



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 18, 2019 – 10:48 PM EDT

PDB ID : 6CSD
Title : V308E mutant of cytochrome P450 2D6 complexed with prinomastat
Authors : Yang, Y.T.; Fujita, K.; Wang, P.F.; Im, S.C.; Pearl, N.M.; Meagher, J.;
Stuckey, J.; Waskell, L.
Deposited on : 2018-03-20
Resolution : 2.39 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.0 (224370), CSD as540be (2019)
Xtriage (Phenix) : 1.13
EDS : rb-20031633
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031633

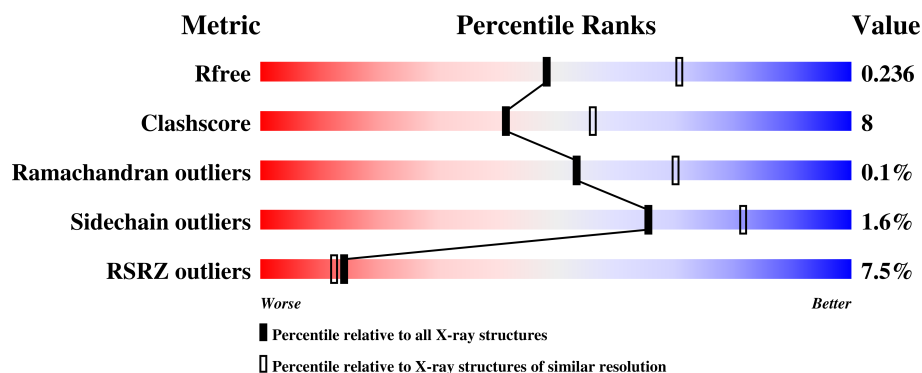
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	3481 (2.40-2.40)
Clashscore	122126	3956 (2.40-2.40)
Ramachandran outliers	120053	3897 (2.40-2.40)
Sidechain outliers	120020	3898 (2.40-2.40)
RSRZ outliers	108989	3386 (2.40-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	479	<div> <div>7%</div> <div>82%</div> <div>14%</div> <div>• •</div> </div>
1	B	479	<div> <div>8%</div> <div>78%</div> <div>15%</div> <div>• 5%</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7727 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 2D6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	462	Total	C	N	O	S	0	3	0
			3570	2276	637	643	14			
1	B	454	Total	C	N	O	S	0	3	0
			3548	2270	627	637	14			

There are 32 discrepancies between the modelled and reference sequences:

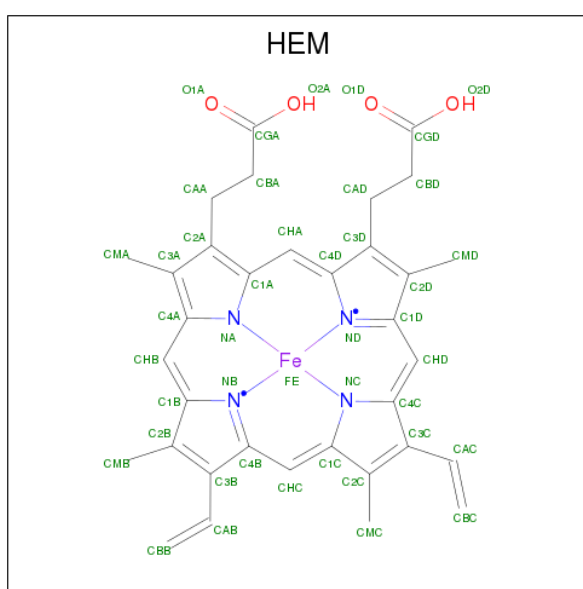
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	initiating methionine	UNP P10635
A	24	ALA	-	expression tag	UNP P10635
A	25	LYS	-	expression tag	UNP P10635
A	26	LYS	-	expression tag	UNP P10635
A	27	THR	-	expression tag	UNP P10635
A	28	SER	-	expression tag	UNP P10635
A	29	SER	-	expression tag	UNP P10635
A	30	LYS	-	expression tag	UNP P10635
A	31	GLY	-	expression tag	UNP P10635
A	32	LYS	-	expression tag	UNP P10635
A	33	LEU	-	expression tag	UNP P10635
A	308	GLU	VAL	engineered mutation	UNP P10635
A	498	HIS	-	expression tag	UNP P10635
A	499	HIS	-	expression tag	UNP P10635
A	500	HIS	-	expression tag	UNP P10635
A	501	HIS	-	expression tag	UNP P10635
B	23	MET	-	initiating methionine	UNP P10635
B	24	ALA	-	expression tag	UNP P10635
B	25	LYS	-	expression tag	UNP P10635
B	26	LYS	-	expression tag	UNP P10635
B	27	THR	-	expression tag	UNP P10635
B	28	SER	-	expression tag	UNP P10635
B	29	SER	-	expression tag	UNP P10635
B	30	LYS	-	expression tag	UNP P10635
B	31	GLY	-	expression tag	UNP P10635

Continued on next page...

Continued from previous page...

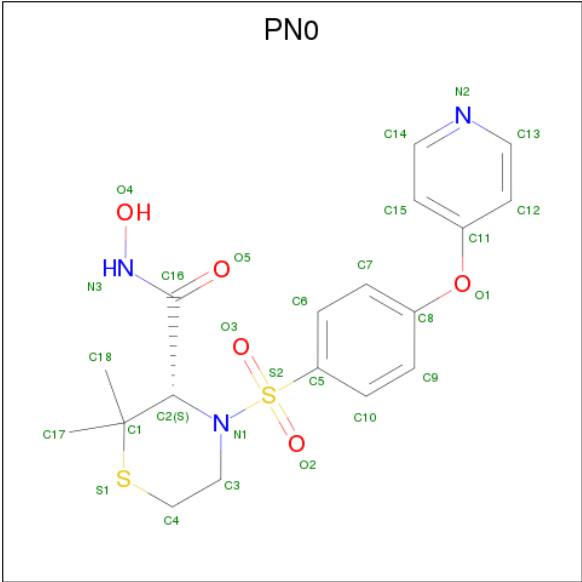
Chain	Residue	Modelled	Actual	Comment	Reference
B	32	LYS	-	expression tag	UNP P10635
B	33	LEU	-	expression tag	UNP P10635
B	308	GLU	VAL	engineered mutation	UNP P10635
B	498	HIS	-	expression tag	UNP P10635
B	499	HIS	-	expression tag	UNP P10635
B	500	HIS	-	expression tag	UNP P10635
B	501	HIS	-	expression tag	UNP P10635

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



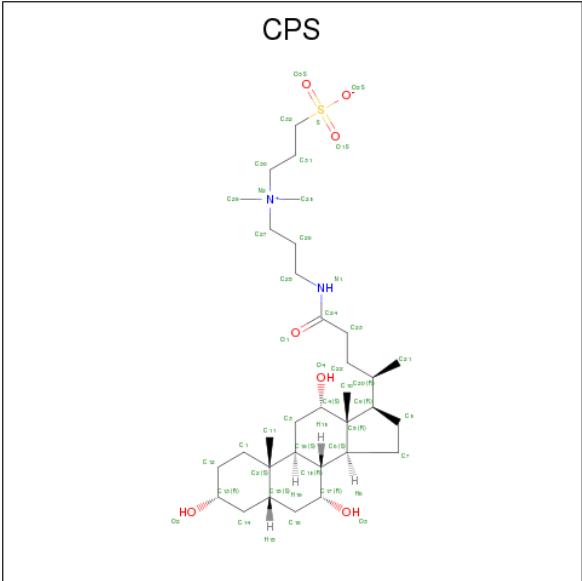
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	
			43	34	1	4	4	0
2	B	1	Total	C	Fe	N	O	
			43	34	1	4	4	0

- Molecule 3 is Prinomastat (three-letter code: PN0) (formula: $C_{18}H_{21}N_3O_5S_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			28	18	3	5	2		
3	B	1	Total	C	N	O	S	0	0
			28	18	3	5	2		

- Molecule 4 is 3-[(3-CHOLAMIDOPROPYL)DIMETHYLAMMONIO]-1-PROPANESULFO NATE (three-letter code: CPS) (formula: C₃₂H₅₈N₂O₇S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			24	21	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			24	21	3		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	3	Total	Zn	0	0
			3	3		
5	A	3	Total	Zn	0	0
			3	3		

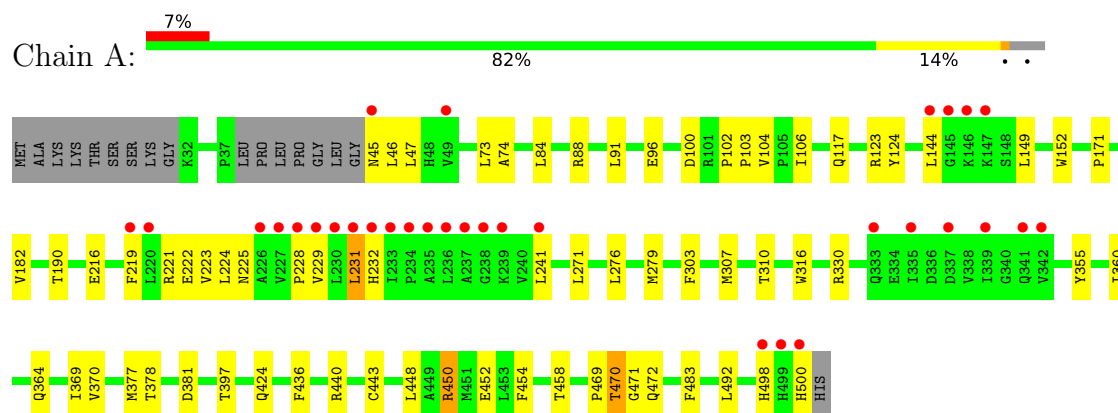
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	226	Total	O	0	0
			226	226		
6	B	187	Total	O	0	0
			187	187		

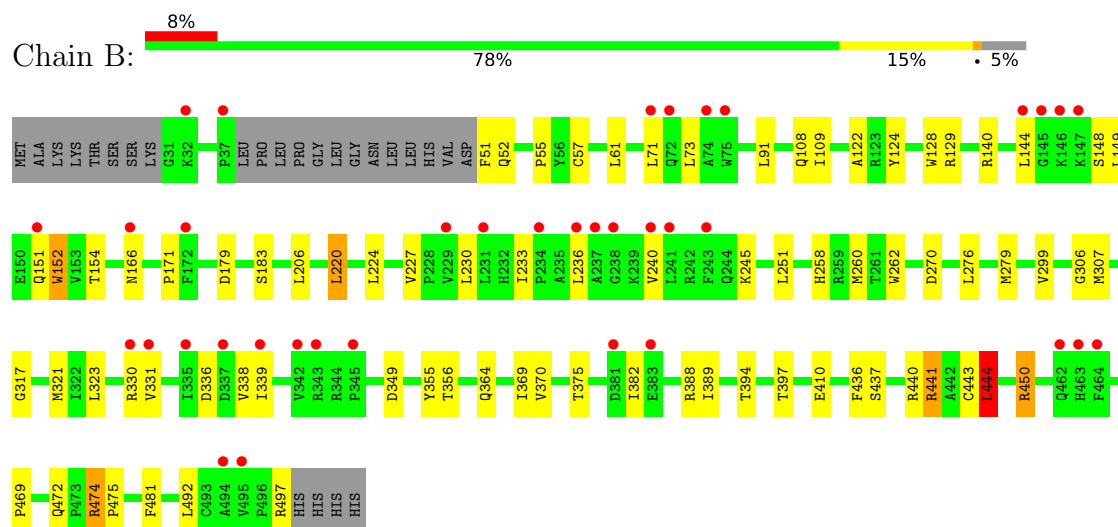
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 2D6



• Molecule 1: Cytochrome P450 2D6



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	57.47 Å 126.53 Å 192.26 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.06 – 2.39 48.06 – 2.39	Depositor EDS
% Data completeness (in resolution range)	94.4 (48.06-2.39) 94.6 (48.06-2.39)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 2.39 Å)	Xtriage
Refinement program	PHENIX (1.12_2829)	Depositor
R, R_{free}	0.198 , 0.236 0.200 , 0.236	Depositor DCC
R_{free} test set	2628 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	33.3	Xtriage
Anisotropy	0.728	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 57.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7727	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.86 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.2373e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: PN0, ZN, CPS, HEM

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	0.42	0/3673	0.60	1/5005 (0.0%)
1	B	0.40	0/3654	0.58	2/4973 (0.0%)
All	All	0.41	0/7327	0.59	3/9978 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	231	LEU	CB-CG-CD2	7.46	123.69	111.00
1	B	441	ARG	NE-CZ-NH2	-5.52	117.54	120.30
1	B	444	LEU	CA-CB-CG	5.51	127.97	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	3570	0	3437	51	0
1	B	3548	0	3469	55	1
2	A	43	0	30	2	0
2	B	43	0	30	7	0
3	A	28	0	21	1	0
3	B	28	0	21	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	24	0	34	2	0
4	B	24	0	34	0	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
6	A	226	0	0	3	1
6	B	187	0	0	4	1
All	All	7727	0	7076	110	2

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 8.

The worst 5 of 110 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:A:228:PRO:O	1:A:231:LEU:HD23	1.44	1.13
1:B:129:ARG:NH1	6:B:701:HOH:O	2.03	0.91
1:B:122:ALA:O	1:B:441:ARG:NH2	2.15	0.77
1:A:450:ARG:NH1	6:A:702:HOH:O	2.17	0.77
1:A:104:VAL:HA	1:A:225:ASN:HD21	1.50	0.76

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:753:HOH:O	6:A:881:HOH:O[2_885]	1.96	0.24
1:B:260:MET:O	6:B:701:HOH:O[4_475]	2.16	0.04

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	461/479 (96%)	443 (96%)	17 (4%)	1 (0%)	49	65
1	B	454/479 (95%)	447 (98%)	7 (2%)	0	100	100
All	All	915/958 (96%)	890 (97%)	24 (3%)	1 (0%)	53	70

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	470	THR

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	369/409 (90%)	365 (99%)	4 (1%)	76	88
1	B	374/409 (91%)	366 (98%)	8 (2%)	56	75
All	All	743/818 (91%)	731 (98%)	12 (2%)	65	82

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	152	TRP
1	B	166	ASN
1	B	444	LEU
1	B	108	GLN
1	B	388	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	117	GLN
1	A	472	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 for Mogul analysis.

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$
2	HEM	A	601	1,3	27,50,50	1.79	6 (22%)	17,82,82	1.80	5 (29%)
3	PN0	A	602	2	28,30,30	0.21	0	35,44,44	0.42	0
4	CPS	A	603	-	26,27,45	0.41	0	41,44,70	0.82	1 (2%)
2	HEM	B	601	1,3	27,50,50	1.86	5 (18%)	17,82,82	1.75	6 (35%)
3	PN0	B	602	2	28,30,30	0.28	0	35,44,44	0.50	0
4	CPS	B	603	-	26,27,45	0.21	0	41,44,70	0.47	0

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
2	HEM	A	601	1,3	-	0/6/54/54	0/0/8/8
3	PN0	A	602	2	-	0/21/39/39	0/3/3/3
4	CPS	A	603	-	-	0/2/67/90	0/4/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	B	601	1,3	-	0/6/54/54	0/0/8/8
3	PN0	B	602	2	-	0/21/39/39	0/3/3/3
4	CPS	B	603	-	-	0/2/67/90	0/4/4/4

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	HEM	C3B-C2B	-4.26	1.34	1.40
2	B	601	HEM	C3C-C2C	-4.18	1.34	1.40
2	A	601	HEM	C3C-C2C	-3.41	1.35	1.40
2	A	601	HEM	C3B-C2B	-3.32	1.35	1.40
2	A	601	HEM	CMB-C2B	2.13	1.56	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	HEM	CAA-CBA-CGA	-3.67	106.39	112.66
2	B	601	HEM	CMD-C2D-C1D	-3.65	122.85	128.46
2	A	601	HEM	CBD-CAD-C3D	-2.93	106.88	112.47
2	A	601	HEM	CMD-C2D-C1D	-2.71	124.30	128.46
2	B	601	HEM	CBD-CAD-C3D	-2.45	107.80	112.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	HEM	2	0
3	A	602	PN0	1	0
4	A	603	CPS	2	0
2	B	601	HEM	7	0
3	B	602	PN0	1	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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(Å ²)	Q<0.9
1	A	462/479 (96%)	0.18	32 (6%) 17 15	20, 40, 85, 109	0
1	B	454/479 (94%)	0.30	37 (8%) 12 10	24, 44, 79, 95	0
All	All	916/958 (95%)	0.24	69 (7%) 14 12	20, 42, 80, 109	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	235	ALA	6.5
1	B	145	GLY	6.2
1	A	236	LEU	6.0
1	B	75	TRP	4.9
1	A	45	ASN	4.9

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

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(\AA^2)	Q<0.9
5	ZN	B	605	1/1	0.82	0.19	120,120,120,120	0
4	CPS	A	603	24/42	0.83	0.25	49,69,77,79	0
5	ZN	A	605	1/1	0.84	0.08	77,77,77,77	0
5	ZN	B	606	1/1	0.94	0.51	103,103,103,103	0
4	CPS	B	603	24/42	0.95	0.10	43,53,64,66	0
3	PN0	B	602	28/28	0.97	0.11	27,42,52,57	0
3	PN0	A	602	28/28	0.97	0.11	22,39,48,51	0
2	HEM	B	601	43/43	0.98	0.17	22,28,32,34	0
2	HEM	A	601	43/43	0.98	0.13	17,25,30,33	0
5	ZN	A	606	1/1	0.99	0.14	42,42,42,42	1
5	ZN	A	604	1/1	1.00	0.10	29,29,29,29	0
5	ZN	B	604	1/1	1.00	0.14	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.