



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 11:15 PM UTC

PDB ID : 6PCZ / pdb_00006pcz
Title : Crystal structure of the bacterial cellulose synthase subunit G (BcsG) catalytic domain from Escherichia coli, selenomethionine variant
Authors : Anderson, A.C.; Brenner, T.; Weadge, J.T.
Deposited on : 2019-06-18
Resolution : 1.44 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

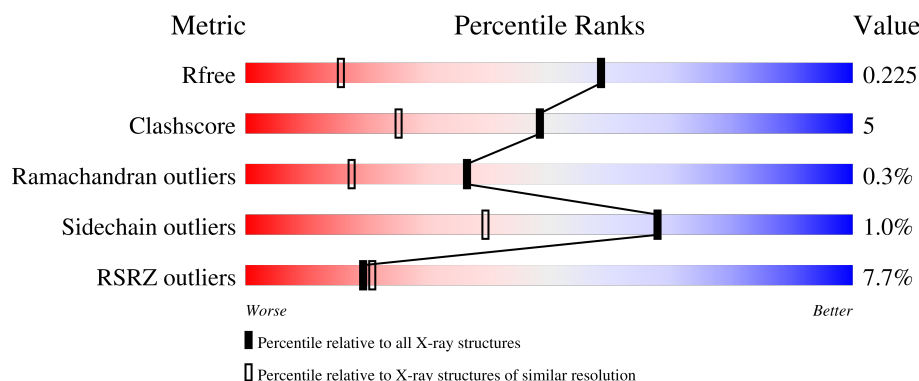
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.44 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	3234 (1.46-1.42)
Clashscore	190562	3289 (1.46-1.42)
Ramachandran outliers	187476	3248 (1.46-1.42)
Sidechain outliers	187428	3248 (1.46-1.42)
RSRZ outliers	180081	3234 (1.46-1.42)

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	417	7% 81% 6% 13%
1	B	417	6% 77% 10% 13%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6376 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 Cellulose biosynthesis protein BcsG.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	363	Total	C	N	O	S	Se	0	0	0
			2833	1802	481	539	3	8			
1	B	364	Total	C	N	O	S	Se	0	0	0
			2838	1805	482	540	3	8			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	MSE	-	initiating methionine	UNP P37659
A	144	GLY	-	expression tag	UNP P37659
A	145	SER	-	expression tag	UNP P37659
A	146	SER	-	expression tag	UNP P37659
A	147	HIS	-	expression tag	UNP P37659
A	148	HIS	-	expression tag	UNP P37659
A	149	HIS	-	expression tag	UNP P37659
A	150	HIS	-	expression tag	UNP P37659
A	151	HIS	-	expression tag	UNP P37659
A	152	HIS	-	expression tag	UNP P37659
A	153	SER	-	expression tag	UNP P37659
A	154	SER	-	expression tag	UNP P37659
A	155	GLY	-	expression tag	UNP P37659
A	156	LEU	-	expression tag	UNP P37659
A	157	VAL	-	expression tag	UNP P37659
A	158	PRO	-	expression tag	UNP P37659
A	159	ARG	-	expression tag	UNP P37659
A	160	GLY	-	expression tag	UNP P37659
A	161	SER	-	expression tag	UNP P37659
A	162	HIS	-	expression tag	UNP P37659
A	163	MSE	-	expression tag	UNP P37659
B	143	MSE	-	initiating methionine	UNP P37659
B	144	GLY	-	expression tag	UNP P37659
B	145	SER	-	expression tag	UNP P37659
B	146	SER	-	expression tag	UNP P37659

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	147	HIS	-	expression tag	UNP P37659
B	148	HIS	-	expression tag	UNP P37659
B	149	HIS	-	expression tag	UNP P37659
B	150	HIS	-	expression tag	UNP P37659
B	151	HIS	-	expression tag	UNP P37659
B	152	HIS	-	expression tag	UNP P37659
B	153	SER	-	expression tag	UNP P37659
B	154	SER	-	expression tag	UNP P37659
B	155	GLY	-	expression tag	UNP P37659
B	156	LEU	-	expression tag	UNP P37659
B	157	VAL	-	expression tag	UNP P37659
B	158	PRO	-	expression tag	UNP P37659
B	159	ARG	-	expression tag	UNP P37659
B	160	GLY	-	expression tag	UNP P37659
B	161	SER	-	expression tag	UNP P37659
B	162	HIS	-	expression tag	UNP P37659
B	163	MSE	-	expression tag	UNP P37659

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

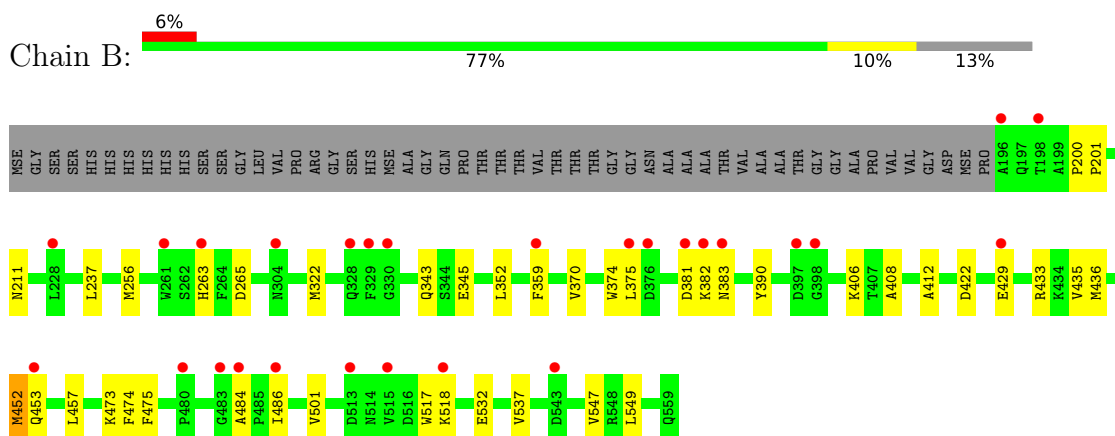
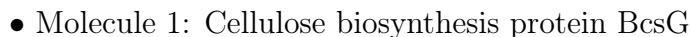
- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	385	Total O 385 385	0	0
4	B	317	Total O 317 317	0	0

- Molecule 1: Cellulose biosynthesis protein BcsG



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	78.71Å 94.58Å 105.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.97 – 1.44 45.97 – 1.44	Depositor EDS
% Data completeness (in resolution range)	100.0 (45.97-1.44) 100.0 (45.97-1.44)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 1.44Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.196 , 0.208 0.215 , 0.225	Depositor DCC
R_{free} test set	7108 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	19.1	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 25.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6376	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.36% 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: MG, ZN

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	0.39	1/2900 (0.0%)	0.57	1/3933 (0.0%)
1	B	0.51	3/2905 (0.1%)	0.60	0/3940
All	All	0.46	4/5805 (0.1%)	0.58	1/7873 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	256	MSE	C-N	-7.75	1.23	1.33
1	B	343	GLN	C-N	-7.07	1.24	1.33
1	A	436	MSE	C-O	-5.04	1.18	1.24
1	B	435	VAL	C-O	-5.04	1.19	1.24

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	477	MSE	O-C-N	5.83	129.74	122.86

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	2833	0	2746	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2838	0	2749	31	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
4	A	385	0	0	8	0
4	B	317	0	0	4	0
All	All	6376	0	5495	60	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 5.

All (60) 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:452:MSE:HE3	1:A:557:TYR:CE2	1.93	1.03
1:B:486:ILE:HD11	1:B:517:TRP:HB3	1.51	0.92
1:A:452:MSE:CE	1:A:557:TYR:CD2	2.54	0.90
1:A:404:VAL:HG12	1:A:406:LYS:HD2	1.58	0.86
1:A:404:VAL:CG1	1:A:406:LYS:HD2	2.07	0.85
1:B:429:GLU:OE2	1:B:475:PHE:HD2	1.64	0.80
1:A:452:MSE:HE2	1:A:558:PRO:HD2	1.64	0.79
1:A:452:MSE:CE	1:A:557:TYR:CE2	2.67	0.77
1:A:290:CYS:HG	1:A:306:CYS:HG	1.32	0.76
1:A:477:MSE:HE3	1:A:509:ILE:HG12	1.66	0.75
1:A:452:MSE:HE1	1:A:557:TYR:CD2	2.26	0.70
1:B:322:MSE:HE2	1:B:370:VAL:CG1	2.22	0.70
1:A:212:ASN:ND2	4:A:705:HOH:O	2.30	0.64
1:A:304:ASN:ND2	4:A:701:HOH:O	2.20	0.64
1:B:322:MSE:HE2	1:B:370:VAL:HG12	1.82	0.62
1:B:322:MSE:HE1	1:B:374:TRP:HB2	1.83	0.60
1:B:322:MSE:HG3	1:B:390:TYR:CD1	2.37	0.58
1:B:453:GLN:HG2	1:B:457:LEU:HB2	1.84	0.58
1:A:345:GLU:HG3	4:A:752:HOH:O	2.03	0.58
1:B:322:MSE:CE	4:B:866:HOH:O	2.52	0.56
1:A:269:LYS:HD3	1:A:487:VAL:HG11	1.88	0.56
1:B:322:MSE:CE	1:B:374:TRP:HB2	2.35	0.56
1:B:408:ALA:HB1	1:B:412:ALA:HB3	1.88	0.56
1:A:404:VAL:HG11	1:A:406:LYS:HD2	1.88	0.55
1:B:429:GLU:OE2	1:B:475:PHE:CD2	2.53	0.55
1:A:406:LYS:HD3	1:A:406:LYS:O	2.07	0.54
1:B:263:HIS:ND1	1:B:422:ASP:OD1	2.37	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:486:ILE:CD1	1:B:517:TRP:HB3	2.29	0.52
1:B:452:MSE:CE	1:B:549:LEU:HD11	2.38	0.52
1:B:322:MSE:HE2	1:B:370:VAL:HG13	1.92	0.52
1:B:406:LYS:HG2	4:B:789:HOH:O	2.09	0.52
1:A:404:VAL:HG12	1:A:406:LYS:H	1.75	0.52
1:B:322:MSE:O	1:B:390:TYR:HA	2.09	0.52
1:A:322:MSE:SE	4:A:924:HOH:O	2.77	0.51
1:A:322:MSE:O	1:A:390:TYR:HA	2.11	0.50
1:A:401:TYR:H	1:A:406:LYS:NZ	2.09	0.50
1:B:237:LEU:HD12	1:B:436:MSE:O	2.11	0.50
1:A:559:GLN:NE2	4:A:712:HOH:O	2.46	0.49
1:A:355:ILE:HG23	1:A:366:ASP:HB2	1.95	0.48
1:B:381:ASP:O	1:B:383:ASN:N	2.46	0.47
1:A:518:LYS:HE3	4:A:898:HOH:O	2.15	0.47
1:B:532:GLU:HG2	1:B:537:VAL:HG22	1.97	0.46
1:A:518:LYS:HD2	1:A:522:SER:OG	2.15	0.45
1:B:436:MSE:HG3	1:B:474:PHE:CE1	2.52	0.45
1:B:322:MSE:HE1	1:B:370:VAL:O	2.18	0.44
1:B:452:MSE:SE	1:B:547:VAL:HG11	2.67	0.44
1:B:436:MSE:HE1	1:B:501:VAL:HA	2.01	0.43
1:A:290:CYS:CB	1:A:306:CYS:HG	2.31	0.43
1:B:322:MSE:HE3	4:B:866:HOH:O	2.17	0.43
1:B:322:MSE:HG3	1:B:390:TYR:HD1	1.84	0.43
1:B:375:LEU:O	1:B:433:ARG:NH2	2.51	0.43
1:A:452:MSE:HE3	1:A:557:TYR:CD2	2.26	0.42
1:B:453:GLN:HG3	4:B:739:HOH:O	2.19	0.42
1:A:269:LYS:HE2	1:A:489:GLU:OE1	2.20	0.42
1:B:200:PRO:HA	1:B:201:PRO:HD3	1.85	0.42
1:B:265:ASP:OD2	1:B:473:LYS:HD2	2.19	0.42
1:A:269:LYS:HD2	4:A:749:HOH:O	2.19	0.41
1:B:486:ILE:HD11	1:B:517:TRP:CB	2.37	0.41
1:A:269:LYS:HD3	1:A:487:VAL:CG1	2.50	0.40
1:A:540:GLN:CD	4:A:715:HOH:O	2.63	0.40

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	361/417 (87%)	354 (98%)	7 (2%)	0	100	100
1	B	362/417 (87%)	349 (96%)	11 (3%)	2 (1%)	21	6
All	All	723/834 (87%)	703 (97%)	18 (2%)	2 (0%)	36	17

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	382	LYS
1	B	484	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	311/337 (92%)	311 (100%)	0	100	100
1	B	311/337 (92%)	305 (98%)	6 (2%)	50	16
All	All	622/674 (92%)	616 (99%)	6 (1%)	68	39

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	211	ASN
1	B	345	GLU
1	B	352	LEU
1	B	359	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	452	MSE
1	B	518	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	241	ASN
1	A	328	GLN
1	A	391	ASN
1	A	482	GLN
1	A	526	GLN
1	B	241	ASN
1	B	303	ASN
1	B	391	ASN
1	B	415	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	355/417 (85%)	0.55	28 (7%) 18 20	12, 20, 32, 39	0
1	B	356/417 (85%)	0.80	27 (7%) 20 21	15, 24, 34, 39	0
All	All	711/834 (85%)	0.68	55 (7%) 19 21	12, 22, 34, 39	0

All (55) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	359	PHE	6.9
1	A	406	LYS	5.5
1	B	484	ALA	5.4
1	A	559	GLN	5.2
1	B	261	TRP	5.1
1	A	404	VAL	5.1
1	B	480	PRO	4.7
1	B	196	ALA	4.3
1	B	513	ASP	4.1
1	B	483	GLY	3.7
1	B	486	ILE	3.7
1	A	397	ASP	3.7
1	B	382	LYS	3.5
1	A	483	GLY	3.4
1	B	383	ASN	3.4
1	B	398	GLY	3.4
1	A	198	THR	3.3
1	B	397	ASP	3.2
1	B	263	HIS	3.2
1	A	448	LYS	3.2
1	A	405	SER	3.2
1	A	261	TRP	3.1
1	A	558	PRO	3.1
1	B	429	GLU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	200	PRO	2.9
1	A	513	ASP	2.9
1	A	199	ALA	2.9
1	B	453	GLN	2.8
1	A	514	ASN	2.8
1	A	263	HIS	2.6
1	B	328	GLN	2.6
1	B	543	ASP	2.6
1	B	329	PHE	2.6
1	B	376	ASP	2.5
1	B	304	ASN	2.5
1	A	543	ASP	2.5
1	B	330	GLY	2.4
1	A	290	CYS	2.4
1	A	482	GLN	2.4
1	B	518	LYS	2.4
1	B	381	ASP	2.3
1	A	504	VAL	2.3
1	B	198	THR	2.3
1	A	329	PHE	2.2
1	A	201	PRO	2.2
1	A	408	ALA	2.2
1	A	518	LYS	2.2
1	A	535	ASN	2.1
1	A	484	ALA	2.1
1	A	557	TYR	2.1
1	B	228	LEU	2.0
1	A	489	GLU	2.0
1	A	555	VAL	2.0
1	B	515	VAL	2.0
1	B	375	LEU	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($\text{\AA}^2$)	Q<0.9
2	ZN	A	601	1/1	0.99	0.33	20,20,20,20	0
2	ZN	B	601	1/1	0.99	0.48	20,20,20,20	0
3	MG	A	602	1/1	0.99	0.05	20,20,20,20	0

6.5 Other polymers [i](#)

There are no such residues in this entry.