



Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 05:06 pm BST

PDB ID : 1VZY
Title : Crystal structure of the Bacillus subtilis HSP33
Authors : Janda, I.K.; Devedjiev, Y.; Derewenda, U.; Dauter, Z.; Bielnicki, J.; Cooper, D.R.; Joachimiak, A.; Derewenda, Z.S.; Midwest Center for Structural Genomics (MCSG)
Deposited on : 2004-05-29
Resolution : 1.97 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

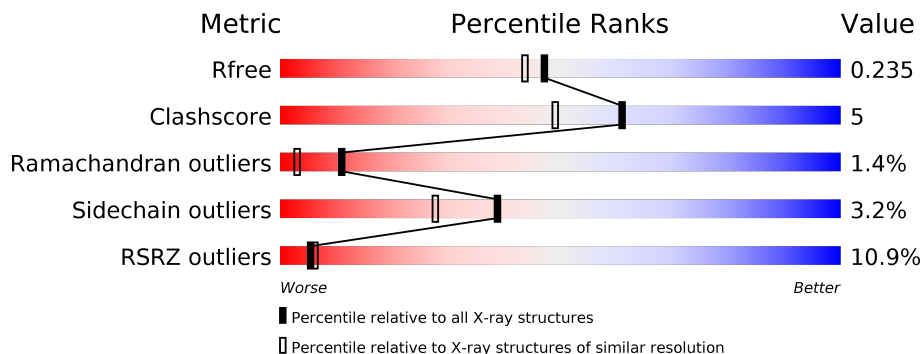
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 on 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	291	 9% 84% 14% .
1	B	291	 13% 85% 11% ..

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 4685 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 33 KDA CHAPERONIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	290	2209	1386	374	437	12	0	0	0
1	B	286	2173	1367	366	428	12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	ALA	GLU	engineered mutation	UNP P37565
A	101	ALA	GLN	engineered mutation	UNP P37565
B	100	ALA	GLU	engineered mutation	UNP P37565
B	101	ALA	GLN	engineered mutation	UNP P37565

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



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

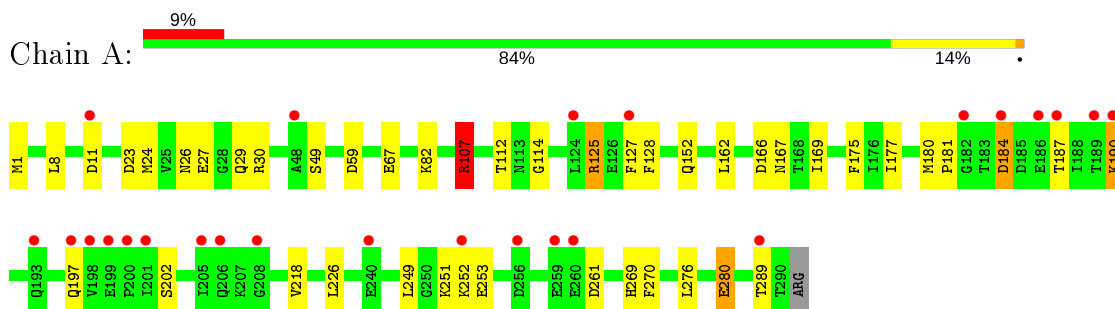
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	146	Total O 146 146	0	0
4	B	131	Total O 131 131	0	0

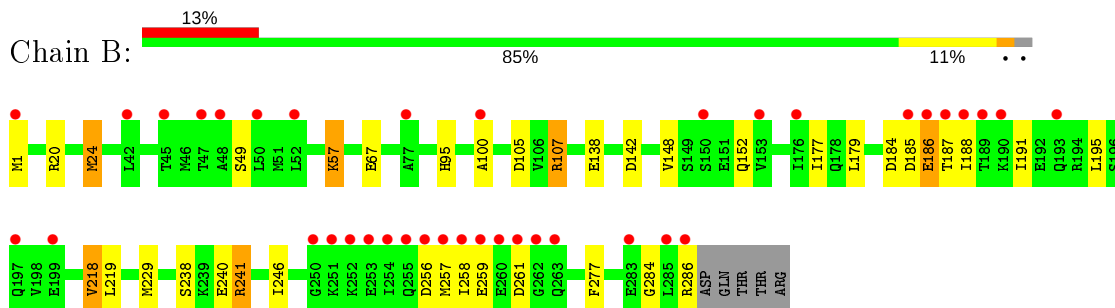
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: 33 KDA CHAPERONIN



- Molecule 1: 33 KDA CHAPERONIN



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	115.29Å 115.29Å 106.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 1.97 28.92 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.00-1.97) 99.7 (28.92-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.05 (at 1.91Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.197 , 0.225 0.210 , 0.235	Depositor DCC
R_{free} test set	1103 reflections (1.71%)	wwPDB-VP
Wilson B-factor (Å ²)	29.7	Xtrriage
Anisotropy	0.221	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 44.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.018 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4685	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.85% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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, ACT

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	1.13	4/2241 (0.2%)	1.03	9/3028 (0.3%)
1	B	1.08	5/2205 (0.2%)	0.97	10/2980 (0.3%)
All	All	1.11	9/4446 (0.2%)	1.00	19/6008 (0.3%)

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	24	MET	CG-SD	7.96	2.01	1.81
1	B	218	VAL	CB-CG1	-6.50	1.39	1.52
1	A	1	MET	CG-SD	6.09	1.97	1.81
1	B	138	GLU	CD-OE1	-5.58	1.19	1.25
1	A	24	MET	CG-SD	5.56	1.95	1.81
1	A	175	PHE	CD1-CE1	5.26	1.49	1.39
1	B	229	MET	SD-CE	-5.26	1.48	1.77
1	A	49	SER	CB-OG	5.08	1.48	1.42
1	B	49	SER	CB-OG	5.08	1.48	1.42

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	24	MET	CG-SD-CE	-7.78	87.75	100.20
1	A	261	ASP	CB-CG-OD2	7.58	125.12	118.30
1	A	30	ARG	NE-CZ-NH2	-7.13	116.73	120.30
1	B	107	ARG	NE-CZ-NH1	6.71	123.65	120.30
1	B	142	ASP	CB-CG-OD2	6.59	124.23	118.30
1	A	30	ARG	NE-CZ-NH1	5.97	123.29	120.30
1	B	107	ARG	NE-CZ-NH2	-5.96	117.32	120.30
1	B	138	GLU	OE1-CD-OE2	-5.82	116.32	123.30
1	B	185	ASP	CB-CG-OD2	5.65	123.39	118.30
1	B	20	ARG	NE-CZ-NH2	-5.61	117.50	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	105	ASP	CB-CG-OD2	5.47	123.23	118.30
1	A	11	ASP	CB-CG-OD2	5.40	123.16	118.30
1	A	184	ASP	CB-CG-OD2	5.36	123.12	118.30
1	A	107	ARG	NE-CZ-NH1	5.35	122.97	120.30
1	B	24	MET	CG-SD-CE	-5.30	91.73	100.20
1	A	125	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	B	184	ASP	CB-CG-OD2	5.25	123.02	118.30
1	B	261	ASP	CB-CG-OD2	5.22	123.00	118.30
1	A	23	ASP	CB-CG-OD2	5.06	122.85	118.30

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	2209	0	2210	21	0
1	B	2173	0	2175	18	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	16	0	12	0	0
3	B	8	0	6	1	0
4	A	146	0	0	5	1
4	B	131	0	0	2	0
All	All	4685	0	4403	38	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 5.

All (38) 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:24:MET:SD	1:B:24:MET:CG	2.01	1.46
1:A:190:LYS:HD3	4:A:2006:HOH:O	1.79	0.83
1:B:95:HIS:HD2	4:B:2053:HOH:O	1.67	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:162:LEU:HG	1:A:202:SER:HB2	1.69	0.74
1:B:24:MET:CE	1:B:24:MET:CG	2.66	0.73
1:A:27:GLU:HG3	4:A:2020:HOH:O	1.89	0.73
1:A:59:ASP:OD1	1:A:82:LYS:NZ	2.22	0.72
1:A:27:GLU:CG	4:A:2020:HOH:O	2.36	0.72
1:A:269:HIS:HD2	4:A:2135:HOH:O	1.73	0.72
1:A:125:ARG:HA	1:A:127:PHE:CE2	2.26	0.70
1:A:27:GLU:CD	4:A:2020:HOH:O	2.30	0.70
1:B:24:MET:CB	1:B:24:MET:SD	2.81	0.68
1:B:240:GLU:O	1:B:241:ARG:CB	2.48	0.61
1:B:177:ILE:HD13	1:B:219:LEU:HD21	1.84	0.60
1:A:26:ASN:ND2	1:A:169:ILE:H	2.03	0.56
1:A:166:ASP:O	1:A:167:ASN:HB2	2.07	0.54
1:B:100:ALA:H	3:B:1289:ACT:H2	1.72	0.53
1:B:95:HIS:HE1	4:B:2056:HOH:O	1.93	0.52
1:A:184:ASP:OD1	1:A:187:THR:HG23	2.11	0.51
1:A:8:LEU:HG	1:A:226:LEU:HD11	1.92	0.51
1:B:195:LEU:CD2	1:B:218:VAL:HG13	2.41	0.51
1:A:26:ASN:HD22	1:A:169:ILE:H	1.59	0.49
1:B:240:GLU:O	1:B:241:ARG:HB2	2.14	0.48
1:A:177:ILE:CD1	1:A:218:VAL:HG11	2.45	0.46
1:B:179:LEU:HD13	1:B:188:ILE:HG12	1.98	0.45
1:B:177:ILE:HD13	1:B:219:LEU:CD2	2.45	0.45
1:A:67:GLU:O	1:A:114:GLY:HA3	2.18	0.44
1:A:249:LEU:HD22	1:A:253:GLU:OE1	2.18	0.44
1:B:246:ILE:CD1	1:B:277:PHE:CE2	3.00	0.44
1:A:251:LYS:HE2	1:A:289:THR:O	2.17	0.43
1:A:152:GLN:OE1	1:B:57:LYS:HG2	2.19	0.42
1:A:180:MET:HB3	1:A:181:PRO:HD2	2.01	0.42
1:B:148:VAL:O	1:B:152:GLN:HA	2.20	0.42
1:A:107:ARG:CZ	1:A:112:THR:HG21	2.51	0.41
1:B:238:SER:O	1:B:241:ARG:HB3	2.21	0.40
1:B:24:MET:CB	1:B:24:MET:CE	2.99	0.40
1:A:280:GLU:CD	1:A:280:GLU:H	2.24	0.40
1:B:191:ILE:O	1:B:195:LEU:HG	2.21	0.40

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)
4:A:2084:HOH:O	4:A:2084:HOH:O[4_555]	0.99	1.21

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	288/291 (99%)	278 (96%)	9 (3%)	1 (0%)	41	29
1	B	284/291 (98%)	269 (95%)	8 (3%)	7 (2%)	5	1
All	All	572/582 (98%)	547 (96%)	17 (3%)	8 (1%)	11	3

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	197	GLN
1	B	186	GLU
1	B	241	ARG
1	B	257	MET
1	B	259	GLU
1	B	256	ASP
1	B	284	GLY
1	B	258	ILE

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	238/239 (100%)	230 (97%)	8 (3%)	37	25
1	B	233/239 (98%)	226 (97%)	7 (3%)	41	29
All	All	471/478 (98%)	456 (97%)	15 (3%)	39	28

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	GLN
1	A	107	ARG
1	A	128	PHE
1	A	190	LYS
1	A	252	LYS
1	A	270	PHE
1	A	276	LEU
1	A	280	GLU
1	B	1	MET
1	B	57	LYS
1	B	67	GLU
1	B	107	ARG
1	B	186	GLU
1	B	187	THR
1	B	286	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	ASN
1	A	234	HIS
1	A	269	HIS
1	B	95	HIS
1	B	164	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](#)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 for Mogul analysis.

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	B	1289	-	1,3,3	0.46	0	0,3,3	0.00	-
3	ACT	A	1292	-	1,3,3	1.24	0	0,3,3	0.00	-
3	ACT	B	1288	-	1,3,3	0.29	0	0,3,3	0.00	-
3	ACT	A	1295	-	1,3,3	0.71	0	0,3,3	0.00	-
3	ACT	A	1294	-	1,3,3	1.74	0	0,3,3	0.00	-
3	ACT	A	1293	-	1,3,3	2.49	1 (100%)	0,3,3	0.00	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1293	ACT	CH3-C	2.49	1.51	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1289	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	290/291 (99%)	0.21	25 (8%) 10 11	21, 30, 61, 69	10 (3%)
1	B	286/291 (98%)	1.20	38 (13%) 3 3	23, 34, 81, 96	27 (9%)
All	All	576/582 (98%)	0.70	63 (10%) 5 6	21, 32, 72, 96	37 (6%)

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	254	ILE	28.5
1	B	261	ASP	26.0
1	B	262	GLY	22.4
1	B	257	MET	20.0
1	B	258	ILE	20.0
1	B	251	LYS	18.2
1	B	253	GLU	17.8
1	B	256	ASP	17.6
1	B	255	GLN	16.1
1	B	260	GLU	15.8
1	B	250	GLY	13.9
1	B	259	GLU	13.5
1	B	252	LYS	12.6
1	B	286	ARG	7.3
1	B	189	THR	4.0
1	A	127	PHE	3.9
1	B	185	ASP	3.5
1	A	198	VAL	3.5
1	A	184	ASP	3.5
1	B	48	ALA	3.4
1	A	193	GLN	3.4
1	B	100	ALA	3.3
1	B	186	GLU	3.3
1	A	187	THR	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	189	THR	3.2
1	A	260	GLU	3.2
1	A	289	THR	3.1
1	B	187	THR	3.0
1	B	193	GLN	3.0
1	A	182	GLY	3.0
1	B	188	ILE	2.9
1	A	240	GLU	2.8
1	B	197	GLN	2.8
1	A	186	GLU	2.7
1	A	190	LYS	2.7
1	B	50	LEU	2.7
1	B	190	LYS	2.7
1	A	205	ILE	2.6
1	B	283	GLU	2.6
1	B	153	VAL	2.6
1	A	124	LEU	2.6
1	B	285	LEU	2.6
1	B	42	LEU	2.5
1	A	252	LYS	2.5
1	A	200	PRO	2.4
1	A	11	ASP	2.4
1	A	208	GLY	2.4
1	A	206	GLN	2.4
1	B	47	THR	2.4
1	A	199	GLU	2.3
1	B	263	GLN	2.3
1	B	77	ALA	2.3
1	A	259	GLU	2.2
1	B	176	ILE	2.2
1	B	150	SER	2.2
1	B	199	GLU	2.2
1	A	201	ILE	2.1
1	A	256	ASP	2.1
1	B	45	THR	2.1
1	B	1	MET	2.0
1	B	52	LEU	2.0
1	A	48	ALA	2.0
1	A	197	GLN	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 carbohydrates 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(\AA^2)	Q<0.9
3	ACT	A	1293	4/4	0.65	0.18	61,61,61,62	0
3	ACT	A	1295	4/4	0.75	0.26	58,58,59,59	0
3	ACT	B	1289	4/4	0.84	0.17	54,54,54,55	0
3	ACT	A	1292	4/4	0.89	0.19	34,36,36,39	0
3	ACT	A	1294	4/4	0.91	0.18	59,59,59,61	0
3	ACT	B	1288	4/4	0.96	0.17	55,56,56,56	0
2	ZN	A	1291	1/1	0.99	0.07	29,29,29,29	0
2	ZN	B	1287	1/1	0.99	0.13	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.