



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 03:18 PM UTC

PDB ID : 5G5T / pdb_00005g5t
Title : Structure of the Argonaute protein from *Methanocaldococcus janaschii* in complex with guide DNA
Authors : Schneider, S.; Oellig, C.A.; Keegan, R.; Grohmann, D.; Zander, A.; Willkomm, S.
Deposited on : 2016-06-03
Resolution : 2.85 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

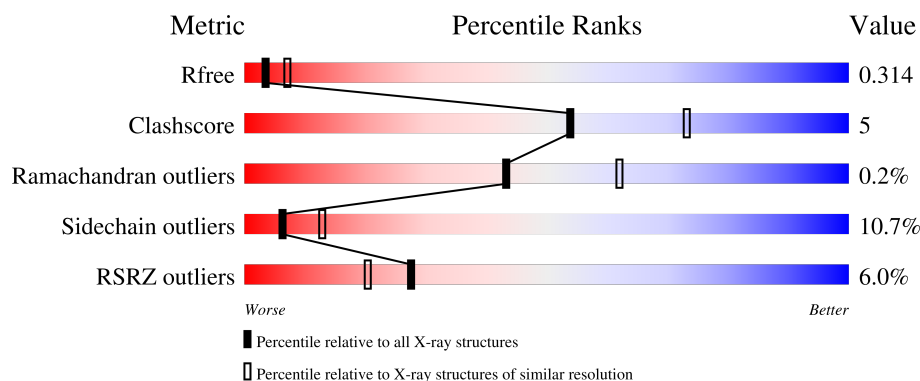
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.85 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	1407 (2.88-2.84)
Clashscore	190562	1446 (2.88-2.84)
Ramachandran outliers	187476	1406 (2.88-2.84)
Sidechain outliers	187428	1407 (2.88-2.84)
RSRZ outliers	180081	1408 (2.88-2.84)

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	713	6% 74% 19% 6%
2	D	21	5% 29% 24% 48%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5413 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 ARGONAUTE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	670	Total	C	N	O	S	0	0	0
			5239	3416	839	964	20			

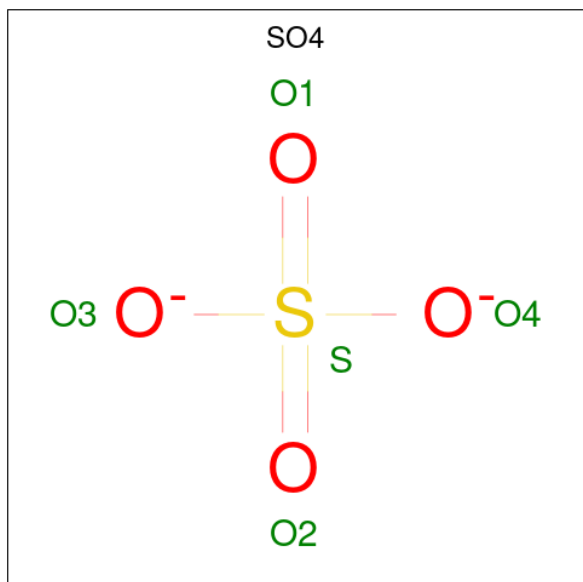
- Molecule 2 is a DNA chain called GUIDE DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	11	Total	C	N	O	P	0	0	0
			138	55	17	55	11			

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is SULFATE ION (CCD ID: SO4) (formula: O₄S).

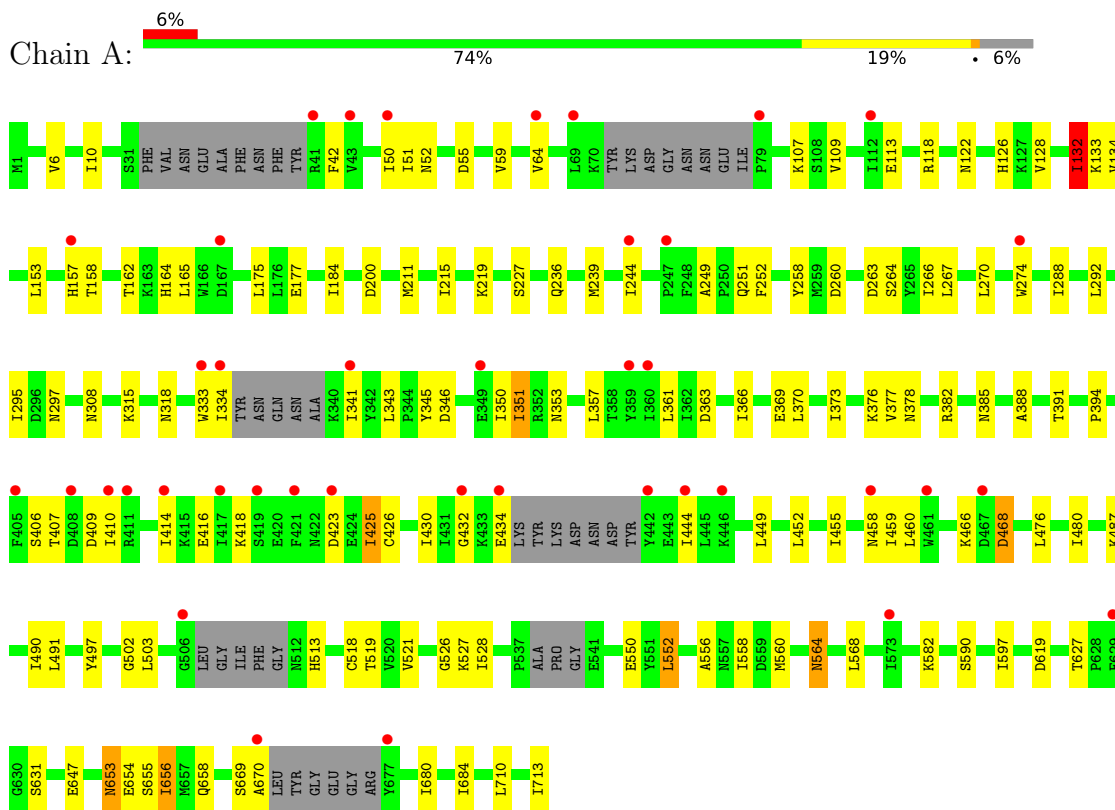


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		

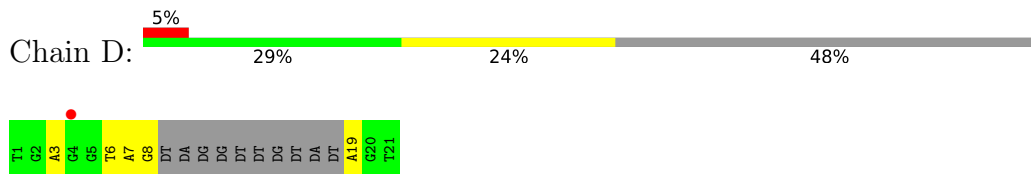
3 Residue-property plots [i](#)

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: ARGONAUTE



• Molecule 2: GUIDE DNA



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	117.95Å 117.95Å 134.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.11 – 2.85 49.11 – 2.85	Depositor EDS
% Data completeness (in resolution range)	99.4 (49.11-2.85) 99.4 (49.11-2.85)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.42 (at 2.86Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.229 , 0.285 0.253 , 0.314	Depositor DCC
R_{free} test set	1147 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	88.6	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 91.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5413	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% 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: SO4, MG

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	0.76	0/5361	1.32	17/7295 (0.2%)
2	D	0.76	1/147 (0.7%)	1.29	1/221 (0.5%)
All	All	0.76	1/5508 (0.0%)	1.32	18/7516 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	7	DA	O3'-P	-8.41	1.48	1.61

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	654	GLU	N-CA-C	-7.46	103.13	112.90
1	A	297	ASN	N-CA-C	5.73	119.53	112.54
1	A	264	SER	CA-C-N	5.68	127.89	120.28
1	A	264	SER	C-N-CA	5.68	127.89	120.28
1	A	50	ILE	N-CA-C	5.48	116.12	108.89
1	A	619	ASP	CA-CB-CG	5.48	118.08	112.60
1	A	52	ASN	N-CA-C	5.47	118.22	110.50
1	A	200	ASP	CA-CB-CG	5.42	118.02	112.60
1	A	654	GLU	CA-C-N	5.30	127.65	120.65
1	A	654	GLU	C-N-CA	5.30	127.65	120.65
1	A	132	ILE	N-CA-C	5.23	116.84	108.89
1	A	631	SER	N-CA-C	5.21	116.83	110.41
1	A	64	VAL	CA-C-N	5.12	127.45	120.54
1	A	64	VAL	C-N-CA	5.12	127.45	120.54
2	D	19	DA	N9-C1'-C2'	5.10	121.14	113.50
1	A	653	ASN	CA-CB-CG	5.06	117.66	112.60
1	A	200	ASP	N-CA-C	5.03	116.33	109.18
1	A	55	ASP	N-CA-C	-5.01	101.00	109.07

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	5239	0	4912	51	0
2	D	138	0	70	3	0
3	A	1	0	0	0	0
4	A	35	0	0	0	0
All	All	5413	0	4982	51	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 5.

All (51) 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:406:SER:HB2	1:A:409:ASP:CB	1.86	1.04
1:A:406:SER:CB	1:A:409:ASP:HB2	1.89	1.00
1:A:357:LEU:HB3	1:A:426:CYS:SG	2.14	0.88
1:A:406:SER:HB2	1:A:409:ASP:HB2	0.95	0.87
1:A:418:LYS:HA	1:A:452:LEU:HD11	1.65	0.78
1:A:468:ASP:OD1	1:A:468:ASP:N	2.22	0.71
1:A:249:ALA:HB3	1:A:252:PHE:HD2	1.61	0.65
1:A:351:ILE:CD1	1:A:425:ILE:HD13	2.30	0.62
1:A:134:VAL:HG11	1:A:274:TRP:HE1	1.64	0.61
1:A:363:ASP:HB2	1:A:432:GLY:HA2	1.83	0.61
1:A:132:ILE:HD11	1:A:270:LEU:HD11	1.84	0.59
1:A:308:ASN:H	1:A:658:GLN:HE22	1.55	0.54
1:A:653:ASN:HB2	1:A:656:ILE:HG13	1.90	0.52
1:A:680:ILE:HD12	1:A:684:ILE:HG23	1.91	0.52
1:A:434:GLU:HG3	1:A:460:LEU:HD13	1.91	0.52
1:A:249:ALA:HB3	1:A:252:PHE:CD2	2.44	0.52
1:A:361:LEU:HD12	1:A:430:ILE:HG12	1.91	0.52
1:A:552:LEU:HA	1:A:556:ALA:HB3	1.92	0.52
1:A:351:ILE:HD12	1:A:425:ILE:HD13	1.91	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:165:LEU:HD23	1:A:239:MET:HE1	1.93	0.49
1:A:164:HIS:ND1	1:A:251:GLN:HA	2.28	0.49
1:A:491:LEU:HD12	1:A:526:GLY:HA3	1.95	0.49
1:A:669:SER:HB3	1:A:713:ILE:HD12	1.94	0.49
1:A:363:ASP:OD2	1:A:407:THR:CG2	2.60	0.49
1:A:133:LYS:HE3	1:A:162:THR:HG23	1.96	0.47
1:A:627:THR:HG21	2:D:6:DT:H5'	1.97	0.46
1:A:263:ASP:HB3	1:A:266:ILE:HD12	1.96	0.46
1:A:343:LEU:HB3	1:A:391:THR:HG22	1.97	0.46
1:A:497:TYR:HA	1:A:564:ASN:HB3	1.99	0.45
1:A:351:ILE:HG12	1:A:394:PRO:HG3	1.98	0.44
1:A:385:ASN:HA	1:A:388:ALA:HB3	1.99	0.44
1:A:333:TRP:CD1	1:A:334:ILE:H	2.36	0.44
1:A:653:ASN:HB3	1:A:655:SER:H	1.83	0.44
1:A:346:ASP:HB3	1:A:487:LYS:HB2	1.99	0.44
1:A:710:LEU:O	1:A:713:ILE:HG12	2.17	0.44
1:A:157:HIS:HB2	2:D:8:DG:O5'	2.18	0.44
1:A:315:LYS:HD2	1:A:345:TYR:HA	2.00	0.43
1:A:126:HIS:HD2	1:A:128:VAL:HG23	1.83	0.42
1:A:568:LEU:HB3	1:A:597:ILE:HD12	2.01	0.42
1:A:466:LYS:O	1:A:468:ASP:OD1	2.38	0.42
1:A:656:ILE:HG13	1:A:656:ILE:H	1.72	0.42
1:A:292:LEU:HD23	1:A:295:ILE:HG12	2.01	0.41
1:A:582:LYS:HD3	1:A:647:GLU:H	1.85	0.41
1:A:487:LYS:HG2	1:A:527:LYS:HG2	2.01	0.41
1:A:153:LEU:HD21	1:A:288:ILE:HG22	2.03	0.41
1:A:308:ASN:N	1:A:658:GLN:HE22	2.18	0.41
1:A:215:ILE:O	1:A:219:LYS:HA	2.21	0.41
1:A:378:ASN:O	1:A:382:ARG:HB2	2.21	0.40
1:A:670:ALA:HB3	2:D:3:DA:H4'	2.04	0.40
1:A:502:GLY:O	1:A:519:THR:HA	2.22	0.40
1:A:552:LEU:HB3	1:A:560:MET:SD	2.61	0.40

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	654/713 (92%)	608 (93%)	45 (7%)	1 (0%)	43 62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	410	ILE

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	532/662 (80%)	475 (89%)	57 (11%)	6 13

All (57) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	VAL
1	A	10	ILE
1	A	42	PHE
1	A	51	ILE
1	A	59	VAL
1	A	107	LYS
1	A	109	VAL
1	A	113	GLU
1	A	118	ARG
1	A	122	ASN
1	A	132	ILE
1	A	158	THR
1	A	175	LEU
1	A	177	GLU
1	A	184	ILE
1	A	211	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	227	SER
1	A	236	GLN
1	A	244	ILE
1	A	258	TYR
1	A	260	ASP
1	A	267	LEU
1	A	318	ASN
1	A	341	ILE
1	A	350	ILE
1	A	351	ILE
1	A	353	ASN
1	A	366	ILE
1	A	369	GLU
1	A	370	LEU
1	A	373	ILE
1	A	376	LYS
1	A	377	VAL
1	A	414	ILE
1	A	416	GLU
1	A	423	ASP
1	A	425	ILE
1	A	444	ILE
1	A	449	LEU
1	A	455	ILE
1	A	458	ASN
1	A	459	ILE
1	A	468	ASP
1	A	476	LEU
1	A	480	ILE
1	A	490	ILE
1	A	503	LEU
1	A	513	HIS
1	A	518	CYS
1	A	521	VAL
1	A	528	ILE
1	A	550	GLU
1	A	552	LEU
1	A	558	ILE
1	A	564	ASN
1	A	590	SER
1	A	656	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16)

such sidechains are listed below:

Mol	Chain	Res	Type
1	A	4	ASN
1	A	110	HIS
1	A	122	ASN
1	A	126	HIS
1	A	170	ASN
1	A	320	ASN
1	A	355	ASN
1	A	383	ASN
1	A	385	ASN
1	A	513	HIS
1	A	544	HIS
1	A	632	ASN
1	A	643	ASN
1	A	653	ASN
1	A	658	GLN
1	A	685	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 1 is monoatomic - leaving 7 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	SO4	A	1717	-	4,4,4	0.26	0	6,6,6	0.15	0
4	SO4	A	1716	-	4,4,4	0.28	0	6,6,6	0.14	0
4	SO4	A	1721	-	4,4,4	0.27	0	6,6,6	0.16	0
4	SO4	A	1718	-	4,4,4	0.26	0	6,6,6	0.10	0
4	SO4	A	1720	-	4,4,4	0.21	0	6,6,6	0.13	0
4	SO4	A	1719	-	4,4,4	0.24	0	6,6,6	0.20	0
4	SO4	A	1715	-	4,4,4	0.21	0	6,6,6	0.24	0

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	670/713 (93%)	0.55	40 (5%) 27 20	60, 94, 134, 171	0
2	D	11/21 (52%)	1.32	1 (9%) 15 11	95, 132, 144, 151	0
All	All	681/734 (92%)	0.57	41 (6%) 27 20	60, 94, 136, 171	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	50	ILE	4.4
1	A	458	ASN	4.2
1	A	414	ILE	3.9
1	A	467	ASP	3.8
1	A	677	TYR	3.8
1	A	408	ASP	3.7
1	A	334	ILE	3.5
1	A	670	ALA	3.5
1	A	43	VAL	3.3
1	A	461	TRP	3.2
1	A	112	ILE	3.2
1	A	629	PHE	3.1
1	A	333	TRP	2.8
1	A	442	TYR	2.8
1	A	274	TRP	2.7
1	A	410	ILE	2.7
1	A	244	ILE	2.6
1	A	423	ASP	2.6
1	A	64	VAL	2.5
1	A	421	PHE	2.5
1	A	79	PRO	2.5
1	A	69	LEU	2.4
1	A	411	ARG	2.4
1	A	349	GLU	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	360	ILE	2.3
1	A	341	ILE	2.3
1	A	417	ILE	2.3
2	D	4	DG	2.3
1	A	41	ARG	2.3
1	A	419	SER	2.3
1	A	359	TYR	2.3
1	A	157	HIS	2.2
1	A	434	GLU	2.2
1	A	573	ILE	2.1
1	A	247	PRO	2.1
1	A	167	ASP	2.1
1	A	444	ILE	2.1
1	A	506	GLY	2.1
1	A	446	LYS	2.1
1	A	405	PHE	2.0
1	A	432	GLY	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
4	SO4	A	1717	5/5	0.67	0.10	139,139,139,139	0
4	SO4	A	1718	5/5	0.80	0.09	165,165,165,166	0
4	SO4	A	1720	5/5	0.80	0.11	145,146,146,147	0
4	SO4	A	1719	5/5	0.86	0.08	123,123,124,124	0
4	SO4	A	1715	5/5	0.88	0.09	112,113,114,115	0
4	SO4	A	1716	5/5	0.91	0.15	93,95,96,96	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
4	SO4	A	1721	5/5	0.92	0.07	116,117,118,119	0
3	MG	A	1714	1/1	0.95	0.07	74,74,74,74	0

6.5 Other polymers [i](#)

There are no such residues in this entry.