



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 3, 2021 – 10:03 PM EDT

PDB ID : 7LDG  
Title : Crystal structure of the MEILB2-BRCA2 complex  
Authors : Nandakumar, J.; Pendlebury, D.F.  
Deposited on : 2021-01-13  
Resolution : 2.56 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](mailto: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 : 2.23.1  
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) : 2.23.1

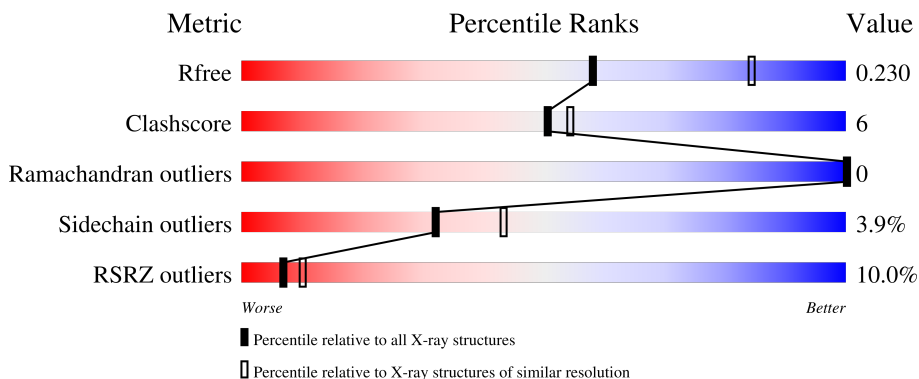
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.56 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	1279 (2.58-2.54)
Clashscore	141614	1327 (2.58-2.54)
Ramachandran outliers	138981	1312 (2.58-2.54)
Sidechain outliers	138945	1312 (2.58-2.54)
RSRZ outliers	127900	1269 (2.58-2.54)

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  $\geq 3$ , 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	253	 6% 75% 11% 13%
1	C	253	 11% 68% 16% 16%
2	B	66	 6% 55% 18% 24%
2	D	66	 9% 89%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3812 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 Heat shock factor 2-binding protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	220	Total	C	N	O	S	Se	0	0	0
			1699	1089	280	320	5	5			
1	C	213	Total	C	N	O	S	Se	0	0	0
			1646	1053	273	310	5	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	SER	-	expression tag	UNP O75031
C	82	SER	-	expression tag	UNP O75031

- Molecule 2 is a protein called Breast cancer type 2 susceptibility protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
2	B	50	Total	C	N	O	S	Se	0	0	0
			410	258	75	75	1	1			
2	D	7	Total	C	N	O			0	0	0
			51	35	7	9					

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	2270	MSE	-	initiating methionine	UNP P51587
D	2270	MSE	-	initiating methionine	UNP P51587

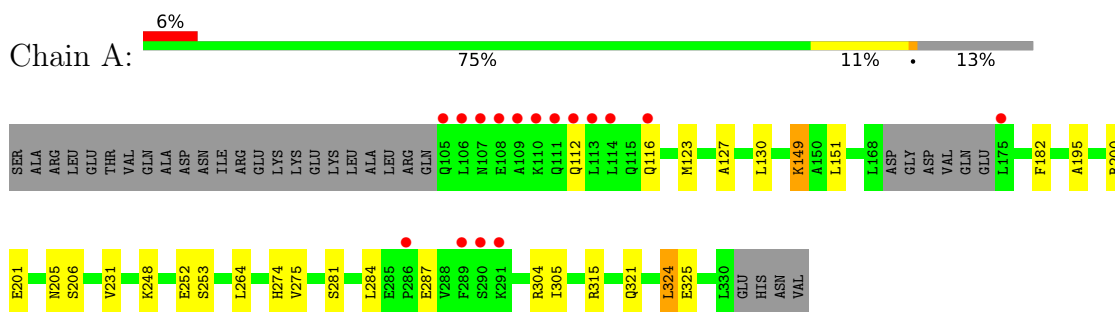
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	3	Total	O	0	0
			3	3		
3	C	3	Total	O	0	0
			3	3		

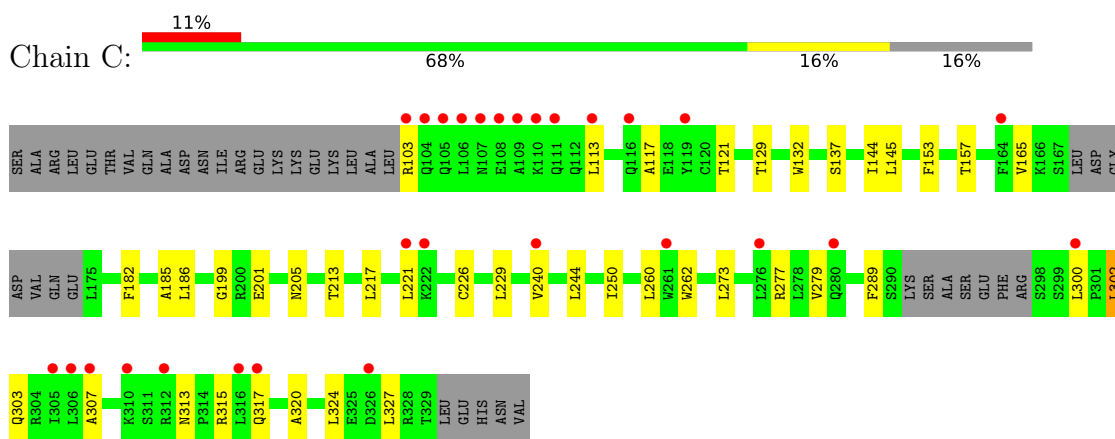
### 3 Residue-property plots [i](#)

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.

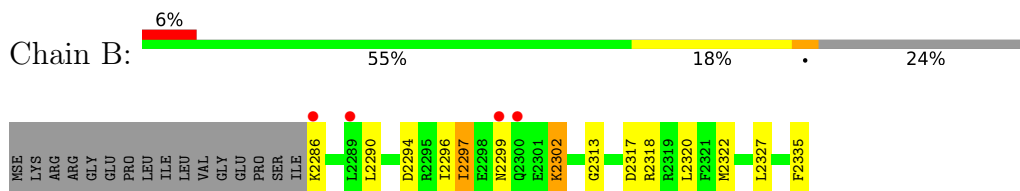
- Molecule 1: Heat shock factor 2-binding protein



- Molecule 1: Heat shock factor 2-binding protein



- Molecule 2: Breast cancer type 2 susceptibility protein



- Molecule 2: Breast cancer type 2 susceptibility protein



MSE	PRO
LYS	SER
ARG	ILE
ARG	LYS
GLY	ARG
GLU	ASN
P2276	LEU
L2279	LEU
E2282	ASN
PRO	GLU
SER	PHE
ILE	ASP
LYS	ARG
ARG	ILE
ASN	ILE
LEU	GLU
ASN	GLN
GLU	GLN
LYS	LYS
SER	SER
LEU	LEU
LYS	LYS
ALA	ALA
SER	SER
LYS	LYS
SER	SER
THR	THR
PRO	PRO
ASP	ASP
GLY	GLY
THR	THR
ILE	ILE
LYS	LYS
ASP	ASP
ARG	ARG
ARG	ARG
LEU	LEU
PHE	PHE
MSE	MSE
HIS	HIS
HIS	HIS
VAL	VAL
SER	SER
LEU	LEU
GLU	GLU
PRO	PRO
ILE	ILE
THR	THR
CYS	
VAL	
PRO	
PHE	

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	179.89Å 179.89Å 179.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	31.10 – 2.56 89.94 – 2.56	Depositor EDS
% Data completeness (in resolution range)	100.0 (31.10-2.56) 100.0 (89.94-2.56)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 2.55Å)	Xtrriage
Refinement program	PHENIX 1.17.1-3600	Depositor
R, $R_{free}$	0.211 , 0.231 0.211 , 0.230	Depositor DCC
$R_{free}$ test set	1714 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.9	Xtrriage
Anisotropy	0.147	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 74.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3812	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	96.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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](#)

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	0.27	0/1718	0.41	0/2315
1	C	0.27	0/1663	0.42	0/2240
2	B	0.28	0/416	0.47	0/556
2	D	0.28	0/51	0.45	0/68
All	All	0.27	0/3848	0.42	0/5179

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	1699	0	1760	16	0
1	C	1646	0	1702	21	0
2	B	410	0	420	13	0
2	D	51	0	58	1	0
3	B	3	0	0	1	0
3	C	3	0	0	0	0
All	All	3812	0	3940	44	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.

All (44) 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:C:201:GLU:OE2	1:C:205:ASN:ND2	2.21	0.74
2:B:2286:LYS:HG2	2:B:2296:ILE:HD11	1.76	0.68
1:A:123:MSE:HE2	1:A:127:ALA:HB2	1.78	0.66
1:A:305:ILE:HG22	1:A:324:LEU:HD22	1.83	0.61
1:C:300:LEU:HD11	1:C:327:LEU:HD22	1.83	0.60
1:C:260:LEU:HD13	1:C:279:VAL:HG22	1.85	0.57
1:A:149:LYS:HG3	1:C:129:THR:HG21	1.87	0.56
1:A:284:LEU:HD22	2:B:2322:MSE:HE1	1.87	0.56
1:C:303:GLN:O	1:C:307:ALA:N	2.29	0.56
1:A:248:LYS:HE2	1:A:252:GLU:OE2	2.06	0.54
1:C:300:LEU:O	1:C:302:LEU:N	2.39	0.54
1:C:165:VAL:HG11	1:C:217:LEU:HD22	1.90	0.53
1:A:195:ALA:O	1:A:200:ARG:NH2	2.42	0.53
1:A:123:MSE:HE1	1:C:153:PHE:CD1	2.43	0.52
1:C:273:LEU:HD21	1:C:315:ARG:HE	1.77	0.49
1:A:130:LEU:HD22	1:C:144:ILE:HG23	1.95	0.48
1:A:231:VAL:HG13	1:A:274:HIS:CD2	2.48	0.48
1:A:281:SER:OG	2:B:2318:ARG:NH2	2.46	0.48
2:B:2297:ILE:HG22	2:B:2299:ASN:H	1.79	0.48
1:C:273:LEU:O	1:C:277:ARG:HG3	2.14	0.47
2:B:2320:LEU:HD13	2:B:2327:LEU:HB2	1.96	0.47
1:C:213:THR:O	1:C:217:LEU:HG	2.15	0.46
1:A:287:GLU:OE1	1:A:287:GLU:N	2.48	0.46
1:A:321:GLN:O	1:A:325:GLU:HG2	2.16	0.46
1:A:201:GLU:OE2	1:A:205:ASN:ND2	2.48	0.45
1:A:116:GLN:NE2	1:C:113:LEU:HD11	2.32	0.44
2:B:2335:PHE:HE1	2:D:2279:LEU:HB2	1.81	0.44
1:C:226:CYS:HB3	1:C:229:LEU:HB3	1.99	0.44
2:B:2297:ILE:HG21	2:B:2302:LYS:O	2.18	0.44
2:B:2294:ASP:OD1	3:B:2401:HOH:O	2.21	0.44
1:C:117:ALA:O	1:C:121:THR:HG23	2.18	0.42
2:B:2318:ARG:NH1	2:B:2320:LEU:HD21	2.34	0.42
2:B:2290:LEU:HD23	2:B:2290:LEU:HA	1.80	0.42
1:C:132:TRP:CE3	1:C:185:ALA:HA	2.54	0.42
1:C:240:VAL:CG2	1:C:250:ILE:HD11	2.50	0.42
2:B:2302:LYS:HB3	2:B:2302:LYS:HE2	1.89	0.41
1:C:221:LEU:O	1:C:262:TRP:HH2	2.03	0.41
2:B:2313:GLY:HA2	2:B:2317:ASP:HB2	2.03	0.41
1:C:320:ALA:O	1:C:324:LEU:HG	2.20	0.41
1:C:145:LEU:HD11	1:C:199:GLY:HA2	2.03	0.41
1:A:151:LEU:HD21	1:A:206:SER:HB3	2.03	0.40
1:A:264:LEU:HG	1:A:275:VAL:HG11	2.03	0.40

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:157:THR:HG23	1:C:186:LEU:HB3	2.03	0.40
2:B:2313:GLY:HA2	2:B:2317:ASP:OD2	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	216/253 (85%)	214 (99%)	2 (1%)	0	100	100
1	C	207/253 (82%)	200 (97%)	7 (3%)	0	100	100
2	B	48/66 (73%)	45 (94%)	3 (6%)	0	100	100
2	D	5/66 (8%)	4 (80%)	1 (20%)	0	100	100
All	All	476/638 (75%)	463 (97%)	13 (3%)	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	192/216 (89%)	185 (96%)	7 (4%)	35	47
1	C	186/216 (86%)	178 (96%)	8 (4%)	29	39
2	B	48/60 (80%)	46 (96%)	2 (4%)	30	40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	6/60 (10%)	6 (100%)	0	100	100
All	All	432/552 (78%)	415 (96%)	17 (4%)	32	43

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	112	GLN
1	A	149	LYS
1	A	182	PHE
1	A	253	SER
1	A	304	ARG
1	A	315	ARG
1	A	324	LEU
2	B	2297	ILE
2	B	2302	LYS
1	C	103	ARG
1	C	137	SER
1	C	182	PHE
1	C	244	LEU
1	C	289	PHE
1	C	302	LEU
1	C	313	ASN
1	C	317	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	313	ASN

### 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	215/253 (84%)	0.73	16 (7%) 14 19	60, 81, 145, 167	0
1	C	208/253 (82%)	1.00	28 (13%) 3 4	63, 95, 151, 167	0
2	B	49/66 (74%)	0.83	4 (8%) 11 16	62, 84, 158, 172	0
2	D	7/66 (10%)	1.07	0 100 100	76, 92, 122, 122	0
All	All	479/638 (75%)	0.86	48 (10%) 7 10	60, 87, 151, 172	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	105	GLN	8.8
1	A	106	LEU	8.7
1	C	106	LEU	6.6
1	C	119	TYR	6.6
1	A	108	GLU	6.3
1	A	290	SER	5.7
1	A	111	GLN	5.2
1	A	109	ALA	5.1
1	A	105	GLN	4.9
1	A	175	LEU	4.7
1	C	109	ALA	4.4
1	C	316	LEU	4.0
1	C	110	LYS	3.9
1	C	305	ILE	3.7
1	A	112	GLN	3.7
1	C	310	LYS	3.5
1	A	289	PHE	3.5
1	C	261	TRP	3.4
1	C	104	GLN	3.4
1	C	221	LEU	3.3
1	A	107	ASN	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	103	ARG	3.3
1	C	300	LEU	3.2
1	A	110	LYS	3.2
1	C	111	GLN	3.1
1	C	164	PHE	3.0
1	C	317	GLN	2.9
2	B	2299	ASN	2.8
1	A	113	LEU	2.8
1	C	312	ARG	2.8
1	A	114	LEU	2.8
1	C	116	GLN	2.7
1	A	291	LYS	2.7
1	C	222	LYS	2.7
1	C	280	GLN	2.6
1	C	307	ALA	2.6
1	C	108	GLU	2.5
1	C	107	ASN	2.4
1	C	306	LEU	2.3
1	C	113	LEU	2.2
1	C	326	ASP	2.2
2	B	2300	GLN	2.2
1	A	286	PRO	2.2
1	C	240	VAL	2.1
1	A	116	GLN	2.0
2	B	2286	LYS	2.0
1	C	276	LEU	2.0
2	B	2289	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 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.