



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 30, 2024 – 04:15 PM JST

PDB ID : 9JPJ  
Title : Crystal structure of DhdR in complex with DNA  
Authors : Sun, P.; Wang, B.  
Deposited on : 2024-09-26  
Resolution : 3.72 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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
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

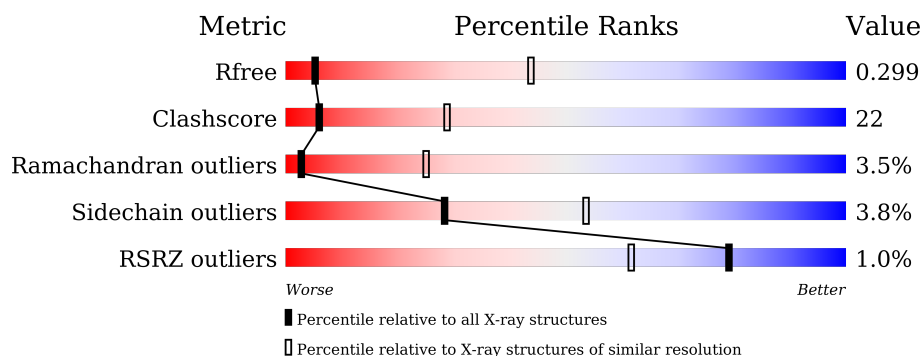
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.72 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))
$R_{free}$	164625	1058 (3.84-3.60)
Clashscore	180529	1114 (3.84-3.60)
Ramachandran outliers	177936	1095 (3.84-3.60)
Sidechain outliers	177891	1091 (3.84-3.60)
RSRZ outliers	164620	1058 (3.84-3.60)

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	27	<div> <div>41%</div> <div>56%</div> <div>.</div> </div>
1	G	27	<div> <div>19%</div> <div>78%</div> <div>.</div> </div>
2	B	27	<div> <div>26%</div> <div>70%</div> <div>.</div> </div>
2	H	27	<div> <div>41%</div> <div>48%</div> <div>11%</div> </div>
3	C	238	<div> <div>2%</div> <div>54%</div> <div>34%</div> <div>10%</div> </div>
3	D	238	<div> <div>52%</div> <div>33%</div> <div>5%</div> <div>11%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	238	<div><div><div></div><div></div><div></div></div><div>%</div><div><div></div><div></div><div></div></div><div>52%35%11%</div></div>
3	F	238	<div><div><div></div><div></div><div></div></div><div>%</div><div><div></div><div></div><div></div></div><div>52%34%12%</div></div>
3	I	238	<div><div><div></div><div></div><div></div></div><div>2%</div><div><div></div><div></div><div></div></div><div>48%39%11%</div></div>
3	J	238	<div><div><div></div><div></div><div></div></div><div></div><div><div></div><div></div><div></div></div><div>51%35%11%</div></div>
3	K	238	<div><div><div></div><div></div><div></div></div><div>%</div><div><div></div><div></div><div></div></div><div>54%35%7%</div></div>
3	L	238	<div><div><div></div><div></div><div></div></div><div></div><div><div></div><div></div><div></div></div><div>47%38%13%</div></div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 14339 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 DNA chain called DNA (27-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	27	Total	C	N	O	P	0	0	0
			557	267	111	153	26			
1	A	27	Total	C	N	O	P	0	0	0
			557	267	111	153	26			

- Molecule 2 is a DNA chain called DNA (27-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	27	Total	C	N	O	P	0	0	0
			548	265	86	170	27			
2	B	27	Total	C	N	O	P	0	0	0
			548	265	86	170	27			

- Molecule 3 is a protein called Pyruvate dehydrogenase complex repressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	215	Total	C	N	O	S	0	0	0
			1500	928	285	285	2			
3	D	213	Total	C	N	O	S	0	0	0
			1509	926	292	289	2			
3	E	213	Total	C	N	O	S	0	0	0
			1512	932	289	289	2			
3	F	209	Total	C	N	O	S	0	0	0
			1514	935	288	289	2			
3	I	212	Total	C	N	O	S	0	0	0
			1484	917	283	281	3			
3	J	212	Total	C	N	O	S	0	0	0
			1516	935	289	289	3			
3	K	221	Total	C	N	O	S	0	0	0
			1579	971	309	296	3			
3	L	208	Total	C	N	O	S	0	0	0
			1507	923	290	292	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
C	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
D	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
D	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
E	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
E	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
F	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
F	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
I	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
I	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
J	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
J	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
K	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
K	134	ASP	GLU	conflict	UNP A0A6N0JVZ6
L	120	ALA	VAL	conflict	UNP A0A6N0JVZ6
L	134	ASP	GLU	conflict	UNP A0A6N0JVZ6

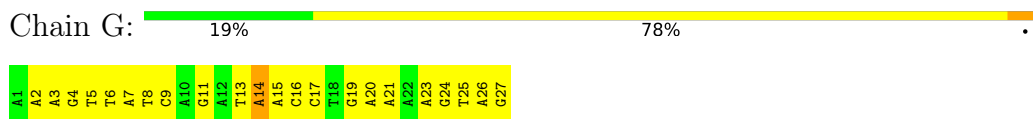
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total Zn 1 1	0	0
4	D	1	Total Zn 1 1	0	0
4	E	1	Total Zn 1 1	0	0
4	F	1	Total Zn 1 1	0	0
4	I	1	Total Zn 1 1	0	0
4	J	1	Total Zn 1 1	0	0
4	K	1	Total Zn 1 1	0	0
4	L	1	Total Zn 1 1	0	0

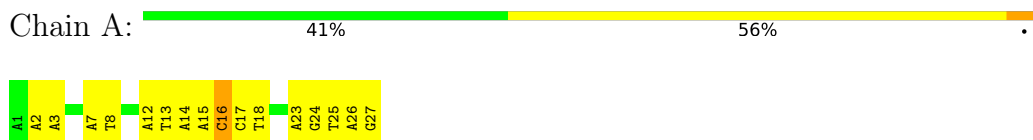
### 3 Residue-property plots

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.

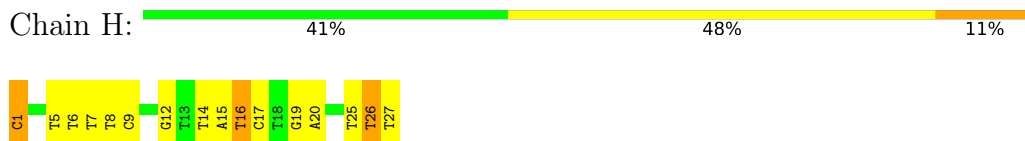
- Molecule 1: DNA (27-MER)



- Molecule 1: DNA (27-MER)



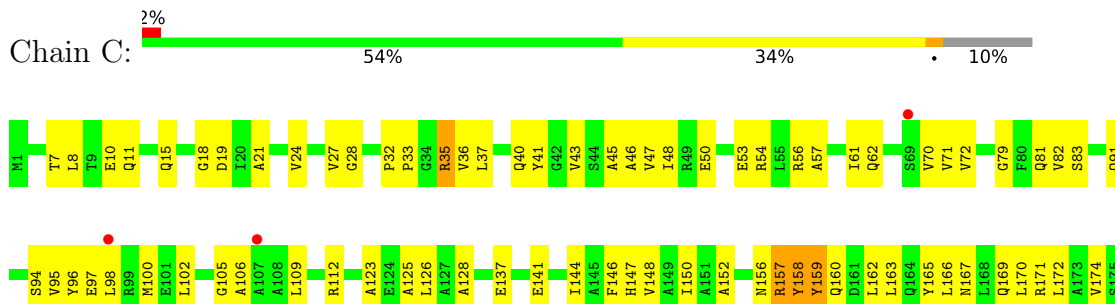
- Molecule 2: DNA (27-MER)



- Molecule 2: DNA (27-MER)



- Molecule 3: Pyruvate dehydrogenase complex repressor

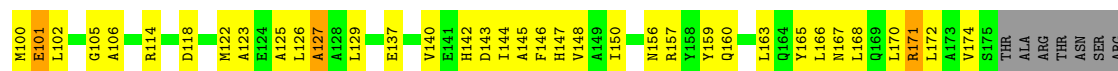




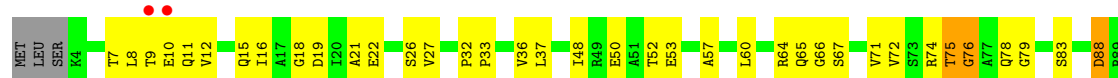
• Molecule 3: Pyruvate dehydrogenase complex repressor



• Molecule 3: Pyruvate dehydrogenase complex repressor

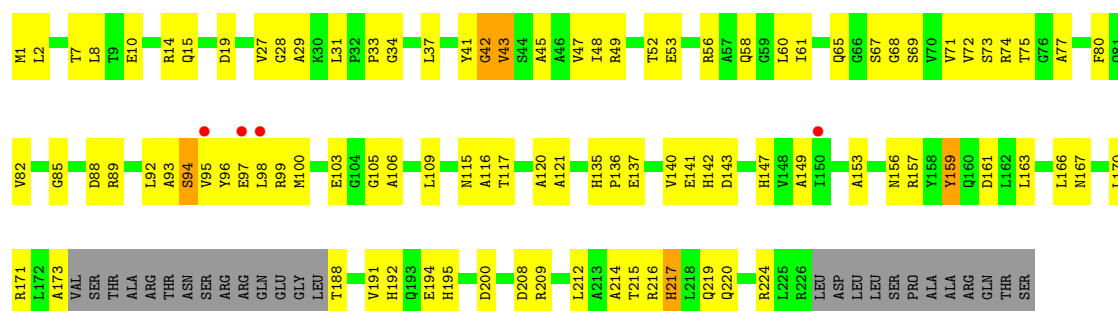


• Molecule 3: Pyruvate dehydrogenase complex repressor



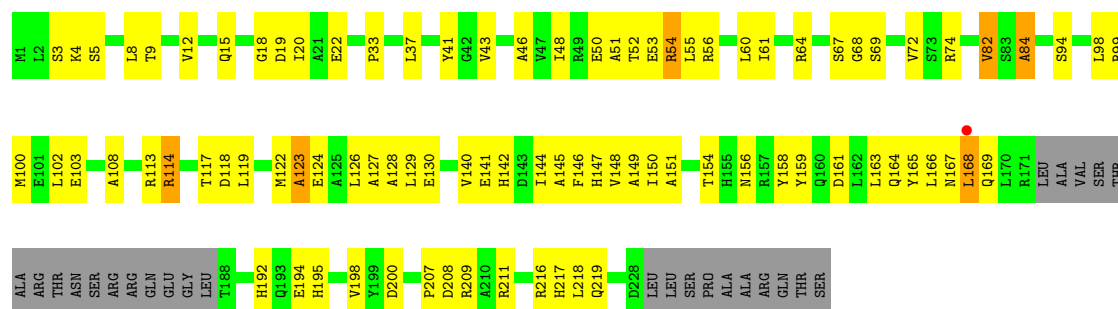
• Molecule 3: Pyruvate dehydrogenase complex repressor



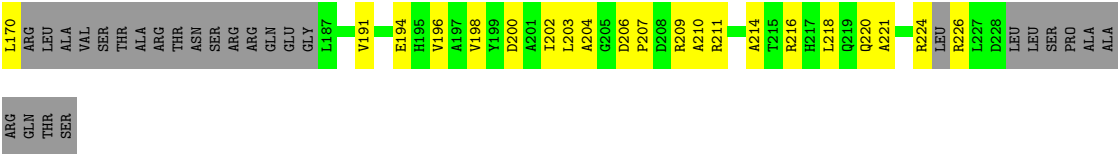


• Molecule 3: Pyruvate dehydrogenase complex repressor

Chain J: 51% 35% 11%







ARG  
GLN  
THR  
SER

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.46Å 87.04Å 133.88Å 91.48° 102.48° 117.55°	Depositor
Resolution (Å)	39.26 – 3.72 39.26 – 3.72	Depositor EDS
% Data completeness (in resolution range)	93.0 (39.26-3.72) 93.3 (39.26-3.72)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.59 (at 3.77Å)	Xtriage
Refinement program	PHENIX 1.14-3260	Depositor
R, $R_{free}$	0.250 , 0.299 0.250 , 0.299	Depositor DCC
$R_{free}$ test set	1569 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	110.3	Xtriage
Anisotropy	0.723	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 161.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	14339	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	162.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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: ZN

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	0.76	0/628	1.04	4/968 (0.4%)
1	G	0.79	1/628 (0.2%)	1.03	2/968 (0.2%)
2	B	0.92	1/610 (0.2%)	1.20	1/937 (0.1%)
2	H	0.88	1/610 (0.2%)	1.21	2/937 (0.2%)
3	C	0.39	0/1517	0.57	0/2069
3	D	0.36	0/1524	0.65	2/2072 (0.1%)
3	E	0.27	0/1528	0.49	0/2078
3	F	0.27	0/1531	0.48	0/2081
3	I	0.28	0/1499	0.50	0/2040
3	J	0.28	0/1533	0.50	0/2084
3	K	0.27	0/1596	0.50	0/2168
3	L	0.28	0/1521	0.51	0/2065
All	All	0.44	3/14725 (0.0%)	0.68	11/20467 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	DC	OP3-P	-10.64	1.48	1.61
2	H	1	DC	OP3-P	-10.46	1.48	1.61
1	G	14	DA	C3'-O3'	-5.17	1.37	1.44

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	218	LEU	CB-CG-CD2	-13.79	87.55	111.00
2	H	26	DT	O4'-C4'-C3'	-5.97	102.11	104.50
3	D	92	LEU	CA-CB-CG	5.62	128.23	115.30
1	G	17	DC	C1'-O4'-C4'	-5.57	104.53	110.10
1	A	12	DA	O4'-C4'-C3'	-5.35	102.36	104.50

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	557	0	305	19	0
1	G	557	0	305	36	0
2	B	548	0	311	29	0
2	H	548	0	311	23	0
3	C	1500	0	1417	77	0
3	D	1509	0	1442	98	1
3	E	1512	0	1456	72	0
3	F	1514	0	1473	66	0
3	I	1484	0	1427	78	0
3	J	1516	0	1454	79	0
3	K	1579	0	1524	67	1
3	L	1507	0	1457	78	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	K	1	0	0	0	0
4	L	1	0	0	0	0
All	All	14339	0	12882	600	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 22.

The worst 5 of 600 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)
3:D:217:HIS:CD2	3:D:218:LEU:HD12	1.69	1.26
3:D:217:HIS:HD2	3:D:218:LEU:CD1	1.76	0.99
3:D:169:GLN:OE1	3:D:170:LEU:O	1.90	0.89
3:J:20:ILE:HG22	3:J:74:ARG:HA	1.59	0.81
3:F:21:ALA:O	3:F:74:ARG:NH2	2.13	0.81

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:D:114:ARG:NH1	3:K:205:GLY:O[1_556]	2.13	0.07

## 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	
3	C	211/238 (89%)	180 (85%)	28 (13%)	3 (1%)	9	39
3	D	209/238 (88%)	173 (83%)	28 (13%)	8 (4%)	2	23
3	E	209/238 (88%)	173 (83%)	28 (13%)	8 (4%)	2	23
3	F	205/238 (86%)	172 (84%)	26 (13%)	7 (3%)	3	26
3	I	208/238 (87%)	165 (79%)	35 (17%)	8 (4%)	2	23
3	J	208/238 (87%)	169 (81%)	32 (15%)	7 (3%)	3	26
3	K	217/238 (91%)	178 (82%)	27 (12%)	12 (6%)	1	17
3	L	202/238 (85%)	166 (82%)	31 (15%)	5 (2%)	4	30
All	All	1669/1904 (88%)	1376 (82%)	235 (14%)	58 (4%)	3	25

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	82	VAL
3	C	137	GLU
3	D	4	LYS
3	E	82	VAL
3	I	43	VAL

### 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	
3	C	129/179 (72%)	125 (97%)	4 (3%)	35	58
3	D	133/179 (74%)	124 (93%)	9 (7%)	13	40
3	E	135/179 (75%)	129 (96%)	6 (4%)	24	50
3	F	140/179 (78%)	136 (97%)	4 (3%)	37	59
3	I	130/179 (73%)	125 (96%)	5 (4%)	28	53
3	J	135/179 (75%)	133 (98%)	2 (2%)	60	75
3	K	141/179 (79%)	138 (98%)	3 (2%)	48	67
3	L	139/179 (78%)	131 (94%)	8 (6%)	17	44
All	All	1082/1432 (76%)	1041 (96%)	41 (4%)	28	53

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	54	ARG
3	L	114	ARG
3	J	114	ARG
3	K	159	TYR
3	L	157	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	132	ASN
3	I	147	HIS
3	L	142	HIS
3	J	15	GLN
3	E	65	GLN

### 5.3.3 RNA [i](#)

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 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	27/27 (100%)	-0.37	0 100 100	71, 99, 153, 160	0
1	G	27/27 (100%)	-0.33	0 100 100	66, 86, 138, 146	0
2	B	27/27 (100%)	-0.16	0 100 100	71, 99, 158, 166	0
2	H	27/27 (100%)	-0.18	0 100 100	74, 96, 145, 153	0
3	C	215/238 (90%)	-0.03	4 (1%) 66 48	64, 226, 258, 271	0
3	D	213/238 (89%)	-0.05	1 (0%) 87 72	66, 164, 204, 226	0
3	E	213/238 (89%)	-0.12	2 (0%) 81 64	86, 181, 232, 247	0
3	F	209/238 (87%)	-0.27	3 (1%) 73 54	69, 193, 225, 236	0
3	I	212/238 (89%)	-0.07	4 (1%) 66 48	67, 228, 270, 290	0
3	J	212/238 (89%)	-0.06	1 (0%) 87 72	54, 219, 250, 257	0
3	K	221/238 (92%)	-0.13	3 (1%) 73 54	81, 155, 197, 232	0
3	L	208/238 (87%)	-0.22	1 (0%) 87 72	79, 175, 224, 234	0
All	All	1811/2012 (90%)	-0.13	19 (1%) 79 62	54, 170, 256, 290	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	K	98	LEU	3.7
3	F	9	THR	3.2
3	K	97	GLU	3.0
3	I	95	VAL	2.8
3	D	167	ASN	2.8

### 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
4	ZN	C	301	1/1	0.81	0.08	230,230,230,230	0
4	ZN	I	301	1/1	0.82	0.17	261,261,261,261	0
4	ZN	E	301	1/1	0.92	0.11	202,202,202,202	0
4	ZN	L	301	1/1	0.92	0.07	168,168,168,168	0
4	ZN	D	301	1/1	0.96	0.04	191,191,191,191	0
4	ZN	J	301	1/1	0.96	0.06	219,219,219,219	0
4	ZN	F	301	1/1	0.96	0.06	204,204,204,204	0
4	ZN	K	301	1/1	0.98	0.08	152,152,152,152	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.