PAPER

Check for updates

Mechanisms of allosteric and mixed mode aromatase inhibitors†

Samson A. Souza, [ID] [a] Abby Held,[b] Wenjie J. Lu,[c] Brendan Drouhard,[a] Bryant Avila,[a] Raul Leyva-Montes,[a] Michelle Hu, [ID] [c] Bill R. Miller 3rd[b] and Ho Leung Ng [ID] *[a]

Aromatase (CYP19) catalyzes the last biosynthetic step of estrogens in mammals and is a primary drug target for hormone-related breast cancer. However, treatment with aromatase inhibitors is often associated with adverse effects and drug resistance. In this study, we used virtual screening targeting a predicted cytochrome P450 reductase binding site on aromatase to discover four novel non-steroidal aromatase inhibitors. The inhibitors have potencies comparable to the noncompetitive tamoxifen metabolite, endoxifen. Our two most potent inhibitors, AR11 and AR13, exhibit both mixed-type and competitive-type inhibition. The cytochrome P450 reductase-CYP19 coupling interface likely acts as a transient binding site. Our modeling shows that our inhibitors bind better at different sites near the catalytic site. Our results predict the location of multiple ligand binding sites on aromatase. The combination of modeling and experimental results supports the important role of the reductase binding interface as a low affinity, promiscuous ligand binding site. Our new inhibitors may be useful as alternative chemical scaffolds that may show different adverse effects profiles than current clinically used aromatase inhibitors.

Introduction

Aromatase (CYP19) catalyzes the three-step transformation of androgens to estrogens in mammals by a mechanism that is well-characterized. The first two steps produce a C-19 gem-diol.[1] 1β-Hydrogen abstraction and electron transfer furnish the A-ring aromatization and loss of a formate molecule to complete a single turnover.[2] As such, aromatase inhibition is a common strategy to treat patients with hormone-dependent breast cancer.[3] Third-generation aromatase inhibitors (AIs) are used as primary therapeutics and long-term adjuvants in postmenopausal women with breast cancer.

AIs in clinical use include the non-steroidal AIs (NSAIs), anastrozole and letrozole, and the steroidal AI, exemestane, which interacts covalently at the active site.[4,5] The most potent NSAI, letrozole, binds a distinct site exhibiting both noncompetitive and mixed-mode type Michaelis inhibition.[6] This type of behavior is also observed for the noncompetitive AI endoxifen (K_i = 4 μM), a potent metabolite of tamoxifen, a clinically-used estrogen receptor α (ERα) antagonist.[7] Metabolic *N*-demethylation of endoxifen produces a competitive AI, norendoxifen

(K_i = 35 nM).[8] This demonstrates two important points. First, minor chemical modifications of inhibitors (letrozole to anastrozole, and endoxifen to norendoxifen) can change the inhibitory mechanism in unexpected ways. Second, CYP19 can be allosterically modulated with high potency. The mechanism of mixed mode inhibition for CYP19 and other cytochrome P450 (CYP450) enzymes is unclear.

Steroidal AIs interact at the active site, also known as the distal heme site, with high efficacy. However, steroidal analogs frequently exert similar adverse effects. CYP19 inhibition at an alternative allosteric site provides opportunities for the discovery of novel NSAIs chemically distinct from current therapeutics and with different toxicity profiles.

In this work, we investigate the binding of small molecule inhibitors to the CYP19 heme distal site, the heme proximal site (the predicted site of CPR binding), and the substrate access channel through a combination of experimental and computational approaches. These sites have been proposed by other investigators as potential therapeutic targets.[9,10] We first predict and model the interfacial contacts of cytochrome P450 reductase (CPR) with CYP19. We performed virtual screening against a library of over a million compounds to identify potential inhibitors. We use molecular dynamics simulations to model how they interact at the CPR–CYP19 interface. We then provide experimental enzyme inhibition data for four new NSAIs we discovered, AR11, AR13, AR19, and AR20. In addition, we use optical absorption spectra to characterize the effects of inhibition on the heme chemical environment. We characterize

[a] Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, KS, USA. E-mail: hng@ksu.edu

[b] Department of Chemistry, Truman State University, Kirksville, MO, USA

[c] Department of Chemistry, University of Hawai'i at Mānoa, Honolulu, HI, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cb00046b

the inhibition modes of AR11, AR13, AR19, and AR20, and their likeliest binding sites closer to the catalytic site. The cytochrome P450 reductase binding site is likely a transient, low-affinity binding site for multiple ligands.

Materials and methods

Protein–protein docking

For protein–protein docking, we used the crystal structures of human CYP19[11] (PDB 4KQ8) with a human-yeast chimeric enzyme[12] (PDB 3FJO) in an open conformation. There is currently no full-length human CPR crystal structure with an exposed FMN-containing face. Human CPR has an intrinsically dynamic hinge region causing the N-and C-terminal domains to adopt a similar open conformation.[13,14] The chimeric protein maintains functionality reducing both cytochrome *c* and human P450s.[12] The Haddock[15] webserver was used for the interface prediction-driven docking[16] of CYP19 (PDB 4KQ8) to CPR. Haddock uses the CPORT[16–21] ensemble with multiple interface predictors to determine the interfacial residues to input as restraints to return high scoring clusters of complexes. Residues that are solvent-exposed and predicted to reside at the CYP19–CPR interface were defined as ''Active'' for Haddock docking, and surrounding solvent-exposed residues were defined as "Passive" (Table S1, ESI†). The Gibbs free energy and dissociation constant estimates were calculated by the PRODIGY prediction webserver.[22,23] Conserved CYP19 surfaces were calculated using Bayes' theorem with the ConSurf server.[24,25] Human CYP19 (PDB 4KQ8) was queried against 150 homologs across different P450 families by setting a 30% maximum and 10% minimum sequence identity cutoff.

Virtual screening and docking

Idock[26,27] was used to virtually screen a library of over one million drug-like molecules from the ZINC database.[28] The highest scoring molecules were rescored with the DSX knowledge-based scoring function, which has been shown to be more predictive of binding affinity than docking scoring functions.[29] Compounds were assessed by the interactions made at the proximal heme site and their DSX scores. The top hits were selected for *in vitro* enzyme inhibition screening. Table S2 (ESI†) lists the vendors and physicochemical descriptions of the inhibitors used here. Compounds that exhibited moderate to potent anti-aromatase activity (except AR13, for which we had identified the binding site by optical absorption assays) were docked against a crystal structure of CYP19 (PDB 3S79) with a more stringent protocol for simulations, involving 10 independent docking runs with Autodock Vina using different random seeds.[30] Autodock Vina exhaustiveness was set to 128, and the number of binding modes was capped at 20 per run. Visual inspection by VMD[31] of clustered docking results was used to predict the best binding mode to be used for MD simulations. The stereochemistry of AR13 was ambiguous from the vendor's (Enamine) molecular description. Both the *cis*- and

trans-cyclopropane forms were used for docking and further modeling.

Molecular dynamics simulations

The AMBER16[32] molecular dynamics package was used to produce MD simulations. Parameters for the unliganded heme described by Shahrock *et al.* were used.[33] The FF14SB[34] and GAFF[35] forcefields were used for protein and ligands, respectively. Simulations were carried out with the TIP3P explicit solvent model[36] in an octahedral box truncated 11 Å from the protein surface. Chloride ions were added for charge neutralization of the system. Energy minimizations were achieved with an 8 Å cutoff for the non-bonded energy term for each atom. A subsequent 2 ns heat step to 310 K with the SHAKE algorithm fixed bond lengths involving hydrogen atoms. A 3.5 ns equilibration step preceded the 1 μs production runs. Simulations were run in triplicate under constant temperature (310 K, Langevin thermostat) and pressure (1 atm, Berendsen barostat) conditions using periodic boundaries. Trajectories were processed using the AMBER cpptraj tools.[37] The binding free energies and decomposition free energies were calculated using MMPBSA.py.[38]

Molecular graphics depictions and regression analyses

Standard molecular graphics were visualized with Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) Conservation maps were generated with UCSF Chimera[39] software. All regression analyses were performed in GraphPad Prism version 9.0.0 for Windows, GraphPad Software, San Diego, California USA. GraphPad functions used for curve-fitting are reported in the appropriate methods sections.

IC$_{50}$ assays of top candidates

Compounds were purchased from vendors including Enamine and ChemBridge. We resolved the stereochemistry of compound AR13 using 1D and 2D NMR, which supported a trans configuration at the cyclopropane moiety. In brief, very weak through-space coupling and a ^3J value of 3.4 supported protons with an uneclipsed dihedral (Fig. S1 and S2, ESI†).

Enzyme activity was measured by monitoring the conversion of exogenous substrate 7-methoxy-4-(trifluoromethyl)coumarin (MFC) to its fluorescent product 7-hydroxy-4-(trifluoromethyl)-coumarin (HFC). A CYP19/MFC high-throughput screening kit (Corning) was used to measure inhibition of aromatase activity for compounds AR11 and AR13. Briefly, 2× NADPH regeneration system (16.25 μM NADP$^+$, 825 μM MgCl$_2$, 825 μM glucose-6-phosphate, 0.4 U mL^{-1} G6P dehydrogenase) was prewarmed with inhibitors (1 : 2 serial dilutions) at 37 °C for 10 minutes in black 96-well plates. Reactions were initiated with prewarmed 2× enzyme–substrate mix (15 nM P450 microsomes enriched with oxidoreductase, 50 μM MFC) and incubated at 37 °C for 30 minutes. The 200 μL reaction mixtures were terminated with 75 μL 0.5 M Tris-base (80% ACN). A FluoDia T70 plate reader measured HFC product formation with excitation/emission filters of 400/530 nm. Reactions were performed in duplicate then repeated twice more on separate days. Assays were performed in-parallel with ketoconazole and endoxifen since the

© 2021 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2021, **2**, 892–905 | **893**

inhibitory activities are documented at the experimental conditions we detail here.[7,40]

The reaction conditions were repeated for measuring the inhibitory activity of compounds AR19 and AR20. Modifications are highlighted here. Supersomes containing CYP19 + CPR (Corning), NADPH regeneration system (Corning), MFC (Chemodex), and ketoconazole (Selleck Chemicals) were used for the reaction mixture. Temperature-controlled incubations of the 96-well plate were carried out with a dry-plate. A Tecan fluorescent plate reader measured the fluorescent product in the circle-read mode at the optimum gain with excitation/emission filters of 405/535 nm. Reactions were measured in duplicate.

Reaction blanks were used for data corrections and these results were normalized to the fluorescence response in the absence of inhibitor. Analyses were performed in GraphPad Prism 9 software, and data were fit to a 4-parameter (variable slope) logistic model on semilog axes.

Soret shifts by absorption spectroscopy

Codon-optimized CYP19A1 cDNA in the pCW expression vector was a generous gift from the F. Peter Guengerich Lab (Vanderbilt University). Preparation of the DNA construct is detailed by Sohl and Guengerich.[1] Recombinant protein was produced, purified, and characterized by methods detailed in the ESI,† section.

Purified CYP19 was diluted with 100 mM potassium phosphate buffer (pH 7.4) to 50 μL to a final concentration of 2–3 μM P450. Inhibitors were titrated such that the endpoint would not exceed 3% ACN. Absorption scans were read with a single-beam Agilent 8453 UV-Vis spectrophotometer after resuspension and a 10 minutes incubation period at 25 °C. Buffer A components, endoxifen, AR11, AR13, AR19, and AR20 did not contribute to hyperchromic shifts in the Soret peak region.

AR11 and AR13 kinetics assay

CYP19 + reductase microsomal preparations (BTI-TN-5B1-4, Corning) were used to measure the conversion of substrate MFC to HFC. Reactions contained 96 μL of NADPH regeneration solution and 4 μL of inhibitor from serial stocks. Mixtures were prewarmed at 37 °C in black 96-well round-bottom plates. Reactions were initiated with a single stream of 100 μL prewarmed enzyme–substrate solution in sequence. Parafilm and aluminum foil were applied to plates before incubation within the linear range at 37 °C for 20 minutes. The reactions were quenched with 75 μL of 0.5 M Tris-base (80% ACN). HFC was measured with a Tecan plate reader in circle-read mode at 52 gain with excitation/emission filters of 405/535 nm. Data was generated from 10 reads and a 40 μs integration time. Reaction mixtures contained 100 mM PPB (pH 7.4), 10 nM P450, 0.325 mM NADP$^+$, 0.825 mM glucose 6-phosphate, 0.825 MgCl$_2$, 0.1 U mL^{-1} glucose 6-phosphate dehydrogenase, MFC (9.9, 22.2, 33.3, and 50 μM), and various concentrations of inhibitor. Ligand sequestration was minimized by working at the lowest enzyme/microsomal concentration that would return a quantifiable fluorescence response. At steady-state conditions, the effect on MFC is negligible since the MFC concentration is greater than 1000-fold that of

the enzyme. We used an NADPH regeneration system to minimize the effects of NADPH depletion on the fluorescence response.

5×4 Lineweaver–Burk and Dixon plots were used to diagnose AR11 and AR13 inhibitory modes. Independent and dependent variables were precalculated and entered into GraphPad Prism 9 software for linear regression analyses. Nonlinear regression curves were fit to mixed- or competitive-type Michaelis functions to return kinetic constants defined by R. A. Copeland.[41]

Statistical analyses

Statistical analyses were performed in GraphPad 9.0.0 for Windows GraphPad Software, San Diego, California USA. The standard error of estimate (S$y.x$) was used as a measure of the goodness-of-fit. We report the ratio of this value from a 3-parameter to 4-parameter model for dose–response curves. Values greater than 1 infer a better fit to a 4-parameter logistic model, otherwise generating a curve with a Hill slope of 1 would result in underfitting. Nonlinear regression curves fit to noncompetitive, competitive, and mixed-type kinetic models were assessed by Akaike's method. Here the probability that a model is correct among two is quantified from a difference in Akaike's Information Criterion (AIC) value[42] for a least squares regression.[43] The standard error of the mean is reported for all activity constants and 95% confidence intervals were calculated for all reported means.

Results

Predicted CYP19–CPR interactions

P450s share a conserved fold with low sequence identity between different enzymes. Nearly all the conserved residues are in the buried regions. We used ConSurf, which identifies conserved surfaces,[24] to find four major surface-exposed sites that are conserved across different P450 families (Fig. 1).

Only two of the sites were predicted by the protein–protein interface algorithm, CPORT,[16] to actively participate in protein–protein interactions (sites 2 and 4 in Fig. 1). They correspond to the N-terminal region (αA′ and β1–2) and the proximal heme site. The likeliest CPR binding surface was selected based on three
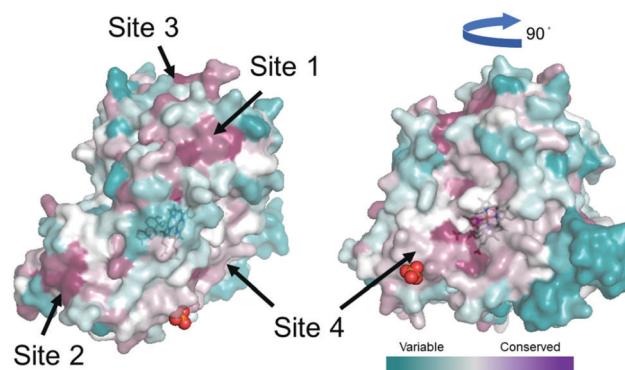
Fig. 1 Surface representation of variable and conserved regions of aromatase coded as a cyan to magenta color-gradient. The four conserved regions are labeled as Site 1–4. Site 2 is the N-terminus linked to the transmembrane segment. Site 4 corresponds to the proximal heme site.

criteria. Firstly, the proximity of the N5 atom of FMN to the iron center of the heme group should be within a distance that is physiologically sound.[44] Secondly, the orientation of the N-terminus of the reductase and CYP19 should be positioned in the same direction since they are truncations of transmembrane segments tethered to the cytoplasmic face of the endoplasmic reticulum.[45] Lastly, only the highest-scoring clusters from docking aromatase with CPR with Haddock[15] were assessed. Only site 4, corresponding to the proximal heme region, fulfilled all three criteria.

The probable binding mode contrasts with that of the crystal structure of *B. megaterium* P450BM3 fusion (PDB 1BVY). The N5 to iron distance in our structure is 14.4 Å, whereas the distance is 22.7 Å in P450BM3. This is within the 14–15 Å threshold limit for electron transfer in most physiological processes.[46] The possibility of through-bond tunneling at much longer distances in P450 BM3 was previously refuted due to faster experimental kinetic rates than predicted from theoretical models.[44] Physico-chemical descriptions of the modeled and fusion P450–CPR complexes are compared in Table S2 (ESI†). In docking CYP19 against CPR in a closed conformation (PDB 3QE2), the closest N5 to iron distance was 34.9 Å. The negative electrostatic potential of the FMN domain interacts with the positive potential of the FAD/NADPH domains. In this conformation, $FADH_2$ is in a closer proximity to reduce the FMN cofactor. In the proposed end-on interaction, 3FJO adopts an extended conformation to expose the buried FMN-binding interface for the reduction of CYP19 at the proximal heme site (Fig. 2 and Fig. S3, ESI†).

At the CYP19–CPR interface there are 17 total polar–polar, polar–charged, and charged–charged bond pairs. The most important of these contacts involve sidechains from the CYP19 residues K108, Y424, K440, and Y441 (Fig. 3 and Table S2, ESI†). Salt bridges between the proximal heme lysines K440/K108 and the reductase acidic residues D125/E73 are projected to be major contributors to enyme-coupling.

IC$_{50}$ values of top hits

The DSX (knowledge-based scoring function)[29] scores and chemical structures of the top four hit compounds are provided in Fig. 4 and Table S3 (ESI†).

Fig. 3 Predicted electrostatic contacts (yellow dotted trace) between CYP19 (cyan) and the CPR FMN domain (magenta). The FMN cofactor (pink) and heme (gray) are within the 15 Å limit of electron tunneling processes.

These compounds showed IC$_{50}$ values < 75 μM. AR13 showed 3-fold increased potency over the control inhibitor, ketoconazole, an antifungal with IC$_{50}$ = 3.08 μM. The IC$_{50}$ value we report for AR19, 72 μM, was fit to a Hill coefficient of 1 with a projected 11% activity at saturation. The S*y.x* value at 0.073 is suboptimal relative to the other inhibitors due to its poor solubility in buffer to interpolate a sigmoidal regression. There was no evidence of an inadequate model returned from a replicates test (Table S4, ESI†). AR11 and AR20 exhibited potencies comparable to an active tamoxifen metabolite, endoxifen. Dose–response curves and activity data are presented in Fig. 5 and Table 1. Values with 95% confidence intervals are provided in Table S4 (ESI†).

Fig. 2 Electrostatic potential surface map of the proposed end-on binding mode of CYP19 (PDB 4KQ8) in complex with CPR (PDB 3FJO) in an open conformation. Negatively charged potentials (red) to positive potentials (blue) are represented as a color gradient with neutral (gray) regions.
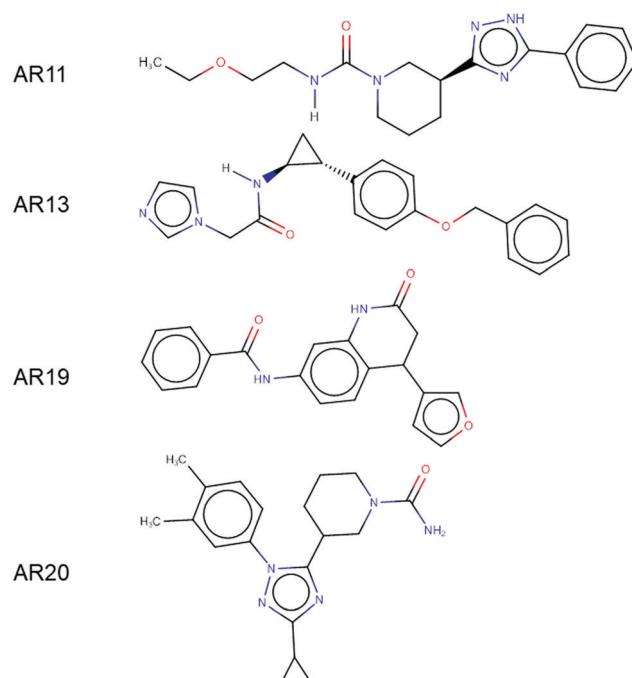
Fig. 4 Four most active aromatase inhibitors from virtual screening.

© 2021 The Author(s). Published by the Royal Society of Chemistry

*RSC Chem. Biol.*, 2021, **2**, 892–905 | **895**
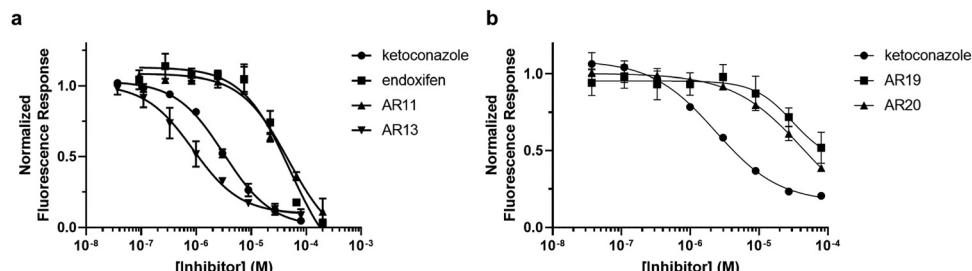
Fig. 5 Dose–response curves at 7.5 nM P450 and 25 μM MFC substrate with varying concentrations of inhibitors (a) AR11 and AR13, and (b) AR19 and AR20. Semi-log curves are fit to a 4-parameter logistic model with data from positive control inhibitors ketoconazole and endoxifen.

Table 1 Activity data collected from active compounds fit to a 4-parameter logistic model

Inhibitor	IC$_{50}$ (μM)	Hill slope	Fractional activity at saturation	Goodness of fit (S$y.x$)	S$y.x_{3p}$/S$y.x_{4p}$
Ketoconazole	3.1 ± 0.2	1.12 ± 0.07	0.03 ± 0.01	0.02	1.059
Endoxifen	30.9 ± 3.1	2.27 ± 0.41	0.02 ± 0.02	0.06	1.596
AR11	31.1 ± 5.4	1.54 ± 0.33	0.08 ± 0.02	0.06	1.103
AR13	0.8 ± 0.2	0.86 ± 0.15	0.07 ± 0.06	0.06	0.998
AR19[a]	72.9 ± 66.6	1	0.11 ± 0.42	0.07	0.976
AR20	42.8 ± 22.2	0.86 ± 0.15	0.03 ± 0.01	0.02	0.994

[a] 3-Parameter model reported with a Hill coefficient of 1.

A Hill slope of 2 for AR11 and endoxifen indicates that multiple binding sites are probable at saturation. The ratio of S$y.x$ values of the 3-parameter to the variable slope functions are 1.596 and 1.103, respectively. This indicates that the error of the residuals is reduced when the data is fit to a 4-parameter function where the Hill slopes are 2-fold greater.

Optical absorption properties of active compounds

P450s exhibit signature Soret peaks that are detectable in the 390–460 nm range.[47] Perturbations to the heme environment yield absorption shifts that are conserved. In the absence of substrate, water typically occupies the sixth site of the iron octahedral complex. This is evidenced by a Soret peak typically in the 415–417 nm range.[47] Interactions between the heme iron and stronger field ligands will induce absorption shifts at longer wavelengths. The addition of the native substrate, androstenedione (ASD), causes water displacement due to the C-19 protrusion from the ASD backbone.[48] This is indicated by a hypsochromic shift typically in the 390–394 nm range due to iron's adoption of a 5-coordinate high-spin state (5-CHS).[47] We report that the heme cofactor in recombinant CYP19 displays a Soret peak shift in 100 mM PPB (pH 7.4) at 416 nm (blue trace) and 395 nm in the absence and presence of ASD, respectively. Titration of AR13 induces an 8 nm bathochromic shift to 424 nm in the absence of ASD (Fig. 6a). This indicates an interaction with the iron by a stronger field ligand than water. Additionally, the enzyme may reversibly adopt the 5-CHS in the presence of AR13 with a 25 : 1 molar excess of ASD to inhibitor (Fig. 6b). Further, the CYP19–AR13 adduct can be reduced with dithionite to bind CO reversibly. The subsequent addition of AR13 fully recovers the 424 nm peak after 1 hour.

Titration of up to 100 μM of endoxifen did not induce a Soret peak shift from 395 nm in the CYP19 ASD-bound state (Fig. 7a).

This indicates that ASD remains in the active site in the presence of a 50-fold molar excess of endoxifen. This behavior is expected of a noncompetitive inhibitor where the K_i is unchanged in the presence of substrate. At a 3-fold (6 μM) molar excess, AR11 prompted the appearance of a peak shoulder near 416 nm. Fig. 7b illustrates the gradual increase of the peak shoulder at 416 nm with an increase in the concentration of inhibitor. This indicates that AR11 causes the enzyme to favor a shift to the 6-coordinate low-spin state and the displacement of ASD in the active site. At 100 μM AR11, the inhibitor-bound 6-coordinate low-spin state is apparent at 416 nm (Fig. 7d). The inset in Fig. 7d illustrates that AR11 causes CYP19 to favor the low-spin state in the absence of its native substrate. The Soret peak is red shifted 2 nm to 418 nm. This is likely a symptom of greater orbital overlap between the water ligand and iron induced by a readjustment of the active site.

AR19 and AR20 share the same Soret peak trends in the presence and absence of 2 μM ASD. Both compounds induce a gradual shift to the iron low-spin state indicated by simultaneous hyperchromic and red shifting towards 420 nm. This indicates that the population of CYP19 with iron in the 6-coordinate state increases in the presence of inhibitor. The red shift from 416 nm (blue trace) to 420 nm (red trace) is apparent in Fig. 8c and d. At 100 μM AR19, 2 μM ASD introduces a faint peak shoulder at 395 nm. Although there is a drop in the absorption at 420 nm, a peak shoulder is not apparent at this concentration. Higher concentrations of AR19 are required to induce the same Soret peak effects as AR20, suggesting that AR20 is a more potent inhibitor.

AR11 and AR13 steady-state kinetic assays

Kinetic assays of our two most potent inhibitors were performed under steady-state conditions to determine their Michaelis–Menten and inhibition constants (Fig. 9). These assays were
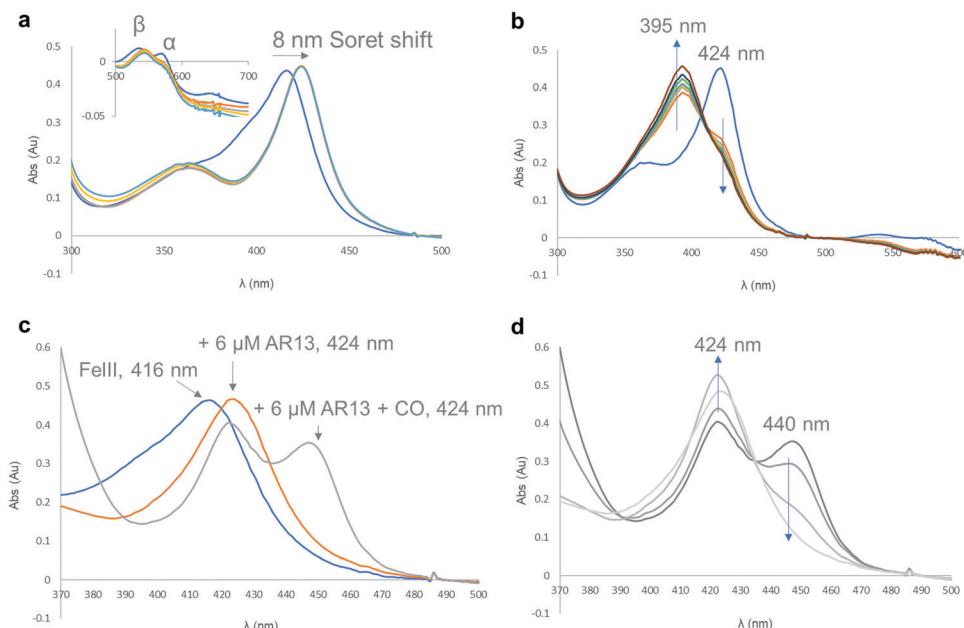
Fig. 6 Soret peak shifts in the presence of inhibitor AR13. (a) Bathochromic shift of Soret peak maximum from 416 to 424 nm upon titration with AR13 at 6, 12, 60, and 120 μM AR13. Inset shows a 5 nm shift in the β band from 535 to 545 nm and loss of the α band at 570 nm. (b) Time-dependent increase of P450 in the high-spin state after the addition of 70 μM ASD to a reaction mixture with 2.5 μM AR13. (c) P450 peak is observed after dithionite and CO addition to a reaction mixture of 3 μM P450 and 6 μM AR13. Subsequent addition of 6 μM AR13 results in the recovery of the 424 nm Soret peak and complete loss of CO-bound enzyme in a time-dependent manner (d).
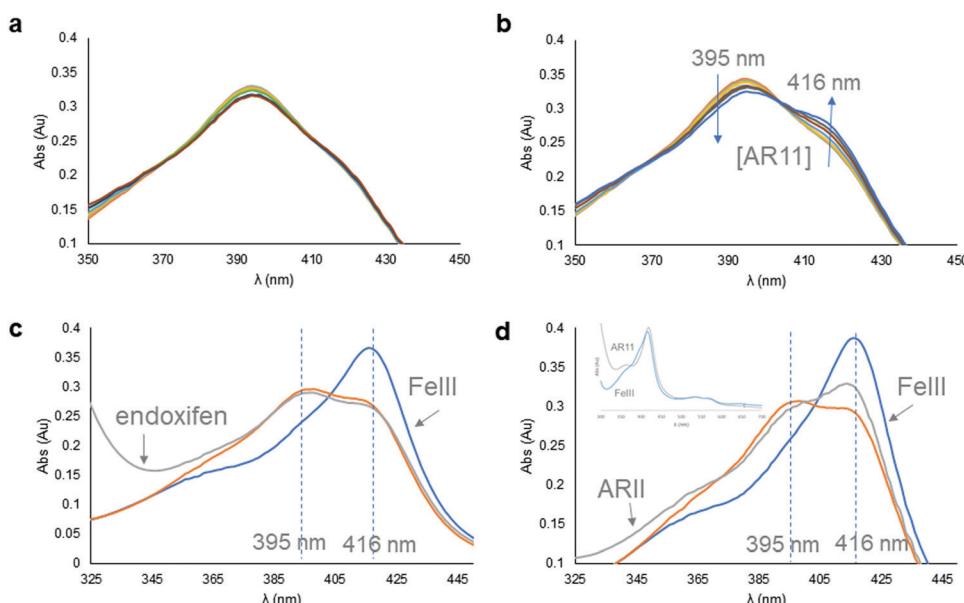
Fig. 7 Optical absorption spectra of reaction mixtures in the presence of 3 μM P450, 2 μM ASD, and various concentrations of endoxifen (a and c), and AR11 (b and d). Titration with inhibitor from 6–36 μM (a and b), and at 100 μM (c and d). The addition of AR11 favors a transition to the 6-coordinate low spin state indicated by a hyperchromic shift at 416 nm. Blue arrows in panel b indicate the absorption trend as the concentration of AR11 increases. Inset in panel d shows CYP19 in the absence and presence of 100 μM AR11.

used as a diagnostic tool to confirm whether AR11 and AR13 interact at the active site or a distinct site. They were performed at various concentrations of MFC up to $2 \times K_m$. Nonlinear regression analyses yielded substrate MFC V_{max} and K_m values of 0.731 pmol HFC min^{-1} pmol^{-1} P450 and 28.48 μM.

Lineweaver–Burk plots at various concentrations of AR11 yielded functions that intersected the y-axis at different inhibitor concentrations (Fig. 10a). These corresponded to different apparent V_{max} values, suggesting that it does not act competitively. In Fig. 10b, curves that correspond to AR13 concentrations

© 2021 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2021, **2**, 892–905 | **897**
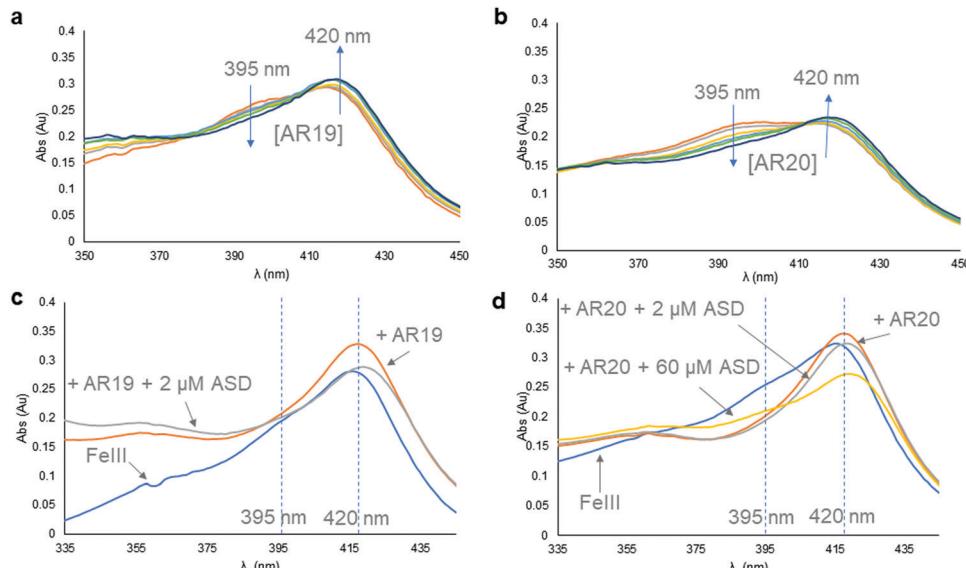
Fig. 8 Optical absorption spectra of reaction mixtures in the presence of 2 μM P450, 2 μM androstenedione (ASD), and various concentrations of AR19 (a and c), and AR20 (b and d). (a and b) Titration with inhibitor 6–36 μM favors a transition to the 6-coordinate low spin state. Blue arrows indicate the absorption trends as the concentration of inhibitor increases. (c and d) Addition of 100 μM inhibitor to CYP19 induces a 4 nm shift from 416 nm to 420 nm. A 395 nm peak shoulder is apparent after the subsequent addition of the native substrate, ASD.
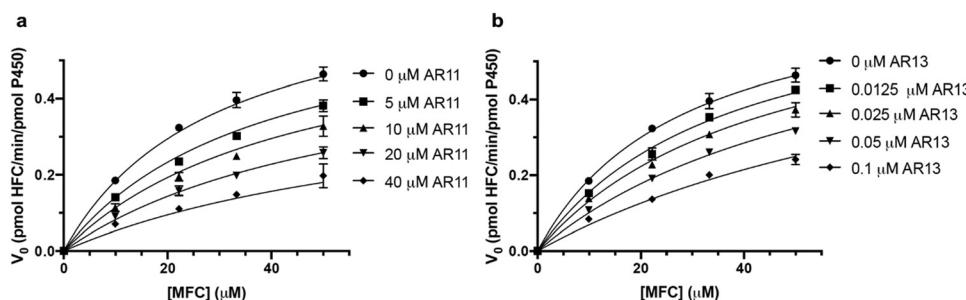
Fig. 9 Non-linear regression curves for AR11 (a) and AR13 (b) fit to single-site mixed and competitive-type Michaelis–Menten functions.

0–0.05 μM nearly intersect the *y*-axis at the same point. However, at 0.1 μM, this convergence is not apparent.

In Dixon-type analyses, the intersection of regression curves in Fig. 11 panels A and B correspond to the K_i values for AR11 and AR13, respectively. The intersection of the curves in panels C and D correspond to a K_i' or inhibitor dissociation constant in the presence of substrate MFC. This value infers non-mutual exclusive binding. Upon visual inspection of panel A, AR11 behaves noncompetitively at low MFC concentrations but exhibits competitive or mixed-type behavior as curves intersect above the *x*-axis at higher concentrations. Inclusive of an apparent V_{max} decrease (Fig. 10a) and a Soret peak shift to 418 nm at a 3:1 androstenedione to inhibitor ratio, AR11 behaves as a mixed-type inhibitor. A hallmark of this inhibition type is the intersection of curves below the *x* axis that would be extrapolated from Fig. 11c. An Akaike's model favored a nonlinear regression fit to a mixed model over a noncompetitive model. The ratio of these probabilities was 126. As such, AR11 was fit to a mixed-type nonlinear function.

AR13 exhibits similar trends as AR11 in a reciprocal velocity by inhibitor concentration plot (Fig. 11b). On the contrary, near-

parallel slopes in panel d suggests an indeterminate K_i' value whereby MFC and AR13 are mutually exclusive. This behavior is demonstrated at substrate concentrations near or greater than the MFC K_m of 28 μM. Data at 9.9 μM MFC perplexes this interpretation since at a low concentration (9.9 μM), AR13 appears to act noncompetitively. Isolated noncompetitive behavior at low substrate concentrations has been reported with the inhibitor-pesticide glyphosate while demonstrating mixed-mode inhibition at higher substrate concentrations.[49] Here, the difference between a hybrid-type mechanism and a single-type mechanism is more transparent. Competitive inhibitors should not preferentially interact at a distinct site at low substrate concentrations. Rather, the inverse may hold true. Distinguishing the difference between a noncompetitive and competitive inhibitor from linear transformations[50] is often unclear and has been documented in Dixon plots.[51] In an Akaike's test, a competitive model was favored over a noncompetitive fit with a 138 ratio of probabilities. Additionally, competitive-type behavior was apparent at a 1:2 enzyme:inhibitor ratio. Experiments at the lowest inhibitor concentration were carried out at a comparable ratio.
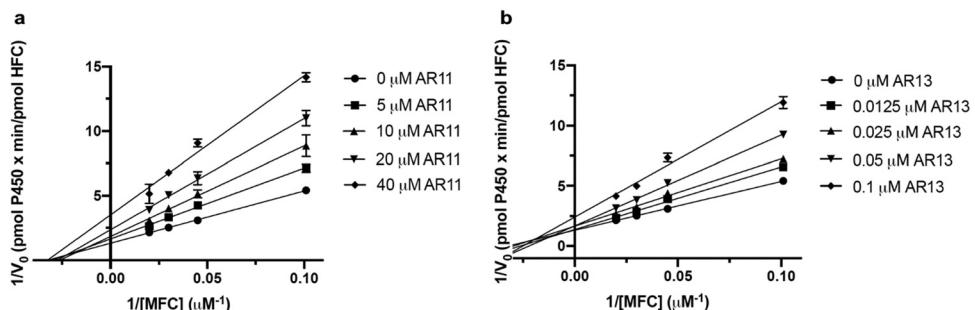
Fig. 10 Lineweaver–Burk plots at 10 nM P450 and various concentrations of substrate MFC with inhibitors AR11 (a) and AR13 (b).
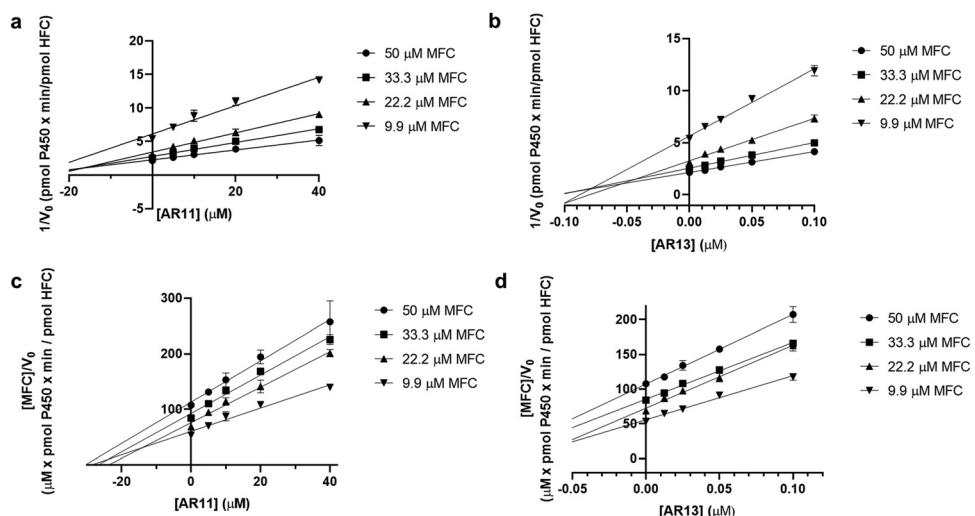
Fig. 11 Dixon-type inhibition plots for AR11 (a and c) and AR13 (b and d).

AR11 and AR13 were fit to nonlinear functions defined by mixed- and competitive-type kinetics, correspondingly. The K_i of AR13 is 42 nM, 39–57 nM for a 95% confidence interval. This is within an order of magnitude from the working enzyme concentration and may be an indicator of partial tight-binding behavior. In a Lineweaver–Burk plot, the $1/V_0$ values of tight-binders converge at high substrate concentrations and curvature is most recognizable at high inhibitor concentrations.[50] This may provide a rationale for a distinct V_{max} (y intercept) at 0.1 μM AR13 in Fig. 10b by which the data deviates from linearity.

Table 2 presents the effects of each inhibitor on these kinetic constants. The apparent K_m increases in the presence of inhibitor, indicating that more substrate MFC is needed to achieve $1/2V_{max}$. Meanwhile, only AR11 decreased the maximum velocity, indicating that MFC saturation will not retain full enzyme activity. Therefore, both substrate and AR11 bind CYP19 at discrete sites. In contrast, the ability to achieve the V_{max} with a molar excess of MFC over AR13 indicates that both compounds bind the active site. In Table 2, we report the AR11 K_i' as a product of the means of K_i and α. Propagation of the

Table 2 Steady-state kinetic constants in the presence of inhibitors AR11 and AR13 at 10 nM P450. Inapplicable quantities are indicated by a dash (—)

	$V_{max,app}$ (pmol HFC min^{-1} pmol^{-1} P450)	$K_{m,app}$ (μM)	R^2	K_i (μM)	α	K_i' (μM)
[AR11] (μM)				12.8 ± 1.8	5.0 ± 2.9	64.6
0	0.73 ± 0.03	29.4 ± 2.8	0.994			
5	0.68	38.0	0.993			
10	0.63	45.4	0.983			
20	0.56	57.6	0.989			
40	0.45	75.0	0.937			
[AR13] (nM)				42 ± 2 nM	—	—
0	0.73 ± 0.02	28.42 ± 1.62	0.995			
0.0125	—	36.78	0.994			
0.025	—	45.14	0.993			
0.05	—	61.86	0.996			
0.1	—	95.31	0.982			

Table 3 Docking and simulation metrics of endoxifen and novel CYP19 inhibitors

Ligand	Proximal heme site average dissociation time (ns)	Proximal heme site average binding free energy (kcal mol^{-1})	Substrate access average binding free energy (kcal mol^{-1})	Active site average binding free energy (kcal mol^{-1})
E-end	<200	−17 ± 4	—	—
Z-end	<500	−24 ± 3	—	—
AR11	>1000	−32 ± 4	−29 ± 8	—
AR13 (1R,2S)	<100	NR	—	−36 ± 4
AR13 (1S,2R)	—	—	—	−36 ± 2
AR19	400[b]	NR	−33 ± 3	—
AR20	800[a]	−22 ± 0[a]	—	—

[a] Denotes a value from 1 replicate run. [b] Denotes a value from 2 replicate runs. NR and a dash (—) indicate unreported and unmeasured quantities, respectively.

standard error of the means yields K_i' values between 0.6 and 1.8-fold greater. Kinetic constants with 95% confidence intervals are supplemented (Table S5, ESI†).

MD simulations of compounds docked against CYP19

Molecular dynamics simulations were performed to observe the residence time of the inhibitors in the proximal heme binding pocket. All our active compounds, except AR11, dissociated from the proximal heme site in less than 1 μs suggesting modest binding strength. AR13 exhibited the weakest interaction at the proximal heme site. It remained in the pocket for less than 100 ns in all replicates (Table 3).

Since the Soret shifts supported an iron–imidazole interaction, AR13 was docked to the catalytic site in 25 independent runs. The binding mode of each enantiomer yielding the closest iron–imidazole distance was used for MD simulations. They bound with free energies of −36 ± 4 kcal mol^{-1} (1R,2S) and −36 ± 2 kcal mol^{-1} (1S,2R) as calculated by MM-PBSA. Both enantiomers are projected to form a π–π interaction with F221 in the substrate access channel and a low energy π–π interaction between the terminal azole and the heme's porphyrin system. The Fe–N distances for each compound averaged 6 Å, indicating an indirect heme interaction.

More rigorous global redocking of AR11 located it in the substrate access channel more often than in the proximal heme site. AR11 formed π–π interactions with F221 and W224 and a hydrogen bond between the triazole and D309. This suggests that the triazole partially obstructs a space occupied by ASD since D309 is involved in a critical contact with the C3–carbonyl of androgens. Production runs yielded an average free energy of −29 ± 8 kcal mol^{-1}. In the proximal heme site, the free energy was −32 ± 4 kcal mol^{-1}, despite a destabilizing effect from E357. We report the putative binding modes at both the proximal heme site and substrate access channel in Fig. 12 since their binding energies have overlapping confidence intervals. We include the decomposition scores of the interactions with the greatest contributions to these energies.

AR19 and AR20 dock to the proximal heme site but dissociate after developing unfavorable interactions with E357, as demonstrated by positive decomposition scores (Fig. S4 and S5, ESI†). We also note that control inhibitors, E- and Z-endoxifen, disassociate due to positive free energy contributions from E357. In 1 out of 3 simulations, AR20 re-associates and remains in

the site >1 μs with −22 kcal mol^{-1} average binding free energy (Fig. S6, ESI†). We do not report decomposition scores for AR20 because it only transiently remains in a single binding mode throughout the production run. AR19 formed stable interactions in the substrate access channel with F221 and D309. Further, it had a proclivity to migrate closer into the active site for the duration of the production runs (Fig. 13a). The average free energy was −33 kcal mol^{-1}.

Discussion

MD simulations support that none of the inhibitors identified bind stably at the P450 reductase binding site. Rather, AR13 binds the enzymatic active site, as our inhibition kinetics data demonstrates, correlating well with the absorption data we present. The 8 nm Soret peak shift in the presence of AR13 indicates an interaction with the heme iron in the active site. Strong red shifts from a water-bound heme typify occupation of a strong, nitrogenous Sigma-donor ligand. This is likely an interaction with AR13's terminal imidazole moiety. A direct Fe–N interaction was confounded by our MD simulations that predicts a 6 Å distance. This would indicate a water-bridged state. Therein lies the possibility that there could be multiple energetically favorable 6-coordinate states in equilibrium. Although there are reports of P450-azole ternary complexes,[53,54] the compounds under study appear to be limited to triazoles that induce weaker Soret peak shifts.[53] Incontrovertible evidence warrants further study.

CYP19's ability to bind ASD and CO in the presence of AR13 demonstrates its capacity to maintain its functionality. Above all, the need for a 28-fold molar excess of ASD to induce a blue shift to 395 nm supports the inhibitor's high potency. Among all our inhibitors, AR13 was the most potent. This correlates with its predicted average binding energy being the greatest over AR11, AR19, AR20, and endoxifen. AR13 has a 40-fold lower IC$_{50}$ value than that of endoxifen and interacts with CYP19 on the same order of magnitude as norendoxifen (K_i = 35 nM). A major contributor to its potency could be through stabilizing a rigid protein-monomer. Substrate- and inhibitor-bound CYP19 complexes reduce the enzyme's intrinsic flexibility.[9,55] Additionally, AR13 interaction at the active site may favor disassociation of the reductase. We deduce this from a recent study whereby

a

AR11 group	Residue- Group	Decomposition score (kcal/mol)
Ph	F221- Ar	-2.00 ± 0.08
Az-NH	Q218- CO	-1.86 ± 0.03
Az	H480- Ar	-1.48 ± 0.05
Az-NH	D309- COO⁻	-1.05 ± 0.04
Ph	W224- Ar	-0.23 ± 0.00

b

| AR11 group	Residue- Group	Decomposition score (kcal/mol)
Az-NH	F418- CO	-2.09 ± 0.01
carbamide-N	Y361- OH	-1.65 ± 0.02
Az	N421	-1.48 ± 0.00
	E357	0.76 ± 0.00
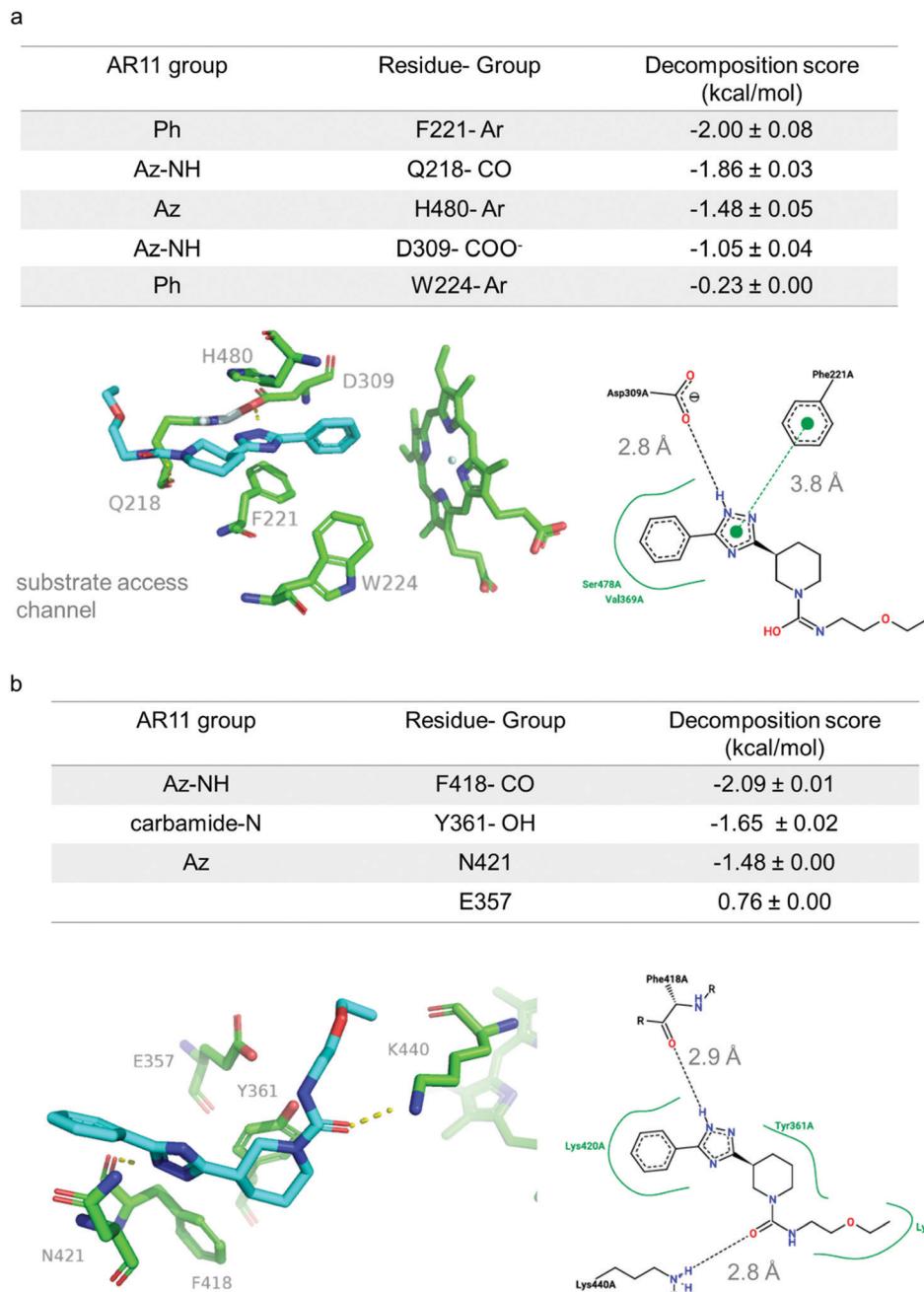
Fig. 12 3-D and 2-D interaction diagrams for AR11 with decomposition scores of the projected protein–ligand interactions at the substrate access channel (a) and the proximal heme site (b). 2-D depictions were generated with PoseView software.[52] Az-azole, Ph-phenyl, Ar-aromatic groups.

androstenedione reduced the reductase binding affinity by 5-fold.[56] We project that AR13 may be useful as a scaffold to design new NSAIs due to its remarkably distinct structure from current third generation NSAIs.

Titration with AR11 against androstenedione induces a peak shoulder between 414–416 nm, indicating that AR11 displaces ASD while substrate access to the heme iron is maintained. The average binding free energies at the proximal heme site and substrate access channel have overlapping confidence intervals and the Hill coefficient from a dose–response curve is greater than 1.5. As such, we project that both sites presented in this work are occupied at saturating concentrations of AR11. Our MD simulations suggest that AR11 preferentially interacts at the substrate access channel since E357 is projected to be destabilizing.

Note that our control inhibitor, E/Z-endoxifen, exhibits similar trends as AR11 in this study. Both inhibitors dock to the proximal heme site and yield a Hill coefficient that suggests the presence of multiple binding sites. However, there are two differences that infer distinct inhibitory mechanisms. Firstly, MD simulations predict that E- and Z-endoxifen do not remain in the proximal heme site, which is consistent with work

© 2021 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2021, 2, 892–905 | 901

a

AR19 group	Residue- Group	Decomposition score (kcal/mol)
Ph	F221- Ar	-3.33 ± 0.04
amide-NH	Q225- CONH$_2$	-2.02 ± 0.05
Ph	H480- Ar	-1.87 ± 0.01
lactam-CO	Q218- CONH$_2$	-1.47 ± 0.01
Ph	W224- Ar	-0.91 ± 0.01

b

AR20 group	Residue- Group	Decomposition score (kcal/mol)
Az-N	K440- NH^{3+}	-
carbamide-CO	Y361- OH	-
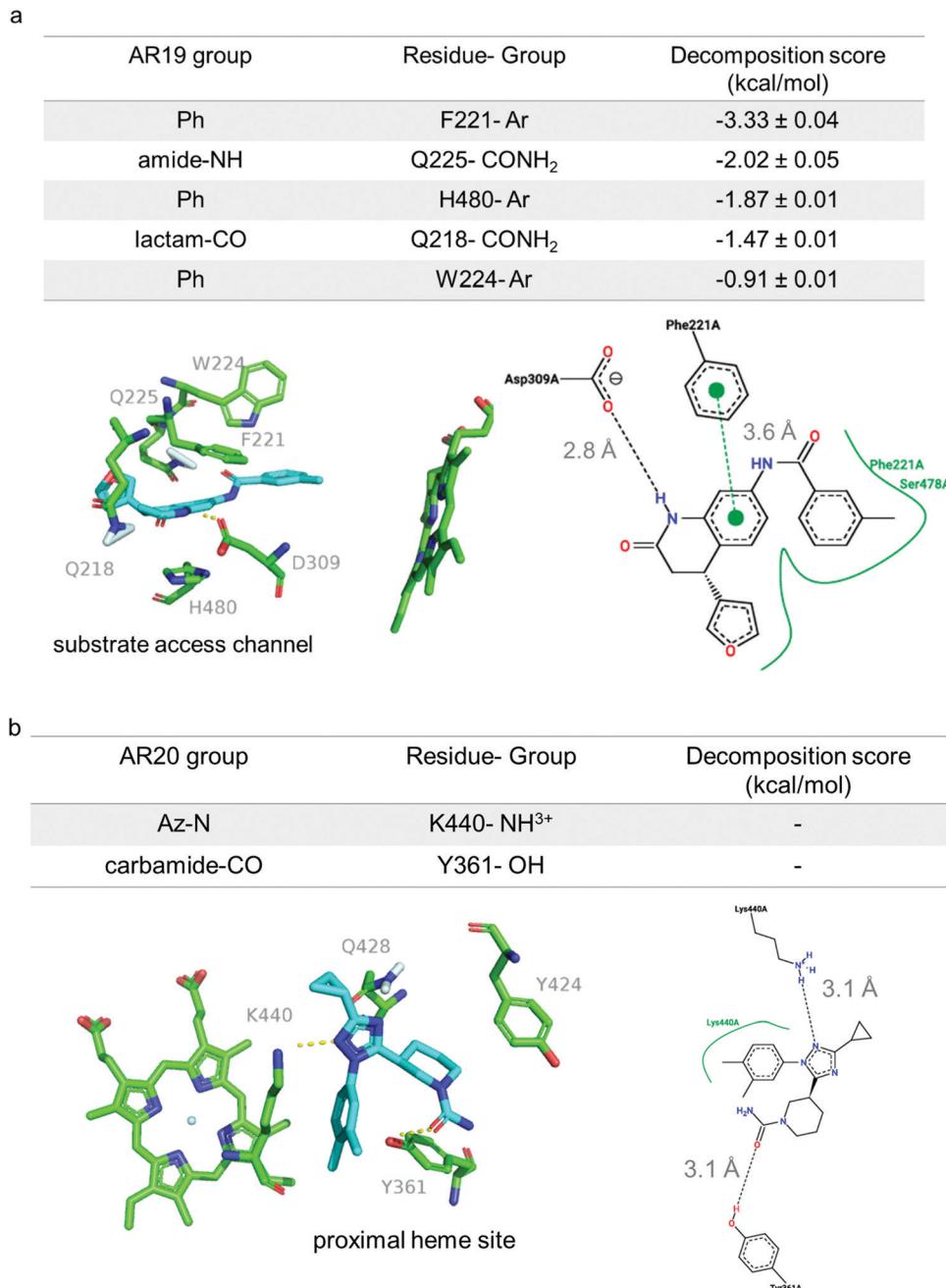
Fig. 13 3-D and 2-D interaction diagrams with decomposition scores of the projected protein−ligand interactions for AR19 (a) and AR20 (b). 2-D depictions were generated with PoseView software.[52] AR20 binding mode represented here is for replicate 2, and >280 ns. Decomposition scores are not available for AR20. Az-azole, Ph-phenyl, Ar-aromatic sidechain.

by Sgrignani *et al.*[9] Secondly, *E/Z*-endoxifen does not cause disassociation of ASD from the active site. It likely occupies a distinct locale, whereas AR11 protrudes into a space that is occupied by ASD.

Titration with AR19 and AR20 induces a 4 nm red shift from a water-bound iron to 420 nm. Production runs revealed that AR19 was mobile in the substrate access channel and active site. It has a terminal furan capable of directly interacting with the heme iron. Additionally, it may form a water-bridged ternary complex to induce greater water–iron orbital overlap.

AR20 is a 1,2,4-triazole that may also directly interact with iron or water. P450 ternary complexes have been reported with 1,2,3- and 1,2,4-triazoles.[53] Alternatively, heme perturbation induced by allosteric interactions may cause such Soret peak shifts. This would correlate with MD simulations that project transient AR20-proximal heme interactions.

Altogether, we conclude that *E/Z*-endoxifen, AR11, and AR19 preferentially occupy the substrate access channel, inhibiting CYP19 on the same order of magnitude. Since the substrate access channel is dynamic, compounds forming weaker

902 | *RSC Chem. Biol.*, 2021, **2**, 892–905

© 2021 The Author(s). Published by the Royal Society of Chemistry

interactions such as AR19 may migrate towards the active site to form a direct or ternary complex with iron. *E/Z*-endoxifen may bind the substrate access channel with high affinity and a second site at saturating levels of inhibitor. We expect that AR11 interacts at the substrate access channel and the proximal heme site at saturating levels. In the substrate access channel, AR11 can bind deeper into the pocket such that it partially occupies space by ASD near the catalytically active D309. AR20 is projected to transiently interact at the proximal heme site.

MD simulations project that the proximal heme site is acting as a low-affinity binding site for multiple inhibitors. It is an attractive target to modulate P450 activity because it is a conserved feature among P450s and distinct from the catalytic core. Despite class 2 P450s having less than a 40% sequence identity, all of them are expected to partner with CPR at this site. CYP19 may serve as a model to study its druggability for many reasons. Firstly, a crystal structure with the pentameric PEG bound at the proximal heme site is available and PEG inhibited enzyme activity in a time-dependent manner.[57] Secondly, by conformational selection, CYP19-reductase coupling affects the on and off rate of substrate at the active site.[56] Thirdly, CYP19 has a well-defined cavity of 584 Å^3, roughly 200 Å^3 larger than the active site, that can accommodate a larger library of compounds [57]. Lastly, the loop region between helices k'' and L contains the meander region of 21 residues long. Roughly 30–40% of P450s have a loop 14 or 15 residues long-all of which are class 2 P450 enzymes. This disparity offers a niche to selectively target CYP19 as less than 5% of P450s have a meander loop as lengthy.[58] It is likely that higher affinity inhibitors to the proximal heme site can be identified.

Other P450s that may be allosterically inhibited at the proximal heme site include CYP3A4 and CYP1A2. They are the major enzymes in drug metabolism and have well-documented heterotrophic effectors.[59–61] It is possible that these P450s can be modulated at the proximal heme site as the cavities are large enough to accommodate ligand-binding. The CYP3A4 and CYP1A2 pockets are 1.1 and 0.7 times the size of the CYP19 proximal cavity.[57] More research is needed to investigate the druggability of alternative ligand binding sites in P450s.[62–65]

In a broader context, nearly all exogenous compounds are subject to redox chemistry. Azoles and furan-containing xenobiotics, such as those presented here, are often P450 substrates.[66] Many of these are important azole-containing drugs used to treat fungal infections.[67] Profiling distinct chemical scaffolds can aid in predicting off-target P450 interactions that may lead to toxicity or adverse drug interactions.

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

References

1 C. D. Sohl and F. P. Guengerich, Kinetic analysis of the three-step steroid aromatase reaction of human cytochrome P450 19A1, *J. Biol. Chem.*, 2010, **285**, 17734–17743.

2 F. K. Yoshimoto and F. P. Guengerich, Mechanism of the third oxidative step in the conversion of androgens to estrogens by cytochrome P450 19A1 steroid aromatase, *J. Am. Chem. Soc.*, 2014, **136**, 15016–15025.

3 P. Fan and V. C. Jordan, New insights into acquired endocrine resistance of breast cancer, *Cancer Drug Resist.*, 2019, **2**, 198–209.

4 S. Maurelli, M. Chiesa, E. Giamello, G. Di Nardo, V. E. V. Ferrero, G. Gilardi and S. Van Doorslaer, Direct spectroscopic evidence for binding of anastrozole to the iron heme of human aromatase. Peering into the mechanism of aromatase inhibition, *Chem. Commun.*, 2011, **47**, 10737.

5 D. Ghosh, J. Lo, D. Morton, D. Valette, J. Xi, J. Griswold, S. Hubbell, C. Egbuta, W. Jiang, J. An and H. M. L. Davies, Novel aromatase inhibitors by structure-guided design, *J. Med. Chem.*, 2012, **55**, 8464–8476.

6 C. Egbuta, J. Lo and D. Ghosh, Mechanism of inhibition of estrogen biosynthesis by azole fungicides, *Endocrinology*, 2014, **155**, 4622–4628.

7 W. J. Lu, Z. Desta and D. A. Flockhart, Tamoxifen metabolites as active inhibitors of aromatase in the treatment of breast cancer, *Breast Cancer Res. Treat.*, 2012, **131**, 473–481.

8 W. J. Lu, C. Xu, Z. Pei, A. S. Mayhoub, M. Cushman and D. A. Flockhart, The tamoxifen metabolite norendoxifen is a potent and selective inhibitor of aromatase (CYP19) and a potential lead compound for novel therapeutic agents, *Breast Cancer Res. Treat.*, 2012, **133**, 99–109.

9 J. Sgrignani, M. Bon, G. Colombo and A. Magistrato, Computational approaches elucidate the allosteric mechanism of human aromatase inhibition: A novel possible route to small-molecule regulation of CYP450s activities?, *J. Chem. Inf. Model.*, 2014, **54**, 2856–2868.

10 A. Spinello, S. Martini, F. Berti, M. Pennati, M. Pavlin, J. Sgrignani, G. Grazioso, G. Colombo, N. Zaffaroni and A. Magistrato, Rational design of allosteric modulators of the aromatase enzyme: An unprecedented therapeutic strategy to fight breast cancer, *Eur. J. Med. Chem.*, 2019, **168**, 253–262.

11 J. Lo, G. Di Nardo, J. Griswold, C. Egbuta, W. Jiang, G. Gilardi and D. Ghosh, Structural basis for the functional roles of critical residues in human cytochrome P450 aromatase, *Biochemistry*, 2013, **52**, 5821–5829.

© 2021 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2021, **2**, 892–905 | 903

12 L. Aigrain, D. Pompon, S. Moréra and G. Truan, Structure of the open conformation of a functional chimeric NADPH cytochrome P450 reductase, *EMBO Rep.*, 2009, **10**, 742–747.

13 D. Campelo, T. Lautier, P. Urban, F. Esteves, S. Bozonnet, G. Truan and M. Kranendonk, The hinge segment of human NADPH-cytochrome P450 reductase in conformational switching: the critical role of ionic strength, *Front. Pharmacol.*, 2017, **8**, 755.

14 D. Hamdane, C. Xia, S.-C. Im, H. Zhang, J.-J. P. Kim and L. Waskell, Structure and function of an NADPH-cytochrome P450 oxidoreductase in an open conformation capable of reducing cytochrome P450, *J. Biol. Chem.*, 2009, **284**, 11374–11384.

15 G. C. P. van Zundert, J. P. G. L. M. Rodrigues, M. Trellet, C. Schmitz, P. L. Kastritis, E. Karaca, A. S. J. Melquiond, M. van Dijk, S. J. de Vries and A. M. J. J. Bonvin, The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes, *J. Mol. Biol.*, 2016, **428**, 720–725.

16 S. J. de Vries and A. M. J. J. Bonvin, CPORT: A consensus interface predictor and its performance in prediction-driven docking with HADDOCK, *PLoS One*, 2011, **6**, e17695.

17 S. Liang, C. Zhang, S. Liu and Y. Zhou, Protein binding site prediction using an empirical scoring function, *Nucleic Acids Res.*, 2006, **34**, 3698–3707.

18 H. Neuvirth, R. Raz and G. Schreiber, ProMate: a structure based prediction program to identify the location of protein–protein binding sites, *J. Mol. Biol.*, 2004, **338**, 181–199.

19 I. Kufareva, L. Budagyan, E. Rausch, M. Totrov and R. Abagyan, PIER: Protein interface recognition for structural proteomics, *Proteins: Struct., Funct., Bioinf.*, 2007, **67**, 400–417.

20 H. Chen and H.-X. Zhou, Prediction of interface residues in protein–protein complexes by a consensus neural network method: Test against NMR data, *Proteins: Struct., Funct., Bioinf.*, 2005, **61**, 21–35.

21 A. Porollo and J. Meller, Prediction-based fingerprints of protein–protein interactions, *Proteins: Struct., Funct., Bioinf.*, 2007, **66**, 630–645.

22 A. Vangone and A. M. Bonvin, Contacts-based prediction of binding affinity in protein–protein complexes, *eLife*, 2015, **4**, e07454.

23 L. C. Xue, J. P. Rodrigues, P. L. Kastritis, A. M. Bonvin and A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein–protein complexes, *Bioinformatics*, 2016, btw514.

24 H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko and N. Ben-Tal, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules, *Nucleic Acids Res.*, 2016, **44**, W344–W350.

25 M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures, *Nucleic Acids Res.*, 2005, **33**, W299–W302.

26 H. Li, K.-S. Leung, P. J. Ballester and M.-H. Wong, istar: a web platform for large-scale protein-ligand docking, *PLoS One*, 2014, **9**, e85678.

27 H. Li, K.-S. Leung and M.-H. Wong, *IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE, San Diego, CA, USA, 2012, pp. 77–84.

28 J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, ZINC: a free tool to discover chemistry for biology, *J. Chem. Inf. Model.*, 2012, **52**, 1757–1768.

29 G. Neudert and G. Klebe, DSX: A knowledge-based scoring function for the assessment of protein–ligand complexes, *J. Chem. Inf. Model.*, 2011, **51**, 2731–2745.

30 O. Trott and A. J. Olson, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.*, 2009, **31**, 455–461.

31 W. Humphrey, S. Schulten and A. Dalke, VMD: visual molecular dynamics, *J. Mol. Graphics*, 1996, **14**, 33–38.

32 D. A. Case, R. M. Betz, D. S. Cerutti, T. Cheatham, T. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Götz, N. Homeyer, S. Izadi, P. A. Janowski, J. Kaus, A. Kovalenko, T.-S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. E. Roitberg, C. Sagui, C. Simmerling, W. M. Botello-Smith, J. M. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao and P. A. Kollman, *Amber 16*, University of California, San Francisco, 2016, DOI: 10.13140/RG.2.2.27958.70729.

33 K. Shahrokh, A. Orendt, G. S. Yost and T. E. Cheatham, Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle, *J. Comput. Chem.*, 2012, **33**, 119–133.

34 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

35 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, Development and testing of a general amber force field, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

36 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**, 926–935.

37 D. R. Roe and T. E. Cheatham, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

38 B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, MMPBSA.py: an efficient program for end-state free energy calculations, *J. Chem. Theory Comput.*, 2012, **8**, 3314–3321.

39 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, UCSF Chimera-A visualization system for exploratory research and analysis, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

40 CYP19/MFC High Throughput Inhibitor Screening Kit, Corning Inc.

41 R. A. Copeland, *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*, John Wiley and Sons, Inc., 2nd edn, 2013.

42 H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control*, 1974, **19**, 716–723.

43 H. Motulsky, Interpreting nonlinear regression results, https://www.graphpad.com/guides/prism/latest/curve-fitting/reg_nonlinear_regression_results.htm, accessed 25 November 2020.

44 A. W. Munro, D. G. Leys, K. J. McLean, K. R. Marshall, T. W. B. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek, C. C. Moser, C. C. Page and P. L. Dutton, P450 BM3: the very model of a modern flavocytochrome, *Trends Biochem. Sci.*, 2002, **27**, 250–257.

45 F. P. Guengerich, Cytochrome P450 enzymes, *Am. Sci.*, 1993, **81**, 8.

46 C. C. Page, C. C. Moser, X. Chen and P. L. Dutton, Natural engineering principles of electron tunnelling in biological oxidation–reduction, *Nature*, 1999, **402**, 47–52.

47 A. Luthra, I. G. Denisov and S. G. Sligar, Spectroscopic features of cytochrome P450 reaction intermediates, *Arch. Biochem. Biophys.*, 2011, **507**, 26–35.

48 N. Kagawa, H. Hori, M. R. Waterman and S. Yoshioka, Characterization of stable human aromatase expressed in E. coli, *Steroids*, 2004, **69**, 235–243.

49 C. Zhang, T. Schilirò, M. Gea, S. Bianchi, A. Spinello, A. Magistrato, G. Gilardi and G. D. Nardo, Molecular basis for endocrine disruption by pesticides targeting aromatase and estrogen receptor, *Int. J. Environ. Res. Public Health*, 2020, **17**(16), 5664.

50 R. A. Copeland, *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, John Wiley and Sons, Inc., 2nd edn, 2000.

51 V. Kolhatkar and J. E. Polli, Reliability of inhibition models to correctly identify type of inhibition, *Pharm. Res.*, 2010, **27**, 2433–2445.

52 K. Stierand and M. Rarey, Drawing the PDB: Protein-ligand complexes in two dimensions, *ACS Med. Chem. Lett.*, 2010, **1**, 540–545.

53 K. P. Conner, P. Vennam, C. M. Woods, M. D. Krzyaniak, M. K. Bowman and W. M. Atkins, 1,2,3-triazole–heme interactions in cytochrome P450: Functionally competent triazole–water–heme complexes, *Biochemistry*, 2012, **51**, 6441–6457.

54 K. P. Conner, A. A. Cruce, M. D. Krzyaniak, A. M. Schimpf, D. J. Frank, P. Ortiz de Montellano, W. M. Atkins and M. K. Bowman, Drug modulation of water–heme interactions in low-spin P450 complexes of CYP2C9d and CYP125A1, *Biochemistry*, 2015, **54**, 1198–1207.

55 G. Di Nardo, M. Breitner, S. J. Sadeghi, S. Castrignanò, G. Mei, A. Di Venere, E. Nicolai, P. Allegra and G. Gilardi, Dynamics and flexibility of human aromatase probed by FTIR and time resolved fluorescence spectroscopy, *PLoS One*, 2013, **8**, e82118.

56 C. Zhang, G. Catucci, G. Di Nardo and G. Gilardi, Effector role of cytochrome P450 reductase for androstenedione binding to human aromatase, *Int. J. Biol. Macromol.*, 2020, **164**, 510–517.

57 D. Ghosh, C. Egbuta and J. Lo, Testosterone complex and non-steroidal ligands of human aromatase, *J. Steroid Biochem. Mol. Biol.*, 2018, **181**, 11–19.

58 D. Sirim, M. Widmann, F. Wagner and J. Pleiss, Prediction and analysis of the modular structure of cytochrome P450 monooxygenases, *BMC Struct. Biol.*, 2010, **10**, 34.

59 V. Polic and K. Auclair, Allosteric activation of cytochrome P450 3A4 *via* progesterone bioconjugation, *Bioconjugate Chem.*, 2017, **28**, 885–889.

60 C. M. Woods, C. Fernandez, K. L. Kunze and W. M. Atkins, Allosteric activation of cytochrome P450 3A4 by α-naphthoflavone: branch point regulation revealed by isotope dilution analysis, *Biochemistry*, 2011, **50**, 10041–10051.

61 Y. Satomi and H. Nishino, Inhibition of the enzyme activity of cytochrome P450 1A1, 1A2 and 3A4 by fucoxanthin, a marine carotenoid, *Oncol. Lett.*, 2013, **6**, 860–864.

62 K. W. Anderson, N. Mast, J. W. Hudgens, J. B. Lin, I. V. Turko and I. A. Pikuleva, Mapping of the allosteric site in cholesterol hydroxylase CYP46A1 for efavirenz, a drug that stimulates enzyme activity, *J. Biol. Chem.*, 2016, **291**, 11876–11886.

63 I. G. Denisov, Y. V. Grinkova, P. Nandigrami, M. Shekhar, E. Tajkhorshid and S. G. Sligar, Allosteric interactions in human cytochrome P450 CYP3A4: the role of phenylalanine 213, *Biochemistry*, 2019, **58**, 1411–1422.

64 H. Du, J. Li, Y. Cai, H. Zhang, G. Liu, Y. Tang and W. Li, Computational investigation of ligand binding to the peripheral site in CYP3A4: conformational dynamics and inhibitor discovery, *J. Chem. Inf. Model.*, 2017, **57**, 616–626.

65 M. Redhair, J. C. Hackett, R. D. Pelletier and W. M. Atkins, Dynamics and location of the allosteric midazolam site in cytochrome P4503A4 in lipid nanodiscs, *Biochemistry*, 2020, **59**, 766–779.

66 L. A. Peterson, Reactive metabolites in the biotransformation of molecules containing a furan ring, *Chem. Res. Toxicol.*, 2013, **26**, 6–25.

67 R. J. M. Brüggemann, J. C. Alffenaar, N. M. A. Blijlevens, E. M. Billaud, J. G. W. Kosterink, P. E. Verweij and D. M. Burger, Clinical relevance of the pharmacokinetic interactions of azole antifungal drugs with other coadministered agents, *Clin. Infect. Dis.*, 2009, **48**, 1441–1458.

© 2021 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2021, **2**, 892–905 | **905**