Full wwPDB X-ray Structure Validation Report (i)

Dec 16, 2023 – 12:28 PM EST

PDB ID	:	4P40
Title	:	Chlamydia pneumoniae CopN
Authors	:	Nawrotek, A.; Guimaraes, B.G.; Knossow, M.; Gigant, B.
Deposited on	:	2014-03-10
Resolution	:	1.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

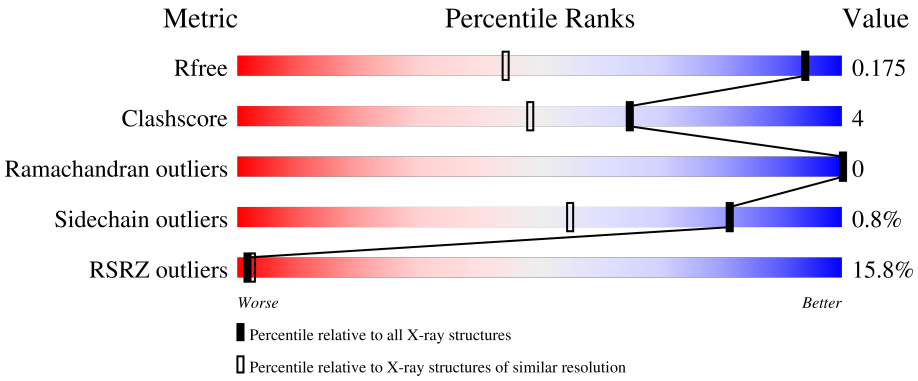
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.20 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	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)
RSRZ outliers	127900	1200 (1.22-1.18)

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	425	11% 61% 6% 32%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5089 atoms, of which 2296 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.

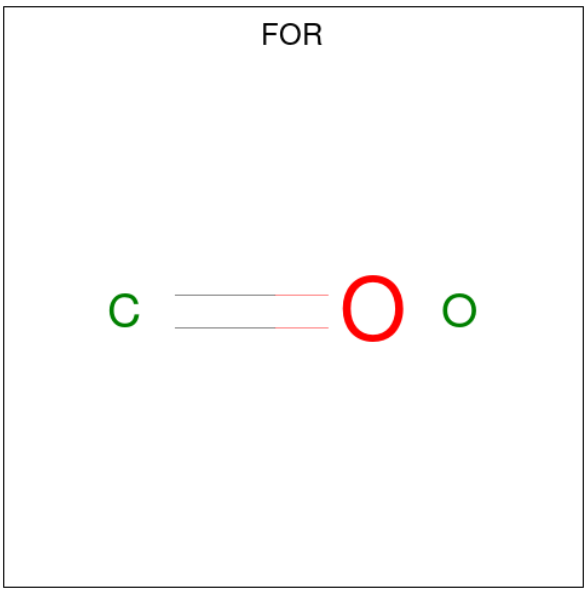
- Molecule 1 is a protein called CopN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	288	4596	1452	2279	387	469	9	0	18	0

There are 26 discrepancies between the modelled and reference sequences:

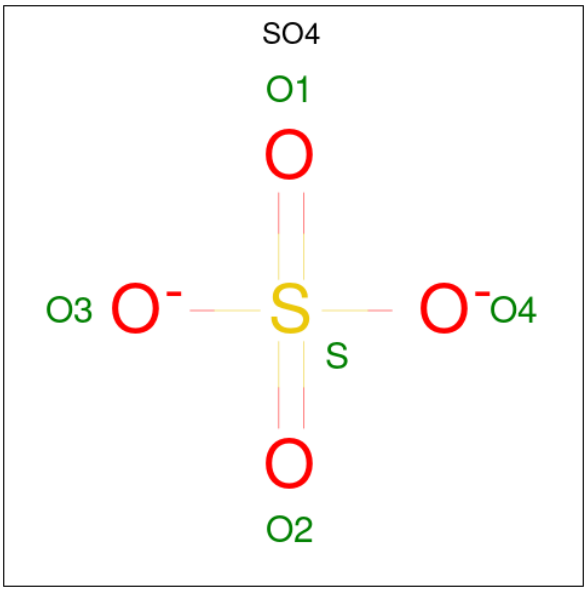
Chain	Residue	Modelled	Actual	Comment	Reference
A	-25	MET	-	initiating methionine	UNP Q9Z8L4
A	-24	LYS	-	expression tag	UNP Q9Z8L4
A	-23	HIS	-	expression tag	UNP Q9Z8L4
A	-22	HIS	-	expression tag	UNP Q9Z8L4
A	-21	HIS	-	expression tag	UNP Q9Z8L4
A	-20	HIS	-	expression tag	UNP Q9Z8L4
A	-19	HIS	-	expression tag	UNP Q9Z8L4
A	-18	HIS	-	expression tag	UNP Q9Z8L4
A	-17	PRO	-	expression tag	UNP Q9Z8L4
A	-16	MET	-	expression tag	UNP Q9Z8L4
A	-15	SER	-	expression tag	UNP Q9Z8L4
A	-14	ASP	-	expression tag	UNP Q9Z8L4
A	-13	TYR	-	expression tag	UNP Q9Z8L4
A	-12	ASP	-	expression tag	UNP Q9Z8L4
A	-11	ILE	-	expression tag	UNP Q9Z8L4
A	-10	PRO	-	expression tag	UNP Q9Z8L4
A	-9	THR	-	expression tag	UNP Q9Z8L4
A	-8	THR	-	expression tag	UNP Q9Z8L4
A	-7	GLU	-	expression tag	UNP Q9Z8L4
A	-6	ASN	-	expression tag	UNP Q9Z8L4
A	-5	LEU	-	expression tag	UNP Q9Z8L4
A	-4	TYR	-	expression tag	UNP Q9Z8L4
A	-3	PHE	-	expression tag	UNP Q9Z8L4
A	-2	GLN	-	expression tag	UNP Q9Z8L4
A	-1	GLY	-	expression tag	UNP Q9Z8L4
A	0	ALA	-	expression tag	UNP Q9Z8L4

- Molecule 2 is FORMYL GROUP (three-letter code: FOR) (formula: CH_2O).



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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

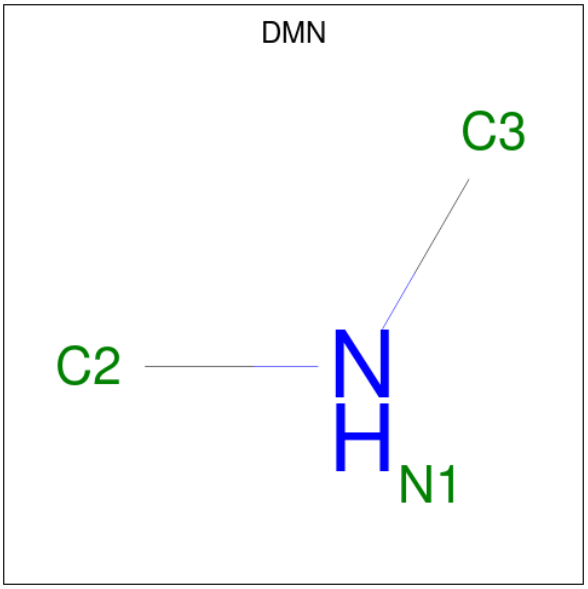


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	4	Total Cl 4 4	0	0

- Molecule 5 is DIMETHYLAMINE (three-letter code: DMN) (formula: C_2H_7N).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	N	0	0
			10	2	7	1		

- Molecule 6 is water.

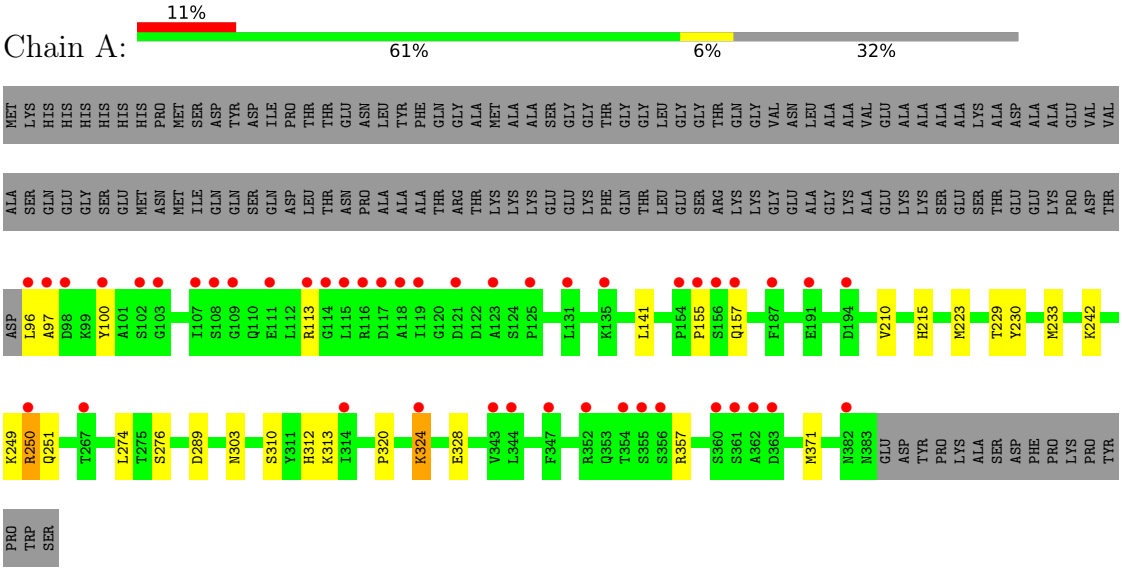
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	449	Total	O	0	0
			449	449		

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: CopN

Chain A:



4 Data and refinement statistics (i)

Property	Value	So
Space group	P 21 21 21	Dep
Cell constants a, b, c, α, β, γ	66.80Å 67.97Å 85.04Å 90.00° 90.00° 90.00°	Dep
Resolution (Å)	41.56 – 1.20 42.52 – 1.20	Dep E
% Data completeness (in resolution range)	97.4 (41.56-1.20) 97.4 (42.52-1.20)	Dep E
R_{merge}	(Not available)	Dep
R_{sym}	(Not available)	Dep
$< I/\sigma(I) >$ [1]	50.00 (at 1.20Å)	Xtr
Refinement program	PHENIX (phenix.refine: 1.8.4_1496), SHELXL, BUSTER-TNT	Dep
R, R_{free}	0.143 , 0.162 0.160 , 0.175	Dep D
R_{free} test set	5946 reflections (5.03%)	wwPl
Wilson B-factor (Å^2)	17.4	Xtr
Anisotropy	0.040	Xtr
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.45 , 62.9	E
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtr
Estimated twinning fraction	0.000 for k,h,-l	Xtr
F_o,F_c correlation	0.97	E
Total number of atoms	5089	wwPl
Average B, all atoms (Å^2)	24.0	wwPl

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% 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: MLY, MLZ, SO4, CL, FOR, DMN

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	0.68	2/2349 (0.1%)	0.77	4/3187 (0.1%)

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	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	324	LYS	C-N	11.57	1.60	1.34
1	A	155	PRO	N-CD	5.76	1.55	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	324	LYS	O-C-N	-14.19	99.99	122.70
1	A	324	LYS	CA-C-N	-5.47	105.17	117.20
1	A	250	ARG	O-C-N	-5.46	113.96	122.70
1	A	357	ARG	NE-CZ-NH2	-5.15	117.73	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	324	LYS	Mainchain

5.2 Too-close contacts (i)

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	2317	2279	2291	18	0
2	A	10	10	0	1	0
3	A	10	0	0	0	0
4	A	4	0	0	1	0
5	A	3	7	7	0	0
6	A	449	0	0	6	0
All	All	2793	2296	2298	19	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.

All (19) 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)
4:A:409:CL:CL	6:A:947:HOH:O	2.51	0.65
1:A:320:PRO:HB2	1:A:371[A]:MET:HE1	1.77	0.65
1:A:242:MLY:HH22	6:A:635:HOH:O	1.98	0.62
1:A:250:ARG:O	1:A:251:GLN:HB2	2.11	0.50
1:A:312:HIS:HD2	6:A:665:HOH:O	1.94	0.50
1:A:97:ALA:CB	1:A:113:ARG:HG3	2.42	0.49
1:A:276[B]:SER:OG	1:A:312:HIS:HE1	1.97	0.47
1:A:303:ASN:HD22	2:A:403:FOR:C	2.18	0.46
1:A:233[A]:MET:HE2	1:A:274:LEU:HD21	1.99	0.44
1:A:289[A]:ASP:OD1	6:A:891:HOH:O	2.21	0.44
1:A:96:LEU:HD12	1:A:97:ALA:H	1.82	0.43
1:A:276[B]:SER:OG	1:A:312:HIS:CE1	2.71	0.43
1:A:141:LEU:HD12	6:A:846:HOH:O	2.19	0.43
1:A:229:THR:O	1:A:233[A]:MET:HG3	2.19	0.42
1:A:230:TYR:CD1	1:A:233[B]:MET:CE	3.03	0.42
1:A:310[B]:SER:OG	1:A:328:GLU:HG3	2.20	0.42
1:A:157:GLN:NE2	6:A:837:HOH:O	2.54	0.41
1:A:276[A]:SER:OG	1:A:312:HIS:HE1	2.02	0.41
1:A:210:VAL:O	1:A:215:HIS:HD2	2.04	0.41

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	301/425 (71%)	297 (99%)	4 (1%)	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	256/351 (73%)	253 (99%)	3 (1%)	71 37

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	100[A]	TYR
1	A	100[B]	TYR
1	A	223	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	215	HIS
1	A	312	HIS
1	A	353	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

3 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	MLZ	A	313	1	8,9,10	0.93	1 (12%)	4,9,11	1.33	1 (25%)
1	MLY	A	242	1	9,10,11	0.62	0	6,11,13	0.74	0
1	MLY	A	249	1	9,10,11	0.47	0	6,11,13	1.70	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	MLZ	A	313	1	-	0/7/8/10	-
1	MLY	A	242	1	-	0/8/9/11	-
1	MLY	A	249	1	-	3/8/9/11	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	313	MLZ	O-C	2.42	1.29	1.19

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	249	MLY	CH2-NZ-CH1	2.62	116.52	109.73
1	A	313	MLZ	CM-NZ-CE	2.49	119.13	111.95
1	A	249	MLY	CD-CE-NZ	2.28	119.95	113.79

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	249	MLY	O-C-CA-CB
1	A	249	MLY	CD-CE-NZ-CH2
1	A	249	MLY	CD-CE-NZ-CH1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	242	MLY	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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	DMN	A	412	-	2,2,2	0.15	0	1,1,1	0.00	0
3	SO4	A	407	-	4,4,4	0.22	0	6,6,6	0.76	0
3	SO4	A	406	-	4,4,4	0.22	0	6,6,6	1.95	1 (16%)
2	FOR	A	402	-	0,1,1	-	-	-		
2	FOR	A	405	-	0,1,1	-	-	-		
2	FOR	A	404	-	0,1,1	-	-	-		
2	FOR	A	403	-	0,1,1	-	-	-		
2	FOR	A	401	-	0,1,1	-	-	-		

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	406	SO4	O4-S-O3	-3.75	93.06	109.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	403	FOR	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	324:LYS	C	325:VAL	N	1.60

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	285/425 (67%)	1.18	45 (15%)	2 2	13, 18, 34, 44	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	156	SER	7.5
1	A	96	LEU	7.2
1	A	356	SER	6.4
1	A	103	GLY	6.4
1	A	155	PRO	5.6
1	A	100[A]	TYR	5.5
1	A	97	ALA	5.2
1	A	154	PRO	4.7
1	A	187[A]	PHE	4.4
1	A	118	ALA	4.3
1	A	157	GLN	4.0
1	A	116	ARG	3.8
1	A	107	ILE	3.7
1	A	115	LEU	3.7
1	A	360[A]	SER	3.6
1	A	355	SER	3.5
1	A	361	SER	3.4
1	A	98	ASP	3.3
1	A	131	LEU	3.3
1	A	123	ALA	3.2
1	A	113	ARG	3.1
1	A	363[A]	ASP	3.0
1	A	267	THR	3.0
1	A	324	LYS	2.9
1	A	119	ILE	2.8
1	A	108	SER	2.8
1	A	362	ALA	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	117	ASP	2.7
1	A	347	PHE	2.7
1	A	194	ASP	2.7
1	A	135	LYS	2.6
1	A	314	ILE	2.6
1	A	382	ASN	2.6
1	A	109	GLY	2.5
1	A	250	ARG	2.5
1	A	125	PRO	2.4
1	A	114	GLY	2.3
1	A	111	GLU	2.3
1	A	344	LEU	2.3
1	A	352	ARG	2.3
1	A	354	THR	2.1
1	A	102	SER	2.0
1	A	343	VAL	2.0
1	A	191	GLU	2.0
1	A	121	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	MLY	A	242	11/12	0.90	0.22	13,29,40,40	0
1	MLZ	A	313	10/11	0.91	0.15	15,25,34,34	0
1	MLY	A	249	11/12	0.94	0.12	13,24,34,34	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	FOR	A	405	2/2	0.34	0.21	47,47,56,56	0
5	DMN	A	412	3/3	0.68	0.14	33,39,40,40	0
4	CL	A	411	1/1	0.69	0.12	72,72,72,72	0
2	FOR	A	401	2/2	0.71	0.26	46,47,56,56	0
2	FOR	A	403	2/2	0.77	0.12	54,54,64,64	0
4	CL	A	410	1/1	0.90	0.09	58,58,58,58	0
2	FOR	A	404	2/2	0.90	0.12	36,37,44,44	0
2	FOR	A	402	2/2	0.90	0.11	39,40,48,48	0
4	CL	A	408	1/1	0.93	0.33	60,60,60,60	0
4	CL	A	409	1/1	0.94	0.07	44,44,44,44	0
3	SO4	A	407	5/5	0.96	0.18	33,35,36,38	0
3	SO4	A	406	5/5	0.96	0.16	21,21,24,25	5

6.5 Other polymers (i)

There are no such residues in this entry.