



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:44 PM GMT

PDB ID : 1YGR
Title : Crystal structure of the tandem phosphatase domain of RPTP CD45
Authors : Nam, H.J.; Poy, F.; Saito, H.; Frederick, C.A.
Deposited on : 2005-01-05
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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

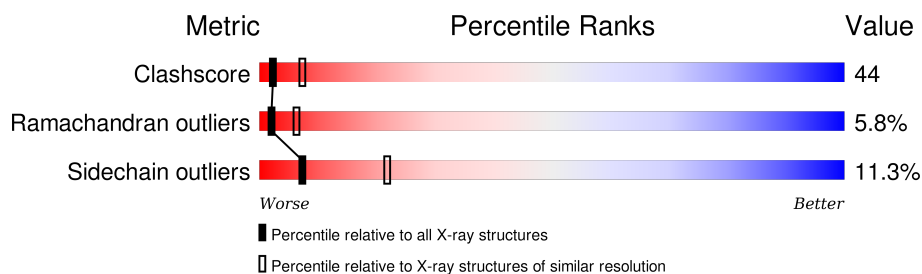
1 Overall quality at a glance

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))
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	610	 35% 49% 10% • 6%
1	B	610	 34% 48% 10% • 6%
2	C	6	 17% 67% 17%
2	D	6	 33% 67%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PTR	C	2004	X	-	-	-
2	PTR	D	2004	X	-	-	-

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9537 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 CD45 Protein Tyrosine Phosphatase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	574	Total	C	N	O	S	Se	0	0	0
			4709	2998	821	870	7	13			
1	B	574	Total	C	N	O	S	Se	0	0	0
			4706	2995	820	871	7	13			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	603	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	627	PRO	LEU	VARIANT	UNP P08575
A	714	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	725	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	744	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	828	SER	CYS	ENGINEERED	UNP P08575
A	844	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	870	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	907	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	999	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	1007	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	1024	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	1035	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	1115	MSE	MET	MODIFIED RESIDUE	UNP P08575
A	1184	LEU	PRO	VARIANT	UNP P08575
A	1186	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	603	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	627	PRO	LEU	VARIANT	UNP P08575
B	714	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	725	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	744	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	828	SER	CYS	ENGINEERED	UNP P08575
B	844	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	870	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	907	MSE	MET	MODIFIED RESIDUE	UNP P08575

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	999	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	1007	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	1024	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	1035	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	1115	MSE	MET	MODIFIED RESIDUE	UNP P08575
B	1184	LEU	PRO	VARIANT	UNP P08575
B	1186	MSE	MET	MODIFIED RESIDUE	UNP P08575

- Molecule 2 is a protein called T-cell Receptor CD3 zeta ITAM-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	P	0	0	0
			61	34	9	17	1			
2	D	6	Total	C	N	O	P	0	0	0
			61	34	9	17	1			

There are 2 discrepancies between the modelled and reference sequences:

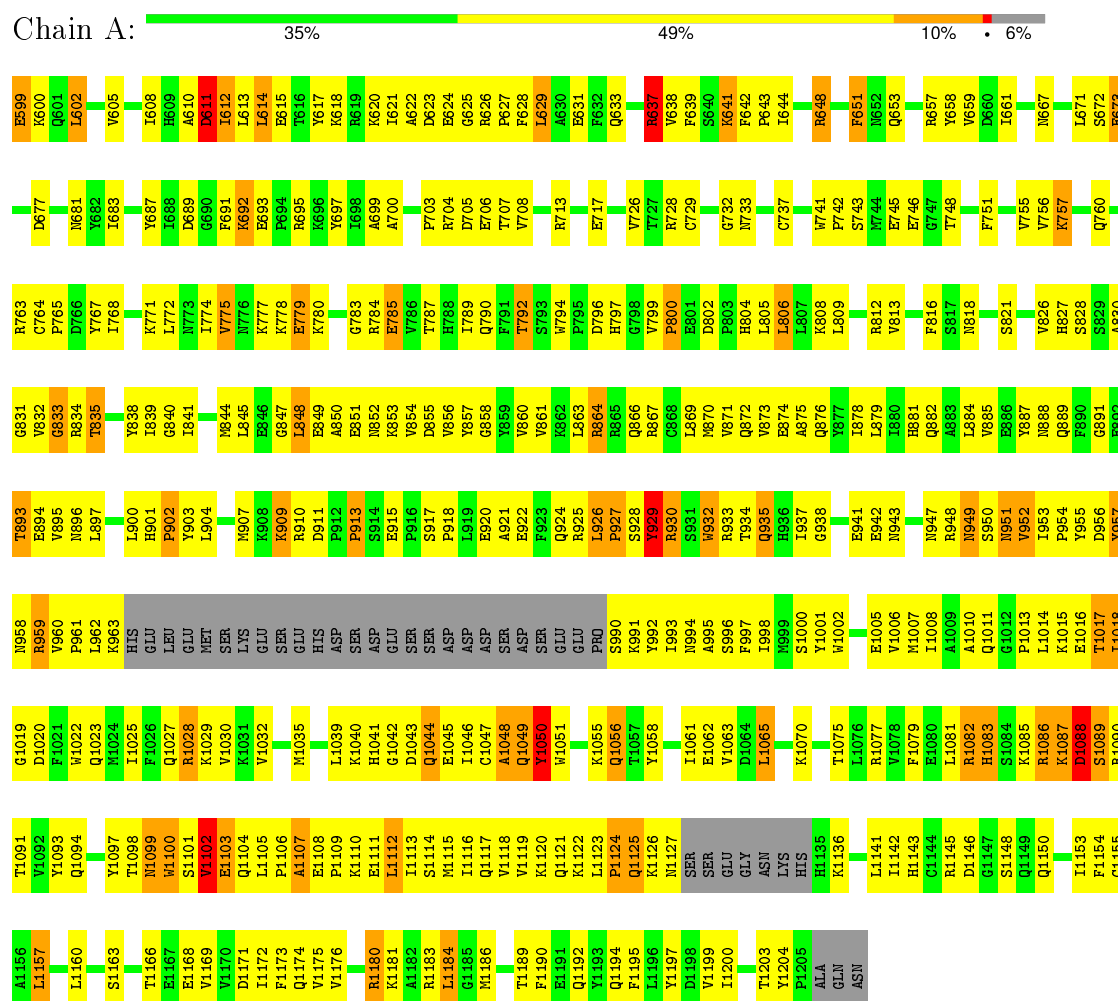
Chain	Residue	Modelled	Actual	Comment	Reference
C	2004	PTR	TYR	MODIFIED RESIDUE	UNP P20963
D	2004	PTR	TYR	MODIFIED RESIDUE	UNP P20963

3 Residue-property plots

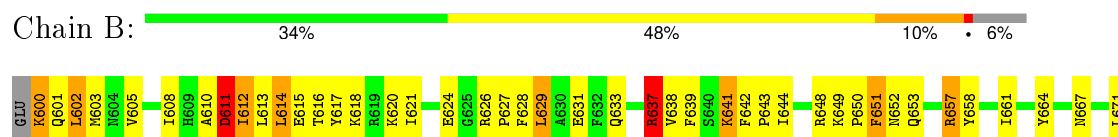
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 ($\text{RSRZ} > 2$). 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.

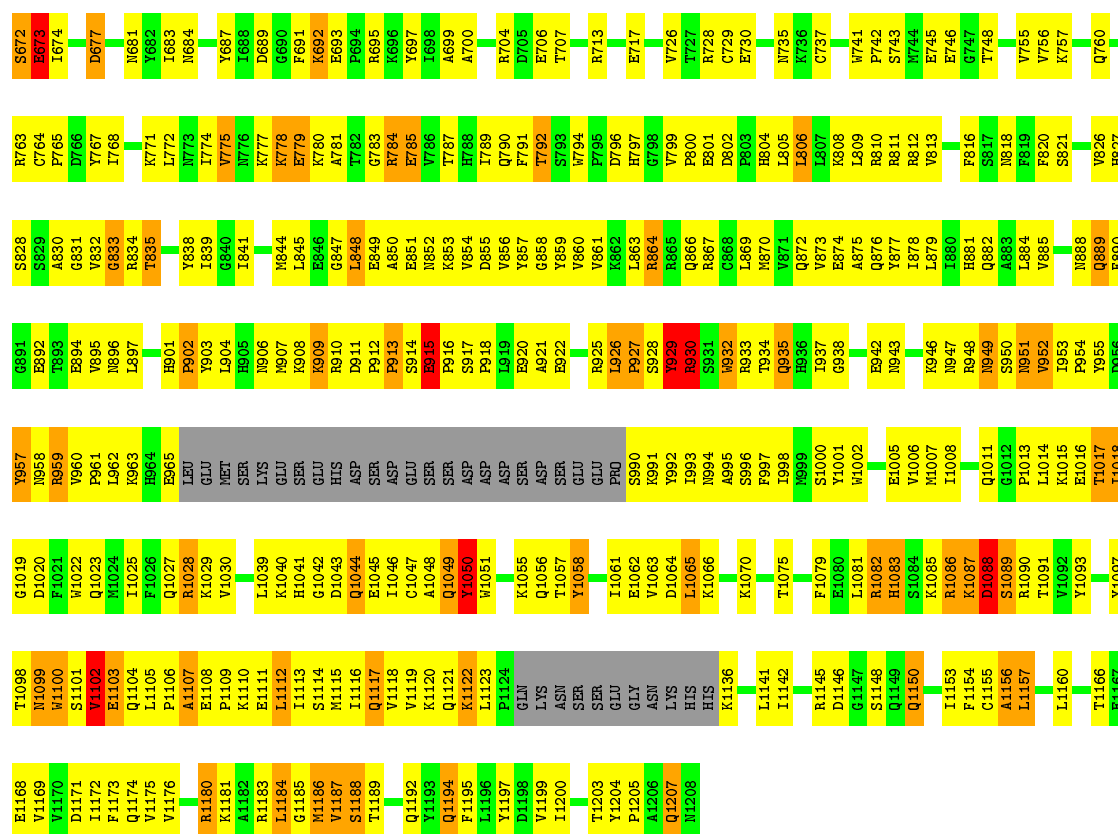
Note EDS was not executed.

• Molecule 1: CD45 Protein Tyrosine Phosphatase



• Molecule 1: CD45 Protein Tyrosine Phosphatase





4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	86.01 Å 59.70 Å 160.53 Å 90.00° 99.53° 90.00°	Depositor
Resolution (Å)	29.85 – 2.90	Depositor
% Data completeness (in resolution range)	81.5 (29.85-2.90)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.255 , 0.312	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9537	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

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: PTR

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	0.51	0/4805	0.72	1/6474 (0.0%)
1	B	0.53	0/4802	0.71	1/6470 (0.0%)
2	C	0.74	0/43	0.79	0/53
2	D	0.79	0/43	0.75	0/53
All	All	0.52	0/9693	0.71	2/13050 (0.0%)

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
2	C	1	0
2	D	1	0
All	All	2	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1088	ASP	N-CA-C	-5.36	96.53	111.00
1	A	1088	ASP	N-CA-C	-5.25	96.81	111.00

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	2004	PTR	CA
2	D	2004	PTR	CA

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	4709	0	4670	398	0
1	B	4706	0	4662	422	0
2	C	61	0	43	8	0
2	D	61	0	43	3	0
All	All	9537	0	9418	826	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 44.

The worst 5 of 826 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:929:TYR:HB3	1:A:932:TRP:HB2	1.22	1.16
1:B:929:TYR:HB3	1:B:932:TRP:HB2	1.23	1.12
1:B:637:ARG:HB3	1:B:637:ARG:HH11	1.07	1.11
1:A:637:ARG:HB3	1:A:637:ARG:HH11	1.09	1.11
1:A:911:ASP:O	1:A:913:PRO:HD3	1.53	1.07

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	568/610 (93%)	457 (80%)	80 (14%)	31 (6%)	2 7
1	B	568/610 (93%)	450 (79%)	84 (15%)	34 (6%)	2 6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	3/6 (50%)	2 (67%)	0	1 (33%)	0	0
2	D	3/6 (50%)	2 (67%)	1 (33%)	0	100	100
All	All	1142/1232 (93%)	911 (80%)	165 (14%)	66 (6%)	2	6

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	779	GLU
1	A	929	TYR
1	A	932	TRP
1	A	1015	LYS
1	A	1086	ARG

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	517/538 (96%)	461 (89%)	56 (11%)	8	24
1	B	516/538 (96%)	455 (88%)	61 (12%)	6	19
2	C	5/5 (100%)	5 (100%)	0	100	100
2	D	5/5 (100%)	4 (80%)	1 (20%)	1	5
All	All	1043/1086 (96%)	925 (89%)	118 (11%)	7	22

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1184	LEU
1	B	673	GLU
1	B	1117	GLN
1	A	1189	THR
1	B	614	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1121	GLN
1	B	770	GLN
1	B	1117	GLN
1	A	1150	GLN
1	B	797	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\# Z > 2$
2	PTR	C	2004	2	14,16,17	1.66	4 (28%)	18,22,24	1.48	3 (16%)
2	PTR	D	2004	2	14,16,17	1.69	4 (28%)	18,22,24	2.57	6 (33%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PTR	C	2004	2	1/1/2/3	0/9/11/13	0/1/1/1
2	PTR	D	2004	2	1/1/2/3	0/9/11/13	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2004	PTR	P-O3P	-2.92	1.44	1.54
2	D	2004	PTR	P-O3P	-2.91	1.44	1.54
2	C	2004	PTR	P-O2P	-2.43	1.46	1.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2004	PTR	CD2-CG	2.08	1.43	1.38
2	D	2004	PTR	CD2-CG	2.31	1.43	1.38

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	2004	PTR	OH-CZ-CE1	-4.21	106.24	119.22
2	D	2004	PTR	CE2-CD2-CG	-2.32	117.86	121.04
2	C	2004	PTR	O3P-P-O1P	-2.28	103.23	110.58
2	C	2004	PTR	OH-CZ-CE2	2.09	125.67	119.22
2	C	2004	PTR	O3P-P-O2P	2.37	116.39	107.38

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	2004	PTR	CA
2	D	2004	PTR	CA

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2004	PTR	1	0
2	D	2004	PTR	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates 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 ⓘ

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.