wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 08:59 PM EDT

PDB ID	:	4NAS
Title	:	The crystal structure of a rubisco-like protein (MtnW) from Alicyclobacillus acidocaldarius subsp. acidocaldarius DSM 446
Authors	:	Tan, K.; Li, H.; Clancy, S.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2013-10-22
Resolution	:	1.92 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 (1)) were used in the production of this report:

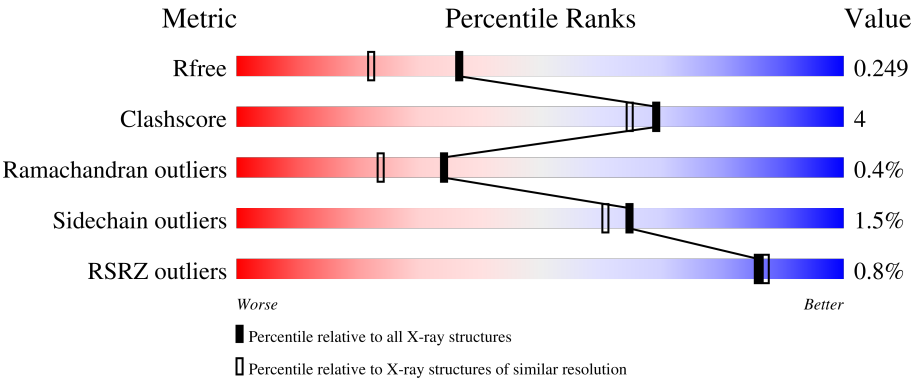
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.92 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}	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	409	87% 10% ·
1	B	409	90% 8% ·
1	C	409	% 86% 10% ·
1	D	409	% 83% 12% ·

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12421 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.

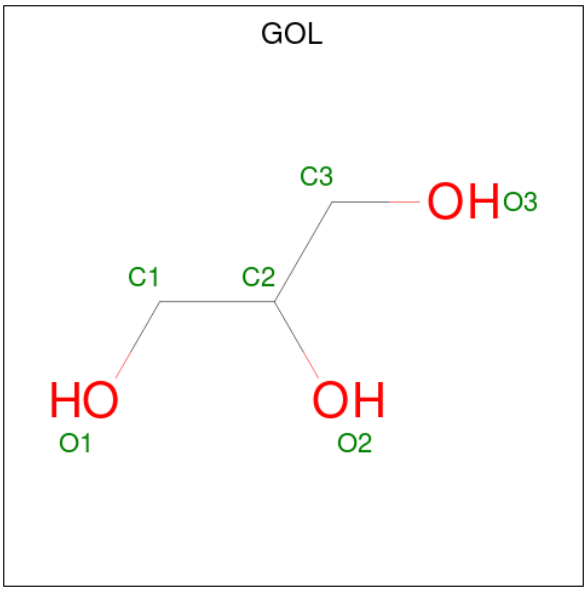
- Molecule 1 is a protein called Ribulose-bisphosphate carboxylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	397	Total	C	N	O	S	Se	0	3	0
			2988	1894	540	546	2	6			
1	B	399	Total	C	N	O	S	Se	0	2	0
			3012	1910	543	551	2	6			
1	C	396	Total	C	N	O	S	Se	0	0	0
			2909	1853	516	532	2	6			
1	D	392	Total	C	N	O	S	Se	0	0	0
			2853	1815	505	525	2	6			

There are 12 discrepancies between the modelled and reference sequences:

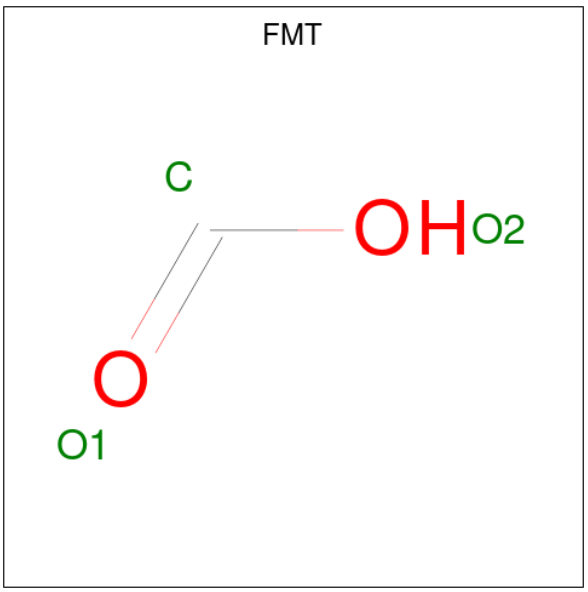
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP C8WQ56
A	1	ASN	-	EXPRESSION TAG	UNP C8WQ56
A	2	ALA	-	EXPRESSION TAG	UNP C8WQ56
B	0	SER	-	EXPRESSION TAG	UNP C8WQ56
B	1	ASN	-	EXPRESSION TAG	UNP C8WQ56
B	2	ALA	-	EXPRESSION TAG	UNP C8WQ56
C	0	SER	-	EXPRESSION TAG	UNP C8WQ56
C	1	ASN	-	EXPRESSION TAG	UNP C8WQ56
C	2	ALA	-	EXPRESSION TAG	UNP C8WQ56
D	0	SER	-	EXPRESSION TAG	UNP C8WQ56
D	1	ASN	-	EXPRESSION TAG	UNP C8WQ56
D	2	ALA	-	EXPRESSION TAG	UNP C8WQ56

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0

- Molecule 3 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 3 1 2	0	0
3	B	1	Total C O 3 1 2	0	0
3	B	1	Total C O 3 1 2	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Ca 2 2	0	0
4	B	2	Total Ca 2 2	0	0
4	C	2	Total Ca 2 2	0	0
4	D	1	Total Ca 1 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Cl 1 1	0	0

- Molecule 6 is water.

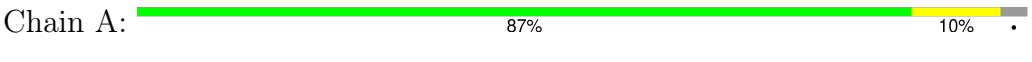
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	193	Total O 193 193	0	0
6	B	203	Total O 203 203	0	0
6	C	133	Total O 133 133	0	0
6	D	107	Total O 107 107	0	0

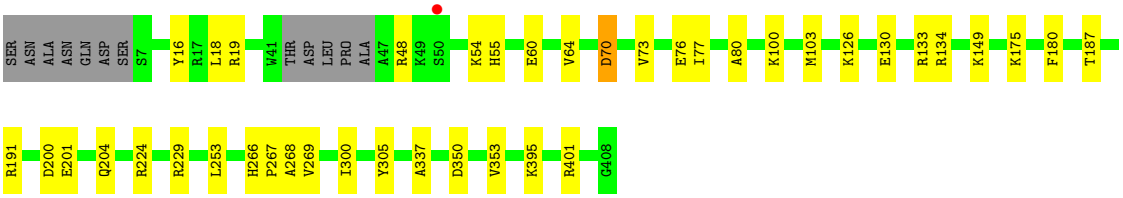
3 Residue-property plots (i)

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: Ribulose-bisphosphate carboxylase

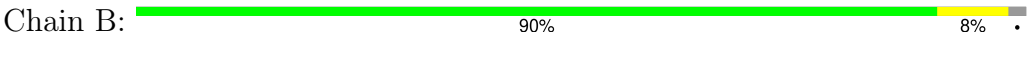
Chain A:

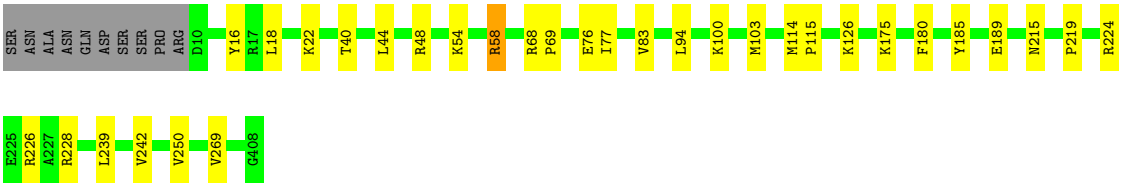




- Molecule 1: Ribulose-bisphosphate carboxylase

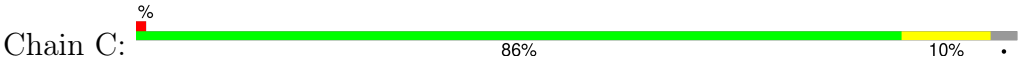
Chain B:

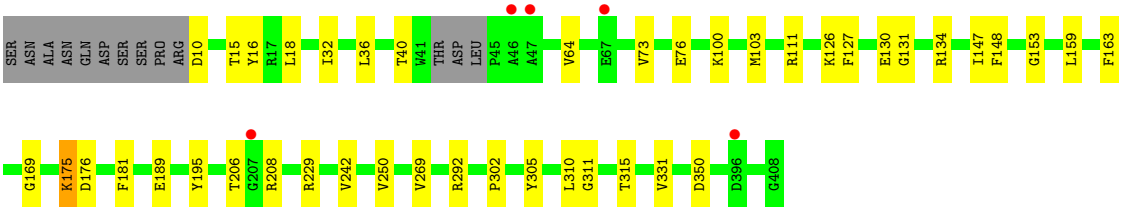




- Molecule 1: Ribulose-bisphosphate carboxylase

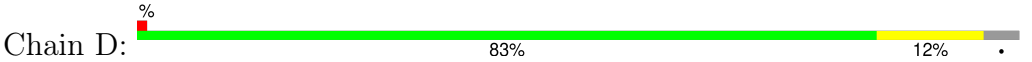
Chain C:

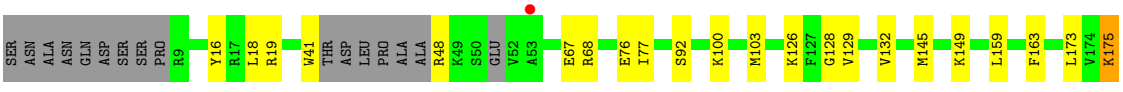


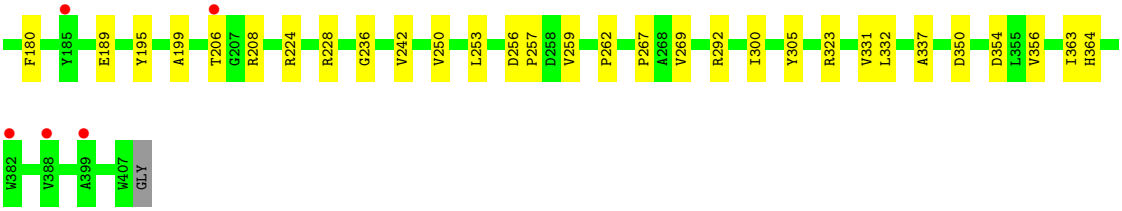


- Molecule 1: Ribulose-bisphosphate carboxylase

Chain D:






F180
Y185
E189
Y195
A199
T206
G207
R208
R224
R228
G236
V242
V250
L253
D256
P257
D258
V259
P262
P267
A268
V269
R292
I300
Y305
R323
V331
L332
A337
D350
D354
L355
V356
I363
H364
W382
V388
A399
W407
GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	56.76Å 175.67Å 83.65Å 90.00° 107.67° 90.00°	Depositor
Resolution (Å)	29.50 – 1.92 29.51 – 1.92	Depositor EDS
% Data completeness (in resolution range)	95.7 (29.50-1.92) 95.6 (29.51-1.92)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.48 (at 1.92Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.195 , 0.251 0.195 , 0.249	Depositor DCC
R_{free} test set	5685 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	28.7	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 37.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.057 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12421	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, KCX, CA, FMT, CL

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	0/3035	0.56	0/4116
1	B	0.42	0/3060	0.57	0/4149
1	C	0.34	0/2950	0.52	0/4007
1	D	0.32	0/2890	0.50	0/3927
All	All	0.38	0/11935	0.54	0/16199

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2988	0	2956	22	0
1	B	3012	0	3005	19	0
1	C	2909	0	2851	22	0
1	D	2853	0	2753	28	0
2	A	6	0	8	0	0
3	A	3	0	1	1	0
3	B	6	0	2	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	2	0	0	0	0
4	D	1	0	0	0	0
5	B	1	0	0	0	0
6	A	193	0	0	1	0
6	B	203	0	0	1	0
6	C	133	0	0	1	0
6	D	107	0	0	0	0
All	All	12421	0	11576	86	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 4.

The worst 5 of 86 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:B:18:LEU:HD21	1:B:77:ILE:HD12	1.62	0.80
1:A:18:LEU:HD21	1:A:77:ILE:HD12	1.65	0.78
1:A:266:HIS:HD2	1:A:268:ALA:H	1.39	0.70
1:C:292:ARG:HD3	1:C:331:VAL:HG23	1.74	0.69
1:C:206:THR:HG22	1:C:208:ARG:H	1.56	0.69

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	395/409 (97%)	384 (97%)	9 (2%)	2 (0%)	29 18
1	B	398/409 (97%)	386 (97%)	11 (3%)	1 (0%)	41 31
1	C	391/409 (96%)	376 (96%)	14 (4%)	1 (0%)	41 31
1	D	385/409 (94%)	364 (94%)	19 (5%)	2 (0%)	29 18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1569/1636 (96%)	1510 (96%)	53 (3%)	6 (0%)	34 24

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	269	VAL
1	B	269	VAL
1	C	269	VAL
1	D	269	VAL
1	A	337	ALA

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	289/304 (95%)	284 (98%)	5 (2%)	60 55
1	B	294/304 (97%)	290 (99%)	4 (1%)	67 63
1	C	274/304 (90%)	271 (99%)	3 (1%)	73 72
1	D	264/304 (87%)	259 (98%)	5 (2%)	57 51
All	All	1121/1216 (92%)	1104 (98%)	17 (2%)	65 61

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	180	PHE
1	D	363	ILE
1	B	126	LYS
1	B	180	PHE
1	C	10	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	266	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)

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	KCX	B	175	4,1	10,11,12	2.01	2 (20%)	6,12,14	2.93	2 (33%)
1	KCX	A	175	4,1	10,11,12	2.37	2 (20%)	6,12,14	1.86	1 (16%)
1	KCX	D	175	4,1	10,11,12	2.33	3 (30%)	6,12,14	2.35	1 (16%)
1	KCX	C	175	4,1	10,11,12	2.15	3 (30%)	6,12,14	2.94	2 (33%)

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
1	KCX	B	175	4,1	-	2/9/10/12	-
1	KCX	A	175	4,1	-	1/9/10/12	-
1	KCX	D	175	4,1	-	0/9/10/12	-
1	KCX	C	175	4,1	-	3/9/10/12	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	175	KCX	CX-NZ	5.83	1.45	1.35
1	D	175	KCX	CX-NZ	5.66	1.45	1.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	175	KCX	CX-NZ	5.15	1.44	1.35
1	B	175	KCX	CX-NZ	5.08	1.44	1.35
1	A	175	KCX	OQ1-CX	3.53	1.28	1.21

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	175	KCX	CE-NZ-CX	-6.16	111.52	121.98
1	B	175	KCX	CE-NZ-CX	-5.91	111.94	121.98
1	D	175	KCX	CE-NZ-CX	-5.42	112.79	121.98
1	A	175	KCX	CE-NZ-CX	-4.38	114.55	121.98
1	B	175	KCX	OQ1-CX-NZ	-3.89	119.01	124.92

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	175	KCX	O-C-CA-CB
1	C	175	KCX	OQ1-CX-NZ-CE
1	C	175	KCX	OQ2-CX-NZ-CE
1	B	175	KCX	CE-CD-CG-CB
1	A	175	KCX	CE-CD-CG-CB

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	175	KCX	1	0
1	C	175	KCX	2	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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	GOL	A	501	-	5,5,5	0.45	0	5,5,5	0.19	0
3	FMT	A	502	-	2,2,2	0.58	0	1,1,1	0.10	0
3	FMT	B	502	-	2,2,2	0.69	0	1,1,1	0.24	0
3	FMT	B	501	-	2,2,2	0.72	0	1,1,1	0.20	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	GOL	A	501	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	GOL	O1-C1-C2-C3
2	A	501	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	FMT	1	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

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, 95^{th} 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	390/409 (95%)	-0.60	1 (0%) 94 94	15, 29, 52, 76	0
1	B	392/409 (95%)	-0.61	0 100 100	14, 30, 55, 71	0
1	C	389/409 (95%)	-0.26	5 (1%) 77 79	24, 44, 69, 92	0
1	D	385/409 (94%)	-0.13	6 (1%) 72 74	24, 51, 84, 96	0
All	All	1556/1636 (95%)	-0.40	12 (0%) 86 87	14, 38, 71, 96	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	46	ALA	4.2
1	D	399	ALA	3.9
1	D	382	TRP	3.4
1	C	396	ASP	3.1
1	D	388	VAL	2.8

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

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, 95^{th} 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
1	KCX	D	175	12/13	0.94	0.11	44,48,55,57	0
1	KCX	C	175	12/13	0.95	0.10	29,36,49,53	0
1	KCX	A	175	12/13	0.96	0.08	16,22,32,33	0
1	KCX	B	175	12/13	0.97	0.09	18,21,35,40	0

6.3 Carbohydrates (i)

There are no monosaccharides 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, 95^{th} 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
2	GOL	A	501	6/6	0.88	0.11	42,45,49,50	0
5	CL	B	505	1/1	0.88	0.07	57,57,57,57	0
3	FMT	B	501	3/3	0.89	0.17	42,42,50,53	0
3	FMT	A	502	3/3	0.93	0.10	23,23,24,29	0
3	FMT	B	502	3/3	0.96	0.05	34,34,36,37	0
4	CA	C	502	1/1	0.97	0.04	44,44,44,44	0
4	CA	D	501	1/1	0.97	0.04	51,51,51,51	0
4	CA	B	504	1/1	0.97	0.04	40,40,40,40	0
4	CA	C	501	1/1	0.98	0.03	44,44,44,44	0
4	CA	A	504	1/1	0.98	0.08	49,49,49,49	0
4	CA	B	503	1/1	0.99	0.04	34,34,34,34	0
4	CA	A	503	1/1	0.99	0.03	32,32,32,32	0

6.5 Other polymers (i)

There are no such residues in this entry.