



wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 05:38 PM GMT

PDB ID : 1M5H
Title : Formylmethanofuran:tetrahydromethanopterinformyltransferase from
Archaeoglobus fulgidus
Authors : Mamat, B.; Roth, A.; Grimm, C.; Ermler, U.; Tziatzios, C.; Schubert, D.;
Thauer, R.K.; Shima, S.
Deposited on : 2002-07-09
Resolution : 2.00 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

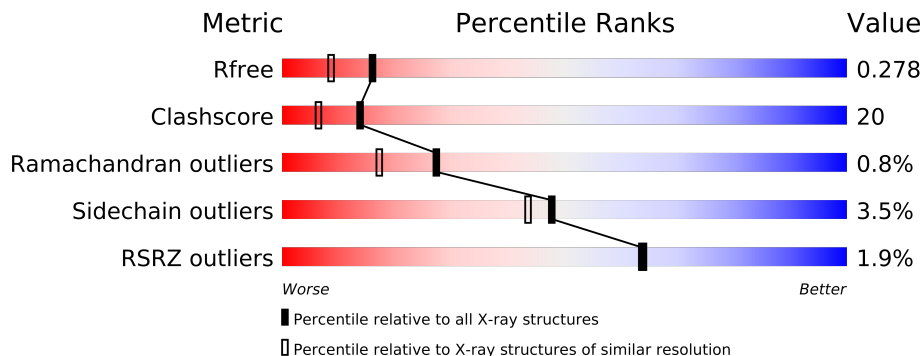
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.00 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}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	297	
1	B	297	
1	C	297	
1	D	297	
1	E	297	
1	F	297	
1	G	297	
1	H	297	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 19658 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Formylmethanofuran--tetrahydromethanopterinformyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	B	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	C	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	D	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	E	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	F	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	G	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			
1	H	297	Total	C	N	O	S	0	0	0
			2230	1429	365	426	10			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	115	ASP	PHE	CONFLICT	UNP O28076
A	129	GLN	GLU	CONFLICT	UNP O28076
A	239	ALA	GLU	CONFLICT	UNP O28076
B	1115	ASP	PHE	CONFLICT	UNP O28076
B	1129	GLN	GLU	CONFLICT	UNP O28076
B	1239	ALA	GLU	CONFLICT	UNP O28076
C	2115	ASP	PHE	CONFLICT	UNP O28076
C	2129	GLN	GLU	CONFLICT	UNP O28076
C	2239	ALA	GLU	CONFLICT	UNP O28076
D	3115	ASP	PHE	CONFLICT	UNP O28076
D	3129	GLN	GLU	CONFLICT	UNP O28076
D	3239	ALA	GLU	CONFLICT	UNP O28076

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	4115	ASP	PHE	CONFLICT	UNP O28076
E	4129	GLN	GLU	CONFLICT	UNP O28076
E	4239	ALA	GLU	CONFLICT	UNP O28076
F	5115	ASP	PHE	CONFLICT	UNP O28076
F	5129	GLN	GLU	CONFLICT	UNP O28076
F	5239	ALA	GLU	CONFLICT	UNP O28076
G	6115	ASP	PHE	CONFLICT	UNP O28076
G	6129	GLN	GLU	CONFLICT	UNP O28076
G	6239	ALA	GLU	CONFLICT	UNP O28076
H	7115	ASP	PHE	CONFLICT	UNP O28076
H	7129	GLN	GLU	CONFLICT	UNP O28076
H	7239	ALA	GLU	CONFLICT	UNP O28076

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	3	Total K 3 3	0	0
2	D	3	Total K 3 3	0	0
2	E	3	Total K 3 3	0	0
2	H	3	Total K 3 3	0	0
2	B	3	Total K 3 3	0	0
2	C	3	Total K 3 3	0	0
2	A	3	Total K 3 3	0	0
2	F	3	Total K 3 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	231	Total O 231 231	0	0
3	B	225	Total O 225 225	0	0
3	C	226	Total O 226 226	0	0

Continued on next page...

Continued from previous page...

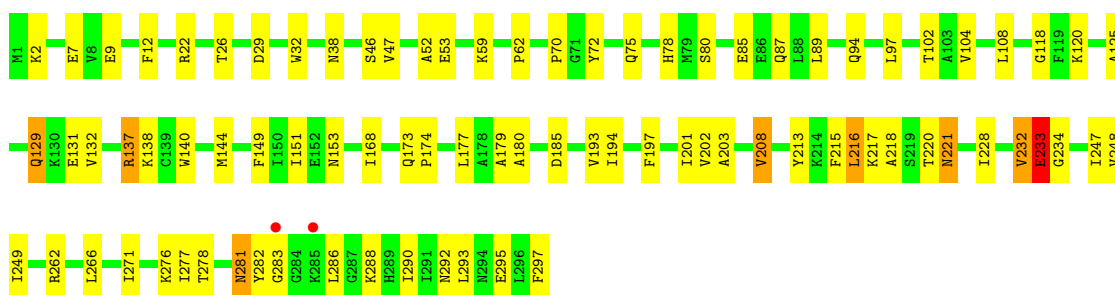
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	215	Total 215	O 215	0	0
3	E	239	Total 239	O 239	0	0
3	F	226	Total 226	O 226	0	0
3	G	206	Total 206	O 206	0	0
3	H	226	Total 226	O 226	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

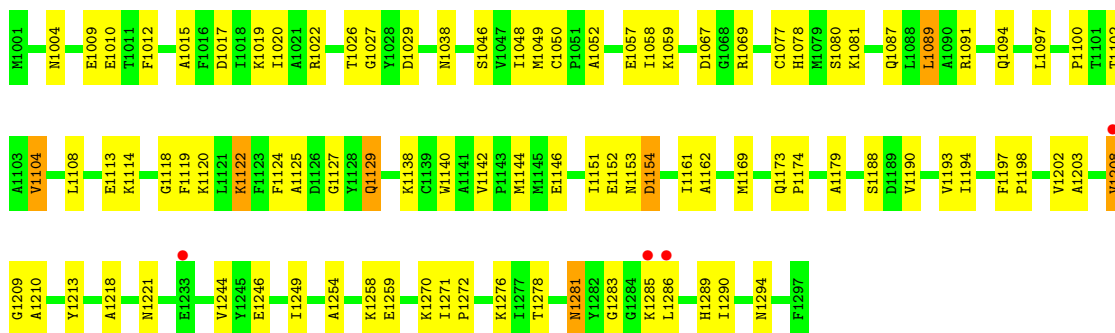
- Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

Chain A: 



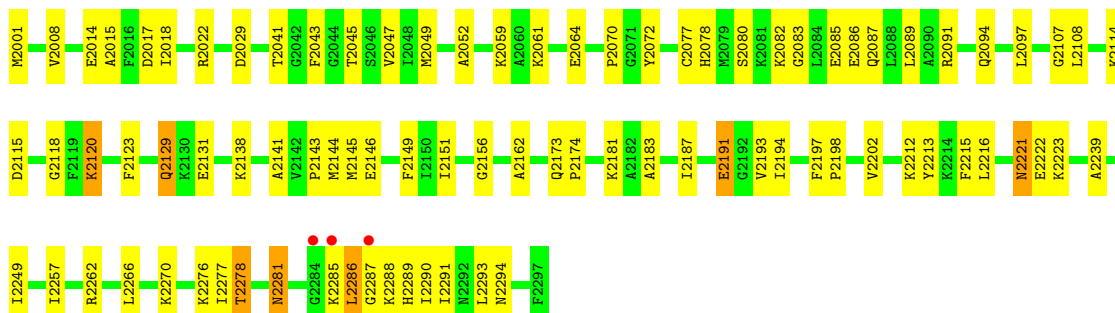
- Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

Chain B: 



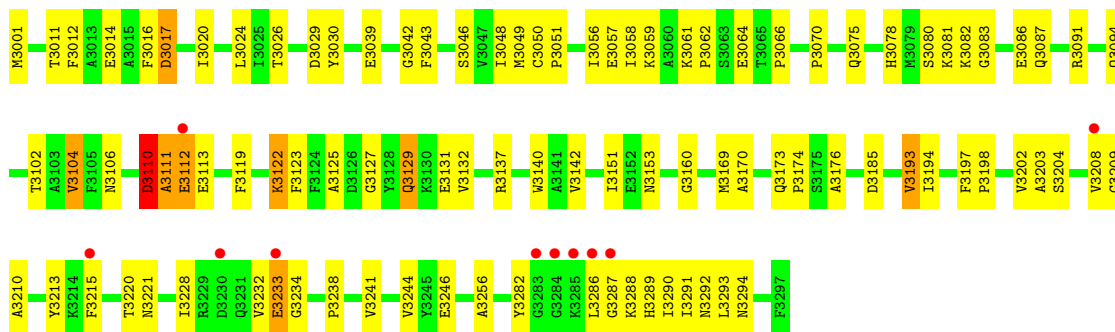
- Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

Chain C: 



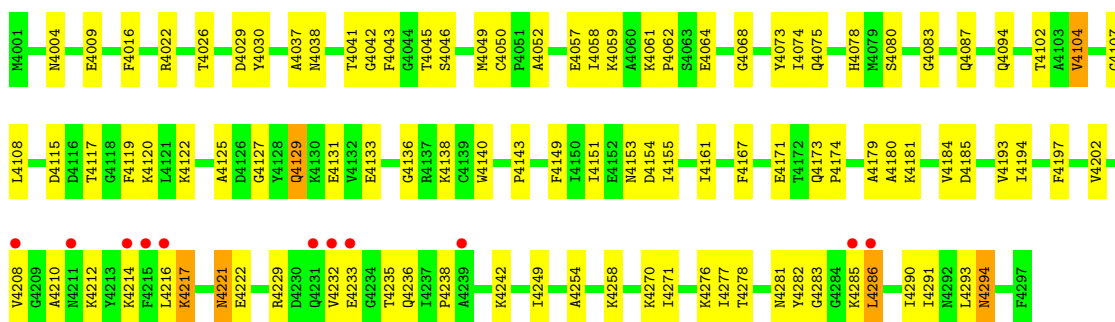
● Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

Chain D:



● Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

Chain E:



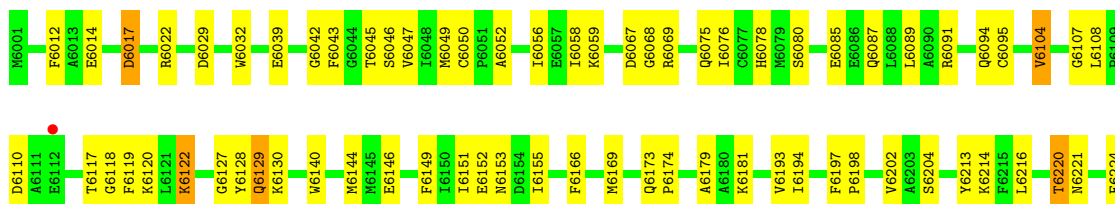
● Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

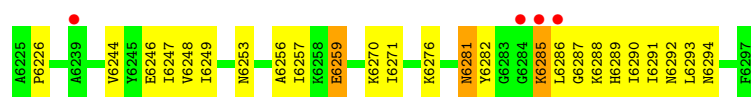
Chain F:



● Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

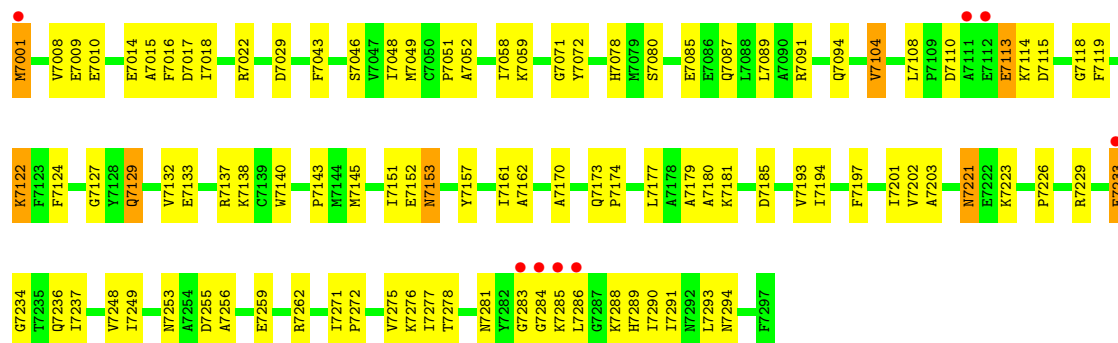
Chain G:





- Molecule 1: Formylmethanofuran--tetrahydromethanopterinformyltransferase

Chain H:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	77.10Å 81.81Å 99.10Å 90.10° 110.04° 93.75°	Depositor
Resolution (Å)	29.25 – 2.00 29.25 – 2.00	Depositor EDS
% Data completeness (in resolution range)	82.2 (29.25-2.00) 82.3 (29.25-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.58 (at 2.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.229 , 0.282 0.225 , 0.278	Depositor DCC
R_{free} test set	6405 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	13.4	Xtriage
Anisotropy	0.277	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 42.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 131922 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	19658	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.90% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section:
K

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	0.35	0/2274	0.63	0/3078
1	B	0.34	0/2274	0.63	0/3078
1	C	0.36	0/2274	0.63	0/3078
1	D	0.35	0/2274	0.63	0/3078
1	E	0.34	0/2274	0.63	0/3078
1	F	0.35	0/2274	0.62	0/3078
1	G	0.33	0/2274	0.63	0/3078
1	H	0.36	0/2274	0.62	0/3078
All	All	0.35	0/18192	0.63	0/24624

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2230	0	2234	98	0
1	B	2230	0	2231	94	0
1	C	2230	0	2231	104	0
1	D	2230	0	2231	87	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2230	0	2231	101	0
1	F	2230	0	2231	88	0
1	G	2230	0	2231	95	0
1	H	2230	0	2231	106	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
2	C	3	0	0	0	0
2	D	3	0	0	0	0
2	E	3	0	0	0	0
2	F	3	0	0	0	0
2	G	3	0	0	0	0
2	H	3	0	0	0	0
3	A	231	0	0	18	0
3	B	225	0	0	13	0
3	C	226	0	0	14	0
3	D	215	0	0	8	0
3	E	239	0	0	7	0
3	F	226	0	0	8	0
3	G	206	0	0	10	0
3	H	226	0	0	16	0
All	All	19658	0	17851	713	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 20.

The worst 5 of 713 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:3104:VAL:HG22	1:D:3151:ILE:HG22	1.31	1.12
1:E:4125:ALA:HB3	1:E:4129:GLN:HG3	1.31	1.09
1:F:5125:ALA:HB3	1:F:5129:GLN:HG3	1.43	0.98
1:D:3140:TRP:HE1	1:D:3153:ASN:HD22	1.08	0.98
1:E:4104:VAL:HG22	1:E:4151:ILE:HG22	1.46	0.97

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	295/297 (99%)	283 (96%)	8 (3%)	4 (1%)	16	7
1	B	295/297 (99%)	282 (96%)	12 (4%)	1 (0%)	50	44
1	C	295/297 (99%)	282 (96%)	12 (4%)	1 (0%)	50	44
1	D	295/297 (99%)	273 (92%)	18 (6%)	4 (1%)	16	7
1	E	295/297 (99%)	276 (94%)	16 (5%)	3 (1%)	22	12
1	F	295/297 (99%)	284 (96%)	9 (3%)	2 (1%)	30	20
1	G	295/297 (99%)	278 (94%)	15 (5%)	2 (1%)	30	20
1	H	295/297 (99%)	275 (93%)	19 (6%)	1 (0%)	50	44
All	All	2360/2376 (99%)	2233 (95%)	109 (5%)	18 (1%)	27	17

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1208	VAL
1	D	3208	VAL
1	E	4286	LEU
1	F	5208	VAL
1	A	208	VAL

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	227/227 (100%)	220 (97%)	7 (3%)	52	49
1	B	227/227 (100%)	220 (97%)	7 (3%)	52	49
1	C	227/227 (100%)	220 (97%)	7 (3%)	52	49
1	D	227/227 (100%)	217 (96%)	10 (4%)	39	32
1	E	227/227 (100%)	220 (97%)	7 (3%)	52	49
1	F	227/227 (100%)	224 (99%)	3 (1%)	80	82
1	G	227/227 (100%)	216 (95%)	11 (5%)	35	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	H	227/227 (100%)	216 (95%)	11 (5%)	35	28
All	All	1816/1816 (100%)	1753 (96%)	63 (4%)	48	43

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	3221	ASN
1	E	4221	ASN
1	H	7153	ASN
1	D	3233	GLU
1	E	4104	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	3173	GLN
1	E	4281	ASN
1	H	7129	GLN
1	D	3221	ASN
1	E	4094	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 24 ligands modelled in this entry, 24 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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(Å ²)	Q<0.9
1	A	297/297 (100%)	-0.09	2 (0%) 84 85	6, 13, 26, 42	0
1	B	297/297 (100%)	-0.02	4 (1%) 74 74	7, 14, 32, 46	0
1	C	297/297 (100%)	-0.05	3 (1%) 79 80	7, 14, 28, 41	0
1	D	297/297 (100%)	0.03	10 (3%) 43 43	7, 14, 40, 55	0
1	E	297/297 (100%)	0.10	11 (3%) 39 39	8, 15, 40, 50	0
1	F	297/297 (100%)	-0.01	2 (0%) 84 85	6, 15, 29, 40	0
1	G	297/297 (100%)	0.02	5 (1%) 67 67	8, 15, 32, 44	0
1	H	297/297 (100%)	0.01	8 (2%) 52 52	7, 14, 30, 49	0
All	All	2376/2376 (100%)	-0.00	45 (1%) 64 64	6, 14, 32, 55	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	7283	GLY	9.9
1	H	7285	LYS	5.0
1	D	3286	LEU	4.9
1	B	1208	VAL	4.7
1	H	7284	GLY	4.6

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 carbohydrates 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. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	K	C	8001	1/1	0.05	-2.05	11,11,11,11	0
2	K	D	8003	1/1	0.05	-2.21	10,10,10,10	0
2	K	F	8009	1/1	0.06	-2.37	11,11,11,11	0
2	K	A	8002	1/1	0.06	-2.48	10,10,10,10	0
2	K	H	8023	1/1	0.07	-2.48	17,17,17,17	0
2	K	B	8015	1/1	0.08	-2.63	13,13,13,13	0
2	K	G	8021	1/1	0.07	-2.90	15,15,15,15	0
2	K	D	8006	1/1	0.03	-2.93	9,9,9,9	0
2	K	H	8004	1/1	0.03	-2.95	11,11,11,11	0
2	K	E	8024	1/1	0.08	-2.99	16,16,16,16	0
2	K	B	8017	1/1	0.04	-3.18	11,11,11,11	0
2	K	F	8010	1/1	0.06	-3.19	13,13,13,13	0
2	K	C	8007	1/1	0.04	-3.21	11,11,11,11	0
2	K	C	8014	1/1	0.09	-3.33	15,15,15,15	0
2	K	A	8011	1/1	0.06	-3.37	11,11,11,11	0
2	K	B	8005	1/1	0.03	-3.40	11,11,11,11	0
2	K	G	8013	1/1	0.04	-3.42	12,12,12,12	0
2	K	F	8022	1/1	0.04	-3.48	14,14,14,14	0
2	K	E	8018	1/1	0.04	-3.67	12,12,12,12	0
2	K	E	8016	1/1	0.06	-3.76	17,17,17,17	0
2	K	H	8020	1/1	0.05	-3.85	14,14,14,14	0
2	K	G	8008	1/1	0.07	-4.04	10,10,10,10	0
2	K	A	8019	1/1	0.03	-4.14	14,14,14,14	0
2	K	D	8012	1/1	0.06	-4.91	12,12,12,12	0

6.5 Other polymers ⓘ

There are no such residues in this entry.