



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 14, 2026 – 11:24 AM UTC

PDB ID : 7Q8I / pdb_00007q8i
Title : Peptide AVAEKQ in complex with human cathepsin V C25S mutant
Authors : Loboda, J.; Sosnowski, P.; Tusar, L.; Vidmar, R.; Vizovisek, M.; Horvat, J.;
Kosec, G.; Impens, F.; Demol, H.; Turk, B.; Gevaert, K.; Turk, D.
Deposited on : 2021-11-11
Resolution : 1.59 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

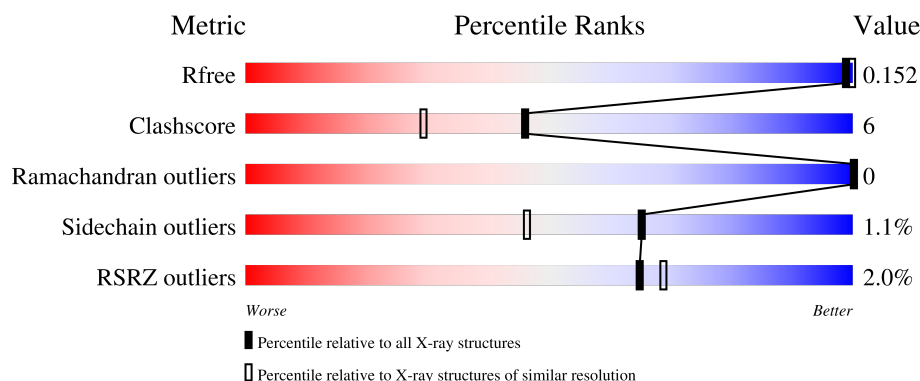
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.59 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}	180053	4673 (1.60-1.60)
Clashscore	190562	4931 (1.60-1.60)
Ramachandran outliers	187476	4831 (1.60-1.60)
Sidechain outliers	187428	4830 (1.60-1.60)
RSRZ outliers	180081	4672 (1.60-1.60)

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	AA	222	 92% 8%
1	BA	222	 4% 89% 9% .
2	PA	6	 17% 83% 17%
2	PB	6	 83% 17%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8077 atoms, of which 4149 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 Cathepsin L2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AA	222	Total	C	H	N	O	S	1635	0	0
			3325	1069	1627	291	328	10			
1	BA	221	Total	C	H	N	O	S	1639	3	0
			3351	1077	1638	294	331	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AA	26	SER	CYS	engineered mutation	UNP O60911
AA	109	GLN	ASN	engineered mutation	UNP O60911
AA	180	GLN	ASN	engineered mutation	UNP O60911
BA	247	SER	CYS	engineered mutation	UNP O60911
BA	330	GLN	ASN	engineered mutation	UNP O60911
BA	401	GLN	ASN	engineered mutation	UNP O60911

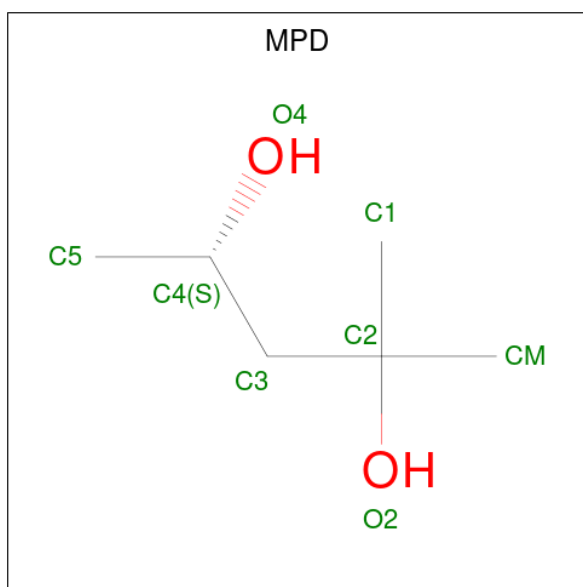
- Molecule 2 is a protein called AVAEKQ peptide.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	PA	5	Total	C	H	N	O		37	0	0
			72	22	37	6	7				
2	PB	6	Total	C	H	N	O		45	0	0
			90	27	45	8	10				

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AA	2	Total	Cl	0	0
			2	2		
3	BA	1	Total	Cl	0	0
			1	1		

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: C₆H₁₄O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AA	1	Total C O 8 6 2	0	0
4	AA	1	Total C O 8 6 2	0	0
4	AA	1	Total C O 8 6 2	0	0
4	BA	1	Total C O 8 6 2	0	0

- Molecule 5 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	BA	1	Total Na 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AA	207	Total H O 621 414 207	414	0
6	BA	184	Total H O 552 368 184	368	0
6	PA	2	Total H O 6 4 2	4	0
6	PB	8	Total H O 24 16 8	16	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: Cathepsin L2

Chain AA:  92% 8%



- Molecule 1: Cathepsin L2

Chain BA:  4% 89% 9%



- Molecule 2: AVAEKQ peptide

Chain PA:  17% 83% 17%



- Molecule 2: AVAEKQ peptide

Chain PB:  83% 17%



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	93.48Å 93.48Å 125.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.74 – 1.59 46.74 – 1.59	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.74-1.59) 100.0 (46.74-1.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.14 (at 1.59Å)	Xtriage
Refinement program	MAIN	Depositor
R, R_{free}	0.169 , 0.189 (Not available) , 0.152	Depositor DCC
R_{free} test set	2100 reflections (2.78%)	wwPDB-VP
Wilson B-factor (Å ²)	15.8	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 45.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8077	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NA, CL, MPD

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	AA	1.14	1/1740 (0.1%)	1.20	2/2351 (0.1%)
1	BA	1.08	1/1755 (0.1%)	1.25	5/2371 (0.2%)
2	PA	1.16	0/34	0.85	0/44
2	PB	0.85	0/44	0.91	0/56
All	All	1.10	2/3573 (0.1%)	1.22	7/4822 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AA	150	SER	CA-C	-5.76	1.45	1.52
1	BA	402	ASN	CA-C	-5.12	1.46	1.52

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	BA	271	SER	N-CA-C	6.55	119.50	108.23
1	BA	401	GLN	CA-C-N	-6.11	112.99	122.59
1	BA	401	GLN	C-N-CA	-6.11	112.99	122.59
1	AA	50	SER	N-CA-C	5.74	117.65	108.41
1	AA	110	ASP	CA-CB-CG	5.64	118.25	112.60
1	BA	419	ASN	CA-CB-CG	-5.51	107.09	112.60
1	BA	312	PRO	N-CA-C	5.07	119.44	111.38

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	AA	1698	1627	1626	10	0
1	BA	1713	1638	1633	23	0
2	PA	35	37	37	0	0
2	PB	45	45	45	1	0
3	AA	2	0	0	0	0
3	BA	1	0	0	1	0
4	AA	24	0	42	5	0
4	BA	8	0	14	3	0
5	BA	1	0	0	0	0
6	AA	207	414	0	4	1
6	BA	184	368	0	4	1
6	PA	2	4	0	0	0
6	PB	8	16	0	0	0
All	All	3928	4149	3397	40	1

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 6.

All (40) 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:AA:304:MPD:HM2	4:AA:304:MPD:H52	1.33	1.07
1:BA:290:GLY:O	4:BA:503:MPD:H13	1.76	0.84
4:AA:304:MPD:HM2	4:AA:304:MPD:C5	2.14	0.76
1:BA:317[B]:ASP:OD1	6:BA:601:HOH:O	2.02	0.76
4:AA:304:MPD:H52	4:AA:304:MPD:CM	2.16	0.74
1:AA:69:GLY:O	4:AA:302:MPD:H13	1.88	0.72
3:BA:501:CL:CL	6:BA:603:HOH:O	2.51	0.66
1:BA:223:LEU:HD23	1:BA:223:LEU:N	2.14	0.61
1:BA:401:GLN:HE22	1:BA:404:LYS:NZ	1.99	0.61
1:BA:243[A]:GLN:CD	1:BA:243[A]:GLN:H	2.08	0.61
1:BA:386:HIS:HD1	2:PB:449:ALA:N	2.00	0.59
4:AA:304:MPD:C5	4:AA:304:MPD:CM	2.78	0.58
1:BA:223:LEU:HB2	1:BA:224:PRO:HD2	1.87	0.57
1:BA:243[B]:GLN:H	1:BA:243[B]:GLN:CD	2.12	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AA:16:PRO:HD3	6:AA:451:HOH:O	2.05	0.55
1:BA:274:ASN:ND2	1:BA:306:ASP:H	2.05	0.54
1:AA:8:TRP:CE2	1:AA:132:GLY:HA2	2.44	0.53
4:BA:503:MPD:H11	4:BA:503:MPD:H53	1.91	0.53
1:BA:401:GLN:HE22	1:BA:404:LYS:HZ1	1.57	0.52
1:BA:223:LEU:N	1:BA:223:LEU:CD2	2.74	0.50
1:BA:243[A]:GLN:CD	1:BA:243[A]:GLN:N	2.71	0.49
1:BA:402:ASN:O	1:BA:403:SER:HB2	2.13	0.49
1:AA:116:VAL:O	1:AA:118:PRO:HD3	2.13	0.48
1:BA:401:GLN:NE2	1:BA:404:LYS:NZ	2.61	0.48
1:BA:243[B]:GLN:HG2	1:BA:243[B]:GLN:O	2.14	0.48
1:AA:207:LYS:HE3	6:AA:462:HOH:O	2.13	0.47
1:BA:236:THR:HB	1:BA:237:PRO:HD2	1.98	0.46
1:BA:343:LYS:HD3	6:BA:625:HOH:O	2.15	0.46
1:AA:9:ARG:HD3	1:AA:200:TYR:CE2	2.52	0.45
1:BA:337:VAL:HB	1:BA:438:ALA:HB3	1.99	0.45
1:BA:229:TRP:CE2	1:BA:353:GLY:HA2	2.51	0.44
1:BA:274:ASN:HD22	1:BA:306:ASP:H	1.64	0.44
1:AA:120:LYS:HD2	6:AA:401:HOH:O	2.17	0.44
1:BA:401:GLN:NE2	1:BA:404:LYS:HZ1	2.15	0.44
4:BA:503:MPD:HM3	4:BA:503:MPD:C5	2.47	0.44
1:AA:59:ARG:N	1:AA:60:PRO:HD2	2.34	0.42
1:AA:175:GLU:HG3	1:AA:182:SER:HB2	2.02	0.41
1:BA:237:PRO:HD3	6:BA:681:HOH:O	2.20	0.41
1:AA:120:LYS:NZ	6:AA:401:HOH:O	2.32	0.40
1:BA:414:PRO:O	1:BA:419:ASN:HA	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:AA:533:HOH:O	6:BA:675:HOH:O[3_554]	2.06	0.14

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	AA	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
1	BA	222/222 (100%)	216 (97%)	6 (3%)	0	100	100
2	PA	3/6 (50%)	3 (100%)	0	0	100	100
2	PB	4/6 (67%)	4 (100%)	0	0	100	100
All	All	449/456 (98%)	439 (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	AA	180/180 (100%)	178 (99%)	2 (1%)	65	46
1	BA	182/180 (101%)	180 (99%)	2 (1%)	65	46
2	PA	3/4 (75%)	3 (100%)	0	100	100
2	PB	4/4 (100%)	4 (100%)	0	100	100
All	All	369/368 (100%)	365 (99%)	4 (1%)	65	46

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AA	158	CYS
1	AA	175	GLU
1	BA	223	LEU
1	BA	379	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	AA	67	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	AA	221	ASN
1	BA	274	ASN
1	BA	401	GLN
1	BA	419	ASN
1	BA	442	ASN
2	PB	454	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 4 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	$\# Z > 2$
4	MPD	AA	304	-	7,7,7	0.67	0	9,10,10	0.63	0
4	MPD	AA	302	-	7,7,7	0.32	0	9,10,10	0.23	0
4	MPD	AA	303	-	7,7,7	0.50	0	9,10,10	0.65	0
4	MPD	BA	503	-	7,7,7	0.75	0	9,10,10	0.88	1 (11%)

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
4	MPD	AA	304	-	-	1/5/5/5	-
4	MPD	AA	302	-	-	1/5/5/5	-
4	MPD	AA	303	-	-	3/5/5/5	-
4	MPD	BA	503	-	-	2/5/5/5	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	BA	503	MPD	C5-C4-C3	2.00	120.97	111.67

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	BA	503	MPD	C2-C3-C4-C5
4	AA	304	MPD	O2-C2-C3-C4
4	BA	503	MPD	O2-C2-C3-C4
4	AA	303	MPD	CM-C2-C3-C4
4	AA	303	MPD	O2-C2-C3-C4
4	AA	302	MPD	C1-C2-C3-C4
4	AA	303	MPD	C1-C2-C3-C4

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	AA	304	MPD	4	0
4	AA	302	MPD	1	0
4	BA	503	MPD	3	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	AA	222/222 (100%)	-0.28	0 100 100	5, 9, 15, 27	4 (1%)
1	BA	221/222 (99%)	0.09	8 (3%) 46 48	5, 10, 21, 45	4 (1%)
2	PA	5/6 (83%)	1.23	1 (20%) 3 2	9, 10, 33, 35	0
2	PB	6/6 (100%)	-0.03	0 100 100	7, 9, 13, 14	0
All	All	454/456 (99%)	-0.08	9 (1%) 65 69	5, 10, 18, 45	8 (1%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BA	223	LEU	6.0
1	BA	398	ALA	4.5
1	BA	243[A]	GLN	3.8
2	PA	448	LYS	3.0
1	BA	397	GLY	2.5
1	BA	399	ASN	2.5
1	BA	401	GLN	2.4
1	BA	343	LYS	2.3
1	BA	402	ASN	2.2

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
5	NA	BA	502	1/1	0.81	0.12	69,69,69,69	0
4	MPD	AA	304	8/8	0.82	0.27	62,64,67,68	0
4	MPD	AA	303	8/8	0.84	0.19	43,47,62,68	0
4	MPD	AA	302	8/8	0.85	0.18	35,40,46,46	0
4	MPD	BA	503	8/8	0.86	0.23	40,48,57,60	0
3	CL	BA	501	1/1	0.97	0.07	21,21,21,21	0
3	CL	AA	301	1/1	0.98	0.05	18,18,18,18	0
3	CL	AA	305	1/1	0.99	0.10	24,24,24,24	0

6.5 Other polymers [i](#)

There are no such residues in this entry.