



HAL
open science

The Binding of Lignans, Flavonoids, Coumestrol to CYP450 Aromatase. A Molecular Modelling Study

Sampo Karkola, Kristiina Wähälä

► **To cite this version:**

Sampo Karkola, Kristiina Wähälä. The Binding of Lignans, Flavonoids, Coumestrol to CYP450 Aromatase. A Molecular Modelling Study. *Molecular and Cellular Endocrinology*, 2009, 301 (1-2), pp.235. 10.1016/j.mce.2008.10.003 . hal-00532094

HAL Id: hal-00532094

<https://hal.science/hal-00532094>

Submitted on 4 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Title: The Binding of Lignans, Flavonoids, Coumestrol to CYP450 Aromatase. A Molecular Modelling Study

Authors: Sampo Karkola, Kristiina Wähälä

PII: S0303-7207(08)00462-0
DOI: doi:10.1016/j.mce.2008.10.003
Reference: MCE 7021

To appear in: *Molecular and Cellular Endocrinology*

Received date: 20-9-2008
Revised date: 6-10-2008
Accepted date: 8-10-2008

Please cite this article as: Karkola, S., Wähälä, K., The Binding of Lignans, Flavonoids, Coumestrol to CYP450 Aromatase. A Molecular Modelling Study, *Molecular and Cellular Endocrinology* (2008), doi:10.1016/j.mce.2008.10.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



1

2 **The Binding of Lignans, Flavonoids and Coumestrol**
3 **to CYP450 Aromatase. A Molecular Modelling Study.**

4

5 Sampo Karkola and Kristiina Wähälä¹

6

7 Laboratory of Organic Chemistry, Department of Chemistry, Faculty of Science, P.

8 O. Box 55, FIN-00014 University of Helsinki, Finland

9

¹Corresponding author. Email kristiina.wahala@helsinki.fi. Tel. +358 919150356,

Fax. +358 919150357

1 **Abstract**

2

3 Androgens are transformed into aromatic estrogens by CYP450 aromatase in a three-
4 step reaction consuming three equivalents of oxygen and three equivalents of
5 NADPH. Estrogens are substrates for nuclear estrogen receptors and play a key role
6 in estrogen-dependent tumour cell formation and proliferation. Natural
7 phytoestrogens are proved to be competitive inhibitors of aromatase enzyme at IC₅₀
8 values in micromolar levels. In order to understand the mechanisms involved in the
9 binding of various phytoestrogens, we used our model of CYP450 aromatase to study
10 the binding of phytoestrogens using molecular dynamics simulations with a bound
11 phytoestrogen. The simulation trajectory was analysed to find the essential
12 interactions which take place upon binding and a representative structure of the
13 trajectory was minimized for docking studies. Sets of phytoestrogens, such as lignans,
14 flavonoids/isoflavonoids and coumestrol, were docked into the aromatase active site
15 and the binding modes were studied.

16

17 **Keywords**

18 CYP450, aromatase, phytoestrogen, molecular modelling, lignan, enterolactone,
19 flavone, isoflavone

20

1 **1. Introduction**

2

3 Aromatase (CYP450_{arom}) is an essential enzyme in estrogen biosynthesis converting
4 the aliphatic androgens testosterone and androstenedione to the aromatic estrogens
5 estradiol and estrone, respectively (Figure 1). Estrogens play a key role in normal cell
6 proliferation by binding to the nuclear estrogen receptor (ER) and triggering a
7 sequence of reactions leading to cell division. (Mosselman et al., 1996) Estrogens are
8 also a key factor in hormone-dependent (ER-positive) tumour development. (Muti et
9 al., 2006) One approach to treat and/or prevent hormone-dependent tumour
10 development is to decrease the level of circulating estrogens and local tumour
11 estrogen production by inhibiting estrogen producing enzymes. (Santen, 2002)

12 Phytoestrogens are a group of polyphenolic compounds found in many plants present
13 in human diet, e.g. soy, flaxseed, grains, berries, vegetables and fruits. (Adlercreutz,
14 1995; Mazura et al, 2000; Milder et al., 2004, 2005; Kiely et al., 2003) They can
15 mimic the size and physico-chemical properties of estrogens and thus can interact
16 with the same receptors and enzymes as endogenous estrogens controlling biological
17 functions. Several studies have shown that populations with a high intake of
18 phytoestrogens (such as lignans and isoflavones) in diet have lower risk for breast
19 cancer. (Lof and Weiderpass, 2006) A recent review describes the potency of
20 phytoestrogens as aromatase inhibitors. (Wähälä et al., 2008)

21 Dietary lignans include lignanolactones and lignanodiols (Figure 2) and they exist
22 mainly as glycosides in the nature. The most studied dietary lignans matairesinol
23 (MAT) and secoisolariciresinol (SI) are converted to mammalian lignans
24 enterolactone (ENL) and enterodiol (END), respectively, by various bacterial strains
25 in the gut microflora. (Borriello et al., 1995; Clavel et al., 2006; Wang et al., 2000) In
26 the lignanolactone series, MAT is converted to 4,4'-dihydroxyenterolactone (4,4'-
27 diOH-ENL) by demethylation and further to ENL by dehydroxylation in the gut.
28 (Borriello et al., 1995; Clavel et al., 2006; Raffaelli et al., 2006) 3-
29 Demethoxymatairesinol (DMM), 3,3'-didemethoxymatairesinol (DDMM) and 3-
30 demethoxy-3'-O-demethylmatairesinol (DMDM) are possible metabolites of MAT
31 and they are all precursors of ENL. In the lignanodiols series, SI is converted to 3'-

1 demethylsecoisolariciresinol (ODSI) and further to 3,3'-
2 didemethylsecoisolariciresinol (DDSI) and 3'-demethoxysecoisolariciresinol (DMSI).
3 (Wang et al., 2000) The last two are then metabolised to END via dehydroxylation.
4 Furthermore, END can be converted to ENL.
5 To our knowledge, the binding mode of lignans to the aromatase enzyme has not been
6 reported earlier. Inhibition studies have identified lignans that competitively inhibit
7 aromatase function with IC_{50} values at low micromolar levels in human placental
8 microsome based assays and at high micromolar levels in cell based assays.
9 (Adlercreutz et al., 1993; Wang et al., 1994) The most potent lignan in human
10 placental aromatase based assay was 4,4'-di-OH-ENL, a metabolite of MAT and a
11 precursor of ENL, inhibiting aromatase activity with an IC_{50} value of 6 μ M.
12 (Adlercreutz et al., 1993) Other active lignans in the same assay were the antioxidant
13 nordihydroguaiaretic acid (NDGA), ENL and its suggested precursor DDMM, with
14 IC_{50} values up to 15 μ M. Cell-based assays have also been performed to identify
15 lignans with inhibitory activity. (Adlercreutz et al., 1993; Wang et al., 1994; Brooks
16 and Thompson, 2005) In a human preadipocyte based assay, DDMM was shown to be
17 the most potent lignan tested (IC_{50} value of 60 μ M). (Wang et al., 1994) DDMM had
18 inhibitory activity comparable to ENL and DDMM. END and its precursors ODSI,
19 DDSI and DMSI were shown to be weak inhibitors of aromatase. (Wang et al., 1994)
20 Flavonoids and isoflavonoids have a rigid benzopyran-4-one ring system extended
21 with an aryl ring at C2 or C3 of the benzopyranone skeleton (Figure 2). Hydroxy
22 and/or methoxy groups are common substituents in dietary flavonoids/isoflavonoids.
23 The inhibitory activity of these types of phytoestrogens against aromatase has been
24 studied with assays utilizing human placental microsomes, recombinant aromatase
25 enzyme and various cell lines. (Wähälä et al., 2008) In general, flavonoids are the
26 most potent inhibitors of aromatase in different assays, with IC_{50} values ranging from
27 sub-micromolar to high-micromolar level. The most potent inhibitor is α -
28 naphthoflavone (ANF) with an IC_{50} value of 0.07 μ M. (Kellis and Vickery, 1984)
29 Isoflavonoids are weaker inhibitors, the best being biochanin A (5,7-dihydroxy-4'-
30 methoxyisoflavone). Flavonoid skeletons have been used as lead compounds in drug
31 discovery projects, producing nanomolar level inhibitors of aromatase. (Recanatini,

1 2001; Brueggemeier and Whetstone, 2001; Kim et al., 2004; Hackett et al., 2005;
2 Cavalli et al., 2005; Su et al., 2005; Gobbi et al., 2006)

3 The inhibitor profiles and the binding modes of flavonoids, isoflavonoids and
4 coumestrol (Figure 2) have been studied using site-directed mutagenesis studies and
5 molecular modelling (Chen et al., 1997; Kao et al., 1998; Hong et al., 2008), where
6 homology models of aromatase (Laughton et al., 1993; Kao et al., 1996) were used.
7 Based on these studies, the resulting hypothesis was that the A and C rings of the
8 flavonoids mimic the C and D rings of the steroid core (Figures 1 and 2). It has been
9 suggested that the flavonoids interact with the heme iron *via* the carbonyl oxygen at
10 C4. (Kellis and Vickery 1984; Kao et al., 1998; Sanderson et al., 2004) Experimental
11 inhibitory data also provide evidence for the importance of the carbonyl group in
12 phytoestrogen binding. The most potent natural (Stermitz et al., 2003) flavonoid
13 tested to date, ANF, lost its binding affinity when the six-membered pyran-4-one ring
14 was replaced with a furan ring. (Kellis et al., 1986) Additionally, the reduction of a
15 flavone or a flavanone to a 4-flavanol reduced the activity dramatically (Ibrahim and
16 Abul-Hajj, 1990) and catechins, which lack the C4 carbonyl function, did not inhibit
17 aromatase at a detectable level. (Sanderson et al., 2004) It has been shown that the
18 coordination of oxygen to the heme iron produces a different type of change in
19 absorption spectrum compared to nitrogen or sulphur coordination. (Dawson et al.,
20 1982; White and Coon, 1982) The difference spectra measured for ANF, its
21 metabolite 9-hydroxy- α -naphthoflavone (9-OH-ANF) and ENL indicate that the
22 phytoestrogen replaces the natural substrate ASD from the active site and converts a
23 high-spin iron into a low-spin iron producing a spectrum change typical for oxygen
24 coordination to the iron. (Adlercreutz et al., 1993; Kellis and Vickery, 1984; Kellis et
25 al., 1986; Kellis and Vickery, 1987) Additionally, two crystal structures of CYP450
26 enzymes show coordination from an oxygen to the heme iron (PDB entry codes 2CIX
27 and 1NOO). (Kuhnel et al., 2006; Li et al., 1995)

28

29 In this study we provide an atomic level explanation for the binding of lignans,
30 flavonoid/isoflavonoids and coumestrol to the aromatase active site. This is the first
31 study to report the binding mode of phytoestrogens to the aromatase enzyme using
32 molecular dynamics simulations (MDS) and ligand-protein docking. The emphasis

1 was given to lignans due to the lack of studies of the binding mode, although lignans
2 have been shown to inhibit the aromatase enzyme function. (Adlercreutz et al., 1998;
3 Wang et al., 1994; Brooks and Thompson, 2005) We performed molecular dynamics
4 simulations for the lignan(ENL)-aromatase and flavonoid(ANF)-aromatase
5 complexes using our aromatase model (Karkola et al., 2007), which explains the
6 majority of the results from mutagenesis studies, to mimic the dynamic process of
7 ligand binding. The trajectories of the two complexes were analyzed, and sets of
8 lignans, flavones/isoflavones and coumestrol were docked into the structures obtained
9 from the molecular dynamics simulations to find the plausible binding modes of these
10 compounds. The compounds for the study were selected based on the structural
11 variability and the inhibition potency.

12

13 **2. Materials and methods**

14

15 Molecular dynamics simulations (MDS) were performed with Gromacs version 3.3.1
16 (Van Der Spoel, 2005) and dockings with GOLD version 3.1.1. (Jones et al., 1997)
17 All calculations were done with an IBM-compatible desktop workstation or a HP
18 ProLiant DL145 supercluster. For the dockings, the all-atom protein structures were
19 generated with Insight II. (Accelrys Inc., 2005) Ligands were sketched and minimised
20 with Sybyl version 7.3. (Tripos Inc.) Graphics were generated with PyMOL software.
21 (DeLano, 2002)

22

23 **2.1 Aromatase model refinement**

24

25 The active site of an enzyme is represented as a rigid or a semi-rigid cavity in docking
26 programs, and we therefore used MDS to mimic the dynamic process of ligand
27 binding. In MDS, the system (protein with or without a bound ligand) is placed into a
28 box of water (and ions). Subsequent minimisation produces a relaxed starting
29 structure, for which the initial atom velocities at a certain temperature are calculated.
30 A small change in time is applied and new atomic positions are calculated for the new

1 time step. The process is repeated until energy is conserved and e.g. the protein
2 backbone movement is stabilised. The calculation of atomic movements as a function
3 of time produces an ensemble of conformations (a trajectory) of the system. From this
4 trajectory, the properties of the system and the representative conformation can be
5 analysed.

6

7 The lignan skeleton is more flexible than the rigid flavonoid/isoflavonoid structure,
8 but can also mimic the functions of estrogens and bind to the estrogen receptor and
9 aromatase enzyme (Figure 2). (Adlercreutz et al., 1993; Wang et al., 1994; Mueller et
10 al., 2004) Due to this flexibility, we generated two complexes; one with the lignan
11 ENL and one with the flavonoid ANF. ENL and ANF were manually inserted into the
12 active site of the aromatase model. The orientation of the ligands was optimised for
13 the carbonyl oxygen – heme iron coordination to take place, as suggested by Kellis et
14 al., Kao et al. and Sanderson et al. (Kellis and Vickery, 1984; Kao et al., 1998;
15 Sanderson et al., 2004)

16 Due to the lack of published parameterisation for the coordination between carbonyl
17 oxygen and iron, we searched the PDB database for heme-containing proteins with
18 oxygen coordinated to the iron using the public Relibase server. (Hendlich, 1998) In
19 the search query, the coordinating oxygen was defined to have a double bond to the
20 adjacent heavy atom in the ligand. The distance from the iron to the oxygen was
21 defined to be 0.05—0.35 nm and the angle (iron-oxygen-heavy atom) 90—180°. The
22 search resulted in 16 complexes where the coordinating oxygen was double bonded to
23 a carbon, nitrogen or sulphur. The distance distribution among the 16 complexes was
24 0.2–0.35 nm and the angle distribution was 100–150°. The Relibase results also
25 showed 162 complexes having a diatomic oxygen molecule bound to the heme iron.
26 All complexes considered together, the average distance from the coordinating
27 oxygen to the iron was 0.19 nm and the average angle was 131°. Based on this data,
28 the GROMOS96 force field (ffG43a1) in Gromacs 3.3.1 was modified to include a
29 Morse potential between the iron and the oxygen with a distance of 0.19 nm. The
30 application of a Morse potential was successfully used earlier in MDS describing the
31 octahedral coordination between the heme iron and the fifth (thiolate) and sixth (basic
32 nitrogen) ligand. (Rupp et al., 2005) We used the same well-depth and steepness

1 parameters for the iron-oxygen coordination as was used for the iron-sulphur
2 coordination. (Rupp et al., 2005) The angle Fe-N-C of the protoporphyrin ring
3 (already defined in the GROMACS force field as 125°) was used for the Fe-O=C
4 angle. Truncated octahedron simulation boxes were generated for the two aromatase
5 complexes (with ANF and ENL). Water and ions were added to neutralise the system
6 and to represent physiological saline. Both complexes consisted of ~9400 water
7 molecules and 32 Na⁺ and 33 Cl⁻ ions with a total volume of ~360 nm³. The
8 simulation systems were minimised to acquire a relaxed starting structure for MDS.
9 The physiological temperature of 310 K was applied to the systems along with
10 periodic boundary conditions and particle-mesh Ewald method for long-range
11 electrostatics. Hydrogen containing bonds were converted to constraints with LINCS
12 algorithm and temperature and pressure couplings were treated with Berendsen
13 algorithms. Time step used was 2 fs. First, 1000 ps equilibration simulations were
14 performed with position restraints (1000 kJ/mol restrain force) applied to the
15 backbone of the protein, the heme and the ligand. The purpose was to let the
16 sidechains adjust to the surroundings and to soak the simulation systems.
17 Subsequently, an unrestrained simulation was performed for both complexes until the
18 systems stabilised. The stable parts of the trajectories were clustered and
19 representative structures from the biggest clusters were extracted and minimised *in*
20 *vacuo* in Gromacs to produce structures for the docking studies.

21

22 **2.2 The dockings of the flavonoid/isoflavonoid and lignan sets**

23

24 GOLD was selected for the docking studies, because it was able to reproduce the
25 exact or very close ligand conformations of four cytochrome enzyme – ligand
26 complexes extracted from the PDB (PDB ids 1suo, 1z11, 1w0g and 2j0c, data not
27 shown). GOLD uses the so-called Genetic Algorithm (GA), where, initially, a number
28 of possible solutions (i.e. conformations) of the ligand and (partially) the active site
29 are created. During the docking procedure (“evolution”), these solutions
30 (“chromosomes”) are combined and mutated to produce offspring that have better

1 qualities than their parents. Finally, the “fittest” survive and provide a solution to the
2 docking problem.

3 The all-atom enzyme structures were prepared for GOLD by extracting the ligand
4 from the active site, adding hydrogens to the protein and minimising the structure
5 with fixed heavy atoms for 500 steps using steepest descent method in Discovery
6 Studio 1.7. The phytoestrogen structures were sketched and minimized in Sybyl 7.3.

7 GOLD can be used with modified metal coordination parameters, derived from a
8 database search from PDB and Cambridge Structural Database (CSD) for a direct
9 coordination between carbonyl oxygen and an iron. (Kirton et al., 2005) Due to the
10 lack of such complexes in the databases (at that time) the authors created additional
11 metal coordination parameters for GOLD, where carbonyl oxygen is not preferred as
12 a coordinative counterpart for an iron. (Kirton et al., 2005) However, as mentioned
13 earlier, experimental data suggest carbonyl oxygen – iron interaction (see
14 Introduction), and therefore these modified parameters were not used in this study.

15 Goldscore scoring function was used for the flavonoids/isoflavonoids and Chemscore
16 for the lignans, based on the reproduction of the ligand orientations from the MDS.
17 Ring corners of the ligands were allowed to flip and internal hydrogen bonds were
18 taken into account. Positional or distance restraints were not used during the
19 dockings. The lignans were docked into a minimised all-atom aromatase structure
20 from the ENL-aromatase-trajectory, and the active site was defined as a 15 Å sphere
21 around a side chain hydrogen atom of Thr310 with cavity detection enabled. In the
22 case of the flavonoids, the structures were docked into the minimised all-atom
23 aromatase structure from the ANF-aromatase-trajectory and the active site was
24 defined as a 15 Å sphere around a side chain hydrogen atom of His480 with cavity
25 detection enabled. The centre atoms in defining the active sites are different because
26 an atom close to the active site centre was chosen in both structures, and these atoms
27 are different due to the adaptation of the active site residues to the ligand. Coumestrol
28 has rigid skeleton and similar spatial and physico-chemical properties as flavonoids
29 and was therefore docked to the aromatase – ANF complex structure with the same
30 parameters. The fitness parameters were left at default values in all dockings. The
31 compounds used in this study are listed in Tables 1 and 2.

32

1 **3. Results and discussion**

2
3 We have studied the stability of enzyme complexes between aromatase and a lignan
4 or a flavone using molecular dynamics simulations. The enzyme structures obtained
5 from the MDS were used to dock sets of aromatase inhibiting lignans,
6 flavonoids/isoflavonoids and coumestrol into the active sites of the enzyme and the
7 interactions between the ligands and the enzyme were investigated.
8

9 **3.1 Molecular dynamics simulations of the aromatase-ENL-complex**

10
11 The aromatase-ENL-complex was simulated for a total time of ~15 ns, including the
12 restrained equilibration in the beginning (see Materials and Methods). The enzyme
13 backbone root mean square deviation (rmsd) stabilised after 12 ns (Figure 3). The
14 ENL rmsd stabilised soon after an increase at the very beginning of the unrestrained
15 simulation. Only an occasional hydrogen bond from ENL to the backbone carbonyl of
16 Glu302 was observed during the simulation and only a few water molecules are
17 present occasionally inside the active site cavity, emphasising the lipophilic nature of
18 the active site. The potential energy of the system stabilised after a decrease in the
19 beginning. The coordination from the ligand carbonyl oxygen to the heme iron
20 showed an average distance of 0.19 nm and an average angle of 128° with a deviation
21 of 5.4°. In the representative structure of the trajectory the active site residues in
22 contact with ENL (within 6 Å) are Ile133, Phe134, Ala226, Leu227, Ile229, Lys230,
23 Tyr244, Glu302, Ile305, Ala306, Ala307, Pro308, Asp309, Thr310, Met311, Val313,
24 Ser314, Phe430, Cys437, Ala438, Gly439, Lys440, Leu479, His480 and Pro481, in
25 addition to the heme. The active site rmsd between the representative structure of the
26 trajectory and the initial structure was 2.7 Å.
27

28 **3.2 Molecular dynamics simulations of the aromatase-ANF-complex**

29

1 The aromatase-ANF-complex required a relatively long simulation time (20 ns) for
2 the system to stabilise. The backbone rmsd was stable after 17 ns until the end of the
3 simulation (Figure 4). The rmsd for ANF shows fluctuation due to the rotation of the
4 phenyl group at C2. The rotation causes the rmsd to fluctuate even though the
5 chemical and physical properties of the rotating phenyl group remain the same. This
6 is due to the method of calculating rmsd, where each atom is individually compared to
7 its starting position. The potential energy of the complex stabilised after a decrease in
8 the beginning. During the dynamics there were only a few occasional water molecules
9 providing hydrogen bonding counterparts to the active site residues Glu302, Lys230
10 and His480. The interaction between the carbonyl oxygen and the heme iron shows an
11 average distance of 0.19 nm and an average angle of 121° with a deviation of 4.6° .
12 The representative structure of the stable trajectory shows that the active site residues
13 near ANF (within 6 Å) are Leu122, Ile132, Ile133, Phe134, Phe148, Lys230, Cys299,
14 Glu302, Met303, Ala306, Ala307, Thr310, Met374, Phe430, Cys437, Ala438,
15 Gly439, Leu477, Ser478, Leu479 and His480 in addition to the heme. The active site
16 rmsd between the representative structure of the trajectory and the initial structure
17 was 2.8 Å.

18 **3.3 Docking of phytoestrogens**

19
20 To clarify the structural features contributing to the interactions during the binding
21 and to the coordination geometry, lignans (10 compounds), flavonoids (18
22 compounds) and coumestrol were docked into the active sites of the ENL- and
23 ANF-complexes. The ligand sets included the compounds used in the MDS.
24 Compounds used in this study are listed in Tables 1 and 2 with their inhibition values.
25

26 **3.3.1 The binding modes of the lignans**

27
28 The orientation of ENL from MDS was reproduced except that the aromatic rings
29 were rotated 180° . This is due to the difference in calculating the optimal orientation
30 between Gromacs and Goldscore scoring function. A similar orientation with the

1 MDS was observed in one of the docked solutions, but with a lower score than the
2 presented one. We decided to follow the ranking from Goldscore (and use the highest
3 scoring orientation with every compound) to make the results between the compounds
4 comparable. Among the 10 docked lignans there are five compounds with a lactone
5 structure (ENL, DMDM, DDMM, DMM and 4,4'-diOH-ENL), and therefore the
6 possibility to coordinate to the heme iron *via* the carbonyl oxygen. Based on the
7 docking results, all these five compounds exhibit this interaction. Optimal
8 coordination position and orientation would be achieved if the plane defined by the
9 lactone group was perpendicular to the heme plane. However, based on the docking, a
10 40° tilting from the optimum 90° was observed. The distance from the iron to the
11 coordinating oxygen was 2.4 Å, and the angle Fe—O=C was 129° (optimal 120°). It
12 has been shown by density functional theory (DFT) calculations that the strength of
13 the coordination to the heme iron is highly dependent on the tilting and torsion angles
14 of the coordinating system. (Rupp et al., 2005) Therefore, even a small change from
15 the optimal coordinating geometry can lead to weaker binding. Based on the docked
16 ENL, the plane tilting, Fe—O=C angle and the Fe—O distance together result in
17 weakened coordination and to poorer inhibition compared to ANF or coumestrol. The
18 same effect can be seen in a recent crystal structure of a chloroperoxidase enzyme
19 (PDB entry code 2CIX) (Kuhnel et al., 2006), which shows a cyclopentadione ligand
20 coordinated to the heme iron via the carbonyl oxygen. NMR titration experiments
21 (Wang and Hoff, 1997) have produced a high dissociation constant (K_d 33 mM) for
22 the cyclopentadione binding to the chloroperoxidase enzyme, indicating a very weak
23 interaction despite the coordination between the oxygen and the iron. Together these
24 studies (Kuhnel et al., 2006; Wang and Hoff, 1997) provide an explanation for the
25 micromolar level IC_{50} values measured for the lignans despite the probable oxygen –
26 iron coordination. The increased potency observed with commercial third generation
27 aromatase inhibitors, bearing nitrogen as the coordinating atom, is probably due to a
28 better geometry of the coordinative system.

29 The best ranked pose (Figure 5) for ENL shows, in addition to the coordination to the
30 heme iron, hydrogen bonds between the phenolic hydroxy group at C3 (ring A, Figure
31 2) and the backbone carbonyl of Thr310, and the hydroxy group at C3' (ring B,
32 Figure 2) and the backbone carbonyl of Glu302. All compounds with a phenolic

1 hydroxyl at C4 form a hydrogen bond to the backbone carbonyl of Pro481. Similarly,
2 compounds with a phenolic hydroxyl at C3 form a hydrogen bond to the backbone
3 carbonyl of Thr310. More variation in the hydrogen bonding pattern was observed
4 when the substituents of ring B were studied. The highest activity observed
5 experimentally for 4,4'-diOH-ENL is probably due to an additional hydrogen bond at
6 ring A (from the hydroxyl at C3 to the backbone carbonyl of Thr310) and possibly, a
7 water-mediated hydrogen bond from ring B hydroxyl to Lys230 or Ile305. DDMM is
8 almost superimposable with 4,4'-diOH-ENL, but lacks the additional hydrogen bond
9 from ring A to Thr310 and is therefore less active. Ring B of DMDM is shifted
10 compared to 4,4'-diOH-ENL, but forms a hydrogen bond to the backbone carbonyl of
11 Glu302. Ring B of the less active DMM is shifted compared to 4,4'-di-OH-ENL and
12 forms a hydrogen bond from the hydroxyl at C3' to the backbone carbonyl of Ile305.
13 The methoxy group at C3' might block a water-mediated hydrogen bond and lead to a
14 decreased activity.

15 The best inhibitor lacking a lactone structure was the naturally occurring antioxidant
16 NDGA. Based on the docking results, the heme-coordinating atom is one of the
17 catecholic hydroxy oxygens, which is not internally hydrogen bonded to the adjacent
18 phenolic hydroxy group. The high activity of NDGA is probably due to a tight
19 hydrogen bonding pattern between the catecholic hydroxy groups and the side chains
20 of Lys230 and Asp309 and the backbone carbonyl of Ile305. Moreover, the branched
21 hydrocarbon chain is complementary in shape to the active site's hydrophobic pocket.
22 END coordinates to the heme iron from one of the aliphatic hydroxy groups. The
23 docking results indicate that the weaker binding affinity is due to an internal hydrogen
24 bond from the coordinating oxygen to the other aliphatic hydroxy group. DMSI has a
25 similar binding mode to END, but neither of the aliphatic hydroxyls is oriented to
26 have a proper coordination to the iron. ODSI and DDSI also coordinate to the heme
27 iron from one of the phenolic oxygens. However, next to the iron-coordinating
28 phenolic hydroxyl of ODSI, a methoxy group lies very close to the heme ring and
29 probably weakens the coordination strength. Additionally, both compounds have
30 aliphatic hydroxy groups in the area of hydrophobic residues of the active site. The
31 docked poses of the lignans are displayed in Figure 6a.

32

1 3.3.2 The binding modes of the flavonoids

2

3 As a result from the dockings, the orientation of ANF from MDS was reproduced.
4 ANF coordinates to the heme iron via C4 carbonyl (Figure 7). The naphthyl moiety of
5 ANF is located near residues Leu122, Ile133, Phe134, Lys230, Val373, Met374,
6 Ser478 and Leu479 and the phenyl B-ring near residues Ile132, Phe148 and Cys299.
7 The hydrophobic skeleton of ANF is therefore accommodated by two hydrophobic
8 areas in the active site. A mutagenesis study involving seven phytoestrogens was
9 performed to study the role of certain aromatase residues. (Kao et al., 1998) Mutant
10 I133Y² significantly enhanced the affinity towards ANF, chrysin, galangin, genistein
11 and biochanin A, but the effect towards 7,8-di-OH-flavone was just the opposite.
12 Based on our dockings, a tyrosine at 133 probably leads to favourable π - π interactions
13 between the aromatic tyrosine ring and the B-ring of the flavonoid, especially with the
14 induced-fit binding mode. The negative effect observed with 7,8-di-OH-flavone could
15 result from the disruption of the perpendicular π - π interactions by the catechol
16 hydroxyls. Other mutations reported in the study (Kao et al., 1998) were I395F,
17 I474Y, P308F, D309A and T310S. I395F and I474Y are located outside the active
18 site cavity and, therefore, their influence to the binding affinity is difficult to predict.
19 Pro308 forms the kink in the I-helix and does not contribute to the active site. The
20 effect of losing the kink in P308F is not clear, and the K_i values differ not more than
21 4-fold from the values for the wild type enzyme for all tested compounds. (Kao et al.,
22 1998) Although Asp309 is located in the I-helix, it points “upwards” from the heme in
23 our model and is not directly involved in inhibitor binding. The effect of D309A
24 could be seen as a disruption of hydrogen bonding network. T310S decreased the
25 affinity to ANF and biochanin A and increased the affinity to chrysin, 7,8-di-OH-
26 flavone, galangin and genistein. (Kao et al., 1998) The side chain hydroxyl of Thr310
27 probably has a dual role in the active site. It serves as a hydrogen bonding partner for
28 the water molecules that are needed in the catalytic reaction and, on the other hand, it
29 stabilises the kink (formed by Pro308) in the I-helix by hydrogen bonding to the
30 carbonyl group of Ala306. Based on our model, losing one methyl group from the
31 side chain of a threonin at 310 gives His480 the possibility to turn towards the serine

1 at 310 and form a hydrogen bond with it. This would lead to a small increase in the
2 active site volume and to a better orientation of the carbonyl group in chrysin, 7,8-di-
3 OH-flavone, galangin and genistein, in respect to the iron. ANF, due to its bulky
4 naphthyl moiety is anchored to its place more tightly and has less room for
5 reorientation. Biochanin A, incapable of a good coordination, loses its parallel π - π
6 interaction with His480. Similar inhibition profiles were observed in another study for
7 chrysin and 7,8-di-OH-flavone. (Chen et al., 1997) The docking poses of the flavones
8 are displayed in Figure 6b.

9

10 The difference in coordination strength between flavones and isoflavones was clearly
11 observed comparing the docking results of unsubstituted flavone and isoflavone. The
12 steric effects of ring B attached to C2 in isoflavone shifts the carbonyl to an
13 unfavourable position and angle compared to the carbonyl in flavone (Figure 8). The
14 distance from the carbonyl oxygen to the heme iron was 2.0 Å in flavone complex
15 and 2.2 Å in isoflavone complex and the angle Fe-O=C was 136° in flavone complex
16 and 160° in isoflavone complex. The average values from Relibase search were 1.9 Å
17 for the distance and 131° for the angle. The optimal theoretical angle is 120°. The
18 tilting between the carbonyl and the heme planes was 27° for flavone and 30° for
19 isoflavone from the optimum 90°. Therefore, the coordination of isoflavone is weaker
20 compared to flavone, leading to a decreased activity.

21

22 A hydroxy group at C7 of the flavone core increases the inhibitory potency
23 significantly. This can be seen when comparing the IC₅₀ values of flavone (from 8
24 µM to 48 µM) and 7-hydroxyflavone (from 0.21 µM to 4 µM). (Sanderson et al.,
25 2004; Ibrahim and Abul-Hajj, 1990; Le Bail et al., 1998) Based on the dockings, the
26 positive effect is due to a hydrogen bond from C7 hydroxyl to Ser478. Mutations at
27 Ser478 have been performed but only to study the effect of steroidal and non-steroidal
28 inhibitors and coumestrol (see below). (Hong et al., 2008; Kao et al., 2001) A
29 methoxy group at C7 is less active than a hydroxy group (Le Bail et al., 1998), due to
30 the loss of hydrogen bonding capacity. Hydroxy groups at C3, C5 and C6 lead to
31 decreased inhibitory activity. This can be clearly seen from the IC₅₀ values of 3-

² Mutants are described with one letter codes of amino acid residues for clarity.

1 hydroxyflavone (140 μM), 5-hydroxyflavone (100 μM) and 6-hydroxyflavone (80
2 μM) compared to flavone (10 μM). (Ibrahim and Abul-Hajj, 1990) 3- and 5-
3 hydroxyflavones can form an intramolecular hydrogen bond to the carbonyl oxygen at
4 C4 and weaken the coordination capacity of the carbonyl. (Kao et al., 1998; Stermitz
5 et al., 2003) However, such a dramatic effect was not observed when comparing 7-
6 hydroxyflavone (IC_{50} value 0.21 μM) and chrysin (5,7-dihydroxyflavone, IC_{50} value
7 0.7 μM) (Le Bail et al., 1998), indicating that the hydroxyl at C7 is an important
8 factor to the binding affinity. The hydroxyl at C6 is too far to form a hydrogen bond
9 to Ser478 and is therefore less active. The importance of the hydroxy group at C7 was
10 also shown in a recent inhibition and molecular modelling study, where a 3D-QSAR
11 model was generated. (Neves et al., 2007) However, the results of the inhibition
12 studies vary from the previous ones and the coordination to the heme iron was not
13 mentioned.

14 A hydroxy group at ring B also leads to decreased activity. The effect can be seen
15 comparing the IC_{50} values of 7,4'-dihydroxyflavone (IC_{50} value 2 μM) and 7-
16 hydroxyflavone (0.5 μM) (Ibrahim and Abul-Hajj, 1990) or apigenin (5,7,4'-
17 trihydroxyflavone, 2.9 μM) and chrysin (5,7-dihydroxyflavone, 0.7 μM). (Le Bail et
18 al., 1998) Based on the dockings, this is probably due to unfavourable interactions of
19 a B-ring hydroxy group in a hydrophobic environment. The C4' hydroxyl is facing
20 the hydrophobic residues Ile132, Phe148 and Cys299. It can be concluded that a
21 hydrophobic skeleton, a sterically unhindered carbonyl functionality and a hydrogen
22 bonding substituent at C7, and not in C4', of the flavone skeleton are necessary
23 factors for effective binding.

24 3.3.3 The binding mode of coumestrol

25

26 Coumestrol was docked to the structure obtained from the ANF – aromatase
27 trajectory. The IC_{50} value for coumestrol in a cell-based assay was 17 μM and K_i
28 value was 1.3 μM making it a better inhibitor than any of the lignans tested in the
29 same assay. (Wang et al., 1994) Based on our dockings, coumestrol is located in the
30 active site with coordination to the heme iron. A 23° tilting from the optimum 90°
31 was observed for the coordinating carbonyl group, resulting in better coordination

1 compared to the lignans (40° tilting). Mutagens S478T and H480Q showed
2 significantly decreased affinity towards coumestrol compared to the wild type enzyme
3 and mutants E302D, D309A and T310S. (Hong et al., 2008) Our docking results show
4 hydrogen bonds from Ser478 side chain hydroxyl and Cys299 backbone carbonyl to
5 the phenolic hydroxy groups in coumestrol (Figure 9). The mutation of a serine at 478
6 to a threonine (Hong et al., 2008) probably alters the position and orientation of the
7 hydroxy group and leads to the loss of hydrogen bond. His480 is located near the
8 aromatic skeleton of coumestrol and the mutation to a glutamine leads to the loss of
9 π - π interactions and to a decrease of inhibition potency. E302D is a conservative
10 mutation and T310S probably leads to similar effect as observed with flavones (see
11 3.3.2). As mentioned earlier, Asp309 is not directly involved in phytoestrogen binding
12 and therefore the mutation to an alanine does not directly affect the binding of
13 coumestrol. Further evidence for the coordination of coumestrol to the iron would
14 have been obtained from the absorbance spectrum of the binding, but the spectrum
15 was not reported. (Hong et al., 2008) Our docking results provide a plausible binding
16 mode and a good explanation for the inhibition profiles observed in mutagenesis
17 studies. (Wang et al., 1994; Hong et al., 2008)

18

19 **Conclusions**

20

21 We have studied the dynamic process of binding and the binding modes of lignans,
22 flavonoids/isoflavonoids and coumestrol into aromatase enzyme using molecular
23 dynamics simulations and ligand–protein docking. Based on the hypothesis for
24 oxygen – iron coordination, the differences in the binding affinity towards various
25 phytoestrogens were clarified based on their binding modes. It was found that the
26 inhibition potency is dependent on the Fe-O distance, Fe-O=C angle and the tilting
27 angle between the planar carbonyl group of phytoestrogens and the cofactor heme
28 plane.

29 This is the first study to provide an atomic level explanation of the binding of lignans
30 to the aromatase enzyme. It was found that the lactone ring is an important but not the
31 only structural requirement to evoke aromatase inhibition. In lignanodiols, as seen in

1 the case of NDGA, the coordination can take place from a phenolic hydroxy group, as
2 long as a tight hydrogen bonding network to Lys230, Asp309 and Ile305, and
3 favourable hydrophobic interactions are achieved.

4 The hydroxy group at C7 of the flavone skeleton was found to interact with Ser478
5 and thus be important for the binding. Further, the substitutions at ring B usually lead
6 to decreased activity. It was found that this decreased activity is due to unfavourable
7 interactions with the hydrophobic active site residues.

8 In the case of the rigid and planar coumestrol, we have shown that mutations of the
9 serine at 478 and the histidine at 480 lead to the loss of hydrogen bond and π - π
10 interactions and therefore to a decreased affinity.

11 Together these findings provide valuable information on the binding process of
12 phytoestrogens to the aromatase active site and reveal the structural requirements
13 needed for lignan binding. This will create good possibilities in drug design research
14 concerning aromatase inhibition.

15

1 **Acknowledgements**

2

3 The Finnish IT Center for Science (CSC) is acknowledged for computer time
4 allocation. The Finnish Academy is acknowledged for funding, grants 78226, 78253,
5 210633. Mrs. Eija Leppälä and Dr. Petri Heinonen are thanked for advice.

6

Accepted Manuscript

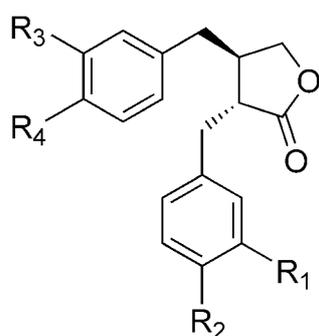
1 **Tables**

2

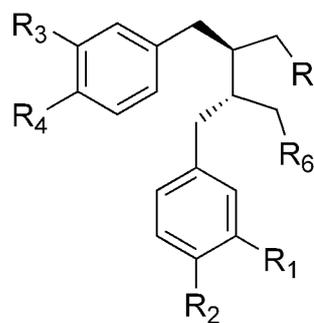
3

Table 1. Lignan structures used in the study and their aromatase inhibition values.

4



Lignan lactones



Lignan diols

Lignan lactones	R ¹	R ²	R ³	R ⁴	IC ₅₀ (μM)	K _i (μM)	Reference	
MAT	OMe	OH	OMe	OH	n/a	n/a	(Adlercreutz et al., 1993; Wang et al., 1994)	
ENL	OH	H	OH	H	11, 74*	14.4*	(Adlercreutz et al., 1993)	
4,4'-diOH-ENL	OH	OH	OH	OH	6		(Adlercreutz et al., 1993)	
DDMM	H	OH	H	OH	15, 60*	7.3*	(Adlercreutz et al., 1993; Wang et al., 1994)	
DMDM	H	OH	OH	OH	84*	5.0*	(Adlercreutz et al., 1993; Wang et al., 1994)	
DMM	H	OH	OMe	OH	37		(Adlercreutz et al., 1993)	
Lignan diols	R ¹	R ²	R ³	R ⁴	R ⁵	R ⁶	Reference	
SI	OMe	OH	OMe	OH	OH	OH	n/a	n/a
NDGA	OH	OH	OH	OH	H	H	11	(Adlercreutz et al., 1993)
END	OH	H	OH	H	OH	OH	30, >100*	(Adlercreutz et al., 1993; Wang et al., 1994)
ODSI	OMe	OH	OH	OH	OH	OH	>100*	(Wang et al., 1994)
DDSI	OH	OH	OH	OH	OH	OH	>100*	Wang et al., 1994)
DMSI	OMe	OH	H	OH	OH	OH	>100*	(Wang et al., 1994)

* Cell-based assay

MAT=matairesinol; ENL=enterolactone; 4,4'-diOH-ENL=4,4'-dihydroxyenterolactone; DDMM=didemethoxymatairesinol; DMDM=3-demethoxy-3'-O-demethylmatairesinol; DMM=3-demethoxymatairesinol; SI=secoisolariciresinol; NDGA=nordihydroguaiaretic acid; END=enterodiol; ODSI=3'-demethylsecoisolariciresinol; DDSI=3,3'-didemethylsecoisolariciresinol; DMSI=3'-demethoxysecoisolariciresinol

5

6

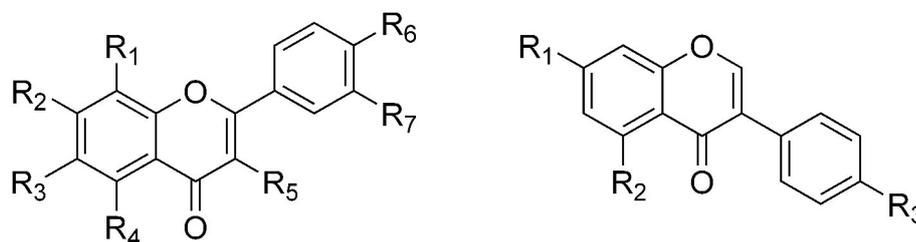
7

8

9

10

1 **Table 2.** Flavonoid and isoflavonoid structures used in the study and their aromatase
 2 inhibition values.



Flavones

Isoflavones

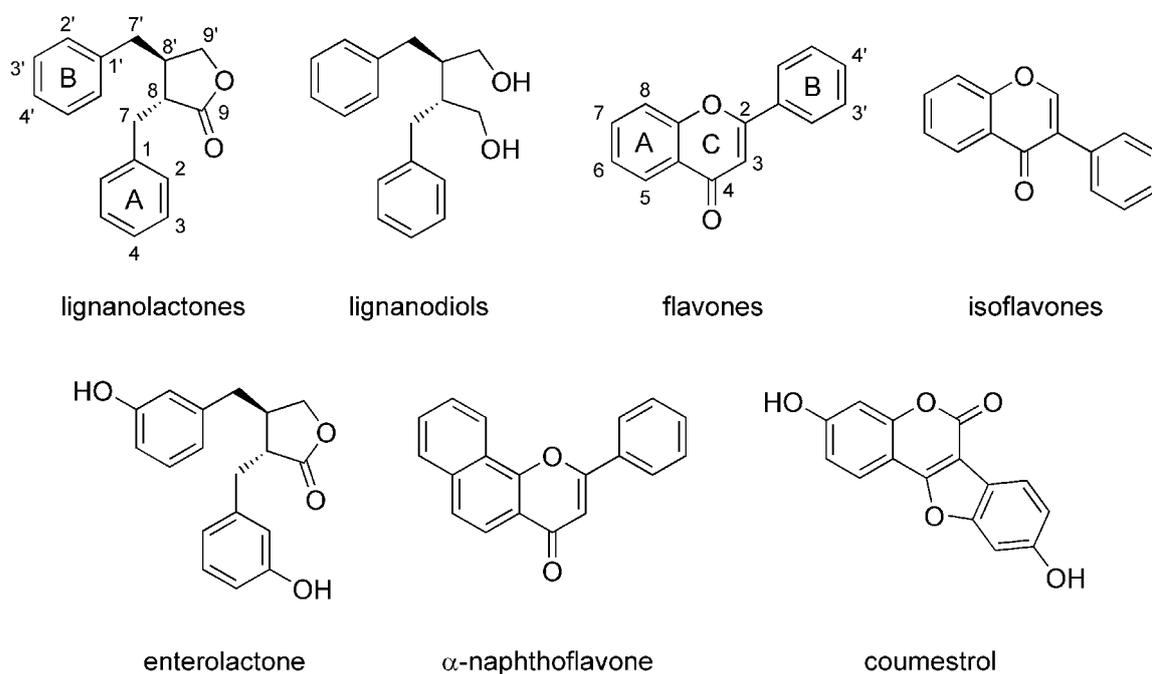
Flavones	R ¹	R ²	R ³	R ⁴	R ⁵	R ⁶	R ⁷	IC ₅₀ (μ M)	K _i (μ M)	Reference
α -naphthoflavone		‡	H	H	H	H	H	0.07		(Kellis and Vickery, 1984)
								0.18		(Stresser et al., 2000)
									0.2	(Campbell and Kurzer, 1993)
									2.2	(Kao et al., 1998)
7-OH-flavone	H	OH	H	H	H	H	H	0.21		(Le Bail et al., 1998)
								0.35		(Saarinen et al., 2001)
								0.5		(Ibrahim and Abul-Hajj, 1990)
								4		(Sanderson et al., 2004)
chrysin	H	OH	H	OH	H	H	H	0.5		(Kellis and Vickery, 1984; 56]
								0.7		(Le Bail et al., 1998; Stresser et al., 2000)
								4.3		(Jeong et al., 1999)
								7		(Sanderson et al., 2004)
									2.6	(Kao et al., 1998)
									1	(Edmunds et al., 2005)
									2.4	(Campbell and Kurzer, 1993)
7,4'-di-OH-flavone	H	OH	H	H	H	OH	H	2		(Ibrahim and Abul-Hajj, 1990)
apigenin	H	OH	H	OH	H	OH	H	0.18		(Saarinen et al., 2001)
								1.2		(Kellis and Vickery, 1984)

								2.9	(Le Bail et al., 1998)
								3.3	(Le Bail et al., 1998)
								20	(Sanderson et al., 2004)
7-OMe-flavone	H	OMe	H	H	H	H	H	3.2	(Le Bail et al., 1998)
7,8-di-OH-flavone	OH	OH	H	H	H	H	H	8	(Le Bail et al., 1998)
								8.7	(Jeong et al., 1999)
								10	(Kao et al., 1998)
flavone	H	H	H	H	H	H	H	8	(Kellis and Vickery, 1984)
								10	(Ibrahim and Abul-Hajj, 1990)
								48	(Le Bail et al., 1998)
								22	(Campbell and Kurzer, 1993)
galangin	H	OH	H	OH	OH	H	H	12	(Saarinen et al., 2001)
								95	(Kao et al., 1998)
quercetin	H	OH	H	OH	OH	OH	OH	12	(Kellis and Vickery, 1984)
								80	(Ibrahim and Abul-Hajj, 1990)
6-OH-flavone	H	H	OH	H	H	H	H	5.5	(Saarinen et al., 2001)
								90	(Ibrahim and Abul-Hajj, 1990)
6,4'-di-OH-flavone	H	H	OH	H	H	OH	H	100	(Ibrahim and Abul-Hajj, 1990)
5-OH-flavone	H	H	H	OH	H	H	H	120	(Ibrahim and Abul-Hajj, 1990)
5,4'-di-OH-flavone	H	H	H	OH	H	OH	H	140	(Ibrahim and Abul-Hajj, 1990)
3-OH-flavone	H	H	H	H	OH	H	H		
<hr/>									
Isoflavones	R ¹	R ²	R ³						
biochanin A	OH	OH	OMe					36	(Jeong et al., 1999)
								49	(Le Bail et al., 1998)
								12	(Kao et al., 1998)
								49	(Campbell and Kurzer, 1993)
genistein	OH	OH	OH					123	(Kao et al., 1998)
								>50	(Edmunds)

	isoflavone	H	H	H	>200	et al., 2005) (Ibrahim and Abul-Hajj, 1990)
	coumestrol				17	1.3 (Wang et al., 1994)
1	‡ R ¹ + R ² = -CH=CH-CH=CH-					
2						

Accepted Manuscript

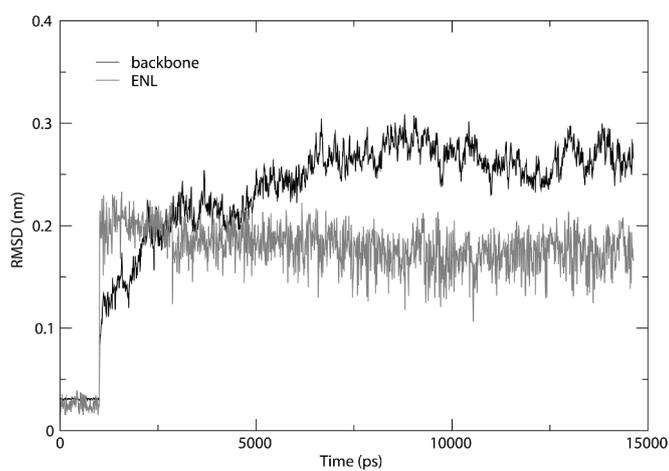
1



2

3 Figure 2. The general structures of lignanolactones, lignanodiols, flavones and
 4 isoflavones with atom numbering and ring labelling shown. The structures of
 5 enterolactone (ENL) and α -naphthoflavone (ANF) used in the molecular dynamics
 6 simulations (MDS), and the structure of coumestrol.

7



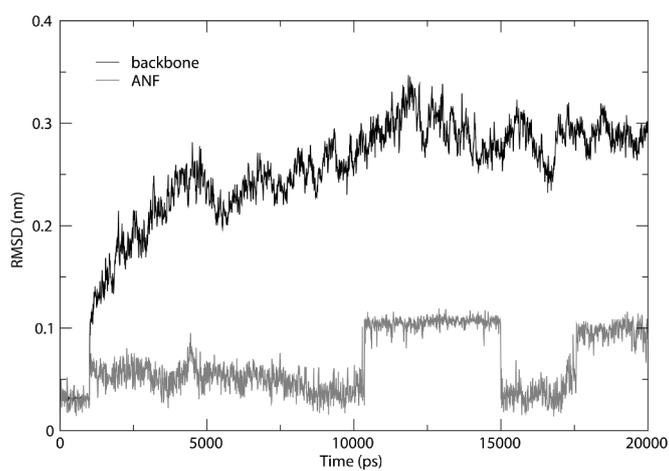
1

2 Figure 3. The root mean square deviations (rmsd) during the MDS of the aromatase—

3 ENL complex. The black curve represents the backbone of the aromatase enzyme and

4 the grey curve represents the lignan ENL.

5



1

2 Figure 4. The rmsds during the MDS of the aromatase—ANF complex. The black
3 curve represents the backbone of the aromatase enzyme and the grey curve represents
4 the flavone ANF.

5

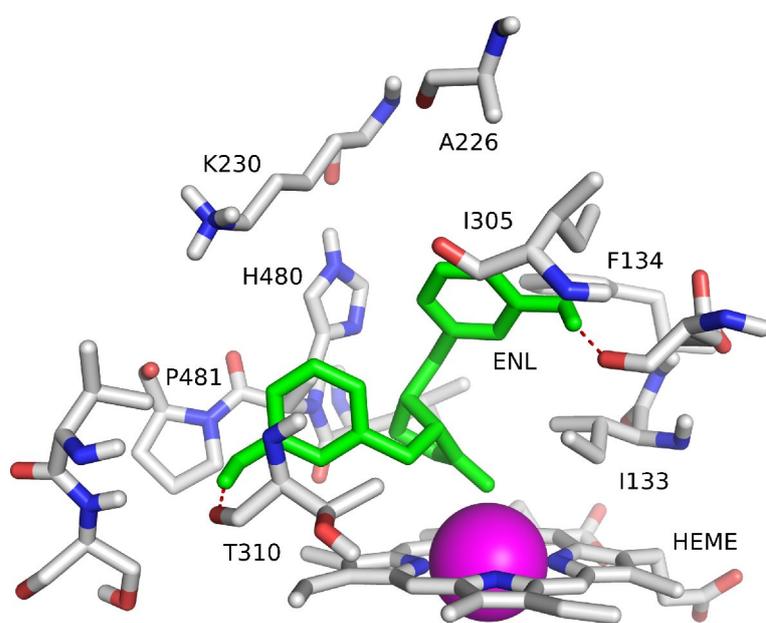
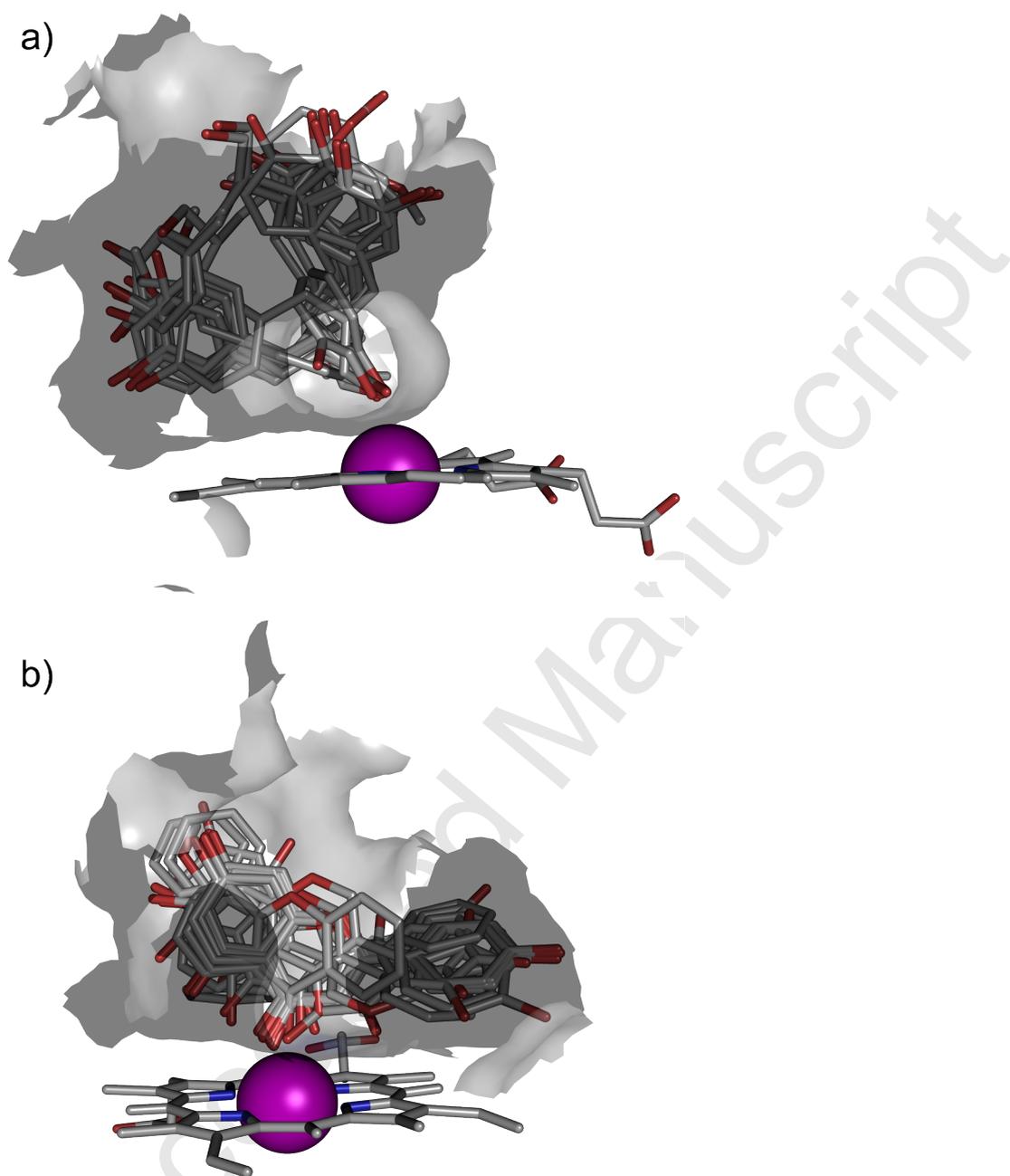
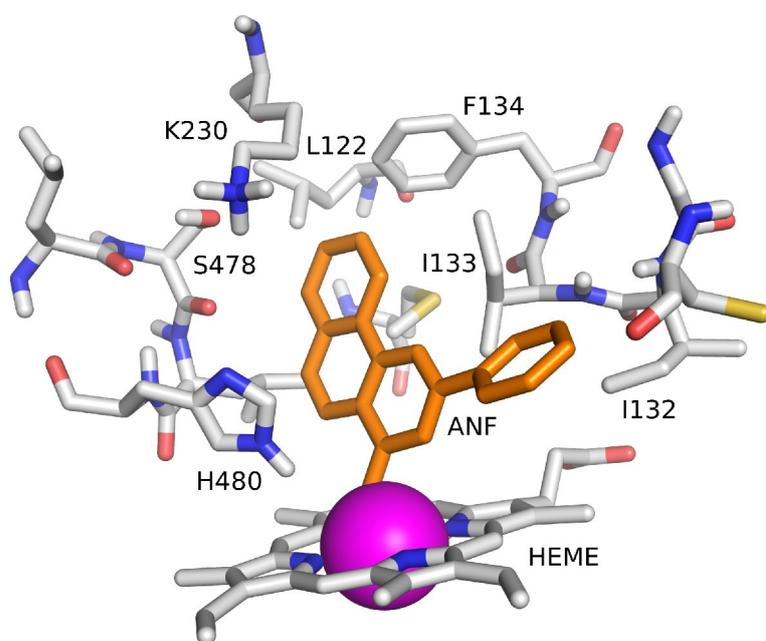


Figure 5. ENL docked into the active site of the representative structure of the aromatase – ENL trajectory. ENL is coloured green and the heme iron is represented as a magenta sphere. Only selected active site residues are displayed for clarity.



1
2
3
4
5
6
7

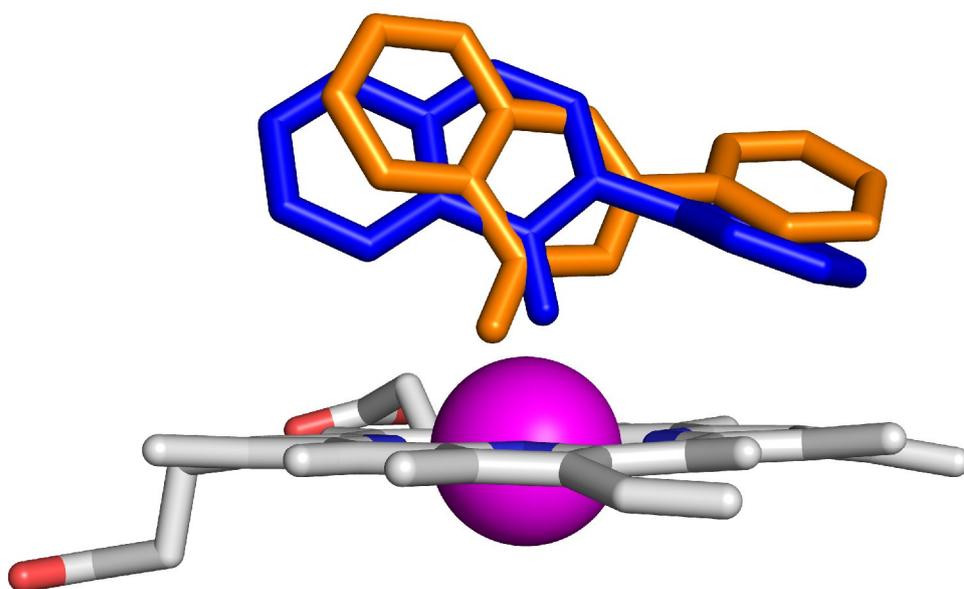
Figure 6. The docked lignan set and flavonoids/isoflavonoid set in the aromatase active site. Only heavy atoms are displayed for clarity. The active site is represented as a transparent surface.



1

2 Figure 7. ANF docked into the active site of the representative structure of the
3 aromatase – ANF trajectory. ANF is coloured orange and the heme iron is represented
4 as a magenta sphere. Only selected active site residues are displayed for clarity.

5

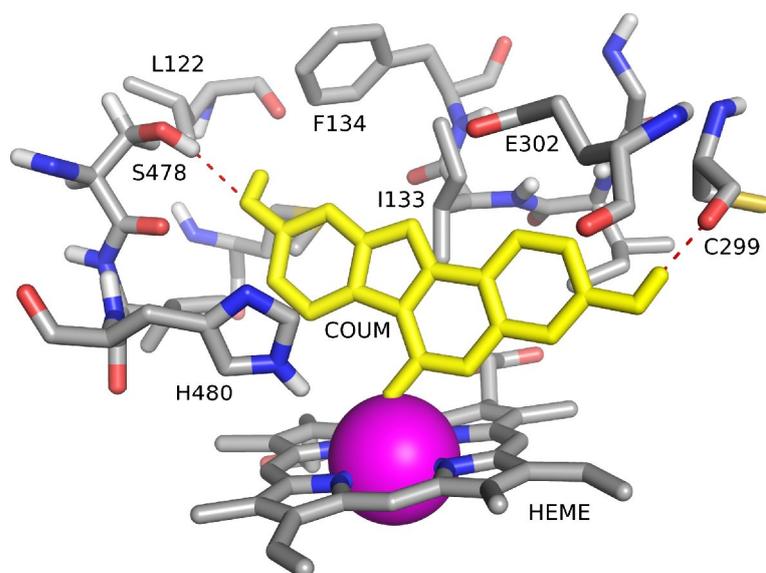


1

2 Figure 8. The interaction between flavone and isoflavone and the heme iron. Flavone
3 is coloured orange, isoflavone is coloured blue and the heme iron is represented as a
4 magenta sphere. Only heavy atoms are shown for clarity.

5

6



1
2 Figure 9. Coumestrol docked into the representative structure of the aromatase – ANF
3 trajectory. Coumestrol is coloured yellow and the heme iron is represented as a
4 magenta sphere. Only selected residues are displayed for clarity.

5

References

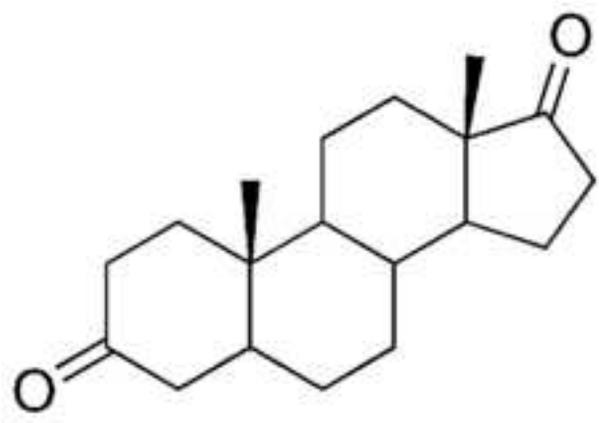
- 1
2
- 3 Accelrys Inc., Insight II version 2005, 10188 Telesis Court, Suite 100 San Diego, CA 92121,
4 USA.
- 5 Adlercreutz, H., 1995. Phytoestrogens: epidemiology and a possible role in cancer protection.
6 *Environ. Health Perspect.* 103, 103-112.
- 7 Adlercreutz, H., Bannwart, C., Wähälä, K., Mäkelä, T., Brunow, G., Hase, T., Arosemena,
8 P.J., Kellis Jr., J.T., Vickery, L.E., 1993. Inhibition of human aromatase by mammalian
9 lignans and isoflavonoid phytoestrogens. *J. Steroid Biochem. Mol. Biol.* 44, 147-153.
- 10 Borriello, S.P., Setchell, K.D., Axelson, M., Lawson, A.M., 1995. Production and metabolism
11 of lignans by the human faecal flora. *J. Appl. Bacteriol.* 58, 37-43.
- 12 Brooks, J.D., Thompson, L.U., 2005. Mammalian lignans and genistein decrease the activities
13 of aromatase and 17 β -hydroxysteroid dehydrogenase in MCF-7 cells, *J. Steroid Biochem.*
14 *Mol. Biol.* 94, 461-467.
- 15 Brueggemeier, R.W., Whetstone, J.L., 2001. Effects of phytoestrogens and synthetic
16 combinatorial libraries on aromatase, estrogen biosynthesis, and metabolism, *Ann. N.Y. Acad.*
17 *Sci.* 948, 51-66.
- 18 Campbell, D.R., Kurzer, M.S., 1993. Flavonoid inhibition of aromatase enzyme activity in
19 human preadipocytes. *J. Steroid Biochem. Mol. Biol.* 46, 381-388.
- 20 Cavalli, A., Bisi, A., Bertucci, C., Rosini, C., Paluszczak, A., Gobbi, S., Giorgio, E., Rampa,
21 A., Belluti, F., Piazza, L., Valenti, P., Hartmann, R.W., Recanatini, M., 2005. Enantioselective
22 nonsteroidal aromatase inhibitors identified through a multidisciplinary medicinal chemistry
23 approach, *J. Med. Chem.* 48, 7282-7289.
- 24 Chen, S., Kao, Y.C., Laughton, C.A., 1997. Binding characteristics of aromatase inhibitors
25 and phytoestrogens to human aromatase, *J. Steroid Biochem. Mol. Biol.* 61, 107-115.
- 26 Clavel, T., Borrmann, D., Braune, A., Blaut, J.D.M., 2006. Occurrence and activity of human
27 intestinal bacteria involved in the conversion of dietary lignans, *Anaerobe* 12, 140-147.
- 28 Dawson, J., Andersson, L., Sono, M., 1982. Spectroscopic investigations of ferric cytochrome
29 P-450-CAM ligand complexes. Identification of the ligand trans to cysteinate in the native
30 enzyme, *J. Biol. Chem.* 257, 3606-3617.
- 31 DeLano, W.L., 2002. The PyMOL molecular graphics system, DeLano Scientific, Palo Alto,
32 CA, USA.
- 33 Edmunds, K.M., Holloway, A.C., Crankshaw, D.J., Agarwal, S.K., Foster, W.G., 2005. The
34 effects of dietary phytoestrogens on aromatase activity in human endometrial stromal cells.
35 *Reprod. Nutr. Dev.* 45, 709-720.
- 36 Gobbi, S., Cavalli, A., Rampa, A., Belluti, F., Piazza, L., Paluszczak, A., Hartmann, R.W.,
37 Recanatini, M., Bisi, A., 2006. Lead optimization providing a series of flavone derivatives as
38 potent nonsteroidal inhibitors of the cytochrome P450 aromatase enzyme, *J. Med. Chem.* 49,
39 4777-4780.

- 1 Hackett, J.C., Kim, Y., Su, B., Brueggemeier, R.W., 2005. Synthesis and characterization of
2 azole isoflavone inhibitors of aromatase, *Bioorg. Med. Chem.* 13, 4063-4070.
- 3 Hendlich, M., 1998. Databases for protein-ligand complexes, *Acta Crystallogr. D* 54, 1178-
4 1182.
- 5 Hong, Y., Cho, M., Yuan, Y., Chen, S., 2008. Molecular basis for the interaction of four
6 different classes of substrates and inhibitors with human aromatase, *Biochem. Pharmacol.* 75,
7 1161-1169.
- 8 Ibrahim, A.R., Abul-Hajj, Y.J., 1990. Aromatase inhibition by flavonoids. *J. Steroid Biochem.*
9 *Mol. Biol.* 37, 257-260.
- 10 Jeong, H.J., Shin, Y.G., Kim, I.H., Pezzuto, J.M., 1999. Inhibition of aromatase activity by
11 flavonoids. *Arch. Pharmacol. Res.* 22, 309-312.
- 12 Jones, G., Willet, P., Glen, R.C., Leach, A.R., Taylor, R., 1997. Development and validation
13 of a genetic algorithm for flexible docking, *J. Mol. Biol.* 267, 727-748.
- 14 Kao, Y., Cam, L.L., Laughton, C.A., Zhou, D., Chen, S., 1996. Binding characteristics of
15 seven inhibitors of human aromatase: A site-directed mutagenesis study, *Cancer Res.* 56,
16 3451-3460.
- 17 Kao, Y., Korzekwa, K.R., Laughton, C.A., Chen, S., 2001. Evaluation of the mechanism of
18 aromatase cytochrome P450. A site-directed mutagenesis study, *Eur. J. Biochem.* 268, 243-
19 251.
- 20 Kao, Y., Zhou, C., Sherman, M., Laughton, C.A., Chen, S., 1998. Molecular basis of the
21 inhibition of human aromatase (estrogen synthetase) by flavone and isoflavone
22 phytoestrogens: A site-directed mutagenesis study, *Environ. Health Perspect.* 106, 85-92.
- 23 Karkola, S., Höltje, H., Wähälä, K., 2007. A three-dimensional model of CYP19 aromatase
24 for structure-based drug design, *J. Steroid Biochem. Mol. Biol.* 105, 63-70.
- 25 Kellis Jr., J.T., Nesnow, S., Vickery, L.E., 1986. Inhibition of aromatase cytochrome P-450
26 (estrogen synthetase) by derivatives of α -naphthoflavone, *Biochem. Pharmacol.* 35, 2887-
27 2891.
- 28 Kellis Jr., J.T., Vickery, L.E., 1984. Inhibition of human estrogen synthetase (aromatase) by
29 flavones, *Science* 225, 1032-1034.
- 30 Kellis Jr., J.T., Vickery, L.E., 1987. Purification and characterization of human placental
31 aromatase cytochrome P-450, *J. Biol. Chem.* 262, 4413-4420.
- 32 Kiely, M., Faughnan, M., Wähälä, K., Brants, H., 2003. Phyto-oestrogen levels in foods: the
33 design and construction of the VENUS database. *Br. J. Nutr.* 89, S19-S23.
- 34 Kim, Y.W., Hackett, J.C., Brueggemeier, R.W., 2004. Synthesis and aromatase inhibitory
35 activity of novel pyridine-containing isoflavones, *J. Med. Chem.* 47, 4032-4040.
- 36 Kirton, S.B., Murray, C.W., Verdonk, M.L., Taylor, R.D., 2005. Prediction of binding modes
37 for ligands in the cytochromes P450 and other heme-containing proteins. *Proteins* 58, 836-
38 844.
- 39 Kuhnel, K., Blankenfeldt, W., Ternner, J., Schlichting, I., 2006. Crystal structures of
40 chloroperoxidase with its bound substrates and complexed with formate, acetate, and nitrate.
41 *J. Biol. Chem.* 281, 23990-23998.

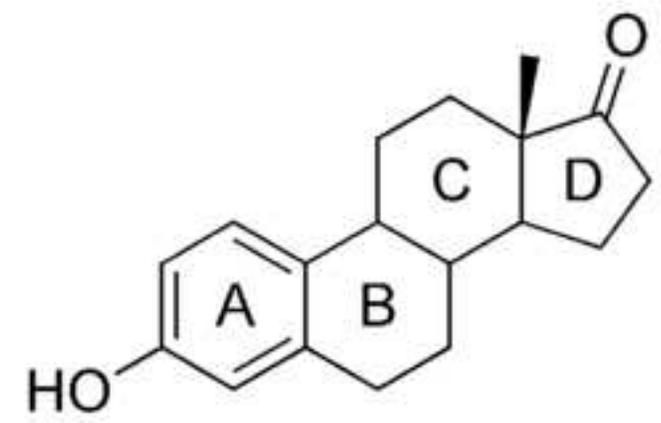
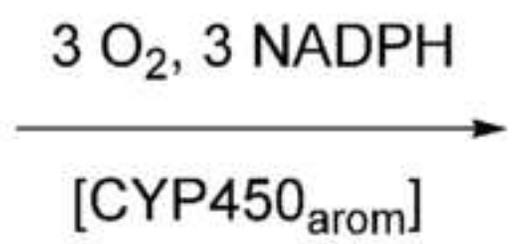
- 1 Laughton, C.A., Zvelebil, M.J., Neidle, S., 1993. A detailed molecular model for human
2 aromatase. *J. Steroid Biochem. Mol. Biol.* 44, 399-407.
- 3 Le Bail, J.C., Laroche, T., Marre-Fournier, F., Habrioux, G., 1998. Aromatase and 17 β -
4 hydroxysteroid dehydrogenase inhibition by flavonoids, *Cancer Lett.* 133, 101-106.
- 5 Li, H., Narasimhulu, S., Havran, L.M., Winkler, J.D., Poulos, T.L., 1995. Crystal structure of
6 cytochrome P450cam complexed with its catalytic product, 5-exo-hydroxycamphor, *J. Am.*
7 *Chem. Soc.* 117, 6297-6299.
- 8 Lof, M., Weiderpass, E., 2006. Epidemiologic evidence suggests that dietary phytoestrogen
9 intake is associated with reduced risk of breast cancer. *Nutr. Res.* 26, 609-619.
- 10 Mazura, W.M., Uehara, M., Adlercreutz, K.W.H., 2000. Phyto-oestrogen content of
11 berries, and plasma concentrations and urinary excretion of enterolactone after a single
12 strawberry-meal in human subjects. *Br. J. Nutr.* 83, 381-387.
- 13 Milder, I.E., Feskens, E.J., Arts, I.C., de Mesquita, H.B., Hollman, P.C., Kromhout, D., 2005.
14 Intake of the plant lignans secoisolariciresinol, matairesinol, lariciresinol, and pinoresinol in
15 Dutch men and women. *J. Nutr.* 135, 1202-1207.
- 16 Milder, I.E.J., Arts, I.C.W., Venema, D.P., Lasaroms, J.J.P., Wahala, K., Hollman, P.C.H.,
17 2004. Optimization of a liquid chromatography-tandem mass spectrometry method for
18 quantification of the plant lignans secoisolariciresinol, matairesinol, lariciresinol, and
19 pinoresinol in foods, *J. Agric. Food Chem.* 52, 4643-4651.
- 20 Mosselman, S., Polman, J., Dijkema, R., 1996. ER β : Identification and characterization of a
21 novel human estrogen receptor, *FEBS Lett.* 392, 49-53.
- 22 Mueller, S.O., Simon, S., Chae, K., Metzler, M., Korach, K.S., 2004. Phytoestrogens and their
23 human metabolites show distinct agonistic and antagonistic properties on estrogen receptor α
24 (ER α) and ER β in human cells, *Toxicol. Sci.* 80, 14-25.
- 25 Muti, P., Rogan, E., Cavalieri, E., 2006. Androgens and estrogens in the etiology and
26 prevention of breast cancer, *Nutr. Cancer* 56, 247-252.
- 27 Neves, M.A.C., Dinis, T.C.P., Colombo, G., Sa e Melo, M.L., 2007. Combining
28 computational and biochemical studies for a rationale on the anti-aromatase activity of natural
29 polyphenols, *ChemMedChem* 2, 1750-1762.
- 30 Raffaelli, B., Leppälä, E., Chappuis, C., Wähälä, K., 2006. New potential mammalian lignan
31 metabolites of environmental phytoestrogens, *Environ. Chem. Lett.* 4, 1-9.
- 32 Recanatini, M., Bisi, A., Cavalli, A., Belluti, F., Gobbi, S., Rampa, A., Valenti, P., Palzer, M.,
33 Paluszczak, A., Hartmann, R.W., 2001. A new class of nonsteroidal aromatase inhibitors:
34 design and synthesis of chromone and xanthone derivatives and inhibition of the P450
35 enzymes aromatase and 17 α -hydroxylase/C17,20-lyase, *J. Med. Chem.* 44, 672-680.
- 36 Rupp, B., Raub, S., Marian, C., Hölftje, H., 2005. Molecular design of two sterol 14 α -
37 demethylase homology models and their interactions with the azole antifungals ketoconazole
38 and bifonazole, *J. Comput.-Aided Mol. Des.* 19, 149-163.
- 39 Saarinen, N., Joshi, S.C., Ahotupa, M., Li, X., Ämmälä, J., Mäkelä, S., Santti, R., 2001. No
40 evidence for the in vivo activity of aromatase-inhibiting flavonoids, *J. Steroid Biochem. Mol.*
41 *Biol.* 78, 231-239.

- 1 Sanderson, J.T., Hordijk, J., Denison, M.S., Springsteel, M.F., Nantz, M.H., van den Berg,
2 M., 2004. Induction and inhibition of aromatase (CYP19) activity by natural and synthetic
3 flavonoid compounds in H295R human adrenocortical carcinoma cells. *Toxicol. Sci.* 82, 70-
4 79.
- 5 Santen, R., 2002. To block estrogen's synthesis or action: That is the question. *J. Clin.*
6 *Endocrinol. Metab.* 87, 3007-3012.
- 7 Stermitz, F.R., Bais, H.P., Foderaro, T.A., Vivanco, J.M., 2003. 7,8-Benzoflavone: a
8 phytotoxin from root exudates of invasive Russian knapweed, *Phytochemistry* 64, 493-497.
- 9 Stresser, D.M., Turner, S.D., McNamara, J., Stocker, P., Miller, V.P., Crespi, C.L., Patten,
10 C.J., 2000. A high-throughput screen to identify inhibitors of aromatase (CYP19), *Anal.*
11 *Biochem.* 284, 427-430.
- 12 Su, B., Hackett, J.C., Diaz Cruz, E.S., Kim, Y.W., Brueggemeier, R.W., 2005. Lead
13 optimization of 7-benzyloxy 2-(4'-pyridylmethyl)thioisoflavone aromatase inhibitors, *Bioorg.*
14 *Med. Chem.* 13, 6571-6577.
- 15 Tripos Inc., Sybyl 7.3, St. Louis, MO, USA.
- 16 Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H. J., 2005.
17 GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26, 1701-1718.
- 18 Wang, C., Mäkelä, T., Hase, T., Adlercreutz, H., Kurzer, M.S., 1994. Lignans and flavonoids
19 inhibit aromatase enzyme in human preadipocytes. *J. Steroid Biochem. Mol. Biol.* 50, 205-
20 212.
- 21 Wang, L.Q., Meselhy, M.R., Li, Y., Qin, G.W., Hattori, M., 2000. Human intestinal bacteria
22 capable of transforming secoisolariciresinol diglucoside to mammalian lignans, enterodiol and
23 enterolactone. *Chem. Pharm. Bull. (Tokyo)* 48, 1606-1610.
- 24 Wang, X., Goff, H.M., 1997. A nuclear paramagnetic relaxation study of the interaction of the
25 cyclopentanedione substrate with chloroperoxidase, *Biochim. Biophys. Acta* 1339, 88-96.
- 26 White, R., Coon, M., 1982. Heme ligand replacement reactions of cytochrome P-450.
27 Characterization of the bonding atom of the axial ligand trans to thiolate as oxygen, *J. Biol.*
28 *Chem.* 257, 3073-3083.
- 29 Wähälä, K., Karkola, S., Lilienkampf, A., 2008. Phytoestrogens in drug discovery. In: Daayf,
30 F. and Lattanzio, V. (eds), *Recent advances in polyphenol research*. Blackwell Publishing,
31 UK.
- 32

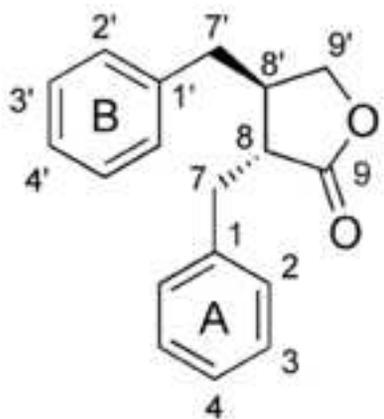
Manuscript



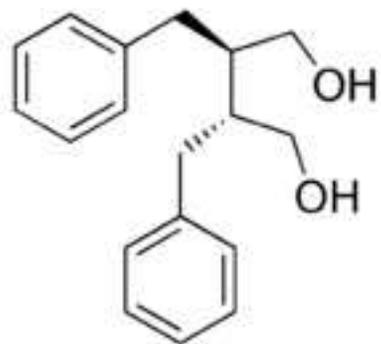
androstenedione



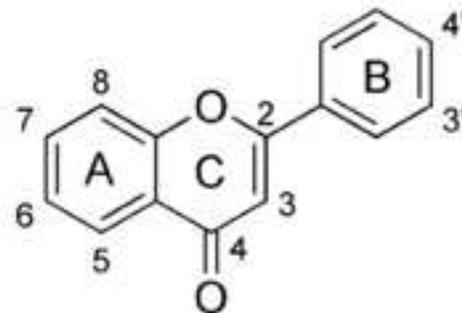
estrone



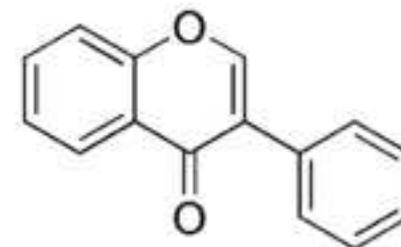
lignan lactones



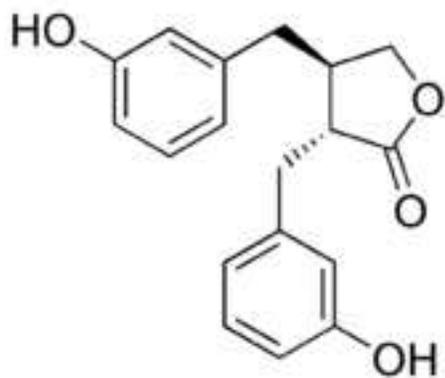
lignan diols



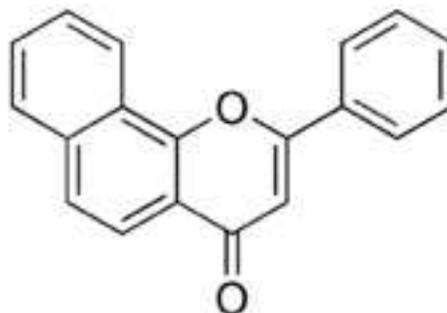
flavones



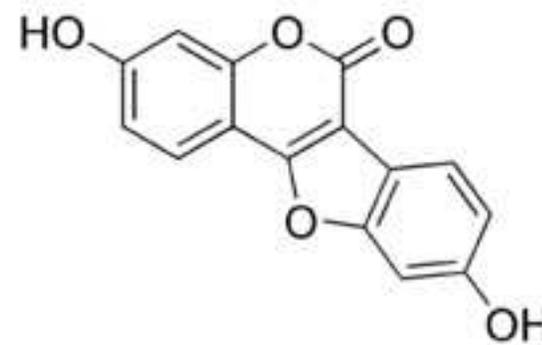
isoflavones



enterolactone



α -naphthoflavone



coumestrol

Figure3

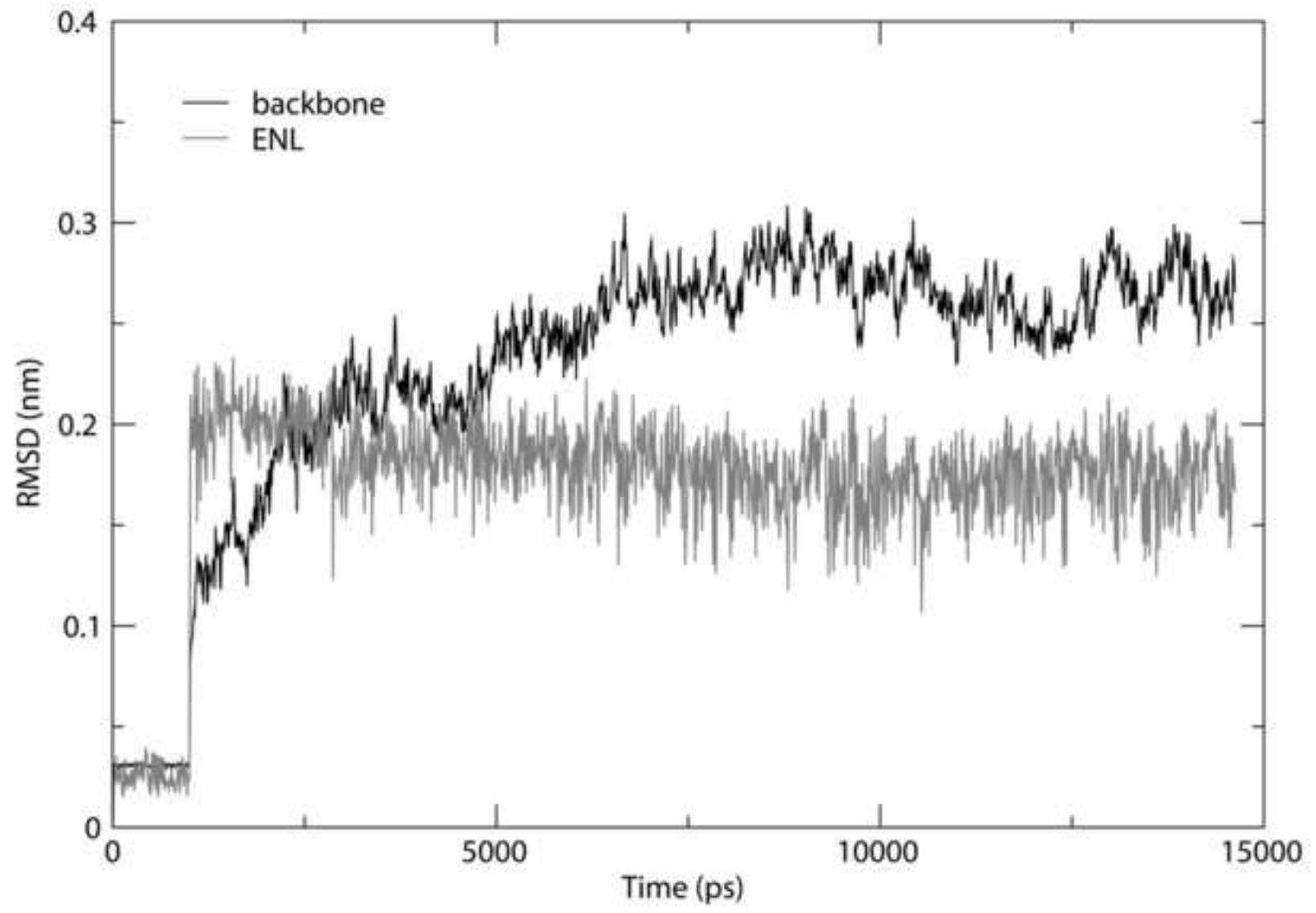


Figure4

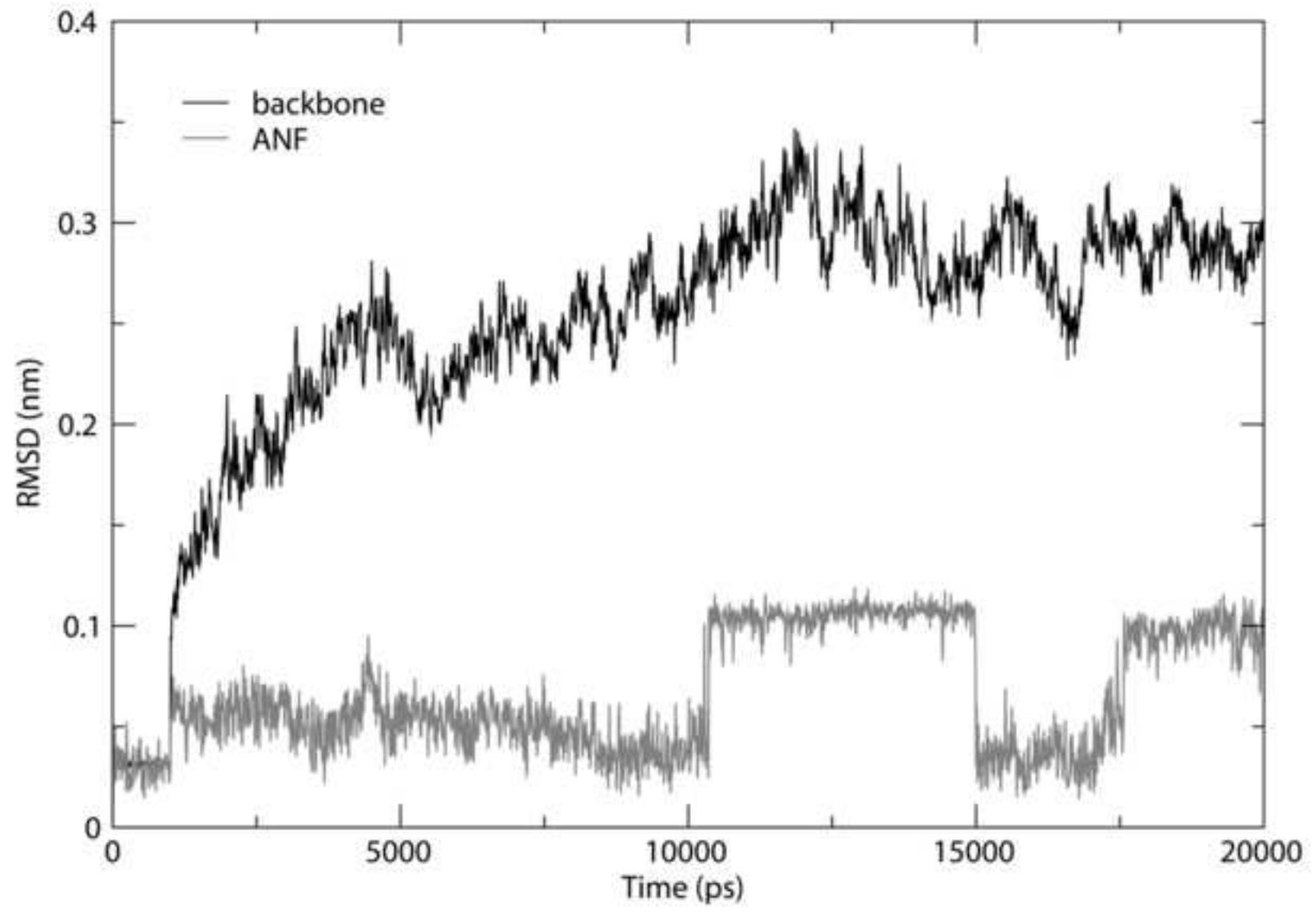
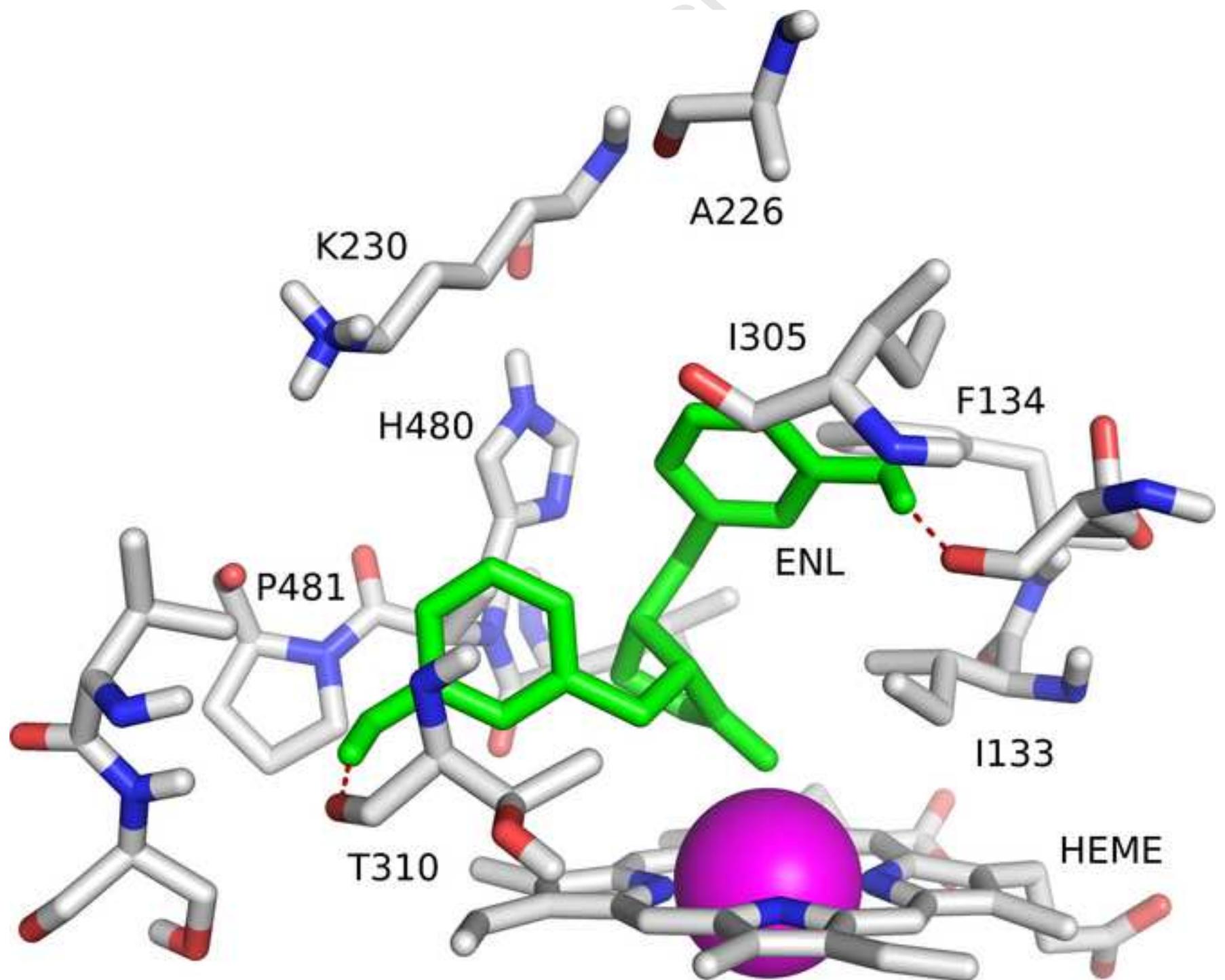


Figure 5



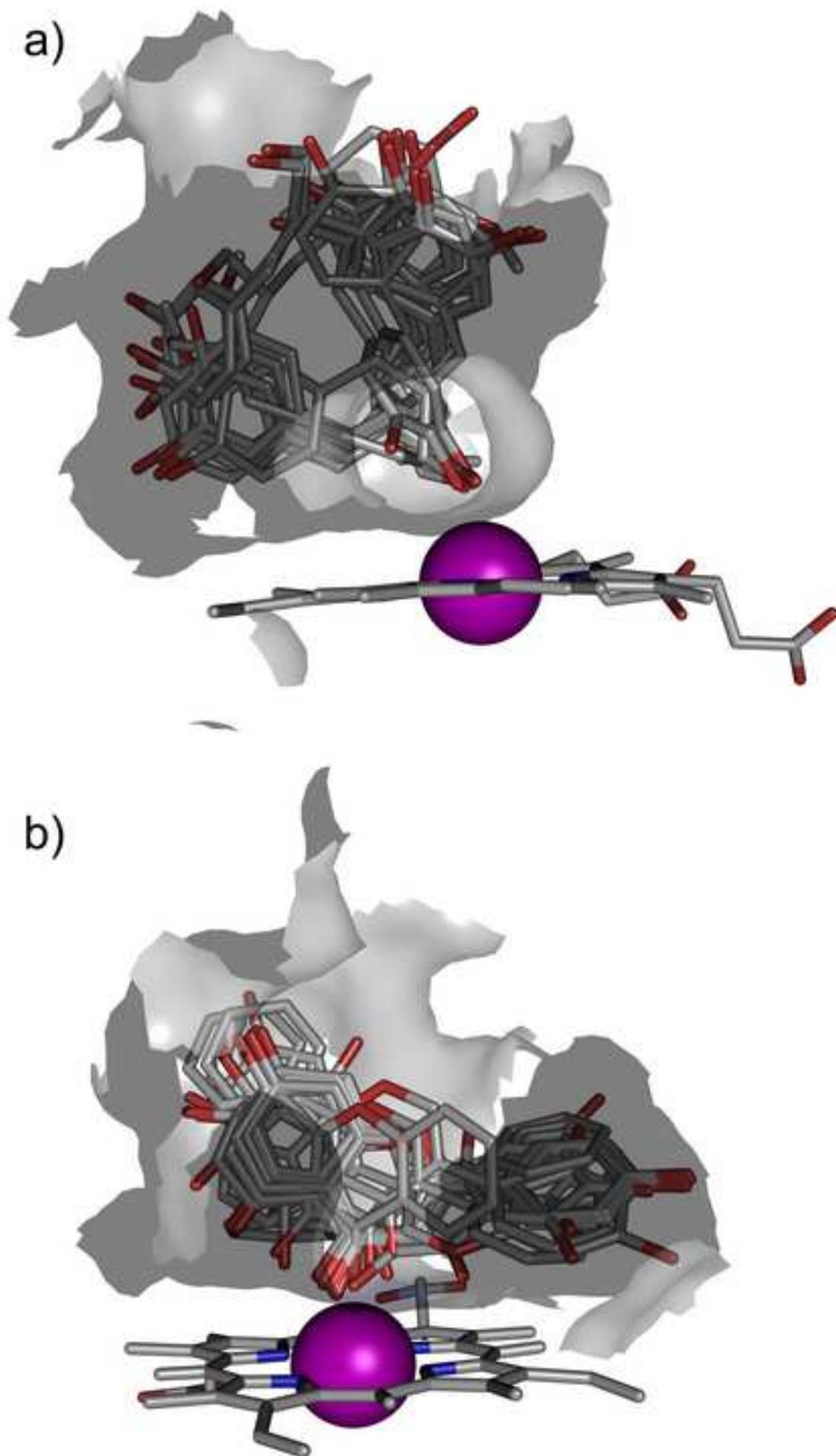
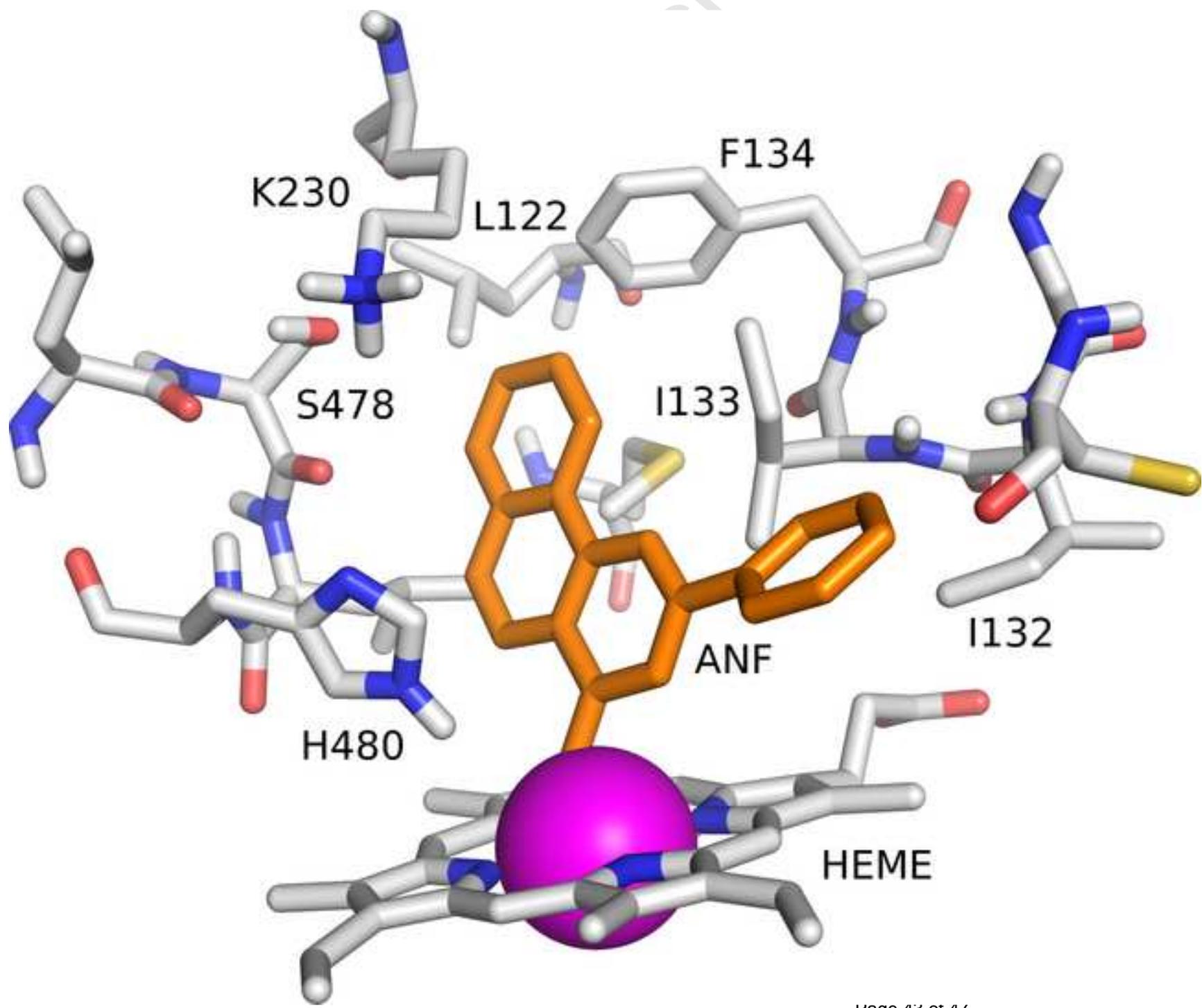


Figure 7



scrip

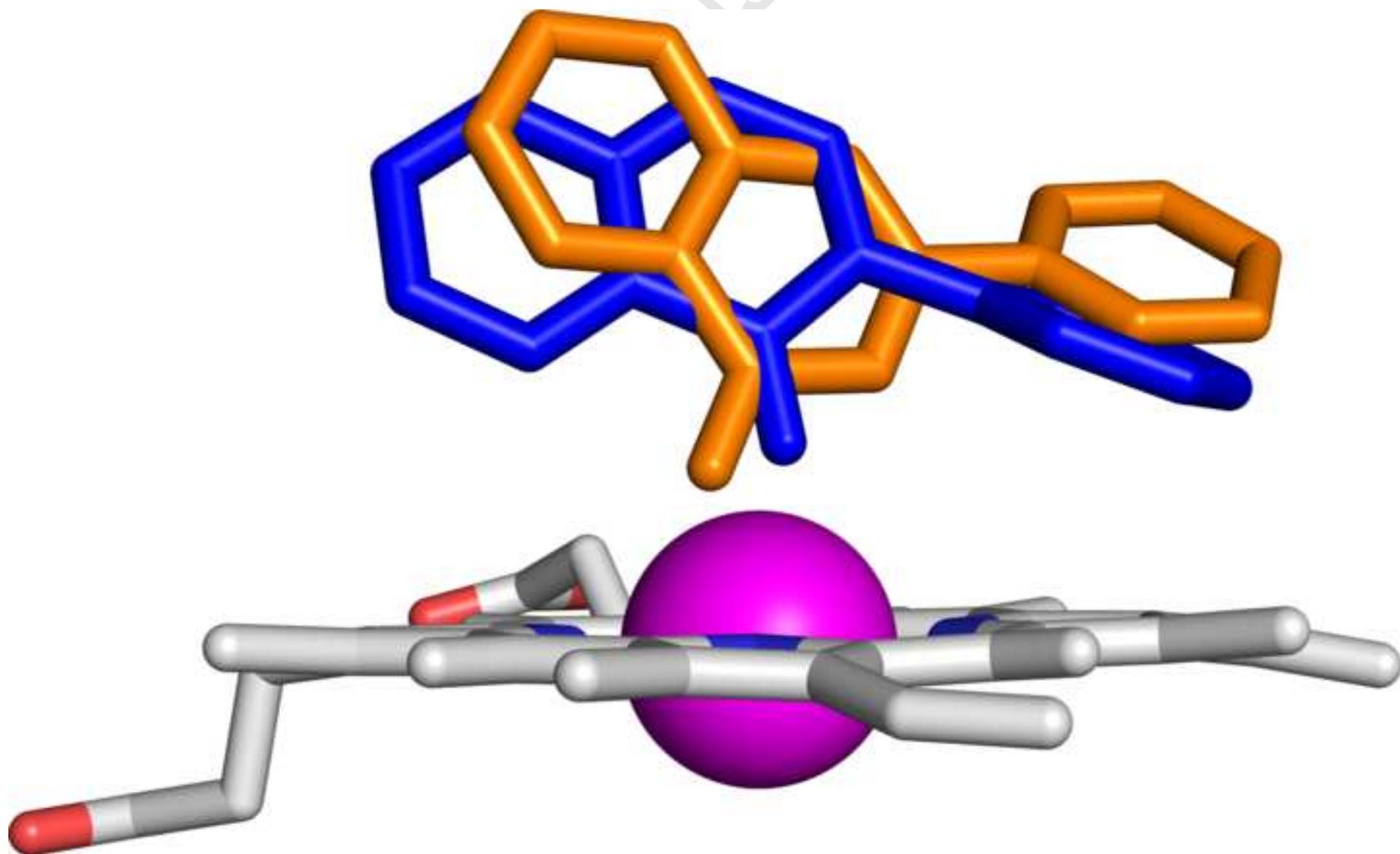
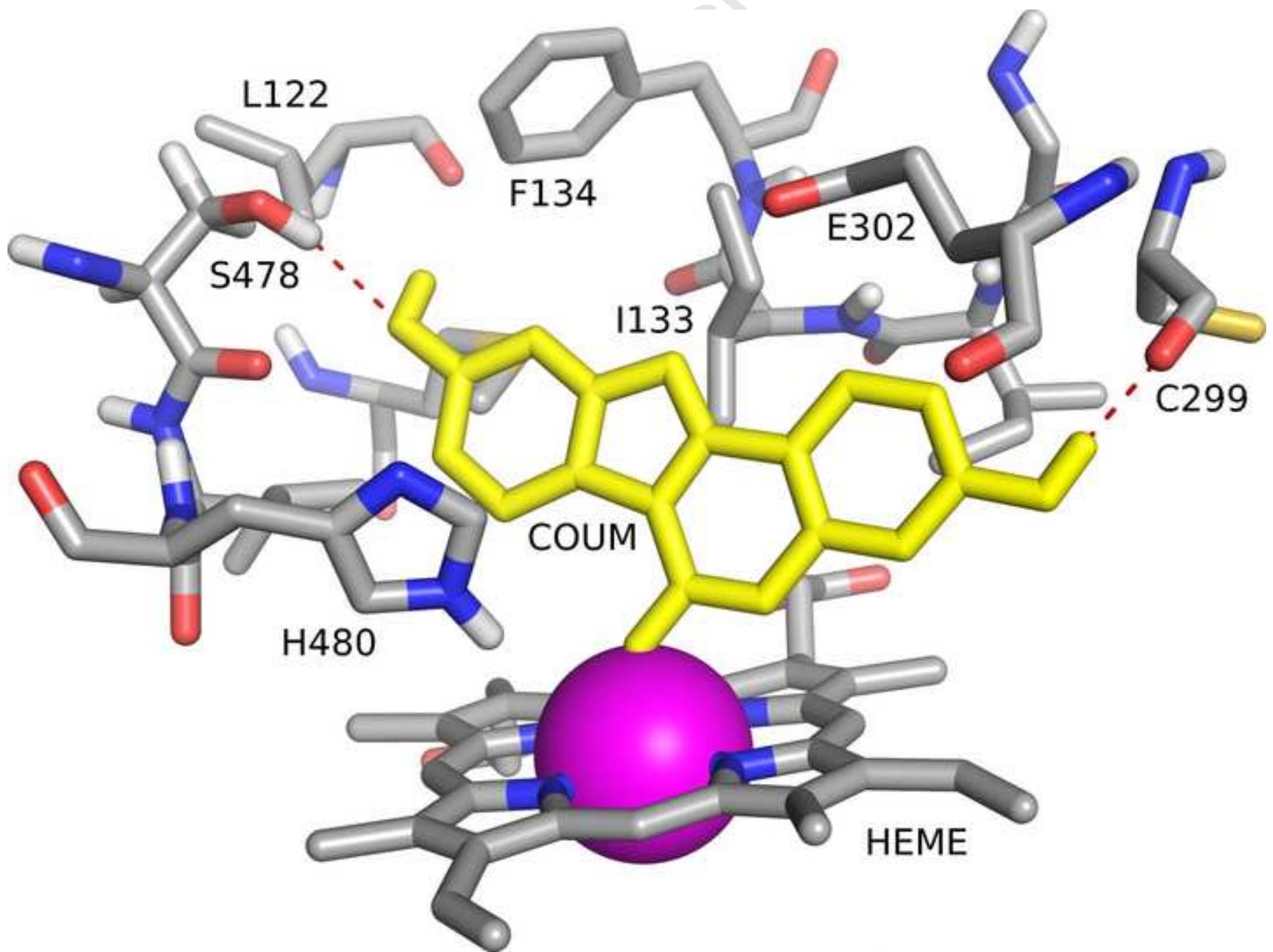
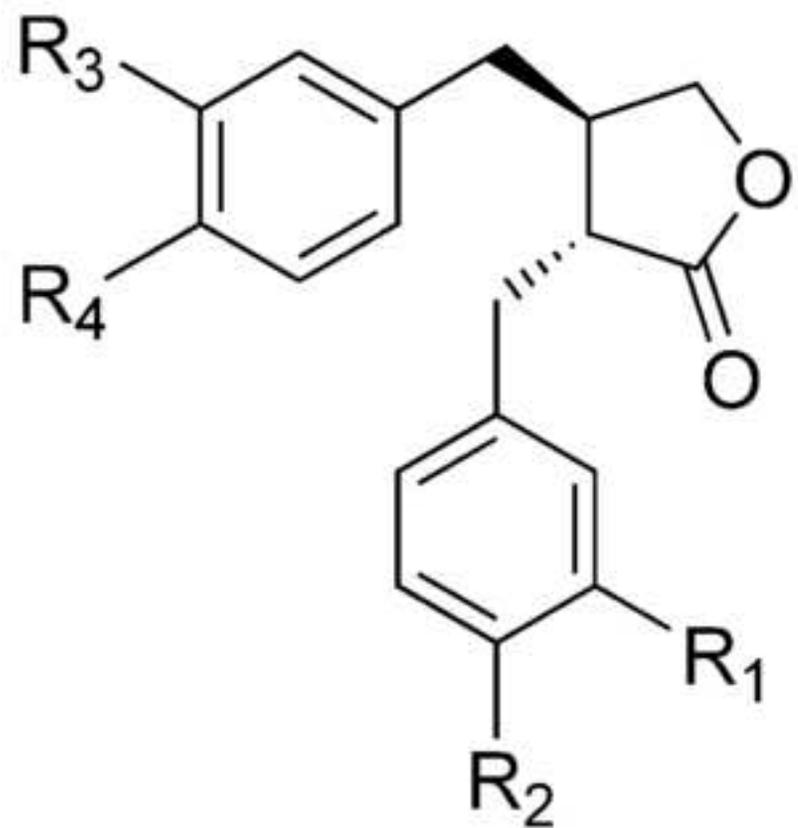
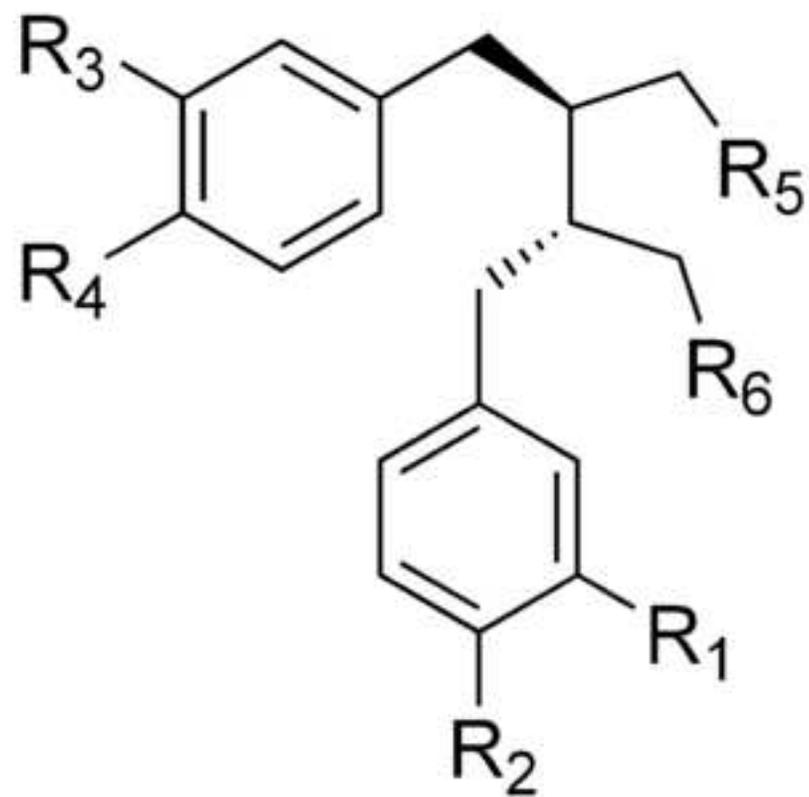


Figure9

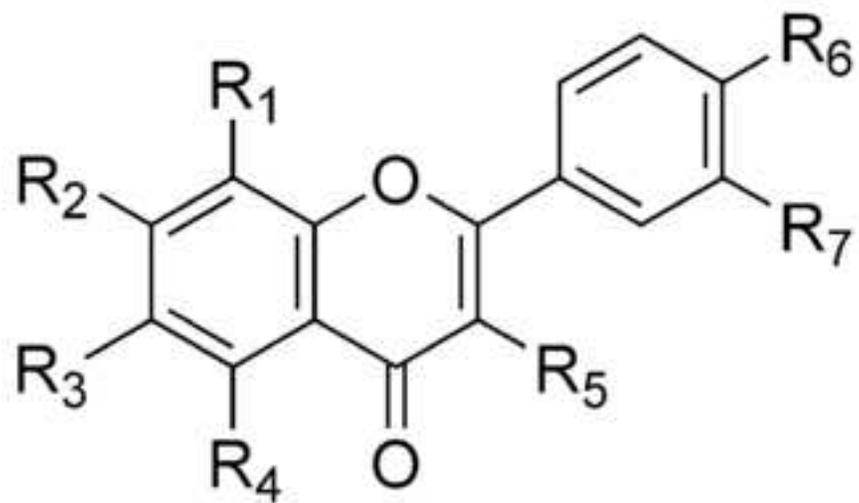




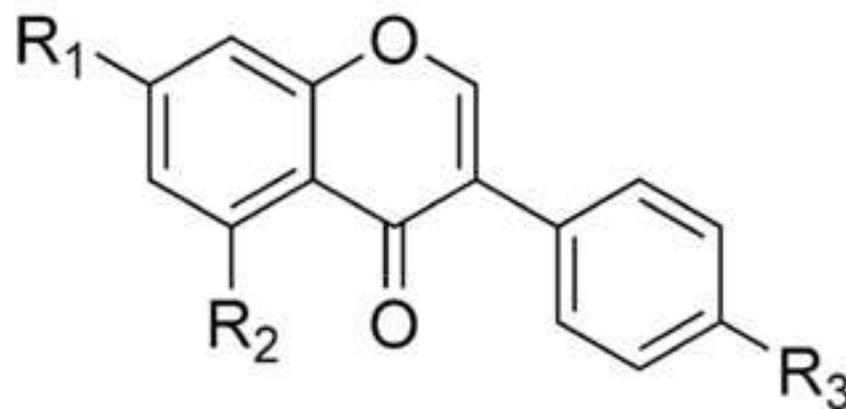
Lignanlactones



Lignanodiols



Flavones



Isoflavones