



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 4, 2024 – 10:43 AM JST

PDB ID : 8XR2
Title : Crystal structure of AKRtyl-apo1
Authors : Lin, S.; Dai, S.; Xiao, Z.
Deposited on : 2024-01-06
Resolution : 2.39 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

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 : 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

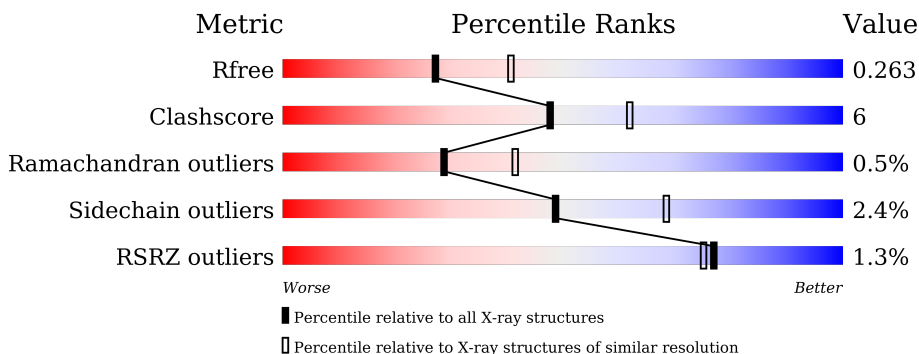
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 ≥ 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	331	
1	B	331	
1	C	331	
1	D	331	
1	E	331	
1	F	331	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	331	<p>2% 71% 18% 10%</p>
1	H	331	<p>2% 83% 13% ..</p>
1	I	331	<p>3% 73% 18% 9%</p>
1	J	331	<p>% 81% 14% ..</p>
1	K	331	<p>% 75% 13% • 11%</p>
1	L	331	<p>% 83% 11% 5%</p>
1	M	331	<p>% 72% 17% • 11%</p>
1	N	331	<p>% 82% 12% • 5%</p>
1	O	331	<p>3% 75% 16% 8%</p>
1	P	331	<p>2% 81% 15% •</p>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 39692 atoms, of which 0 are hydrogens and 0 are deuteriums.

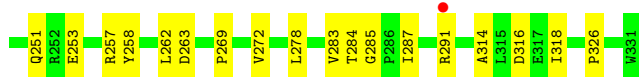
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Aldo/keto reductase.

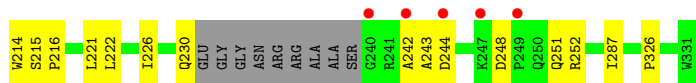
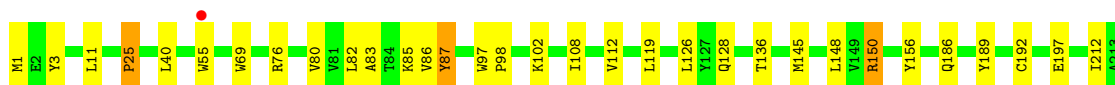
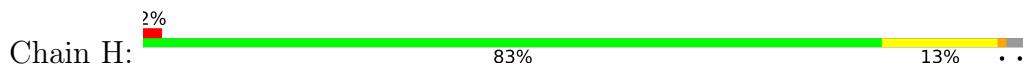
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	294	Total 2290	C 1441	N 411	O 432	S 6	0	0	0
1	B	318	Total 2468	C 1554	N 440	O 466	S 8	0	0	0
1	C	295	Total 2294	C 1445	N 411	O 432	S 6	0	0	0
1	D	313	Total 2429	C 1531	N 434	O 456	S 8	0	0	0
1	E	295	Total 2297	C 1447	N 413	O 430	S 7	0	0	0
1	F	315	Total 2446	C 1542	N 436	O 460	S 8	0	0	0
1	G	297	Total 2311	C 1452	N 415	O 437	S 7	0	0	0
1	H	322	Total 2493	C 1568	N 447	O 470	S 8	0	0	0
1	I	300	Total 2334	C 1470	N 418	O 440	S 6	0	0	0
1	J	317	Total 2460	C 1550	N 439	O 463	S 8	0	0	0
1	K	294	Total 2293	C 1445	N 410	O 431	S 7	0	0	0
1	L	314	Total 2437	C 1537	N 435	O 457	S 8	0	0	0
1	M	295	Total 2297	C 1447	N 413	O 430	S 7	0	0	0
1	N	315	Total 2446	C 1542	N 436	O 460	S 8	0	0	0
1	O	303	Total 2355	C 1479	N 422	O 447	S 7	0	0	0
1	P	319	Total 2469	C 1555	N 440	O 466	S 8	0	0	0

- Molecule 2 is water.

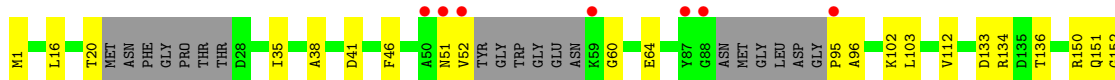
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	79	Total O 79 79	0	0
2	B	105	Total O 105 105	0	0
2	C	86	Total O 86 86	0	0
2	D	111	Total O 111 111	0	0
2	E	84	Total O 84 84	0	0
2	F	85	Total O 85 85	0	0
2	G	71	Total O 71 71	0	0
2	H	111	Total O 111 111	0	0
2	I	81	Total O 81 81	0	0
2	J	110	Total O 110 110	0	0
2	K	68	Total O 68 68	0	0
2	L	122	Total O 122 122	0	0
2	M	108	Total O 108 108	0	0
2	N	129	Total O 129 129	0	0
2	O	87	Total O 87 87	0	0
2	P	136	Total O 136 136	0	0



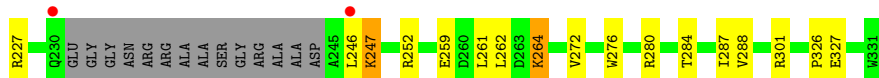
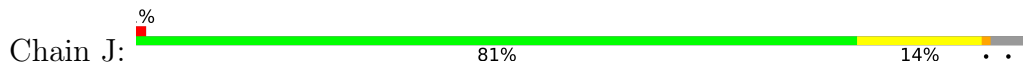
- Molecule 1: Aldo/keto reductase



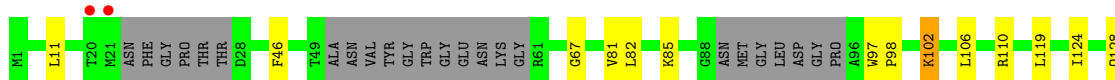
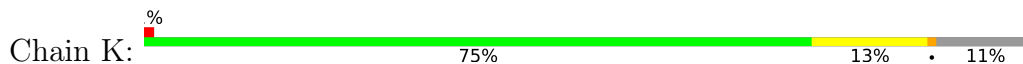
- Molecule 1: Aldo/keto reductase



- Molecule 1: Aldo/keto reductase



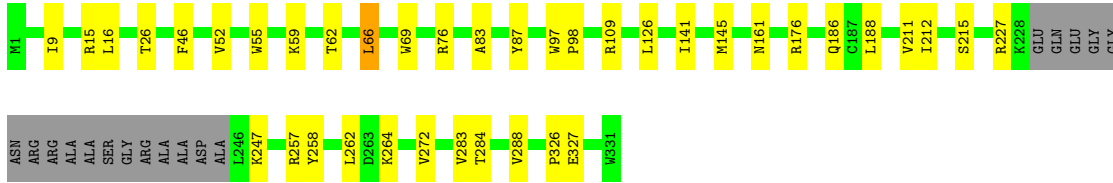
- Molecule 1: Aldo/keto reductase





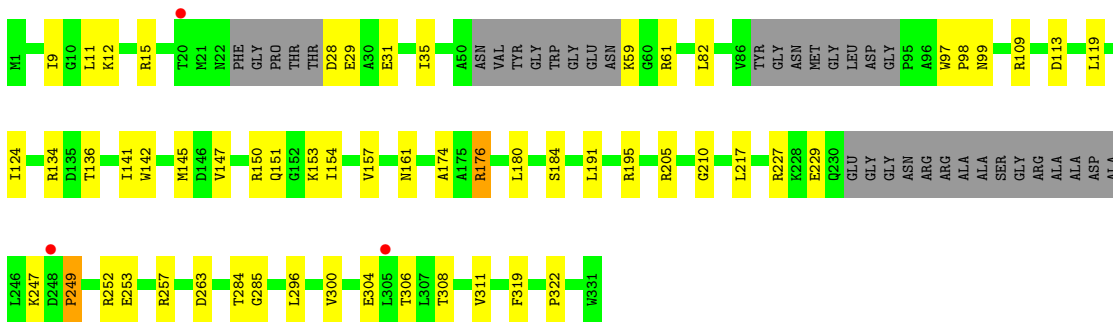
- Molecule 1: Aldo/keto reductase

Chain L: 83% 11% 5%



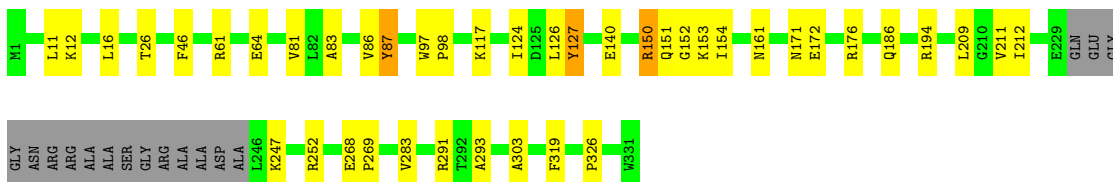
- Molecule 1: Aldo/keto reductase

Chain M: 72% 17% 11%



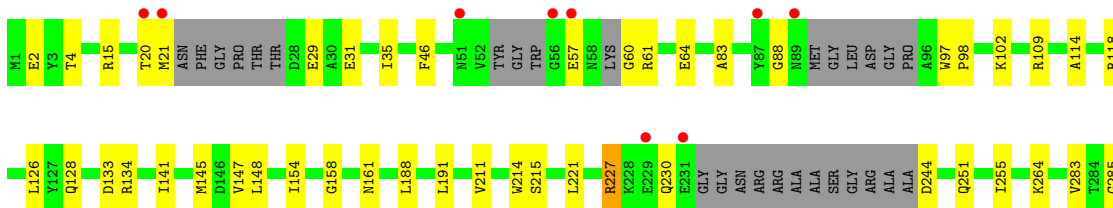
- Molecule 1: Aldo/keto reductase

Chain N: 82% 12% 5%



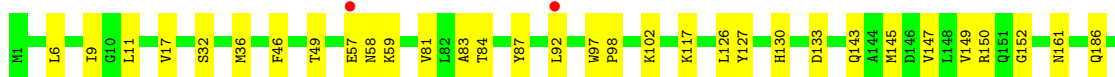
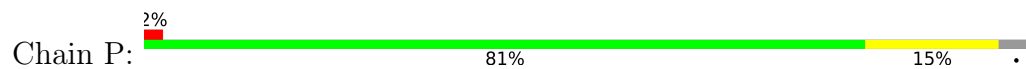
- Molecule 1: Aldo/keto reductase

Chain O: 3% 75% 16% 8%





- Molecule 1: Aldo/keto reductase



4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	110.78Å 110.78Å 561.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.67 – 2.39 49.67 – 2.39	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.67-2.39) 100.0 (49.67-2.39)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.85 (at 2.39Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.222 , 0.263 0.222 , 0.263	Depositor DCC
R_{free} test set	13209 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	27.7	Xtrriage
Anisotropy	0.031	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 21.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtrriage
Estimated twinning fraction	0.067 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	39692	wwPDB-VP
Average B, all atoms (Å ²)	31.0	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 22.77 % 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 5.3278e-03.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.52	0/2336	0.65	0/3167
1	B	0.47	0/2523	0.62	0/3427
1	C	0.40	0/2340	0.57	0/3173
1	D	0.47	0/2484	0.62	0/3374
1	E	0.45	0/2343	0.60	0/3175
1	F	0.44	0/2501	0.58	0/3397
1	G	0.45	0/2355	0.62	0/3191
1	H	0.48	0/2548	0.62	0/3460
1	I	0.48	0/2381	0.63	0/3228
1	J	0.44	0/2515	0.61	0/3416
1	K	0.42	0/2339	0.58	0/3171
1	L	0.48	0/2492	0.63	0/3385
1	M	0.45	0/2343	0.61	0/3175
1	N	0.47	0/2501	0.60	0/3397
1	O	0.46	0/2400	0.61	0/3252
1	P	0.48	0/2524	0.63	0/3429
All	All	0.46	0/38925	0.61	0/52817

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	2290	0	2241	17	0
1	B	2468	0	2408	44	0
1	C	2294	0	2249	35	0
1	D	2429	0	2374	26	0
1	E	2297	0	2264	31	0
1	F	2446	0	2391	43	0
1	G	2311	0	2263	41	0
1	H	2493	0	2434	31	0
1	I	2334	0	2291	37	0
1	J	2460	0	2404	30	0
1	K	2293	0	2250	29	0
1	L	2437	0	2385	20	0
1	M	2297	0	2264	34	0
1	N	2446	0	2391	29	0
1	O	2355	0	2299	33	0
1	P	2469	0	2410	29	0
2	A	79	0	0	2	0
2	B	105	0	0	0	0
2	C	86	0	0	0	0
2	D	111	0	0	0	0
2	E	84	0	0	0	0
2	F	85	0	0	1	0
2	G	71	0	0	1	0
2	H	111	0	0	1	0
2	I	81	0	0	1	0
2	J	110	0	0	1	0
2	K	68	0	0	2	0
2	L	122	0	0	0	0
2	M	108	0	0	2	0
2	N	129	0	0	3	0
2	O	87	0	0	0	0
2	P	136	0	0	0	0
All	All	39692	0	37318	460	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 6.

The worst 5 of 460 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:218:HIS:HD2	1:B:221:LEU:HD12	1.56	0.70
1:D:150:ARG:HH11	1:F:150:ARG:HD2	1.57	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:188:LEU:HD13	1:L:215:SER:HB2	1.77	0.67
1:H:82:LEU:HD21	1:H:119:LEU:HD13	1.78	0.65
1:I:152:GLY:O	1:M:109:ARG:NH1	2.30	0.65

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	284/331 (86%)	271 (95%)	13 (5%)	0	100	100
1	B	314/331 (95%)	297 (95%)	16 (5%)	1 (0%)	41	55
1	C	285/331 (86%)	274 (96%)	11 (4%)	0	100	100
1	D	309/331 (93%)	292 (94%)	14 (4%)	3 (1%)	15	23
1	E	285/331 (86%)	269 (94%)	14 (5%)	2 (1%)	22	32
1	F	311/331 (94%)	297 (96%)	11 (4%)	3 (1%)	15	23
1	G	286/331 (86%)	277 (97%)	9 (3%)	0	100	100
1	H	318/331 (96%)	302 (95%)	13 (4%)	3 (1%)	17	25
1	I	290/331 (88%)	282 (97%)	8 (3%)	0	100	100
1	J	313/331 (95%)	298 (95%)	14 (4%)	1 (0%)	41	55
1	K	284/331 (86%)	273 (96%)	10 (4%)	1 (0%)	34	48
1	L	310/331 (94%)	294 (95%)	13 (4%)	3 (1%)	15	23
1	M	285/331 (86%)	276 (97%)	8 (3%)	1 (0%)	34	48
1	N	311/331 (94%)	297 (96%)	13 (4%)	1 (0%)	41	55
1	O	291/331 (88%)	278 (96%)	12 (4%)	1 (0%)	41	55
1	P	315/331 (95%)	304 (96%)	9 (3%)	2 (1%)	25	36
All	All	4791/5296 (90%)	4581 (96%)	188 (4%)	22 (0%)	29	41

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	87	TYR
1	D	227	ARG
1	E	249	PRO
1	J	87	TYR
1	M	249	PRO

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	234/258 (91%)	228 (97%)	6 (3%)	46	66
1	B	252/258 (98%)	242 (96%)	10 (4%)	31	49
1	C	234/258 (91%)	229 (98%)	5 (2%)	53	72
1	D	248/258 (96%)	239 (96%)	9 (4%)	35	54
1	E	236/258 (92%)	233 (99%)	3 (1%)	69	84
1	F	250/258 (97%)	246 (98%)	4 (2%)	62	79
1	G	237/258 (92%)	231 (98%)	6 (2%)	47	67
1	H	253/258 (98%)	248 (98%)	5 (2%)	55	74
1	I	239/258 (93%)	231 (97%)	8 (3%)	38	57
1	J	251/258 (97%)	246 (98%)	5 (2%)	55	74
1	K	235/258 (91%)	232 (99%)	3 (1%)	69	84
1	L	249/258 (96%)	242 (97%)	7 (3%)	43	63
1	M	236/258 (92%)	230 (98%)	6 (2%)	47	67
1	N	250/258 (97%)	246 (98%)	4 (2%)	62	79
1	O	241/258 (93%)	234 (97%)	7 (3%)	42	62
1	P	251/258 (97%)	245 (98%)	6 (2%)	49	68
All	All	3896/4128 (94%)	3802 (98%)	94 (2%)	49	68

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	246	LEU
1	M	28	ASP
1	J	252	ARG
1	L	46	PHE
1	M	304	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	186	GLN
1	J	128	GLN
1	N	22	ASN
1	K	295	GLN
1	D	295	GLN

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 [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, 95th 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 ²)	Q<0.9
1	A	294/331 (88%)	-0.17	5 (1%) 70 68	18, 32, 59, 71	0
1	B	318/331 (96%)	-0.24	0 100 100	17, 30, 61, 81	0
1	C	295/331 (89%)	-0.19	8 (2%) 54 52	19, 32, 58, 73	0
1	D	313/331 (94%)	-0.31	2 (0%) 89 88	15, 27, 53, 89	0
1	E	295/331 (89%)	-0.19	5 (1%) 70 68	18, 32, 57, 82	0
1	F	315/331 (95%)	-0.23	1 (0%) 94 93	17, 32, 53, 73	0
1	G	297/331 (89%)	-0.23	6 (2%) 65 63	18, 30, 56, 72	0
1	H	322/331 (97%)	-0.27	6 (1%) 66 64	17, 27, 60, 88	0
1	I	300/331 (90%)	-0.25	10 (3%) 46 45	14, 27, 57, 77	0
1	J	317/331 (95%)	-0.41	3 (0%) 84 82	15, 24, 48, 75	0
1	K	294/331 (88%)	-0.29	3 (1%) 82 80	15, 30, 57, 74	0
1	L	314/331 (94%)	-0.46	0 100 100	16, 23, 41, 81	0
1	M	295/331 (89%)	-0.30	3 (1%) 82 80	17, 29, 53, 80	0
1	N	315/331 (95%)	-0.45	0 100 100	15, 25, 43, 64	0
1	O	303/331 (91%)	-0.22	9 (2%) 50 49	17, 29, 60, 75	0
1	P	319/331 (96%)	-0.41	5 (1%) 72 70	13, 22, 46, 89	0
All	All	4906/5296 (92%)	-0.29	66 (1%) 77 75	13, 28, 56, 89	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	95	PRO	4.7
1	I	231	GLU	4.7
1	M	305	LEU	4.6
1	H	240	GLY	4.4
1	O	21	MET	4.4

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.