

Whitepaper:

# Cryo-EM Structure of CYP3A4 Unliganded and In Complex with a Substrate

By Claudio Catalano, Lindsey Gottler, Vincent Shaw, Jason Haineault, Kevin Cannon, Ashlyn Farwell, Kyle Lucier, Reid Luksic, Arnold Muccini, Skerdi Senko

## Introduction

### P450 / CYP3A4

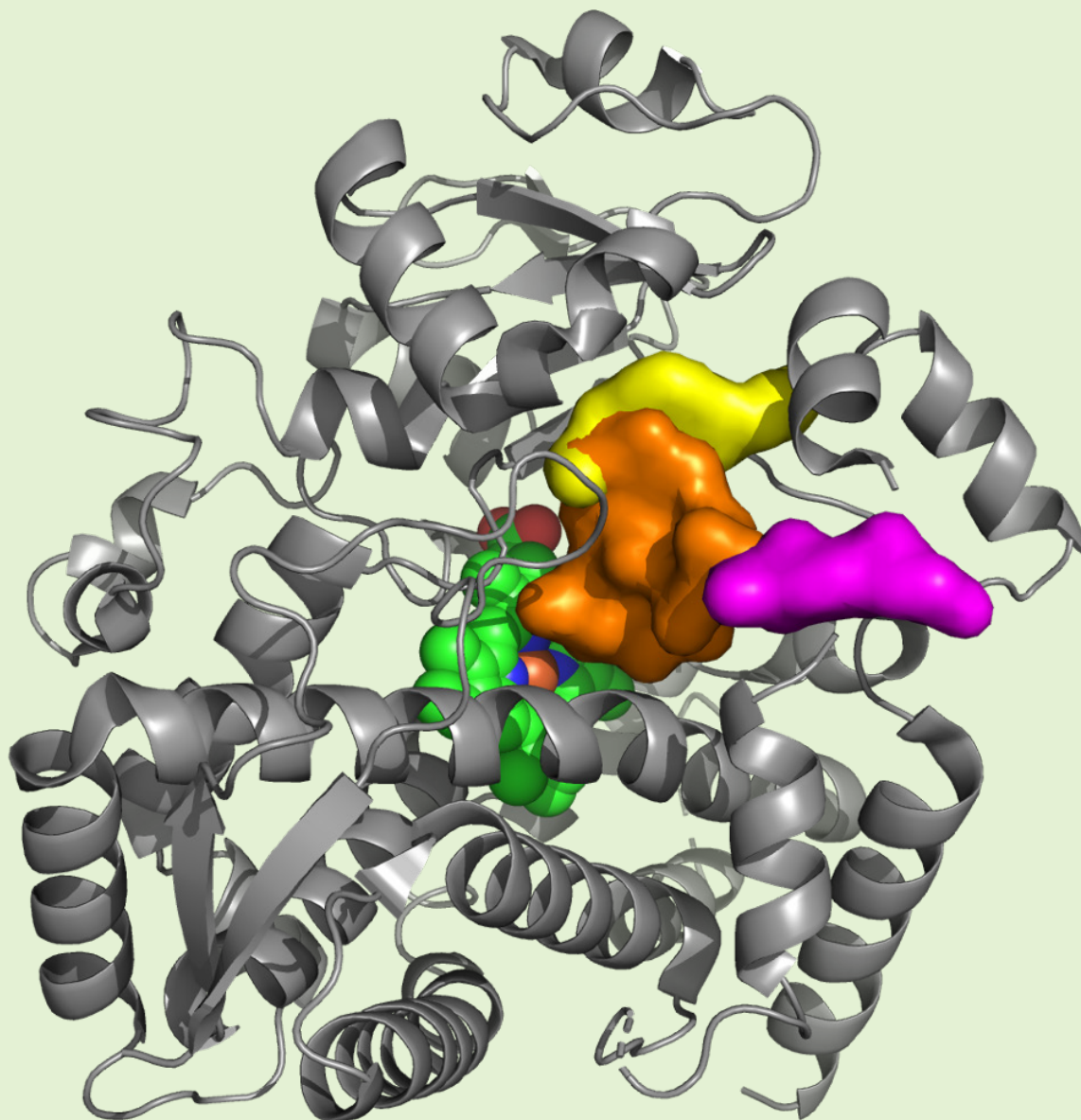
Cytochrome P450 enzymes (or CYPs) represent a diverse group of heme-containing proteins found in almost all organisms. Based on amino acid sequence similarity, CYPs are arranged into families (>40% identity) and subfamilies (>55% identity). Humans have 57 CYP genes, arranged into 18 families and 42 subfamilies. The 3A4 isoform (CYP3A4) is the most abundant and is responsible for metabolizing the majority of drugs. CYP3A4 shows extreme promiscuity in substrate specificity and binding, which often leads to undesirable drug-drug interactions and toxic side effects. CYP3A4 is a 57 kDa, single pass transmembrane protein that is highly expressed in the liver and small intestine<sup>1</sup>. Because of its role in drug development, CYP3A4 has been the most extensively studied mammalian P450 enzyme<sup>2</sup>. The three-dimensional structure of the soluble domain of CYP3A4 has been extensively documented.

In the PDB, there are over 120 crystal structures of the enzyme in complex with a variety of substrates and inhibitors aimed at understanding the complexity of its large and highly flexible active site, and the ability to accommodate more than one molecule in or near the substrate-binding pocket where they can act either as substrates or as modulators of substrate metabolism. The mammalian CYP structures all adopt the general CYP fold first described in 1987<sup>3</sup> and are known to undergo notable conformational changes following binding of substrates and inhibitors. This wealth of structural information led to the identification of three major binding sites in P450: the active region, a substrate access channel, and an outer surface site known as the peripheral site<sup>4</sup> (**Figure 1, page 2**). The most common biological assembly suggested for crystalline CYP3A4 is monomeric, although higher order oligomers are sometime present in the crystals' asymmetric unit. In vivo, oligomerization of P450 on the microsomal membrane surfaces has been proposed to have functional consequences<sup>5</sup>.

To learn about our cryo-EM services, visit [nanoimagingervices.com](https://nanoimagingervices.com) or email [info@nanoimagingervices.com](mailto:info@nanoimagingervices.com)

For more about protein production at Proteos, visit [proteos.com](https://proteos.com) or email [info@proteos.com](mailto:info@proteos.com)

Given the role of CYPs in drug metabolism and the potential liabilities they may contribute to, CYP activity of drug candidates is routinely tested during the drug discovery process. While in many cases these off-target effects can be resolved by a combination of extensive screening chemical, and/or computational approaches, access to structural information (such as the 3D structure of the target bound to the off-target of interest) may significantly reduce the time required to de-risk the candidate. This is the process we refer to as structural toxicology.



**Figure 1** Cyp3A4 structure with the location of the three major binding sites: the active site (orange), the substrate access channel (yellow) and the peripheral site (magenta). The protein is shown as grey ribbon; the heme as green spheres

### Why Cryo-EM?

A wealth of structural information regarding P450 in general, and CYP3A4 in particular, has been obtained thus far using X-ray crystallography. The truncated version of the protein used in these structural studies is soluble, medium-sized and amenable to crystallization experiments. We recently reported<sup>6</sup> efforts to establish reproducible cryo-EM workflows for samples below 80 kDa, the value generally accepted as the lowest molecular weight limit for routine cryo-EM single particle analysis (SPA) work. Given its relatively low molecular weight (57 kDa), CYP3A4 was considered a good candidate to test our processes, and we decided to add it to our list of “small proteins” solved by cryo-EM. In addition, cryo-EM allows for visualization of flexibility, oligomeric states, and conformational changes in the sample that would not be allowed in a crystallographic setting, thus potentially providing more insights into the mechanism and activity of the sample.

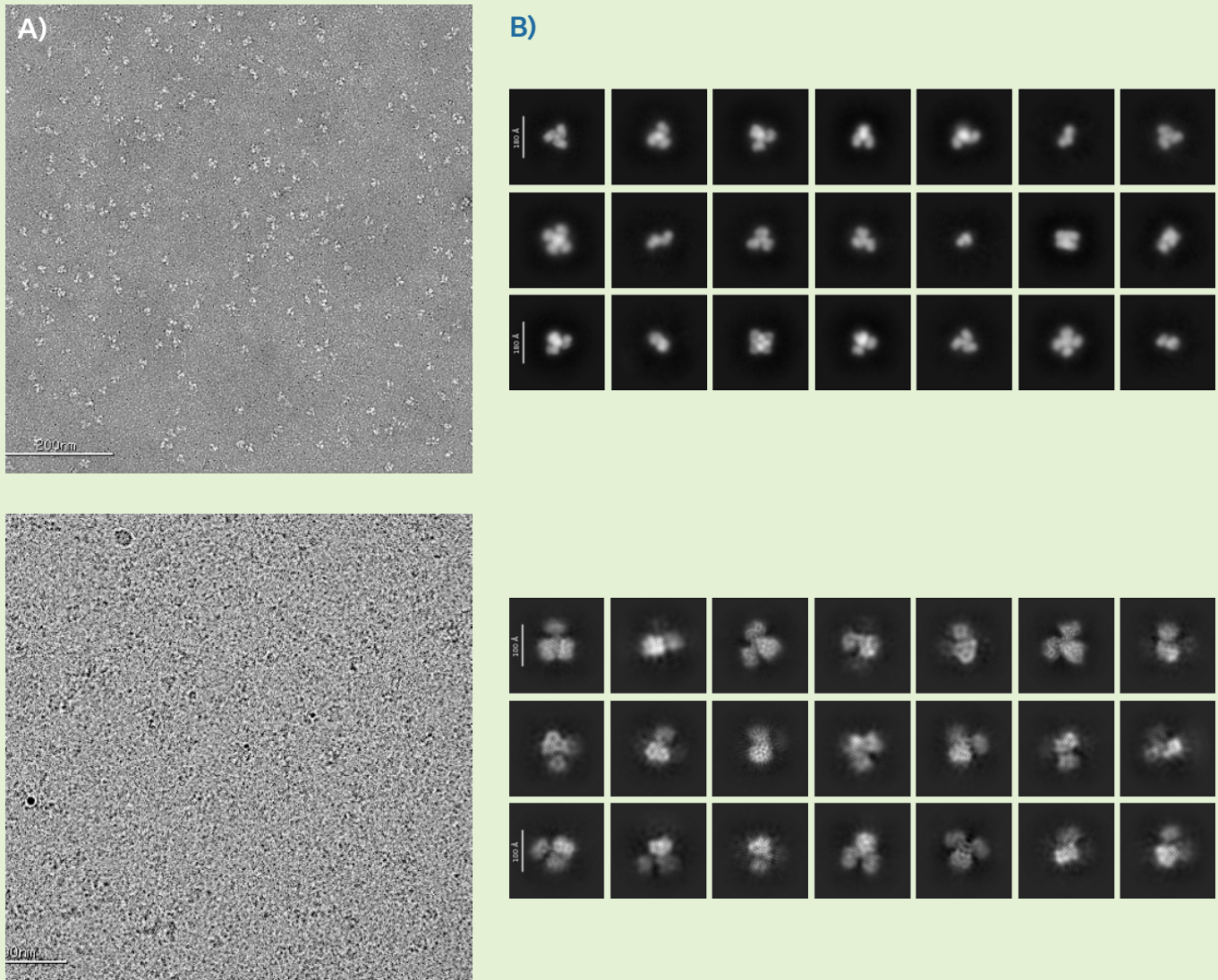
## Results & Discussion

### Protein Expression, Purification, and Characterization

Under optimized conditions, the soluble domain of CYP3A4 was successfully expressed in *E. coli* and purified using a three-step chromatographic workflow. The protein was first captured via Ni-NTA affinity chromatography, followed by further purification using ion exchange chromatography (IEX). Final polishing was performed by preparative size exclusion chromatography (SEC) to isolate a monodisperse form. The purity and integrity of the protein were confirmed by SDS-PAGE, HPLC-SEC, and intact mass spectrometry, which collectively verified high purity of the final CYP3A4 preparation. Preparative and analytical SEC suggest dimeric or trimeric assembly of the protein in solution, respectively.

### CYP3A4 Cryo-EM Characterization

The purified protein was initially analyzed with negative staining, to assess protein suitability for cryo-EM. The images confirmed that the sample was monodispersed and suggested that oligomeric forms were present (**Figure 2A, page 4**). The oligomerization state was confirmed by the initial 2D classification results obtained in negative staining and cryo-EM experiments (**Figure 2B, top and bottom, respectively, page 4**). The most common quaternary structure observed was a trimer, although dimers and tetramers are also visible. Given that CYP3A4 is relatively small for cryo-EM analysis, the presence of trimers could be useful both for the increased size of the particles as well as for the introduction of additional symmetry, as long as the binding sites remain accessible and the oligomerization does not affect the activity of the enzyme. One remaining problem that became evident from this initial work was the presence of a strong preferred orientation (PO), which needed to be addressed to visualize the atomic details necessary to answer questions about oligomerization state and ligand binding.

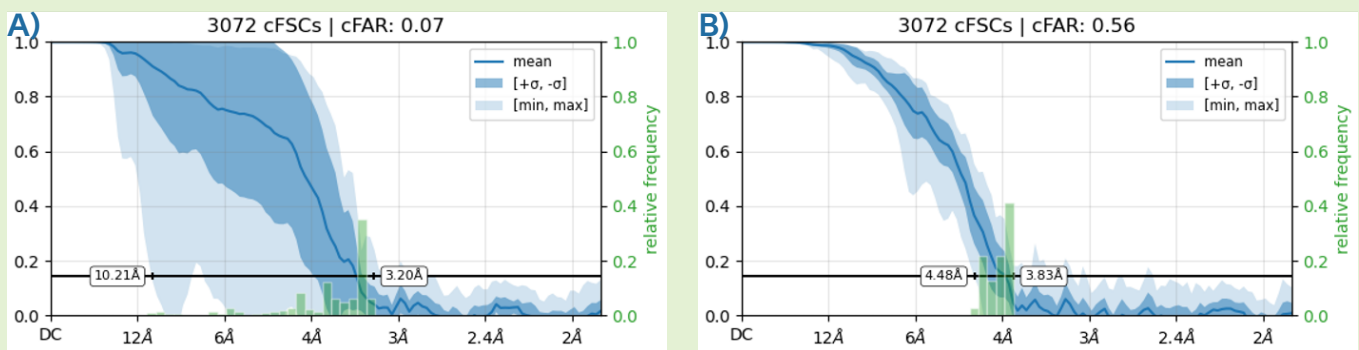


**Figure 2** Initial characterization of the sample. **A)** A representative negative staining image (**top**) and cryo-EM image (**bottom**) of the sample at a concentration of 0.015 and 1.5 mg/ml, respectively. **B)** Representative 2D classes from negative staining (**top**) and cryo-EM (**bottom**) showing the presence of different oligomers, with a prevalence of trimeric species.

## How to Address Preferred Orientation

The first set of good conditions identified during vitrification screening involved using the sample at ~ 1.5 mg/ml and UltrAuFoil 1.2x1.3 grids. With gold grids, tilted data collection was seen as a viable path to overcome the PO, and indeed both the structures reported here were solved with data collected on a Krios with a tilt of 30° and 20° degrees for the apo structure and the liganded structure, respectively (**Table 1, page 11**). Tilted data collections nevertheless have some shortcomings: increased ice thickness may reduce the signal-to-noise ratio, an effect particularly problematic when working with small particles; greater beam-induced movement can result in blurred data; and different angles and focus gradients, as well as potential particle overlap on the grid require more complex image processing analysis. These issues can limit the resolution of the final structure, though advanced computational methods and careful data collection strategies can mitigate them<sup>7, 8</sup>.

Therefore, our focus shifted to exploring a range of conditions, including variations in protein concentration, grid type, and the use of additives, to identify the most effective parameters to mitigate the preferred orientation issue and to improve the behavior of CYP3A4 at the Air-Water Interface (AWI). To learn more about AWI, [read our blog](#). All conditions underwent screening, and short data collections (~1,200 images on a Glacios microscope) were performed on suitable grids to assess vitrification conditions and particle distribution. To evaluate the impact of the different vitrification processes on preferred orientation, 2D classification -followed by 3D classification and 3D reconstruction - were used to compare the different conditions. This analysis identified lauryl maltose neopentyl glycol (LMNG) and fos-choline-8, fluorinated (FFC8) as the two best vitrification conditions. Both reduced PO significantly, and the results suggest these vitrification conditions could be used in the future to avoid the need for tilted data collection during high-resolution studies (**Figure 3, page 5**).

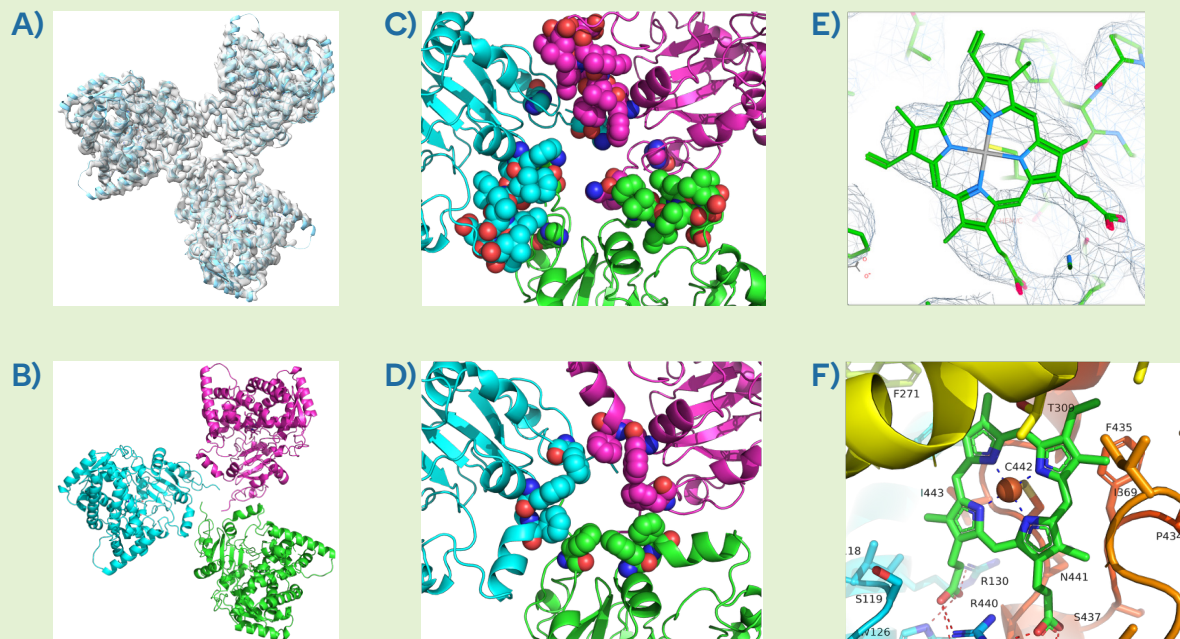


**Figure 3** Conical FSC summary plots for 3D reconstructions of CYP3A4 **A)** under initial vitrification conditions and **B)** Using FFC8 as an additive during vitrification. The plots were generated in Cryosparc V4.6.2 and show very similar information to the figure generated via the legacy 3DFSC job<sup>9</sup> higher the number for the conical FSC Area Ratio, or cFAR, the least orientation bias is present.

### 3D Structure of the Unliganded CYP3A4

The structure of the unliganded form of CYP3A4 (soluble domain, residues 24-503) was solved using data collected on a TFS Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan Quantum 967 LS imaging filter and Gatan K3 Direct Detection Camera (data collection parameters can be found in **Table 1, page 11**) using Legion software<sup>10</sup>. The data processing followed strategies similar to those described in previously published work<sup>6, 11</sup> and was carried out in CryoSPARC v4.6.2<sup>12</sup>, resulting in a final reconstruction at a nominal resolution of 3.4 Å (**Figure 4A**).

Model building and refinement were initiated with a published model for CYP3A4 (*PDB ID:1W0E*). The initial structure was placed into the sharpened density maps using the ChimeraX “fit to map” utility<sup>13</sup>. Iterative rounds of model building and refinement were performed in PHENIX v.1.19.2<sup>14, 15</sup> and COOT v.0.9.6 EL<sup>16</sup>. Refinement statistics are summarized in **Table 2, page 12**. As mentioned above, the protein forms symmetrical trimers with a central cavity (**Figure 4B**). Each pair of monomers (“A,B”, “A,C”, “B,C”) share about 400 Å<sup>2</sup> of surface (according to PISA)<sup>17</sup>; within these surfaces, polar and hydrophobic interactions are located between the helix spanning residues Phe228-Leu236 of one molecule and residues Leu44-Phe46 of the neighboring one; and the N-terminal residues (Ala24-Tyr25) of one molecule and His28 of the neighboring one (**Figure 4C**). A second set of interactions involves a ring of 9 Phe residues (3 from each monomer), that also surrounds the central cavity (**Figure 4D**).



**Figure 4** CYP3A4 APO structure. **A)** The cryo-EM map at 3.44 Å; **B)** The trimer derived from the cryo-EM experiment; **C)** and **D)** the two sets of interactions among the monomers in the trimer. **E)** The density for the heme; **F)** The heme binding site

This is the first time that a symmetric trimer is observed as a quaternary structure for CYP3A4. Dimers and dimers of dimers are frequently observed in the crystalline structures (often generated by crystallographic symmetry but sometimes present as non-crystallographic complexes in the asymmetric unit); there is one instance of an asymmetric trimer/AU (*PDB ID: 5TE8*), although a symmetrical one is generated by crystallography symmetry. The arrangement of the monomers is different from the one observed in the cryo-EM generated structure. Oligomerization of P450 on the microsomal membrane surfaces has been proposed to have functional consequences<sup>5</sup>. Based on cross-linking and luminescence resonance energy transfer analysis (LRET), the conclusion from the paper is that CYP3A4 is an oligomer and that the most likely number of subunits in the oligomer is a multiple of 3.

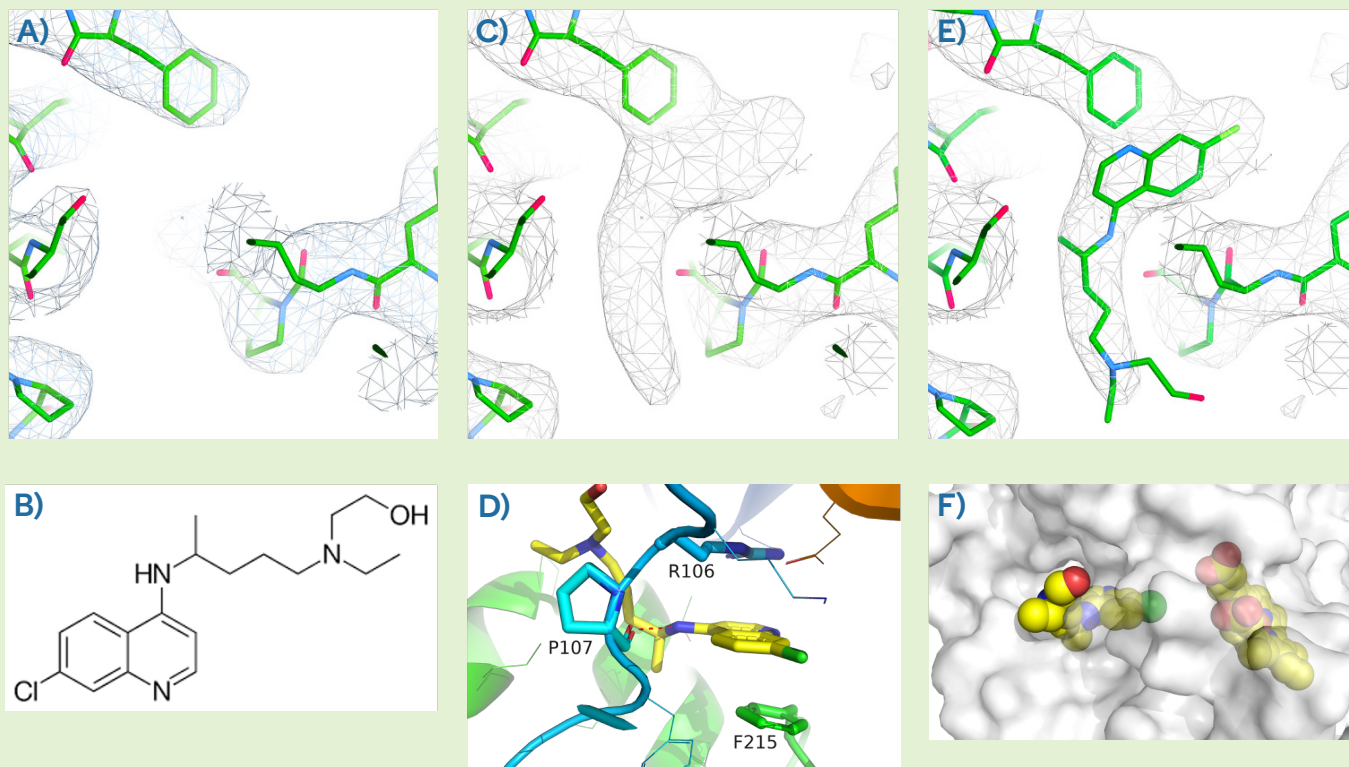
## Heme Binding Site

The cryo-EM map showed clear density for the bound heme (**Figure 4E, page 6**). The heme group is located within the enzyme's large active site cavity, and it is accessible to outside solvent by the substrate access channel. The heme iron is ligated by a conserved cysteine (Cys442) and the heme propionates interact with the side chains of Arg105, Trp126, Arg130, Arg375 and Arg440 (**Figure 4F, page 6**). The binding site observed in the cryo-EM structure is identical to the site described in all previous crystallographic structures.

## Structure of CYP3A4 with Hydroxychloroquine

To test the robustness of the cryo-EM approach to generate high-resolution structures of CYP3A4 with ligands bound, we attempted to solve the structure of the complex between CYP3A4 and hydroxychloroquine (HCQ). HCQ is a medication used to prevent and treat malaria and may be used for the treatment of rheumatoid arthritis, lupus, and porphyria cutanea tarda. HCQ is both a substrate for CYP3A4, and a time-dependent inhibitor<sup>18</sup>. No crystal structure of HCQ bound to CYP3A4 has ever been reported, although structures of substrates bound to CYP3A4 are available (*PDB ID: 8DYC; 8SO1; 8SO2*). The drug-bound sample was prepared by adding hydroxychloroquine to CYP3A4, yielding a final concentration of 50 mM (1:600 molar ratio). The mixture was incubated for 1 hour at 4°C. Following a procedure similar to that used for the unliganded sample, the HCQ-bound CYP3A4 structure was reconstructed to a nominal resolution of 3.1 Å using CryoSPARC v4.6.2<sup>12</sup>.

Model building and refinement were initiated with the experimentally obtained apo CYP3A4 structure. Cartesian coordinates and geometry restraints for HCQ were generated with PHENIX 14. Iterative rounds of model building and refinement were performed in PHENIX v.1.19.2<sup>14, 15</sup> and COOT v. 0.9.6 EL<sup>16</sup>. The overall structure of the bound protein is identical to the one of the apo protein (RMSD on C $\alpha$  is 0.28 Å), with clear density for the heme. Map inspection, and comparison with the apo sample, revealed additional density in the substrate access channel (**Figures 5A-C, page 8**). This additional density was used to model the bound ligand.



**Figure 4** A) The substrate access channel in the apo structure; B) Density was observed in the same area in the HCQ bound sample and C) The observed density fitted with the ligand. D) Chemical structure of HCQ. E) Interactions of HCQ with the residues of the substrate access channel. F) Relative orientation of HCQ and Heme (represented as sphere) in the substrate access channel. The protein is shown as a semi-transparent grey surface.

## Description of Binding Site

HCQ is bound in the substrate access channel, which has also been identified as a high-affinity binding site<sup>19, 20</sup>. The quinoline ring is stacked between the side chains of Phe215 and Arg106. There is one hydrogen bond between the nitrogen of the quinoline and the main chain oxygen of Pro107 (**Figure 5E**). A very similar binding mode (compound heterocyclic ring stacked between Phe125 and Arg 106) is observed in the crystal structures of Fluorol-CYP3A4 (*PDB ID: 8DYC*) and caffeine-CYP3A4 (*PDB ID: 8SO1, 8SO2*). The chlorine atom is facing the heme, while the tail of the molecule, which is the site of metabolic processing<sup>18, 19</sup>, is facing the outside solvent (**Figure 5F**). This suggests that this binding mode is inhibitory, since a 180-degree rotation of the molecule is necessary for the metabolization reaction to occur. This is in agreement with the CYP3A4 time-dependent inhibition observed for HCQ and its metabolites<sup>18</sup>.

## Conclusions

We report here the first cryo-EM structures for the P450 CYP3A4, in the unliganded form and bound to a substrate (HCQ).

The soluble domain of CYP3A4 was expressed under optimized conditions and purified using a three-step chromatographic workflow. Preparative and analytical SEC, as well as the cryo-EM results, indicate that the sample is in an oligomeric form (trimer), which has been suggested to have functional consequences<sup>5</sup>. HCQ is found binding in the substrate access channel in a nonproductive orientation, in agreement with its observed time-dependent inhibition of CYP3A4<sup>18</sup>.

The specific oligomeric state observed during the cryo-EM analysis has not been previously observed using other structural techniques and confirms the importance of using alternative methods to study even well characterized systems.

To learn more about AWI (page 5), check out our blog at:  
<https://www.nanoimagingervices.com/about/blog/a-guide-to-cryo-grid-production-materials-interfaces-and-their-impact-on-sample-prep>  
or scan the QR code.



## References

- Klyushova, L. S., Perepechaeva, M. L. & Grishanova, A. Y. The Role of CYP3A in Health and Disease. *Biomedicines* **10**, 2686 (2022).
- Sevrioukova, I. F. & Poulos, T. L. Understanding the mechanism of cytochrome P450 3A4: recent advances and remaining problems. *Dalton Trans.* **42**, 3116–3126 (2012).
- Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700 (1987).
- Williams, P. A. et al. Crystal Structures of Human Cytochrome P450 3A4 Bound to Metyrapone and Progesterone. *Science* **305**, 683–686 (2004).
- Davydov, D. R., Davydova, N. Y., Sineva, E. V. & Halpert, J. R. Interactions among Cytochromes P450 in Microsomal Membranes OLIGOMERIZATION OF CYTOCHROMES P450 3A4, 3A5, AND 2E1 AND ITS FUNCTIONAL CONSEQUENCES\*. *J. Biol. Chem.* **290**, 3850–3864 (2015).
- Catalano, C. et al. The CryoEM structure of human serum albumin in complex with ligands. *J. Struct. Biol.* **216**, 108105 (2024).
- Aiyer, S. et al. Overcoming resolution attenuation during tilted cryo-EM data collection. *Nat. Commun.* **15**, 389 (2024).
- Aiyer, S., Strutzenberg, T. S., Bowman, M. E., Noel, J. P. & Lyumkis, D. Single-Particle Cryo-EM Data Collection with Stage Tilt using Leginon. *J. Vis. Exp.* (2022). doi:10.3791/64136
- Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* **14**, 793–796 (2017).
- Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41–60 (2005).
- Tran, N. L. et al. High-Resolution Cryo-Electron Microscopy Structure Determination of Haemophilus influenzae Tellurite-Resistance Protein A via 200 kV Transmission Electron Microscopy. *Int. J. Mol. Sci.* **25**, 4528
- Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296 (2017).
- Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14–25 (2018).
- Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074–80 (2009).
- Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531–544 (2018).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
- Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797 (2007).
- Paludetto, M.-N. et al. Hydroxychloroquine is Metabolized by Cytochrome P450 2D6, 3A4, and 2C8, and Inhibits Cytochrome P450 2D6, while its Metabolites also Inhibit Cytochrome P450 3A in vitro. *Drug Metab. Dispos.* **51**, 293–305 (2023).
- Sevrioukova, I. F. Crystal Structure of CYP3A4 Complexed with Fluorol Identifies the Substrate Access Channel as a High-Affinity Ligand Binding Site. *Int. J. Mol. Sci.* **23**, 12591 (2022).

To learn more about cryo-EM services,  
visit [nanoimagingervices.com](https://nanoimagingervices.com) or  
email [info@nanoimagingervices.com](mailto:info@nanoimagingervices.com)

To learn more about protein production  
at Proteos, visit [proteos.com](https://proteos.com)  
or email [info@proteos.com](mailto:info@proteos.com)

Table 1. Summary of Data Collection

Sample	CYP3A4	CYP3A4:HCQ
Electron Voltage	300 kV	300 kV
Nominal Defocus	-1.8 $\mu\text{m}$ to -4.2 $\mu\text{m}$	-1.8 $\mu\text{m}$ to -4.2 $\mu\text{m}$
Nominal Magnification	105000	105000
Pixel Size	0.834 $\text{\AA}$	0.834 $\text{\AA}$
Total Dose	51.25 e-/ $\text{\AA}^2$	51.25 e-/ $\text{\AA}^2$
Total Exposure Time	1.4 sec	1.4 sec
Frame Rate	40 ms	40 ms
Total Number of Frames	35	35
Energy Filter Slit Width	20 eV	20 eV
Imaging	Image shift	Image shift
Total Number of Images	12,858	8,219
Total Number of High Mag Images	11,648	7,362
Stage Tilt Angle	30.0°	20.0°

Table 2. Refinement Statistics

Data	CYP3A4	CYP3A4:HCQ
Total extracted particles (no.)	1,450,000	1,059,000
Final particles (no.)	100,726	108,382
Symmetry	C3	C3
Nominal Resolution (FSC 0.143)	3.4 Å	3.4 Å
Model		
Chains	3	3
Atoms	10701 (Hydrogens: 0)	10770 (Hydrogens: 0)
Residues	Protein: 1404	Protein: 1404
Water	0	0
Ligands	HEM: 3	HEM: 3; AIH: 3
Bonds (RMSD)		
Length (Å) (# > 4sigma)	0.004 (0)	0.003 (0)
Angles (°) (# > 4sigma)	0.720 (2)	0.695 (7)
MolProbity score	1.93	1.84
Clash score	8.42	7.23
Ramachandran Plot (%)		
Outliers	0.65	0.36
Allowed	2.80	2.73
Favored	96.55	96.91
Ramachandran Plot Z-score (RMSD)		
Whole (N=1392)	1.04 (0.23)	0.30 (0.23)
Helix (N=783)	1.50 (0.19)	0.74 (0.19)
Sheet (N=84)	0.02 (0.55)	0.06 (0.54)
Loop (N=525)	-0.35 (0.29)	-0.47 (0.29)
Rotamer outliers (%)	2.11	2.20
Cbeta outliers (%)	0.00	0.00
Peptide plane (%)		
Cis proline/general	0.0/0.0	0.0/0.0
Twisted proline/general	0.0/0.0	0.0/0.0
CaBLAM outliers (%)	2.10	2.54
ADP (B-factors)		
Iso/Aniso	10701/0	10770/0
Protein (min/max/average)	44.48/163.60/95.86	28.94/118.69/62.74
Ligand (min/max/average)	61.12/62.35/61.7	35.44/41.98/38.13

Data	CYP3A4	CYP3A4:HCQ
Data		
Box pixel size	256/256/256	256/256/256
Angles (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Supplied Resolution (Å)	3.4	3.1
Resolution Estimates (Å)	Masked/Unmasked	Masked/Unmasked
d FSC (half maps; 0.143)	3.6/3.7	3.2/3.3
d 99 (full/half1/half2)	3.7/2.0/2.0	3.3/1.7/1/7//3.2/1/7/1/7
d model	3.8/3.8	3.4/3.4
d FSC model (0/0.143/0.5)	3.3/3.4/3.7//3.4/3.4/3.9	3.0/3/1/3.4//3.1/3.1/3.6
Map min/max/mean	-0.41/0.74/0.01	-0.41/0.58/0.01
Model vs. Data		
CC (mask)	0.82	0.83
CC (box)	0.74	0.84
CC (peaks)	0.66	0.76
CC (volume)	0.80	0.80
Mean CC for ligands	0.88	0.83