WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2023 – 07:35 pm GMT

PDB ID	:	4CCN
Title	:	60S ribosomal protein L8 histidine hydroxylase (NO66 L299C/C300S) in complex with Mn(II), N-oxalylglycine (NOG) and 60S ribosomal protein L8 (RPL8 G220C) peptide fragment (complex-2)
Authors	:	Chowdhury, R.; Schofield, C.J.
Deposited on	:	2013-10-23
Resolution	:	2.23 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 (i)) were used in the production of this report:

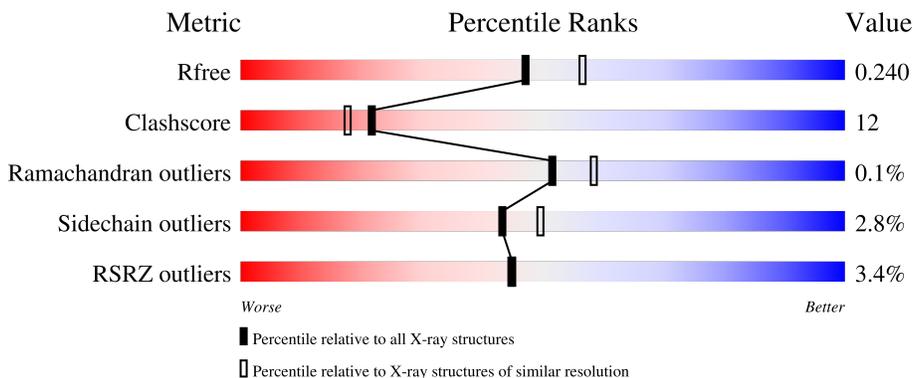|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.23 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	Percentile Ranks	Value
Rfree		0.240
Clashscore		12
Ramachandran outliers		0.1%
Sidechain outliers		2.8%
RSRZ outliers		3.4%

Worse ▬▬▬▬▬▬▬▬▬▬▬▬ *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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	467	3% / 75% green, 23% yellow, •
1	B	467	3% / 72% green, 24% yellow, • •
2	C	35	3% / 20%, 6%, • 71%
2	D	35	6% / 9%, 20%, 71%

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 7936 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 BIFUNCTIONAL LYSINE-SPECIFIC DEMETHYLASE AND HISTIDYL-HYDROXYLASE NO66.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	458	Total	C	N	O	S	0	8	0
			3681	2343	657	666	15			
1	B	459	Total	C	N	O	S	0	10	0
			3711	2362	664	670	15			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	MET	-	expression tag	UNP Q9H6W3
A	642	ALA	-	expression tag	UNP Q9H6W3
A	643	GLU	-	expression tag	UNP Q9H6W3
A	644	ASN	-	expression tag	UNP Q9H6W3
A	645	LEU	-	expression tag	UNP Q9H6W3
A	646	TYR	-	expression tag	UNP Q9H6W3
A	647	PHE	-	expression tag	UNP Q9H6W3
A	648	GLN	-	expression tag	UNP Q9H6W3
A	299	CYS	LEU	engineered mutation	UNP Q9H6W3
A	300	SER	CYS	engineered mutation	UNP Q9H6W3
A	364	ALA	VAL	engineered mutation	UNP Q9H6W3
B	182	MET	-	expression tag	UNP Q9H6W3
B	642	ALA	-	expression tag	UNP Q9H6W3
B	643	GLU	-	expression tag	UNP Q9H6W3
B	644	ASN	-	expression tag	UNP Q9H6W3
B	645	LEU	-	expression tag	UNP Q9H6W3
B	646	TYR	-	expression tag	UNP Q9H6W3
B	647	PHE	-	expression tag	UNP Q9H6W3
B	648	GLN	-	expression tag	UNP Q9H6W3
B	299	CYS	LEU	engineered mutation	UNP Q9H6W3
B	300	SER	CYS	engineered mutation	UNP Q9H6W3
B	364	ALA	VAL	engineered mutation	UNP Q9H6W3

- Molecule 2 is a protein called 60S RIBOSOMAL PROTEIN L8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	10	Total 71	C 42	N 16	O 12	S 1	0	0	0
2	D	10	Total 74	C 44	N 16	O 12	S 2	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	220	CYS	GLY	engineered mutation	UNP P62917
D	220	CYS	GLY	engineered mutation	UNP P62917

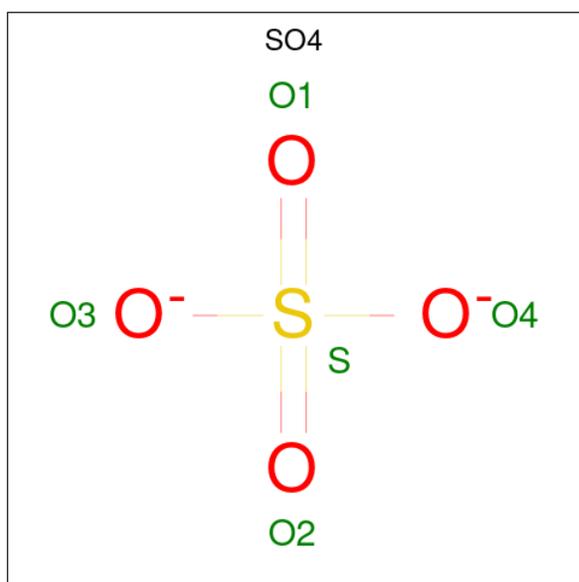- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Mn 1	0	0
3	B	1	Total 1	Mn 1	0	0

- Molecule 4 is N-OXALYLGLYCINE (three-letter code: OGA) (formula: $C_4H_5NO_5$).
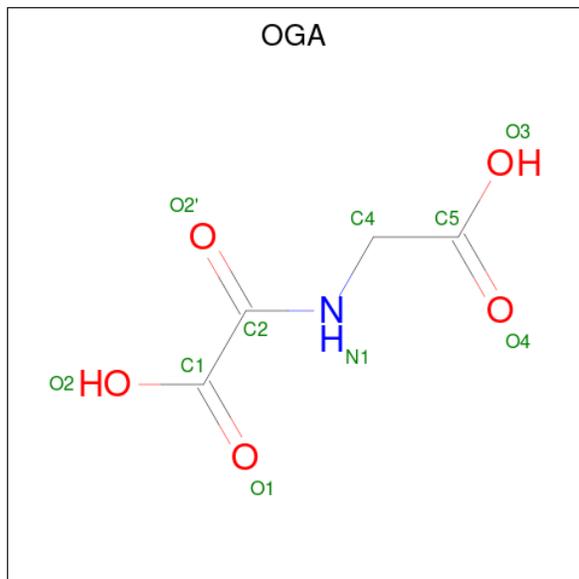
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 10	C 4	N 1	O 5	0	0
4	B	1	Total 10	C 4	N 1	O 5	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O$_4$S).

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

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C$_2$H$_6$O$_2$).
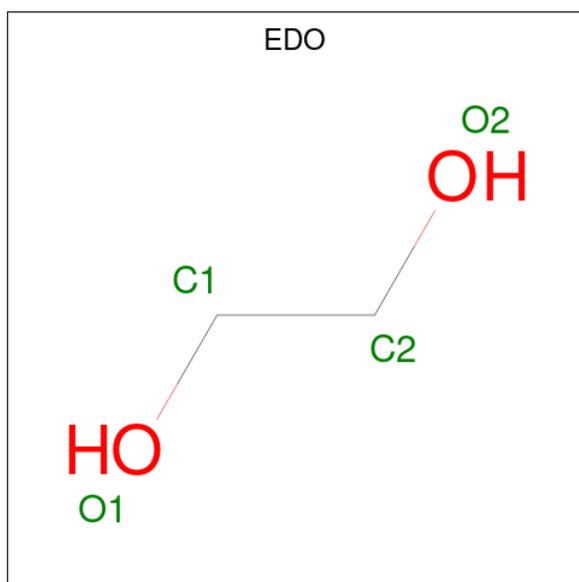
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total C O 4 2 2			0	0

- Molecule 7 is water.

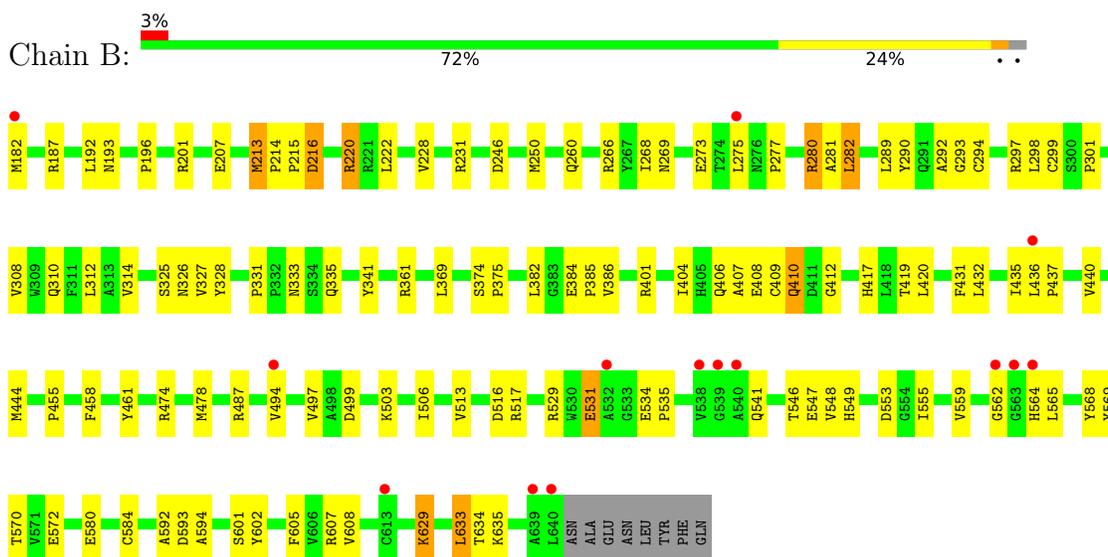| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	192	Total 192	O 192	0	0
7	B	169	Total 169	O 169	0	0
7	C	1	Total 1	O 1	0	0
7	D	1	Total 1	O 1	0	0

3 Residue-property plots ⓘ

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: BIFUNCTIONAL LYSINE-SPECIFIC DEMETHYLASE AND HISTIDYL-HYDROXYLASE NO66
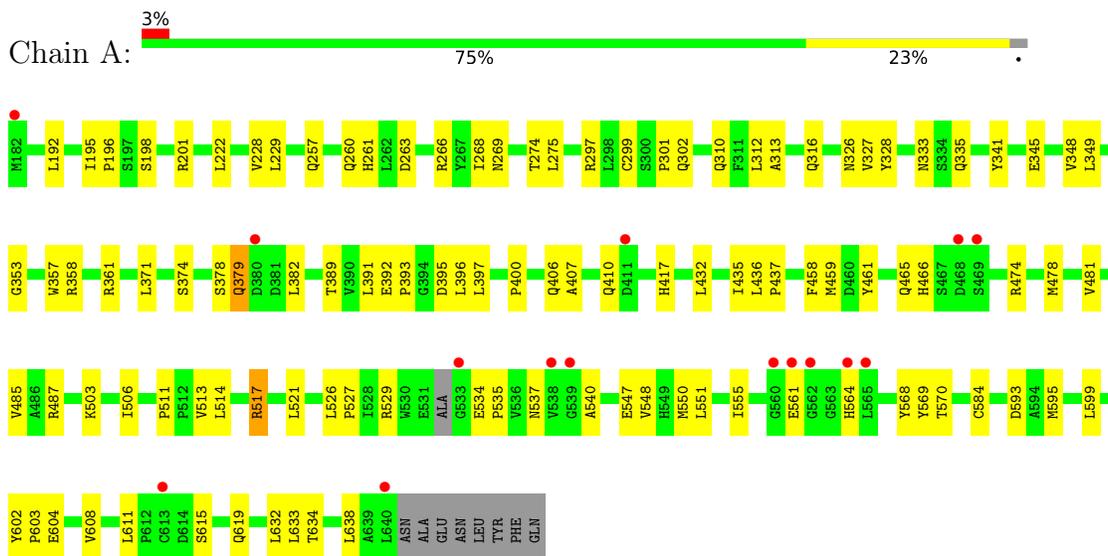
• Molecule 1: BIFUNCTIONAL LYSINE-SPECIFIC DEMETHYLASE AND HISTIDYL-HYDROXYLASE NO66

• Molecule 2: 60S RIBOSOMAL PROTEIN L8

Chain C:

3%

20% 6% · 71%

ASN PRO VAL GLU HIS PRO PHE GLY **G213 G214 N215 H216 Q217 P222** SER THR ILE ARG ARG ASP ALA PRO ALA GLY ARG LYS VAL GLY LEU ILE ALA

• Molecule 2: 60S RIBOSOMAL PROTEIN L8

Chain D:

6%

9% 20% 71%

ASN PRO VAL GLU HIS PRO PHE GLY **G213 G214 N215 H216 Q217 H218 I219 C220 K221 P222** SER THR ILE ARG ARG ASP ALA PRO ALA GLY ARG LYS VAL GLY LEU ILE ALA

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	155.16Å 83.72Å 97.03Å 90.00° 100.35° 90.00°	Depositor
Resolution (Å)	44.19 – 2.23 44.19 – 2.23	Depositor EDS
% Data completeness (in resolution range)	98.2 (44.19-2.23) 98.2 (44.19-2.23)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 2.22Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.238 , 0.248 0.232 , 0.240	Depositor DCC
R_{free} test set	2956 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	35.7	Xtriage
Anisotropy	0.360	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 42.8	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7936	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.65% 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: OGA, EDO, MN, 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	0.41	0/3799	0.63	0/5168
1	B	0.40	0/3836	0.62	0/5219
2	C	0.25	0/73	1.09	1/98 (1.0%)
2	D	0.57	0/79	1.03	1/106 (0.9%)
All	All	0.41	0/7787	0.64	2/10591 (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	A	0	6
1	B	0	2
All	All	0	8

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	215	ASN	CB-CA-C	8.52	127.43	110.40
2	C	215	ASN	CB-CA-C	8.48	127.35	110.40

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	260[A]	GLN	Mainchain
1	A	260[B]	GLN	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	379[A]	GLN	Mainchain
1	A	379[B]	GLN	Mainchain
1	A	487[A]	ARG	Mainchain

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	3681	0	3597	71	0
1	B	3711	0	3644	103	0
2	C	71	0	57	2	0
2	D	74	0	62	7	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	10	0	3	0	0
4	B	10	0	3	1	0
5	A	5	0	0	1	0
5	B	5	0	0	0	0
6	B	4	0	6	2	0
7	A	192	0	0	6	0
7	B	169	0	0	6	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
All	All	7936	0	7372	178	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 12.

The worst 5 of 178 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:333:ASN:HD21	1:B:412:GLY:H	1.21	0.86
1:B:266[B]:ARG:NH2	1:B:294:CYS:SG	2.51	0.84
1:B:280[B]:ARG:HG2	1:B:282:LEU:HD13	1.60	0.83
1:A:268:ILE:HG22	1:A:269[B]:ASN:ND2	2.00	0.76
1:B:266[A]:ARG:HD2	1:B:268:ILE:HD11	1.64	0.76

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	462/467 (99%)	452 (98%)	10 (2%)	0	100	100
1	B	467/467 (100%)	451 (97%)	15 (3%)	1 (0%)	47	53
2	C	8/35 (23%)	6 (75%)	2 (25%)	0	100	100
2	D	9/35 (26%)	8 (89%)	1 (11%)	0	100	100
All	All	946/1004 (94%)	917 (97%)	28 (3%)	1 (0%)	51	58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	196	PRO

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	387/399 (97%)	379 (98%)	8 (2%)	53	60
1	B	392/399 (98%)	377 (96%)	15 (4%)	33	36
2	C	7/27 (26%)	6 (86%)	1 (14%)	3	1
2	D	8/27 (30%)	8 (100%)	0	100	100
All	All	794/852 (93%)	770 (97%)	24 (3%)	43	47

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	282	LEU
1	B	432	LEU
1	B	410	GLN
1	B	517	ARG
1	A	593	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	379	GLN
1	B	441	GLN
1	B	417	HIS
1	B	466	HIS
1	A	465	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 monosaccharides in this entry.

5.6   Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	SO4	A	903	-	4,4,4	0.31	0	6,6,6	0.12	0
5	SO4	B	903	-	4,4,4	0.27	0	6,6,6	0.06	0
6	EDO	B	904	-	3,3,3	0.47	0	2,2,2	0.33	0
4	OGA	B	902	3	9,9,9	1.05	0	10,11,11	1.29	0
4	OGA	A	902	3	9,9,9	1.05	0	10,11,11	1.30	1 (10%)

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
6	EDO	B	904	-	-	1/1/1/1	-
4	OGA	B	902	3	-	0/8/9/9	-
4	OGA	A	902	3	-	0/8/9/9	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	902	OGA	O2'-C2-C1	-2.04	118.44	121.32

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	904	EDO	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	903	SO4	1	0
6	B	904	EDO	2	0
4	B	902	OGA	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 ⓘ

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(Å²)	Q<0.9
1	A	458/467 (98%)	0.25	15 (3%)	46	46	17, 35, 62, 73	6 (1%)
1	B	459/467 (98%)	0.31	14 (3%)	49	49	16, 35, 66, 77	2 (0%)
2	C	10/35 (28%)	0.85	1 (10%)	7	6	32, 38, 67, 72	2 (20%)
2	D	10/35 (28%)	1.35	2 (20%)	1	1	33, 39, 63, 67	2 (20%)
All	All	937/1004 (93%)	0.30	32 (3%)	45	45	16, 35, 65, 77	12 (1%)

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	562	GLY	4.6
2	D	213	GLY	4.0
1	B	639	ALA	4.0
2	D	222	PRO	3.9
1	B	532	ALA	3.7

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 monosaccharides 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. 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($Å^2$)	Q<0.9
5	SO4	A	903	5/5	0.46	0.33	98,100,101,101	5
5	SO4	B	903	5/5	0.78	0.29	102,102,103,103	0
4	OGA	A	902	10/10	0.90	0.15	33,35,42,44	0
6	EDO	B	904	4/4	0.91	0.22	29,31,33,37	0
4	OGA	B	902	10/10	0.94	0.14	22,27,35,35	0
3	MN	B	901	1/1	0.97	0.13	19,19,19,19	0
3	MN	A	901	1/1	0.99	0.13	26,26,26,26	0

6.5 Other polymers ⓘ

There are no such residues in this entry.