



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 11, 2016 – 02:20 PM EDT

PDB ID : 5FRQ
Title : Crystal Structure of Helicobacter pylori beta clamp bound to DNA ligase peptide
Authors : Pandey, P.; Gourinath, S.
Deposited on : 2015-12-21
Resolution : 2.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

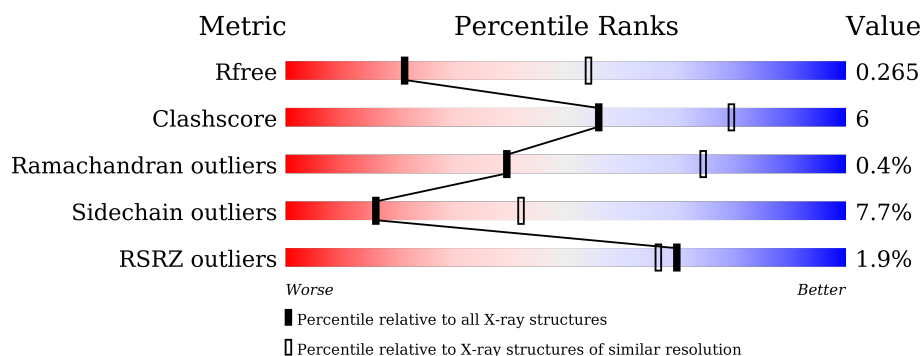
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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. 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	384	<div> <div>2%</div> <div>79% 12% • 7%</div> </div>
1	B	384	<div> <div>2%</div> <div>77% 13% • • 5%</div> </div>
1	C	384	<div> <div>3%</div> <div>77% 15% • 7%</div> </div>
1	D	384	<div> <div>%</div> <div>78% 12% • • 7%</div> </div>
2	G	8	<div> <div>38% 38% 25%</div> </div>
2	L	8	<div> <div>13% 38% 25% 25%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11507 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 DNA POLYMERASE III SUBUNIT BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	358	Total	C	N	O	S	0	0	0
			2837	1826	450	550	11			
1	B	364	Total	C	N	O	S	0	0	0
			2882	1854	458	559	11			
1	C	358	Total	C	N	O	S	0	0	0
			2840	1829	451	549	11			
1	D	357	Total	C	N	O	S	0	0	0
			2828	1820	448	549	11			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	EXPRESSION TAG	UNP O25242
A	0	ALA	-	EXPRESSION TAG	UNP O25242
A	375	GLU	-	EXPRESSION TAG	UNP O25242
A	376	LEU	-	EXPRESSION TAG	UNP O25242
A	377	HIS	-	EXPRESSION TAG	UNP O25242
A	378	HIS	-	EXPRESSION TAG	UNP O25242
A	379	HIS	-	EXPRESSION TAG	UNP O25242
A	380	HIS	-	EXPRESSION TAG	UNP O25242
A	381	HIS	-	EXPRESSION TAG	UNP O25242
A	382	HIS	-	EXPRESSION TAG	UNP O25242
B	-1	MET	-	EXPRESSION TAG	UNP O25242
B	0	ALA	-	EXPRESSION TAG	UNP O25242
B	375	GLU	-	EXPRESSION TAG	UNP O25242
B	376	LEU	-	EXPRESSION TAG	UNP O25242
B	377	HIS	-	EXPRESSION TAG	UNP O25242
B	378	HIS	-	EXPRESSION TAG	UNP O25242
B	379	HIS	-	EXPRESSION TAG	UNP O25242
B	380	HIS	-	EXPRESSION TAG	UNP O25242
B	381	HIS	-	EXPRESSION TAG	UNP O25242
B	382	HIS	-	EXPRESSION TAG	UNP O25242
C	-1	MET	-	EXPRESSION TAG	UNP O25242

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	ALA	-	EXPRESSION TAG	UNP O25242
C	375	GLU	-	EXPRESSION TAG	UNP O25242
C	376	LEU	-	EXPRESSION TAG	UNP O25242
C	377	HIS	-	EXPRESSION TAG	UNP O25242
C	378	HIS	-	EXPRESSION TAG	UNP O25242
C	379	HIS	-	EXPRESSION TAG	UNP O25242
C	380	HIS	-	EXPRESSION TAG	UNP O25242
C	381	HIS	-	EXPRESSION TAG	UNP O25242
C	382	HIS	-	EXPRESSION TAG	UNP O25242
D	-1	MET	-	EXPRESSION TAG	UNP O25242
D	0	ALA	-	EXPRESSION TAG	UNP O25242
D	375	GLU	-	EXPRESSION TAG	UNP O25242
D	376	LEU	-	EXPRESSION TAG	UNP O25242
D	377	HIS	-	EXPRESSION TAG	UNP O25242
D	378	HIS	-	EXPRESSION TAG	UNP O25242
D	379	HIS	-	EXPRESSION TAG	UNP O25242
D	380	HIS	-	EXPRESSION TAG	UNP O25242
D	381	HIS	-	EXPRESSION TAG	UNP O25242
D	382	HIS	-	EXPRESSION TAG	UNP O25242

- Molecule 2 is a protein called DNA LIGASE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	6	Total	C	N	O	0	0	0
			55	39	9	7			
2	L	6	Total	C	N	O	0	0	0
			55	39	9	7			

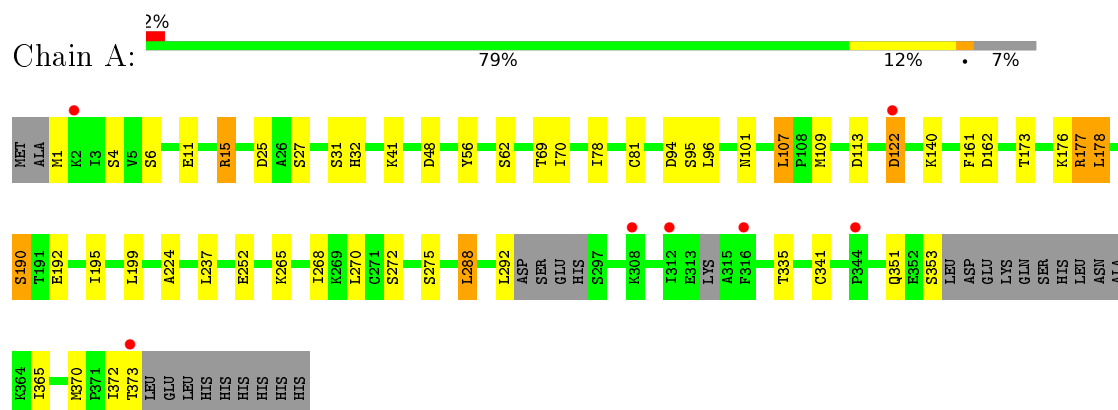
- Molecule 3 is water.

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

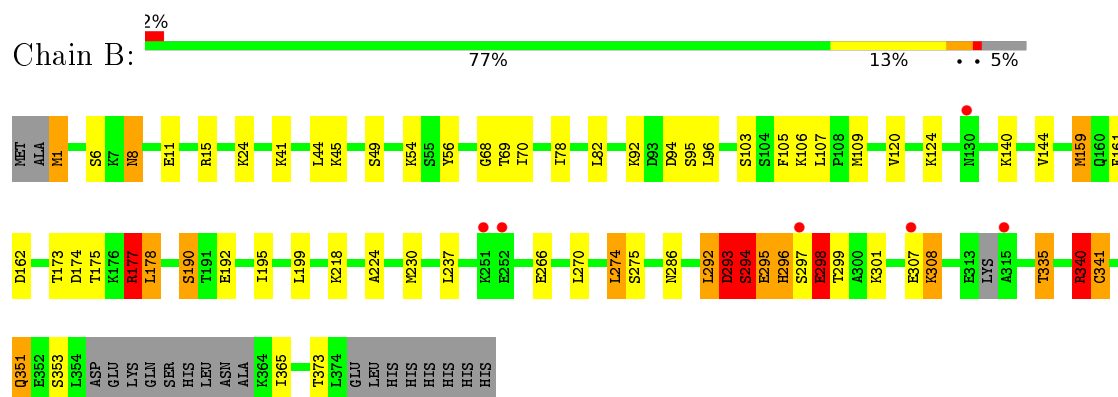
3 Residue-property plots [i](#)

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

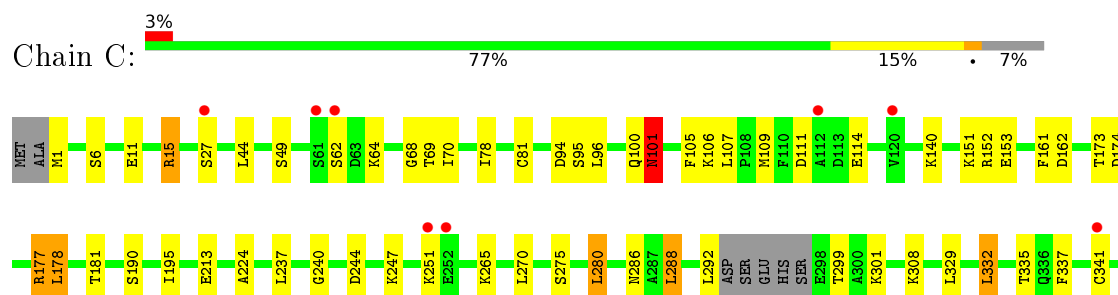
• Molecule 1: DNA POLYMERASE III SUBUNIT BETA



• Molecule 1: DNA POLYMERASE III SUBUNIT BETA

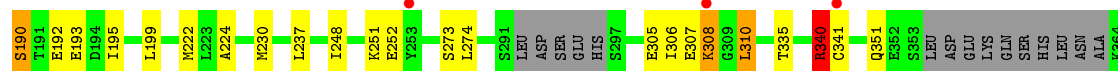
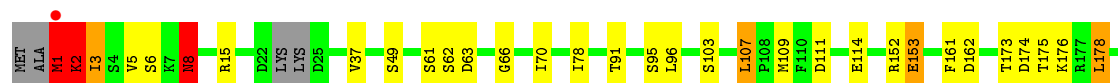
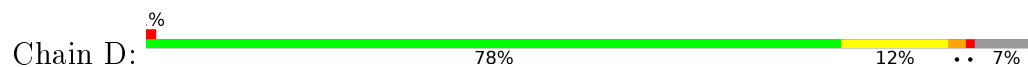


• Molecule 1: DNA POLYMERASE III SUBUNIT BETA

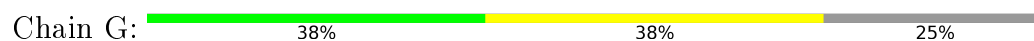




• Molecule 1: DNA POLYMERASE III SUBUNIT BETA



• Molecule 2: DNA LIGASE



• Molecule 2: DNA LIGASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	64.83Å 146.15Å 179.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	113.23 – 2.90 42.64 – 2.90	Depositor EDS
% Data completeness (in resolution range)	94.3 (113.23-2.90) 94.3 (42.64-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.30 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.231 , 0.262 0.236 , 0.265	Depositor DCC
R_{free} test set	1842 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	46.5	Xtriage
Anisotropy	0.268	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 27.6	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	0.90	EDS
Total number of atoms	11507	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% 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](#)

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	0.66	0/2883	0.88	12/3876 (0.3%)
1	B	0.67	0/2930	0.89	13/3942 (0.3%)
1	C	0.67	0/2887	0.91	11/3882 (0.3%)
1	D	0.71	1/2874 (0.0%)	0.92	10/3865 (0.3%)
2	G	0.67	0/56	0.98	0/73
2	L	0.84	0/56	1.13	1/73 (1.4%)
All	All	0.68	1/11686 (0.0%)	0.90	47/15711 (0.3%)

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
1	D	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	153	GLU	CD-OE1	5.16	1.31	1.25

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	1	MET	CG-SD-CE	11.81	119.10	100.20
1	B	341	CYS	N-CA-C	9.69	137.16	111.00
1	C	280	LEU	CA-CB-CG	9.33	136.76	115.30
1	C	332	LEU	CB-CG-CD2	8.48	125.41	111.00
1	B	1	MET	CG-SD-CE	8.36	113.57	100.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	1	MET	Peptide
1	D	8	ASN	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	2837	0	2901	22	0
1	B	2882	0	2942	58	0
1	C	2840	0	2910	37	0
1	D	2828	0	2888	43	0
2	G	55	0	57	3	0
2	L	55	0	57	5	0
3	A	3	0	0	0	0
3	B	2	0	0	0	0
3	C	3	0	0	0	0
3	D	2	0	0	0	0
All	All	11507	0	11755	144	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.

The worst 5 of 144 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:D:1:MET:CE	1:D:2:LYS:O	1.78	1.32
1:D:1:MET:HE3	1:D:2:LYS:O	1.47	1.13
1:B:298:GLU:N	1:B:298:GLU:OE1	1.94	0.99
1:C:70:ILE:HD11	1:C:107:LEU:HD13	1.61	0.83
1:B:70:ILE:HD11	1:B:107:LEU:HD13	1.61	0.80

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	350/384 (91%)	334 (95%)	16 (5%)	0	100	100
1	B	358/384 (93%)	338 (94%)	16 (4%)	4 (1%)	17	51
1	C	352/384 (92%)	335 (95%)	15 (4%)	2 (1%)	30	67
1	D	349/384 (91%)	331 (95%)	18 (5%)	0	100	100
2	G	4/8 (50%)	4 (100%)	0	0	100	100
2	L	4/8 (50%)	4 (100%)	0	0	100	100
All	All	1417/1552 (91%)	1346 (95%)	65 (5%)	6 (0%)	39	74

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	293	ASP
1	B	294	SER
1	B	295	GLU
1	B	296	HIS
1	C	101	ASN

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	327/351 (93%)	305 (93%)	22 (7%)	20	50
1	B	332/351 (95%)	303 (91%)	29 (9%)	13	36
1	C	327/351 (93%)	301 (92%)	26 (8%)	15	40

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	326/351 (93%)	304 (93%)	22 (7%)	20	50
2	G	6/8 (75%)	5 (83%)	1 (17%)	3	8
2	L	6/8 (75%)	4 (67%)	2 (33%)	0	1
All	All	1324/1420 (93%)	1222 (92%)	102 (8%)	16	42

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	340	ARG
1	C	95	SER
1	D	335	THR
1	B	341	CYS
1	C	6	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	8	ASN
1	B	165	HIS
1	D	130	ASN
1	D	241	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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, 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	358/384 (93%)	0.18	7 (1%) 68 64	49, 59, 69, 77	0
1	B	364/384 (94%)	0.17	6 (1%) 74 72	53, 64, 81, 90	0
1	C	358/384 (93%)	0.24	10 (2%) 56 50	54, 64, 75, 85	0
1	D	357/384 (92%)	0.13	4 (1%) 82 80	51, 60, 71, 78	0
2	G	6/8 (75%)	0.33	0 100 100	62, 66, 70, 75	0
2	L	6/8 (75%)	0.72	0 100 100	75, 77, 80, 86	0
All	All	1449/1552 (93%)	0.18	27 (1%) 70 66	49, 62, 75, 90	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1	MET	4.2
1	A	344	PRO	3.5
1	A	316	PHE	3.2
1	C	252	GLU	3.2
1	A	312	ILE	3.2

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.