



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 14, 2023 – 09:50 PM EDT

PDB ID : 8CZG  
Title : Human BAK in complex with the dF3 peptide  
Authors : Aguilar, F.; Keating, A.E.  
Deposited on : 2022-05-24  
Resolution : 1.99 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

<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 : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

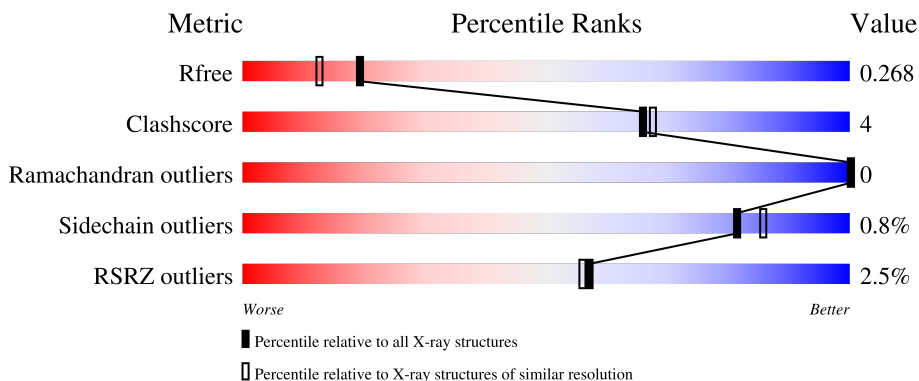
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.99 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.



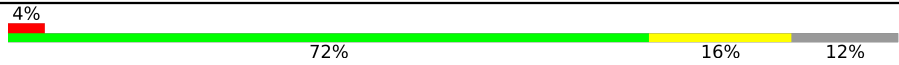

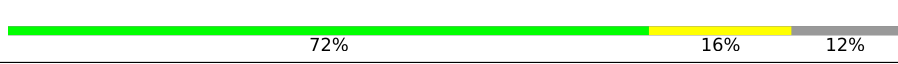
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	170	
1	B	170	
1	C	170	
1	D	170	
2	E	25	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	25	 4% 72% 16% 12%
2	G	25	 4% 76% 16% 8%
2	H	25	 4% 72% 16% 12%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 11698 atoms, of which 5573 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 Bcl-2 homologous antagonist/killer.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	157	2480	806	1216	221	232	5	0	0	0
1	B	155	2424	793	1182	216	229	4	0	0	0
1	C	160	2509	816	1227	224	238	4	0	0	0
1	D	160	2501	815	1219	224	238	5	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	GLY	-	expression tag	UNP Q16611
A	18	PRO	-	expression tag	UNP Q16611
A	19	LEU	-	expression tag	UNP Q16611
A	20	GLY	-	expression tag	UNP Q16611
A	21	SER	-	expression tag	UNP Q16611
A	22	MET	-	expression tag	UNP Q16611
A	166	SER	CYS	engineered mutation	UNP Q16611
B	17	GLY	-	expression tag	UNP Q16611
B	18	PRO	-	expression tag	UNP Q16611
B	19	LEU	-	expression tag	UNP Q16611
B	20	GLY	-	expression tag	UNP Q16611
B	21	SER	-	expression tag	UNP Q16611
B	22	MET	-	expression tag	UNP Q16611
B	166	SER	CYS	engineered mutation	UNP Q16611
C	17	GLY	-	expression tag	UNP Q16611
C	18	PRO	-	expression tag	UNP Q16611
C	19	LEU	-	expression tag	UNP Q16611
C	20	GLY	-	expression tag	UNP Q16611
C	21	SER	-	expression tag	UNP Q16611
C	22	MET	-	expression tag	UNP Q16611
C	166	SER	CYS	engineered mutation	UNP Q16611

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	17	GLY	-	expression tag	UNP Q16611
D	18	PRO	-	expression tag	UNP Q16611
D	19	LEU	-	expression tag	UNP Q16611
D	20	GLY	-	expression tag	UNP Q16611
D	21	SER	-	expression tag	UNP Q16611
D	22	MET	-	expression tag	UNP Q16611
D	166	SER	CYS	engineered mutation	UNP Q16611

- Molecule 2 is a protein called dF3 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
2	E	22	Total	C	H	N	O	0	0	0
			342	110	170	28	34			
2	F	22	Total	C	H	N	O	0	1	0
			365	116	185	29	35			
2	G	23	Total	C	H	N	O	0	0	0
			376	118	190	30	38			
2	H	22	Total	C	H	N	O	0	0	0
			361	113	184	29	35			

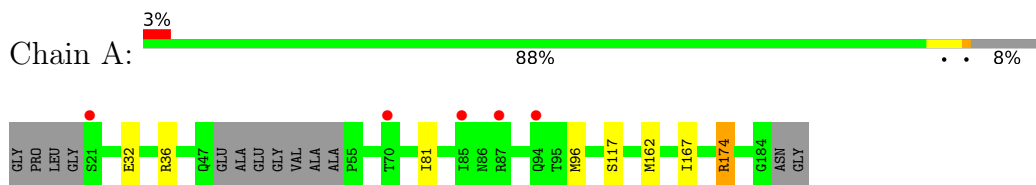
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	74	Total	O	0	0
			74	74		
3	B	81	Total	O	0	0
			81	81		
3	C	84	Total	O	0	0
			84	84		
3	D	81	Total	O	0	0
			81	81		
3	E	2	Total	O	0	0
			2	2		
3	F	2	Total	O	0	0
			2	2		
3	G	9	Total	O	0	0
			9	9		
3	H	7	Total	O	0	0
			7	7		

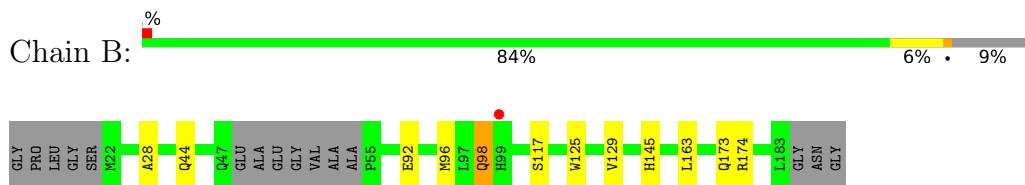
### 3 Residue-property plots [i](#)

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.

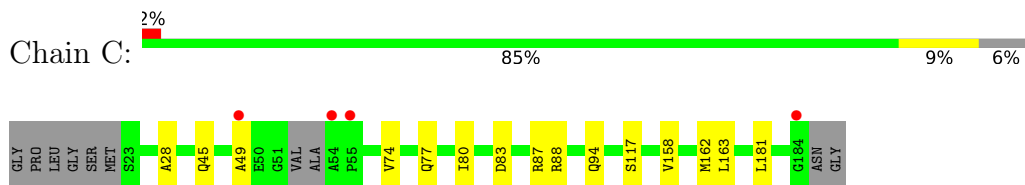
- Molecule 1: Bcl-2 homologous antagonist/killer



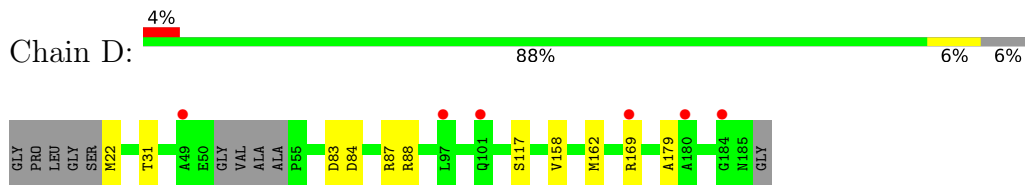
- Molecule 1: Bcl-2 homologous antagonist/killer



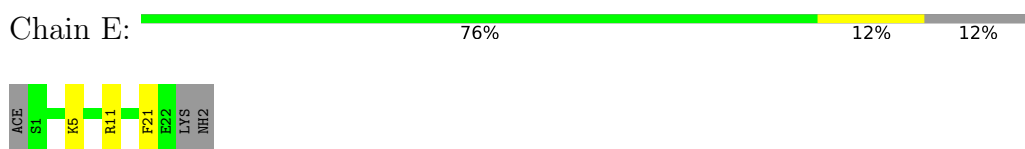
- Molecule 1: Bcl-2 homologous antagonist/killer



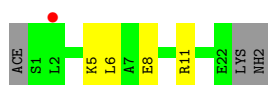
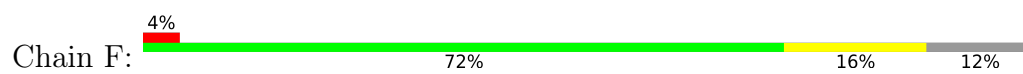
- Molecule 1: Bcl-2 homologous antagonist/killer



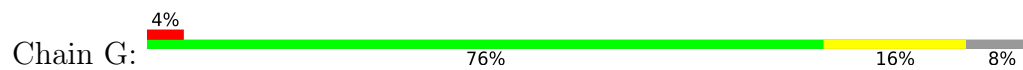
- Molecule 2: dF3 peptide



- Molecule 2: dF3 peptide



- Molecule 2: dF3 peptide



- Molecule 2: dF3 peptide



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.29Å 65.26Å 111.57Å 90.00° 102.12° 90.00°	Depositor
Resolution (Å)	41.85 – 1.99 41.85 – 1.99	Depositor EDS
% Data completeness (in resolution range)	92.0 (41.85-1.99) 92.3 (41.85-1.99)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.96 (at 1.98Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.228 , 0.267 0.228 , 0.268	Depositor DCC
$R_{free}$ test set	1995 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtrriage
Anisotropy	0.825	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11698	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 79.61 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.3780e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: ACE

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	0.26	0/1295	0.48	0/1754
1	B	0.27	0/1273	0.48	0/1727
1	C	0.27	0/1313	0.48	0/1780
1	D	0.27	0/1313	0.48	0/1779
2	E	0.26	0/172	0.43	0/229
2	F	0.28	0/183	0.41	0/244
2	G	0.26	0/184	0.46	0/244
2	H	0.26	0/175	0.47	0/232
All	All	0.27	0/5908	0.48	0/7989

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 [i](#)

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	1264	1216	1215	7	0
1	B	1242	1182	1180	11	0
1	C	1282	1227	1225	10	0
1	D	1282	1219	1218	12	0
2	E	172	170	170	4	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	180	185	188	5	0
2	G	186	190	193	3	0
2	H	177	184	187	4	0
3	A	74	0	0	1	1
3	B	81	0	0	2	0
3	C	84	0	0	3	1
3	D	81	0	0	3	0
3	E	2	0	0	0	0
3	F	2	0	0	0	0
3	G	9	0	0	1	0
3	H	7	0	0	0	0
All	All	6125	5573	5576	45	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 4.

All (45) 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:88:ARG:O	3:C:201:HOH:O	2.07	0.73
1:D:179:ALA:O	3:D:201:HOH:O	2.07	0.72
1:B:44:GLN:NE2	3:B:202:HOH:O	2.22	0.70
1:D:83:ASP:OD2	1:D:87:ARG:NH1	2.25	0.69
1:D:31:THR:HG21	1:D:162:MET:HE3	1.74	0.68
1:D:31:THR:HG21	1:D:162:MET:CE	2.24	0.68
1:C:117:SER:OG	2:G:11:ARG:NH2	2.28	0.66
1:D:169:ARG:NH2	3:D:202:HOH:O	2.29	0.65
1:C:49:ALA:O	3:C:202:HOH:O	2.17	0.59
1:D:158:VAL:HG12	1:D:162:MET:HE2	1.86	0.57
1:A:81:ILE:HD12	2:E:21:PHE:HE1	1.70	0.57
1:B:96:MET:HE2	2:F:6:LEU:HG	1.87	0.57
1:B:28:ALA:HB2	1:B:163:LEU:HD13	1.88	0.54
1:A:174:ARG:NH1	3:A:201:HOH:O	2.41	0.54
1:B:96:MET:HE1	2:F:5:LYS:C	2.28	0.54
1:D:22:MET:N	3:D:208:HOH:O	2.42	0.52
2:F:8:GLU:OE1	2:F:11:ARG:NH1	2.43	0.51
1:C:77:GLN:O	1:C:80:ILE:HG13	2.10	0.50
1:D:84:ASP:OD1	1:D:87:ARG:NH2	2.45	0.50
1:A:117:SER:OG	2:E:11:ARG:NH2	2.46	0.49
1:B:117:SER:OG	2:F:11:ARG:NH2	2.46	0.49
2:G:0:ACE:H3	3:G:109:HOH:O	2.12	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:117:SER:OG	2:H:11:ARG:NH2	2.46	0.48
1:A:81:ILE:HD12	2:E:21:PHE:CE1	2.49	0.47
1:A:32:GLU:OE2	1:A:36:ARG:HD2	2.14	0.47
1:D:158:VAL:HG12	1:D:162:MET:CE	2.45	0.47
1:C:83:ASP:OD2	1:C:87:ARG:NH1	2.45	0.46
1:C:94:GLN:NE2	3:C:206:HOH:O	2.47	0.45
1:B:98:GLN:OE1	1:B:145:HIS:NE2	2.47	0.45
1:D:88:ARG:NH2	2:H:16:GLU:OE1	2.43	0.44
1:C:28:ALA:HB2	1:C:163:LEU:HD13	1.99	0.44
1:D:31:THR:HG21	1:D:162:MET:HE1	1.99	0.43
1:B:125:TRP:O	1:B:129:VAL:HG23	2.18	0.43
1:B:173:GLN:OE1	3:B:201:HOH:O	2.21	0.43
1:C:45:GLN:OE1	1:C:80:ILE:HA	2.17	0.43
1:C:158:VAL:HG12	1:C:162:MET:HE3	2.01	0.43
2:G:13:LEU:HD23	2:G:13:LEU:O	2.19	0.42
1:B:96:MET:HE1	2:F:6:LEU:N	2.35	0.42
1:C:74:VAL:HG22	1:C:181:LEU:HG	2.02	0.41
1:B:92:GLU:O	1:B:96:MET:HG3	2.20	0.41
1:B:92:GLU:OE1	1:B:92:GLU:N	2.46	0.41
2:H:13:LEU:HD23	2:H:13:LEU:C	2.41	0.41
2:H:13:LEU:HD23	2:H:13:LEU:O	2.21	0.41
1:A:96:MET:HE3	2:E:5:LYS:HB3	2.03	0.40
1:A:162:MET:HA	1:A:167:ILE:HB	2.04	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)
3:A:271:HOH:O	3:C:284:HOH:O[2_545]	2.16	0.04

## 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	153/170 (90%)	152 (99%)	1 (1%)	0	100	100
1	B	151/170 (89%)	151 (100%)	0	0	100	100
1	C	156/170 (92%)	156 (100%)	0	0	100	100
1	D	156/170 (92%)	156 (100%)	0	0	100	100
2	E	20/25 (80%)	20 (100%)	0	0	100	100
2	F	21/25 (84%)	21 (100%)	0	0	100	100
2	G	21/25 (84%)	21 (100%)	0	0	100	100
2	H	20/25 (80%)	19 (95%)	1 (5%)	0	100	100
All	All	698/780 (90%)	696 (100%)	2 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	132/138 (96%)	131 (99%)	1 (1%)	81	86
1	B	128/138 (93%)	126 (98%)	2 (2%)	62	67
1	C	132/138 (96%)	132 (100%)	0	100	100
1	D	132/138 (96%)	132 (100%)	0	100	100
2	E	17/21 (81%)	17 (100%)	0	100	100
2	F	19/21 (90%)	19 (100%)	0	100	100
2	G	20/21 (95%)	19 (95%)	1 (5%)	24	20
2	H	19/21 (90%)	18 (95%)	1 (5%)	22	18
All	All	599/636 (94%)	594 (99%)	5 (1%)	81	86

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

Mol	Chain	Res	Type
1	A	174	ARG
1	B	98	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	174	ARG
2	G	2	LEU
2	H	2	LEU

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](#)

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	157/170 (92%)	0.42	5 (3%) 47 46	19, 29, 44, 60	0
1	B	155/170 (91%)	0.35	1 (0%) 89 88	16, 26, 43, 54	0
1	C	160/170 (94%)	0.30	4 (2%) 57 56	19, 26, 40, 56	0
1	D	160/170 (94%)	0.35	6 (3%) 40 39	15, 28, 49, 60	0
2	E	22/25 (88%)	0.42	0 100 100	24, 34, 57, 66	0
2	F	22/25 (88%)	0.33	1 (4%) 33 32	22, 31, 45, 60	0
2	G	22/25 (88%)	0.27	1 (4%) 33 32	22, 33, 41, 67	0
2	H	21/25 (84%)	0.10	0 100 100	21, 31, 42, 44	0
All	All	719/780 (92%)	0.34	18 (2%) 57 56	15, 28, 47, 67	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	49	ALA	2.9
1	C	184	GLY	2.6
1	A	85	ILE	2.6
1	C	55	PRO	2.5
1	A	94	GLN	2.3
1	C	54	ALA	2.3
1	A	87	ARG	2.2
1	A	21	SER	2.2
1	D	97	LEU	2.2
2	F	2[A]	LEU	2.2
1	A	70	THR	2.2
1	D	184	GLY	2.1
1	D	180	ALA	2.1
1	B	99	HIS	2.1
1	D	169	ARG	2.1
2	G	22	GLU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	49	ALA	2.1
1	D	101	GLN	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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.