



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 28, 2020 – 12:18 PM EDT

PDB ID : 5ULY  
Title : Crystal Structure of Human Protocadherin-15 EC2-3  
Authors : Wimalasena, L.N.; Sotomayor, M.  
Deposited on : 2017-01-25  
Resolution : 2.64 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.14.6  
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) : 2.14.6

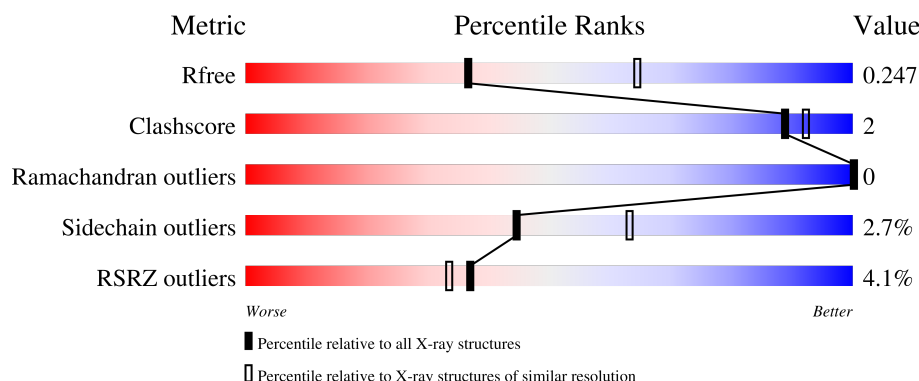
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.64 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	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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 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	259	<div> <div>2%</div> <div> <div></div> <div>78%</div> <div>7%</div> <div>14%</div> </div> </div>
1	B	259	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>7%</div> <div>17%</div> </div> </div>
1	C	259	<div> <div>4%</div> <div> <div></div> <div>80%</div> <div>5%</div> <div>15%</div> </div> </div>
1	D	259	<div> <div>7%</div> <div> <div></div> <div>81%</div> <div>8%</div> <div>10%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7274 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 Protocadherin-15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	224	Total	C	N	O	S	0	0	0
			1785	1140	295	345	5			
1	B	215	Total	C	N	O	S	0	0	0
			1722	1102	284	331	5			
1	C	221	Total	C	N	O	S	0	0	0
			1764	1128	291	340	5			
1	D	233	Total	C	N	O	S	0	0	0
			1864	1184	315	360	5			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	119	MET	-	initiating methionine	UNP Q96QU1
A	120	ALA	-	expression tag	UNP Q96QU1
A	121	SER	-	expression tag	UNP Q96QU1
A	370	LEU	-	expression tag	UNP Q96QU1
A	371	GLU	-	expression tag	UNP Q96QU1
A	372	HIS	-	expression tag	UNP Q96QU1
A	373	HIS	-	expression tag	UNP Q96QU1
A	374	HIS	-	expression tag	UNP Q96QU1
A	375	HIS	-	expression tag	UNP Q96QU1
A	376	HIS	-	expression tag	UNP Q96QU1
A	377	HIS	-	expression tag	UNP Q96QU1
B	119	MET	-	initiating methionine	UNP Q96QU1
B	120	ALA	-	expression tag	UNP Q96QU1
B	121	SER	-	expression tag	UNP Q96QU1
B	370	LEU	-	expression tag	UNP Q96QU1
B	371	GLU	-	expression tag	UNP Q96QU1
B	372	HIS	-	expression tag	UNP Q96QU1
B	373	HIS	-	expression tag	UNP Q96QU1
B	374	HIS	-	expression tag	UNP Q96QU1
B	375	HIS	-	expression tag	UNP Q96QU1
B	376	HIS	-	expression tag	UNP Q96QU1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	377	HIS	-	expression tag	UNP Q96QU1
C	119	MET	-	initiating methionine	UNP Q96QU1
C	120	ALA	-	expression tag	UNP Q96QU1
C	121	SER	-	expression tag	UNP Q96QU1
C	370	LEU	-	expression tag	UNP Q96QU1
C	371	GLU	-	expression tag	UNP Q96QU1
C	372	HIS	-	expression tag	UNP Q96QU1
C	373	HIS	-	expression tag	UNP Q96QU1
C	374	HIS	-	expression tag	UNP Q96QU1
C	375	HIS	-	expression tag	UNP Q96QU1
C	376	HIS	-	expression tag	UNP Q96QU1
C	377	HIS	-	expression tag	UNP Q96QU1
D	119	MET	-	initiating methionine	UNP Q96QU1
D	120	ALA	-	expression tag	UNP Q96QU1
D	121	SER	-	expression tag	UNP Q96QU1
D	370	LEU	-	expression tag	UNP Q96QU1
D	371	GLU	-	expression tag	UNP Q96QU1
D	372	HIS	-	expression tag	UNP Q96QU1
D	373	HIS	-	expression tag	UNP Q96QU1
D	374	HIS	-	expression tag	UNP Q96QU1
D	375	HIS	-	expression tag	UNP Q96QU1
D	376	HIS	-	expression tag	UNP Q96QU1
D	377	HIS	-	expression tag	UNP Q96QU1

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Ca 2 2	0	0
2	A	2	Total Ca 2 2	0	0
2	D	2	Total Ca 2 2	0	0
2	C	2	Total Ca 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	37	Total O 37 37	0	0

*Continued on next page...*

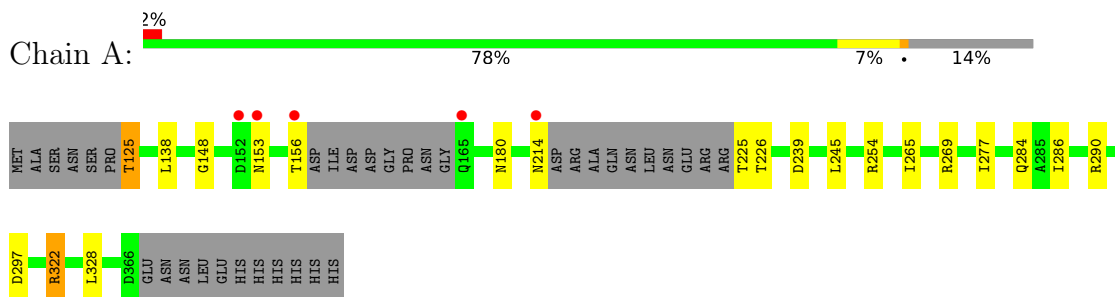
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	32	Total 32	O 32	0	0
3	C	30	Total 30	O 30	0	0
3	D	32	Total 32	O 32	0	0

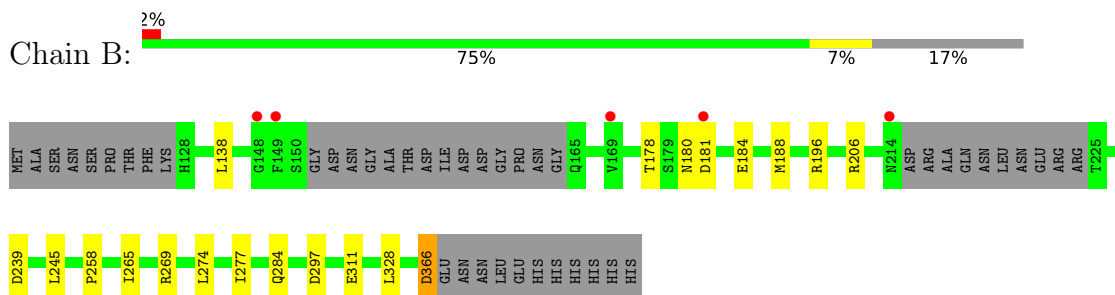
### 3 Residue-property plots [i](#)

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.

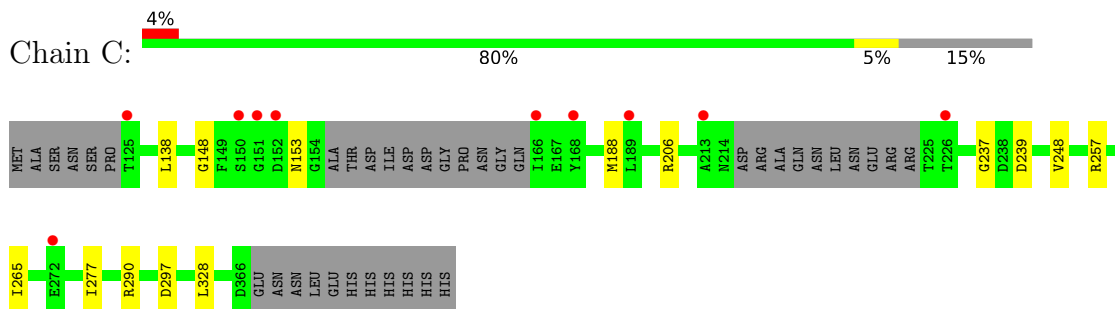
- Molecule 1: Protocadherin-15



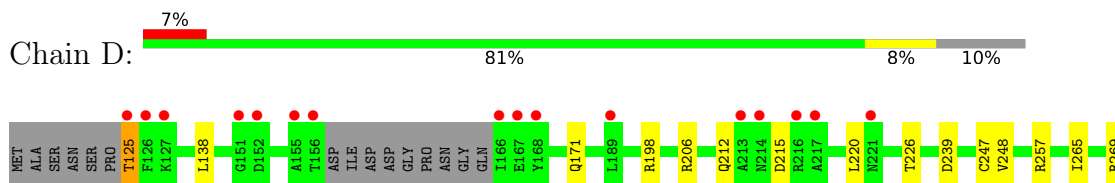
- Molecule 1: Protocadherin-15

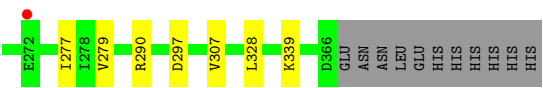


- Molecule 1: Protocadherin-15



- Molecule 1: Protocadherin-15





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.76 Å 147.62 Å 77.03 Å 90.00° 101.16° 90.00°	Depositor
Resolution (Å)	50.00 – 2.64 36.91 – 2.63	Depositor EDS
% Data completeness (in resolution range)	96.1 (50.00-2.64) 96.1 (36.91-2.63)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 2.65 Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.215 , 0.246 0.217 , 0.247	Depositor DCC
$R_{free}$ test set	1962 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.7	Xtriage
Anisotropy	0.578	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 30.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7274	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.06% 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

### 5.1 Standard geometry

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

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	0.75	0/1830	0.88	3/2505 (0.1%)
1	B	0.79	1/1766 (0.1%)	0.89	4/2419 (0.2%)
1	C	0.67	0/1809	0.86	4/2476 (0.2%)
1	D	0.68	0/1910	0.87	3/2613 (0.1%)
All	All	0.72	1/7315 (0.0%)	0.87	14/10013 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	311	GLU	CG-CD	5.02	1.59	1.51

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	290	ARG	NE-CZ-NH2	7.96	124.28	120.30
1	A	322	ARG	NE-CZ-NH1	-7.76	116.42	120.30
1	D	198	ARG	NE-CZ-NH1	-7.13	116.74	120.30
1	B	206	ARG	NE-CZ-NH1	-6.87	116.87	120.30
1	D	206	ARG	NE-CZ-NH1	-5.94	117.33	120.30
1	C	206	ARG	NE-CZ-NH2	5.75	123.18	120.30
1	B	181	ASP	CB-CG-OD1	5.75	123.47	118.30
1	B	206	ARG	NE-CZ-NH2	5.60	123.10	120.30
1	D	257	ARG	NE-CZ-NH1	5.50	123.05	120.30
1	B	366	ASP	N-CA-CB	5.48	120.47	110.60
1	C	188	MET	CG-SD-CE	5.40	108.83	100.20
1	C	206	ARG	NE-CZ-NH1	-5.37	117.62	120.30
1	A	254	ARG	NE-CZ-NH2	5.08	122.84	120.30
1	A	290	ARG	NE-CZ-NH2	5.07	122.83	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.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1785	0	1741	10	0
1	B	1722	0	1684	6	0
1	C	1764	0	1721	6	0
1	D	1864	0	1819	10	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	37	0	0	0	0
3	B	32	0	0	0	0
3	C	30	0	0	1	0
3	D	32	0	0	0	0
All	All	7274	0	6965	27	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 (27) 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:B:184:GLU:HG3	1:B:196:ARG:HD2	1.82	0.61
1:A:125:THR:N	1:A:226:THR:HG1	1.99	0.61
1:D:277:ILE:HD11	1:D:328:LEU:HB2	1.83	0.60
1:D:125:THR:N	1:D:226:THR:HG1	1.99	0.60
1:C:277:ILE:HD11	1:C:328:LEU:HB2	1.83	0.60
1:C:138:LEU:HD21	1:C:239:ASP:HB3	1.84	0.60
1:D:138:LEU:HD21	1:D:239:ASP:HB3	1.84	0.59
1:A:277:ILE:HD11	1:A:328:LEU:HB2	1.84	0.59
1:B:277:ILE:HD11	1:B:328:LEU:HB2	1.85	0.58
1:A:245:LEU:HD11	1:D:248:VAL:HG22	1.84	0.58
1:A:284:GLN:OE1	1:D:247:CYS:N	2.26	0.58
1:D:171:GLN:OE1	1:D:212:GLN:HB2	2.03	0.58

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:138:LEU:HD21	1:B:239:ASP:HB3	1.87	0.56
1:A:138:LEU:HD21	1:A:239:ASP:HB3	1.87	0.56
1:D:265:ILE:HD11	1:D:277:ILE:HD13	1.90	0.53
1:C:265:ILE:HD11	1:C:277:ILE:HD13	1.89	0.53
1:A:265:ILE:HD11	1:A:277:ILE:HD13	1.91	0.51
1:B:265:ILE:HD11	1:B:277:ILE:HD13	1.94	0.48
1:A:322:ARG:NH1	1:D:279:VAL:O	2.44	0.48
1:C:248:VAL:HG13	1:C:257:ARG:HB3	1.95	0.48
1:A:245:LEU:CD1	1:D:248:VAL:HG22	2.44	0.47
1:A:286:ILE:HD13	1:D:248:VAL:HG21	1.97	0.45
1:B:245:LEU:HB2	1:B:284:GLN:HG3	1.99	0.43
1:A:148:GLY:O	1:A:153:ASN:HB3	2.19	0.42
1:C:237:GLY:HA2	3:C:1106:HOH:O	2.21	0.41
1:B:269:ARG:HB2	1:B:274:LEU:HD11	2.03	0.40
1:C:148:GLY:O	1:C:153:ASN:HB3	2.21	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	218/259 (84%)	215 (99%)	3 (1%)	0	100	100
1	B	209/259 (81%)	208 (100%)	1 (0%)	0	100	100
1	C	215/259 (83%)	212 (99%)	3 (1%)	0	100	100
1	D	229/259 (88%)	225 (98%)	4 (2%)	0	100	100
All	All	871/1036 (84%)	860 (99%)	11 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	203/234 (87%)	196 (97%)	7 (3%)	37	53
1	B	197/234 (84%)	191 (97%)	6 (3%)	41	59
1	C	201/234 (86%)	200 (100%)	1 (0%)	88	94
1	D	211/234 (90%)	203 (96%)	8 (4%)	33	50
All	All	812/936 (87%)	790 (97%)	22 (3%)	44	63

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	125	THR
1	A	156	THR
1	A	180	ASN
1	A	214	ASN
1	A	225	THR
1	A	269	ARG
1	A	297	ASP
1	B	178	THR
1	B	180	ASN
1	B	188	MET
1	B	258	PRO
1	B	297	ASP
1	B	366	ASP
1	C	297	ASP
1	D	125	THR
1	D	215	ASP
1	D	220	LEU
1	D	269	ARG
1	D	290	ARG
1	D	297	ASP
1	D	307	VAL
1	D	339	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	180	ASN
1	A	212	GLN
1	B	180	ASN
1	B	214	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 8 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	224/259 (86%)	-0.14	5 (2%) 62 58	32, 48, 89, 131	0
1	B	215/259 (83%)	-0.00	5 (2%) 60 57	31, 49, 82, 97	0
1	C	221/259 (85%)	0.02	10 (4%) 33 30	32, 52, 94, 118	0
1	D	233/259 (89%)	0.17	17 (7%) 15 12	32, 56, 113, 130	0
All	All	893/1036 (86%)	0.01	37 (4%) 37 33	31, 51, 96, 131	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	156	THR	7.1
1	C	166	ILE	5.2
1	D	216	ARG	4.4
1	D	155	ALA	4.3
1	D	166	ILE	4.1
1	A	152	ASP	4.1
1	A	156	THR	3.9
1	A	153	ASN	3.8
1	D	221	ASN	3.6
1	D	217	ALA	3.6
1	A	214	ASN	3.4
1	D	152	ASP	3.2
1	B	214	ASN	3.2
1	D	272	GLU	3.0
1	C	152	ASP	2.9
1	D	127	LYS	2.8
1	C	168	TYR	2.7
1	D	125	THR	2.7
1	C	150	SER	2.7
1	B	181	ASP	2.7
1	C	272	GLU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	189	LEU	2.7
1	C	151	GLY	2.6
1	D	126	PHE	2.6
1	D	167	GLU	2.6
1	C	213	ALA	2.5
1	D	151	GLY	2.5
1	C	189	LEU	2.4
1	C	226	THR	2.4
1	A	165	GLN	2.3
1	B	149	PHE	2.3
1	D	213	ALA	2.2
1	D	168	TYR	2.2
1	C	125	THR	2.2
1	B	169	VAL	2.2
1	D	214	ASN	2.1
1	B	148	GLY	2.1

## 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<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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	A	1002	1/1	0.97	0.17	46,46,46,46	0
2	CA	B	1002	1/1	0.98	0.18	38,38,38,38	0
2	CA	C	1001	1/1	0.99	0.15	43,43,43,43	0
2	CA	D	1002	1/1	0.99	0.10	39,39,39,39	0
2	CA	C	1002	1/1	0.99	0.16	39,39,39,39	0
2	CA	D	1001	1/1	0.99	0.12	39,39,39,39	0
2	CA	B	1001	1/1	0.99	0.17	37,37,37,37	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	A	1001	1/1	1.00	0.13	35,35,35,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.