SUPPLEMENTAL

Selection and suggestion of mutants

The selection of mutants was entirely based on the largest clusters from the single CBZ molecular dynamics simulation (Table S1). The prime notion was that the effect of methylation of small molecules, which is done in iterations of drug optimizations to increase potency, can also be achieved by increasing the hydrophobic surface of the protein, since the CBZ had to stay the same. Table S1 shows the centre of mass distances for the CBZ molecule in the largest cluster and the surrounding residues. While side chains featuring a hydrogen bond donor/acceptor were excluded from mutation due to unforeseeable influences on the charge distribution of the active site, residues with a hydrophobic side chain in a certain distance were considered. F residues (already at maximal hydrophobic surface) were not considered either. Only residues close to the 5H-dibenzo [b,f] azepine featuring a side chain which would be closer after mutation could potentially increase the contact of vdW interaction. Considering these criteria, only A370 and I369 were left for mutagenesis. The S119A mutant was suggested to disrupt the only present hydrogen bond and thus the overall stability of the binding mode. I301, A305, and G306 were not considered for mutation since an increase in vdW side chain was assumed to introduce steric clashes and thus obstruct the binding mode. I120 was on the side opposite to the 5H-dibenzo [b,f] azepine and too close to the hydrogen bond with S119 to be considered for mutation.

Table S1: Mean **distance of center of mass (COM) for the respective residues to the COM of carbamazepine in the largest cluster*.**

residue	mean	residue	mean
Arg105	8.97	Ala297	12.72
Arg106	14.09	Gln298	12.1
Pro107	14.41	Ser299	13.19
Phe108	9.46	Ile300	11.47
Gly109	14.28	Ile301	6.85
Phe110	16.28	Phe302	9.95
Val111	12.12	Ile303	11.43
Gly112	15.41	Phe304	6.01
Phe113	15.19	Ala305	5.45
Met114	10.86	Gly306	9.15
Ile118	9.82	Tyr307	11.96
Ser119	5.83	Glu308	8.28
Ile120	8.5	Thr309	7.85
Ala121	11.79	Leu366	20.52
Leu210	10.36	Phe367	16.9
Leu211	11.05	Pro368	13.16
Arg212	6.38	Ile369	8.01
Phe213	8.81	Ala370	8.74
Asp214	13.47	Met371	12.88
Phe215	13.62	Arg372	14.33
Leu221	20.75	Leu373	11.59
Ser222	22.02	Leu479	21.44
Ile223	18.59	Gly480	16.69
Thr224	19.09	Gly481	13.45
Val225	23.88	Leu482	10.76
Phe226	25.67	Leu483	13.48
Pro227	23.87	Gln484	14.24
Phe228	27.1	Pro485	17.97

* Based on the binding mode, the decision was made to investigate the increase of hydrophobic contact without introducing either steric clashes or changing the biochemical property of the residue (e.g. change a Thr to a Phe). This only left A370 and I369 for consideration since other residues were either too close to CBZ or already large enough in terms of vdW size of the side chain.

Derivation of an equation that combines an allowance for partial substrate inhibition with an approximation of cooperative interactions in terms of the Hill equation

The Hill equationwas first introduced in 1910 by A. Hill (Hill, A.V. (1910), J. Physiol. 40 (Suppl): iv–vii) to depict the interactions of hemoglobin with oxygen. Itis based on a simple model where the complex of a protein with ligand is formed via simultaneous binding of h molecules of ligand. Although the initial model considered the interactions of a protein oligomer (h-mer) with h molecules of ligand, the assumption of oligomerization of the protein is not necessary for this formalism to be applicable. The Hill equation in its modern representation may be obtained considering an equilibrium of simultaneous binding of h molecules of ligand S to protein E (either monomer or oligomer, or a subunit of an oligomer) with the formation of complex X:

$$E + hS \underset{}{\overset{K_D}{\rightleftharpoons}} X$$

The equilibrium constant K_D that defines the steady-state concentration of the complex $[X]$ may be represented as:

$$K_D = \frac{([E]_0 - [X]) \cdot [S]^h}{[X]} \tag{1}$$

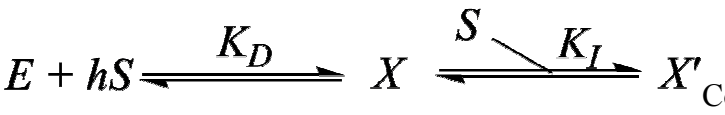
Resolving this equation for $[X]$ for the case where the total concentration of the ligand ($[S]_0$) is much higher than the total concentration of the protein ($[E]_0$), so that $[X] << [S]_0$ and $[S] \approx [S]_0$, we obtain:

$$[X] = \frac{[E]_0 \cdot [S]_0^h}{K_D + [S]_0^h}$$

Now introducing a new parameter S_{50}, defined as $S_{50} = \sqrt[h]{K_D}$ we get the Hill equation in its most common representation:

$$[X] = \frac{[E]_0 \cdot [S]_0^h}{S_{50}^h + [S]_0^h} \tag{2}$$

Let us complement this model with a binding of one more substrate molecule to the complex *X* resulting in the formation of complex X′:



Considering the case where $[S]_0 >> [E]_0$ so that $[S] \approx [S]_0$, we obtain the following relationship for the dissociation constant K_D of the complex *X*:

$$K_D = \frac{([E]_0 - [X] - [X']) \cdot [S]^h}{[X]} \tag{3}$$

The dissociation constant of the (inhibitory) complex *X*′ (K_I) may be defined as:

$$K_I = \frac{[X] \cdot [S]}{[X']} \tag{4}$$

Resolving the system of equations (3) and (4) for [*X*] and [*X*′] and introducing the parameter S_{50}, which is defined as $S_{50} = \sqrt[h]{K_D}$, we obtain the following equations:

$$[X] = \frac{K_I \cdot [E]_0 \cdot [S]^h}{[S]^h \cdot (K_I + [S]) + S_{50}^h \cdot K_I} \tag{5}$$

and

$$[X'] = \frac{[E]_0 \cdot [S]^{h+1}}{[S]^h \cdot (K_I + [S]) + S_{50}^h \cdot K_I} \tag{6}$$

Let's now consider the case when the protein E is an enzyme, ligand S is a substrate, the complex *X* is the enzyme-substrate complex with the maximal rate of catalytic turnover, and the complex *X*′ is an inhibitory complex, where the rate of catalysis is decreased due to the binding of an additional (inhibitory) molecule of substrate. We define the parameter α as a ratio of the turnover numbers (apparent catalytic constants) characteristic to the complexes *X* (k_{cat}) and *X*′ (k'_{cat}):

$$\alpha = \frac{k'_{cat}}{k_{cat}}$$

We may represent the overall rate of catalysis (*V*) as a function of [*X*] and [*X*′]:

$$V = k_{cat} \cdot [X] + k'_{cat} \cdot [X'] = V_{max} \cdot ([X] + \alpha \cdot [X']) \tag{7}$$

Combining equations (5), (6) and (7) we obtain:

$$V_S = \frac{k_{cat} \cdot [E]_0 \cdot [S]^h \cdot (\alpha \cdot [S] + K_I)}{[S]^h \cdot (K_I + [S]) + S_{50}^h \cdot K_I} \tag{8}$$

This equation is identical to the Equation 2 in the manuscript. Considering the case when cooperativity of the interactions is eliminated (h = 1) and replacing S_{50} with K_M we can obtain Equation 1 of the manuscript.

Similar to the regular Hill equation, equation (8) is based on a presumption that the binding of several ligand molecules to the protein takes place simultaneously. Due to this outermost simplification, the parameter h (the Hill coefficient) cannot be straightforwardly considered as a number of binding sites involved in the formation of complex X. However, similar to the case of the regular Hill equation, it may be shown that, for a mechanism with a positive cooperativity (h > 1), this parameter is less or equal to the actual number of binding sites involved in the interactions (see: A. Cornish-Bowden, Fundamental of Enzyme Kinetics, Butterworths, 1979, Chapter 7).

Distances of the CBZ C10,11 COM to the heme iron in MD Simulations with 2-4 CBZ molecules

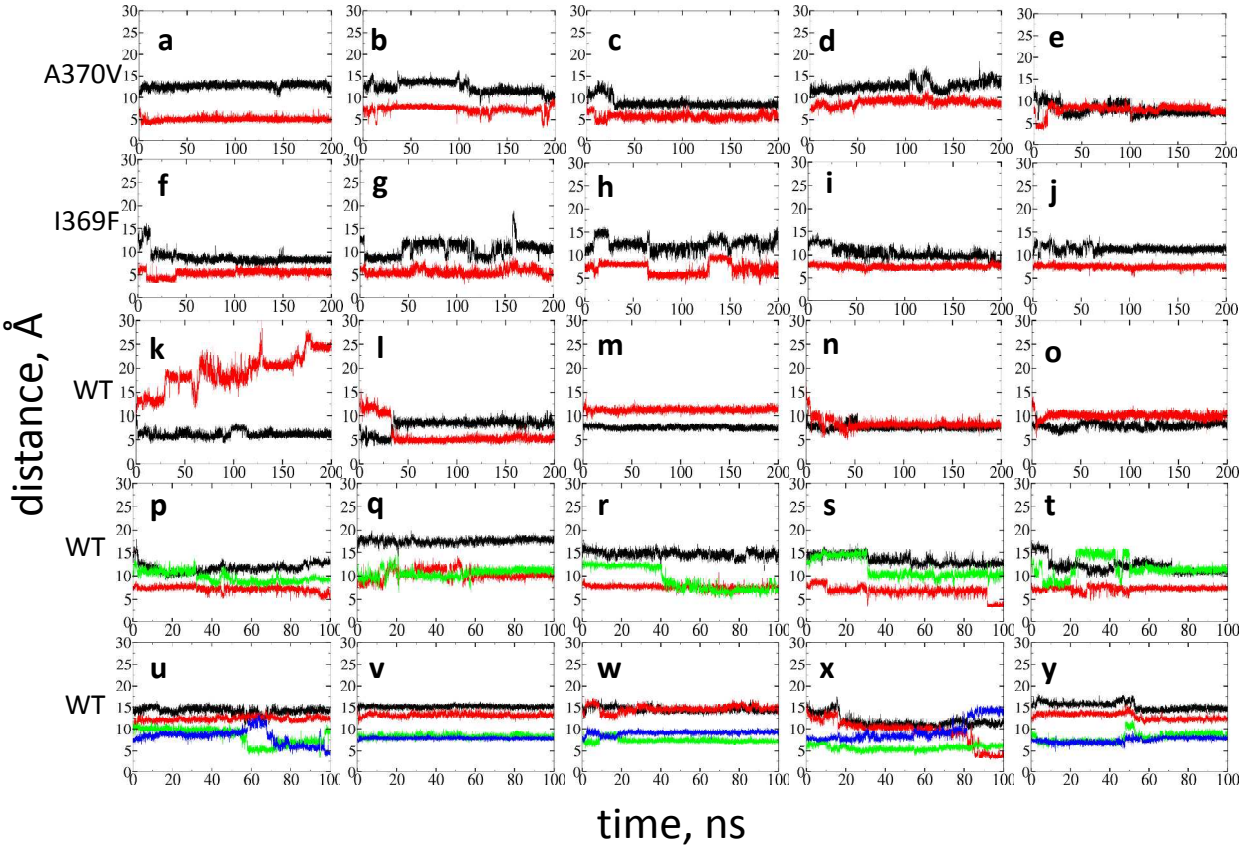


Figure S1: Time series of the distance between the center of the C10-C11 bond in CBZ and the heme iron. Each color corresponds to a single CBZ. Five 200 ns simulations were performed with two CBZs with the A370V mutant (a-e), the I369F mutant (f-j), and the wild-type CYP3A4 (k-o); five 100-ns simulations were performed with the wild-type with three (p-t) and four (u-y) CBZ molecules. All figures were generated with xmgrace. No full dissociation event is observed in the MD runs. The largest separation of CBZ from the heme is observed in panel (k) and a snapshot from the corresponding MD run is shown in Fig. 7 in the main section.

Distances between the side-chains of the residues 369 and Phe304

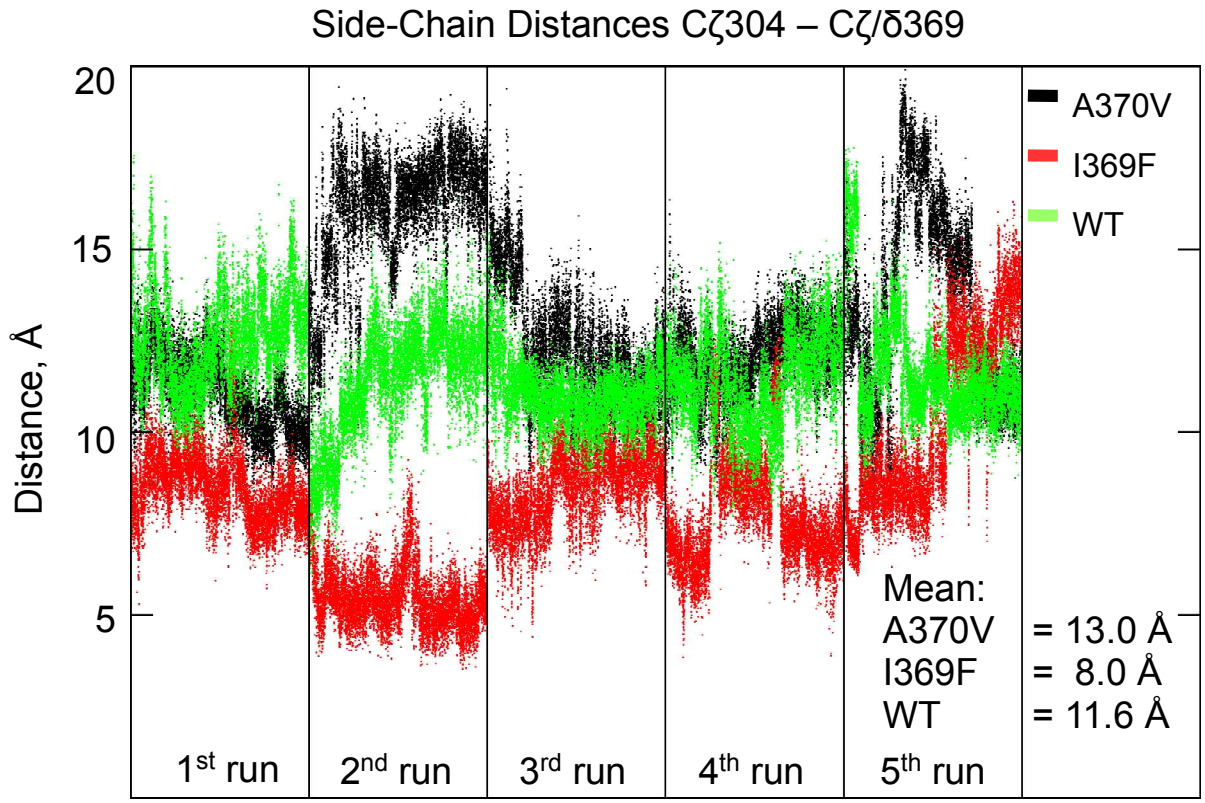


Figure S2: Time-series of distances between the side chains of residue 369 and Phe304 of the five simulations of a complex of two CBZ molecules with CYP3A4 wild-type (light green) and the mutants A370V (black) and I369F (red). The data shown here are based on the identical simulations as the evaluations of Fig. S1 (a-o).