WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2024 – 01:42 PM EST

|		
PDB ID	:	4JKM
Title	:	Crystal Structure of Clostridium perfringens beta-glucuronidase
Authors	:	Wallace, B.D.; Redinbo, M.R.
Deposited on	:	2013-03-09
Resolution	:	2.26 Å(reported)

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 2.26 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.
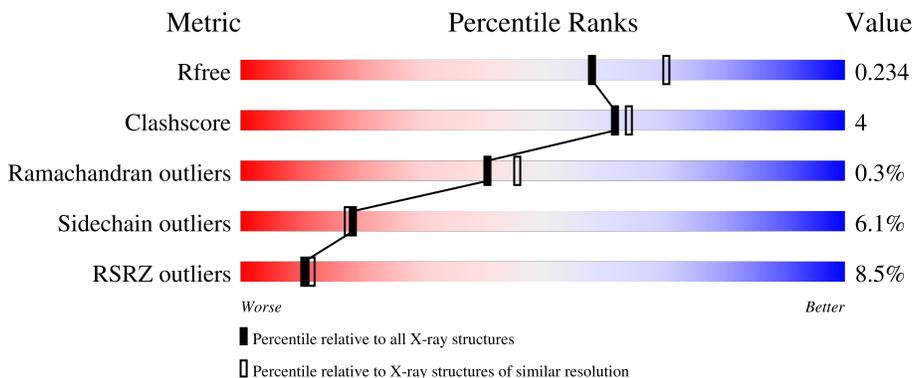
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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 <=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	602	3% / 86% / 13% / ·
1	B	602	4% / 86% / 12% / ·
2	C	400	18% / 50% / 12% / · / 37%
2	D	400	10% / 80% / 8% / · / 10%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 14718 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 Beta-glucuronidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	602	Total 4863	C 3107	N 811	O 926	S 19	0	1	0
1	B	602	Total 4851	C 3098	N 808	O 926	S 19	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q8XP19
A	-1	ASN	-	expression tag	UNP Q8XP19
A	0	ALA	-	expression tag	UNP Q8XP19
B	-2	SER	-	expression tag	UNP Q8XP19
B	-1	ASN	-	expression tag	UNP Q8XP19
B	0	ALA	-	expression tag	UNP Q8XP19

- Molecule 2 is a protein called Maltose-binding periplasmic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	362	Total 2653	C 1693	N 433	O 521	S 6	0	0	0
2	C	252	Total 1839	C 1168	N 302	O 363	S 6	0	0	0

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	MET	-	expression tag	UNP P0AEX9
D	2	LYS	-	expression tag	UNP P0AEX9
D	3	HIS	-	expression tag	UNP P0AEX9
D	4	HIS	-	expression tag	UNP P0AEX9
D	5	HIS	-	expression tag	UNP P0AEX9
D	6	HIS	-	expression tag	UNP P0AEX9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	7	HIS	-	expression tag	UNP P0AEX9
D	8	HIS	-	expression tag	UNP P0AEX9
D	9	GLY	-	expression tag	UNP P0AEX9
D	25	ALA	GLY	SEE REMARK 999	UNP P0AEX9
D	376	ASN	-	expression tag	UNP P0AEX9
D	377	SER	-	expression tag	UNP P0AEX9
D	378	SER	-	expression tag	UNP P0AEX9
D	379	SER	-	expression tag	UNP P0AEX9
D	380	ASN	-	expression tag	UNP P0AEX9
D	381	ASN	-	expression tag	UNP P0AEX9
D	382	ASN	-	expression tag	UNP P0AEX9
D	383	ASN	-	expression tag	UNP P0AEX9
D	384	ASN	-	expression tag	UNP P0AEX9
D	385	ASN	-	expression tag	UNP P0AEX9
D	386	ASN	-	expression tag	UNP P0AEX9
D	387	ASN	-	expression tag	UNP P0AEX9
D	388	ASN	-	expression tag	UNP P0AEX9
D	389	ASN	-	expression tag	UNP P0AEX9
D	390	ARG	-	expression tag	UNP P0AEX9
D	391	ASP	-	expression tag	UNP P0AEX9
D	392	LEU	-	expression tag	UNP P0AEX9
D	393	GLY	-	expression tag	UNP P0AEX9
D	394	THR	-	expression tag	UNP P0AEX9
D	395	GLU	-	expression tag	UNP P0AEX9
D	396	ASN	-	expression tag	UNP P0AEX9
D	397	LEU	-	expression tag	UNP P0AEX9
D	398	TYR	-	expression tag	UNP P0AEX9
D	399	PHE	-	expression tag	UNP P0AEX9
D	400	GLN	-	expression tag	UNP P0AEX9
C	1	MET	-	expression tag	UNP P0AEX9
C	2	LYS	-	expression tag	UNP P0AEX9
C	3	HIS	-	expression tag	UNP P0AEX9
C	4	HIS	-	expression tag	UNP P0AEX9
C	5	HIS	-	expression tag	UNP P0AEX9
C	6	HIS	-	expression tag	UNP P0AEX9
C	7	HIS	-	expression tag	UNP P0AEX9
C	8	HIS	-	expression tag	UNP P0AEX9
C	9	GLY	-	expression tag	UNP P0AEX9
C	25	ALA	GLY	SEE REMARK 999	UNP P0AEX9
C	376	ASN	-	expression tag	UNP P0AEX9
C	377	SER	-	expression tag	UNP P0AEX9
C	378	SER	-	expression tag	UNP P0AEX9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	379	SER	-	expression tag	UNP P0AEX9
C	380	ASN	-	expression tag	UNP P0AEX9
C	381	ASN	-	expression tag	UNP P0AEX9
C	382	ASN	-	expression tag	UNP P0AEX9
C	383	ASN	-	expression tag	UNP P0AEX9
C	384	ASN	-	expression tag	UNP P0AEX9
C	385	ASN	-	expression tag	UNP P0AEX9
C	386	ASN	-	expression tag	UNP P0AEX9
C	387	ASN	-	expression tag	UNP P0AEX9
C	388	ASN	-	expression tag	UNP P0AEX9
C	389	ASN	-	expression tag	UNP P0AEX9
C	390	ARG	-	expression tag	UNP P0AEX9
C	391	ASP	-	expression tag	UNP P0AEX9
C	392	LEU	-	expression tag	UNP P0AEX9
C	393	GLY	-	expression tag	UNP P0AEX9
C	394	THR	-	expression tag	UNP P0AEX9
C	395	GLU	-	expression tag	UNP P0AEX9
C	396	ASN	-	expression tag	UNP P0AEX9
C	397	LEU	-	expression tag	UNP P0AEX9
C	398	TYR	-	expression tag	UNP P0AEX9
C	399	PHE	-	expression tag	UNP P0AEX9
C	400	GLN	-	expression tag	UNP P0AEX9

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	265	Total 265	O 265	0	0
3	B	174	Total 174	O 174	0	0
3	D	49	Total 49	O 49	0	0
3	C	24	Total 24	O 24	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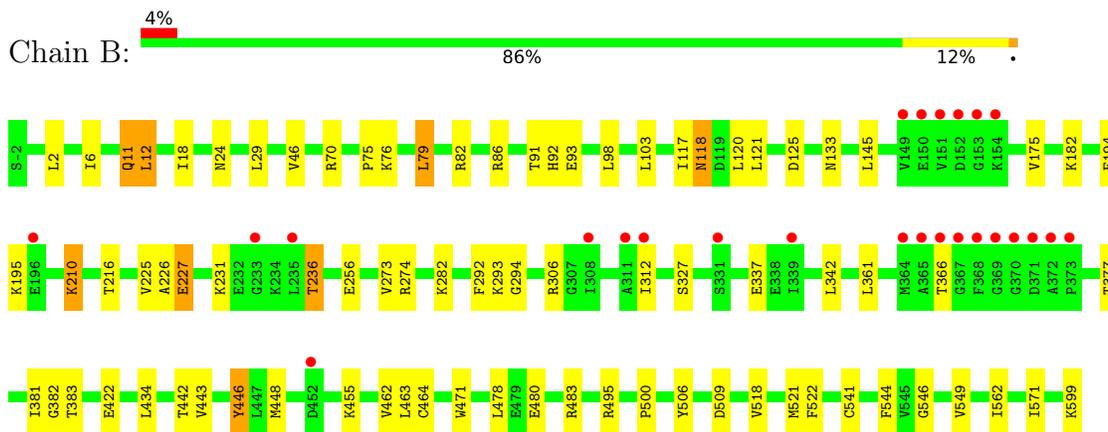 to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
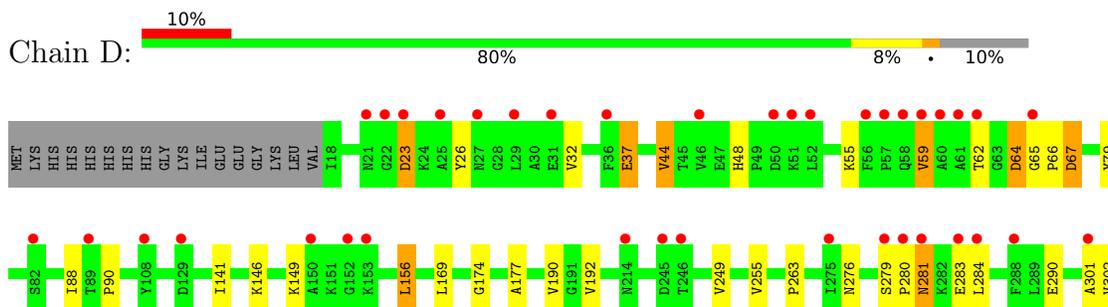
- Molecule 1: Beta-glucuronidase
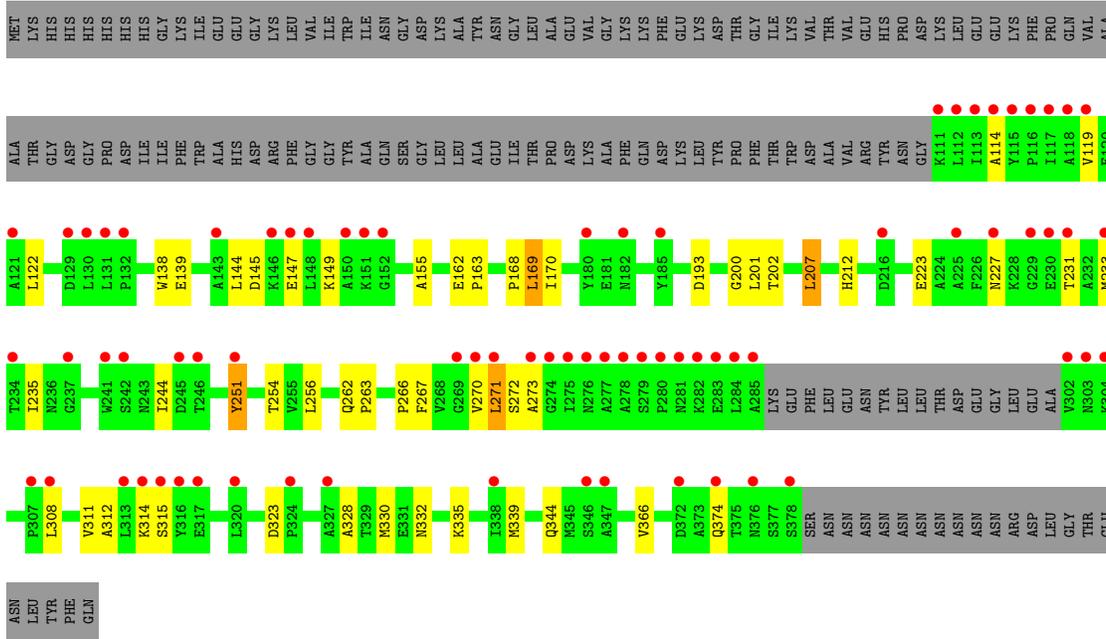
Chain A:

- Molecule 1: Beta-glucuronidase

Chain B:

- Molecule 2: Maltose-binding periplasmic protein

Chain D:

N303 K304 R325 K335 Q344 Q374 S379 ASN ASN ASN ASN ASN ASN ASN ASN ASN ARG ASP LEU GLY THR GLU ASN LEU TYR PHE GLN

• Molecule 2: Maltose-binding periplasmic protein

Chain C:

18%

50%　　12%　　•　　37%

MET LYS HIS HIS HIS HIS HIS GLY LYS ILE GLU GLY GLY LYS LEU VAL ILE TRP ILE ASN GLY ASP LYS ALA TYR ASN GLY LEU ALA GLU VAL GLY LYS LYS PHE GLU LYS ASP THR GLY ILE LYS VAL THR VAL GLU HIS PRO ASP LYS LEU GLU GLU LYS PHE PRO GLN VAL ALA

ALA THR GLY ASP GLY PRO ASP ILE ILE PHE TRP ALA HIS ASP ARG PHE GLY GLY TYR ALA GLN SER GLY LEU LEU ALA GLU ILE THR PRO ASP LYS ALA PHE GLN ASP LYS LEU TYR PRO PHE THR TRP ASP ALA VAL ARG TYR ASN GLY K111 L112 I113 A114 Y115 P116 I117 A118 V119 E120

A121 L122 D129 L130 L131 P132 W138 E139 A143 L144 D145 K146 E147 L148 K149 A150 K151 G152 A155 E162 P163 P168 L169 I170 Y180 E181 N182 Y185 D193 G200 L201 T202 L207 H212 D216 E223 A224 A225 F226 N227 K228 G229 E230 T231 A232 M233

T234 I235 N236 G237 W241 S242 N243 I244 D245 T246 Y251 T254 V255 L256 Q262 P263 P266 F267 V268 G269 V270 L271 S272 A273 G274 I275 N276 A277 A278 S279 P280 N281 K282 E283 L284 A285 LYS GLU PHE LEU GLU ASN TYR LEU LEU THR ASP GLU GLY LEU GLU ALA V302 N303 K304

P307 L308 V311 A312 L313 K314 S315 Y316 E317 L320 D323 P324 A327 A328 T329 M330 E331 N332 K335 I338 M339 Q344 M345 S346 A347 V366 D372 A373 Q374 T375 N376 S377 S378 SER ASN ASN ASN ASN ASN ASN ASN ASN ASN ARG ASP LEU GLY THR GLU

ASN LEU TYR PHE GLN

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	71.51Å 292.61Å 239.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.79 – 2.26 47.79 – 2.26	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.79-2.26) 99.6 (47.79-2.26)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 2.27Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.207 , 0.234 0.207 , 0.234	Depositor DCC
R_{free} test set	5862 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	45.8	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.7	EDS
L-test for twinning[2]	$< |L| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14718	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.37% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

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	0.26	1/4977 (0.0%)	0.45	1/6740 (0.0%)
1	B	0.23	0/4965	0.45	0/6728
2	C	0.22	0/1879	0.42	0/2569
2	D	0.23	0/2717	0.44	0/3718
All	All	0.24	1/14538 (0.0%)	0.45	1/19755 (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
2	D	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	373	PRO	N-CD	5.32	1.55	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	372	ALA	C-N-CD	5.42	139.77	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	59	VAL	Peptide

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	4863	0	4766	41	0
1	B	4851	0	4733	44	0
2	C	1839	0	1730	27	0
2	D	2653	0	2441	19	0
3	A	265	0	0	4	0
3	B	174	0	0	4	0
3	C	24	0	0	3	0
3	D	49	0	0	1	0
All	All	14718	0	13670	123	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.

The worst 5 of 123 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:361:LEU:H	1:B:377:THR:HG22	1.57	0.70
2:D:280:PRO:HD2	2:D:283:GLU:HB3	1.76	0.67
2:D:66:PRO:O	2:D:67:ASP:HB2	1.94	0.66
2:D:59:VAL:HG21	2:D:302:VAL:O	1.96	0.66
1:B:216:THR:HG22	1:B:227:GLU:HG3	1.78	0.65

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	601/602 (100%)	573 (95%)	28 (5%)	0	100	100
1	B	601/602 (100%)	568 (94%)	32 (5%)	1 (0%)	47	55
2	C	248/400 (62%)	227 (92%)	20 (8%)	1 (0%)	34	37
2	D	360/400 (90%)	338 (94%)	18 (5%)	4 (1%)	14	10
All	All	1810/2004 (90%)	1706 (94%)	98 (5%)	6 (0%)	41	46

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	270	VAL
2	D	37	GLU
2	D	44	VAL
2	D	67	ASP
2	D	177	ALA

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	529/531 (100%)	499 (94%)	30 (6%)	20	20
1	B	526/531 (99%)	494 (94%)	32 (6%)	18	18
2	C	179/325 (55%)	167 (93%)	12 (7%)	16	15
2	D	251/325 (77%)	233 (93%)	18 (7%)	14	12
All	All	1485/1712 (87%)	1393 (94%)	92 (6%)	18	17

5 of 92 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	463	LEU
2	D	249	VAL
1	B	495	ARG
2	D	55	LYS
2	D	284	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

There are no ligands in this entry.

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, 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(Å²)	Q<0.9
1	A	602/602 (100%)	0.00	18 (2%)	50	53	18, 29, 55, 162	0
1	B	602/602 (100%)	0.16	25 (4%)	36	38	24, 43, 79, 171	0
2	C	252/400 (63%)	1.41	74 (29%)	0	0	56, 82, 119, 139	0
2	D	362/400 (90%)	0.43	38 (10%)	6	6	28, 61, 104, 117	0
All	All	1818/2004 (90%)	0.33	155 (8%)	10	12	18, 44, 105, 171	0

The worst 5 of 155 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	368	PHE	13.5
1	B	369	GLY	13.5
1	A	371	ASP	11.0
2	C	280	PRO	11.0
1	A	365	ALA	10.2

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 monosaccharides in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.