



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:30 AM GMT

PDB ID : 2R03  
Title : Crystal Structure of ALIX/AIP1 in complex with the YPDL Late Domain  
Authors : Hill, C.P.; Zhai, Q.; Fisher, R.D.  
Deposited on : 2007-08-17  
Resolution : 2.59 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

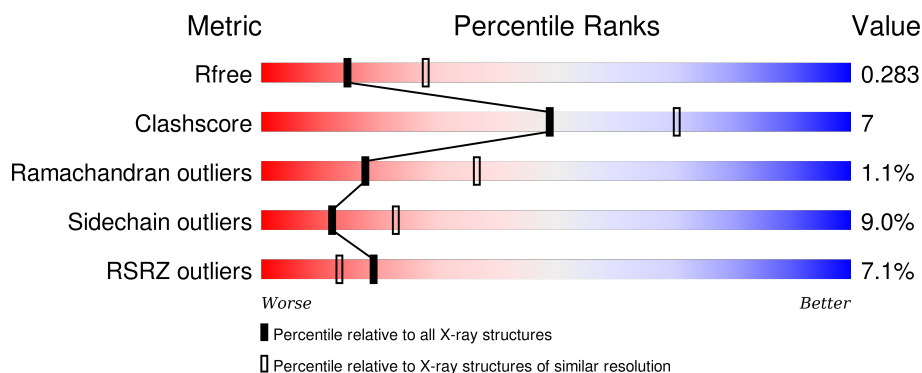
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.59 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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 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	697	<div> <div>7%</div> <div>76%</div> <div>22%</div> <div>.</div> </div>
2	B	8	<div> <div>100%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5575 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 Programmed cell death 6-interacting protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	697	Total	C	N	O	S	0	0	0
			5486	3461	938	1069	18			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	268	TYR	LYS	ENGINEERED	UNP Q8WUM4
A	269	TYR	LYS	ENGINEERED	UNP Q8WUM4

- Molecule 2 is a protein called p6-Gag.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	8	Total	C	N	O	0	0	0
			66	42	9	15			

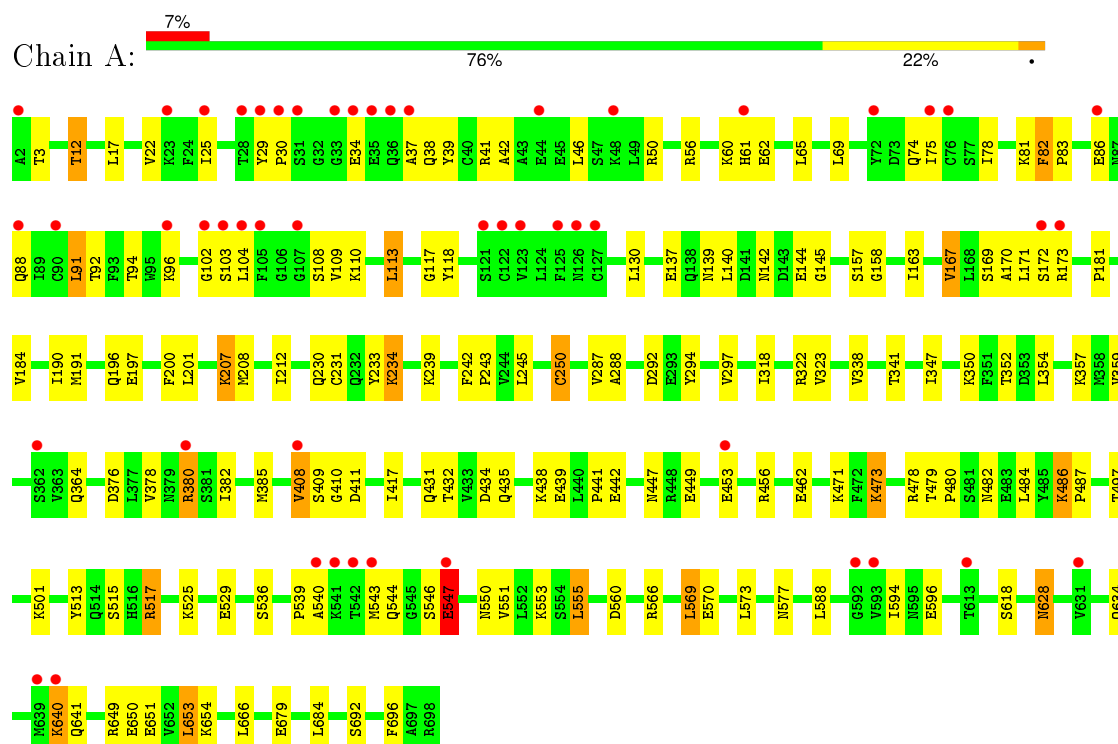
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	23	Total	O	0	0
			23	23		

### 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: Programmed cell death 6-interacting protein



- Molecule 2: p6-Gag



There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.38Å 98.57Å 72.80Å 90.00° 107.07° 90.00°	Depositor
Resolution (Å)	50.00 – 2.59 32.45 – 2.59	Depositor EDS
% Data completeness (in resolution range)	96.1 (50.00-2.59) 96.2 (32.45-2.59)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.56 (at 2.57Å)	Xtriage
Refinement program	REFMAC 5.3.0008	Depositor
R, $R_{free}$	0.223 , 0.286 0.219 , 0.283	Depositor DCC
$R_{free}$ test set	1488 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	62.5	Xtriage
Anisotropy	0.226	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 56.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 29643 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5575	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% 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

### 5.1 Standard geometry

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	0.90	6/5570 (0.1%)	0.69	3/7523 (0.0%)
2	B	0.46	0/67	0.65	0/91
All	All	0.89	6/5637 (0.1%)	0.68	3/7614 (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	A	0	1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	547	GLU	CD-OE1	41.03	1.70	1.25
1	A	547	GLU	CD-OE2	29.45	1.58	1.25
1	A	550	ASN	CG-ND2	6.17	1.48	1.32
1	A	250	CYS	CB-SG	-5.99	1.72	1.81
1	A	550	ASN	CG-OD1	5.70	1.36	1.24
1	A	231	CYS	CB-SG	-5.22	1.73	1.81

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	547	GLU	OE1-CD-OE2	10.72	136.16	123.30
1	A	547	GLU	CG-CD-OE1	-6.87	104.56	118.30
1	A	666	LEU	CA-CB-CG	5.05	126.93	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	547	GLU	Sidechain

## 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	5486	0	5532	80	0
2	B	66	0	58	0	0
3	A	23	0	0	0	0
All	All	5575	0	5590	80	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 7.

All (80) 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:A:547:GLU:OE1	1:A:547:GLU:CD	1.70	1.30
1:A:517:ARG:HH11	1:A:517:ARG:HG3	1.32	0.94
1:A:447:ASN:HD22	1:A:684:LEU:HD12	1.42	0.84
1:A:382:ILE:HD12	1:A:570:GLU:HG3	1.60	0.83
1:A:376:ASP:O	1:A:380:ARG:HB2	1.82	0.80
1:A:525:LYS:HD3	1:A:529:GLU:OE2	1.82	0.79
1:A:486:LYS:HG3	1:A:487:PRO:HD3	1.71	0.73
1:A:190:ILE:HG12	1:A:245:LEU:HD21	1.73	0.69
1:A:74:GLN:O	1:A:78:ILE:HD13	1.93	0.68
1:A:692:SER:O	1:A:696:PHE:HD2	1.78	0.66
1:A:517:ARG:CG	1:A:517:ARG:HH11	2.06	0.65
1:A:517:ARG:NH1	1:A:517:ARG:HG3	2.10	0.65
1:A:163:ILE:O	1:A:167:VAL:HG22	1.99	0.63
1:A:378:VAL:O	1:A:382:ILE:HG12	1.99	0.62
1:A:142:ASN:HD22	1:A:145:GLY:H	1.47	0.62
1:A:25:ILE:HD13	1:A:39:TYR:O	2.02	0.60
1:A:447:ASN:ND2	1:A:684:LEU:HD12	2.17	0.57
1:A:29:TYR:HB3	1:A:30:PRO:HD2	1.85	0.57
1:A:113:LEU:HD11	1:A:173:ARG:HG3	1.85	0.57
1:A:431:GLN:NE2	1:A:435:GLN:HE21	2.03	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:208:MET:HE2	1:A:212:ILE:HG21	1.86	0.56
1:A:3:THR:HG21	1:A:294:TYR:O	2.05	0.56
1:A:3:THR:HG23	1:A:3:THR:O	2.06	0.55
1:A:462:GLU:OE1	1:A:482:ASN:ND2	2.40	0.55
1:A:385:MET:HE1	1:A:569:LEU:HB3	1.87	0.55
1:A:171:LEU:HD23	1:A:173:ARG:H	1.71	0.54
1:A:96:LYS:HG2	1:A:110:LYS:HG2	1.90	0.54
1:A:118:TYR:HD1	1:A:171:LEU:HD12	1.71	0.54
1:A:34:GLU:HA	1:A:37:ALA:HB3	1.89	0.53
1:A:517:ARG:CG	1:A:517:ARG:NH1	2.68	0.53
1:A:547:GLU:OE1	1:A:547:GLU:CG	2.55	0.53
1:A:692:SER:O	1:A:696:PHE:CD2	2.59	0.53
1:A:359:VAL:HB	1:A:364:GLN:HE22	1.73	0.53
1:A:453:GLU:HG2	1:A:456:ARG:HH12	1.74	0.52
1:A:588:LEU:HB2	1:A:594:ILE:HD11	1.92	0.52
1:A:157:SER:HB2	1:A:191:MET:HB2	1.91	0.52
1:A:242:PHE:HB3	1:A:243:PRO:CD	2.40	0.51
1:A:60:LYS:HD3	1:A:137:GLU:OE2	2.10	0.51
1:A:233:TYR:O	1:A:234:LYS:HG3	2.10	0.51
1:A:322:ARG:NH1	1:A:323:VAL:O	2.44	0.51
1:A:61:HIS:CG	1:A:62:GLU:H	2.29	0.51
1:A:242:PHE:HB3	1:A:243:PRO:HD3	1.93	0.51
1:A:540:ALA:HB3	1:A:640:LYS:HE3	1.94	0.50
1:A:409:SER:OG	1:A:411:ASP:HB2	2.12	0.49
1:A:22:VAL:HA	1:A:25:ILE:HD12	1.95	0.49
1:A:158:GLY:HA3	1:A:341:THR:O	2.13	0.48
1:A:434:ASP:OD1	1:A:513:TYR:OH	2.19	0.48
1:A:438:LYS:O	1:A:441:PRO:HD2	2.14	0.48
1:A:431:GLN:HE22	1:A:435:GLN:HE21	1.60	0.47
1:A:410:GLY:O	1:A:536:SER:HA	2.14	0.47
1:A:649:ARG:HG2	1:A:653:LEU:HD22	1.97	0.47
1:A:17:LEU:HD13	1:A:46:LEU:HG	1.97	0.47
1:A:497:THR:O	1:A:501:LYS:HD3	2.14	0.47
1:A:628:ASN:HD22	1:A:628:ASN:C	2.18	0.46
1:A:167:VAL:HG21	1:A:181:PRO:HD3	1.97	0.46
1:A:233:TYR:O	1:A:234:LYS:CG	2.65	0.45
1:A:250:CYS:HB2	1:A:287:VAL:HG21	1.99	0.45
1:A:473:LYS:HD2	1:A:473:LYS:H	1.82	0.45
1:A:102:GLY:C	1:A:104:LEU:H	2.20	0.45
1:A:409:SER:OG	1:A:411:ASP:CB	2.65	0.44
1:A:139:ASN:O	1:A:145:GLY:HA3	2.18	0.44

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:82:PHE:HA	1:A:83:PRO:HD3	1.93	0.44
1:A:196:GLN:O	1:A:197:GLU:C	2.55	0.43
1:A:117:GLY:HA3	1:A:170:ALA:HB1	2.00	0.43
1:A:74:GLN:O	1:A:78:ILE:CD1	2.65	0.43
1:A:12:THR:HG21	1:A:50:ARG:HH22	1.83	0.42
1:A:12:THR:CG2	1:A:94:THR:O	2.67	0.42
1:A:65:LEU:HB2	1:A:130:LEU:HD23	2.01	0.42
1:A:42:ALA:HB1	1:A:354:LEU:HB3	2.01	0.42
1:A:69:LEU:HD22	1:A:347:ILE:HD12	2.02	0.42
1:A:385:MET:HE2	1:A:566:ARG:CB	2.50	0.42
1:A:439:GLU:HA	1:A:442:GLU:OE2	2.19	0.42
1:A:172:SER:O	1:A:173:ARG:HG2	2.20	0.42
1:A:409:SER:HA	1:A:539:PRO:HG3	2.02	0.42
1:A:479:THR:HA	1:A:480:PRO:HD3	1.95	0.41
1:A:75:ILE:HD13	1:A:75:ILE:HA	1.91	0.41
1:A:17:LEU:HD23	1:A:91:LEU:HD21	2.03	0.40
1:A:551:VAL:O	1:A:555:LEU:HB2	2.21	0.40
1:A:288:ALA:O	1:A:292:ASP:HB3	2.22	0.40
1:A:38:GLN:HA	1:A:41:ARG:HD3	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	695/697 (100%)	645 (93%)	42 (6%)	8 (1%)	16	33
2	B	6/8 (75%)	5 (83%)	1 (17%)	0	100	100
All	All	701/705 (99%)	650 (93%)	43 (6%)	8 (1%)	17	36

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	207	LYS
1	A	81	LYS
1	A	543	MET
1	A	546	SER
1	A	230	GLN
1	A	318	ILE
1	A	408	VAL
1	A	82	PHE

### 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	600/600 (100%)	545 (91%)	55 (9%)	11	21
2	B	8/8 (100%)	8 (100%)	0	100	100
All	All	608/608 (100%)	553 (91%)	55 (9%)	12	23

All (55) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	THR
1	A	56	ARG
1	A	86	GLU
1	A	88	GLN
1	A	91	LEU
1	A	92	THR
1	A	103	SER
1	A	108	SER
1	A	109	VAL
1	A	113	LEU
1	A	140	LEU
1	A	144	GLU
1	A	167	VAL
1	A	169	SER
1	A	184	VAL
1	A	200	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	201	LEU
1	A	207	LYS
1	A	234	LYS
1	A	239	LYS
1	A	297	VAL
1	A	338	VAL
1	A	350	LYS
1	A	352	THR
1	A	357	LYS
1	A	380	ARG
1	A	408	VAL
1	A	417	ILE
1	A	432	THR
1	A	449	GLU
1	A	471	LYS
1	A	473	LYS
1	A	478	ARG
1	A	484	LEU
1	A	486	LYS
1	A	515	SER
1	A	517	ARG
1	A	544	GLN
1	A	553	LYS
1	A	555	LEU
1	A	560	ASP
1	A	569	LEU
1	A	573	LEU
1	A	577	ASN
1	A	596	GLU
1	A	618	SER
1	A	628	ASN
1	A	634	GLN
1	A	640	LYS
1	A	641	GLN
1	A	650	GLU
1	A	651	GLU
1	A	653	LEU
1	A	654	LYS
1	A	679	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	GLN
1	A	133	GLN
1	A	142	ASN
1	A	230	GLN
1	A	364	GLN
1	A	384	GLN
1	A	415	GLN
1	A	431	GLN
1	A	447	ASN
1	A	482	ASN
1	A	531	ASN
1	A	571	ASN
1	A	628	ASN
1	A	647	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](#)

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	697/697 (100%)	0.55	50 (7%) 18 13	33, 81, 136, 172	2 (0%)
2	B	8/8 (100%)	0.77	0 100 100	90, 95, 105, 107	0
All	All	705/705 (100%)	0.55	50 (7%) 19 13	33, 82, 135, 172	2 (0%)

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	542	THR	5.6
1	A	102	GLY	4.6
1	A	593	VAL	4.5
1	A	31	SER	4.4
1	A	540	ALA	4.4
1	A	543	MET	4.2
1	A	29	TYR	4.2
1	A	103	SER	3.9
1	A	107	GLY	3.8
1	A	33	GLY	3.7
1	A	105	PHE	3.7
1	A	36	GLN	3.6
1	A	104	LEU	3.6
1	A	592	GLY	3.6
1	A	122	CYS	3.5
1	A	631	VAL	3.4
1	A	90	CYS	3.3
1	A	640	LYS	3.2
1	A	547	GLU	3.2
1	A	28	THR	3.2
1	A	173	ARG	3.0
1	A	88	GLN	3.0
1	A	61	HIS	3.0
1	A	30	PRO	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	127	CYS	2.8
1	A	639	MET	2.8
1	A	2	ALA	2.8
1	A	123	VAL	2.8
1	A	453	GLU	2.7
1	A	408	VAL	2.6
1	A	96	LYS	2.6
1	A	35	GLU	2.5
1	A	86	GLU	2.5
1	A	75	ILE	2.4
1	A	76	CYS	2.4
1	A	48	LYS	2.3
1	A	125	PHE	2.3
1	A	37	ALA	2.3
1	A	34	GLU	2.3
1	A	25	ILE	2.3
1	A	613	THR	2.2
1	A	72	TYR	2.2
1	A	44	GLU	2.2
1	A	121	SER	2.2
1	A	380	ARG	2.2
1	A	126	ASN	2.2
1	A	541	LYS	2.1
1	A	172	SER	2.1
1	A	362	SER	2.1
1	A	23	LYS	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.