



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 7, 2017 – 09:10 AM EST

PDB ID : 4JRX  
Title : Crystal Structure of CA5 TCR-HLA B\*3505-LPEP complex  
Authors : Liu, Y.C.; Rossjohn, J.; Gras, S.  
Deposited on : unknown  
Resolution : 2.30 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  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

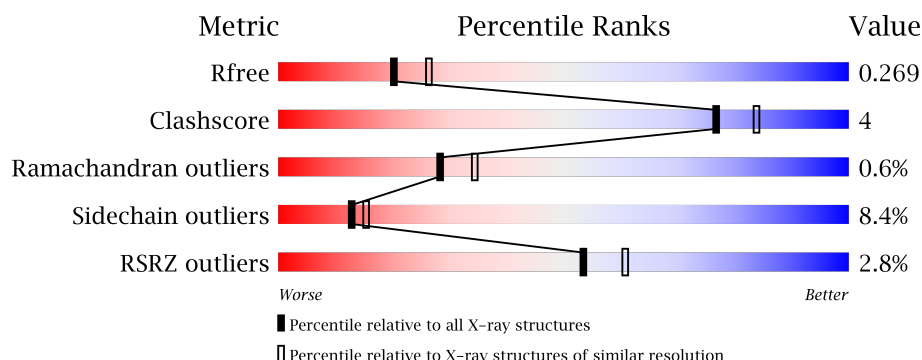
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
$R_{free}$	100719	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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	276	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>11%</div> <div>.</div> </div> </div>
2	B	99	<div> <div>%</div> <div> <div></div> <div>86%</div> <div>12%</div> <div>.</div> </div> </div>
3	C	13	<div> <div></div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>
4	D	204	<div> <div>4%</div> <div> <div></div> <div>82%</div> <div>16%</div> <div>.</div> </div> </div>
5	E	238	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>14%</div> <div>.</div> </div> </div>

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 6865 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2257	1405	414	431	7			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			

- Molecule 3 is a protein called Trans-activator protein BZLF1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	0	0	0
			101	66	15	20			

- Molecule 4 is a protein called CA5 TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	204	Total	C	N	O	S	0	1	0
			1646	1041	266	332	7			

- Molecule 5 is a protein called CA5 TCR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	238	Total	C	N	O	S	0	0	0
			1889	1185	329	366	9			

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total 1	Na 1	0	0
6	E	1	Total 1	Na 1	0	0

- Molecule 7 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	3	Total 3	I 3	0	0
7	D	3	Total 3	I 3	0	0
7	C	1	Total 1	I 1	0	0
7	E	1	Total 1	I 1	0	0

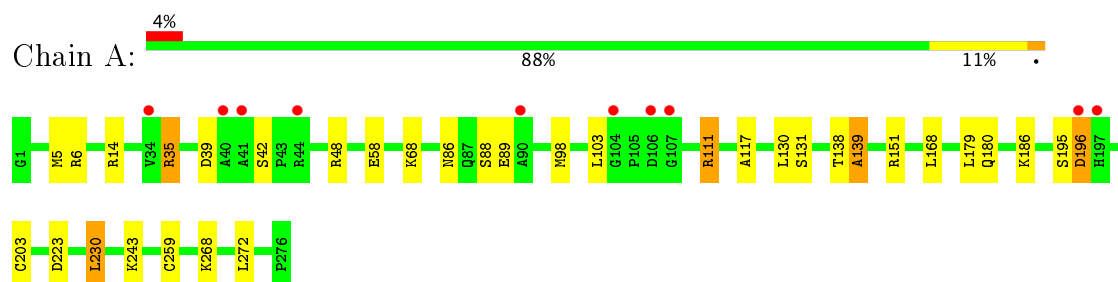
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	37	Total 37	O 37	0	0
8	B	18	Total 18	O 18	0	0
8	C	3	Total 3	O 3	0	0
8	D	36	Total 36	O 36	0	0
8	E	38	Total 38	O 38	0	0

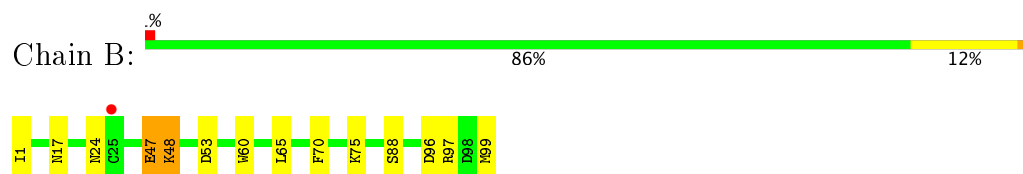
### 3 Residue-property plots [i](#)

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

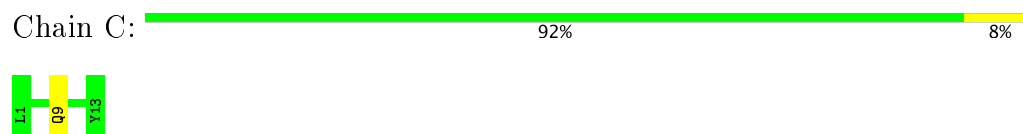
#### • Molecule 1: MHC class I antigen



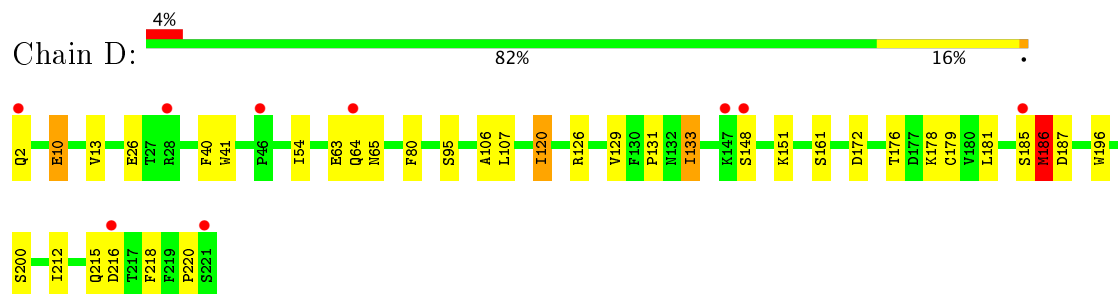
#### • Molecule 2: Beta-2-microglobulin



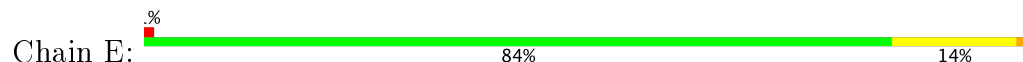
#### • Molecule 3: Trans-activator protein BZLF1

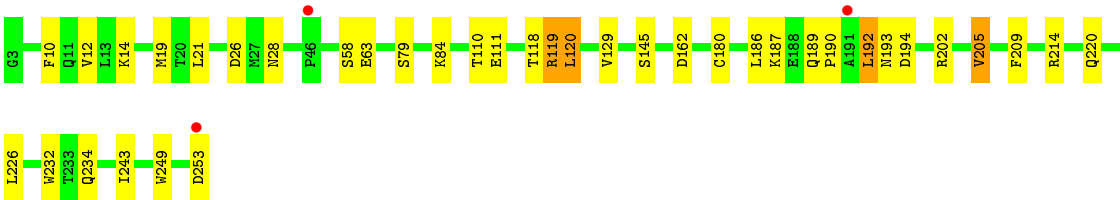


#### • Molecule 4: CA5 TCR alpha chain



#### • Molecule 5: CA5 TCR beta chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.10 Å 78.41 Å 105.34 Å 90.00° 93.06° 90.00°	Depositor
Resolution (Å)	19.80 – 2.30 19.79 – 2.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (19.80-2.30) 100.0 (19.79-2.30)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.20 (at 2.30 Å)	Xtriage
Refinement program	BUSTER-TNT, BUSTER	Depositor
R, $R_{free}$	0.205 , 0.257 0.219 , 0.269	Depositor DCC
$R_{free}$ test set	2001 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.6	Xtriage
Anisotropy	0.567	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 38.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6865	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.77% 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, 0.333 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: NA, IOD

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	0.43	0/2320	0.68	0/3153
2	B	0.42	0/852	0.65	0/1152
3	C	0.44	0/104	0.54	0/142
4	D	0.43	0/1684	0.71	0/2287
5	E	0.41	0/1939	0.70	0/2638
All	All	0.42	0/6899	0.69	0/9372

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	2257	0	2121	11	0
2	B	829	0	796	7	0
3	C	101	0	102	1	0
4	D	1646	0	1555	14	0
5	E	1889	0	1789	14	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
6	E	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	3	0	0	1	0
7	C	1	0	0	0	0
7	D	3	0	0	2	0
7	E	1	0	0	0	0
8	A	37	0	0	0	0
8	B	18	0	0	0	0
8	C	3	0	0	0	0
8	D	36	0	0	0	0
8	E	38	0	0	0	0
All	All	6865	0	6363	42	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 4.

All (42) 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)
4:D:120[A]:ILE:HD11	7:D:302:IOD:I	2.37	0.94
4:D:161:SER:HB2	7:D:303:IOD:I	2.63	0.69
4:D:65:ASN:HD21	4:D:80:PHE:H	1.47	0.60
5:E:19:MET:SD	5:E:120:LEU:HD11	2.43	0.58
4:D:181:LEU:HB3	5:E:180:CYS:HB2	1.85	0.58
4:D:40:PHE:HB2	4:D:106:ALA:HB3	1.86	0.58
2:B:47:GLU:O	2:B:48:LYS:HG2	2.04	0.57
2:B:17:ASN:ND2	2:B:97:ARG:HH12	2.03	0.57
5:E:12:VAL:HG11	5:E:226:LEU:HD12	1.86	0.56
4:D:41:TRP:HB2	4:D:54:ILE:HG22	1.89	0.54
1:A:35:ARG:HG2	1:A:48:ARG:HD3	1.91	0.52
4:D:185:SER:O	4:D:186:MET:HB2	2.09	0.52
4:D:131:PRO:HG2	4:D:133:ILE:HD11	1.94	0.50
4:D:10:GLU:HG3	4:D:126:ARG:HB3	1.94	0.49
1:A:6:ARG:HD2	1:A:98:MET:SD	2.52	0.49
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.47	0.49
1:A:195:SER:OG	1:A:196:ASP:N	2.44	0.49
4:D:65:ASN:ND2	4:D:80:PHE:H	2.10	0.48
1:A:111:ARG:HE	1:A:111:ARG:HB2	1.53	0.47
1:A:138:THR:O	1:A:139:ALA:HB3	2.14	0.47
2:B:24:ASN:HB3	2:B:65:LEU:HD11	1.96	0.46
3:C:9:GLN:HG3	5:E:28:ASN:HD22	1.81	0.46
4:D:218:PHE:CE2	4:D:220:PRO:HG3	2.51	0.46
1:A:5:MET:HB2	1:A:168:LEU:HD13	1.98	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:192:LEU:HD23	5:E:194:ASP:H	1.82	0.44
1:A:35:ARG:NH1	2:B:53:ASP:OD1	2.49	0.44
5:E:129:VAL:HG11	5:E:226:LEU:HD13	2.00	0.43
5:E:205:VAL:HG13	5:E:209:PHE:HB3	1.99	0.43
4:D:13:VAL:HG13	4:D:129:VAL:HB	2.00	0.43
5:E:220:GLN:HG3	5:E:243:ILE:HG23	2.00	0.43
5:E:110:THR:HG23	5:E:111:GLU:HG2	2.00	0.43
1:A:203:CYS:CB	1:A:259:CYS:SG	3.06	0.42
5:E:21:LEU:HD22	5:E:118:THR:HG21	2.01	0.42
4:D:176:THR:HG23	4:D:196:TRP:HZ3	1.85	0.41
1:A:230:LEU:HD23	1:A:243:LYS:HE3	2.02	0.41
2:B:17:ASN:HD21	2:B:97:ARG:HH12	1.66	0.41
5:E:10:PHE:CE1	5:E:119:ARG:HG3	2.56	0.41
1:A:68:LYS:HG2	7:A:304:IOD:I	2.91	0.40
5:E:187:LYS:HD2	5:E:190:PRO:HA	2.03	0.40
5:E:232:TRP:CE2	5:E:234:GLN:HB2	2.56	0.40
2:B:96:ASP:HB3	2:B:99:MET:HG3	2.03	0.40
4:D:176:THR:HG22	5:E:186:LEU:HD11	2.02	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	274/276 (99%)	262 (96%)	10 (4%)	2 (1%)	25	30
2	B	97/99 (98%)	95 (98%)	2 (2%)	0	100	100
3	C	11/13 (85%)	10 (91%)	1 (9%)	0	100	100
4	D	203/204 (100%)	194 (96%)	7 (3%)	2 (1%)	18	20
5	E	236/238 (99%)	231 (98%)	4 (2%)	1 (0%)	38	47
All	All	821/830 (99%)	792 (96%)	24 (3%)	5 (1%)	28	34

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	186	MET
4	D	64	GLN
1	A	139	ALA
5	E	162	ASP
1	A	86	ASN

### 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	234/234 (100%)	214 (92%)	20 (8%)	12	15
2	B	94/94 (100%)	88 (94%)	6 (6%)	20	27
3	C	11/11 (100%)	11 (100%)	0	100	100
4	D	190/189 (100%)	170 (90%)	20 (10%)	8	9
5	E	206/206 (100%)	189 (92%)	17 (8%)	13	16
All	All	735/734 (100%)	672 (91%)	63 (9%)	13	14

All (63) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	ARG
1	A	35	ARG
1	A	39	ASP
1	A	42	SER
1	A	58	GLU
1	A	88	SER
1	A	89	GLU
1	A	103	LEU
1	A	111	ARG
1	A	130	LEU
1	A	131	SER
1	A	151	ARG
1	A	179	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	180	GLN
1	A	186	LYS
1	A	196	ASP
1	A	223	ASP
1	A	230	LEU
1	A	268	LYS
1	A	272	LEU
2	B	1	ILE
2	B	47	GLU
2	B	48	LYS
2	B	70	PHE
2	B	75	LYS
2	B	88	SER
4	D	2	GLN
4	D	10	GLU
4	D	26	GLU
4	D	63	GLU
4	D	95	SER
4	D	107	LEU
4	D	120[A]	ILE
4	D	120[B]	ILE
4	D	133	ILE
4	D	148	SER
4	D	151	LYS
4	D	172	ASP
4	D	178	LYS
4	D	179	CYS
4	D	186	MET
4	D	187	ASP
4	D	200	SER
4	D	212	ILE
4	D	215	GLN
4	D	216	ASP
5	E	14	LYS
5	E	26	ASP
5	E	58	SER
5	E	63	GLU
5	E	79	SER
5	E	84	LYS
5	E	119	ARG
5	E	120	LEU
5	E	145	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	E	189	GLN
5	E	192	LEU
5	E	193	ASN
5	E	202	ARG
5	E	205	VAL
5	E	214	ARG
5	E	249	TRP
5	E	253	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	17	ASN
4	D	2	GLN
4	D	65	ASN
5	E	28	ASN
5	E	128	ASN

### 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](#)

Of 11 ligands modelled in this entry, 11 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, 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	276/276 (100%)	0.08	10 (3%) 43 50	23, 39, 59, 73	43 (15%)
2	B	99/99 (100%)	-0.11	1 (1%) 82 86	25, 37, 54, 59	14 (14%)
3	C	13/13 (100%)	-0.29	0 100 100	28, 32, 38, 40	0
4	D	204/204 (100%)	0.19	9 (4%) 35 42	24, 39, 59, 75	32 (15%)
5	E	238/238 (100%)	0.01	3 (1%) 77 81	20, 36, 53, 96	26 (10%)
All	All	830/830 (100%)	0.06	23 (2%) 53 61	20, 38, 58, 96	115 (13%)

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	107	GLY	5.6
4	D	185	SER	5.5
4	D	148	SER	5.5
1	A	106	ASP	4.5
4	D	64	GLN	4.0
5	E	191	ALA	3.8
4	D	2	GLN	3.7
1	A	41	ALA	3.0
5	E	253	ASP	3.0
4	D	46	PRO	2.7
1	A	40	ALA	2.6
1	A	34	VAL	2.6
1	A	197	HIS	2.5
4	D	147	LYS	2.5
1	A	196	ASP	2.5
1	A	90	ALA	2.4
4	D	216	ASP	2.3
4	D	28	ARG	2.2
5	E	46	PRO	2.2
1	A	44	ARG	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	25	CYS	2.1
4	D	221	SER	2.0
1	A	104	GLY	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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
6	NA	A	301	1/1	0.90	0.13	-0.13	49,49,49,49	0
7	IOD	D	303	1/1	0.91	0.11	-0.85	128,128,128,128	0
7	IOD	D	304	1/1	0.99	0.05	-1.43	64,64,64,64	1
7	IOD	C	101	1/1	0.98	0.07	-2.94	38,38,38,38	1
6	NA	D	301	1/1	0.97	0.07	-3.79	27,27,27,27	0
7	IOD	A	302	1/1	0.98	0.05	-3.83	37,37,37,37	0
7	IOD	D	302	1/1	0.98	0.06	-3.93	67,67,67,67	1
7	IOD	A	304	1/1	0.99	0.03	-	56,56,56,56	0
7	IOD	A	303	1/1	0.99	0.03	-	44,44,44,44	0
7	IOD	E	302	1/1	0.93	0.31	-	109,109,109,109	1
6	NA	E	301	1/1	0.96	0.14	-	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.