



wwPDB X-ray Structure Validation Summary Report

Nov 25, 2014 – 08:01 AM EST

PDB ID : 4OSG
Title : Klebsiella pneumoniae complexed with NADPH and 6-ethyl-5-[(3R)-3-[3-methoxyl-5-(pyridine-4-yl)phenyl]but-1-yn-1-yl]pyrimidine-2,4-diamine(UCP1006)
Authors : Lamb, K.M.; Anderson, A.C.
Deposited on : 2014-02-12
Resolution : 2.70 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>

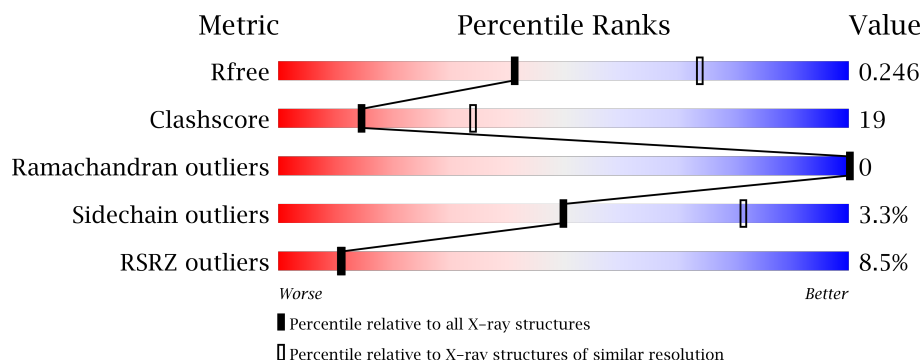
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24195
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.1.3
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24195

1 Overall quality at a glance

The reported resolution of this entry is 2.70 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	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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	165	
1	B	165	
1	C	165	
1	D	165	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	06U	A	402	-	X
3	06U	C	402	-	X
3	06U	D	402	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 5399 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 Dihydrofolate reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	159	Total	C	N	O	S	0	0	0
			1264	802	215	240	7			
1	B	159	Total	C	N	O	S	0	0	0
			1264	802	215	240	7			
1	C	159	Total	C	N	O	S	0	0	0
			1264	802	215	240	7			
1	D	159	Total	C	N	O	S	0	0	0
			1264	802	215	240	7			

There are 24 discrepancies between the modelled and reference sequences:

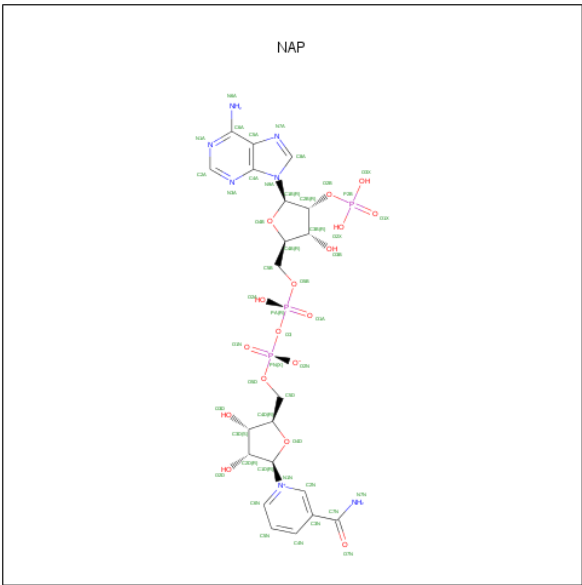
Chain	Residue	Modelled	Actual	Comment	Reference
A	160	HIS	-	EXPRESSION TAG	UNP U5M636
A	161	HIS	-	EXPRESSION TAG	UNP U5M636
A	162	HIS	-	EXPRESSION TAG	UNP U5M636
A	163	HIS	-	EXPRESSION TAG	UNP U5M636
A	164	HIS	-	EXPRESSION TAG	UNP U5M636
A	165	HIS	-	EXPRESSION TAG	UNP U5M636
B	160	HIS	-	EXPRESSION TAG	UNP U5M636
B	161	HIS	-	EXPRESSION TAG	UNP U5M636
B	162	HIS	-	EXPRESSION TAG	UNP U5M636
B	163	HIS	-	EXPRESSION TAG	UNP U5M636
B	164	HIS	-	EXPRESSION TAG	UNP U5M636
B	165	HIS	-	EXPRESSION TAG	UNP U5M636
C	160	HIS	-	EXPRESSION TAG	UNP U5M636
C	161	HIS	-	EXPRESSION TAG	UNP U5M636
C	162	HIS	-	EXPRESSION TAG	UNP U5M636
C	163	HIS	-	EXPRESSION TAG	UNP U5M636
C	164	HIS	-	EXPRESSION TAG	UNP U5M636
C	165	HIS	-	EXPRESSION TAG	UNP U5M636
D	160	HIS	-	EXPRESSION TAG	UNP U5M636
D	161	HIS	-	EXPRESSION TAG	UNP U5M636
D	162	HIS	-	EXPRESSION TAG	UNP U5M636

Continued on next page...

Continued from previous page...

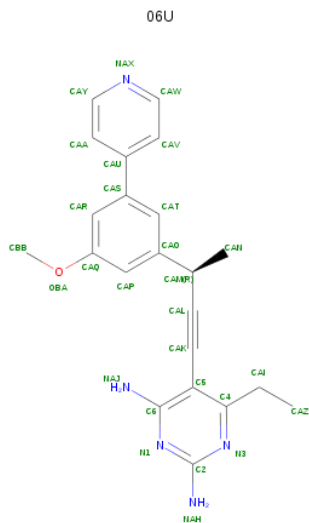
Chain	Residue	Modelled	Actual	Comment	Reference
D	163	HIS	-	EXPRESSION TAG	UNP U5M636
D	164	HIS	-	EXPRESSION TAG	UNP U5M636
D	165	HIS	-	EXPRESSION TAG	UNP U5M636

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDEPHOSPHATE (three-letter code: NAP) (formula: C₂₁H₂₈N₇O₁₇P₃).



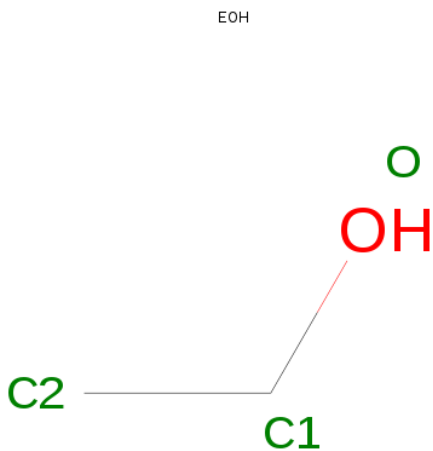
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	D	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 3 is 6-ETHYL-5-{(3R)-3-[3-METHOXY-5-(PYRIDIN-4-YL)PHENYL]BUT-1-YN-1-YL}PYRIMIDINE-2,4-DIAMINE (three-letter code: 06U) (formula: C₂₂H₂₃N₅O).



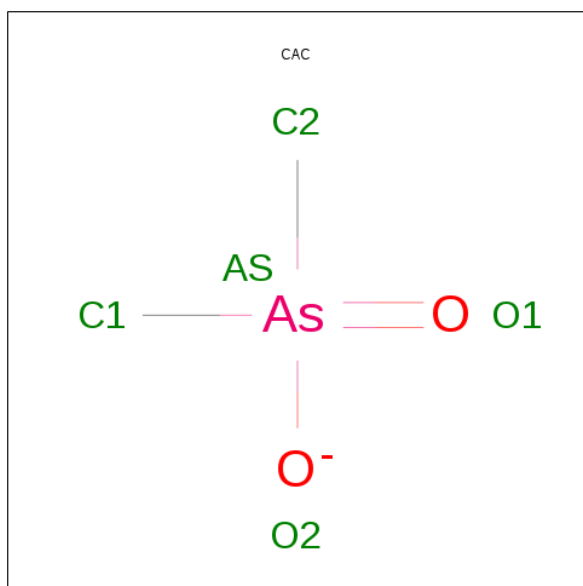
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 28	C 22	N 5	O 1	0	0
3	B	1	Total 28	C 22	N 5	O 1	0	0
3	C	1	Total 28	C 22	N 5	O 1	0	0
3	D	1	Total 28	C 22	N 5	O 1	0	0

- Molecule 4 is ETHANOL (three-letter code: EOH) (formula: $\text{C}_2\text{H}_6\text{O}$).



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

- Molecule 5 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$).

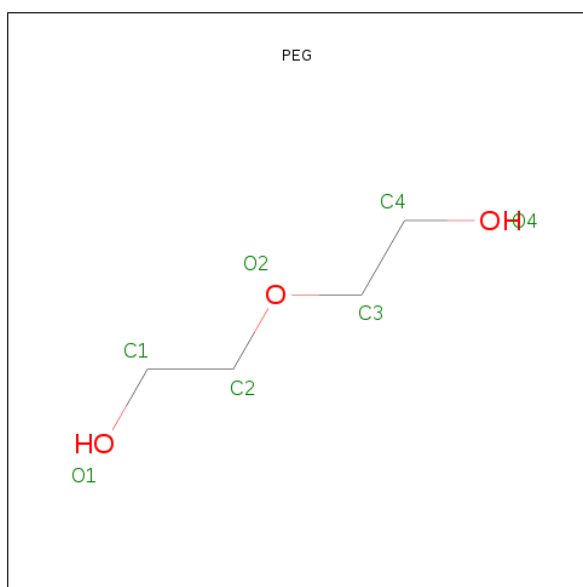


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	As	C	O	0	0
			5	1	2	2		
5	D	1	Total	As	C	O	0	0
			5	1	2	2		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	1	Total	Ca	0	0
			1	1		
8	C	1	Total	Ca	0	0
			1	1		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	5	Total	O	0	0
			5	5		
9	B	1	Total	O	0	0
			1	1		
9	C	3	Total	O	0	0
			3	3		
9	D	4	Total	O	0	0
			4	4		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	36.12Å 74.21Å 82.53Å 67.94° 77.70° 75.92°	Depositor
Resolution (Å)	39.27 – 2.70 39.27 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.2 (39.27-2.70) 92.8 (39.27-2.70)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.95 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.239 , 0.248 0.233 , 0.246	Depositor DCC
R_{free} test set	1982 reflections (10.03%)	DCC
Wilson B-factor (Å ²)	71.8	Xtriage
Anisotropy	0.728	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 54.6	EDS
Estimated twinning fraction	0.448 for h,h-k,h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 19765 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5399	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.69% 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: 06U, CL, CA, EOH, NAP, CAC, PEG

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	0.37	0/1296	0.58	0/1761
1	B	0.30	0/1296	0.56	0/1761
1	C	0.31	0/1296	0.56	0/1761
1	D	0.39	1/1296 (0.1%)	0.60	0/1761
All	All	0.34	1/5184 (0.0%)	0.57	0/7044

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	59	ASN	C-N	-7.11	1.17	1.34

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	1264	0	1220	54	0
1	B	1264	0	1220	48	0
1	C	1264	0	1220	34	0
1	D	1264	0	1219	56	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	48	0	25	7	0
2	B	48	0	25	13	0
2	C	48	0	25	3	0
2	D	48	0	25	8	0
3	A	28	0	23	8	0
3	B	28	0	23	5	0
3	C	28	0	23	9	0
3	D	28	0	23	5	0
4	A	3	0	6	0	0
5	B	5	0	0	0	0
5	D	5	0	0	0	0
6	B	1	0	0	0	0
6	C	2	0	0	0	0
6	D	1	0	0	0	0
7	C	7	0	10	0	0
8	C	1	0	0	0	0
8	D	1	0	0	0	0
9	A	5	0	0	0	0
9	B	1	0	0	0	0
9	C	3	0	0	0	0
9	D	4	0	0	0	0
All	All	5399	0	5087	202	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 19.

The worst 5 of 202 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:45:LEU:HD21	2:B:401:NAP:O5B	1.31	1.27
1:B:45:LEU:HD21	2:B:401:NAP:PA	1.92	1.09
1:C:50:ILE:HG22	1:C:52:ARG:H	1.34	0.91
1:D:65:LYS:HE2	1:D:66:PRO:HD2	1.52	0.90
1:C:54:LEU:H	1:C:59:ASN:HD21	1.19	0.87

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	157/165 (95%)	153 (98%)	4 (2%)	0	100	100
1	B	157/165 (95%)	155 (99%)	2 (1%)	0	100	100
1	C	157/165 (95%)	155 (99%)	2 (1%)	0	100	100
1	D	157/165 (95%)	155 (99%)	2 (1%)	0	100	100
All	All	628/660 (95%)	618 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	136/142 (96%)	132 (97%)	4 (3%)	55	85
1	B	136/142 (96%)	131 (96%)	5 (4%)	45	78
1	C	136/142 (96%)	133 (98%)	3 (2%)	64	90
1	D	136/142 (96%)	130 (96%)	6 (4%)	39	71
All	All	544/568 (96%)	526 (97%)	18 (3%)	50	81

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	158	ARG
1	C	75	VAL
1	D	59	ASN
1	B	52	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	148	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	59	ASN
1	C	147	ASN

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 18 ligands modelled in this entry, 6 are monoatomic - leaving 12 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	NAP	A	401	-	52,52,52	1.39	3 (5%)	80,80,80	1.67	6 (7%)
3	06U	A	402	-	30,30,30	2.00	6 (20%)	41,41,41	2.11	11 (26%)
4	EOH	A	403	-	2,2,2	0.38	0	1,1,1	0.36	0
2	NAP	B	401	-	52,52,52	1.39	3 (5%)	80,80,80	1.76	7 (8%)
3	06U	B	402	-	30,30,30	2.05	5 (16%)	41,41,41	2.03	10 (24%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	CAC	B	403	-	4,4,4	3.57	1 (25%)	6,6,6	5.89	1 (16%)
2	NAP	C	401	-	52,52,52	1.40	3 (5%)	80,80,80	1.73	11 (13%)
3	06U	C	402	-	30,30,30	3.44	8 (26%)	41,41,41	2.05	10 (24%)
7	PEG	C	403	-	6,6,6	0.59	0	5,5,5	0.39	0
2	NAP	D	401	-	52,52,52	1.40	3 (5%)	80,80,80	1.71	7 (8%)
3	06U	D	402	-	30,30,30	3.14	8 (26%)	41,41,41	1.91	9 (21%)
5	CAC	D	405	-	4,4,4	3.57	1 (25%)	6,6,6	5.98	1 (16%)

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	NAP	A	401	-	-	0/35/67/67	0/5/5/5
3	06U	A	402	-	-	0/17/17/17	0/3/3/3
4	EOH	A	403	-	-	0/0/0/0	0/0/0/0
2	NAP	B	401	-	-	0/35/67/67	0/5/5/5
3	06U	B	402	-	-	0/17/17/17	0/3/3/3
5	CAC	B	403	-	-	0/0/0/0	0/0/0/0
2	NAP	C	401	-	-	0/35/67/67	0/5/5/5
3	06U	C	402	-	-	0/17/17/17	0/3/3/3
7	PEG	C	403	-	-	0/4/4/4	0/0/0/0
2	NAP	D	401	-	-	0/35/67/67	0/5/5/5
3	06U	D	402	-	-	0/17/17/17	0/3/3/3
5	CAC	D	405	-	-	0/0/0/0	0/0/0/0

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	402	06U	CAK-CAL	-14.55	1.05	1.19
3	D	402	06U	CAK-CAL	-13.10	1.06	1.19
3	C	402	06U	CAO-CAM	-7.86	1.39	1.52
2	D	401	NAP	O7N-C7N	7.68	1.41	1.24
3	B	402	06U	CAO-CAM	-7.66	1.39	1.52

The worst 5 of 73 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	405	CAC	O2-AS-O1	-14.63	109.38	112.54
5	B	403	CAC	O2-AS-O1	-14.40	109.43	112.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	NAP	N3A-C2A-N1A	-10.16	119.95	128.89
2	A	401	NAP	N3A-C2A-N1A	-9.76	120.30	128.89
2	B	401	NAP	N3A-C2A-N1A	-9.65	120.40	128.89

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	159/165 (96%)	0.77	19 (11%) 5 5	60, 81, 110, 128	0
1	B	159/165 (96%)	0.70	12 (7%) 14 15	60, 86, 118, 132	0
1	C	159/165 (96%)	0.78	13 (8%) 12 12	57, 75, 109, 127	0
1	D	159/165 (96%)	0.69	10 (6%) 19 21	68, 94, 129, 144	0
All	All	636/660 (96%)	0.73	54 (8%) 11 11	57, 84, 119, 144	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	66	PRO	5.9
1	B	131	ASP	5.6
1	C	69	ASP	4.9
1	B	59	ASN	3.8
1	A	60	ILE	3.6

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(\AA^2)	Q<0.9
3	06U	A	402	28/28	0.38	3.38	65,85,102,103	0
3	06U	D	402	28/28	0.37	2.75	67,89,101,102	0
3	06U	C	402	28/28	0.35	2.42	60,78,94,97	0
4	EOH	A	403	3/3	0.32	1.31	90,90,95,100	0
3	06U	B	402	28/28	0.23	0.51	64,75,83,91	0
5	CAC	D	405	5/5	0.20	0.31	127,143,163,181	0
7	PEG	C	403	7/7	0.23	-0.05	50,72,75,86	0
2	NAP	C	401	48/48	0.22	-0.10	58,72,112,125	0
2	NAP	D	401	48/48	0.22	-0.28	74,118,126,127	0
2	NAP	B	401	48/48	0.21	-0.35	67,108,118,126	0
2	NAP	A	401	48/48	0.18	-0.86	68,94,115,117	0
5	CAC	B	403	5/5	0.19	-2.04	132,142,163,181	0
6	CL	D	403	1/1	0.09	-2.31	88,88,88,88	0
6	CL	C	404	1/1	0.07	-2.54	94,94,94,94	0
8	CA	C	405	1/1	0.17	-2.61	82,82,82,82	0
6	CL	B	404	1/1	0.22	-	147,147,147,147	0
6	CL	C	406	1/1	0.12	-	87,87,87,87	0
8	CA	D	404	1/1	0.12	-	108,108,108,108	0

6.5 Other polymers ⓘ

There are no such residues in this entry.