Induced fit describes ligand binding to membrane-associated cytochrome P450 3A4.†

David Tyler Sweeney, Francisco Zárate-Pérez, Kamila Stokowa-Soltys, and John C Hackett

Department of Chemistry and Biochemistry and Biomolecular Sciences Institute, Florida International University (JCH, KSS, FZP)

Department of Physiology and Biophysics and The Massey Cancer Center, Virginia Commonwealth University (DTS)

Department of Biological and Medicinal Chemistry, Faculty of Chemistry, University of Wroclaw (KSS)
Running title: Induced fit ligand binding to CYP3A4

Corresponding Author:

John C Hackett
Department of Chemistry and Biochemistry
Florida International University
11200 SW 8th St. Miami, FL 33122
Telephone: 305.348.4536
jhackett@fiu.edu

Number of Text Pages: 37
Number of Tables: 1
Number of Schemes/Figures: 7
Number of References: 46
Number of Abstract Words: 193/250
Number of Introduction Words: 617/750
Number of Discussion Words: 1255/1500

Nonstandard Abbreviations: Azamulin, AZM; Bromocriptine, BRC; Conformational Selection, CS; Cytochrome P450 3A4, CYP3A4; Induced Fit, IF; Mibrefadil, MIB; Midazolam, MDZ; Nanodisc, ND; One-step, OS; Pleuromutilin, PLM; Radius of gyration, Rg; Retapamulin, RPM; Singular Value Decomposition, SVD; Testosterone, TST
Abstract

Cytochrome P450 3A4 (CYP3A4) is the dominant P450 involved in human xenobiotic metabolism. Competition for CYP3A4 therefore underlies several adverse drug-drug interactions. Despite its clinical significance, the mechanisms CYP3A4 uses to bind diverse ligands remain poorly understood. Highly monodisperse CYP3A4 embedded in anionic lipoprotein nanodiscs containing an equal mixture of POPC and POPG were used to determine which of the limiting kinetic schemes that include protein conformational change, conformational selection (CS) or induced fit (IF), best described the binding of four known irreversible inhibitors. Azamulin, retapamulin, pleuromutilin, and mibefradil binding to CYP3A4 NDs conformed to a single-site binding model. Exponential fits of stopped-flow UV-vis absorption spectroscopy data supported multiple-step binding mechanisms. Trends in the rates of relaxation to equilibrium with increasing ligand concentrations were ambiguous to whether IF or CS was involved; however, global fitting and consideration of the rate constants favored an IF mechanism. In the case of mibefradil, a transient complex was observed in the stopped-flow UV-vis experiment, definitively assigning the presence of IF in ligand binding. While these studies only consider a small region of CYP3A4’s vast ligand space, they provide kinetic evidence that CYP3A4 can utilize an IF mechanism.
Significance Statement

CYP3A4 is capable of oxidizing numerous xenobiotics, including many drugs. Such promiscuity could not be achieved without conformational changes to accommodate diverse substrates. It is unknown whether conformational heterogeneity is present before (conformational selection) or after (induced fit) ligand binding. Stopped-flow measurements of suicide inhibitors binding to nanodisc-embedded CYP3A4 combined with sophisticated numerical analyses support that induced fit describes ligand binding to this important enzyme.
**Introduction**

The Cytochromes P450 (CYPs) metabolize hormones, xenobiotics, and have emerged as pharmacological targets for several diseases. (Rendic 2002, Ortiz de Montellano 2005, Rendic and Guengerich 2012) CYP3A4 metabolizes >50% of drugs, endogenous substrates, and environmental xenobiotics. (Williams, Hyland et al. 2004, Zanger and Schwab 2013) To accommodate the plethora of structural diversity in its ligand space, its active site must be malleable and the protein backbone flexible. This notion is supported by numerous crystal structures of the enzyme that illustrate a range of ligand-induced structural perturbations. (Yano, Wester et al. 2004, Ekroos and Sjogren 2006, Sevrioukova and Poulos 2010, Sevrioukova and Poulos 2012) Hydrogen-deuterium exchange mass spectrometry of CYP3A4 embedded in lipoprotein nanodiscs (NDs) has shown that ligands can induce distinct dynamics that spread over the entire protein fold. (Treuheit, Redhair et al. 2016, Redhair, Hackett et al. 2020) In view of the strong evidence for a continuum of CYP3A4 conformations, a one-step (OS) ligand binding mechanism is insufficient to describe the mechanism of molecular recognition. (Isin and Guengerich 2006, Isin and Guengerich 2007) There are two irreducible ligand binding mechanisms that account for protein conformational change, each differing as to whether the change precedes or follows the ligand binding event. In the former case, known as conformational selection (CS) (Eigen 1957, Boehr, Nussinov et al. 2009) the ligand preferentially binds to one of at least two conformational ensembles:

\[
E_1 \xleftrightarrow[k_{11}]{k_{12}} E_2 \xleftrightarrow[k_{21}]{k_{on}[L]} EL \xleftrightarrow[k_{off}]{k_{off}} E
\]

In the latter case, known as induced fit (IF) (Koshland 1958), the ligand binds to a single ensemble that subsequently induces a conformational change in the protein:
Vogt and Di Cera (Vogt and Di Cera 2012) have demonstrated that the presence of CS in the mechanism can be definitively assigned if a decrease in the relaxation rates ($k_{obs}$), as measured by stopped-flow methods, are observed with increasing ligand concentration. Conversely, they demonstrated that an increase in the relaxation rate with increasing ligand concentration, long accepted as a hallmark of IF, was at best ambiguous and could not be used to assign the operative mechanism. Guengerich and coworkers recently applied this approach to interrogate whether CS or IF best described ligand binding to several human CYPs, concluding that CS was the dominant binding mechanism for many, but the data for CYP3A4 was ambiguous. (Guengerich, Wilkey et al. 2019)

Distinguishing CS from IF in CYP binding mechanisms presents several challenges. (Di Cera 2020) First, both mechanisms assume that there are no equilibria between oligomers of the enzyme or account for potentially different characteristics to bind ligands. Mammalian CYPs, including CYP3A4, are associated with membranes of the smooth endoplasmic reticulum. CYP3A4 is not monodisperse in solution because those same protein domains that mediate lipid interactions nucleate oligomerization in solution. (Davydov, Fernando et al. 2005) Second, the irreducible kinetic schemes of IF and CS exclude the possibilities of multiple ligand binding and cooperativity, which are common features of ligand binding to CYP3A4. Indeed, multiple copies of both of the common probes midazolam, (MDZ, two) and testosterone (TST, three), bind to CYP3A4 NDs with positive homotropic cooperativity. (Khan, He et al. 2002, Denisov, Baas et al. 2007, Hackett 2018, Redhair, Hackett et al. 2020) Experimental designs that overcome these confounding factors are necessary to illuminate the fundamental mechanisms of ligand binding to CYP3A4.
To this end, we describe equilibrium and pre-steady state binding of four known irreversible inhibitors to CYP3A4 embedded in anionic lipoprotein nanodiscs (CYP3A4 NDs) to explore whether CS or IF better describe the ligand binding mechanism. The ND scaffold ensures both monodispersed CYP3A4 and affords ligand binding measurements in a native-like membrane environment. Azamulin (AZM), and mibrefadil (MIB) (Scheme 1) were selected based on observations of single ligand occupancy of active sites in published crystal structures (Sevrioukova 2019). RPM and PLM were selected based on their structural similarity to AZM as well as the recent observation that they bind to the related enzyme, CYP3A5. (Hsu and Johnson 2022) Herein, binding of all ligands to CYP3A4 NDs were confirmed to conform to a single site binding model. Pre-steady state kinetics measurements using stopped-flow UV-vis spectroscopy yielded trends in the relaxations that did not permit immediate assignment of CS or IF; however, global numerical fits to the kinetic schemes favored IF. The direct observation of a transient intermediate unequivocally demonstrates a contribution of IF to MIB binding to CYP3A4.

Materials and Methods

Chemicals and Reagents. AZM, RPM, PLM, and MIB were procured from Sigma-Aldrich (St. Louis, MO). All other reagents were of the highest grade available.

Protein expression and purification. A codon-optimized gene for CYP3A4 truncated at the N-terminus (Δ3-12) and a C-terminal 6×His was synthesized by Genscript and inserted between the NdeI and HindIII restriction sites of pCWOri. The resulting plasmid, pCW-3A4, was used to transform C41(DE3) E. coli cells that were subsequently plated onto LB agar plates containing 100 µg/mL ampicillin. A single colony was used to inoculate 100 mL of LB broth supplemented
with 100 µg/mL ampicillin and this culture was grown overnight at 37 °C. The overnight culture was used to inoculate 6 × 0.5 L of terrific broth supplemented with 100 µg/mL of ampicillin, 0.34g/L of thiamine and 0.25 mL/L of trace element mixture in 2.7 L baffled Fernbach flasks. Cultures were incubated at 37°C and 225 rpm until an O. D. at 600nm of 0.5-0.6 was reached. The flasks were removed from the shaker and cooled at room temperature before inducing expression with the addition of 1 mM IPTG and 1 mM 5-aminolevulinic acid (ALA). Following induction, the cultures were incubated at 27 °C and shaking at 125 rpm for 40 h. Cells were pelleted by centrifugation at 3500 × g for 25 min then resuspended in 4 mL/g cells of ice-cold lysis buffer (0.1 M potassium phosphate [pH 7.4], 20% glycerol, 3% Tergitol NP-11, 50 µM TST, 20 mM imidazole, 5 mM 2-mercaptoethanol). Protease inhibitor cocktail, DNAase, and 2 mg/mL lysozyme were added, and the mixture was stirred at 4 °C for 30 min. Cells were lysed by sonication on ice, stirred for an additional 30 min, then re-sonicated. The resulting lysate was centrifuged at 118,000 × g for 75 min. The red supernatant was pooled and loaded directly onto a HisPrep 16/10 FF Ni-NTA column (GE Healthcare) previously equilibrated with wash buffer 1 (0.05 M potassium phosphate [pH 7.4], 20% glycerol, 0.2% Anapoe C10E9, 50 µM TST, 0.3 M potassium chloride, 20 mM imidazole, and 5 mM 2-mercaptoethanol). The loaded column was then washed with five column volumes (CVs) of wash buffer 1, five CVs of wash buffer 2 (0.05 M potassium phosphate [pH 7.4], 20% glycerol, 0.2% sodium cholate, 50 µM TST, 0.3 M potassium chloride, 50 mM imidazole, and 5 mM 2-mercaptoethanol) and five CV of wash buffer 3 (0.05 M potassium phosphate [pH 7.4], 20% glycerol, 0.2% sodium cholate, 50 µM TST, 50 mM imidazole, and 5 mM 2-mercaptoethanol). Protein was eluted with elution buffer (0.05 M potassium phosphate [pH 7.4], 20% glycerol, 0.2% sodium cholate, 50 µM TST, 400 mM imidazole). Red-colored fractions were pooled and dialyzed overnight against 4 L of
storage buffer (0.1 M potassium phosphate [pH 7.4], 10 % glycerol, 1 mM EDTA, 0.5 mM TCEP). Protein was concentrated to 40-50 µM using a centrifugal concentrator with a 50 kDa MWCO. Aliquots of 0.1 µmole CYP3A were flash frozen in liquid nitrogen and stored at -80˚C until further use.

Membrane scaffold protein 1D1 (MSP1D1) (Denisov, Grinkova et al. 2004) was expressed in BL21 Gold (DE3) E. coli and purified and characterized as previously described (Hagn, Nasr et al. 2018, Sweeney, Krueger et al. 2022). TEV protease (Kapust, Tozser et al. 2001) was used to cleave the N-terminal 7-His tag from MSP1D1 to yield MSP1D1(-). Residual un-cleaved MSP1D1 was removed from the cleaved protein, MSP1D1(-), by capture on a HisPrep 16/10 FF Ni-NTA column.

Assembly of CYP3A4 NDs. Assembly of mixed lipid CYP3A4 NDs used a molar ratio of 0.1:1:43 CYP3A4: MSP1D1(-): lipids. The lipid mixture was equal parts 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1’-rac-glycerol) (POPG). Lipid films were obtained by drying chloroform solutions of the lipids under a N2 stream, then residual solvent was removed by placing them in a vacuum desiccator overnight. Lipid films were initially solubilized with a volume of 0.1 M sodium cholate/0.1 M NaCl solution such that the ratio of cholate:lipid was 2:1. Following the addition of CYP3A4 and MSP1D1(-), additional 0.1 M sodium cholate/0.1 M NaCl solution was added so that the final cholate concentration was 20 mM. Typically, preparations contained 0.1 µmole CYP3A4. The complete assembly mixtures were rocked at 4 °C for 1h. To initiate ND self-assembly, 1g/mL of washed Amberlite XAD-2 beads were added, and the mixture was rocked at 4 °C for an additional 3 h. Following filtration of the Amberlite, empty NDs were sorted from those containing CYP3A4 by loading the reaction mixture onto a 1 mL HisTrap HP Ni2+-NTA column.
previously equilibrated with 40 mM Tris-HCl (pH 7.4), 100 mM NaCl, and 5 mM 2-mercaptoethanol and subsequently washed with 10 CV of the same buffer. CYP3A4 NDs were eluted with same buffer containing 0.25 M imidazole. Fractions containing CYP3A4 NDs were pooled, concentrated to approximately 0.5 mL, and loaded onto a Superdex 10/300 GL Increase size exclusion column previously equilibrated with ND buffer (20 mM Tris-HCl [pH 7.4], 0.1 M NaCl, 5 mM 2-mercaptoethanol, 0.5 mM EDTA) to remove high molecular weight aggregates. ND homogeneity was confirmed prior to spectroscopic experiments using a Superdex 5/150 GL Increase size exclusion column.

Small-Angle Scattering of CYP3A4 NDs. Samples were shipped overnight to Lawrence Berkeley National Laboratory in 96-well plates at 4 °C. Triplicate sets of 30 μL CYP3A4 NDs samples were bracketed by measurements of two protein-free buffer samples to reduce error in subtraction. To ensure the ND-free reference buffers matched those of the corresponding ND samples, buffer flowing through the filters of the Amicon centrifugal devices (50 kDa MWCO) during concentration of the NDs were used. SAXS data were collected at the SIBYLS beamline (12.3.1) at the Advanced Light Source, part of the Lawrence Berkeley National Laboratory (Classen, Hura et al. 2013). Samples were transferred from a 96-well plate at 10 °C to the sample cuvette, where they are exposed to an X-ray beam for a total of 10 seconds (Dyer, Hammel et al. 2014). Scattering images were collected by a PILATUS3 2M detector every 0.3 seconds, for a total of 33 images per sample. The sample-to-detector distance was 1.5 m and the wavelength of the beam was 1.27 Å. Each collected image was circularly integrated, normalized for beam intensity, and buffer subtracted using software at the beamline (Classen, Rodic et al. 2010) to obtain \( I(q) \) vs \( q \) curves, where \( q = \frac{4\pi \sin(\theta)}{\lambda} \) and \( 2\theta \) is the scattering angle. Over the 10 s exposure, there was no evidence of radiation damage to the samples in the SAXS data, so all
buffer-subtracted frames were averaged to generate composite SAXS curves using PRIMUS. (Franke, Petoukhov et al. 2017) The radius of gyration and pair distance distribution function were determined using BayesApp. (Hansen 1991)

*UV-Visible Absorbance Analysis of Equilibrium Ligand Binding.* Measurement of UV-visible absorbance spectra was performed on a Olis Cary-14 Spectrophotometer Conversion (Olis, Inc Athens, GA). Baseline spectra containing only ND buffer were obtained by scanning the range of 350-700 nm. Stocks of concentrated ligands dissolved in ethanol were titrated into each cuvette containing 1 μM CYP3A4 NDs for AZM, RPM, and PLM and 1.8 μM CYP3A4 NDs for MIB, all in ND buffer. The final concentrations of ethanol in the samples after the addition of ligands did not exceed 1% (v/v). Difference spectra were calculated by subtracting the spectra collected in the absence of ligand from all spectra. The maximum absorbance difference, ΔA, at each ligand concentration were fit to $\Delta A = \Delta A_{\text{max}} [L][E]K_a/[E]_0$ while iteratively minimizing the values of $[E] + K_a[E][L] - [E]_0$ and $[L] + K_a[E][L] - [L]_0$ with *lsqcurvefit* and *fsolve*, respectively in MATLAB r2022a. $\Delta A_{\text{max}}$ is the maximum absorbance change, $K_a$ is the association constant, [E] is the free enzyme concentration, [L] is the free ligand concentration, and [E]$_0$ is the total enzyme concentration.

*Stopped-flow UV-vis spectroscopy.* Stopped-flow measurements were made with an OLIS RSM 1000 spectrophotometer equipped with a dual-syringe stopped-flow system maintained at 25 °C. For experiments using AZM, RPM, PLM, one syringe contained 2 μM CYP3A4 NDs while the other contained variable concentrations of the ligands, both dissolved in ND buffer. For the MIB experiments, one syringe contained 4 μM CYP3A4 NDs, while the other contained variable concentrations of MIB. CYP3A4 NDs and ligands were in ND buffer. Spectra were recorded from 330 to 580 nm at a scan rate of 1 ms$^{-1}$ for 3.8 seconds. The first two seconds of stopped
flow data were fit to sums of exponentials as well as globally fit to the systems of differential equations for CS and IF as previously described (Zárate-Pérez and Hackett 2020) in using *fmincon* and the *ode15s* solver in MATLAB r2022a with some differences. Specifically, the nonlinear equality constraints for CS and IF

\[
K_{d,CS}^{app} = \frac{k_{off}(k_{21} + k_{12})}{k_{on}k_{12}}
\]

\[
K_{d,IF}^{app} = \frac{k_{off}k_{21}}{k_{on}(k_{21} + k_{12})}
\]

were applied to limit the global fitting solutions to rate constants consistent with the $K_d$ values determined in the equilibrium binding measurements. In the fitting of data obtained with AZM, RPM, and PLM to the IF model, the intermediate E'L was assumed to be spectroscopically silent, $(\Delta \varepsilon_{E'L} = 0)$ and $\Delta A_{max}/[P] \cdot l$, where $[P]$ and $l$ are the ND concentration and path length, was defined as the lower bound for $\Delta \varepsilon_{EL}$. The upper bounds of first order rate constants were limited by the dead time of the stopped-flow instrument to 250 s$^{-1}$.(Di Cera 2020) The upper bound for the second order rate constant, $k_{on}$, was the diffusion controlled rate, $6.5 \times 10^8$ M$^{-1}$·s$^{-1}$.(van Holde 2002, Schreiber, Haran et al. 2009) When fitting data obtained with MIB, $\Delta \varepsilon_{EL}$ and $\Delta \varepsilon_{E'L}$ were fit alongside the rate constants with a positivity constraint as well as the nonlinear inequality constraint $\Delta \varepsilon_{EL} - \Delta \varepsilon_{E'L} < 0$. The Olis GlobalWorks (Olis, Inc., Athens, GA USA) software was used for singular value decomposition of the data obtained with MIB.

**Results**

*Assembly and Characterization of Anionic CYP3A4 NDs.* Two features of CYP3A4 that may confound interpretation of experiments aimed at delineating ligand binding mechanisms is its
propensity to oligomerize in solution and the possibility of cooperative binding of multiple ligands. The irreducible mechanisms of CS and IF implicitly assume that the enzyme is monodisperse, or that there are no oligomeric equilibria. The mechanisms also assume that a single binding site is present and only accommodates a single ligand. To ensure the preparation was monodisperse, CYP3A4 was incorporated into lipoprotein NDs containing an equal mixture of POPC and POPG. Analytical size exclusion chromatography of the terminal CYP3A4 ND preparation on a Superdex 5/150 GL Increase column resulted in a single peak eluting at 1.84 mL, between aldolase (77 kDa) and conalbumin (158k Da) standards eluting at 1.91 mL and 1.71 mL, respectively. A representative chromatogram is depicted in Figure 1A. Conversely, analytical size exclusion chromatography of CYP3A4 not incorporated into NDs suggests that the enzyme is primarily tetrameric under identical conditions. (Supplemental Figure S1).

Further confirmation was obtained by small angle x-ray scattering. The I(q) vs. q curve illustrated in Figure 1B is like that previously described for CYP3A4 NDs containing only POPC, with a pair of maxima at \(q\sim0.1\text{Å} \) and 0.18 Å.(Skar-Gislinge, Kynde et al. 2015) A fit of the low-q region \((qI(q) < 1.3)\) to the Guinier equation was linear and revealed a radius of gyration \(R_g\) (± S.E.M) of 48.0 ± 0.2 Å. (Figure 1B, inset) The corresponding pair distance distribution function (Figure 1B, inset) decayed to zero at a maximum distance (± S.E.M) of 136.5 ± 0.2 Å. The \(R_g\) (± S.E.M) derived from the reciprocal space analysis was 47.6 ± 0.0 Å, in close agreement to that obtained using the Guinier approximation. Both the \(R_g\) and \(D_{max}\) for the mixed POPC:POPG CYP3A4 NDs were similar to those previously reported for those containing 100% POPC.(Baas, Denisov et al. 2004, Skar-Gislinge, Kynde et al. 2015) Taken together with the analytical size exclusion chromatography, small angle scattering confirms that the CYP3A4 NDs are well-defined, homogeneous, and monodisperse particles.
Equilibrium ligand binding CYP3A4 NDs. To overcome the complexities of interpreting data obtained with ligands that bind to CYP3A4 cooperatively, we identified four ligands to exploit as probes to distinguish between CS and IF in CYP3A4. The binding of two ligands to CYP3A4, AZM and MIB, were previously shown to conform to a single site model and the crystal structures of the ligand complexes confirmed single occupancy of the active site. (Sevrioukova 2019) Binding of AZM, RPM, PLM, and MIB to CYP3A4 NDs were determined using UV-vis absorption spectroscopy that monitors shifts in the heme Soret band resulting from a low to high spin transition of the heme iron resulting from water displacement. (Figure 2) AZM, RPM, and PLM induced nearly complete high-spin (type I) transitions of the heme iron at saturation, whereas MIB induced only a partial shift. Ligand-dependent changes in the difference between 384 nm and 416 nm (ΔA\textsubscript{384-416}) were best fit to a one site model, thereby supporting occupancy of a single binding site by the ligands. Numerical fits were used to avoid the free ligand approximation as well provide the concentration of free ligand associated with each ΔA\textsubscript{384-416}. Fitted K\textsubscript{d} (± S.D.) values were 2.5 ± 0.02 µM, 1.6 ± 0.2 µM, 6.8 ± 1.3 µM, and 2.9 ± 0.5 µM for AZM, RPM, PLM, and MIB, respectively. The fitted values for the maximum absorbance change at ligand saturation, ΔA\textsubscript{max}, were 0.096, 0.087, 0.109, and 0.056. The ΔA\textsubscript{max} values for AZM, RPM, and PLM served as the lower bounds of the extinction coefficients in the global analyses of the stopped flow data. K\textsubscript{d} values of AZM and MIB were like previously reported values. (Sevrioukova 2019)

Stopped-flow UV-Vis Measurements of Ligand Binding. Equal volumes of CYP3A4 NDs and buffer solutions containing varying concentrations of AZM, RPM, PLM, and MIB were mixed in a stopped-flow system and the UV-vis absorption spectrum including the Soret band was collected at 1 ms\textsuperscript{-1}. The temporal dependence of ΔA derived from difference spectra were best
fit to the sum of two exponential functions, $\Delta A(t) = A e^{-k_{obs}^{fast} t} + B e^{-k_{obs}^{slow} t} + C$, for all ligands supporting that multi-step binding mechanisms are operative. (Supplemental Figure S2) The concentration dependence of the slow and fast relaxations, $k_{obs}$, are plotted in Figure 3. Both relaxations increase with increasing concentrations of AZM and RPM. Only the fast $k_{obs}$ is definitively increasing in the cases of PLM and MIB. The slow relaxation with PLM appears to increase; however, the values of $k_{obs}$ are the same within error at all PLM concentrations. In the case of MIB, the $\Delta A$ does not continuously increase at all ligand concentrations. Hence, assignment of the preferred mechanism was not possible by evaluation of the trends in $k_{obs}$, so we turned to a more sophisticated approach.

In view of the ambiguous $k_{obs}$ trends, the temporal changes in $\Delta A$ data at varying ligand concentrations were globally fit to the coupled differential equations for the CS and IF mechanisms. The solutions were limited to rate constants that reproduced the $K_d$ values measured in the equilibrium binding experiments. In global fitting of AZM, RPM, and PLM data to the IF model, we assumed that the intermediate state, $E'L$, did not contribute to ligand induced changes in absorbance ($\Delta e_{E'L} = 0$). If only the terminal bound state, $EL$, contributes to the absorbance change and this state is in equilibrium with a spectroscopically silent species, the apparent extinction coefficient observed at equilibrium serves as a lower bound for the extinction coefficient of pure $EL$. Hence, the $\Delta A_{max}$ values obtained from the equilibrium binding studies served as the lower bounds of the values of $\Delta e_{EL}$. Efforts were also made to fit the data to a sequential CS followed by IF model as previously described (Guengerich, Wilkey et al. 2019, Zárate-Pérez and Hackett 2020); however, despite the application of constraints, consistent, unique solutions to this sequential mechanism could not be found.
The results of global fitting to the CS and IF (top to bottom) for AZM, RPM, and PLM are illustrated in Figure 4. From visual inspection, both the CS and IF appear to describe the AZM and PLM data well; however, the IF is clearly a superior description of the RPM data. Simulations of the species concentrations provide additional insight into the model that more realistically describes the data. (Figure 5) Simulations of mixing of 1 μM CYP3A4 NDs with 100 μM of the ligands using the global fit rate constants indicate that in all cases that a favorable CS fit requires either E1 or E2 behave as a spectator to ligand binding. In contrast, the simulated species concentrations using rate constants derived from the IF model are more credible, where E’L accumulates to >20-60% of the population of the ligand bound states.

MIB offered evaluation of a ligand structurally distinct from the pleuromutilins to evaluate CS and IF. It was clear that when mixing higher concentrations (64 and 128 μM) of MIB with CYP3A4 NDs that an intermediate species with a greater high spin composition accumulates prior to 0.5 s then slowly decays thereafter. (Figure 6A) Since a spin transition can only be attributable to MIB interacting with the enzyme active site and perturbing the heme Soret band, these data are direct evidence of at least two ligand bound states, a distinguishing feature of IF. Accordingly, these data were globally fit to an IF model where both E’L and EL have flexible Δε with the constraint that Δε_E’L > Δε_EL to maintain consistency with the stopped-flow data. This approach yielded rate constants (Table 1) that nicely fit the data, correctly reproducing the initial accumulation and subsequent decay of a higher spin intermediate. (Figure 6A) While the MIB k_on and k_off rates are like those estimates for AZM and RPM, k_12 and k_21 are orders of magnitude smaller. A simulation of the species concentrations following mixing of 1 μM CYP3A4 NDs with 100 μM MIB illustrates that E’L, not EL, is the predominant species contributing to the ΔA during the stopped-flow experiment. (Figure 6B) Since the accumulation
and decay of E’L was most apparent in data obtained with 128 μM data, these individual datasets were subjected to SVD while simultaneously fitting to IF using the rate constants from the global fit as the initial conditions. SVD and fitting of the 128 μM data converged to rate constants $k_{on} = 0.25 \pm 0.02 \, \mu\text{M}^{-1}\cdot\text{s}^{-1}$, $k_{off} = 4.3 \pm 1.6 \, \text{s}^{-1}$, $k_{12} = 1.53 \pm 0.11 \, \text{s}^{-1}$ and $k_{21} = 0.03 \pm 0.02 \, \text{s}^{-1}$, like those obtained from the global fitting. The corresponding basis spectra of E, E’L, and EL are consistent with low-spin ligand-free CYP3A4, a nearly fully high-spin intermediate, and a mixed, but mostly low-spin terminal bound state, respectively. (Figure 6C)

**Discussion**

The most simplistic mechanisms of ligand binding invoking a conformational change, CS and IF, can give rise to complex kinetic behavior. (Chakraborty and Di Cera 2017) Hence, every effort should be made to ensure that the system being studied conforms to the implied features of the models to avoid confounding interpretation of the experimental data. Two major reasons make this challenging for CYP3A4. First, it tends to oligomerize in solution, which can result in multiple non-equivalent binding sites. This was overcome by embedding CYP3A4 in phospholipid NDs containing anionic lipids that like their empty congeners (Sweeney, Krueger et al. 2022), were shown to be highly monodisperse in solution by analytical size exclusion chromatography and small-angle x-ray scattering. Second, CS and IF assume binding of a single ligand to a single site. To date, the only type I ligand evaluated for this purpose is bromocriptine (BRC). (Isin and Guengerich 2006) All ligands evaluated herein for CYP3A4 binding conformed to a single site model, supporting occupancy of a single binding site. X-ray crystal structures of AZM and MIB with CYP3A4 are likewise consistent with single occupancy and the $K_d$ values obtained with CYP3A4 NDs are like those previously reported for the enzyme in solution. (Sevrioukova 2019)
In general, our results favor an IF mechanism for ligand binding to CYP3A4. Global fitting of the kinetic traces to the coupled differential equations describing each mechanism with constraints is perhaps the most rigorous method to distinguish between the possibilities. (Redhair and Atkins 2021) However, the results obtained with AZM and PLM illustrate that it is indeed possible to obtain good fits of both CS and IF to the data, making distinguishing between the ‘best’ model challenging. While IF was clearly a better descriptor of RPM binding, IF was judged to be a more credible fit for AZM and PLM based on the prediction that CS required that only a single conformer participate in the ligand binding event. The MIB data was represented by at least two ligand bound states and could only be fit to IF. To the extent that the pair of states represented by distinct perturbed heme spectra represent reorientation of the ligand and that such is coupled to protein conformational change, these data are perhaps the best kinetic evidence for a contribution to IF in ligand binding to CYP3A4.

CS has been reported to be the primary mode of ligand binding for many human P450s; however, the data for CYP3A4 has been more ambiguous. (Guengerich, Wilkey et al. 2019, Guengerich, Wilkey et al. 2019, Guengerich and Fekry 2020, Mast, Verwilst et al. 2020, Zárate-Pérez and Hackett 2020) Ours is not the first proposal for IF ligand binding to CYP3A4. Isin and Guengerich showed a 1:1 binding stoichiometry for BRC binding to solution CYP3A4 and stopped flow fluorescence experiments with multiple concentrations of BRC were best fit to a three-step binding model, with two of these dedicated to ligand reorientation/protein conformational changes. (Isin and Guengerich 2006) Stopped-flow studies of inhibitors binding to solution CYP3A4 also conformed to the same three step binding model. (Isin and Guengerich 2007) A stopped-flow fluorescence study of morphiceptin binding confirmed the presence an absorbance-silent binding step that did not perturb the heme Soret band position, consistent with
a spectroscopically-silent state like that invoked here for AZM, RPM, and PLM. Stopped-flow absorbance studies mixing multiple concentrations of clotrimazole, was also best described by the three-step model. Stopped flow data previously obtained with MDZ, TST, BRC, and ketoconazole (Isin and Guengerich 2006) and fit to the three step model were reevaluated in light of Vogt and Di Cera’s work (Vogt and Di Cera 2012); however, trends of both relaxations derived from double exponential fits were ambiguous to whether CS or IF was present and global fits to the corresponding differential equations both had deficiencies. (Guengerich, Wilkey et al. 2019) Separate experiments varying both [MDZ] and [CYP3A4] both resulted in increasing relaxation rates and their discordance was interpreted as evidence of CS for CYP3A4. Unfortunately, the use of MDZ as a probe to distinguish CS from IF has a limitation. The active site of CYP3A4 can sequentially accommodate two copies of MDZ and binding is positively cooperative. (Redhair, Hackett et al. 2020) In the case of positive cooperativity, the intermediate state EL is preferentially depleted by binding the second ligand, and the resulting EL₂ complex accounts for most of the heme Soret band perturbation measured in the stopped flow absorbance experiment. Hence, data obtained with MDZ reflect the second ligand distinguishing between the ligand free and the remaining singly occupied enzyme, rather than selection between a pair of interchanging ligand free conformers.

Despite the drawbacks of MDZ as a useful probe to distinguish ligand binding mechanisms in CYP3A4, the possibility that CS is present under some conditions cannot be discounted. Since CYP3A4 is oligomeric in solution, it is indeed possible that multiple conformations of CYP3A4 simultaneously exist within the oligomer. The presence of multiple conformations could result in a scenario where both CS and IF are present, but the former dominates and manifests in the stopped flow data. In a scenario where CS is dependent on
oligomerization, the mono-dispersity enforced by incorporating the protein into NDs would attenuate the CS component and unmask the contribution of IF to the binding mechanism. Testing this hypothesis will require the development membrane mimetic models capable of consistently incorporating homogenous systems of CYP3A4 oligomers.

While NDs are monodisperse and support the protein in a native-like membrane environment, it is possible that the lipid and scaffold protein create additional binding sites that may be transiently occupied prior to interacting with CYP3A4. (Nath, Grinkova et al. 2007) Such interactions would be spectroscopically silent in the stopped-flow UV-vis experiment. Although in the case of MIB, data support that the initial detectable interaction is with the enzyme. Since the $k_{on}$ values for the AZM and RPM are comparable to that for MIB, and that for PLM was greater, a preliminary interaction with the either ND does not occur or is too short lived to be detected by our experiments.

IF is largely consistent with our current understanding of how ligands enter the CYP3A4 active site. Spontaneous binding molecular dynamics simulations of TST binding to CYP3A4 embedded in a POPC bilayer reflect multistep binding. (Hackett 2018) In the predicted pathway, TST originates in the lipid headgroups, followed by occupation of an intermediate binding site where TST is flanked by the F’-helix, G’-helix, and the B-C loop in a configuration apparently stabilized by the Phe-cluster, and terminates in the active site above the heme. Transient occupation of the intermediate binding site is consistent with IF having a spectroscopically silent intermediate, since it is too distant from the heme to induce a shift of the Soret band. A recent crystal structure with fluorol bound provides experimental confirmation that the Phe cluster constitutes a high affinity binding site too distant from the heme to induce a spin shift. (Sevrioukova 2022)
In the crystal structure of the CYP3A4-MIB complex, the isopropyl and fluorophenyl moiety are oriented closest to the heme, yet CYP3A4 is known to catalyze N- and O-demethylation as well hydroxylation of the benzimidazole unit, which are too distant for metabolism in the crystallographic binding mode. (Welker, Wiltshire et al. 1998) Hence, structural and metabolism data likewise support significant reorientation and multiple binding modes for MIB in the CYP3A4 active site.

In conclusion, we analyzed the binding of four irreversible inhibitors of CYP3A4, selected based on their single occupancy of the enzyme active site, to monodisperse CYP3A4 NDs. As in previous analyses of CYP3A4 in solution, trends in $k_{obs}$ could not distinguish between CS and IF. Only global fitting to the differential equations and in some cases, subsequent consideration of the resulting rate constants could afford a diagnosis of IF. The direct observation of multiple MIB bound states is possibly the most convincing spectroscopic evidence of substrate reorientation in the CYP3A4 active site and is consistent with the presence of IF.

**Acknowledgements**

The authors are grateful to Michelle Redhair and William M. Atkins (University of Washington) for fruitful discussions.
Data Availability Statement

The authors declare that all the data supporting the findings of this study are available within the paper and its Supplemental Material.
Authorship Contributions

Participated in research design: Hackett and Sweeney.

Conducted experiments: Hackett, Stokowa-Sołtys, Sweeney, and Zárate-Pérez

Contributed new reagents or analytic tools: Sweeney and Zárate-Pérez

Performed data analysis: Hackett and Sweeney

Wrote or contributed to the writing of the manuscript: Hackett and Sweeney.
Footnotes

†This work was supported by the National Institutes of Health National Institute of General Medical Sciences [Grant R01GM135414]. The Research Excellence Initiative at the University of Wrocław supported K.S.S. while in the Hackett laboratory. This work was also conducted at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the National Institute of Health project ALS-ENABLE [P30GM124169] and a High-End Instrumentation Grant [S10OD018483]. No author has an actual or perceived conflict of interest with the contents of this article.
Figure Legends

Scheme 1. Structures of pleuromutilin antibiotics and mibrafadil used in this study.

Figure 1. A. Analytical size exclusion chromatography of CYP3A4 NDs containing 1:1 POPC:POPG using a Superdex 200 Increase 5/150 GL column. Absorbance was monitored at 280 nm (dotted line) and 414 nm (solid line). CYP3A4 NDs eluted at 1.84 mL. The vertical lines at 1.71 mL and 1.91 mL are the elution volumes of aldolase (158 kDa) and conalbumin (75 kDa) molecular weight standards on the same column. The CYP3A4 NDs consist of two MSP1D1(-) monomers, (22,044 Da each) and Δ3-12 CYP3A4×6His (56,996 Da). B. Small-angle x-ray scattering of CYP3A4 NDs containing 1:1 POPC:POPG. Guinier plot (lower left-hand corner) with residuals of the fit to the Guinier equation and the pair-distance distribution function (upper right-hand corner) are inset.

Figure 2. Determination of equilibrium binding of AZM (A), RPM (B), PLM (C) to 1 μM CYP3A4 NDs and MIB (D) to 1.8 μM CYP3A4 NDs. Data for pairs of titrations (open circles) and numerical fits to a single binding site (solid line) are shown for each ligand. Ligand concentrations are those of the free ligand concentration. The hypsochromic shift of the CYP3A4 Soret band with increasing concentrations of each ligand are inset into each panel. Each point in A-D represents the mean ± S.D. of triplicate experiments.

Figure 3. Fast and slow relaxations (k_{obs}) from fitting changes in absorbance induced by mixing varying concentrations of (A) AZM, (B) RPM, (C) PLM, and (D) MIB with 1 μM CYP3A4 NDs (2 μM CYP3A4 NDs for MIB) to the sum of two exponentials. Each point in A-D represents the mean ± S.D. of triplicate experiments.
**Figure 4.** Global fitting of stopped-flow UV-vis data obtained by mixing varying concentrations of (A) AZM, (B) RPM, and (C) PLM to the CS (first row) and IF (second row). In models including IF, species E’L was assumed to be spectroscopically silent.

**Figure 5.** Simulated concentrations of species resulting from mixing 1 \( \mu M \) CYP3A4 NDs with 100 \( \mu M \) (A) AZM, (B) RPM, and (C) PLM using the rate constants derived from global fitting to the CS (first row) and IF (second row). In simulations of CS, \( E_1 \) is depicted with a dotted line, \( E_2 \) with a dashed line, and EL with a solid line. In simulations of IF, \( E \) is depicted with a dotted line, E’L with a dashed line, and EL with a solid line.

**Figure 6.** (A) Global fitting of stopped-flow UV-vis data obtained by mixing varying concentrations of MIB with 2 \( \mu M \) CYP3A4 NDs to an IF model with spectroscopically observable E’L and EL, each having distinct extinction coefficients. (B) Simulated concentrations of species resulting from mixing 2 \( \mu M \) CYP3A4 NDs with 100 \( \mu M \) MIB. E is depicted with a dotted line, E’L with a dashed line, and EL with a solid line. (C). Representative basis spectra obtained by fitting data obtained by mixing 2 \( \mu M \) CYP3A4 NDs with 128 \( \mu M \) to a two-step, fully reversible IF model in Olis GlobalWorks using the global fit rate constants as initial conditions.
### Table 1. Global Fit Parameters to IF and CS Models

<table>
<thead>
<tr>
<th></th>
<th>$k_{12}$ (s$^{-1}$)</th>
<th>$k_{21}$ (s$^{-1}$)</th>
<th>$k_{on}$ ($\times 10^6$ M$^{-1}$s$^{-1}$)</th>
<th>$k_{off}$ (s$^{-1}$)</th>
<th>$\Delta \varepsilon^{EL}$ ($\times 10^{-2}$ μM$^{-1}$cm$^{-1}$)</th>
<th>$\Delta \varepsilon^{EL}$ ($\times 10^{-2}$ μM$^{-1}$cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AZM</td>
<td>174.6</td>
<td>8.2</td>
<td>0.2</td>
<td>0.4</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>RPM</td>
<td>12.2</td>
<td>105.2</td>
<td>2.3</td>
<td>0.4</td>
<td>8.4</td>
<td></td>
</tr>
<tr>
<td>PLM</td>
<td>16.9</td>
<td>171.0</td>
<td>3.5</td>
<td>1.5</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td><strong>IF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AZM</td>
<td>170.6</td>
<td>28.2</td>
<td>0.2</td>
<td>3.0</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>RPM</td>
<td>125.0</td>
<td>37.3</td>
<td>0.4</td>
<td>2.6</td>
<td>9.3</td>
<td></td>
</tr>
<tr>
<td>PLM</td>
<td>17.1</td>
<td>1.6</td>
<td>4.2</td>
<td>231.3</td>
<td>9.7</td>
<td></td>
</tr>
<tr>
<td>MIB</td>
<td>0.2</td>
<td>0.1</td>
<td>0.4</td>
<td>4.6</td>
<td>5.5</td>
<td>11.7</td>
</tr>
</tbody>
</table>
References


Figure 1.
Supplementary Material

Induced fit is describes ligand binding to membrane-associated cytochrome P450 3A4.†

David Tyler Sweeney, Francisco Zárate-Pérez, Kamila Stokowa-Soltys, and John C Hackett

Department of Chemistry and Biochemistry and Biomolecular Sciences Institute, Florida International University (JCH, KSS, FZP)

Department of Physiology and Biophysics and The Massey Cancer Center, Virginia Commonwealth University (DTS)

Department of Biological and Medicinal Chemistry, Faculty of Chemistry, University of Wroclaw (KSS)
**Figure S1.** Analytical size exclusion chromatography of CYP3A4 using a Superdex 200 Increase 5/150 GL column in nanodisc buffer (20 mM Tris-HCl [pH 7.4], 0.1 M NaCl, 0.5 mM EDTA, 5 mM 2-mercaptoethanol). Absorbance was monitored at 280 nm (dotted line) and 414 nm (solid line). Vertical lines (left to right, increasing retention volume) represent the elution volumes of ferritin (1.44 mL, 440 kDa), aldolase (1.71 mL, 158 kDa), conalbumin (1.91 mL, 75 kDa), and ovalbumin (2.02 mL, 44 kDa). The elution volume of CYP3A4 was 1.66 mL, corresponding to an estimated molecular weight of 204 kDa. This result supports that the CYP3A4 protein used in these studies is largely tetrameric in solution.
Figure S2. Two-exponential fits to the mean absorbance changes observed in SF mixing of ligands with CYP3A4 NDs. A. 1 μM CYP3A4 with 4 μM (blue), 8 μM (orange), 16 μM (yellow), 32 μM (purple), and 64 μM (green) AZM. B. 1 μM CYP3A4 with 4 μM (blue), 8 μM (orange), 16 μM (yellow), 32 μM (purple), and 64 μM (green) RPM. C. 1 μM CYP3A4 with 5 μM (blue), 10 μM (orange), 20 μM (yellow), 40 μM (purple), and 80 μM (green) PLM. D. 2 μM CYP3A4 with 4 μM (blue), 16 μM (orange), 32 μM (yellow), 64 μM (purple), and 128 μM (green) MIB.