



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 3, 2023 – 04:21 PM EDT

PDB ID : 3SPA
Title : Crystal Structure of Human Mitochondrial RNA Polymerase
Authors : Ringel, R.; Sologub, M.; Morozov, Y.I.; Litonin, D.; Cramer, P.; Temiakov, D.
Deposited on : 2011-07-01
Resolution : 2.50 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : 2.35
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) : 2.35

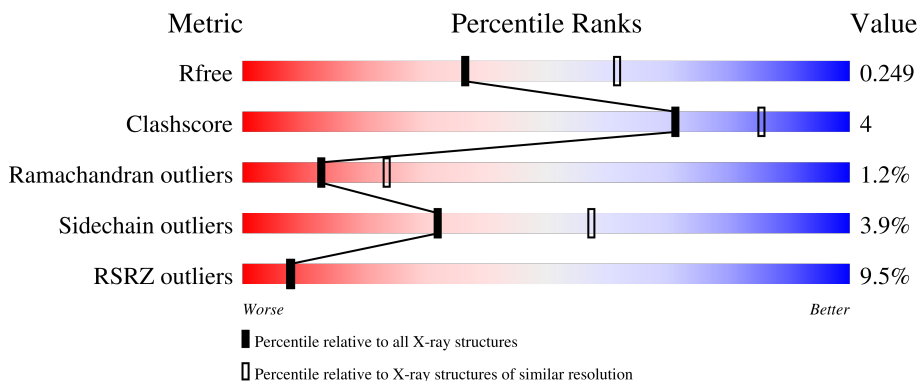
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


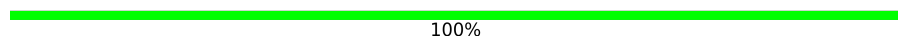
The reported resolution of this entry is 2.50 Å.

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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 $\leq 5\%$. 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	1134	 8% 73% 9% 17%
2	B	9	 100%

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	GOL	A	1	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7666 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 DNA-directed RNA polymerase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	941	7454	4745	1345	1315	49	0	7	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	97	MET	-	expression tag	UNP O00411
A	98	GLY	-	expression tag	UNP O00411
A	99	HIS	-	expression tag	UNP O00411
A	100	HIS	-	expression tag	UNP O00411
A	101	HIS	-	expression tag	UNP O00411
A	102	HIS	-	expression tag	UNP O00411
A	103	HIS	-	expression tag	UNP O00411
A	104	HIS	-	expression tag	UNP O00411
A	555	ALA	GLU	engineered mutation	UNP O00411

- Molecule 2 is a protein called Nonamer peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	9	45	27	9	9	0	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).

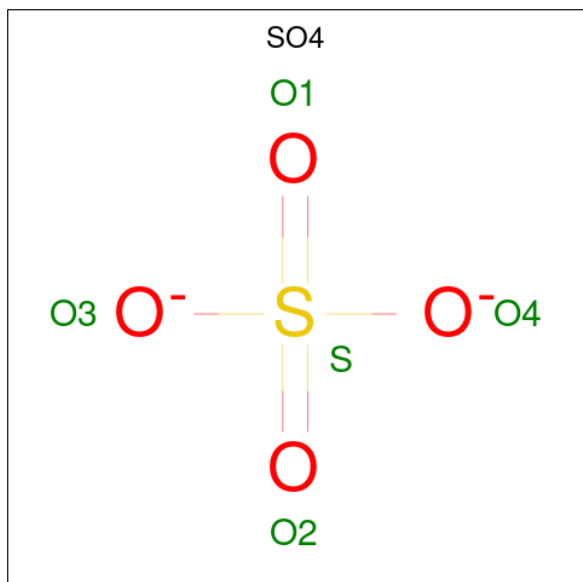


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	5	Total Cl 5 5	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	145	Total	O	0	0
			145	145		

4 Data and refinement statistics i

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	211.26Å 211.26Å 60.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.13 – 2.50 58.13 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (58.13-2.50) 99.8 (58.13-2.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.00 (at 2.51Å)	Xtrriage
Refinement program	BUSTER 2.11.1	Depositor
R, R_{free}	0.185 , 0.231 0.211 , 0.249	Depositor DCC
R_{free} test set	2347 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	61.3	Xtrriage
Anisotropy	0.549	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 73.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.013 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7666	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, GOL, 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	0.47	0/7651	0.69	3/10383 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	356	PRO	N-CA-CB	6.00	110.50	103.30
1	A	735	PRO	N-CA-CB	5.57	109.98	103.30
1	A	1085	ASP	CB-CG-OD2	5.18	122.96	118.30

There are no chirality outliers.

There are no planarity outliers.

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	7454	0	7444	53	0
2	B	45	0	11	0	0
3	A	12	0	16	5	0
4	A	5	0	0	0	0
5	A	5	0	0	0	0
6	A	145	0	0	2	0
All	All	7666	0	7471	53	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 (53) 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:A:218:GLN:O	1:A:218:GLN:HG3	1.77	0.78
1:A:1002:THR:HA	1:A:1003:ARG:O	1.93	0.67
1:A:1002:THR:H	1:A:1003:ARG:HG2	1.60	0.67
1:A:1039:SER:O	1:A:1043:MET:HG2	1.98	0.63
1:A:779:ALA:HB1	3:A:1:GOL:H32	1.82	0.60
1:A:507:ARG:HB3	1:A:511:ARG:HH21	1.68	0.59
1:A:1147:VAL:HG23	1:A:1154:TRP:HB2	1.84	0.58
1:A:397:GLN:HG2	1:A:531:ARG:HG2	1.87	0.56
1:A:248:VAL:HG11	1:A:463:GLY:HA3	1.87	0.56
1:A:218:GLN:O	1:A:218:GLN:CG	2.49	0.56
1:A:564:GLU:HG3	1:A:659:HIS:CD2	2.41	0.55
1:A:854[A]:ARG:HD3	1:A:954:ASP:HB2	1.88	0.54
1:A:1051[B]:GLN:NE2	6:A:1295:HOH:O	2.33	0.53
1:A:645:ASP:HB3	1:A:812:HIS:HD2	1.74	0.53
1:A:467:LEU:HD11	1:A:574:GLU:HB3	1.90	0.52
1:A:653:LEU:HD13	1:A:664:LEU:HD12	1.91	0.52
1:A:699:LEU:HD11	1:A:800:MET:HG3	1.92	0.52
1:A:238:GLN:HE22	1:A:469:PRO:HA	1.75	0.52
1:A:255:LYS:HA	1:A:258:LEU:HD13	1.93	0.50
1:A:803:ARG:HH22	1:A:1003:ARG:HH12	1.59	0.50
1:A:584:VAL:HG22	1:A:606:LEU:HB2	1.93	0.50
1:A:407:GLU:OE2	1:A:664:LEU:HB2	2.13	0.49
1:A:513:VAL:HG22	1:A:563:ARG:HG3	1.94	0.49
1:A:218:GLN:HA	6:A:1271:HOH:O	2.13	0.48
1:A:1007:ARG:HA	1:A:1010:ILE:HD12	1.95	0.48
1:A:255:LYS:O	1:A:258:LEU:HB2	2.13	0.48
1:A:388:LYS:HD2	1:A:538:ALA:O	2.13	0.48
1:A:847:ASN:HD21	1:A:860:ARG:HH11	1.60	0.47
1:A:675:GLU:HG2	1:A:1081:PRO:HB2	1.95	0.47
1:A:412:VAL:HG21	1:A:646:VAL:HG21	1.96	0.47
1:A:272:ALA:HB1	1:A:304:GLN:HB3	1.97	0.47
1:A:396:LEU:HB3	1:A:534:LEU:HD22	1.97	0.47
1:A:1002:THR:HA	1:A:1003:ARG:C	2.36	0.46
1:A:419:LYS:HE3	1:A:722:GLN:HB3	1.97	0.46
1:A:645:ASP:HA	1:A:811:PRO:HD2	1.96	0.46
1:A:403:GLN:HG3	1:A:664:LEU:HD22	1.97	0.46
1:A:473:LEU:HD22	1:A:512:HIS:CE1	2.51	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:820:VAL:HG13	3:A:1:GOL:H2	1.98	0.45
1:A:496:SER:HA	1:A:620:ILE:HA	1.99	0.45
1:A:279:GLU:O	1:A:283:VAL:HG23	2.16	0.45
1:A:782:ARG:HG3	1:A:813:PHE:HD1	1.82	0.45
1:A:639:LEU:HD11	1:A:780:LEU:HD13	1.98	0.45
1:A:645:ASP:HB3	1:A:812:HIS:CD2	2.52	0.44
1:A:670:LEU:HD11	1:A:699:LEU:HD12	1.99	0.44
1:A:826:GLU:OE2	3:A:2:GOL:H12	2.18	0.43
1:A:445:LEU:HD21	1:A:583:LEU:HA	2.02	0.42
1:A:779:ALA:CB	3:A:1:GOL:H32	2.49	0.42
1:A:563:ARG:HA	1:A:563:ARG:HD2	1.91	0.41
1:A:1062:SER:OG	1:A:1112:THR:HG22	2.20	0.41
1:A:948:PRO:HD3	1:A:1221:GLN:HB3	2.02	0.41
1:A:975:MET:HG2	1:A:977:VAL:HG23	2.01	0.41
1:A:1047:THR:O	1:A:1051[A]:GLN:HB2	2.20	0.41
1:A:820:VAL:CG1	3:A:1:GOL:H2	2.51	0.40

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	938/1134 (83%)	891 (95%)	36 (4%)	11 (1%)	13 24

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	359	SER
1	A	561	ALA
1	A	816	LEU
1	A	817	GLY
1	A	1002	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1003	ARG
1	A	382	GLY
1	A	662	ALA
1	A	565	GLN
1	A	1005	GLY
1	A	423	PRO

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	795/982 (81%)	764 (96%)	31 (4%)	32 57

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	218	GLN
1	A	245	LEU
1	A	269	LEU
1	A	297	LEU
1	A	320	LEU
1	A	364	LEU
1	A	384	VAL
1	A	529	HIS
1	A	608	HIS
1	A	620	ILE
1	A	638	THR
1	A	644	VAL
1	A	653	LEU
1	A	664	LEU
1	A	780	LEU
1	A	799	ASN
1	A	803	ARG
1	A	851	LEU
1	A	857	LEU
1	A	858	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	867	VAL
1	A	877	GLN
1	A	915	SER
1	A	920	HIS
1	A	977	VAL
1	A	980	VAL
1	A	1003	ARG
1	A	1004	TYR
1	A	1116	LYS
1	A	1185	ARG
1	A	1198	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	238	GLN
1	A	390	HIS
1	A	485	GLN
1	A	616	GLN
1	A	799	ASN
1	A	812	HIS
1	A	847	ASN
1	A	920	HIS
1	A	1209	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](#)

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 8 ligands modelled in this entry, 5 are monoatomic - leaving 3 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	GOL	A	2	-	5,5,5	0.52	0	5,5,5	0.38	0
3	GOL	A	1	-	5,5,5	0.72	0	5,5,5	1.02	0
5	SO4	A	1233	-	4,4,4	0.45	0	6,6,6	0.55	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	2	-	-	0/4/4/4	-
3	GOL	A	1	-	-	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
3	A	1	GOL	O1-C1-C2-C3
3	A	1	GOL	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2	GOL	1	0
3	A	1	GOL	4	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, 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	941/1134 (82%)	0.70	89 (9%) 8 8	43, 82, 144, 194	0
2	B	0/9	-	-	-	-
All	All	941/1143 (82%)	0.70	89 (9%) 8 8	43, 82, 144, 194	0

All (89) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	562	LEU	7.5
1	A	815	HIS	7.1
1	A	563	ARG	6.9
1	A	285	PHE	6.5
1	A	323	MET	6.4
1	A	325	GLN	5.9
1	A	258	LEU	5.2
1	A	218	GLN	5.2
1	A	353	LYS	5.1
1	A	1085	ASP	5.0
1	A	324	SER	4.9
1	A	336	ALA	4.8
1	A	359	SER	4.8
1	A	312	ASP	4.7
1	A	567	TRP	4.6
1	A	631	LEU	4.4
1	A	316	ILE	4.2
1	A	467	LEU	4.2
1	A	326	GLU	4.2
1	A	355	LYS	4.2
1	A	283	VAL	4.2
1	A	290	ALA	4.1
1	A	297	LEU	4.0
1	A	313	ALA	4.0
1	A	976	ARG	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	314	GLY	3.9
1	A	311	GLN	3.8
1	A	321	GLU	3.8
1	A	447	ARG	3.7
1	A	423	PRO	3.6
1	A	350	ALA	3.5
1	A	284	LEU	3.4
1	A	1151	ASP	3.4
1	A	278	LYS	3.4
1	A	249	HIS	3.3
1	A	616	GLN	3.1
1	A	246	LEU	3.1
1	A	320	LEU	3.1
1	A	421	THR	3.1
1	A	255	LYS	3.0
1	A	257	LYS	3.0
1	A	439	ASP	2.9
1	A	630	LEU	2.9
1	A	432	LYS	2.9
1	A	520	SER	2.9
1	A	633	LYS	2.9
1	A	287	VAL	2.8
1	A	434	LEU	2.8
1	A	611	SER	2.8
1	A	354	VAL	2.8
1	A	433	THR	2.8
1	A	286	MET	2.8
1	A	426	GLU	2.8
1	A	248	VAL	2.8
1	A	428	LYS	2.8
1	A	429	HIS	2.8
1	A	277	PHE	2.7
1	A	310	ASP	2.7
1	A	259	LEU	2.6
1	A	267	VAL	2.6
1	A	812	HIS	2.6
1	A	337	VAL	2.6
1	A	443	LYS	2.6
1	A	305	CYS	2.5
1	A	315	THR	2.5
1	A	303	LEU	2.5
1	A	318	ARG	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	571	VAL	2.4
1	A	252	GLN	2.4
1	A	1005	GLY	2.4
1	A	445	LEU	2.4
1	A	425	LYS	2.3
1	A	264	TYR	2.3
1	A	497	PHE	2.3
1	A	253	ARG	2.2
1	A	219	LEU	2.2
1	A	430	ALA	2.2
1	A	280	LEU	2.2
1	A	299	TYR	2.2
1	A	335	THR	2.2
1	A	627	TYR	2.2
1	A	296	LEU	2.2
1	A	606	LEU	2.1
1	A	587	THR	2.1
1	A	773	HIS	2.0
1	A	241	LEU	2.0
1	A	240	PRO	2.0
1	A	340	SER	2.0
1	A	671[A]	MET	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [\(i\)](#)

There are no non-standard protein/DNA/RNA residues in this entry.

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, 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
4	CL	A	3	1/1	0.51	0.25	122,122,122,122	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	CL	A	4	1/1	0.53	0.20	121,121,121,121	0
4	CL	A	1231	1/1	0.66	0.25	113,113,113,113	0
4	CL	A	5	1/1	0.77	0.15	97,97,97,97	0
3	GOL	A	2	6/6	0.79	0.25	107,108,109,109	0
4	CL	A	1232	1/1	0.86	0.15	102,102,102,102	0
3	GOL	A	1	6/6	0.88	0.24	85,86,86,87	0
5	SO4	A	1233	5/5	0.99	0.17	73,77,79,80	0

6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.