Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 07:35 am BST

PDB ID	:	4RS6
Title	:	Crystal structure of the C domain of Polo like Kinase II in Homo Sapiens
Authors	:	Shan, H.; Quan, J.; Wang, T.
Deposited on	:	2014-11-07
Resolution	:	2.60 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 following versions of software and data (see references ①) were used in the production of this report:

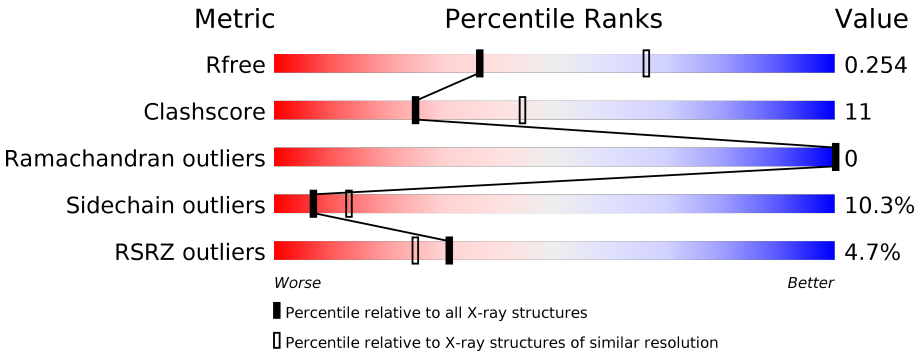
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 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 <=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	242	3% 69% 15% • 12%
1	B	242	5% 70% 13% 5% 12%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3462 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 Serine/threonine-protein kinase PLK2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	0	0
			1724	1100	283	327	14			
1	B	213	Total	C	N	O	S	0	0	0
			1724	1100	283	327	14			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	444	GLY	-	EXPRESSION TAG	UNP Q9NYY3
A	445	PRO	-	EXPRESSION TAG	UNP Q9NYY3
A	446	LEU	-	EXPRESSION TAG	UNP Q9NYY3
A	447	GLY	-	EXPRESSION TAG	UNP Q9NYY3
A	448	SER	-	EXPRESSION TAG	UNP Q9NYY3
A	449	PRO	-	EXPRESSION TAG	UNP Q9NYY3
A	450	GLU	-	EXPRESSION TAG	UNP Q9NYY3
B	444	GLY	-	EXPRESSION TAG	UNP Q9NYY3
B	445	PRO	-	EXPRESSION TAG	UNP Q9NYY3
B	446	LEU	-	EXPRESSION TAG	UNP Q9NYY3
B	447	GLY	-	EXPRESSION TAG	UNP Q9NYY3
B	448	SER	-	EXPRESSION TAG	UNP Q9NYY3
B	449	PRO	-	EXPRESSION TAG	UNP Q9NYY3
B	450	GLU	-	EXPRESSION TAG	UNP Q9NYY3

- Molecule 2 is water.

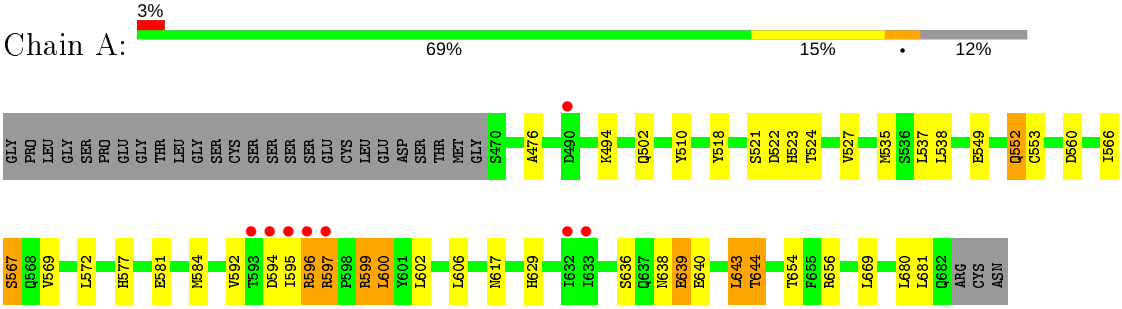
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	11	Total	O	0	0
			11	11		
2	B	3	Total	O	0	0
			3	3		

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 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.

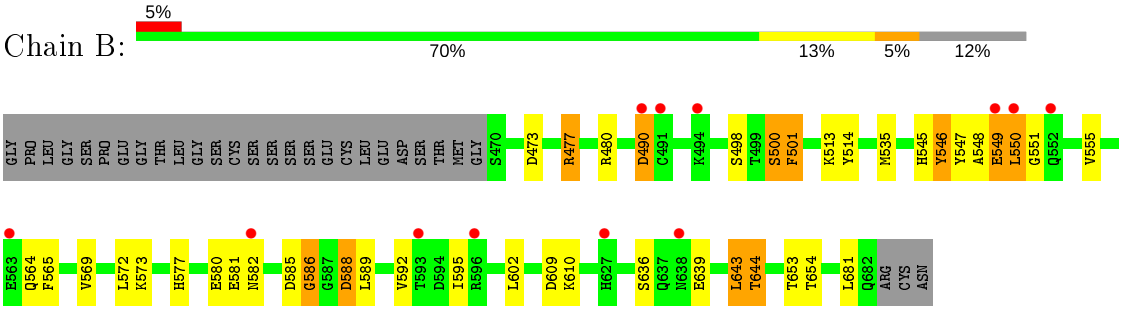
- Molecule 1: Serine/threonine-protein kinase PLK2

Chain A:



- Molecule 1: Serine/threonine-protein kinase PLK2

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	153.07Å 153.07Å 153.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.83 – 2.60 24.83 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.8 (24.83-2.60) 100.0 (24.83-2.60)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	5.31 (at 2.60Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.186 , 0.247 0.213 , 0.254	Depositor DCC
R_{free} test set	948 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	62.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 51.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.027 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3462	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.50% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.80	0/1764	0.88	2/2391 (0.1%)
1	B	0.97	3/1764 (0.2%)	0.97	6/2391 (0.3%)
All	All	0.89	3/3528 (0.1%)	0.92	8/4782 (0.2%)

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	B	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	586	GLY	CA-C	-18.02	1.23	1.51
1	B	585	ASP	CA-CB	-10.01	1.31	1.53
1	B	585	ASP	N-CA	-8.57	1.29	1.46

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	585	ASP	N-CA-CB	9.74	128.14	110.60
1	B	586	GLY	N-CA-C	9.03	135.67	113.10
1	B	588	ASP	N-CA-C	8.23	133.22	111.00
1	B	480	ARG	NE-CZ-NH1	7.41	124.00	120.30
1	A	597	ARG	N-CA-C	-6.80	92.64	111.00
1	B	477	ARG	NE-CZ-NH2	-6.61	117.00	120.30
1	B	589	LEU	C-N-CD	5.87	140.72	128.40
1	A	643	LEU	CA-CB-CG	5.75	128.52	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	586	GLY	Mainchain

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	1724	0	1679	37	0
1	B	1724	0	1679	36	0
2	A	11	0	0	0	0
2	B	3	0	0	0	0
All	All	3462	0	3358	73	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 11.

All (73) 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:592:VAL:CG2	1:A:595:ILE:HD11	1.54	1.35
1:A:592:VAL:CG2	1:A:595:ILE:CD1	2.24	1.16
1:A:592:VAL:HG23	1:A:595:ILE:CD1	1.82	1.06
1:B:550:LEU:H	1:B:550:LEU:HD13	1.24	1.00
1:A:596:ARG:HH11	1:A:596:ARG:HB3	1.27	0.98
1:A:592:VAL:HG21	1:A:595:ILE:CD1	1.98	0.93
1:B:550:LEU:CD1	1:B:550:LEU:N	2.30	0.92
1:A:522:ASP:OD1	1:A:524:THR:OG1	1.90	0.90
1:A:592:VAL:HG23	1:A:595:ILE:HD11	0.90	0.89
1:B:550:LEU:N	1:B:550:LEU:HD13	1.84	0.88
1:A:535:MET:CE	1:A:572:LEU:HD22	2.04	0.87
1:A:599:ARG:HH11	1:A:599:ARG:HG3	1.43	0.83
1:B:546:TYR:CD1	1:B:546:TYR:C	2.51	0.82
1:B:577:HIS:O	1:B:580:GLU:HG2	1.79	0.82
1:A:644:THR:HG22	1:A:654:THR:HG23	1.60	0.82
1:A:592:VAL:HG21	1:A:595:ILE:HD12	1.61	0.80
1:A:596:ARG:CB	1:A:596:ARG:HH11	1.94	0.80

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:549:GLU:OE1	1:B:550:LEU:CD1	2.30	0.80
1:B:549:GLU:OE1	1:B:550:LEU:HD13	1.82	0.79
1:A:596:ARG:NH1	1:A:596:ARG:HB3	2.00	0.75
1:A:595:ILE:HD13	1:A:595:ILE:N	2.04	0.71
1:A:599:ARG:HH11	1:A:599:ARG:CG	2.05	0.68
1:B:581:GLU:HG3	1:B:582:ASN:H	1.59	0.68
1:A:595:ILE:O	1:A:596:ARG:HD2	1.95	0.67
1:A:594:ASP:C	1:A:595:ILE:HD13	2.15	0.66
1:A:577:HIS:O	1:A:581:GLU:HG2	1.96	0.66
1:B:577:HIS:CD2	1:B:580:GLU:OE2	2.50	0.64
1:B:548:ALA:O	1:B:551:GLY:N	2.30	0.62
1:B:535:MET:CE	1:B:572:LEU:HD22	2.29	0.62
1:B:549:GLU:OE1	1:B:550:LEU:HD11	2.00	0.61
1:A:535:MET:HE2	1:A:572:LEU:HD22	1.82	0.60
1:B:546:TYR:HD1	1:B:546:TYR:C	2.03	0.59
1:B:581:GLU:HG3	1:B:582:ASN:N	2.17	0.59
1:B:550:LEU:N	1:B:550:LEU:HD12	2.15	0.59
1:A:535:MET:HE3	1:A:572:LEU:HD22	1.85	0.59
1:A:600:LEU:HA	1:A:617:ASN:OD1	2.04	0.58
1:B:535:MET:HE1	1:B:569:VAL:HA	1.86	0.58
1:A:502:GLN:OE1	1:A:521:SER:HB3	2.04	0.57
1:B:546:TYR:CE2	1:B:565:PHE:CE1	2.93	0.56
1:A:552:GLN:OE1	1:A:552:GLN:HA	2.05	0.56
1:B:581:GLU:CG	1:B:582:ASN:H	2.19	0.56
1:B:546:TYR:HD1	1:B:546:TYR:O	1.89	0.55
1:B:564:GLN:OE1	1:B:564:GLN:N	2.36	0.54
1:A:595:ILE:C	1:A:596:ARG:HD2	2.28	0.54
1:B:535:MET:HE3	1:B:572:LEU:HD22	1.90	0.52
1:A:640:GLU:OE2	1:A:656:ARG:NH1	2.42	0.52
1:A:537:LEU:HB2	1:A:572:LEU:HD21	1.93	0.51
1:B:501:PHE:N	1:B:501:PHE:CD1	2.79	0.51
1:A:552:GLN:OE1	1:A:553:CYS:N	2.44	0.51
1:A:535:MET:HE1	1:A:569:VAL:HA	1.92	0.50
1:A:476:ALA:HB2	1:A:680:LEU:HD13	1.94	0.49
1:B:577:HIS:HD2	1:B:580:GLU:OE2	1.96	0.49
1:A:599:ARG:NH1	1:A:599:ARG:CG	2.73	0.49
1:A:638:ASN:O	1:A:639:GLU:HB2	2.12	0.49
1:B:490:ASP:OD1	1:B:490:ASP:N	2.45	0.49
1:B:513:LYS:HE3	1:B:514:TYR:CZ	2.49	0.48
1:B:581:GLU:CG	1:B:582:ASN:N	2.77	0.48
1:A:566:ILE:O	1:A:567:SER:C	2.52	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:546:TYR:CD1	1:B:547:TYR:N	2.82	0.46
1:B:609:ASP:OD1	1:B:609:ASP:N	2.42	0.46
1:A:535:MET:HE2	1:A:572:LEU:HB2	1.99	0.45
1:B:545:HIS:CD2	1:B:555:VAL:HG22	2.52	0.45
1:A:522:ASP:O	1:A:523:HIS:HB2	2.18	0.44
1:B:473:ASP:O	1:B:477:ARG:HG3	2.18	0.44
1:B:643:LEU:CD1	1:B:643:LEU:C	2.86	0.44
1:B:500:SER:HA	1:B:501:PHE:HD1	1.82	0.43
1:A:600:LEU:N	1:A:600:LEU:HD23	2.34	0.42
1:A:636:SER:HA	1:A:640:GLU:O	2.19	0.42
1:B:535:MET:HE2	1:B:572:LEU:HD22	2.01	0.41
1:B:501:PHE:N	1:B:501:PHE:HD1	2.19	0.41
1:A:518:TYR:CD1	1:A:518:TYR:N	2.89	0.40
1:B:644:THR:HG22	1:B:654:THR:HG23	2.04	0.40
1:B:546:TYR:CE1	1:B:548:ALA:N	2.90	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	211/242 (87%)	196 (93%)	15 (7%)	0	100	100
1	B	211/242 (87%)	196 (93%)	15 (7%)	0	100	100
All	All	422/484 (87%)	392 (93%)	30 (7%)	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	194/218 (89%)	173 (89%)	21 (11%)	6 12
1	B	194/218 (89%)	175 (90%)	19 (10%)	8 15
All	All	388/436 (89%)	348 (90%)	40 (10%)	7 13

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	494	LYS
1	A	510	TYR
1	A	527	VAL
1	A	538	LEU
1	A	549	GLU
1	A	552	GLN
1	A	560	ASP
1	A	567	SER
1	A	584	MET
1	A	596	ARG
1	A	597	ARG
1	A	599	ARG
1	A	600	LEU
1	A	602	LEU
1	A	606	LEU
1	A	629	HIS
1	A	639	GLU
1	A	643	LEU
1	A	644	THR
1	A	669	LEU
1	A	681	LEU
1	B	490	ASP
1	B	498	SER
1	B	500	SER
1	B	501	PHE
1	B	546	TYR
1	B	549	GLU
1	B	550	LEU
1	B	573	LYS
1	B	588	ASP
1	B	592	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	595	ILE
1	B	602	LEU
1	B	610	LYS
1	B	636	SER
1	B	639	GLU
1	B	643	LEU
1	B	644	THR
1	B	653	THR
1	B	681	LEU

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

Mol	Chain	Res	Type
1	A	523	HIS
1	A	545	HIS
1	B	577	HIS
1	B	671	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

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

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, 95^{th} 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	213/242 (88%)	-0.06	8 (3%)	40	33	24, 63, 96, 136	0
1	B	213/242 (88%)	0.11	12 (5%)	24	19	45, 70, 122, 152	0
All	All	426/484 (88%)	0.02	20 (4%)	31	25	24, 67, 112, 152	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	491	CYS	6.3
1	A	595	ILE	5.7
1	A	596	ARG	5.6
1	A	593	THR	5.3
1	A	597	ARG	5.0
1	B	550	LEU	4.2
1	B	490	ASP	3.9
1	B	552	GLN	3.7
1	A	594	ASP	3.6
1	B	494	LYS	3.2
1	B	549	GLU	2.8
1	A	632	ILE	2.5
1	B	627	HIS	2.4
1	B	593	THR	2.4
1	B	638	ASN	2.2
1	A	490	ASP	2.2
1	B	582	ASN	2.2
1	B	596	ARG	2.2
1	B	563	GLU	2.2
1	A	633	ILE	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.