



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 02:15 AM UTC

PDB ID : 2CBV / pdb_00002cbv
Title : Beta-glucosidase from *Thermotoga maritima* in complex with calystegine B2
Authors : Gloster, T.M.; Madsen, R.; Davies, G.J.
Deposited on : 2006-01-09
Resolution : 1.95 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

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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

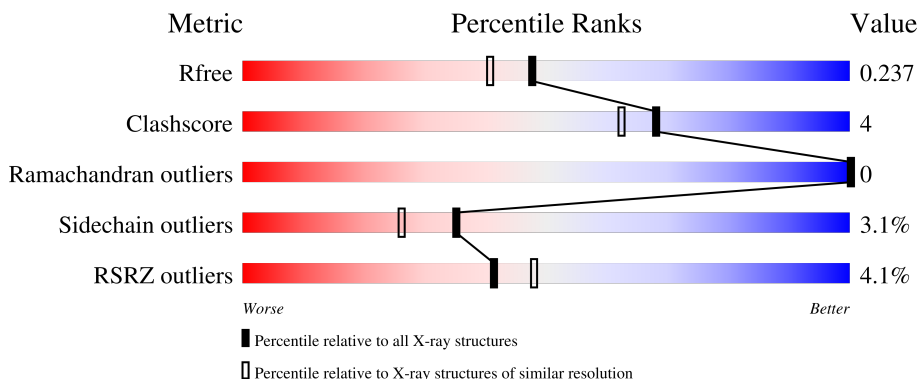
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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}	180053	3494 (1.96-1.96)
Clashscore	190562	3612 (1.96-1.96)
Ramachandran outliers	187476	3587 (1.96-1.96)
Sidechain outliers	187428	3587 (1.96-1.96)
RSRZ outliers	180081	3495 (1.96-1.96)

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	468	
1	B	468	

2 Entry composition [i](#)

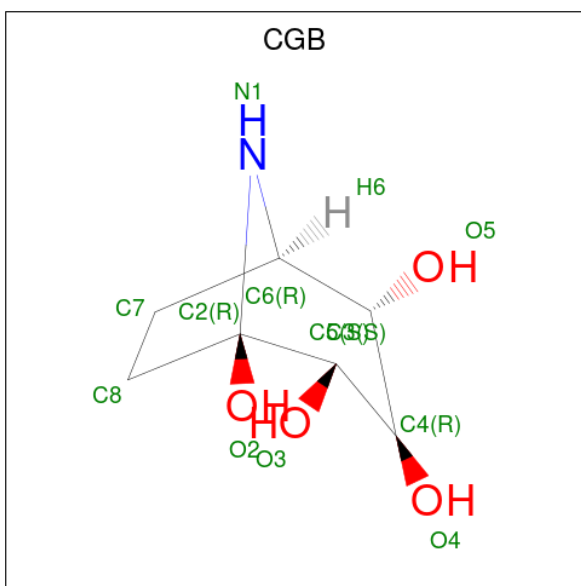
There are 4 unique types of molecules in this entry. The entry contains 8168 atoms, of which 0 are hydrogens and 0 are deuteriums.

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-GLUCOSIDASE A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	443	Total	C	N	O	S	0	20	0
			3759	2450	626	676	7			
1	B	442	Total	C	N	O	S	0	16	0
			3688	2402	617	662	7			

- Molecule 2 is CALYSTEGINE B2 (CCD ID: CGB) (formula: $C_7H_{13}NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total	C	N	O	0	0
			12	7	1	4		
2	B	1	Total	C	N	O	0	0
			12	7	1	4		

- Molecule 3 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf		
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		

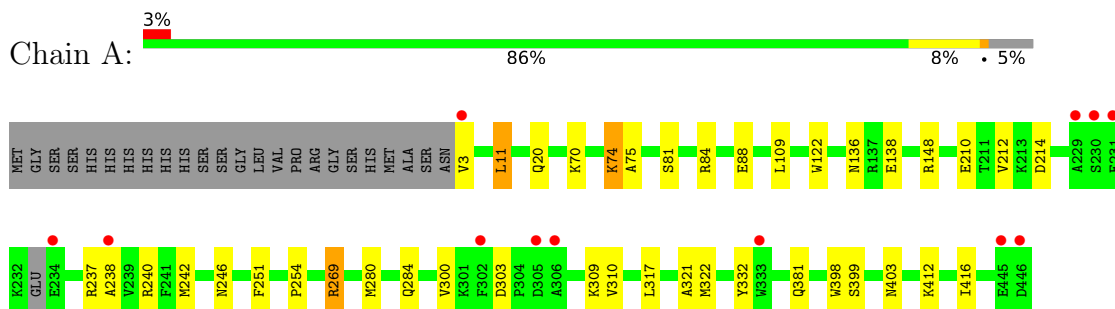
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	
4	A	419	Total	O	0	57
			419	419		
4	B	270	Total	O	0	37
			270	270		

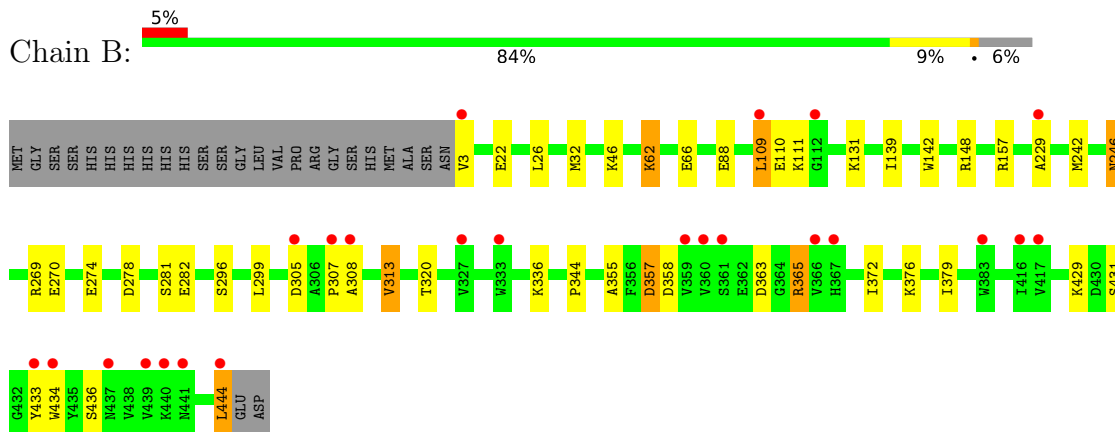
3 Residue-property plots [i](#)

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: BETA-GLUCOSIDASE A



- Molecule 1: BETA-GLUCOSIDASE A



4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	94.26Å 94.31Å 113.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.55 – 1.95 72.51 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.1 (72.55-1.95) 99.1 (72.51-1.95)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.15 (at 1.95Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.183 , 0.238 0.182 , 0.237	Depositor DCC
R_{free} test set	3701 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	28.7	Xtrriage
Anisotropy	0.005	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.012 for k,h,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	8168	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% of the height of the origin peak. No significant pseudotranslation is detected.*

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](#)

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

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	0.83	0/3940	0.87	5/5341 (0.1%)
1	B	0.79	1/3846 (0.0%)	0.88	2/5222 (0.0%)
All	All	0.81	1/7786 (0.0%)	0.87	7/10563 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	344	PRO	CA-C	6.45	1.55	1.51

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	416	ILE	N-CA-C	-7.82	105.08	112.83
1	B	22	GLU	N-CA-C	5.82	117.42	111.14
1	A	238	ALA	N-CA-C	-5.39	105.57	111.82
1	A	214[A]	ASP	N-CA-C	5.21	119.79	113.38
1	A	214[B]	ASP	N-CA-C	5.21	119.79	113.38
1	B	296	SER	N-CA-C	5.19	116.69	108.96
1	A	280	MET	N-CA-C	5.01	116.82	111.36

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	3759	0	3679	21	0
1	B	3688	0	3560	31	0
2	A	12	0	13	1	0
2	B	12	0	13	0	0
3	A	4	0	3	1	0
3	B	4	0	3	1	0
4	A	419	0	0	6	0
4	B	270	0	0	8	0
All	All	8168	0	7271	53	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.

All (53) 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:365:ARG:HH11	1:B:365:ARG:HG2	1.17	1.06
1:A:212:VAL:HG11	4:A:2215:HOH:O	1.69	0.91
1:B:148[A]:ARG:NH1	4:B:2116[A]:HOH:O	2.11	0.83
1:A:70[A]:LYS:HD2	4:A:2111:HOH:O	1.86	0.76
1:B:372:ILE:CD1	1:B:431:SER:HA	2.18	0.72
3:A:1448:ACT:H3	4:A:2013:HOH:O	1.95	0.67
1:B:270[A]:GLU:HG2	4:B:2189:HOH:O	1.94	0.67
1:A:74[A]:LYS:HE3	4:B:2200[A]:HOH:O	1.94	0.66
1:B:278:ASP:O	3:B:1446:ACT:C	2.45	0.65
1:B:365:ARG:HH11	1:B:365:ARG:CG	2.00	0.64
1:B:365:ARG:HG2	1:B:365:ARG:NH1	1.96	0.63
1:B:376:LYS:HB3	1:B:434:TRP:CZ2	2.35	0.62
1:B:372:ILE:HD11	1:B:431:SER:HA	1.80	0.61
1:A:136:ASN:OD1	1:A:138[B]:GLU:HG3	2.01	0.60
1:A:148[B]:ARG:NH2	1:A:210:GLU:OE2	2.34	0.60
1:A:70[A]:LYS:HG3	4:A:2106:HOH:O	2.02	0.59
1:B:372:ILE:HD13	1:B:431:SER:HA	1.84	0.58
1:A:237:ARG:HH22	1:A:240[A]:ARG:HH22	1.51	0.57
1:B:299:LEU:HD23	1:B:313:VAL:HG22	1.87	0.55
1:B:26:LEU:HD21	1:B:32:MET:HG2	1.90	0.54
1:A:242:MET:O	1:A:246[B]:ASN:HB2	2.09	0.53
1:A:269[B]:ARG:HG3	4:A:2133:HOH:O	2.09	0.52
1:B:365:ARG:CG	1:B:365:ARG:NH1	2.65	0.51
1:B:274[A]:GLU:HB3	4:B:2192:HOH:O	2.09	0.51
1:A:300:VAL:HG13	1:A:310:VAL:HG13	1.93	0.50
1:B:66:GLU:HG2	1:B:111:LYS:NZ	2.27	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:307:PRO:O	1:B:308:ALA:HB3	2.14	0.48
1:B:131[B]:LYS:NZ	4:B:2104:HOH:O	2.45	0.48
1:A:20:GLN:O	1:A:403:ASN:HB2	2.14	0.48
1:A:11:LEU:HG	1:A:75:ALA:HB2	1.98	0.46
1:B:358:ASP:OD1	1:B:429:LYS:NZ	2.49	0.45
1:B:62:LYS:HB2	4:B:2053:HOH:O	2.17	0.45
1:B:242:MET:O	1:B:246[B]:ASN:HB2	2.17	0.45
1:B:320:THR:HB	1:B:355:ALA:O	2.17	0.45
1:A:303:ASP:O	1:A:309:LYS:HA	2.17	0.44
1:B:229:ALA:HA	1:B:299:LEU:HD11	2.00	0.44
1:A:321:ALA:O	1:A:412:LYS:HE3	2.19	0.43
1:A:122:TRP:CH2	2:A:1447:CGB:H8C2	2.53	0.43
1:A:322:MET:SD	1:A:412:LYS:HG3	2.59	0.43
1:A:251:PHE:C	1:A:254:PRO:HD2	2.44	0.42
1:B:109:LEU:HD21	1:B:157:ARG:HB3	2.01	0.42
1:B:270[B]:GLU:CD	4:B:2184[B]:HOH:O	2.62	0.42
1:B:379:ILE:HG22	1:B:444:LEU:HD21	2.03	0.41
1:B:433:TYR:O	1:B:436:SER:HB3	2.21	0.41
1:B:139:ILE:HA	1:B:142:TRP:CE3	2.55	0.41
1:A:398:TRP:HA	1:A:399:SER:HA	1.85	0.40
1:A:81:SER:HB3	1:A:84:ARG:HG3	2.02	0.40
1:B:376:LYS:HB3	1:B:434:TRP:HZ2	1.82	0.40
1:A:284:GLN:NE2	4:A:2295:HOH:O	2.51	0.40
1:A:332:TYR:HB3	1:A:381:GLN:HE21	1.86	0.40
1:B:270[B]:GLU:OE1	4:B:2184[B]:HOH:O	2.21	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	462/468 (99%)	452 (98%)	10 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	457/468 (98%)	447 (98%)	10 (2%)	0	100	100
All	All	919/936 (98%)	899 (98%)	20 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	398/399 (100%)	387 (97%)	11 (3%)	38	30
1	B	383/399 (96%)	363 (95%)	20 (5%)	21	9
All	All	781/798 (98%)	750 (96%)	31 (4%)	35	17

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	VAL
1	A	11	LEU
1	A	74[A]	LYS
1	A	74[B]	LYS
1	A	88[A]	GLU
1	A	88[B]	GLU
1	A	88[C]	GLU
1	A	109	LEU
1	A	269[A]	ARG
1	A	269[B]	ARG
1	A	317	LEU
1	B	3	VAL
1	B	46	LYS
1	B	62	LYS
1	B	88	GLU
1	B	109	LEU
1	B	110	GLU
1	B	246[A]	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	246[B]	ASN
1	B	269[A]	ARG
1	B	269[B]	ARG
1	B	269[C]	ARG
1	B	281	SER
1	B	282	GLU
1	B	313	VAL
1	B	336	LYS
1	B	357[A]	ASP
1	B	357[B]	ASP
1	B	363	ASP
1	B	365	ARG
1	B	444	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	59	ASN
1	A	95	GLN
1	A	253	ASN
1	A	343	ASN
1	A	381	GLN
1	A	387	GLN
1	B	253	ASN
1	B	352	ASN
1	B	387	GLN
1	B	441	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 oligosaccharides in this entry.

5.6 Ligand geometry

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	ACT	A	1448	-	3,3,3	1.23	0	3,3,3	0.76	0
2	CGB	B	1445	-	11,13,13	1.15	1 (9%)	12,21,21	0.81	0
2	CGB	A	1447	-	11,13,13	1.29	1 (9%)	12,21,21	0.75	0
3	ACT	B	1446	-	3,3,3	1.08	0	3,3,3	1.73	2 (66%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CGB	B	1445	-	-	-	0/3/2/2
2	CGB	A	1447	-	-	-	0/3/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1447	CGB	C6-N1	-3.52	1.44	1.47
2	B	1445	CGB	C6-N1	-2.73	1.44	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	B	1446	ACT	OXT-C-CH3	2.17	124.13	115.05
3	B	1446	ACT	OXT-C-O	-2.06	114.40	122.03

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1448	ACT	1	0
2	A	1447	CGB	1	0
3	B	1446	ACT	1	0

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, 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	443/468 (94%)	-0.08	12 (2%) 56 62	8, 25, 47, 57	20 (4%)
1	B	442/468 (94%)	0.39	24 (5%) 31 37	10, 32, 49, 56	16 (3%)
All	All	885/936 (94%)	0.15	36 (4%) 41 48	8, 28, 48, 57	36 (4%)

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	VAL	4.5
1	B	441	ASN	3.6
1	B	3	VAL	3.4
1	A	306	ALA	3.0
1	B	360	VAL	2.9
1	A	231	GLU	2.8
1	B	359	VAL	2.7
1	B	307	PRO	2.7
1	A	446	ASP	2.6
1	B	361[A]	SER	2.5
1	B	333	TRP	2.5
1	B	305[A]	ASP	2.5
1	B	416	ILE	2.5
1	A	305	ASP	2.4
1	A	445	GLU	2.4
1	B	439	VAL	2.4
1	B	327	VAL	2.4
1	B	308	ALA	2.3
1	A	333	TRP	2.2
1	A	302	PHE	2.2
1	B	112	GLY	2.2
1	A	234	GLU	2.2
1	B	434	TRP	2.2
1	A	230	SER	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	229	ALA	2.2
1	B	366	VAL	2.2
1	A	238	ALA	2.2
1	B	229	ALA	2.2
1	B	417	VAL	2.1
1	B	109	LEU	2.1
1	B	444	LEU	2.1
1	B	367	HIS	2.1
1	B	437	ASN	2.0
1	B	383	TRP	2.0
1	B	440	LYS	2.0
1	B	433	TYR	2.0

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 oligosaccharides 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, 95th 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(Å ²)	Q < 0.9
3	ACT	B	1446	4/4	0.84	0.13	25,29,29,29	0
3	ACT	A	1448	4/4	0.92	0.12	34,35,35,36	0
2	CGB	A	1447	12/12	0.92	0.07	23,28,30,30	0
2	CGB	B	1445	12/12	0.95	0.06	25,27,29,29	0

6.5 Other polymers [i](#)

There are no such residues in this entry.