



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2021 – 12:17 PM EST

PDB ID : 7KJ1
Title : human peroxiredoxin 2 - C172S mutant
Authors : Kean, K.M.; Karplus, P.A.
Deposited on : 2020-10-25
Resolution : 2.15 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

<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
Xtriage (Phenix) : 1.13
EDS : 2.17.1
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) : 2.17.1

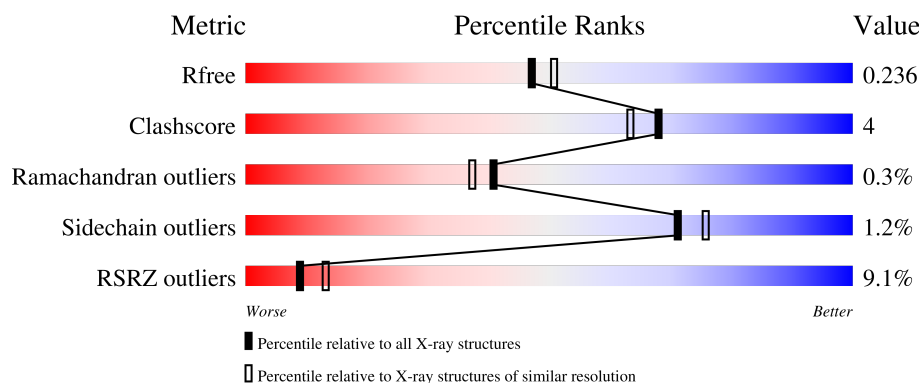
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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.



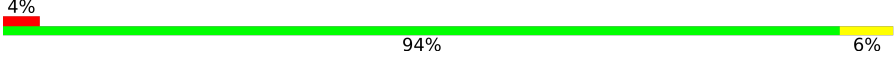
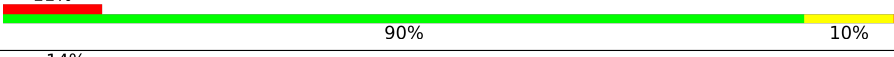
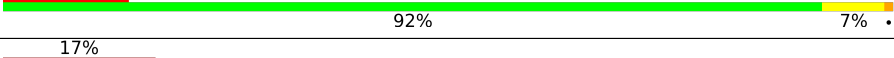

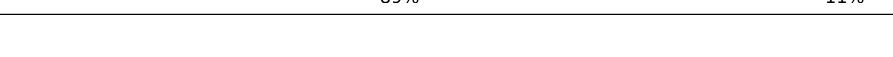
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	197	<div> <div>6%</div> <div>93%</div> <div>7%</div> </div>
1	B	197	<div> <div>5%</div> <div>92%</div> <div>8%</div> </div>
1	C	197	<div> <div>6%</div> <div>92%</div> <div>8%</div> </div>
1	D	197	<div> <div>12%</div> <div>92%</div> <div>7%</div> </div>
1	E	197	<div> <div>6%</div> <div>91%</div> <div>9%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	197	 4% 94% 6%
1	G	197	 11% 90% 10% .
1	H	197	 14% 92% 7% .
1	I	197	 17% 93% 6% .
1	J	197	 12% 89% 11% .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 33452 atoms, of which 15790 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 Peroxiredoxin-2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	197	Total	C	H	N	O	S	0	11	0
			3182	1020	1589	272	299	2			
1	B	197	Total	C	H	N	O	S	0	15	0
			3195	1026	1589	274	304	2			
1	C	197	Total	C	H	N	O	S	0	8	0
			3144	1010	1566	268	298	2			
1	D	197	Total	C	H	N	O	S	0	10	0
			3140	1010	1558	270	300	2			
1	E	197	Total	C	H	N	O	S	0	11	0
			3185	1021	1588	273	301	2			
1	F	197	Total	C	H	N	O	S	0	10	0
			3169	1016	1581	271	299	2			
1	G	197	Total	C	H	N	O	S	0	13	0
			3172	1017	1577	273	303	2			
1	H	197	Total	C	H	N	O	S	0	8	0
			3149	1011	1569	269	298	2			
1	I	197	Total	C	H	N	O	S	0	13	0
			3189	1021	1589	275	302	2			
1	J	197	Total	C	H	N	O	S	0	10	0
			3174	1017	1584	272	299	2			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	172	SER	CYS	conflict	UNP P32119
B	172	SER	CYS	conflict	UNP P32119
C	172	SER	CYS	conflict	UNP P32119
D	172	SER	CYS	conflict	UNP P32119
E	172	SER	CYS	conflict	UNP P32119
F	172	SER	CYS	conflict	UNP P32119
G	172	SER	CYS	conflict	UNP P32119
H	172	SER	CYS	conflict	UNP P32119
I	172	SER	CYS	conflict	UNP P32119

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	172	SER	CYS	conflict	UNP P32119

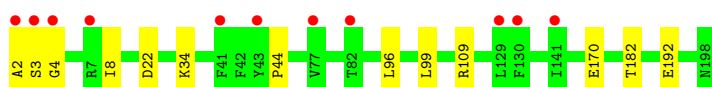
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	196	Total 196	O 196	0	7
2	B	183	Total 183	O 183	0	7
2	C	170	Total 170	O 170	0	6
2	D	171	Total 171	O 171	0	10
2	E	206	Total 206	O 206	0	10
2	F	208	Total 208	O 208	0	8
2	G	137	Total 137	O 137	0	8
2	H	149	Total 149	O 149	0	7
2	I	171	Total 171	O 171	0	6
2	J	162	Total 162	O 162	0	9

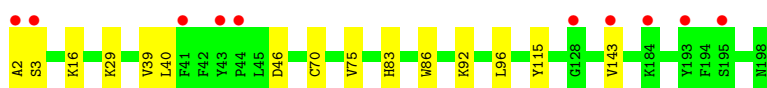
3 Residue-property plots [i](#)

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.

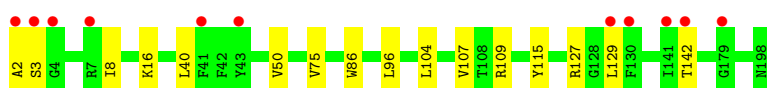
- Molecule 1: Peroxiredoxin-2



- Molecule 1: Peroxiredoxin-2



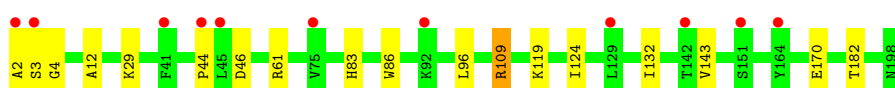
- Molecule 1: Peroxiredoxin-2



- Molecule 1: Peroxiredoxin-2



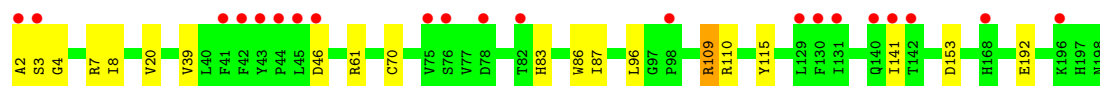
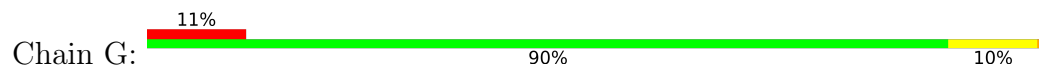
- Molecule 1: Peroxiredoxin-2



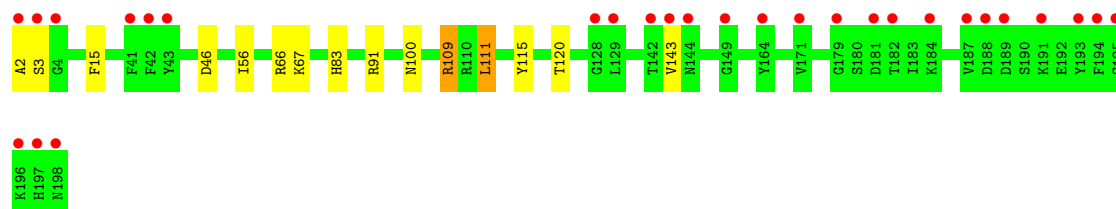
- Molecule 1: Peroxiredoxin-2



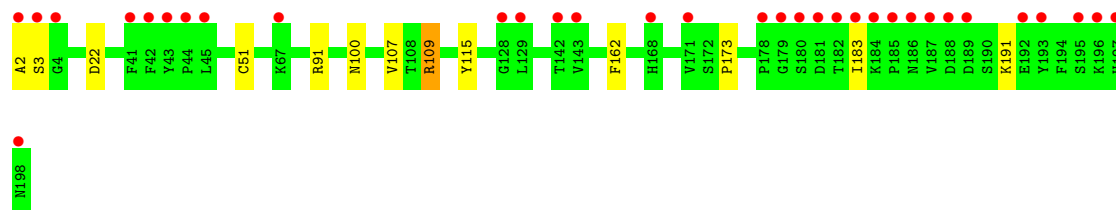
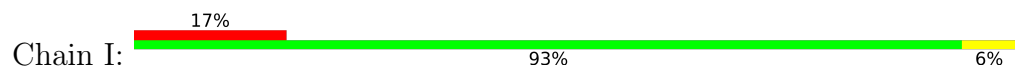
• Molecule 1: Peroxiredoxin-2



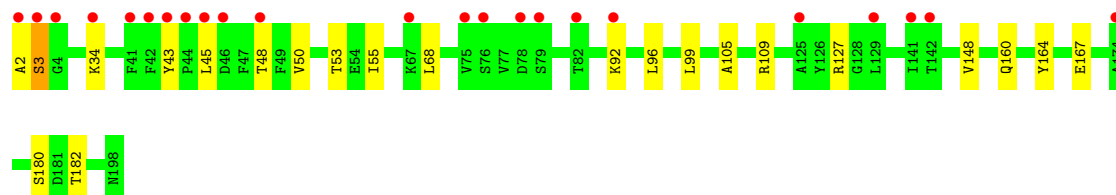
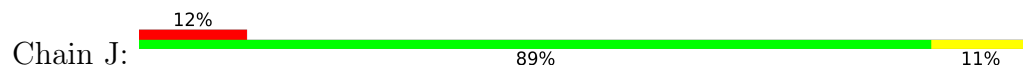
• Molecule 1: Peroxiredoxin-2



• Molecule 1: Peroxiredoxin-2



• Molecule 1: Peroxiredoxin-2



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	233.20Å 88.30Å 125.20Å 90.00° 99.20° 90.00°	Depositor
Resolution (Å)	35.24 – 2.15 35.24 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.6 (35.24-2.15) 98.7 (35.24-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.38 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.17.1 _3660	Depositor
R, R_{free}	0.195 , 0.236 0.195 , 0.236	Depositor DCC
R_{free} test set	6708 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	43.4	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 42.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	33452	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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

5.1 Standard geometry

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	0.25	0/1664	0.44	0/2251
1	B	0.24	0/1682	0.43	0/2278
1	C	0.24	0/1628	0.43	0/2206
1	D	0.25	0/1645	0.43	0/2230
1	E	0.25	0/1658	0.43	0/2245
1	F	0.25	0/1649	0.42	0/2232
1	G	0.25	0/1668	0.43	0/2258
1	H	0.24	0/1630	0.42	0/2208
1	I	0.24	0/1674	0.43	0/2267
1	J	0.25	0/1651	0.43	0/2234
All	All	0.25	0/16549	0.43	0/22409

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

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	1593	1589	1558	11	0
1	B	1606	1589	1551	10	0
1	C	1578	1566	1559	12	0
1	D	1582	1558	1525	16	0
1	E	1597	1588	1575	13	0
1	F	1588	1581	1560	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	1595	1577	1527	16	0
1	H	1580	1569	1563	10	0
1	I	1600	1589	1551	12	0
1	J	1590	1584	1563	16	0
2	A	196	0	0	9	0
2	B	183	0	0	1	1
2	C	170	0	0	6	0
2	D	171	0	0	6	0
2	E	206	0	0	6	0
2	F	208	0	0	8	0
2	G	137	0	0	5	0
2	H	149	0	0	5	0
2	I	171	0	0	6	0
2	J	162	0	0	4	1
All	All	17662	15790	15532	120	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 4.

The worst 5 of 120 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:E:2[A]:ALA:O	2:E:201:HOH:O	1.97	0.82
1:D:75[A]:VAL:HG12	1:D:104:LEU:HB2	1.64	0.80
1:E:2[A]:ALA:N	2:E:202[A]:HOH:O	2.15	0.79
1:F:2[B]:ALA:N	2:F:202[B]:HOH:O	2.16	0.78
1:H:2[B]:ALA:N	2:H:201[B]:HOH:O	2.16	0.78

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)
2:B:375:HOH:O	2:J:245:HOH:O[4_556]	2.08	0.12

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	205/197 (104%)	197 (96%)	5 (2%)	3 (2%)	10	5
1	B	209/197 (106%)	204 (98%)	5 (2%)	0	100	100
1	C	202/197 (102%)	194 (96%)	8 (4%)	0	100	100
1	D	204/197 (104%)	199 (98%)	5 (2%)	0	100	100
1	E	205/197 (104%)	195 (95%)	5 (2%)	5 (2%)	6	1
1	F	204/197 (104%)	199 (98%)	5 (2%)	0	100	100
1	G	207/197 (105%)	197 (95%)	8 (4%)	2 (1%)	15	9
1	H	202/197 (102%)	198 (98%)	4 (2%)	0	100	100
1	I	207/197 (105%)	199 (96%)	8 (4%)	0	100	100
1	J	204/197 (104%)	198 (97%)	6 (3%)	0	100	100
All	All	2049/1970 (104%)	1980 (97%)	59 (3%)	10 (0%)	41	22

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	4[A]	GLY
1	E	4[B]	GLY
1	E	3[A]	SER
1	E	3[B]	SER
1	G	4[A]	GLY

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	173/165 (105%)	172 (99%)	1 (1%)	86	90
1	B	174/165 (106%)	173 (99%)	1 (1%)	86	90
1	C	170/165 (103%)	169 (99%)	1 (1%)	86	90
1	D	171/165 (104%)	167 (98%)	4 (2%)	50	53
1	E	173/165 (105%)	170 (98%)	3 (2%)	60	65
1	F	172/165 (104%)	171 (99%)	1 (1%)	86	90
1	G	174/165 (106%)	170 (98%)	4 (2%)	50	53
1	H	170/165 (103%)	164 (96%)	6 (4%)	36	34
1	I	175/165 (106%)	172 (98%)	3 (2%)	60	65
1	J	172/165 (104%)	166 (96%)	6 (4%)	36	34
All	All	1724/1650 (104%)	1694 (98%)	30 (2%)	71	65

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	110[B]	ARG
1	J	92[A]	LYS
1	H	109[B]	ARG
1	J	167	GLU
1	J	3[A]	SER

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

Mol	Chain	Res	Type
1	B	197	HIS

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 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, 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	197/197 (100%)	0.36	11 (5%) 24 33	34, 43, 60, 74	0
1	B	197/197 (100%)	0.42	10 (5%) 28 36	35, 46, 63, 67	0
1	C	197/197 (100%)	0.37	11 (5%) 24 33	38, 48, 64, 80	0
1	D	197/197 (100%)	0.64	24 (12%) 4 6	38, 47, 62, 70	0
1	E	197/197 (100%)	0.34	11 (5%) 24 33	37, 47, 62, 71	0
1	F	197/197 (100%)	0.31	8 (4%) 37 46	35, 44, 66, 72	0
1	G	197/197 (100%)	0.60	21 (10%) 6 8	41, 53, 68, 75	0
1	H	197/197 (100%)	0.63	28 (14%) 2 3	37, 45, 69, 85	0
1	I	197/197 (100%)	0.80	33 (16%) 1 2	38, 49, 85, 98	0
1	J	197/197 (100%)	0.67	23 (11%) 4 6	41, 54, 71, 81	0
All	All	1970/1970 (100%)	0.52	180 (9%) 9 13	34, 48, 68, 98	0

The worst 5 of 180 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	3[A]	SER	5.9
1	H	3[A]	SER	5.4
1	G	3[A]	SER	5.3
1	A	3[A]	SER	5.2
1	C	3[A]	SER	5.1

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.