM]

[g, h, i] Distances of interactions between substrate substituents (**1a-7a**) and (g) His-311, (h) Val-309 and (i) Gln-215 [Å]

[j, k] Distances of closest C=C-bond to Fe-atom [Å]