wwPDB X-ray Structure Validation Summary Report (i)

Oct 26, 2021 – 05:03 pm BST

PDB ID	:	6Z49
Title	:	Crystal structure of deubiquitinase Mindy2
Authors	:	Abdul Rehman, S.A.; Kulathu, Y.
Deposited on	:	2020-05-23
Resolution	:	2.00 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 following versions of software and data (see references (1)) were used in the production of this report:

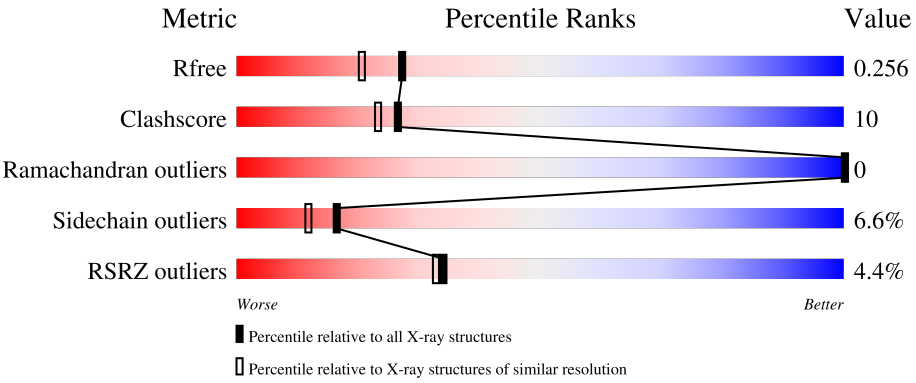
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 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	273	% 74% 16% • 8%
1	B	273	% 78% 14% • 7%
1	C	273	8% 74% 17% • 5%
1	D	273	6% 73% 16% • 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PEG	D	603	-	-	X	-
4	PG4	B	605	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7949 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 Ubiquitin carboxyl-terminal hydrolase MINDY-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	252	Total	C	N	O	S	0	0	0
			1955	1256	316	370	13			
1	B	254	Total	C	N	O	S	0	0	0
			1993	1284	319	377	13			
1	C	258	Total	C	N	O	S	0	0	0
			1990	1279	325	372	14			
1	D	243	Total	C	N	O	S	0	0	0
			1856	1197	305	342	12			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	232	GLY	-	expression tag	UNP Q8NBR6
A	233	PRO	-	expression tag	UNP Q8NBR6
A	234	LEU	-	expression tag	UNP Q8NBR6
A	235	GLY	-	expression tag	UNP Q8NBR6
A	236	SER	-	expression tag	UNP Q8NBR6
A	237	PRO	-	expression tag	UNP Q8NBR6
A	238	GLU	-	expression tag	UNP Q8NBR6
A	239	PHE	-	expression tag	UNP Q8NBR6
A	240	MET	-	expression tag	UNP Q8NBR6
B	232	GLY	-	expression tag	UNP Q8NBR6
B	233	PRO	-	expression tag	UNP Q8NBR6
B	234	LEU	-	expression tag	UNP Q8NBR6
B	235	GLY	-	expression tag	UNP Q8NBR6
B	236	SER	-	expression tag	UNP Q8NBR6
B	237	PRO	-	expression tag	UNP Q8NBR6
B	238	GLU	-	expression tag	UNP Q8NBR6
B	239	PHE	-	expression tag	UNP Q8NBR6
B	240	MET	-	expression tag	UNP Q8NBR6
C	232	GLY	-	expression tag	UNP Q8NBR6
C	233	PRO	-	expression tag	UNP Q8NBR6
C	234	LEU	-	expression tag	UNP Q8NBR6

Continued on next page...

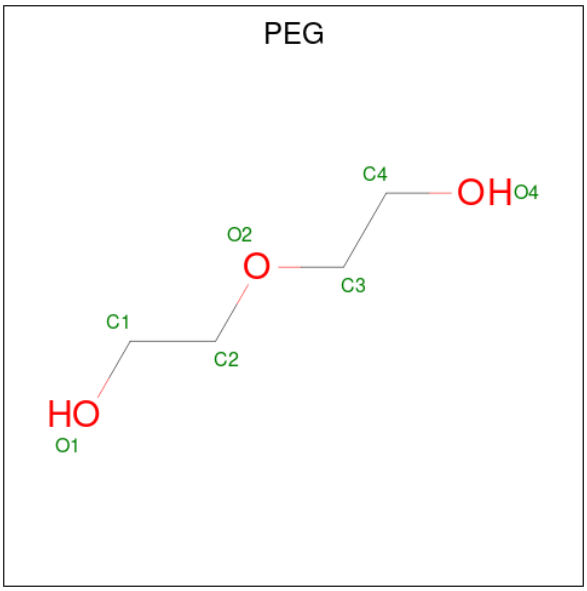
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	235	GLY	-	expression tag	UNP Q8NBR6
C	236	SER	-	expression tag	UNP Q8NBR6
C	237	PRO	-	expression tag	UNP Q8NBR6
C	238	GLU	-	expression tag	UNP Q8NBR6
C	239	PHE	-	expression tag	UNP Q8NBR6
C	240	MET	-	expression tag	UNP Q8NBR6
D	232	GLY	-	expression tag	UNP Q8NBR6
D	233	PRO	-	expression tag	UNP Q8NBR6
D	234	LEU	-	expression tag	UNP Q8NBR6
D	235	GLY	-	expression tag	UNP Q8NBR6
D	236	SER	-	expression tag	UNP Q8NBR6
D	237	PRO	-	expression tag	UNP Q8NBR6
D	238	GLU	-	expression tag	UNP Q8NBR6
D	239	PHE	-	expression tag	UNP Q8NBR6
D	240	MET	-	expression tag	UNP Q8NBR6

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

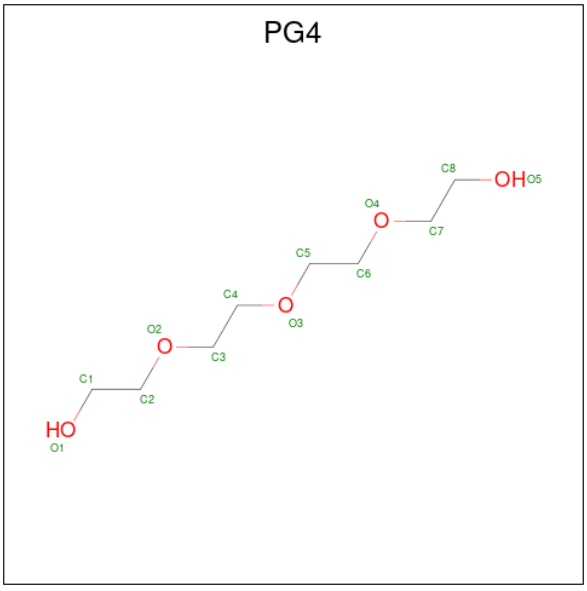
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0
2	B	3	Total Cl 3 3	0	0
2	C	1	Total Cl 1 1	0	0
2	D	2	Total Cl 2 2	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 7 4 3	0	0
3	B	1	Total C O 7 4 3	0	0
3	D	1	Total C O 7 4 3	0	0

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 13 8 5	0	0

- Molecule 5 is water.

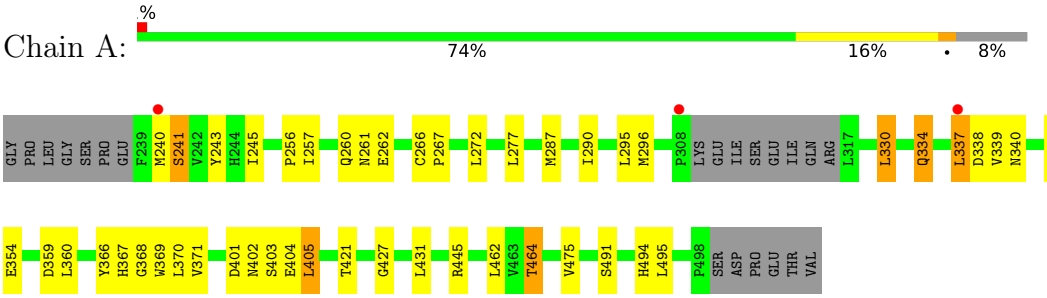
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	35	Total O 35 35	0	0
5	B	42	Total O 42 42	0	0
5	C	20	Total O 20 20	0	0
5	D	17	Total O 17 17	0	0

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.

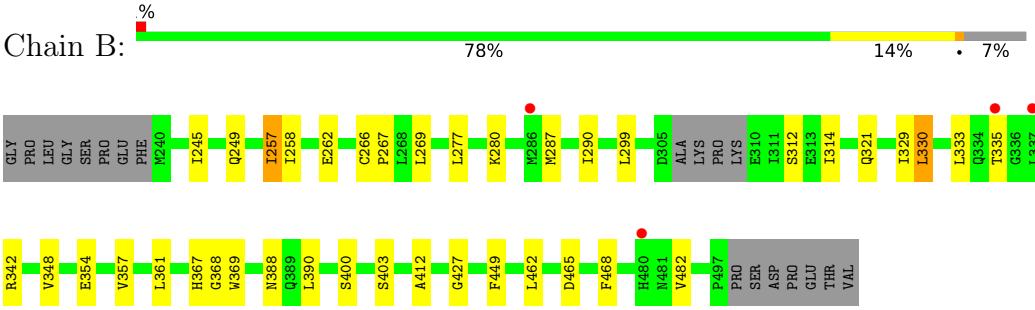
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase MINDY-2

Chain A:



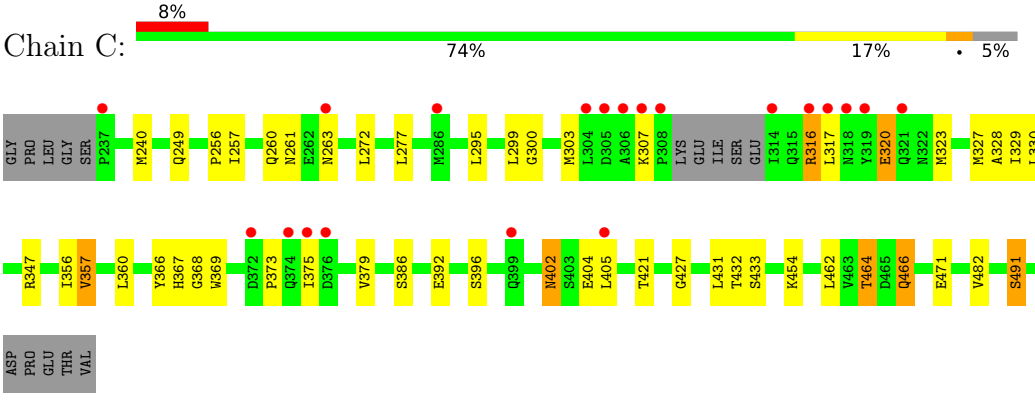
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase MINDY-2

Chain B:



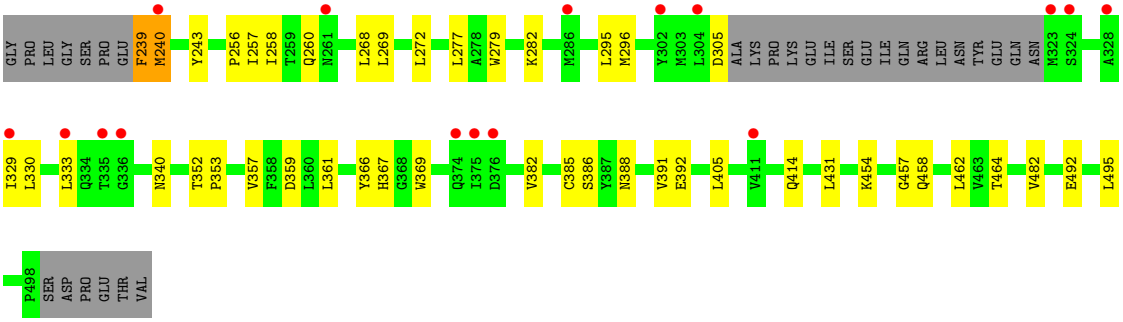
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase MINDY-2

Chain C:



• Molecule 1: Ubiquitin carboxyl-terminal hydrolase MINDY-2

Chain D:


GLY PRO LEU GLY SER PRO GLU F239 M240 Y243 P256 I257 I258 T259 Q260 N261 L268 L269 L272 L277 A278 W279 K282 M286 L295 M296 Y302 M303 L304 D305 ALA LYS PRO LYS GLU ILE SER GLU ILE GLN ARG LEU ASN TYR GLU GLN ASN M323 S324 A328
I329 L330 L333 Q334 T335 G336 N340 T352 P353 V357 F358 D359 L360 L361 Y366 H367 G368 W369 Q374 I375 D376 V382 C385 S386 Y387 N388 V391 E392 L405 V411 Q414 L431 K454 G457 Q458 L462 V463 T464 V482 E492 L495
P498 SER ASP PRO GLU THR VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	67.11Å 120.42Å 78.38Å 90.00° 92.94° 90.00°	Depositor
Resolution (Å)	47.72 – 2.00 47.72 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.72-2.00) 99.6 (47.72-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.215 , 0.251 0.227 , 0.256	Depositor DCC
R_{free} test set	4169 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	34.3	Xtriage
Anisotropy	0.574	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7949	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.84 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.4502e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: PEG, CL, PG4

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	0.71	0/1999	0.83	0/2725
1	B	0.71	0/2038	0.83	0/2776
1	C	0.73	0/2035	0.82	0/2774
1	D	0.72	0/1899	0.81	0/2589
All	All	0.72	0/7971	0.82	0/10864

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	1955	0	1828	35	0
1	B	1993	0	1893	51	0
1	C	1990	0	1870	42	0
1	D	1856	0	1728	30	0
2	A	1	0	0	0	0
2	B	3	0	0	0	0
2	C	1	0	0	0	0
2	D	2	0	0	0	0
3	A	7	0	10	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	7	0	10	0	0
3	D	7	0	10	5	0
4	B	13	0	18	10	0
5	A	35	0	0	0	0
5	B	42	0	0	1	0
5	C	20	0	0	0	0
5	D	17	0	0	0	0
All	All	7949	0	7367	157	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 10.

The worst 5 of 157 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:280:LYS:NZ	4:B:605:PG4:H11	1.71	1.05
1:D:282:LYS:HE2	3:D:603:PEG:H22	1.38	1.01
1:A:421:THR:HG22	1:A:445:ARG:HH21	1.25	0.99
1:A:421:THR:CG2	1:A:445:ARG:HH21	1.75	0.99
1:B:287:MET:HE3	1:B:290:ILE:HB	1.45	0.98

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	248/273 (91%)	243 (98%)	5 (2%)	0	100	100
1	B	250/273 (92%)	242 (97%)	8 (3%)	0	100	100
1	C	254/273 (93%)	244 (96%)	10 (4%)	0	100	100
1	D	239/273 (88%)	231 (97%)	8 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	991/1092 (91%)	960 (97%)	31 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	207/249 (83%)	193 (93%)	14 (7%)	16 11
1	B	214/249 (86%)	204 (95%)	10 (5%)	26 22
1	C	208/249 (84%)	190 (91%)	18 (9%)	10 6
1	D	190/249 (76%)	178 (94%)	12 (6%)	18 13
All	All	819/996 (82%)	765 (93%)	54 (7%)	16 12

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	347	ARG
1	C	433	SER
1	D	414	GLN
1	C	357	VAL
1	C	402	ASN

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

Mol	Chain	Res	Type
1	C	414	GLN
1	C	494	HIS
1	D	414	GLN
1	D	260	GLN
1	C	480	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 7 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	$\#\|Z\| > 2$
3	PEG	D	603	-	6,6,6	0.10	0	5,5,5	0.08	0
4	PG4	B	605	-	12,12,12	0.14	0	11,11,11	0.12	0
3	PEG	B	604	-	6,6,6	0.12	0	5,5,5	0.08	0
3	PEG	A	602	-	6,6,6	0.11	0	5,5,5	0.07	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
3	PEG	D	603	-	-	3/4/4/4	-
4	PG4	B	605	-	-	8/10/10/10	-
3	PEG	B	604	-	-	1/4/4/4	-
3	PEG	A	602	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	605	PG4	C4-C3-O2-C2
4	B	605	PG4	O3-C5-C6-O4
4	B	605	PG4	O2-C3-C4-O3
3	B	604	PEG	O1-C1-C2-O2
3	D	603	PEG	O2-C3-C4-O4

There are no ring outliers.

2 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	603	PEG	5	0
4	B	605	PG4	10	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	252/273 (92%)	-0.03	3 (1%)	79	78	25, 38, 76, 106	0
1	B	254/273 (93%)	-0.08	4 (1%)	72	70	27, 40, 67, 110	0
1	C	258/273 (94%)	0.37	21 (8%)	12	11	28, 49, 89, 123	0
1	D	243/273 (89%)	0.31	16 (6%)	18	17	30, 50, 90, 109	0
All	All	1007/1092 (92%)	0.14	44 (4%)	34	33	25, 44, 82, 123	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	319	TYR	10.7
1	C	318	ASN	9.7
1	C	317	LEU	6.5
1	D	336	GLY	5.8
1	B	337	LEU	5.4

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, 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	CL	B	603	1/1	0.70	0.12	77,77,77,77	0
2	CL	A	601	1/1	0.82	0.17	72,72,72,72	0
4	PG4	B	605	13/13	0.85	0.14	55,65,76,77	0
3	PEG	A	602	7/7	0.89	0.14	50,54,61,61	0
3	PEG	B	604	7/7	0.90	0.16	53,54,61,67	0
2	CL	C	601	1/1	0.90	0.07	69,69,69,69	0
2	CL	D	602	1/1	0.91	0.06	82,82,82,82	0
2	CL	B	602	1/1	0.93	0.11	69,69,69,69	0
2	CL	D	601	1/1	0.93	0.09	56,56,56,56	0
3	PEG	D	603	7/7	0.94	0.09	73,73,76,76	0
2	CL	B	601	1/1	0.97	0.13	53,53,53,53	0

6.5 Other polymers (i)

There are no such residues in this entry.