Full wwPDB X-ray Structure Validation Report ⓘ

Dec 6, 2021 – 01:03 pm GMT

PDB ID	:	7O9Q
Title	:	Crystal structure of the Awp1 (adhesin-like wall protein 1) A-domain from Candida glabrata
Authors	:	Reithofer, V.; de Groot, P.; Essen, L.-O.
Deposited on	:	2021-04-16
Resolution	:	1.85 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 following versions of software and data (see references ⓘ) were used in the production of this report:

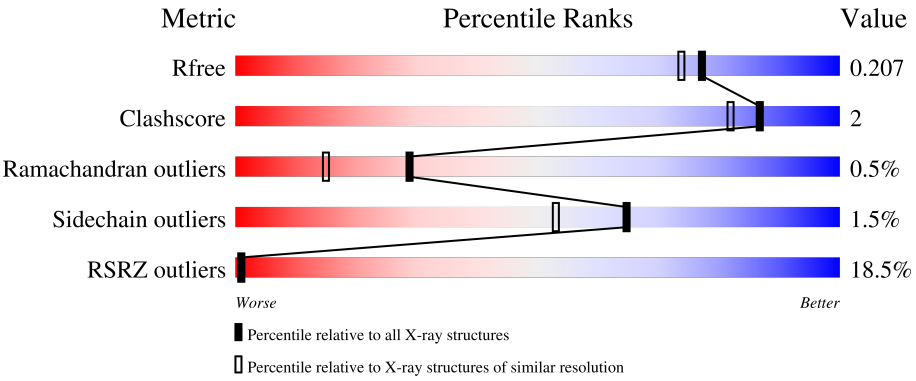
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 <=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	341	25% 86% · 10%
1	B	341	9% 86% · 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9780 atoms, of which 4518 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 Awp1A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	307	Total	C	H	N	O	S	0	8	0
			4539	1438	2246	367	483	5			
1	B	308	Total	C	H	N	O	S	0	2	0
			4496	1423	2224	365	480	4			

There are 66 discrepancies between the modelled and reference sequences:

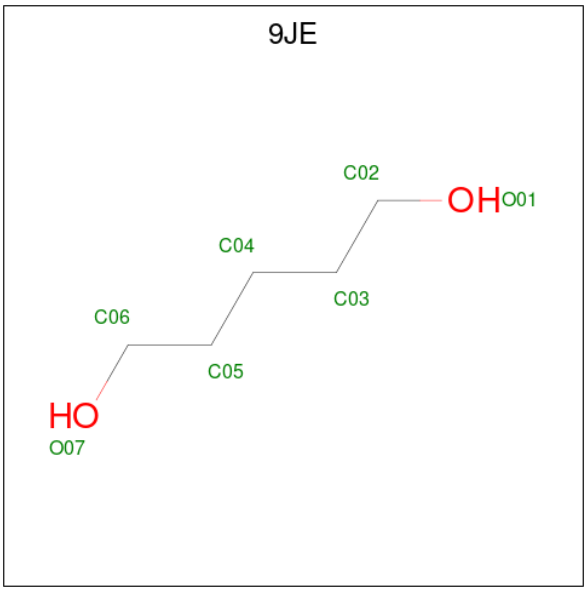
Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	MET	-	initiating methionine	UNP Q6FPN0
A	-14	GLY	-	expression tag	UNP Q6FPN0
A	-13	SER	-	expression tag	UNP Q6FPN0
A	-12	SER	-	expression tag	UNP Q6FPN0
A	-11	HIS	-	expression tag	UNP Q6FPN0
A	-10	HIS	-	expression tag	UNP Q6FPN0
A	-9	HIS	-	expression tag	UNP Q6FPN0
A	-8	HIS	-	expression tag	UNP Q6FPN0
A	-7	HIS	-	expression tag	UNP Q6FPN0
A	-6	HIS	-	expression tag	UNP Q6FPN0
A	-5	SER	-	expression tag	UNP Q6FPN0
A	-4	SER	-	expression tag	UNP Q6FPN0
A	-3	GLY	-	expression tag	UNP Q6FPN0
A	-2	LEU	-	expression tag	UNP Q6FPN0
A	-1	VAL	-	expression tag	UNP Q6FPN0
A	0	PRO	-	expression tag	UNP Q6FPN0
A	1	ARG	-	expression tag	UNP Q6FPN0
A	2	GLY	-	expression tag	UNP Q6FPN0
A	3	SER	-	expression tag	UNP Q6FPN0
A	4	HIS	-	expression tag	UNP Q6FPN0
A	5	MET	-	expression tag	UNP Q6FPN0
A	6	ALA	-	expression tag	UNP Q6FPN0
A	7	SER	-	expression tag	UNP Q6FPN0
A	8	MET	-	expression tag	UNP Q6FPN0
A	9	THR	-	expression tag	UNP Q6FPN0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	GLY	-	expression tag	UNP Q6FPN0
A	11	GLY	-	expression tag	UNP Q6FPN0
A	12	GLN	-	expression tag	UNP Q6FPN0
A	13	GLN	-	expression tag	UNP Q6FPN0
A	14	MET	-	expression tag	UNP Q6FPN0
A	15	GLY	-	expression tag	UNP Q6FPN0
A	16	ARG	-	expression tag	UNP Q6FPN0
A	17	GLY	-	expression tag	UNP Q6FPN0
B	-15	MET	-	initiating methionine	UNP Q6FPN0
B	-14	GLY	-	expression tag	UNP Q6FPN0
B	-13	SER	-	expression tag	UNP Q6FPN0
B	-12	SER	-	expression tag	UNP Q6FPN0
B	-11	HIS	-	expression tag	UNP Q6FPN0
B	-10	HIS	-	expression tag	UNP Q6FPN0
B	-9	HIS	-	expression tag	UNP Q6FPN0
B	-8	HIS	-	expression tag	UNP Q6FPN0
B	-7	HIS	-	expression tag	UNP Q6FPN0
B	-6	HIS	-	expression tag	UNP Q6FPN0
B	-5	SER	-	expression tag	UNP Q6FPN0
B	-4	SER	-	expression tag	UNP Q6FPN0
B	-3	GLY	-	expression tag	UNP Q6FPN0
B	-2	LEU	-	expression tag	UNP Q6FPN0
B	-1	VAL	-	expression tag	UNP Q6FPN0
B	0	PRO	-	expression tag	UNP Q6FPN0
B	1	ARG	-	expression tag	UNP Q6FPN0
B	2	GLY	-	expression tag	UNP Q6FPN0
B	3	SER	-	expression tag	UNP Q6FPN0
B	4	HIS	-	expression tag	UNP Q6FPN0
B	5	MET	-	expression tag	UNP Q6FPN0
B	6	ALA	-	expression tag	UNP Q6FPN0
B	7	SER	-	expression tag	UNP Q6FPN0
B	8	MET	-	expression tag	UNP Q6FPN0
B	9	THR	-	expression tag	UNP Q6FPN0
B	10	GLY	-	expression tag	UNP Q6FPN0
B	11	GLY	-	expression tag	UNP Q6FPN0
B	12	GLN	-	expression tag	UNP Q6FPN0
B	13	GLN	-	expression tag	UNP Q6FPN0
B	14	MET	-	expression tag	UNP Q6FPN0
B	15	GLY	-	expression tag	UNP Q6FPN0
B	16	ARG	-	expression tag	UNP Q6FPN0
B	17	GLY	-	expression tag	UNP Q6FPN0

- Molecule 2 is pentane-1,5-diol (three-letter code: 9JE) (formula: $C_5H_{12}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C H O 19 5 12 2	0	0
2	B	1	Total C H O 19 5 12 2	0	0
2	B	1	Total C H O 19 5 12 2	0	0
2	B	1	Total C H O 19 5 12 2	0	0

- Molecule 3 is water.

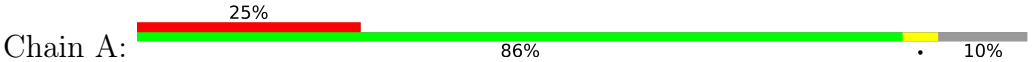
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	274	Total O 274 274	0	0
3	B	394	Total O 395 395	0	1

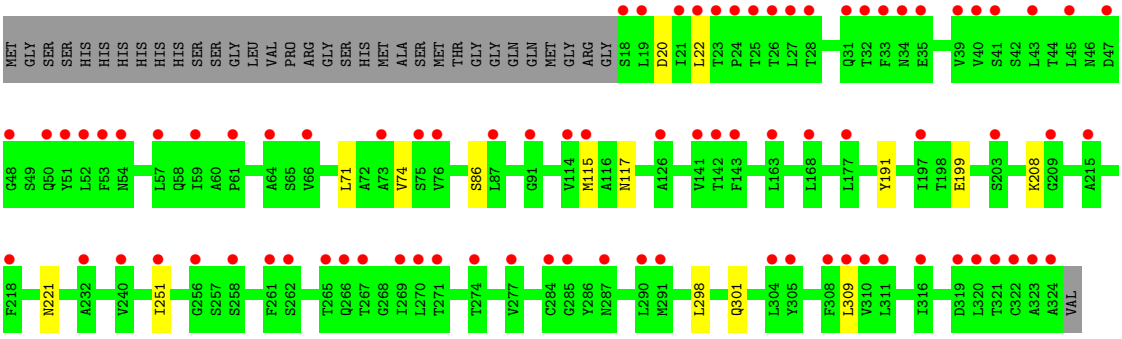
3 Residue-property plots (i)

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: Awp1A

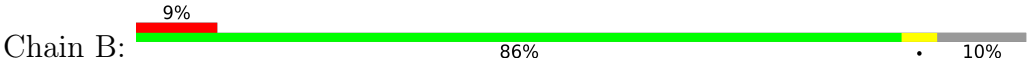
Chain A:

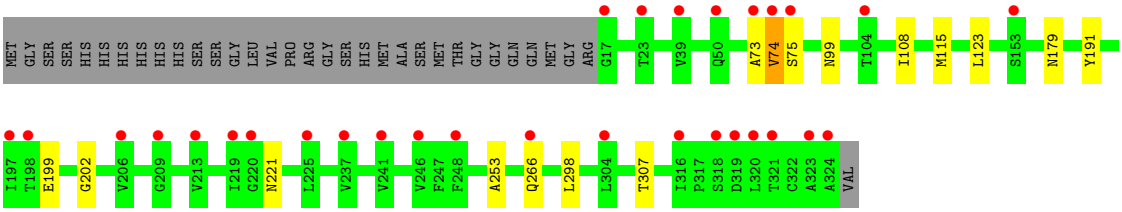




• Molecule 1: Awp1A

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	83.28Å 83.28Å 274.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.81 – 1.85 49.51 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.2 (45.81-1.85) 89.0 (49.51-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.04 (at 1.86Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.187 , 0.207 0.187 , 0.207	Depositor DCC
R_{free} test set	1621 reflections (1.95%)	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.800	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	9780	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.33% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: 9JE

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	0.28	0/2348	0.52	0/3219
1	B	0.30	0/2315	0.55	0/3175
All	All	0.29	0/4663	0.53	0/6394

There are no bond length outliers.

There are no bond angle outliers.

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	2293	2246	2220	7	0
1	B	2272	2224	2217	12	0
2	A	7	12	0	0	0
2	B	21	36	0	0	0
3	A	274	0	0	2	0
3	B	395	0	0	0	0
All	All	5262	4518	4437	19	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 2.

All (19) 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:301:GLN:OE1	3:A:501:HOH:O	2.08	0.70
1:B:74:VAL:HG22	1:B:75:SER:H	1.57	0.69
1:A:115[B]:MET:HE1	1:A:117:ASN:HB2	1.78	0.65
1:B:108:ILE:HD13	1:B:123:LEU:HD11	1.79	0.65
1:A:20:ASP:HB3	1:A:22:LEU:HD21	1.85	0.58
1:B:74:VAL:HG22	1:B:75:SER:N	2.18	0.58
1:B:298:LEU:HD12	1:B:298:LEU:C	2.35	0.47
1:B:74:VAL:CG2	1:B:75:SER:H	2.25	0.46
1:A:208:LYS:NZ	3:A:512:HOH:O	2.49	0.45
1:B:73:ALA:O	1:B:74:VAL:O	2.35	0.45
1:B:253:ALA:HB1	1:B:307:THR:HG21	2.00	0.43
1:B:179:ASN:O	1:B:202:GLY:HA3	2.18	0.42
1:A:71:LEU:HD23	1:A:74:VAL:HG11	2.01	0.42
1:A:309:LEU:HD12	1:A:309:LEU:N	2.35	0.42
1:B:108:ILE:CD1	1:B:123:LEU:HD11	2.47	0.41
1:A:298:LEU:C	1:A:298:LEU:HD12	2.41	0.41
1:B:73:ALA:O	1:B:99:ASN:OD1	2.39	0.41
1:B:74:VAL:CG2	1:B:75:SER:N	2.84	0.40
1:B:73:ALA:C	1:B:74:VAL:HG12	2.42	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	312/341 (92%)	302 (97%)	9 (3%)	1 (0%)	41	26
1	B	308/341 (90%)	298 (97%)	8 (3%)	2 (1%)	25	12
All	All	620/682 (91%)	600 (97%)	17 (3%)	3 (0%)	29	15

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	74	VAL
1	A	199	GLU
1	B	199	GLU

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	271/291 (93%)	267 (98%)	4 (2%)	65	53
1	B	267/291 (92%)	263 (98%)	4 (2%)	65	53
All	All	538/582 (92%)	530 (98%)	8 (2%)	65	53

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	86	SER
1	A	191	TYR
1	A	221	ASN
1	A	251	ILE
1	B	115	MET
1	B	191	TYR
1	B	221	ASN
1	B	266	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

4 ligands are modelled in this entry.

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, 95^{th} 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	307/341 (90%)	1.66	84 (27%) 0 0	37, 53, 82, 114	0
1	B	308/341 (90%)	1.04	30 (9%) 7 7	32, 40, 57, 125	0
All	All	615/682 (90%)	1.35	114 (18%) 1 1	32, 46, 77, 125	0

All (114) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	73	ALA	7.3
1	A	321	THR	6.1
1	A	33	PHE	6.1
1	A	23	THR	5.8
1	A	323	ALA	5.6
1	A	39	VAL	5.5
1	B	17	GLY	5.1
1	A	43	LEU	5.0
1	A	285	GLY	5.0
1	B	324	ALA	4.9
1	A	324	ALA	4.9
1	A	34	ASN	4.6
1	A	290	LEU	4.4
1	A	64	ALA	4.3
1	A	51	TYR	4.2
1	A	75	SER	4.1
1	A	270	LEU	4.1
1	A	22	LEU	4.1
1	A	265	THR	4.0
1	A	18	SER	4.0
1	B	246	VAL	3.9
1	A	24	PRO	3.9
1	A	266	GLN	3.8
1	A	35	GLU	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	284	CYS	3.8
1	A	168	LEU	3.8
1	A	73	ALA	3.7
1	A	309	LEU	3.7
1	B	75	SER	3.7
1	A	66	VAL	3.6
1	A	50	GLN	3.5
1	A	27	LEU	3.5
1	A	54	ASN	3.5
1	B	266	GLN	3.4
1	A	45	LEU	3.4
1	A	143	PHE	3.3
1	A	261	PHE	3.3
1	A	21	ILE	3.3
1	B	74	VAL	3.3
1	A	31	GLN	3.3
1	A	41	SER	3.1
1	A	251	ILE	3.1
1	A	25	THR	3.1
1	A	304	LEU	3.1
1	A	320	LEU	3.0
1	A	267	THR	3.0
1	B	323	ALA	3.0
1	A	322	CYS	3.0
1	A	40	VAL	3.0
1	A	53	PHE	3.0
1	A	310	VAL	3.0
1	A	52	LEU	2.9
1	A	316	ILE	2.9
1	A	87	LEU	2.9
1	A	197	ILE	2.9
1	A	19	LEU	2.9
1	A	258	SER	2.8
1	A	47	ASP	2.7
1	A	277	VAL	2.7
1	B	319	ASP	2.7
1	B	206	VAL	2.7
1	A	142	THR	2.7
1	A	76	VAL	2.7
1	A	203	SER	2.7
1	A	262	SER	2.7
1	A	308	PHE	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	291	MET	2.6
1	A	269	ILE	2.6
1	A	256	GLY	2.6
1	A	177	LEU	2.6
1	A	240	VAL	2.5
1	A	319	ASP	2.4
1	A	287	ASN	2.4
1	A	28	THR	2.4
1	B	304	LEU	2.4
1	B	213	VAL	2.4
1	B	241	VAL	2.4
1	A	218	PHE	2.4
1	B	197	ILE	2.4
1	A	61	PRO	2.4
1	A	305	TYR	2.4
1	B	321	THR	2.3
1	A	209	GLY	2.3
1	B	248	PHE	2.3
1	A	48	GLY	2.3
1	B	220	GLY	2.3
1	B	225	LEU	2.3
1	B	198	THR	2.3
1	A	114	VAL	2.3
1	B	153[A]	SER	2.2
1	A	274	THR	2.2
1	B	209	GLY	2.2
1	B	39	VAL	2.2
1	A	163	LEU	2.2
1	B	237	VAL	2.2
1	B	318	SER	2.1
1	A	115[A]	MET	2.1
1	A	57	LEU	2.1
1	A	271	THR	2.1
1	B	23	THR	2.1
1	A	26	THR	2.1
1	B	104	THR	2.1
1	A	59	ILE	2.1
1	A	311	LEU	2.1
1	A	91	GLY	2.1
1	A	215	ALA	2.1
1	A	141	VAL	2.0
1	A	126	ALA	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	232	ALA	2.0
1	B	50	GLN	2.0
1	B	219	ILE	2.0
1	B	316	ILE	2.0
1	B	320	LEU	2.0
1	A	32	THR	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, 95^{th} 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
2	9JE	B	402	7/7	0.51	0.39	130,157,166,168	0
2	9JE	A	401	7/7	0.61	0.34	94,113,117,117	0
2	9JE	B	401	7/7	0.70	0.26	72,91,105,105	0
2	9JE	B	403	7/7	0.85	0.17	61,79,95,99	0

6.5 Other polymers (i)

There are no such residues in this entry.