Full wwPDB X-ray Structure Validation Report (i)

Nov 11, 2024 – 08:38 AM JST

PDB ID	:	5H9V
Title	:	Crystal structure of Regnase PIN domain, form I
Authors	:	Yokogawa, M.; Tsushima, T.; Adachi, W.; Noda, N.N.; Inagaki, F.
Deposited on	:	2015-12-29
Resolution	:	2.75 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

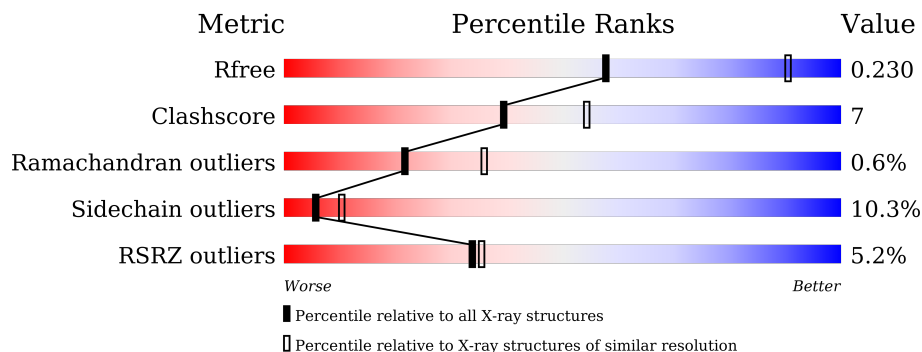|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.75 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		7
Ramachandran outliers		0.6%
Sidechain outliers		10.3%
RSRZ outliers		5.2%

Worse ..... *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	164625	1606 (2.78-2.74)
Clashscore	180529	1689 (2.78-2.74)
Ramachandran outliers	177936	1665 (2.78-2.74)
Sidechain outliers	177891	1665 (2.78-2.74)
RSRZ outliers	164620	1606 (2.78-2.74)

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	210	4% — 60% / 14% / · / 23%
1	B	210	3% — 58% / 15% / · / 25%
1	C	210	4% — 57% / 17% / · / 23%
1	D	210	5% — 59% / 17% / · / 23%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 5289 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 Ribonuclease ZC3H12A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	161	Total	C	N	O	S	Se	0	0	0
			1317	838	238	236	2	3			
1	B	158	Total	C	N	O	S	Se	0	0	0
			1304	833	233	233	2	3			
1	C	162	Total	C	N	O	S	Se	0	0	0
			1334	850	242	237	2	3			
1	D	162	Total	C	N	O	S	Se	0	0	0
			1330	847	241	237	2	3			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	130	GLY	-	expression tag	UNP Q5D1E7
A	131	PRO	-	expression tag	UNP Q5D1E7
A	132	HIS	-	expression tag	UNP Q5D1E7
A	133	MSE	-	expression tag	UNP Q5D1E7
B	130	GLY	-	expression tag	UNP Q5D1E7
B	131	PRO	-	expression tag	UNP Q5D1E7
B	132	HIS	-	expression tag	UNP Q5D1E7
B	133	MSE	-	expression tag	UNP Q5D1E7
C	130	GLY	-	expression tag	UNP Q5D1E7
C	131	PRO	-	expression tag	UNP Q5D1E7
C	132	HIS	-	expression tag	UNP Q5D1E7
C	133	MSE	-	expression tag	UNP Q5D1E7
D	130	GLY	-	expression tag	UNP Q5D1E7
D	131	PRO	-	expression tag	UNP Q5D1E7
D	132	HIS	-	expression tag	UNP Q5D1E7
D	133	MSE	-	expression tag	UNP Q5D1E7

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

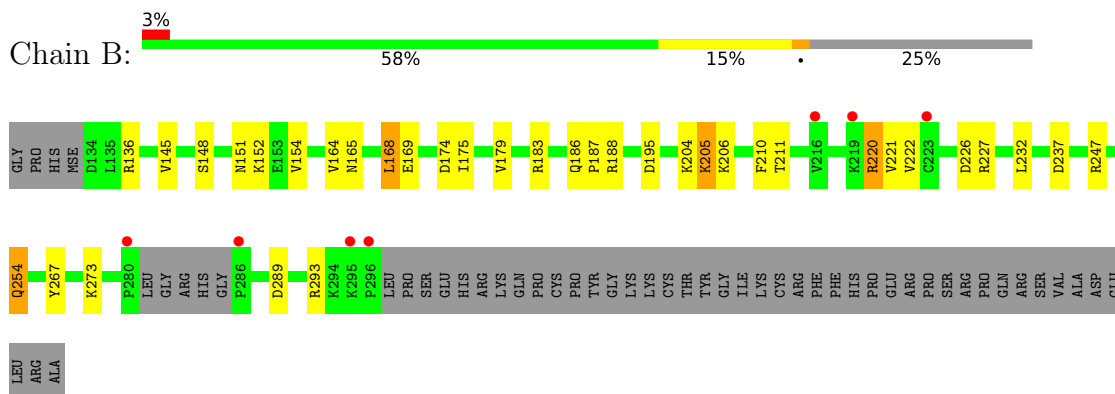| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Na 1	0	0
2	B	1	Total 1	Na 1	0	0
2	C	1	Total 1	Na 1	0	0
2	D	1	Total 1	Na 1	0	0

3 Residue-property plots ⓘ

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.
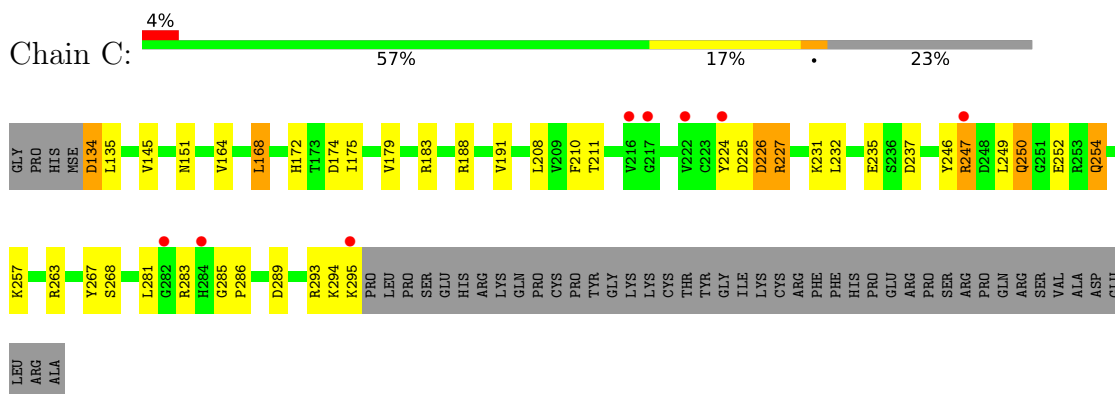
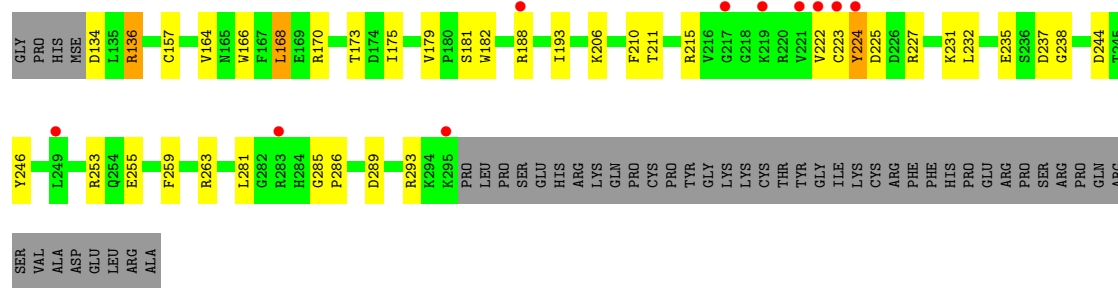• Molecule 1: Ribonuclease ZC3H12A

Chain A:

• Molecule 1: Ribonuclease ZC3H12A

Chain B:

• Molecule 1: Ribonuclease ZC3H12A

Chain C:

• Molecule 1: Ribonuclease ZC3H12A

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	113.37Å 113.37Å 187.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.49 – 2.75 47.49 – 2.75	Depositor EDS
% Data completeness (in resolution range)	90.3 (47.49-2.75) 90.4 (47.49-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.59 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.189 , 0.226 0.196 , 0.230	Depositor DCC
R_{free} test set	1656 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	60.2	Xtriage
Anisotropy	0.782	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 70.4	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.029 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5289	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% 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: NA

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	0.69	0/1346	0.90	3/1816 (0.2%)
1	B	0.77	0/1332	0.87	1/1793 (0.1%)
1	C	0.77	0/1363	0.92	0/1835
1	D	0.71	1/1359 (0.1%)	0.85	0/1831
All	All	0.74	1/5400 (0.0%)	0.89	4/7275 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	182	TRP	CB-CG	-5.08	1.41	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	183	ARG	NE-CZ-NH1	6.92	123.76	120.30
1	B	195	ASP	CB-CG-OD1	-5.58	113.28	118.30
1	A	266	MSE	CA-CB-CG	-5.13	104.57	113.30
1	A	183	ARG	NE-CZ-NH2	-5.03	117.79	120.30

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1317	0	1301	10	0
1	B	1304	0	1306	16	0
1	C	1334	0	1336	29	0
1	D	1330	0	1325	15	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
All	All	5289	0	5268	70	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 7.

All (70) 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:C:246:TYR:O	1:C:250:GLN:NE2	1.90	1.03
1:C:250:GLN:NE2	1:C:250:GLN:H	1.77	0.81
1:C:231:LYS:O	1:C:235:GLU:HG3	1.89	0.73
1:C:247:ARG:HA	1:C:250:GLN:NE2	2.06	0.70
1:B:136:ARG:NH1	1:B:293:ARG:O	2.25	0.69
1:C:281:LEU:HD12	1:C:286:PRO:HD2	1.76	0.68
1:B:186:GLN:HG3	1:B:187:PRO:HD2	1.78	0.66
1:C:250:GLN:H	1:C:250:GLN:HE21	1.41	0.65
1:C:289:ASP:O	1:C:293:ARG:HG3	1.97	0.64
1:D:237:ASP:OD1	1:D:263:ARG:NH2	2.31	0.63
1:D:259:PHE:CE1	1:D:263:ARG:HG3	2.37	0.60
1:C:237:ASP:OD1	1:C:263:ARG:NH2	2.36	0.59
1:B:289:ASP:O	1:B:293:ARG:HG3	2.03	0.58
1:A:139:VAL:HG22	1:A:232:LEU:HD13	1.86	0.58
1:B:254:GLN:HE21	1:B:254:GLN:HA	1.69	0.57
1:C:247:ARG:HA	1:C:250:GLN:CD	2.24	0.57
1:C:226:ASP:HB2	1:C:249:LEU:HD11	1.87	0.57
1:C:281:LEU:HB2	1:C:285:GLY:HA3	1.88	0.55
1:B:220:ARG:HG2	1:B:221:VAL:N	2.23	0.54
1:D:231:LYS:O	1:D:235:GLU:HG3	2.08	0.52
1:C:247:ARG:CA	1:C:250:GLN:NE2	2.73	0.51
1:B:168:LEU:HD13	1:B:175:ILE:CD1	2.41	0.51
1:A:164:VAL:HG12	1:A:168:LEU:HD22	1.92	0.51
1:C:227:ARG:NH2	1:C:252:GLU:OE2	2.44	0.51
1:C:134:ASP:OD1	1:C:134:ASP:N	2.43	0.51
1:C:250:GLN:HB2	1:C:257:LYS:HG3	1.92	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:166:TRP:CH2	1:D:170:ARG:HD2	2.47	0.50
1:A:137:PRO:HB3	1:A:174:ASP:OD2	2.12	0.49
1:B:168:LEU:HD13	1:B:175:ILE:HD12	1.93	0.49
1:B:179:VAL:O	1:B:210:PHE:HA	2.13	0.49
1:C:208:LEU:C	1:C:208:LEU:HD23	2.33	0.49
1:C:250:GLN:HE21	1:C:250:GLN:N	2.10	0.48
1:B:164:VAL:HG12	1:B:168:LEU:HD22	1.96	0.48
1:C:168:LEU:HD13	1:C:175:ILE:HD12	1.97	0.47
1:C:164:VAL:HG12	1:C:168:LEU:HD22	1.96	0.47
1:D:164:VAL:HG12	1:D:168:LEU:HD22	1.96	0.47
1:A:136:ARG:NH2	1:A:263:ARG:CZ	2.79	0.46
1:C:174:ASP:OD1	1:C:174:ASP:C	2.52	0.46
1:D:244:ASP:OD1	1:D:246:TYR:N	2.47	0.45
1:D:281:LEU:HB2	1:D:285:GLY:HA3	1.97	0.45
1:B:148:SER:O	1:B:152:LYS:NZ	2.44	0.45
1:B:174:ASP:OD1	1:B:174:ASP:C	2.55	0.45
1:C:145:VAL:HA	1:C:267:TYR:OH	2.16	0.45
1:D:289:ASP:O	1:D:293:ARG:HG3	2.17	0.45
1:A:143:SER:O	1:A:147:MSE:HG2	2.17	0.44
1:B:151:ASN:HB3	1:B:154:VAL:HB	1.99	0.44
1:C:254:GLN:CA	1:C:254:GLN:HE21	2.31	0.43
1:D:179:VAL:O	1:D:210:PHE:HA	2.18	0.43
1:C:168:LEU:HD13	1:C:175:ILE:CD1	2.48	0.43
1:B:168:LEU:CD1	1:B:175:ILE:HD12	2.49	0.43
1:D:157:CYS:SG	1:D:193:ILE:HD11	2.58	0.43
1:D:136:ARG:HG2	1:D:238:GLY:HA2	2.01	0.43
1:C:247:ARG:C	1:C:250:GLN:NE2	2.72	0.43
1:C:281:LEU:HD12	1:C:286:PRO:CD	2.45	0.42
1:A:165:ASN:HA	1:A:168:LEU:HB2	2.00	0.42
1:C:226:ASP:CB	1:C:249:LEU:HD11	2.48	0.42
1:C:179:VAL:O	1:C:210:PHE:HA	2.19	0.42
1:A:186:GLN:HA	1:A:187:PRO:HA	1.57	0.42
1:A:167:PHE:O	1:A:172:HIS:HB2	2.19	0.42
1:B:145:VAL:HA	1:B:267:TYR:OH	2.20	0.42
1:B:205:LYS:O	1:B:206:LYS:HB2	2.20	0.42
1:B:165:ASN:HA	1:B:168:LEU:HB2	2.01	0.41
1:A:258:ARG:NH2	1:A:261:GLU:OE1	2.53	0.41
1:C:294:LYS:O	1:C:295:LYS:HG2	2.20	0.41
1:D:168:LEU:HD13	1:D:175:ILE:CD1	2.50	0.41
1:D:223:CYS:O	1:D:224:TYR:C	2.59	0.41
1:D:281:LEU:HD12	1:D:286:PRO:CD	2.51	0.41

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:181:SER:HA	1:D:210:PHE:CD1	2.56	0.40
1:C:135:LEU:HD13	1:C:172:HIS:CE1	2.57	0.40
1:A:226:ASP:HB3	1:A:249:LEU:HD11	2.03	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	159/210 (76%)	149 (94%)	10 (6%)	0	100	100
1	B	154/210 (73%)	149 (97%)	4 (3%)	1 (1%)	22	36
1	C	160/210 (76%)	150 (94%)	8 (5%)	2 (1%)	10	17
1	D	160/210 (76%)	150 (94%)	9 (6%)	1 (1%)	22	36
All	All	633/840 (75%)	598 (94%)	31 (5%)	4 (1%)	22	36

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	220	ARG
1	C	224	TYR
1	D	224	TYR
1	C	225	ASP

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	146/188 (78%)	129 (88%)	17 (12%)	4	7
1	B	147/188 (78%)	132 (90%)	15 (10%)	6	10
1	C	149/188 (79%)	134 (90%)	15 (10%)	6	10
1	D	148/188 (79%)	134 (90%)	14 (10%)	7	13
All	All	590/752 (78%)	529 (90%)	61 (10%)	6	10

All (61) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	134	ASP
1	A	153	GLU
1	A	158	ARG
1	A	168	LEU
1	A	183	ARG
1	A	188	ARG
1	A	190	ASP
1	A	204	LYS
1	A	211	THR
1	A	213	SER
1	A	225	ASP
1	A	231	LYS
1	A	232	LEU
1	A	247	ARG
1	A	254	GLN
1	A	258	ARG
1	A	287	SER
1	B	168	LEU
1	B	169	GLU
1	B	183	ARG
1	B	188	ARG
1	B	204	LYS
1	B	205	LYS
1	B	211	THR
1	B	222	VAL
1	B	226	ASP
1	B	227	ARG
1	B	232	LEU
1	B	237	ASP
1	B	247	ARG
1	B	254	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	273	LYS
1	C	134	ASP
1	C	151	ASN
1	C	168	LEU
1	C	183	ARG
1	C	188	ARG
1	C	191	VAL
1	C	211	THR
1	C	226	ASP
1	C	227	ARG
1	C	232	LEU
1	C	247	ARG
1	C	250	GLN
1	C	254	GLN
1	C	268	SER
1	C	283	ARG
1	D	134	ASP
1	D	136	ARG
1	D	168	LEU
1	D	173	THR
1	D	188	ARG
1	D	206	LYS
1	D	211	THR
1	D	215	ARG
1	D	222	VAL
1	D	225	ASP
1	D	227	ARG
1	D	232	LEU
1	D	253	ARG
1	D	255	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	196	GLN
1	B	254	GLN
1	C	250	GLN
1	C	254	GLN
1	C	271	ASN
1	D	196	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 4 ligands modelled in this entry, 4 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 ⓘ

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 ⓘ

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	158/210 (75%)	0.06	8 (5%)	34	36	51, 86, 151, 179	0
1	B	155/210 (73%)	0.07	7 (4%)	39	40	48, 80, 151, 163	0
1	C	159/210 (75%)	-0.04	8 (5%)	35	37	56, 80, 140, 173	0
1	D	159/210 (75%)	0.16	10 (6%)	27	30	58, 89, 150, 176	0
All	All	631/840 (75%)	0.06	33 (5%)	34	35	48, 84, 150, 179	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	217	GLY	4.4
1	B	286	PRO	4.3
1	A	222	VAL	3.9
1	D	295	LYS	3.8
1	A	221	VAL	3.7
1	D	217	GLY	3.4
1	D	222	VAL	3.2
1	B	223	CYS	3.2
1	C	295	LYS	3.1
1	A	294	LYS	3.0
1	B	280	PRO	2.8
1	D	249	LEU	2.7
1	B	295	LYS	2.7
1	B	296	PRO	2.7
1	C	247	ARG	2.6
1	B	216	VAL	2.6
1	D	283	ARG	2.5
1	C	222	VAL	2.4
1	C	224	TYR	2.4
1	A	223	CYS	2.4
1	D	224	TYR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	284	HIS	2.3
1	D	221	VAL	2.3
1	A	189	PRO	2.2
1	A	187	PRO	2.2
1	C	216	VAL	2.2
1	D	219	LYS	2.2
1	D	188	ARG	2.1
1	C	284	HIS	2.1
1	C	282	GLY	2.1
1	B	219	LYS	2.0
1	D	223	CYS	2.0
1	A	134	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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(Å^2)	Q<0.9
2	NA	C	401	1/1	0.96	0.13	85,85,85,85	0
2	NA	D	401	1/1	0.97	0.06	88,88,88,88	0
2	NA	A	401	1/1	0.98	0.05	73,73,73,73	0
2	NA	B	401	1/1	0.98	0.10	79,79,79,79	0

6.5 Other polymers ⓘ

There are no such residues in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK