



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 02:22 AM GMT

PDB ID : 2GTR  
Title : Human chromodomain Y-like protein  
Authors : Min, J.; Wu, H.; Antoshenko, T.; Loppnau, P.; Weigelt, J.; Sundstrom, M.; Arrowsmith, C.H.; Edwards, A.M.; Bochkarev, A.; Plotnikov, A.N.; Structural Genomics Consortium (SGC)  
Deposited on : 2006-04-28  
Resolution : 1.90 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

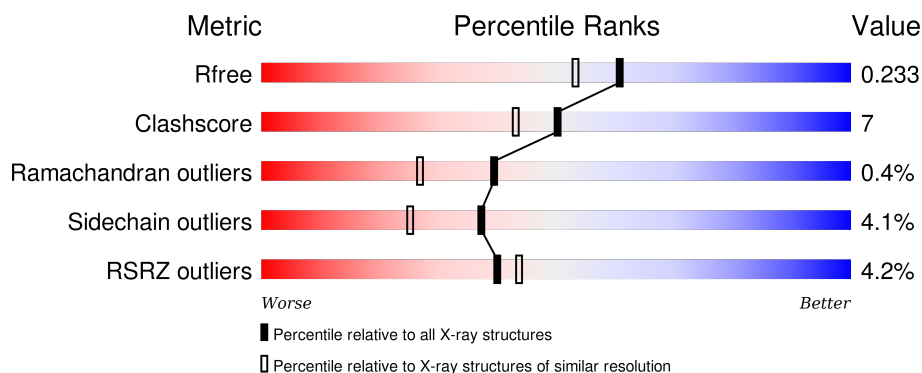
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	261	<div> <div></div> <div>86% 11% ..</div> </div>
1	B	261	<div> <div>3%</div> <div>82% 15% ..</div> </div>
1	C	261	<div> <div>9%</div> <div>79% 14% . .</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6497 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 Chromodomain Y-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	258	Total	C	N	O	S	0	3	0
			2030	1283	350	378	19			
1	B	258	Total	C	N	O	S	0	0	0
			2004	1269	342	375	18			
1	C	250	Total	C	N	O	S	0	0	0
			1931	1220	333	361	17			

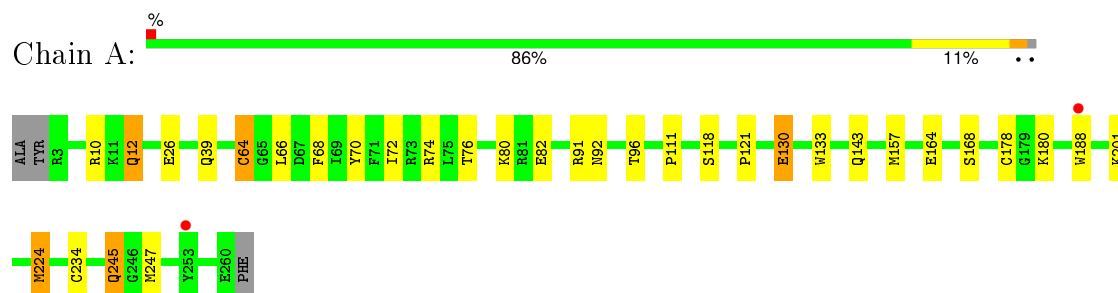
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	236	Total	O	0	0
			236	236		
2	B	158	Total	O	0	0
			158	158		
2	C	138	Total	O	0	0
			138	138		

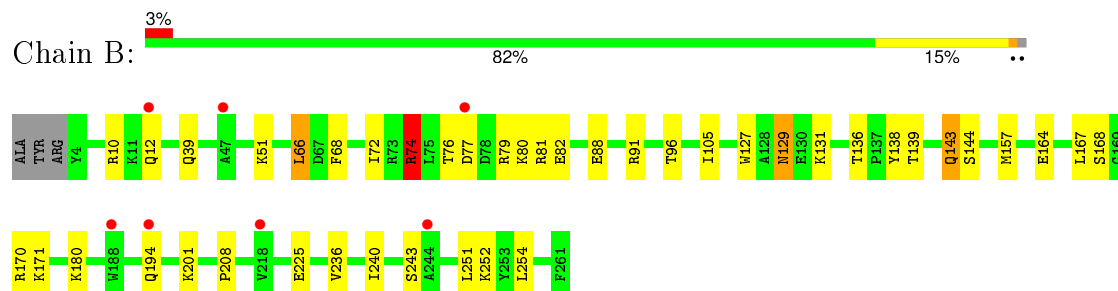
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

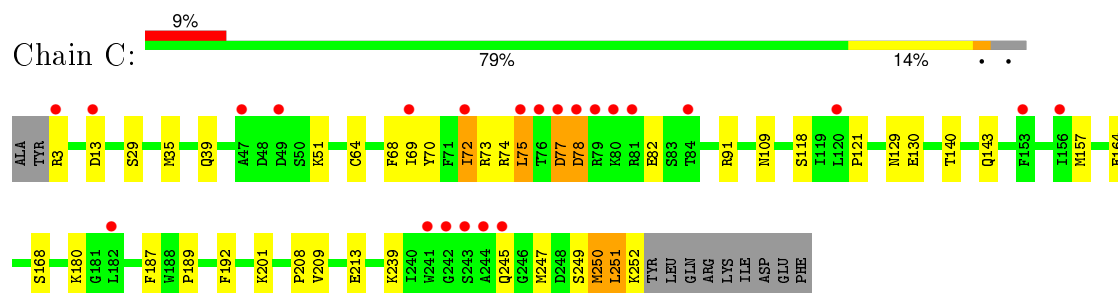
- Molecule 1: Chromodomain Y-like protein



- Molecule 1: Chromodomain Y-like protein



- Molecule 1: Chromodomain Y-like protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.21Å 132.74Å 100.36Å 90.00° 102.92° 90.00°	Depositor
Resolution (Å)	39.37 – 1.90 39.37 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.0 (39.37-1.90) 96.0 (39.37-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.03 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.189 , 0.233 0.188 , 0.233	Depositor DCC
$R_{free}$ test set	3671 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.2	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 50.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 73449 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6497	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	1.05	3/2063 (0.1%)	0.85	0/2779
1	B	0.85	0/2038	0.84	3/2749 (0.1%)
1	C	0.74	1/1963 (0.1%)	0.73	0/2648
All	All	0.89	4/6064 (0.1%)	0.81	3/8176 (0.0%)

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	B	0	1
1	C	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	64	CYS	CB-SG	-9.05	1.66	1.82
1	A	234	CYS	CB-SG	5.80	1.92	1.82
1	C	64	CYS	CB-SG	-5.56	1.72	1.81
1	A	178	CYS	CB-SG	-5.17	1.73	1.81

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	170	ARG	NE-CZ-NH1	8.06	124.33	120.30
1	B	170	ARG	NE-CZ-NH2	-5.81	117.39	120.30
1	B	74	ARG	NE-CZ-NH2	-5.57	117.52	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	251	LEU	Peptide
1	C	250	MET	Peptide

## 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	2030	0	2061	29	1
1	B	2004	0	2025	24	1
1	C	1931	0	1958	31	0
2	A	236	0	0	5	0
2	B	158	0	0	5	0
2	C	138	0	0	10	0
All	All	6497	0	6044	84	1

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 (84) 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:68:PHE:HB2	2:C:387:HOH:O	1.67	0.94
1:A:72:ILE:HD13	1:A:247:MET:HG3	1.51	0.92
1:B:129:ASN:ND2	1:B:131:LYS:H	1.69	0.91
1:B:129:ASN:HD22	1:B:131:LYS:H	1.04	0.91
1:C:251:LEU:HA	2:C:382:HOH:O	1.72	0.89
1:C:109:ASN:OD1	1:C:129:ASN:ND2	2.12	0.81
1:A:130:GLU:HB3	2:A:308:HOH:O	1.79	0.81
1:C:130:GLU:HB3	2:C:378:HOH:O	1.82	0.78
1:A:10:ARG:HB3	1:A:12:GLN:HE22	1.51	0.75
1:B:81:ARG:NH1	1:B:82:GLU:OE2	2.19	0.74
1:B:143:GLN:HG3	1:B:144:SER:O	1.87	0.73
1:B:129:ASN:HD22	1:B:131:LYS:N	1.83	0.72
1:B:10:ARG:HD3	2:B:320:HOH:O	1.90	0.70
1:A:245:GLN:H	1:A:245:GLN:HE21	1.39	0.69
1:C:249:SER:O	1:C:251:LEU:HB2	1.94	0.68
1:A:91:ARG:HH11	1:A:92:ASN:ND2	1.91	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:39:GLN:HG2	1:A:96:THR:OG1	1.96	0.65
1:B:88:GLU:HG2	1:B:91:ARG:HH12	1.62	0.64
1:C:251:LEU:HD22	2:C:382:HOH:O	1.98	0.63
1:A:72:ILE:CD1	1:A:247:MET:HG3	2.29	0.62
1:A:72:ILE:O	1:A:76:THR:HG23	2.00	0.61
1:A:201[B]:LYS:HD3	2:A:355:HOH:O	2.00	0.61
1:A:164:GLU:O	1:A:168:SER:HB2	2.01	0.61
1:A:188:TRP:HE3	2:A:477:HOH:O	1.82	0.60
1:C:77:ASP:O	1:C:78:ASP:HB2	2.01	0.60
1:C:91:ARG:HD2	2:C:299:HOH:O	2.01	0.60
1:C:252:LYS:HG3	2:C:377:HOH:O	2.02	0.60
1:A:91:ARG:HH11	1:A:92:ASN:HD22	1.48	0.59
1:C:73:ARG:HB2	1:C:74:ARG:NH1	2.17	0.59
1:B:72:ILE:O	1:B:76:THR:HG23	2.03	0.58
1:A:10:ARG:HB3	1:A:12:GLN:NE2	2.19	0.58
1:C:118:SER:O	1:C:121:PRO:HD2	2.03	0.57
1:B:194:GLN:HB3	2:B:410:HOH:O	2.06	0.55
1:A:74:ARG:O	1:A:82:GLU:HG3	2.07	0.55
1:C:13:ASP:HB3	2:C:347:HOH:O	2.06	0.55
1:B:225:GLU:HG3	2:B:291:HOH:O	2.07	0.54
1:A:68:PHE:HZ	1:A:143:GLN:NE2	2.06	0.53
1:C:245:GLN:CD	1:C:245:GLN:H	2.12	0.53
1:A:111:PRO:HG3	1:A:133:TRP:CE2	2.45	0.52
1:A:111:PRO:HG3	1:A:133:TRP:NE1	2.25	0.51
1:A:157:MET:HB3	1:A:180:LYS:O	2.11	0.51
1:A:91:ARG:NH1	1:A:92:ASN:ND2	2.58	0.50
1:B:164:GLU:O	1:B:168:SER:HB2	2.12	0.50
1:A:91:ARG:NH1	1:A:92:ASN:HD22	2.10	0.50
1:A:224[B]:MET:HE2	1:A:224[B]:MET:HA	1.95	0.49
1:C:157:MET:HB3	1:C:180:LYS:O	2.11	0.49
1:B:136:THR:HB	1:B:138:TYR:CE2	2.47	0.49
1:B:143:GLN:HB2	2:B:273:HOH:O	2.12	0.48
1:C:252:LYS:HE3	2:C:335:HOH:O	2.12	0.48
1:C:68:PHE:HZ	1:C:143:GLN:NE2	2.11	0.48
1:C:130:GLU:HG2	1:C:189:PRO:HG3	1.95	0.47
1:A:10:ARG:CB	1:A:12:GLN:HE22	2.22	0.47
1:C:35:MET:O	1:C:39:GLN:HG3	2.15	0.46
1:B:236:VAL:HG12	1:B:240:ILE:HD12	1.99	0.45
1:A:68:PHE:HZ	1:A:143:GLN:HE21	1.64	0.45
1:C:51:LYS:HG2	1:C:208:PRO:HB3	1.98	0.45
1:B:88:GLU:HG2	1:B:91:ARG:NH1	2.29	0.45

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:105:ILE:HG21	1:B:127:TRP:CZ3	2.52	0.45
1:C:68:PHE:HZ	1:C:143:GLN:HE21	1.66	0.44
1:C:74:ARG:O	1:C:82:GLU:HG3	2.17	0.44
1:B:39:GLN:HG2	1:B:96:THR:OG1	2.18	0.44
1:C:187:PHE:HB2	1:C:192:PHE:HD1	1.83	0.44
1:B:157:MET:HB3	1:B:180:LYS:O	2.18	0.43
1:B:74:ARG:HB3	1:B:82:GLU:HG3	2.00	0.43
1:B:138:TYR:HD1	1:B:143:GLN:HG2	1.82	0.43
1:A:201[B]:LYS:HD3	1:A:201[B]:LYS:HA	1.78	0.43
1:C:75:LEU:HD13	1:C:247:MET:HG3	1.99	0.43
1:A:201[B]:LYS:CD	2:A:355:HOH:O	2.62	0.43
1:B:139:THR:HG21	1:B:167:LEU:HA	2.01	0.42
1:C:73:ARG:HB2	1:C:74:ARG:HH12	1.84	0.42
1:C:72:ILE:HD13	1:C:72:ILE:HA	1.85	0.42
1:A:70:TYR:CD2	1:A:70:TYR:C	2.93	0.42
1:B:66:LEU:HB3	1:B:68:PHE:CE2	2.54	0.42
1:C:209:VAL:O	1:C:213:GLU:HG2	2.19	0.41
1:B:171:LYS:HE2	2:B:330:HOH:O	2.20	0.41
1:A:72:ILE:HD13	1:A:247:MET:CG	2.36	0.41
1:A:118:SER:O	1:A:121:PRO:HD2	2.20	0.41
1:B:51:LYS:HG2	1:B:208:PRO:HB3	2.03	0.41
1:C:164:GLU:O	1:C:168:SER:HB2	2.21	0.41
1:C:70:TYR:C	1:C:70:TYR:CD2	2.93	0.41
1:C:143:GLN:HB3	2:C:383:HOH:O	2.20	0.40
1:C:130:GLU:HG2	1:C:189:PRO:CG	2.51	0.40
1:C:250:MET:CB	2:C:398:HOH:O	2.70	0.40
1:A:39:GLN:OE1	2:A:496:HOH:O	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:26:GLU:OE2	1:B:74:ARG:NH2[4_444]	2.05	0.15

## 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	259/261 (99%)	255 (98%)	4 (2%)	0	100	100
1	B	256/261 (98%)	248 (97%)	6 (2%)	2 (1%)	24	11
1	C	248/261 (95%)	236 (95%)	11 (4%)	1 (0%)	39	27
All	All	763/783 (97%)	739 (97%)	21 (3%)	3 (0%)	39	27

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	78	ASP
1	B	252	LYS
1	B	77	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	224/224 (100%)	216 (96%)	8 (4%)	42	30
1	B	221/224 (99%)	211 (96%)	10 (4%)	34	21
1	C	213/224 (95%)	203 (95%)	10 (5%)	32	20
All	All	658/672 (98%)	630 (96%)	28 (4%)	37	23

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	GLN
1	A	64	CYS
1	A	66	LEU
1	A	80	LYS
1	A	130	GLU
1	A	224[A]	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	224[B]	MET
1	A	245	GLN
1	B	12	GLN
1	B	66	LEU
1	B	74	ARG
1	B	79	ARG
1	B	80	LYS
1	B	129	ASN
1	B	143	GLN
1	B	201	LYS
1	B	243	SER
1	B	254	LEU
1	C	3	ARG
1	C	29	SER
1	C	69	ILE
1	C	72	ILE
1	C	75	LEU
1	C	77	ASP
1	C	140	THR
1	C	201	LYS
1	C	239	LYS
1	C	251	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	12	GLN
1	A	92	ASN
1	A	143	GLN
1	A	163	ASN
1	A	175	GLN
1	A	245	GLN
1	B	92	ASN
1	B	129	ASN
1	B	143	GLN
1	B	194	GLN
1	C	92	ASN
1	C	163	ASN
1	C	175	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	258/261 (98%)	-0.18	2 (0%) 87 88	17, 24, 40, 55	0
1	B	258/261 (98%)	0.09	7 (2%) 58 61	16, 33, 49, 61	0
1	C	250/261 (95%)	0.58	23 (9%) 11 12	22, 38, 69, 80	0
All	All	766/783 (97%)	0.16	32 (4%) 40 44	16, 32, 55, 80	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	244	ALA	9.2
1	B	244	ALA	5.7
1	C	80	LYS	5.4
1	C	242	GLY	5.3
1	C	3	ARG	5.0
1	B	188	TRP	4.9
1	C	69	ILE	4.8
1	C	75	LEU	4.5
1	C	76	THR	4.4
1	C	77	ASP	4.1
1	C	81	ARG	3.9
1	C	245	GLN	3.8
1	C	78	ASP	3.5
1	B	194	GLN	3.2
1	C	243	SER	3.0
1	B	12	GLN	2.9
1	A	188	TRP	2.7
1	B	47	ALA	2.7
1	C	120	LEU	2.6
1	C	47	ALA	2.6
1	C	156	ILE	2.6
1	C	49	ASP	2.4
1	C	79	ARG	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	13	ASP	2.4
1	B	77	ASP	2.3
1	C	153	PHE	2.3
1	A	253	TYR	2.3
1	C	241	TRP	2.1
1	C	72	ILE	2.1
1	C	182	LEU	2.0
1	C	84	THR	2.0
1	B	218	VAL	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.