



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 17, 2026 – 07:23 PM UTC

PDB ID : 5NGJ / pdb_00005ngj
Title : Crystal structure of pb6, major tail tube protein of bacteriophage T5
Authors : Arnaud, C.-A.; Effantin, G.; Vives, C.; Engilberge, S.; Bacia, M.; Boulanger, P.; Girard, E.; Schoehn, G.; Breyton, C.
Deposited on : 2017-03-17
Resolution : 2.20 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 ⓘ 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
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

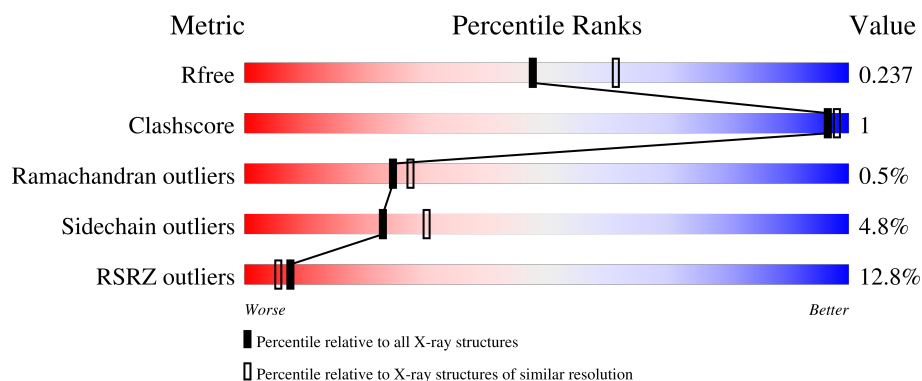
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	180053	6164 (2.20-2.20)
Clashscore	190562	6851 (2.20-2.20)
Ramachandran outliers	187476	6768 (2.20-2.20)
Sidechain outliers	187428	6769 (2.20-2.20)
RSRZ outliers	180081	6166 (2.20-2.20)

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	477	
1	B	477	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13513 atoms, of which 6596 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 Tail tube protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	440	Total	C	H	N	O	S	0	1	0
			6729	2128	3347	563	682	9			
1	B	433	Total	C	H	N	O	S	0	2	1
			6551	2073	3249	552	668	9			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	465	GLU	-	expression tag	UNP Q6QGE2
A	466	ASN	-	expression tag	UNP Q6QGE2
A	467	LEU	-	expression tag	UNP Q6QGE2
A	468	TYR	-	expression tag	UNP Q6QGE2
A	469	PHE	-	expression tag	UNP Q6QGE2
A	470	GLN	-	expression tag	UNP Q6QGE2
A	471	GLY	-	expression tag	UNP Q6QGE2
A	472	HIS	-	expression tag	UNP Q6QGE2
A	473	HIS	-	expression tag	UNP Q6QGE2
A	474	HIS	-	expression tag	UNP Q6QGE2
A	475	HIS	-	expression tag	UNP Q6QGE2
A	476	HIS	-	expression tag	UNP Q6QGE2
A	477	HIS	-	expression tag	UNP Q6QGE2
B	465	GLU	-	expression tag	UNP Q6QGE2
B	466	ASN	-	expression tag	UNP Q6QGE2
B	467	LEU	-	expression tag	UNP Q6QGE2
B	468	TYR	-	expression tag	UNP Q6QGE2
B	469	PHE	-	expression tag	UNP Q6QGE2
B	470	GLN	-	expression tag	UNP Q6QGE2
B	471	GLY	-	expression tag	UNP Q6QGE2
B	472	HIS	-	expression tag	UNP Q6QGE2
B	473	HIS	-	expression tag	UNP Q6QGE2
B	474	HIS	-	expression tag	UNP Q6QGE2
B	475	HIS	-	expression tag	UNP Q6QGE2
B	476	HIS	-	expression tag	UNP Q6QGE2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	477	HIS	-	expression tag	UNP Q6QGE2

- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

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

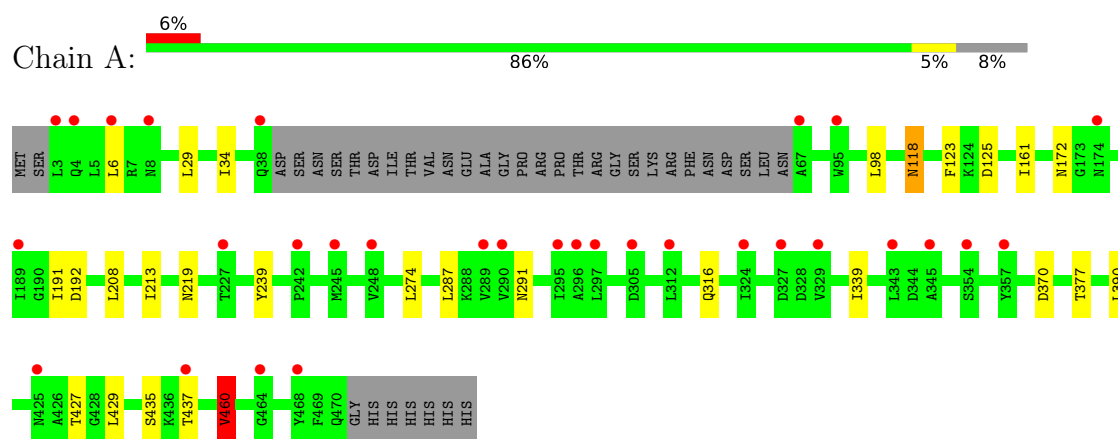
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	145	Total O 145 145	0	0
3	B	83	Total O 83 83	0	0

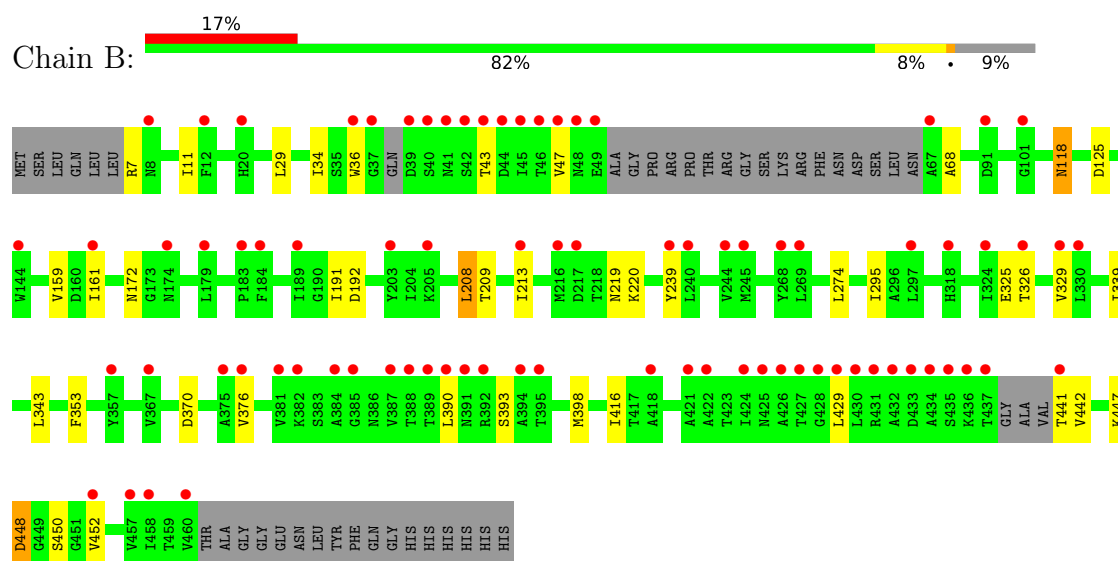
3 Residue-property plots [i](#)

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: Tail tube protein



• Molecule 1: Tail tube protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	76.79Å 115.08Å 83.37Å 90.00° 111.88° 90.00°	Depositor
Resolution (Å)	44.00 – 2.20 44.00 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.00-2.20) 99.7 (44.00-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.73 (at 2.20Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.205 , 0.235 0.207 , 0.237	Depositor DCC
R_{free} test set	3400 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	56.8	Xtriage
Anisotropy	0.495	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 76.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	13513	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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: CL

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	0.88	0/3442	1.09	7/4685 (0.1%)
1	B	0.80	0/3362	1.07	11/4575 (0.2%)
All	All	0.84	0/6804	1.08	18/9260 (0.2%)

There are no bond length outliers.

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	208	LEU	CD1-CG-CD2	7.46	127.21	110.80
1	B	118	ASN	CA-CB-CG	6.85	119.45	112.60
1	A	118	ASN	CA-CB-CG	6.74	119.33	112.60
1	A	208	LEU	CD1-CG-CD2	6.72	125.59	110.80
1	A	370	ASP	CA-CB-CG	5.56	118.16	112.60
1	A	460	VAL	N-CA-CB	5.44	116.64	110.72
1	A	123	PHE	CA-CB-CG	-5.43	108.36	113.80
1	A	291	ASN	CA-CB-CG	5.43	118.03	112.60
1	A	192	ASP	CA-CB-CG	5.39	117.99	112.60
1	B	43	THR	N-CA-C	5.27	117.93	110.50
1	B	447	LYS	CA-C-N	5.27	131.60	121.54
1	B	447	LYS	C-N-CA	5.27	131.60	121.54
1	B	370	ASP	CA-CB-CG	5.25	117.85	112.60
1	B	448	ASP	CA-CB-CG	5.24	117.84	112.60
1	B	325	GLU	CA-C-N	5.06	127.37	120.54
1	B	325	GLU	C-N-CA	5.06	127.37	120.54
1	B	192	ASP	CA-CB-CG	5.05	117.65	112.60
1	B	353	PHE	CA-CB-CG	5.05	118.85	113.80

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	3382	3347	3347	3	0
1	B	3302	3249	3221	10	0
2	A	5	0	0	0	0
3	A	145	0	0	0	0
3	B	83	0	0	0	0
All	All	6917	6596	6568	13	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 1.

All (13) 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:376:VAL:HG12	1:B:452:VAL:HG11	1.67	0.76
1:B:213:ILE:HG22	1:B:295:ILE:HG22	1.82	0.61
1:B:376:VAL:CG1	1:B:452:VAL:HG11	2.32	0.60
1:B:159:VAL:O	1:B:159:VAL:HG12	2.11	0.51
1:B:450:SER:OG	1:B:452:VAL:HG12	2.12	0.49
1:A:390:LEU:O	1:A:460:VAL:HA	2.13	0.48
1:A:172:ASN:HB2	1:A:239:TYR:CE2	2.50	0.47
1:A:316:GLN:HG2	1:A:339:ILE:O	2.17	0.45
1:B:36:TRP:CZ2	1:B:68:ALA:HB3	2.52	0.45
1:B:172:ASN:HB2	1:B:239:TYR:CE2	2.52	0.44
1:B:159:VAL:HG13	1:B:209:THR:HG21	2.02	0.42
1:B:398:MET:HE3	1:B:442:VAL:HG11	2.00	0.42
1:B:11:ILE:HD11	1:B:34:ILE:HG21	2.01	0.42

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	437/477 (92%)	423 (97%)	14 (3%)	0	100	100
1	B	427/477 (90%)	406 (95%)	17 (4%)	4 (1%)	14	14
All	All	864/954 (91%)	829 (96%)	31 (4%)	4 (0%)	24	27

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	448	ASP
1	B	393	SER
1	B	329	VAL
1	B	47	VAL

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	374/406 (92%)	356 (95%)	18 (5%)	23	30
1	B	362/406 (89%)	345 (95%)	17 (5%)	23	31
All	All	736/812 (91%)	701 (95%)	35 (5%)	23	30

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	LEU
1	A	29	LEU
1	A	34	ILE
1	A	98	LEU
1	A	118	ASN
1	A	125	ASP
1	A	161	ILE
1	A	191	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	213	ILE
1	A	219	ASN
1	A	274	LEU
1	A	287	LEU
1	A	377	THR
1	A	427	THR
1	A	429	LEU
1	A	435	SER
1	A	437	THR
1	A	460	VAL
1	B	7	ARG
1	B	29	LEU
1	B	118	ASN
1	B	125	ASP
1	B	161	ILE
1	B	191	ILE
1	B	208	LEU
1	B	219	ASN
1	B	220	LYS
1	B	274	LEU
1	B	326	THR
1	B	339	ILE
1	B	343	LEU
1	B	390	LEU
1	B	416	ILE
1	B	429	LEU
1	B	441	THR

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	154	GLN
1	A	466	ASN
1	B	154	GLN

5.3.3 RNA ⓘ

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	440/477 (92%)	0.53	31 (7%)	22 19	38, 70, 113, 149	1 (0%)
1	B	433/477 (90%)	1.15	81 (18%)	3 2	40, 85, 179, 285	1 (0%)
All	All	873/954 (91%)	0.84	112 (12%)	7 5	38, 76, 142, 285	2 (0%)

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	37	GLY	6.5
1	B	437	THR	6.4
1	B	42	SER	5.6
1	B	432	ALA	5.5
1	B	47	VAL	5.3
1	B	46	THR	5.3
1	B	430	LEU	5.1
1	B	389	THR	4.8
1	B	330	LEU	4.8
1	B	329	VAL	4.8
1	B	428	GLY	4.6
1	B	375	ALA	4.5
1	B	429	LEU	4.2
1	B	40	SER	4.2
1	B	43	THR	4.2
1	B	39	ASP	4.1
1	A	4	GLN	4.1
1	B	189	ILE	4.1
1	B	424	ILE	4.0
1	B	183	PRO	4.0
1	B	460	VAL	3.9
1	B	435	SER	3.9
1	B	318[A]	HIS	3.9
1	B	434	ALA	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	36	TRP	3.7
1	A	3	LEU	3.7
1	B	388	THR	3.7
1	B	41	ASN	3.6
1	B	245	MET	3.5
1	B	161	ILE	3.5
1	B	91	ASP	3.5
1	B	427	THR	3.4
1	B	457	VAL	3.4
1	B	12	PHE	3.3
1	B	390	LEU	3.3
1	A	329	VAL	3.3
1	B	101	GLY	3.3
1	B	452	VAL	3.3
1	B	458	ILE	3.2
1	B	268	TYR	3.1
1	B	436	LYS	3.1
1	B	203	TYR	3.1
1	B	44	ASP	3.0
1	A	95	TRP	3.0
1	B	426	ALA	2.9
1	B	425	ASN	2.9
1	A	327	ASP	2.9
1	A	357	TYR	2.9
1	B	392	ARG	2.9
1	A	312	LEU	2.9
1	A	38	GLN	2.8
1	B	376	VAL	2.8
1	B	431	ARG	2.8
1	A	8	ASN	2.8
1	B	422	ALA	2.8
1	B	216	MET	2.8
1	B	385	GLY	2.7
1	B	205	LYS	2.7
1	B	394	ALA	2.7
1	B	49	GLU	2.7
1	B	395	THR	2.6
1	B	144	TRP	2.6
1	A	324	ILE	2.6
1	B	384	ALA	2.6
1	B	217	ASP	2.6
1	B	67	ALA	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	357	TYR	2.5
1	A	189	ILE	2.5
1	B	367	VAL	2.5
1	B	213	ILE	2.4
1	B	244	VAL	2.4
1	A	305	ASP	2.4
1	A	242	PRO	2.4
1	A	354	SER	2.4
1	B	297	LEU	2.4
1	A	67	ALA	2.4
1	A	290	VAL	2.3
1	A	295	ILE	2.3
1	A	6	LEU	2.3
1	B	179	LEU	2.3
1	A	174	ASN	2.3
1	B	324	ILE	2.3
1	A	248	VAL	2.3
1	A	468	TYR	2.3
1	B	326	THR	2.3
1	B	387	VAL	2.3
1	A	464	GLY	2.2
1	B	418	ALA	2.2
1	B	382	LYS	2.2
1	B	184	PHE	2.2
1	B	240	LEU	2.2
1	B	391	ASN	2.2
1	A	296	ALA	2.2
1	A	345	ALA	2.2
1	A	437	THR	2.2
1	A	343	LEU	2.1
1	B	433	ASP	2.1
1	B	441	THR	2.1
1	B	239	TYR	2.1
1	B	45	ILE	2.1
1	B	269	LEU	2.1
1	B	421	ALA	2.1
1	A	297	LEU	2.1
1	B	48	ASN	2.1
1	B	381	VAL	2.1
1	A	245	MET	2.0
1	B	8	ASN	2.0
1	A	289	VAL	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	227	THR	2.0
1	B	20	HIS	2.0
1	A	425	ASN	2.0
1	B	174	ASN	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 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
2	CL	A	503	1/1	0.87	0.16	107,107,107,107	0
2	CL	A	504	1/1	0.91	0.23	92,92,92,92	0
2	CL	A	501	1/1	0.93	0.17	56,56,56,56	1
2	CL	A	505	1/1	0.96	0.33	91,91,91,91	1
2	CL	A	502	1/1	0.98	0.13	50,50,50,50	1

6.5 Other polymers [i](#)

There are no such residues in this entry.