



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 06:46 AM UTC

PDB ID : 7FP1 / pdb_00007fp1
Title : PanDDA analysis group deposition – Aar2/RNaseH in complex with fragment P08F04 from the F2X-Universal Library
Authors : Barthel, T.; Wollenhaupt, J.; Lima, G.M.A.; Wahl, M.C.; Weiss, M.S.
Deposited on : 2022-08-26
Resolution : 1.43 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

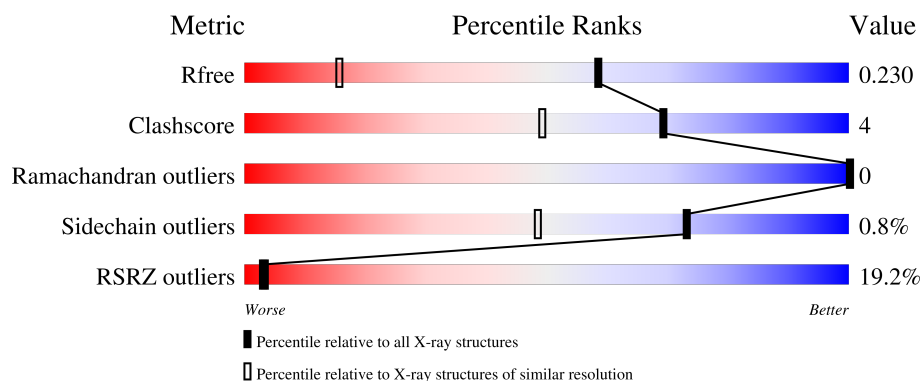
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.43 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	3234 (1.46-1.42)
Clashscore	190562	3289 (1.46-1.42)
Ramachandran outliers	187476	3248 (1.46-1.42)
Sidechain outliers	187428	3248 (1.46-1.42)
RSRZ outliers	180081	3234 (1.46-1.42)

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	258	18% 88% 9% ••
2	B	308	19% 90% 7% •

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9399 atoms, of which 4515 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 Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	253	4188	1363	2060	355	398	12	0	21	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1833	GLY	-	expression tag	UNP P33334
A	1834	ALA	-	expression tag	UNP P33334
A	1835	MET	-	expression tag	UNP P33334

- Molecule 2 is a protein called A1 cistron-splicing factor AAR2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	300	5027	1649	2455	420	484	19	33	16	0

There are 20 discrepancies between the modelled and reference sequences:

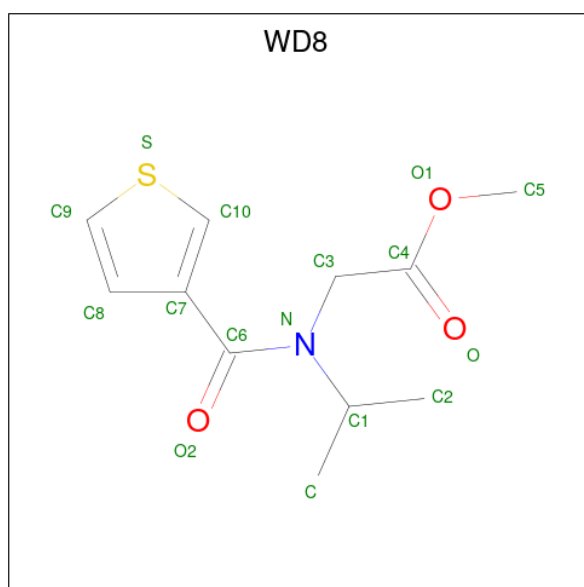
Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP P32357
B	-2	ALA	-	expression tag	UNP P32357
B	-1	MET	-	expression tag	UNP P32357
B	0	ALA	-	expression tag	UNP P32357
B	166	SER	LEU	conflict	UNP P32357
B	167	SER	LYS	conflict	UNP P32357
B	?	-	LEU	deletion	UNP P32357
B	?	-	GLN	deletion	UNP P32357
B	?	-	LYS	deletion	UNP P32357
B	?	-	ALA	deletion	UNP P32357
B	?	-	GLY	deletion	UNP P32357
B	?	-	SER	deletion	UNP P32357
B	?	-	LYS	deletion	UNP P32357

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	MET	deletion	UNP P32357
B	?	-	GLU	deletion	UNP P32357
B	?	-	ALA	deletion	UNP P32357
B	?	-	LYS	deletion	UNP P32357
B	?	-	ASN	deletion	UNP P32357
B	?	-	GLU	deletion	UNP P32357
B	170	SER	ASP	conflict	UNP P32357

- Molecule 3 is methyl N-(propan-2-yl)-N-(thiophene-3-carbonyl)glycinate (CCD ID: WD8) (formula: $C_{11}H_{15}NO_3S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			16	11	1	3	1		
3	A	1	Total	C	N	O	S	0	0
			16	11	1	3	1		

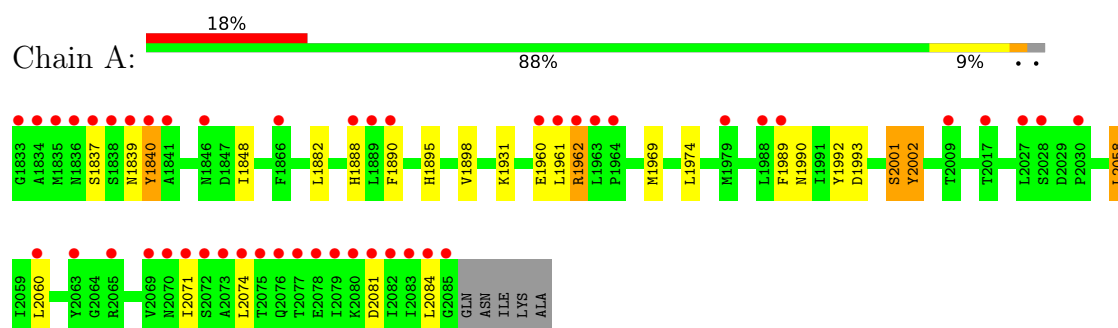
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	80	Total	O	0	0
			80	80		
4	B	72	Total	O	0	0
			72	72		

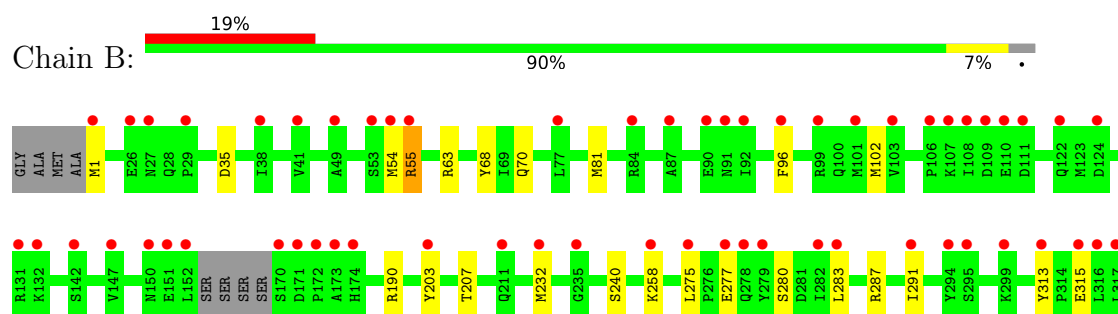
3 Residue-property plots [i](#)

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: Pre-mRNA-splicing factor 8



- Molecule 2: A1 cistron-splicing factor AAR2



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	89.89Å 81.72Å 93.73Å 90.00° 108.91° 90.00°	Depositor
Resolution (Å)	44.67 – 1.43 44.67 – 1.43	Depositor EDS
% Data completeness (in resolution range)	98.1 (44.67-1.43) 98.6 (44.67-1.43)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.96 (at 1.43Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.192 , 0.222 0.221 , 0.230	Depositor DCC
R_{free} test set	2097 reflections (1.80%)	wwPDB-VP
Wilson B-factor (Å ²)	30.8	Xtriage
Anisotropy	0.336	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 38.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9399	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% 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: WD8

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	0.71	8/2269 (0.4%)	0.96	12/3074 (0.4%)
2	B	0.46	3/2731 (0.1%)	0.62	1/3689 (0.0%)
All	All	0.59	11/5000 (0.2%)	0.79	13/6763 (0.2%)

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 (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1961[A]	LEU	N-CA	7.19	1.54	1.46
1	A	1961[B]	LEU	N-CA	7.19	1.54	1.46
1	A	2001[A]	SER	C-N	6.96	1.42	1.33
1	A	2001[B]	SER	C-N	6.96	1.42	1.33
1	A	1960[A]	GLU	C-N	6.56	1.43	1.33
1	A	1960[B]	GLU	C-N	6.56	1.43	1.33
1	A	2002[A]	TYR	N-CA	6.49	1.54	1.46
1	A	2002[B]	TYR	N-CA	6.49	1.54	1.46
2	B	240	SER	C-O	-5.06	1.18	1.24
2	B	55[A]	ARG	N-CA	5.00	1.52	1.46
2	B	55[B]	ARG	N-CA	5.00	1.52	1.46

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1960[A]	GLU	CA-C-N	-9.28	108.36	122.09
1	A	1960[A]	GLU	C-N-CA	-9.28	108.36	122.09
1	A	1960[B]	GLU	CA-C-N	-9.28	108.36	122.09
1	A	1960[B]	GLU	C-N-CA	-9.28	108.36	122.09
1	A	2001[A]	SER	CA-C-N	-8.74	109.08	120.44
1	A	2001[A]	SER	C-N-CA	-8.74	109.08	120.44
1	A	2001[B]	SER	CA-C-N	-8.74	109.08	120.44
1	A	2001[B]	SER	C-N-CA	-8.74	109.08	120.44
1	A	1990	ASN	N-CA-CB	6.33	120.73	110.42
1	A	1890	PHE	CA-C-O	-6.09	114.76	121.58
1	A	1990	ASN	N-CA-C	-6.03	98.90	108.73
1	A	1840	TYR	CB-CA-C	-5.53	101.61	110.79
2	B	54	MET	CA-C-O	-5.23	116.09	122.41

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2001[A]	SER	Mainchain

CLOSE-CONTACTS INFOmissingINFO

5.2 Torsion angles ⓘ

5.2.1 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	274/258 (106%)	267 (97%)	7 (3%)	0	100	100
2	B	314/308 (102%)	306 (98%)	8 (2%)	0	100	100
All	All	588/566 (104%)	573 (97%)	15 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.2.2 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	251/233 (108%)	249 (99%)	2 (1%)	73	48
2	B	293/284 (103%)	290 (99%)	3 (1%)	68	39
All	All	544/517 (105%)	539 (99%)	5 (1%)	73	44

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1962	ARG
1	A	2058	LEU
2	B	55[A]	ARG
2	B	55[B]	ARG
2	B	315	GLU

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

Mol	Chain	Res	Type
2	B	150	ASN
2	B	174	HIS
2	B	206	ASN

5.2.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.4 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.5 Ligand geometry

2 ligands 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
3	WD8	A	2102	-	16,16,16	1.57	2 (12%)	18,21,21	1.11	1 (5%)
3	WD8	A	2101	-	16,16,16	1.66	4 (25%)	18,21,21	1.12	1 (5%)

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
3	WD8	A	2102	-	-	2/18/18/18	0/1/1/1
3	WD8	A	2101	-	-	8/18/18/18	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2101	WD8	C3-C4	-3.54	1.45	1.51
3	A	2101	WD8	C1-N	-3.33	1.44	1.48
3	A	2102	WD8	C7-C6	-3.08	1.45	1.50
3	A	2102	WD8	O1-C5	-2.91	1.38	1.45
3	A	2101	WD8	C7-C6	-2.34	1.46	1.50
3	A	2101	WD8	C3-N	-2.21	1.43	1.45

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2102	WD8	C3-N-C1	-2.51	115.48	118.49
3	A	2101	WD8	O2-C6-C7	2.13	121.95	118.05

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2101	WD8	O2-C6-C7-C10
3	A	2101	WD8	C3-C4-O1-C5
3	A	2102	WD8	O2-C6-C7-C10
3	A	2101	WD8	O-C4-O1-C5
3	A	2101	WD8	N-C3-C4-O1
3	A	2101	WD8	N-C3-C4-O
3	A	2101	WD8	C4-C3-N-C1
3	A	2101	WD8	O2-C6-C7-C8
3	A	2102	WD8	N-C6-C7-C10
3	A	2101	WD8	N-C6-C7-C8

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2102	WD8	1	0

5.6 Other polymers [i](#)

There are no such residues in this entry.

5.7 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	253/258 (98%)	1.29	47 (18%) 3 3	15, 40, 64, 93	12 (4%)
2	B	300/308 (97%)	1.02	59 (19%) 3 3	15, 44, 79, 98	8 (2%)
All	All	553/566 (97%)	1.14	106 (19%) 3 3	15, 42, 73, 98	20 (3%)

All (106) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2017[A]	THR	19.9
2	B	152	LEU	9.0
1	A	2073	ALA	8.0
1	A	2071	ILE	7.4
1	A	2074	LEU	7.4
2	B	170	SER	7.4
1	A	2084	LEU	7.3
1	A	1979[A]	MET	7.2
1	A	2083	ILE	6.5
1	A	2085	GLY	6.3
1	A	2075	THR	6.3
1	A	2072	SER	6.3
1	A	1840	TYR	6.2
1	A	2079	ILE	6.0
1	A	2077	THR	5.9
1	A	2082	ILE	5.8
1	A	1989	PHE	5.7
2	B	173	ALA	5.6
2	B	1	MET	5.4
1	A	2081	ASP	5.3
1	A	2076	GLN	5.0
1	A	1833	GLY	4.9
2	B	172	PRO	4.8
1	A	1834	ALA	4.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	174	HIS	4.7
2	B	279	TYR	4.5
1	A	1961[A]	LEU	4.5
2	B	108	ILE	4.4
2	B	316	LEU	4.3
2	B	203[A]	TYR	4.3
1	A	2070	ASN	4.2
1	A	2078	GLU	4.2
1	A	1835	MET	4.2
2	B	317	LEU	4.0
1	A	2080	LYS	3.9
1	A	1839	ASN	3.7
2	B	275	LEU	3.6
1	A	1960[A]	GLU	3.5
2	B	101	MET	3.5
1	A	1889	LEU	3.5
1	A	1890	PHE	3.4
1	A	1962	ARG	3.4
1	A	1888	HIS	3.4
1	A	2063	TYR	3.3
2	B	41[A]	VAL	3.3
2	B	147	VAL	3.3
2	B	29	PRO	3.3
1	A	1836	ASN	3.2
1	A	2069	VAL	3.2
2	B	96	PHE	3.1
2	B	103	VAL	3.1
1	A	2065	ARG	3.1
2	B	54	MET	3.0
2	B	131	ARG	3.0
2	B	171	ASP	3.0
1	A	1988	LEU	3.0
2	B	282	ILE	2.9
2	B	295	SER	2.9
2	B	38	ILE	2.8
1	A	2027	LEU	2.8
1	A	1837	SER	2.8
1	A	1838	SER	2.8
1	A	2060	LEU	2.8
2	B	55[A]	ARG	2.7
2	B	291	ILE	2.6
2	B	283	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	278	GLN	2.6
2	B	109	ASP	2.6
2	B	87	ALA	2.6
2	B	150	ASN	2.6
2	B	110	GLU	2.5
2	B	77	LEU	2.5
2	B	84	ARG	2.5
1	A	2030	PRO	2.5
2	B	142	SER	2.5
1	A	1964	PRO	2.5
2	B	122[A]	GLN	2.4
2	B	299	LYS	2.4
2	B	235	GLY	2.4
2	B	106	PRO	2.4
2	B	111	ASP	2.4
2	B	26	GLU	2.4
2	B	92	ILE	2.4
1	A	1841	ALA	2.3
2	B	99	ARG	2.3
2	B	315	GLU	2.3
1	A	2009	THR	2.3
2	B	53	SER	2.3
2	B	107	LYS	2.3
2	B	90	GLU	2.2
1	A	2028	SER	2.2
2	B	151	GLU	2.2
2	B	277	GLU	2.2
1	A	1866	PHE	2.2
2	B	132	LYS	2.2
2	B	91	ASN	2.2
1	A	1963	LEU	2.1
2	B	211[A]	GLN	2.1
1	A	1846	ASN	2.1
2	B	124	ASP	2.1
2	B	232	MET	2.1
2	B	49	ALA	2.1
2	B	258	LYS	2.1
2	B	27	ASN	2.1
2	B	294	TYR	2.0
2	B	313	TYR	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
3	WD8	A	2101	16/16	0.61	0.30	20,20,20,20	16
3	WD8	A	2102	16/16	0.73	0.26	20,20,20,20	16

6.5 Other polymers [i](#)

There are no such residues in this entry.