



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:03 AM GMT

PDB ID : 2FDU
Title : Microsomal P450 2A6 with the inhibitor N,N-Dimethyl(5-(pyridin-3-yl)furan-2-yl)methanamine bound
Authors : Yano, J.K.; Stout, C.D.; Johnson, E.F.
Deposited on : 2005-12-14
Resolution : 1.85 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

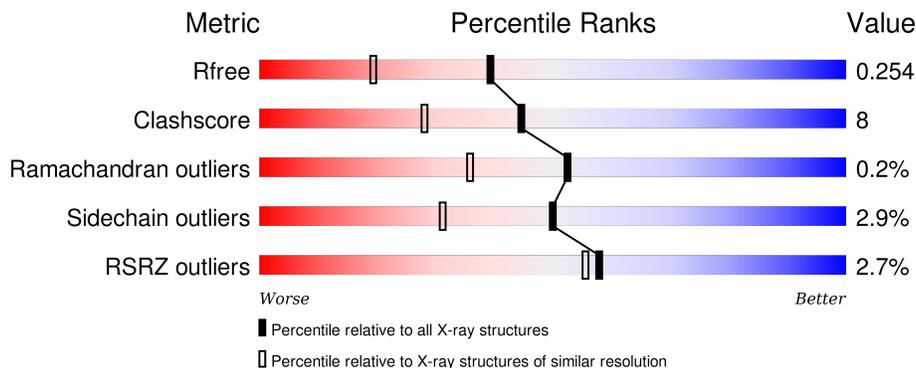
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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}	91344	1745 (1.86-1.86)
Clashscore	102246	1898 (1.86-1.86)
Ramachandran outliers	100387	1875 (1.86-1.86)
Sidechain outliers	100360	1875 (1.86-1.86)
RSRZ outliers	91569	1747 (1.86-1.86)

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	476	 3% 86% 12% ..
1	B	476	 4% 79% 15% ..
1	C	476	 % 87% 9% ..
1	D	476	 3% 84% 13% ..

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 16510 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 2A6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	465	3760	2414	650	678	18	0	0	0
1	B	463	3747	2406	647	676	18	0	0	0
1	C	464	3751	2408	648	677	18	0	0	0
1	D	464	3751	2408	648	677	18	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

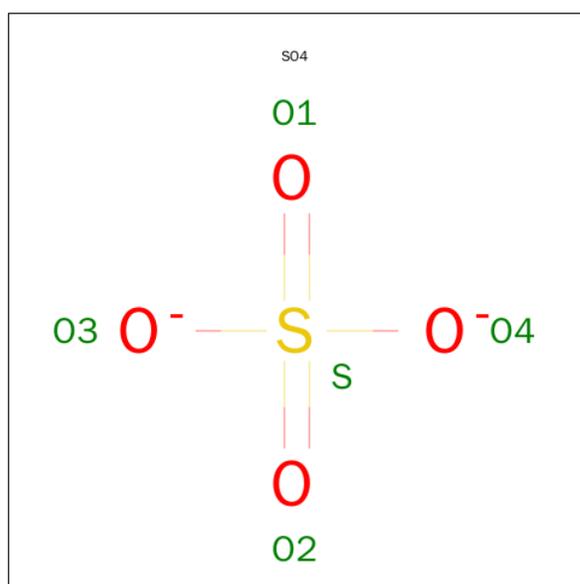
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	CLONING ARTIFACT	UNP P11509
A	24	ALA	-	CLONING ARTIFACT	UNP P11509
A	25	LYS	-	CLONING ARTIFACT	UNP P11509
A	26	LYS	-	CLONING ARTIFACT	UNP P11509
A	27	THR	-	CLONING ARTIFACT	UNP P11509
A	28	SER	-	CLONING ARTIFACT	UNP P11509
A	495	HIS	-	EXPRESSION TAG	UNP P11509
A	496	HIS	-	EXPRESSION TAG	UNP P11509
A	497	HIS	-	EXPRESSION TAG	UNP P11509
A	498	HIS	-	EXPRESSION TAG	UNP P11509
B	23	MET	-	CLONING ARTIFACT	UNP P11509
B	24	ALA	-	CLONING ARTIFACT	UNP P11509
B	25	LYS	-	CLONING ARTIFACT	UNP P11509
B	26	LYS	-	CLONING ARTIFACT	UNP P11509
B	27	THR	-	CLONING ARTIFACT	UNP P11509
B	28	SER	-	CLONING ARTIFACT	UNP P11509
B	495	HIS	-	EXPRESSION TAG	UNP P11509
B	496	HIS	-	EXPRESSION TAG	UNP P11509
B	497	HIS	-	EXPRESSION TAG	UNP P11509
B	498	HIS	-	EXPRESSION TAG	UNP P11509
C	23	MET	-	CLONING ARTIFACT	UNP P11509

Continued on next page...

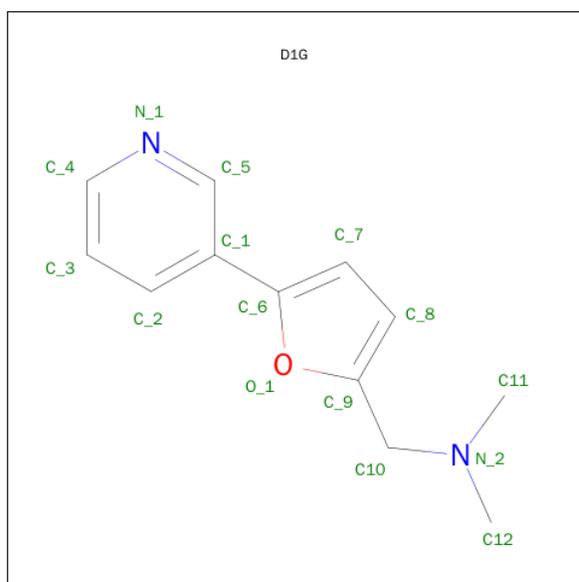
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	24	ALA	-	CLONING ARTIFACT	UNP P11509
C	25	LYS	-	CLONING ARTIFACT	UNP P11509
C	26	LYS	-	CLONING ARTIFACT	UNP P11509
C	27	THR	-	CLONING ARTIFACT	UNP P11509
C	28	SER	-	CLONING ARTIFACT	UNP P11509
C	495	HIS	-	EXPRESSION TAG	UNP P11509
C	496	HIS	-	EXPRESSION TAG	UNP P11509
C	497	HIS	-	EXPRESSION TAG	UNP P11509
C	498	HIS	-	EXPRESSION TAG	UNP P11509
D	23	MET	-	CLONING ARTIFACT	UNP P11509
D	24	ALA	-	CLONING ARTIFACT	UNP P11509
D	25	LYS	-	CLONING ARTIFACT	UNP P11509
D	26	LYS	-	CLONING ARTIFACT	UNP P11509
D	27	THR	-	CLONING ARTIFACT	UNP P11509
D	28	SER	-	CLONING ARTIFACT	UNP P11509
D	495	HIS	-	EXPRESSION TAG	UNP P11509
D	496	HIS	-	EXPRESSION TAG	UNP P11509
D	497	HIS	-	EXPRESSION TAG	UNP P11509
D	498	HIS	-	EXPRESSION TAG	UNP P11509

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0

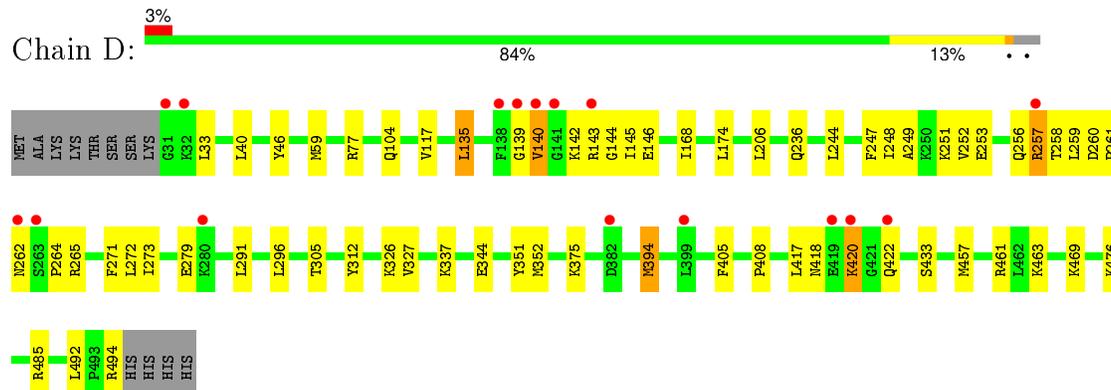


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			15	12	2	1		
4	B	1	Total	C	N	O	0	0
			15	12	2	1		
4	C	1	Total	C	N	O	0	0
			15	12	2	1		
4	D	1	Total	C	N	O	0	0
			15	12	2	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	346	Total	O	0	0
			346	346		
5	B	228	Total	O	0	0
			228	228		
5	C	368	Total	O	0	0
			368	368		
5	D	317	Total	O	0	0
			317	317		

- Molecule 1: Cytochrome P450 2A6



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	69.81Å 157.70Å 103.89Å 90.00° 92.23° 90.00°	Depositor
Resolution (Å)	50.00 – 1.85 47.22 – 1.79	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-1.85) 91.6 (47.22-1.79)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.73 (at 1.79Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.186 , 0.223 0.224 , 0.254	Depositor DCC
R_{free} test set	9103 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	27.9	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 46.7	EDS
Estimated twinning fraction	0.026 for h,-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 191950 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	16510	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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, D1G, SO4

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	0.89	0/3851	0.83	1/5186 (0.0%)
1	B	0.80	1/3838 (0.0%)	0.84	4/5170 (0.1%)
1	C	0.87	1/3842 (0.0%)	0.84	2/5175 (0.0%)
1	D	0.89	2/3842 (0.1%)	0.83	2/5175 (0.0%)
All	All	0.86	4/15373 (0.0%)	0.83	9/20706 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	394	MET	SD-CE	-14.46	0.96	1.77
1	D	59	MET	SD-CE	-7.67	1.34	1.77
1	C	103	GLU	CG-CD	5.86	1.60	1.51
1	B	103	GLU	CG-CD	5.00	1.59	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	234	GLN	CA-CB-CG	6.57	127.86	113.40
1	D	104	GLN	N-CA-C	-6.35	93.85	111.00
1	A	104	GLN	N-CA-C	-6.01	94.78	111.00
1	C	76	ARG	NE-CZ-NH2	-5.96	117.32	120.30
1	B	446	ARG	NE-CZ-NH1	5.93	123.27	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	46	TYR	Sidechain

5.2 Too-close contacts

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	3760	0	3747	51	0
1	B	3747	0	3731	80	0
1	C	3751	0	3734	47	0
1	D	3751	0	3734	61	0
2	A	10	0	0	0	0
3	A	43	0	30	5	0
3	B	43	0	30	1	0
3	C	43	0	30	4	0
3	D	43	0	30	4	0
4	A	15	0	14	3	0
4	B	15	0	14	3	0
4	C	15	0	14	3	0
4	D	15	0	14	4	0
5	A	346	0	0	8	0
5	B	228	0	0	10	0
5	C	368	0	0	5	0
5	D	317	0	0	5	0
All	All	16510	0	15122	241	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 8.

The worst 5 of 241 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:D:394:MET:CE	1:D:394:MET:CG	2.21	1.18
1:B:208:ILE:HD11	1:B:240:LEU:HB2	1.13	1.12
1:D:394:MET:SD	1:D:394:MET:CE	0.96	1.06

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:253:GLU:HG3	1:D:257:ARG:HH21	1.21	1.01
1:D:142:LYS:HG3	1:D:144:GLY:H	1.26	1.01

There are no symmetry-related clashes.

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	463/476 (97%)	452 (98%)	11 (2%)	0	100	100
1	B	461/476 (97%)	446 (97%)	14 (3%)	1 (0%)	52	36
1	C	462/476 (97%)	451 (98%)	11 (2%)	0	100	100
1	D	462/476 (97%)	442 (96%)	18 (4%)	2 (0%)	39	22
All	All	1848/1904 (97%)	1791 (97%)	54 (3%)	3 (0%)	52	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	140	VAL
1	B	261	PRO
1	D	261	PRO

5.3.2 Protein sidechains [i](#)

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	412/422 (98%)	402 (98%)	10 (2%)	57	39
1	B	411/422 (97%)	393 (96%)	18 (4%)	35	15
1	C	411/422 (97%)	403 (98%)	8 (2%)	65	49
1	D	411/422 (97%)	400 (97%)	11 (3%)	52	34
All	All	1645/1688 (98%)	1598 (97%)	47 (3%)	50	31

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	312	TYR
1	B	467	SER
1	D	337	LYS
1	B	417	LEU
1	B	468	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	GLN
1	A	283	ASN
1	C	418	ASN
1	C	422	GLN
1	D	340	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry

10 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
2	SO4	A	2502	-	4,4,4	0.94	0	6,6,6	0.42	0
2	SO4	A	2503	-	4,4,4	1.11	0	6,6,6	0.28	0
3	HEM	A	500	1,4	30,50,50	3.66	13 (43%)	24,82,82	2.61	11 (45%)
4	D1G	A	501	3	13,16,16	1.80	3 (23%)	15,21,21	1.50	1 (6%)
3	HEM	B	500	1,4	30,50,50	3.76	12 (40%)	24,82,82	2.48	9 (37%)
4	D1G	B	501	3	13,16,16	1.81	4 (30%)	15,21,21	1.43	1 (6%)
3	HEM	C	500	1,4	30,50,50	4.05	13 (43%)	24,82,82	2.65	11 (45%)
4	D1G	C	501	3	13,16,16	1.81	2 (15%)	15,21,21	1.30	1 (6%)
3	HEM	D	500	1,4	30,50,50	4.08	13 (43%)	24,82,82	2.55	11 (45%)
4	D1G	D	501	3	13,16,16	1.86	3 (23%)	15,21,21	1.87	2 (13%)

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	SO4	A	2502	-	-	0/0/0/0	0/0/0/0
2	SO4	A	2503	-	-	0/0/0/0	0/0/0/0
3	HEM	A	500	1,4	-	0/10/54/54	0/0/8/8
4	D1G	A	501	3	-	0/5/8/8	0/1/2/2
3	HEM	B	500	1,4	-	0/10/54/54	0/0/8/8
4	D1G	B	501	3	-	0/5/8/8	0/1/2/2
3	HEM	C	500	1,4	-	0/10/54/54	0/0/8/8
4	D1G	C	501	3	-	0/5/8/8	0/1/2/2
3	HEM	D	500	1,4	-	0/10/54/54	0/0/8/8
4	D1G	D	501	3	-	0/5/8/8	0/1/2/2

The worst 5 of 63 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	500	HEM	C3B-C4B	-14.34	1.39	1.51
3	C	500	HEM	C3B-C4B	-13.32	1.40	1.51
3	B	500	HEM	C3B-C4B	-11.26	1.41	1.51
3	A	500	HEM	C3D-C4D	-9.81	1.39	1.51
3	C	500	HEM	C3D-C4D	-9.14	1.39	1.51

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	500	HEM	CMA-C3A-C4A	-2.22	124.69	128.36
3	C	500	HEM	CAA-C2A-C1A	-2.22	124.60	127.01
3	C	500	HEM	C1D-CHD-C4C	2.03	129.22	125.82
4	D	501	D1G	C10-C_9-C_8	2.06	133.08	128.59
3	A	500	HEM	C2D-C3D-C4D	2.10	105.05	101.50

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	500	HEM	5	0
4	A	501	D1G	3	0
3	B	500	HEM	1	0
4	B	501	D1G	3	0
3	C	500	HEM	4	0
4	C	501	D1G	3	0
3	D	500	HEM	4	0
4	D	501	D1G	4	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	465/476 (97%)	0.38	13 (2%) 56 54	17, 28, 46, 66	0
1	B	463/476 (97%)	0.54	17 (3%) 45 43	20, 34, 55, 67	0
1	C	464/476 (97%)	0.06	4 (0%) 85 85	18, 28, 46, 65	0
1	D	464/476 (97%)	0.48	16 (3%) 49 46	18, 29, 55, 68	0
All	All	1856/1904 (97%)	0.37	50 (2%) 58 55	17, 30, 50, 68	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	31	GLY	8.4
1	A	30	LYS	4.7
1	A	31	GLY	4.0
1	D	139	GLY	3.8
1	D	31	GLY	3.6

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
4	D1G	D	501	15/15	0.93	0.15	1.68	29,34,35,35	0
3	HEM	C	500	43/43	0.97	0.13	0.71	18,22,24,26	0
3	HEM	B	500	43/43	0.96	0.14	0.51	19,24,27,31	0
3	HEM	D	500	43/43	0.94	0.15	0.36	17,21,23,30	0
4	D1G	A	501	15/15	0.91	0.16	0.10	23,27,30,30	0
4	D1G	C	501	15/15	0.94	0.12	0.08	29,31,35,35	0
2	SO4	A	2503	5/5	0.91	0.16	0.05	54,55,57,60	0
3	HEM	A	500	43/43	0.97	0.13	0.02	17,19,22,26	0
4	D1G	B	501	15/15	0.92	0.14	-0.16	30,34,36,36	0
2	SO4	A	2502	5/5	0.86	0.19	-	61,62,64,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.