



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:02 am GMT

PDB ID : 1KQF
Title : FORMATE DEHYDROGENASE N FROM E. COLI
Authors : Jormakka, M.; Tornroth, S.; Byrne, B.; Iwata, S.
Deposited on : 2002-01-05
Resolution : 1.60 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

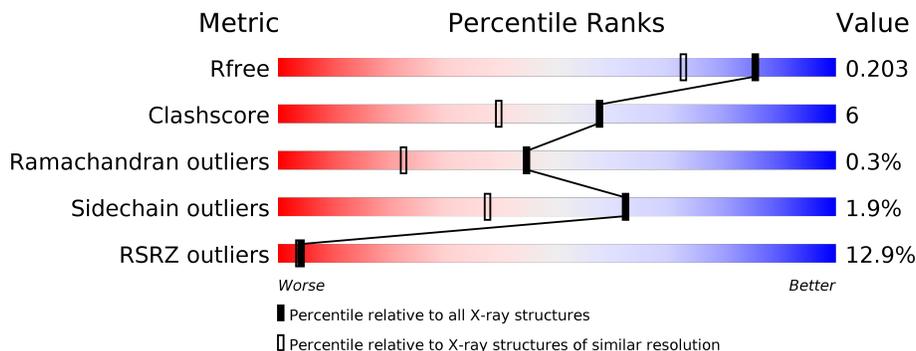
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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}	100719	2696 (1.60-1.60)
Clashscore	112137	2967 (1.60-1.60)
Ramachandran outliers	110173	2887 (1.60-1.60)
Sidechain outliers	110143	2886 (1.60-1.60)
RSRZ outliers	101464	2714 (1.60-1.60)

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	1015	 90% 7% •
2	B	294	 11% 86% 11% ••
3	C	217	 69% 72% 24% •

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 13989 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 FORMATE DEHYDROGENASE, NITRATE-INDUCIBLE, MAJOR SUBUNIT.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	982	7719	4872	1352	1457	37	1	0	0	0

- Molecule 2 is a protein called FORMATE DEHYDROGENASE, NITRATE-INDUCIBLE, IRON-SULFUR SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	289	2207	1383	381	421	22	0	0	0

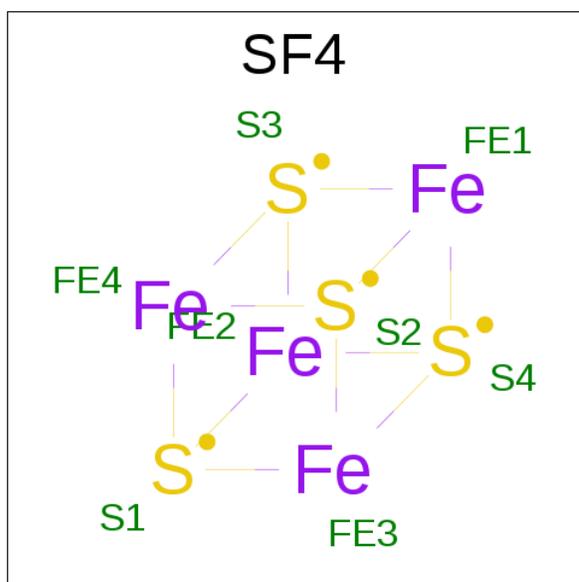
- Molecule 3 is a protein called FORMATE DEHYDROGENASE, NITRATE-INDUCIBLE, CYTOCHROME B556(FDN) SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	216	1783	1192	301	276	14	0	0	0

- Molecule 4 is MOLYBDENUM(VI) ION (three-letter code: 6MO) (formula: Mo).

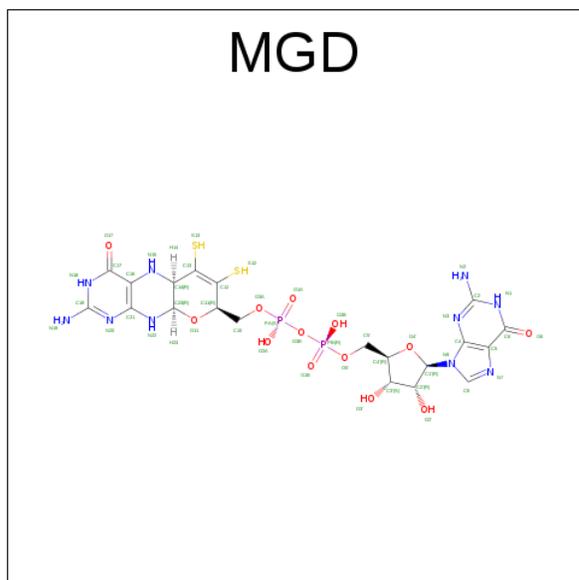
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mo	0	0
			1	1		

- Molecule 5 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



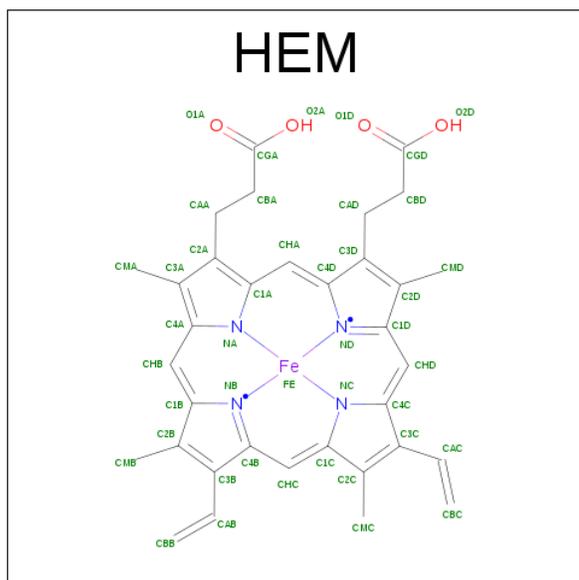
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Fe S	0	0
			8	4 4		
5	B	1	Total	Fe S	0	0
			8	4 4		
5	B	1	Total	Fe S	0	0
			8	4 4		
5	B	1	Total	Fe S	0	0
			8	4 4		

- Molecule 6 is 2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE (three-letter code: MGD) (formula: C₂₀H₂₆N₁₀O₁₃P₂S₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
6	A	1	47	20	10	13	2	2	0	0
6	A	1	47	20	10	13	2	2	0	0

- Molecule 7 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



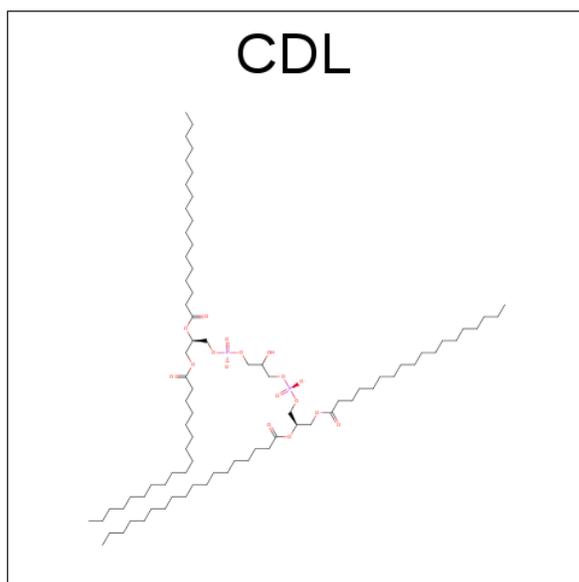
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	Fe	N			O
7	C	1	43	34	1	4	4	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 8 is CARDIOLIPIN (three-letter code: CDL) (formula: $C_{81}H_{156}O_{17}P_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
8	C	1	70	51	17	2	0	0

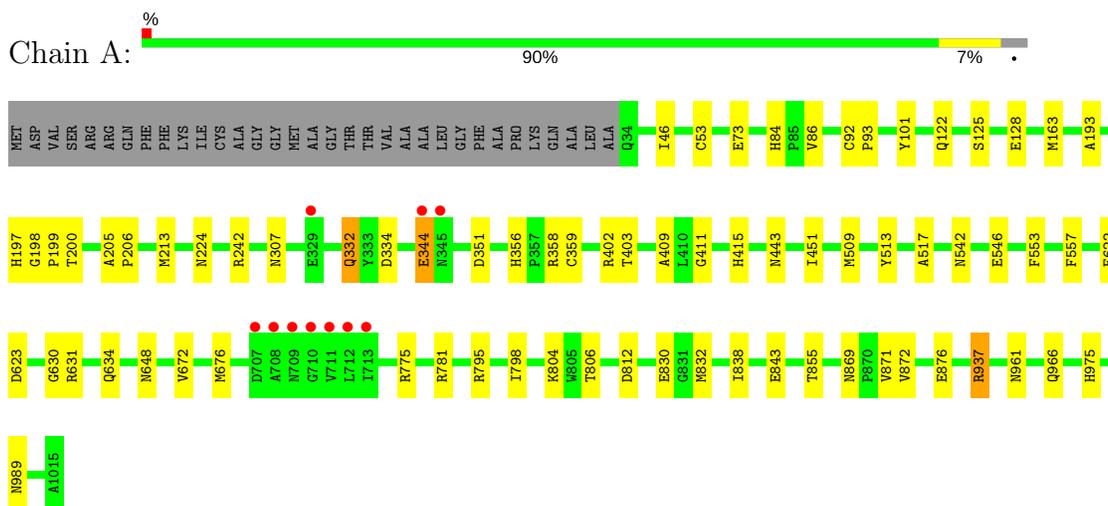
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1497	Total	O	0	0
			1497	1497		
9	B	399	Total	O	0	0
			399	399		
9	C	93	Total	O	0	0
			93	93		

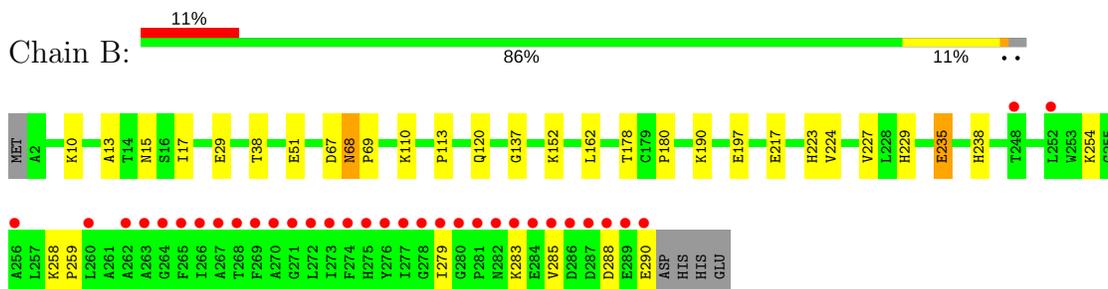
3 Residue-property plots [i](#)

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.

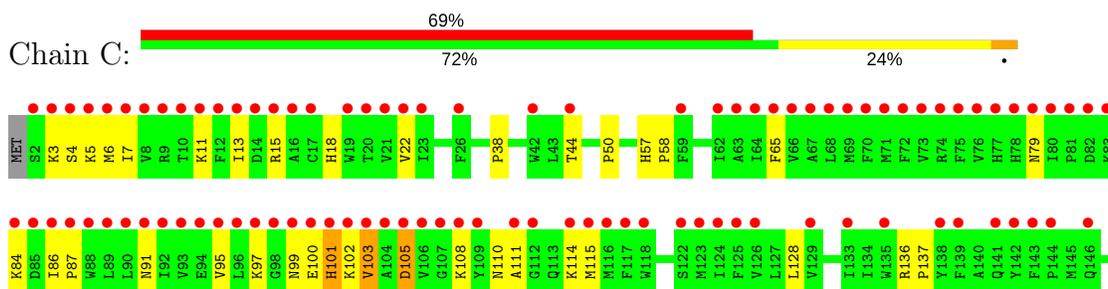
- Molecule 1: FORMATE DEHYDROGENASE, NITRATE-INDUCIBLE, MAJOR SUBUNIT

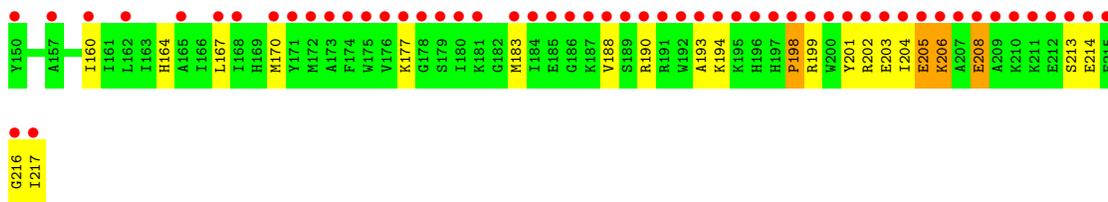


- Molecule 2: FORMATE DEHYDROGENASE, NITRATE-INDUCIBLE, IRON-SULFUR SUBUNIT



- Molecule 3: FORMATE DEHYDROGENASE, NITRATE-INDUCIBLE, CYTOCHROME B556(FDN) SUBUNIT





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	203.00Å 203.00Å 203.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.60 39.81 – 1.60	Depositor EDS
% Data completeness (in resolution range)	94.9 (40.00-1.60) 94.1 (39.81-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.22 (at 1.60Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.177 , 0.195 0.187 , 0.203	Depositor DCC
R_{free} test set	3446 reflections (1.02%)	DCC
Wilson B-factor (Å ²)	17.3	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 60.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.013 for l,-k,h	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	13989	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.73% 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, 0.333 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: 6MO, MGD, CDL, SF4, SEC, 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	0.34	0/7910	0.69	2/10749 (0.0%)
2	B	0.32	0/2255	0.68	2/3056 (0.1%)
3	C	0.33	0/1840	0.62	0/2483
All	All	0.33	0/12005	0.68	4/16288 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	334	ASP	N-CA-C	-6.00	94.79	111.00
2	B	227	VAL	N-CA-C	-5.85	95.21	111.00
1	A	975	HIS	N-CA-C	5.28	125.26	111.00
2	B	224	VAL	N-CA-C	-5.07	97.31	111.00

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	7719	0	7457	60	0
2	B	2207	0	2140	21	0
3	C	1783	0	1836	57	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	8	0	0	0	0
5	B	32	0	0	0	0
6	A	94	0	44	1	0
7	C	86	0	60	2	0
8	C	70	0	83	0	0
9	A	1497	0	0	16	2
9	B	399	0	0	6	0
9	C	93	0	0	10	0
All	All	13989	0	11620	132	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 6.

The worst 5 of 132 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:359:CYS:HB3	9:A:1672:HOH:O	1.66	0.95
1:A:356:HIS:HD2	1:A:358:ARG:H	1.15	0.93
1:A:869:ASN:HB3	1:A:872:VAL:HG23	1.52	0.91
1:A:224:ASN:HD22	1:A:403:THR:H	1.16	0.88
1:A:622:GLU:OE1	1:A:648:ASN:HB2	1.73	0.86

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)
9:A:2324:HOH:O	9:A:2324:HOH:O[6_456]	1.12	1.08
9:A:1746:HOH:O	9:A:1746:HOH:O[6_456]	1.25	0.95

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	979/1015 (96%)	951 (97%)	26 (3%)	2 (0%)	51 27
2	B	287/294 (98%)	278 (97%)	9 (3%)	0	100 100
3	C	214/217 (99%)	206 (96%)	6 (3%)	2 (1%)	20 4
All	All	1480/1526 (97%)	1435 (97%)	41 (3%)	4 (0%)	44 22

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	208	GLU
1	A	409	ALA
1	A	838	ILE
3	C	198	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	815/837 (97%)	807 (99%)	8 (1%)	80 65
2	B	238/243 (98%)	229 (96%)	9 (4%)	38 12
3	C	188/189 (100%)	182 (97%)	6 (3%)	44 17
All	All	1241/1269 (98%)	1218 (98%)	23 (2%)	62 37

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	68	ASN
2	B	162	LEU
3	C	205	GLU
2	B	113	PRO
2	B	178	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	479	GLN
1	A	865	ASN
3	C	79	ASN
1	A	415	HIS
3	C	101	HIS

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 [i](#)

Of 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 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$
5	SF4	A	1017	1	0,12,12	0.00	-	0,24,24	0.00	-
6	MGD	A	1018	4	41,52,52	2.27	10 (24%)	37,81,81	3.33	20 (54%)
6	MGD	A	1019	4	41,52,52	2.68	11 (26%)	37,81,81	2.91	16 (43%)
5	SF4	B	805	2	0,12,12	0.00	-	0,24,24	0.00	-
5	SF4	B	806	2	0,12,12	0.00	-	0,24,24	0.00	-
5	SF4	B	807	2	0,12,12	0.00	-	0,24,24	0.00	-
5	SF4	B	808	2	0,12,12	0.00	-	0,24,24	0.00	-
7	HEM	C	809	3	28,50,50	2.20	9 (32%)	17,82,82	2.04	8 (47%)
7	HEM	C	810	3	28,50,50	2.12	9 (32%)	17,82,82	2.00	5 (29%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
8	CDL	C	812	-	69,69,99	2.92	21 (30%)	71,81,111	2.81	16 (22%)

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
5	SF4	A	1017	1	-	0/0/48/48	0/6/5/5
6	MGD	A	1018	4	-	0/18/66/66	0/6/6/6
6	MGD	A	1019	4	-	0/18/66/66	0/6/6/6
5	SF4	B	805	2	-	0/0/48/48	0/6/5/5
5	SF4	B	806	2	-	0/0/48/48	0/6/5/5
5	SF4	B	807	2	-	0/0/48/48	0/6/5/5
5	SF4	B	808	2	-	0/0/48/48	0/6/5/5
7	HEM	C	809	3	-	0/6/54/54	0/0/8/8
7	HEM	C	810	3	-	0/6/54/54	0/0/8/8
8	CDL	C	812	-	-	0/80/80/110	0/0/0/0

The worst 5 of 60 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	C	812	CDL	OB8-CB6	-14.88	1.12	1.45
6	A	1018	MGD	O3A-C10	-9.64	1.06	1.44
6	A	1019	MGD	O3A-C10	-9.44	1.07	1.44
6	A	1019	MGD	C23-C14	-5.72	1.49	1.53
6	A	1019	MGD	O11-C23	-4.74	1.36	1.43

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	C	812	CDL	CB6-OB8-CB7	-9.07	89.86	117.13
8	C	812	CDL	CA6-CA4-CA3	-7.50	94.93	111.86
6	A	1019	MGD	C21-N22-C23	-6.38	111.17	123.67
8	C	812	CDL	C52-C51-CB5	-6.36	90.36	113.58
6	A	1018	MGD	C4'-O4'-C1'	-6.30	103.07	109.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1018	MGD	1	0
7	C	810	HEM	2	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	981/1015 (96%)	-0.20	10 (1%) 82 83	11, 18, 30, 56	0
2	B	289/294 (98%)	0.46	33 (11%) 6 5	11, 18, 52, 67	0
3	C	216/217 (99%)	3.97	149 (68%) 0 0	30, 49, 63, 65	0
All	All	1486/1526 (97%)	0.53	192 (12%) 4 4	11, 19, 54, 67	0

The worst 5 of 192 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	217	ILE	18.4
3	C	207	ALA	14.5
2	B	285	VAL	13.5
3	C	209	ALA	12.2
3	C	216	GLY	12.1

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
8	CDL	C	812	70/100	0.76	0.26	1.77	44,56,61,61	0
6	MGD	A	1019	47/47	0.96	0.12	0.45	11,14,16,18	0
7	HEM	C	809	43/43	0.92	0.15	0.42	29,33,35,37	0
7	HEM	C	810	43/43	0.87	0.25	0.25	44,45,48,49	0
6	MGD	A	1018	47/47	0.96	0.11	0.11	11,13,14,15	0
5	SF4	B	806	8/8	0.99	0.09	-0.72	12,13,13,13	0
5	SF4	B	807	8/8	0.98	0.08	-0.82	14,15,16,16	0
5	SF4	A	1017	8/8	0.99	0.10	-0.83	11,12,12,12	0
5	SF4	B	805	8/8	0.99	0.09	-1.07	12,13,13,14	0
5	SF4	B	808	8/8	0.98	0.07	-1.40	16,18,19,19	0
4	6MO	A	1016	1/1	0.99	0.09	-5.81	16,16,16,16	0

6.5 Other polymers [i](#)

There are no such residues in this entry.