W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 6, 2023 – 01:26 AM EDT

PDB ID	:	4EJH
Title	:	Human Cytochrome P450 2A13 in complex with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)
Authors	:	DeVore, N.M.; Scott, E.E.
Deposited on	:	2012-04-06
Resolution	:	2.35 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

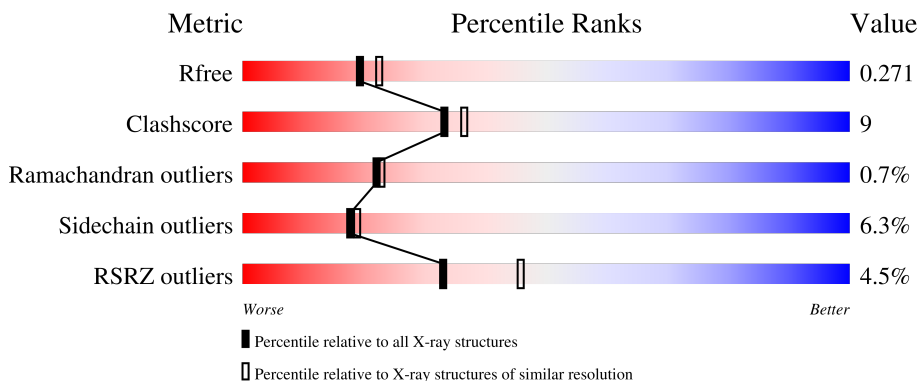|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.271
Clashscore		9
Ramachandran outliers		0.7%
Sidechain outliers		6.3%
RSRZ outliers		4.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	476	82% 13% • •
1	B	476	79% 17% • •
1	C	476	82% 14% • •
1	D	476	81% 15% • •
1	E	476	2% 76% 20% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	476	7% / 71% / 24%
1	G	476	7% / 71% / 23%
1	H	476	18% / 66% / 28%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	0QA	F	502	-	-	X	-

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 30804 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Cytochrome P450 2A13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	464	Total 3777	C 2430	N 653	O 676	S 18	8	2	0
1	B	464	Total 3771	C 2426	N 653	O 674	S 18	0	1	0
1	C	464	Total 3776	C 2429	N 653	O 676	S 18	0	2	0
1	D	464	Total 3771	C 2426	N 653	O 674	S 18	0	1	0
1	E	464	Total 3771	C 2426	N 653	O 674	S 18	0	1	0
1	F	464	Total 3771	C 2426	N 653	O 674	S 18	0	1	0
1	G	463	Total 3737	C 2404	N 642	O 673	S 18	0	0	0
1	H	459	Total 3723	C 2396	N 640	O 669	S 18	0	1	0

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	initiating methionine	UNP Q16696
A	24	ALA	-	expression tag	UNP Q16696
A	25	LYS	-	expression tag	UNP Q16696
A	26	LYS	-	expression tag	UNP Q16696
A	27	THR	-	expression tag	UNP Q16696
A	28	SER	-	expression tag	UNP Q16696
A	29	SER	-	expression tag	UNP Q16696
A	30	LYS	-	expression tag	UNP Q16696
A	495	HIS	-	expression tag	UNP Q16696
A	496	HIS	-	expression tag	UNP Q16696
A	497	HIS	-	expression tag	UNP Q16696
A	498	HIS	-	expression tag	UNP Q16696
B	23	MET	-	initiating methionine	UNP Q16696

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	24	ALA	-	expression tag	UNP Q16696
B	25	LYS	-	expression tag	UNP Q16696
B	26	LYS	-	expression tag	UNP Q16696
B	27	THR	-	expression tag	UNP Q16696
B	28	SER	-	expression tag	UNP Q16696
B	29	SER	-	expression tag	UNP Q16696
B	30	LYS	-	expression tag	UNP Q16696
B	495	HIS	-	expression tag	UNP Q16696
B	496	HIS	-	expression tag	UNP Q16696
B	497	HIS	-	expression tag	UNP Q16696
B	498	HIS	-	expression tag	UNP Q16696
C	23	MET	-	initiating methionine	UNP Q16696
C	24	ALA	-	expression tag	UNP Q16696
C	25	LYS	-	expression tag	UNP Q16696
C	26	LYS	-	expression tag	UNP Q16696
C	27	THR	-	expression tag	UNP Q16696
C	28	SER	-	expression tag	UNP Q16696
C	29	SER	-	expression tag	UNP Q16696
C	30	LYS	-	expression tag	UNP Q16696
C	495	HIS	-	expression tag	UNP Q16696
C	496	HIS	-	expression tag	UNP Q16696
C	497	HIS	-	expression tag	UNP Q16696
C	498	HIS	-	expression tag	UNP Q16696
D	23	MET	-	initiating methionine	UNP Q16696
D	24	ALA	-	expression tag	UNP Q16696
D	25	LYS	-	expression tag	UNP Q16696
D	26	LYS	-	expression tag	UNP Q16696
D	27	THR	-	expression tag	UNP Q16696
D	28	SER	-	expression tag	UNP Q16696
D	29	SER	-	expression tag	UNP Q16696
D	30	LYS	-	expression tag	UNP Q16696
D	495	HIS	-	expression tag	UNP Q16696
D	496	HIS	-	expression tag	UNP Q16696
D	497	HIS	-	expression tag	UNP Q16696
D	498	HIS	-	expression tag	UNP Q16696
E	23	MET	-	initiating methionine	UNP Q16696
E	24	ALA	-	expression tag	UNP Q16696
E	25	LYS	-	expression tag	UNP Q16696
E	26	LYS	-	expression tag	UNP Q16696
E	27	THR	-	expression tag	UNP Q16696
E	28	SER	-	expression tag	UNP Q16696
E	29	SER	-	expression tag	UNP Q16696

Continued on next page...

Continued from previous page...

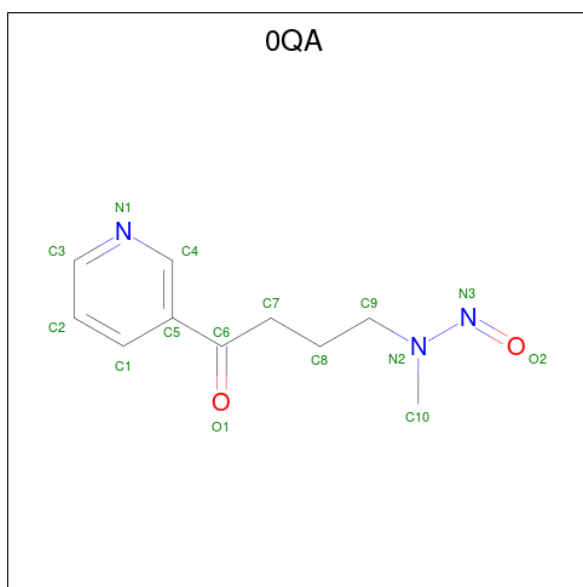| Chain	Residue	Modelled	Actual	Comment	Reference
E	30	LYS	-	expression tag	UNP Q16696
E	495	HIS	-	expression tag	UNP Q16696
E	496	HIS	-	expression tag	UNP Q16696
E	497	HIS	-	expression tag	UNP Q16696
E	498	HIS	-	expression tag	UNP Q16696
F	23	MET	-	initiating methionine	UNP Q16696
F	24	ALA	-	expression tag	UNP Q16696
F	25	LYS	-	expression tag	UNP Q16696
F	26	LYS	-	expression tag	UNP Q16696
F	27	THR	-	expression tag	UNP Q16696
F	28	SER	-	expression tag	UNP Q16696
F	29	SER	-	expression tag	UNP Q16696
F	30	LYS	-	expression tag	UNP Q16696
F	495	HIS	-	expression tag	UNP Q16696
F	496	HIS	-	expression tag	UNP Q16696
F	497	HIS	-	expression tag	UNP Q16696
F	498	HIS	-	expression tag	UNP Q16696
G	23	MET	-	initiating methionine	UNP Q16696
G	24	ALA	-	expression tag	UNP Q16696
G	25	LYS	-	expression tag	UNP Q16696
G	26	LYS	-	expression tag	UNP Q16696
G	27	THR	-	expression tag	UNP Q16696
G	28	SER	-	expression tag	UNP Q16696
G	29	SER	-	expression tag	UNP Q16696
G	30	LYS	-	expression tag	UNP Q16696
G	495	HIS	-	expression tag	UNP Q16696
G	496	HIS	-	expression tag	UNP Q16696
G	497	HIS	-	expression tag	UNP Q16696
G	498	HIS	-	expression tag	UNP Q16696
H	23	MET	-	initiating methionine	UNP Q16696
H	24	ALA	-	expression tag	UNP Q16696
H	25	LYS	-	expression tag	UNP Q16696
H	26	LYS	-	expression tag	UNP Q16696
H	27	THR	-	expression tag	UNP Q16696
H	28	SER	-	expression tag	UNP Q16696
H	29	SER	-	expression tag	UNP Q16696
H	30	LYS	-	expression tag	UNP Q16696
H	495	HIS	-	expression tag	UNP Q16696
H	496	HIS	-	expression tag	UNP Q16696
H	497	HIS	-	expression tag	UNP Q16696
H	498	HIS	-	expression tag	UNP Q16696

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (for-

mula: $C_{34}H_{32}FeN_4O_4$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	E	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	F	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	G	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	H	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is 4-[methyl(nitroso)amino]-1-(pyridin-3-yl)butan-1-one (three-letter code: 0QA) (formula: $C_{10}H_{13}N_3O_2$).

0QA

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			15	10	3	2		
3	B	1	Total	C	N	O	0	0
			15	10	3	2		
3	C	1	Total	C	N	O	0	0
			15	10	3	2		
3	D	1	Total	C	N	O	0	0
			15	10	3	2		
3	E	1	Total	C	N	O	0	0
			15	10	3	2		
3	F	1	Total	C	N	O	0	0
			15	10	3	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

C3

OH O3

C1   C2

HO

O1

OH

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	G	1	Total 6	C 3	O 3	0	0
4	H	1	Total 6	C 3	O 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	56	Total 56	O 56	0	0
5	B	39	Total 39	O 39	0	0
5	C	71	Total 71	O 71	0	0
5	D	49	Total 49	O 49	0	0
5	E	13	Total 13	O 13	0	0
5	F	11	Total 11	O 11	0	0
5	G	12	Total 12	O 12	0	0
5	H	10	Total 10	O 10	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
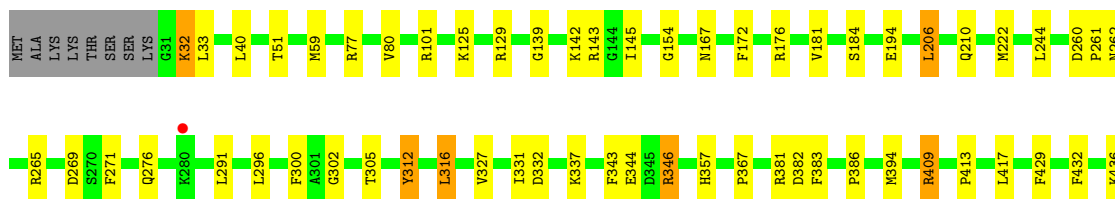
• Molecule 1: Cytochrome P450 2A13

Chain A:

• Molecule 1: Cytochrome P450 2A13

Chain B:
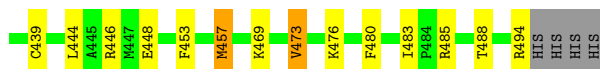

• Molecule 1: Cytochrome P450 2A13

Chain C:

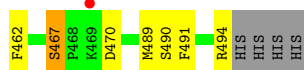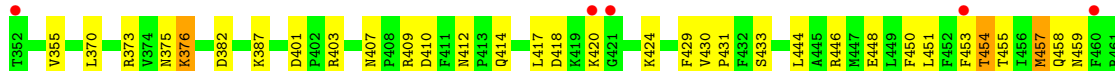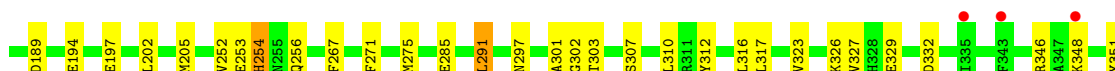C439 L444 A445 R446 M447 E448 F453 M457 K469 V473 K476 F480 I483 P484 R485 T488 R494 HIS HIS HIS

- Molecule 1: Cytochrome P450 2A13

Chain D:

81% 15%

MET ALA LYS LYS THR SER SER LYS G31 K32 L40 P41 F42 Q53 N56 M59 E63 R64 P67 R77 E97 R101 L110 A124 R128 L135 G139 K142 R143 G144 I145 E146 E147 R161 A166 D189 E194 D195 K196 E197

L201 T216 Y220 S224 P233 Q236 L241 D260 P261 F271 L291 N297 F300 Y312 L315 L316 A325 H328 I335 P341 K342 F343 E344 D345 R346 V355 I359 A371 H372 R373 V374 N375 K376 K379 F380 R381 D382 K387

R403 R409 F450 L451 F452 F453 T454 T455 I456 M457 Q458 N459 F460 R461 F462 K463 S467 D470 I471 D472 V473 K476 R485 S490 F491 L492 P493 R494 HIS HIS HIS
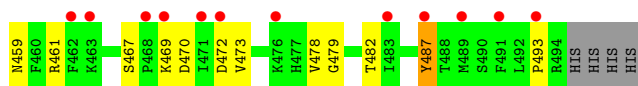
- Molecule 1: Cytochrome P450 2A13

Chain E:

2%

76% 20%

MET ALA LYS LYS THR SER SER LYS G31 L40 Q53 M54 Y55 N56 M59 S62 L73 E96 R101 G115 S119 R123 A124 K125 Q126 L127 R128 L135 K142 R143 E146 E147 Q150 R161 G162 T163 N167 I168 F172 R176 V186

D189 E194 E197 L202 M205 V252 E253 H254 N255 Q256 F267 F271 M275 E285 L291 N297 A301 G302 T303 S307 L310 R311 Y312 L316 L317 V323 K326 V327 H328 E329 D332 I335 F343 R346 A347 K348 Y351

T352 V355 L370 R373 V374 N375 K376 D382 K387 D401 P402 R403 N407 P408 R409 D410 F411 N412 P413 Q414 L417 D418 K419 K420 G421 K424 F429 V430 P431 F432 S433 L444 A445 R446 M447 E448 L449 F450 L451 F452 F453 T454 T455 I456 M457 Q458 N459 F460 R461

F462 S467 P468 K469 D470 M489 S490 F491 R494 HIS HIS HIS

- Molecule 1: Cytochrome P450 2A13

Chain F:

7%

71% 24%

MET ALA LYS LYS THR SER SER LYS G31 P37 L40 I43 Q48 Q53 N56 M59 C82 E89 A90 L91 V92 D93 R101 G102 E103 D108 F118 L135 G139 V140 G141 K142 R143 G144 I145 E146 E147 R148 L156 G162 N167 I168

F172 F173 L174 S175 R176 T177 N180 S184 I185 V186 F187 G188 D189 R190 F191 D192 Y193 K196 E197 F198 L199 S200 L201 L202 L206 F211 T212 T216 K228 G232 P233 Q234 F247 K250 R257 T258 L259 D260 P261 N262 S263 P264 R265 D266 F267 S270 F271

K280 N281 P282 N283 L288 K289 N290 L291 V292 N297 F300 A301 G302 T303 T305 V306 S307 T308 Y312 G313 F314 L315 L316 L317 M318 V323 V327 I331 I335 N338 R339 K342 F343 E344 D345 R346 A347 K348 M349 P350 Y351 A354 V355 M365 L366 L370

A371 H372 R373 V374 N375 R381 D382 K387 G388 V398 N407 P408 R409 D410 F411 L417 K420 Q421 Q422 F423 K424 A428 S433 E442 G443 L444 A445 R446 M447 E448 L449 F450 L451 F452 F453 T454 T455 I456 M457 Q458 R461 K476 R485 S490 F491 L492 P493

R494 HIS HIS HIS HIS

• Molecule 1: Cytochrome P450 2A13

Chain G:

7%

71% 23%

MET ALA LYS LYS THR SER SER LYS GLY K32 L33 L40 T51 E52 Q53 N56 M59 S62 E63 R64 R77 D85 E89 D93 S99 E103 Q104 W109 L110 F111 G115 F118 S119 R123 A124 K125 R129 G137 F138 G139 V140 G141

I145 E151 E152 A153 G154 F155 L156 I157 G165 A166 N167 F173 V178 I182 S183 S184 F187 G188 F191 F198 L199 S200 L201 L202 R203 M204 F209 Q210 F211 T212 A213 T214 S215 T216 G217 Q218 L219 S225 P233 E238 L241 K250 E253 H254 N255

L259 D266 P271 L272 I273 R274 E278 P282 N283 T284 E285 L288 K289 N290 L291 V292 L296 N297 F300 A301 G302 T303 E304 T305 L310 R311 Y312 G313 F314 L315 L316 L317 M318 K319 V323 K326 V327 H328 E329 K337 K342 F343 E344 D345 R346 V351

T352 E353 A354 V355 I359 Q360 R361 D364 M365 L370 A371 H372 N375 F380 R381 K387 L399 P408 F411 D418 K419 K420 G421 Q422 F423 S433 K436 C439 L449 F453 I456 M457 R461 F462 K463 S464 P468 K469 D470 I471 D472 V473

S474 P475 K476 H477 V478 G479 F480 A481 I483 P484 R485 N486 Y487 F491 R494 HIS HIS HIS HIS

• Molecule 1: Cytochrome P450 2A13

Chain H:

18%

66% 28%

MET ALA LYS LYS THR SER SER LYS GLY K32 L33 T51 E52 Q53 M54 Y55 N56 S57 L58 M59 R64 P67 I71 H72 L73 C82 D85 K88 E89 A90 L91 V92 D93 Q94 A95 E96 E97 F98 S99 G100 R101 Q104 L110 Y114 G115 V116 A117 F118 R123

Q126 L127 R128 R129 A133 T134 L135 R136 G137 F138 G139 G144 I145 E146 E147 R148 E151 E152 A153 G154 F155 L156 I157 D158 A159 L160 R161 D169 P170 T171 S175 R176 T177 V178 S179 N180 V181 I182 S183 S184 I185 V186 F187 Y193 E194 D195 M205 F209

Q218 M222 S225 L241 Q242 G243 L244 E245 D246 F247 K251 V252 E253 H254 L259 D260 P261 N262 S263 P264 R265 D266 F267 I268 D269 S270 F271 L272 I273 K280 E285 L288 L291 T294 T295 L296 N297 L298 F299 F300 T303 E304 T305 V306 T309 Y312

L315 L316 L317 M318 K319 H320 P321 E322 V323 E324 A325 K326 E329 E330 I331 D332 R333 V334 I335 N338 R339 Q340 P341 K342 F343 R346 A347 K348 M349 P350 Y351 T352 V355 I356 H357 E358 I359 F362 G363 D364 M365 G369 L370 A371 H372 K376 D382 F383 F384 L385 P386

K387 V391 M394 L395 L399 R400 D401 P402 R403 F404 F405 S406 N407 N412 P413 Q414 H415 F416 L417 D418 K419 K420 G421 Q422 F423 K424 K425 S426 D427 A428 F429 V430 P431 F432 S433 I434 G435 K436 R437 Y438 F440 G441 E442 M447 E448 L449 F450 L451 F452 F453 T454 T455 I456

N459 F460 R461 F462 K463 S467 P468 K469 D470 I471 D472 V473 K476 H477 V478 G479 T482 I483 Y487 T488 M489 S490 F491 L492 P493 R494 HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	71.68Å 119.27Å 153.66Å 100.59° 101.86° 93.56°	Depositor
Resolution (Å)	69.77 – 2.35 69.77 – 2.35	Depositor EDS
% Data completeness (in resolution range)	97.5 (69.77-2.35) 97.5 (69.77-2.35)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.10	Depositor
$<I/\sigma(I)>$ [1]	2.17 (at 2.34Å)	Xtriage
Refinement program	REFMAC 6.1.13	Depositor
R, R_{free}	0.214 , 0.273 0.214 , 0.271	Depositor DCC
R_{free} test set	9958 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	42.3	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 46.8	EDS
L-test for twinning[2]	$<\vert L\vert> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	30804	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.85% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $<\vert L\vert>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, 0QA, GOL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
1	A	1.01	2/3880 (0.1%)	0.90	1/5224 (0.0%)
1	B	0.99	3/3871 (0.1%)	0.88	4/5212 (0.1%)
1	C	1.04	4/3879 (0.1%)	0.89	2/5223 (0.0%)
1	D	0.99	1/3871 (0.0%)	0.87	1/5212 (0.0%)
1	E	0.85	0/3871	0.79	0/5212
1	F	0.80	0/3871	0.76	1/5212 (0.0%)
1	G	0.82	0/3834	0.81	0/5170
1	H	0.77	0/3822	0.75	1/5149 (0.0%)
All	All	0.92	10/30899 (0.0%)	0.83	10/41614 (0.0%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	473	VAL	CB-CG2	5.94	1.65	1.52
1	C	383	PHE	CE2-CZ	5.37	1.47	1.37
1	B	63	GLU	CG-CD	5.28	1.59	1.51
1	B	322	GLU	CG-CD	5.24	1.59	1.51
1	A	133	ALA	CA-CB	5.20	1.63	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	123	ARG	NE-CZ-NH1	-5.97	117.31	120.30
1	C	143	ARG	NE-CZ-NH2	-5.90	117.35	120.30
1	H	288	LEU	CA-CB-CG	5.86	128.78	115.30
1	C	143	ARG	NE-CZ-NH1	5.85	123.22	120.30
1	F	316	LEU	CA-CB-CG	5.79	128.63	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3777	0	3747	50	0
1	B	3771	0	3741	47	0
1	C	3776	0	3745	53	0
1	D	3771	0	3741	56	0
1	E	3771	0	3741	55	0
1	F	3771	0	3741	78	0
1	G	3737	0	3670	85	0
1	H	3723	0	3668	81	0
2	A	43	0	30	2	0
2	B	43	0	30	1	0
2	C	43	0	30	3	0
2	D	43	0	30	2	0
2	E	43	0	30	3	0
2	F	43	0	30	9	0
2	G	43	0	30	6	0
2	H	43	0	30	6	0
3	A	15	0	13	4	0
3	B	15	0	13	5	0
3	C	15	0	13	4	0
3	D	15	0	13	5	0
3	E	15	0	13	3	0
3	F	15	0	13	9	0
4	G	6	0	8	2	0
4	H	6	0	8	0	0
5	A	56	0	0	0	0
5	B	39	0	0	3	0
5	C	71	0	0	4	0
5	D	49	0	0	3	0
5	E	13	0	0	0	0
5	F	11	0	0	0	0
5	G	12	0	0	1	0
5	H	10	0	0	0	0
All	All	30804	0	30128	524	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

The worst 5 of 524 close contacts within the same asymmetric unit are listed below, sorted by

their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:32:LYS:HD2	1:C:33:LEU:H	1.09	1.07
1:D:381:ARG:O	1:D:382:ASP:HB2	1.55	1.03
1:D:143:ARG:HH12	1:D:147:GLU:HB2	1.22	1.03
3:B:502:0QA:H5	3:B:502:0QA:O1	1.58	0.99
1:A:77:ARG:HG2	1:A:77:ARG:HH11	1.26	0.98

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	464/476 (98%)	439 (95%)	24 (5%)	1 (0%)	47	56
1	B	463/476 (97%)	445 (96%)	18 (4%)	0	100	100
1	C	464/476 (98%)	443 (96%)	20 (4%)	1 (0%)	47	56
1	D	463/476 (97%)	444 (96%)	18 (4%)	1 (0%)	47	56
1	E	463/476 (97%)	434 (94%)	28 (6%)	1 (0%)	47	56
1	F	463/476 (97%)	413 (89%)	44 (10%)	6 (1%)	12	10
1	G	461/476 (97%)	417 (90%)	36 (8%)	8 (2%)	9	7
1	H	456/476 (96%)	396 (87%)	53 (12%)	7 (2%)	10	8
All	All	3697/3808 (97%)	3431 (93%)	241 (6%)	25 (1%)	22	23

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	346	ARG
1	G	364	ASP
1	G	468	PRO
1	A	369	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	162	GLY

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	410/419 (98%)	391 (95%)	19 (5%)	27	32
1	B	409/419 (98%)	392 (96%)	17 (4%)	30	36
1	C	410/419 (98%)	394 (96%)	16 (4%)	32	40
1	D	409/419 (98%)	388 (95%)	21 (5%)	24	27
1	E	409/419 (98%)	380 (93%)	29 (7%)	14	15
1	F	409/419 (98%)	383 (94%)	26 (6%)	17	18
1	G	403/419 (96%)	375 (93%)	28 (7%)	15	15
1	H	403/419 (96%)	354 (88%)	49 (12%)	5	4
All	All	3262/3352 (97%)	3057 (94%)	205 (6%)	18	19

5 of 205 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	289	LYS
1	G	266	ASP
1	H	470	ASP
1	F	348	LYS
1	G	52	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	255	ASN
1	G	126	GLN
1	F	422	GLN
1	F	283	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	104	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

16 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	0QA	E	502	2	13,15,15	0.71	0	14,18,18	1.55	3 (21%)
2	HEM	B	501	1	41,50,50	2.19	15 (36%)	45,82,82	1.74	12 (26%)
2	HEM	E	501	1,3	41,50,50	2.05	8 (19%)	45,82,82	1.60	9 (20%)
4	GOL	H	502	-	5,5,5	1.24	1 (20%)	5,5,5	2.56	2 (40%)
3	0QA	D	502	-	13,15,15	0.61	0	14,18,18	1.54	4 (28%)
2	HEM	C	501	1,3	41,50,50	2.24	12 (29%)	45,82,82	1.91	9 (20%)
3	0QA	C	502	2	13,15,15	0.61	0	14,18,18	1.53	4 (28%)
3	0QA	F	502	2	13,15,15	0.53	0	14,18,18	1.44	2 (14%)
2	HEM	F	501	1,3	41,50,50	2.06	9 (21%)	45,82,82	1.69	8 (17%)
3	0QA	A	502	2	13,15,15	0.44	0	14,18,18	2.12	4 (28%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	HEM	D	501	1	41,50,50	2.12	12 (29%)	45,82,82	1.77	7 (15%)
3	0QA	B	502	-	13,15,15	0.60	0	14,18,18	1.96	5 (35%)
2	HEM	A	501	1,3	41,50,50	2.12	10 (24%)	45,82,82	1.72	9 (20%)
4	GOL	G	502	-	5,5,5	0.58	0	5,5,5	0.53	0
2	HEM	H	501	1	41,50,50	2.04	8 (19%)	45,82,82	1.54	6 (13%)
2	HEM	G	501	1	41,50,50	2.30	10 (24%)	45,82,82	1.71	12 (26%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	0QA	E	502	2	-	6/12/12/12	0/1/1/1
2	HEM	B	501	1	-	2/12/54/54	-
2	HEM	E	501	1,3	-	2/12/54/54	-
4	GOL	H	502	-	-	2/4/4/4	-
3	0QA	D	502	-	-	7/12/12/12	0/1/1/1
2	HEM	C	501	1,3	-	4/12/54/54	-
3	0QA	C	502	2	-	7/12/12/12	0/1/1/1
3	0QA	F	502	2	-	6/12/12/12	0/1/1/1
2	HEM	F	501	1,3	-	3/12/54/54	-
3	0QA	A	502	2	-	5/12/12/12	0/1/1/1
2	HEM	D	501	1	-	2/12/54/54	-
3	0QA	B	502	-	-	6/12/12/12	0/1/1/1
2	HEM	A	501	1,3	-	4/12/54/54	-
4	GOL	G	502	-	-	2/4/4/4	-
2	HEM	H	501	1	-	4/12/54/54	-
2	HEM	G	501	1	-	4/12/54/54	-

The worst 5 of 85 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	501	HEM	C3D-C2D	8.65	1.55	1.36
2	G	501	HEM	C3D-C2D	8.42	1.54	1.36
2	E	501	HEM	C3D-C2D	7.94	1.53	1.36
2	H	501	HEM	C3D-C2D	7.92	1.53	1.36
2	A	501	HEM	C3D-C2D	7.67	1.53	1.36

The worst 5 of 96 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	501	HEM	C4D-ND-C1D	7.11	112.42	105.07
2	A	501	HEM	C4D-ND-C1D	5.70	110.96	105.07
2	G	501	HEM	C4D-ND-C1D	5.54	110.80	105.07
2	D	501	HEM	C4B-CHC-C1C	5.29	129.54	122.56
2	D	501	HEM	C4D-ND-C1D	5.22	110.46	105.07

There are no chirality outliers.

5 of 66 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	0QA	C10-N2-N3-O2
3	A	502	0QA	C8-C9-N2-N3
3	A	502	0QA	C6-C7-C8-C9
3	B	502	0QA	C10-N2-N3-O2
3	B	502	0QA	C8-C9-N2-N3

There are no ring outliers.

15 monomers are involved in 57 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	502	0QA	3	0
2	B	501	HEM	1	0
2	E	501	HEM	3	0
3	D	502	0QA	5	0
2	C	501	HEM	3	0
3	C	502	0QA	4	0
3	F	502	0QA	9	0
2	F	501	HEM	9	0
3	A	502	0QA	4	0
2	D	501	HEM	2	0
3	B	502	0QA	5	0
2	A	501	HEM	2	0
4	G	502	GOL	2	0
2	H	501	HEM	6	0
2	G	501	HEM	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand HEM B 501

Bond lengths   Bond angles

Torsions   Rings

Ligand HEM E 501

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM C 501

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM F 501

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM D 501

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM A 501

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM H 501

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM G 501

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	464/476 (97%)	0.21	0	100	100	20, 35, 53, 63	0
1	B	464/476 (97%)	0.21	0	100	100	19, 35, 54, 62	0
1	C	464/476 (97%)	0.16	1 (0%)	95	97	19, 33, 48, 58	0
1	D	464/476 (97%)	0.19	1 (0%)	95	97	17, 35, 55, 63	0
1	E	464/476 (97%)	0.25	11 (2%)	59	68	27, 48, 70, 78	0
1	F	464/476 (97%)	0.63	32 (6%)	16	24	27, 60, 85, 91	0
1	G	463/476 (97%)	0.65	34 (7%)	15	22	29, 59, 78, 88	0
1	H	459/476 (96%)	1.01	86 (18%)	1	2	28, 71, 92, 96	0
All	All	3706/3808 (97%)	0.41	165 (4%)	33	46	17, 44, 81, 96	0

The worst 5 of 165 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	141	GLY	6.4
1	H	491	PHE	6.1
1	F	491	PHE	5.2
1	H	468	PRO	4.9
1	H	453	PHE	4.8

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
3	0QA	E	502	15/15	0.90	0.28	71,74,76,78	0
3	0QA	C	502	15/15	0.91	0.22	48,55,58,67	0
3	0QA	F	502	15/15	0.91	0.35	84,85,86,87	0
3	0QA	B	502	15/15	0.92	0.21	42,51,59,59	0
4	GOL	G	502	6/6	0.92	0.16	54,56,58,58	0
3	0QA	D	502	15/15	0.94	0.21	48,55,58,67	0
3	0QA	A	502	15/15	0.94	0.24	57,62,71,72	0
4	GOL	H	502	6/6	0.94	0.27	33,36,42,50	0
2	HEM	H	501	43/43	0.96	0.16	47,52,60,63	0
2	HEM	F	501	43/43	0.97	0.16	37,49,54,59	0
2	HEM	E	501	43/43	0.98	0.15	30,35,37,39	0
2	HEM	A	501	43/43	0.98	0.17	21,27,31,36	0
2	HEM	G	501	43/43	0.98	0.15	33,38,43,47	0
2	HEM	B	501	43/43	0.98	0.17	18,25,29,32	0
2	HEM	C	501	43/43	0.98	0.15	12,22,27,28	0
2	HEM	D	501	43/43	0.98	0.16	18,26,29,30	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HEM H 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around HEM F 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
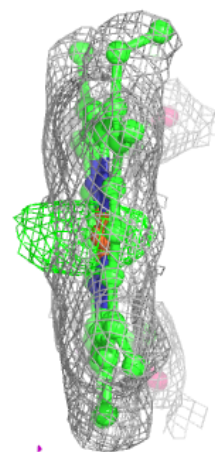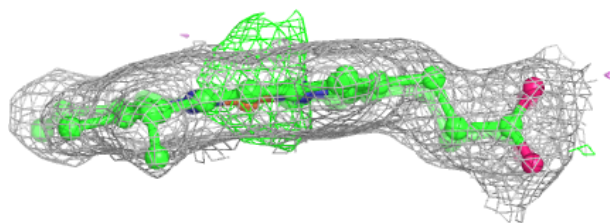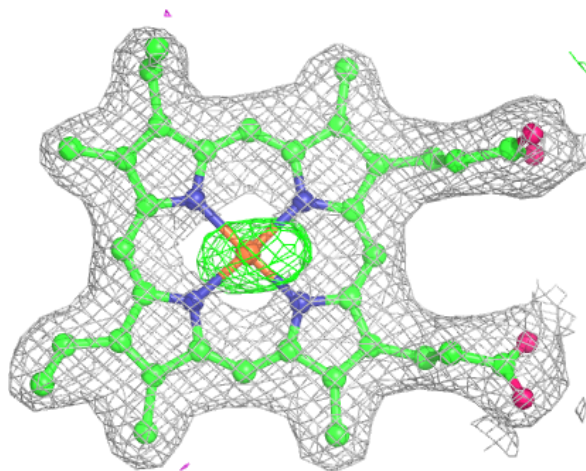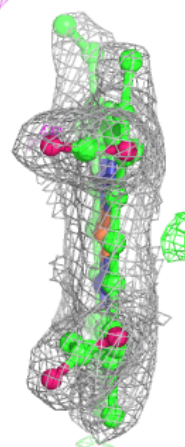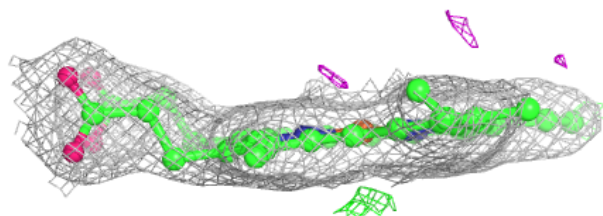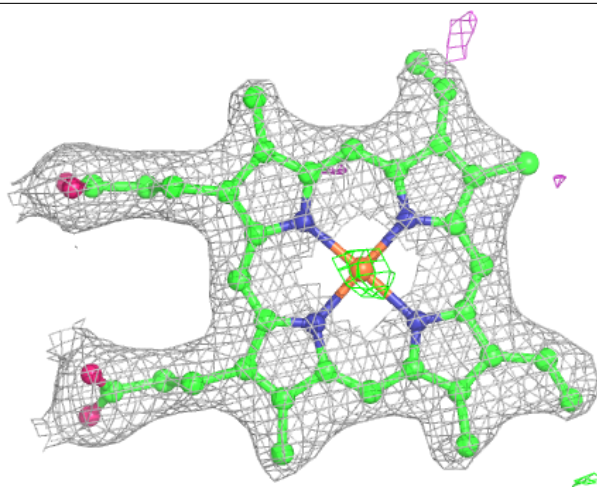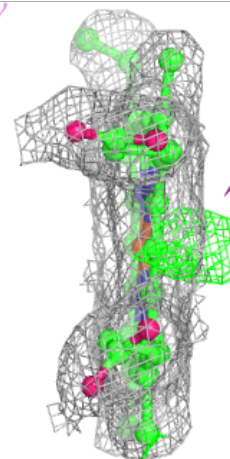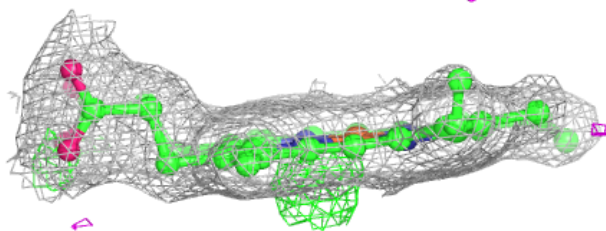$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
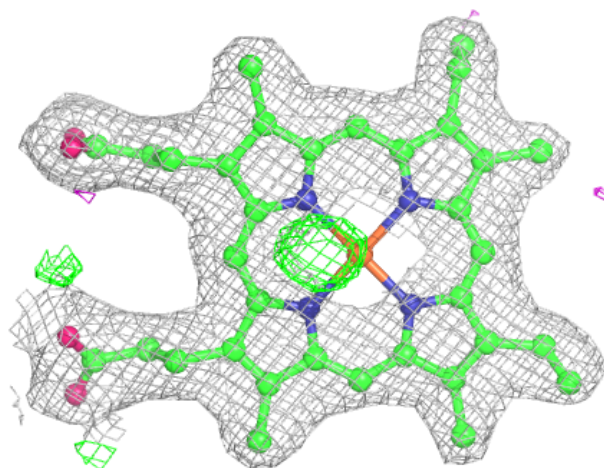and green (positive)

Electron density around HEM E 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around HEM A 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
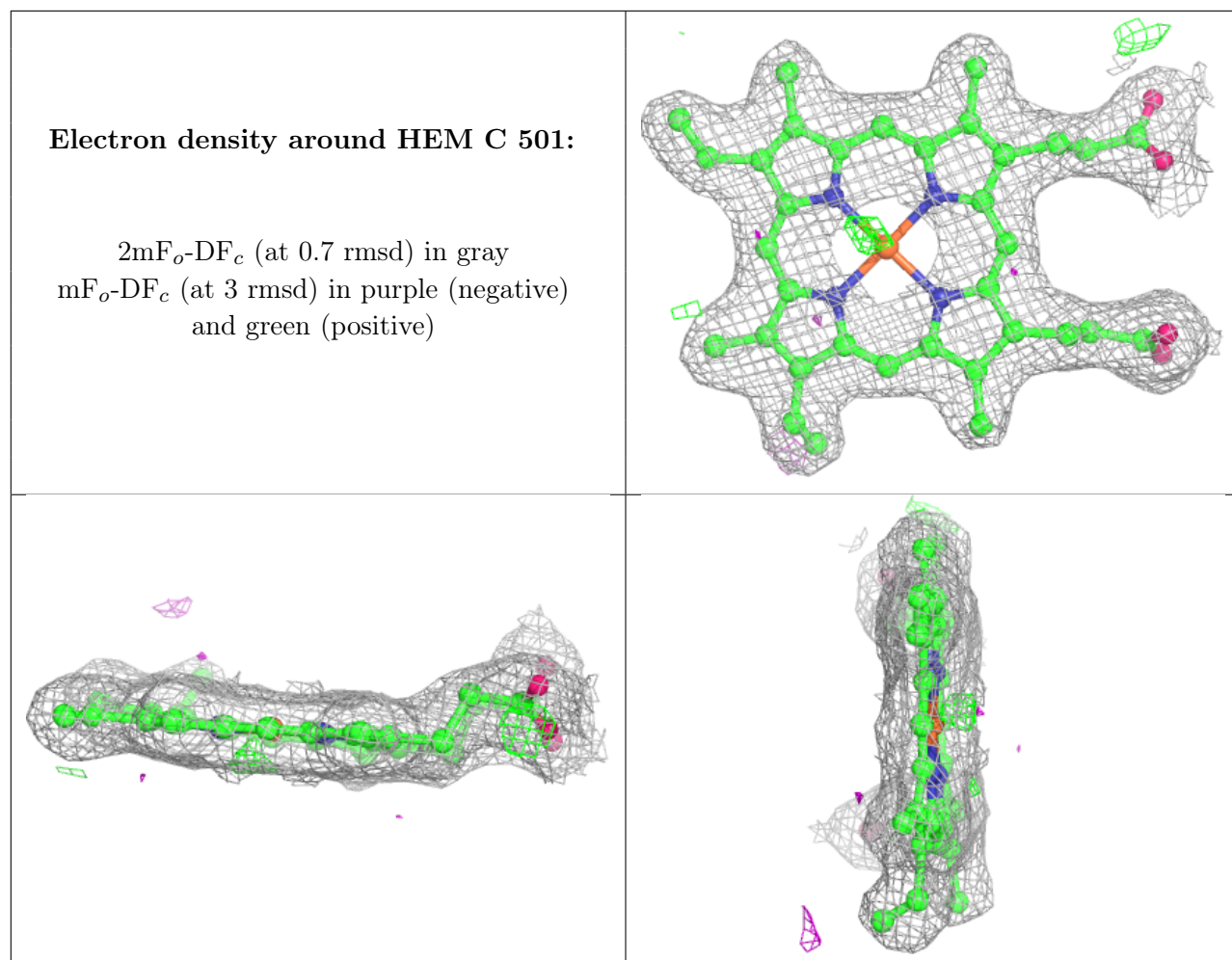and green (positive)

Electron density around HEM G 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around HEM B 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around HEM C 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
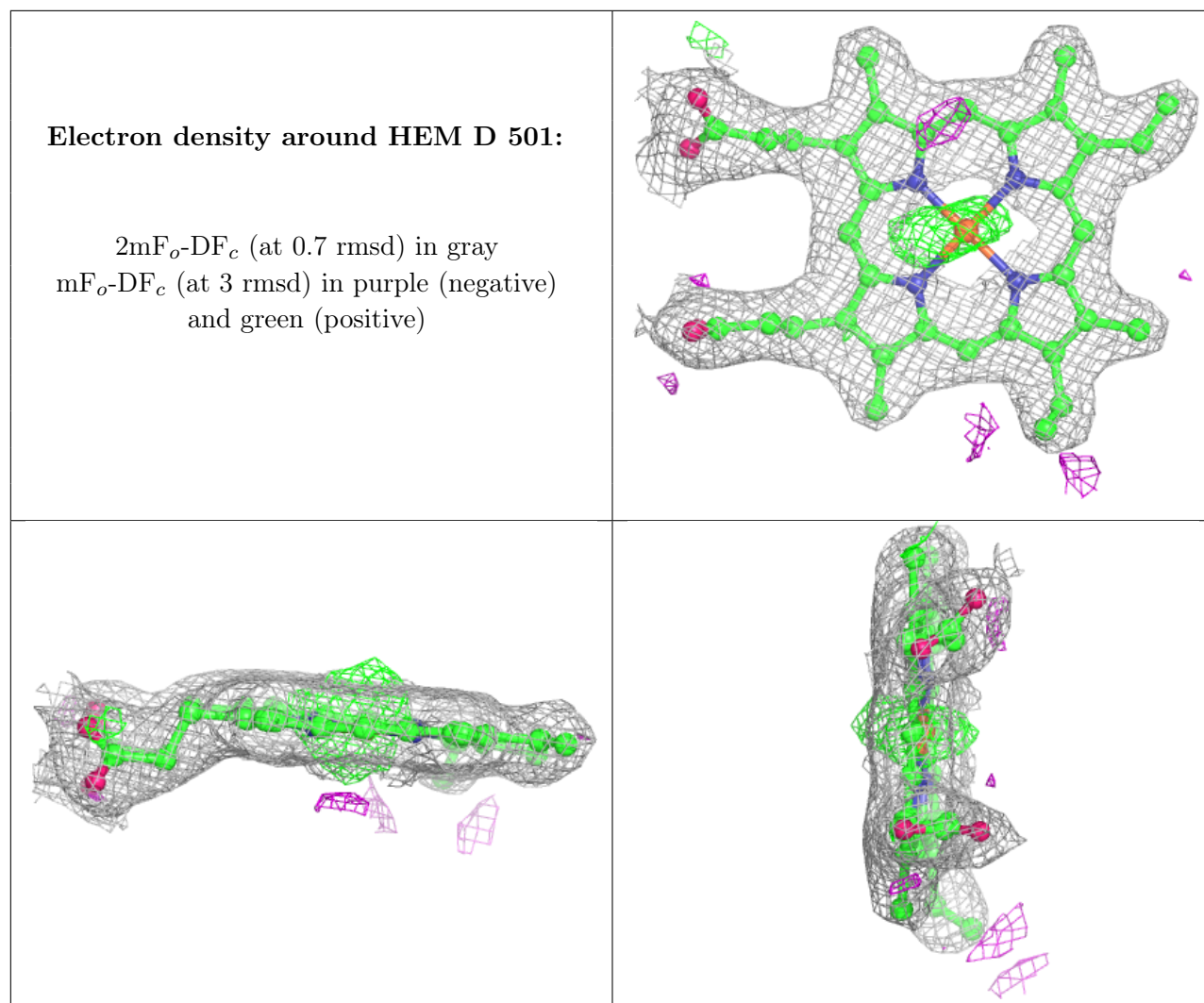and green (positive)

Electron density around HEM D 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers (i)

There are no such residues in this entry.