Full wwPDB X-ray Structure Validation Report (i)

Aug 20, 2026 – 12:08 PM EDT

PDB ID	:	9XYT / pdb_00009xyt
Title	:	Structure of an ancestral glucokinase
Authors	:	Kamalaldinezabadi, S.; Li, H.; Miller, B.
Deposited on	:	2025-08-26
Resolution	:	2.00 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

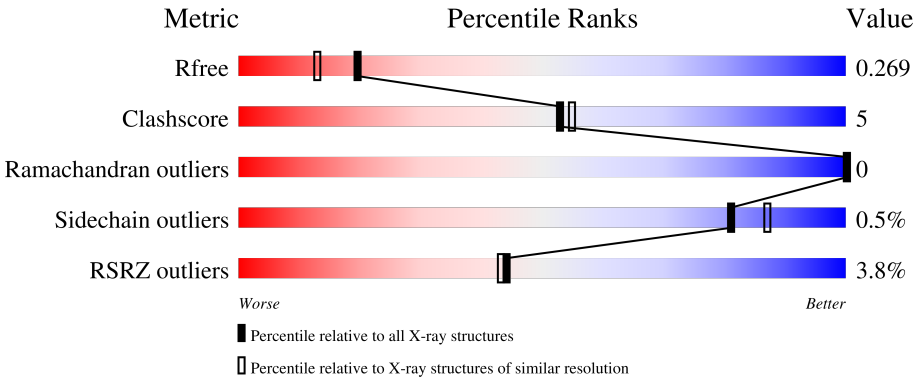
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

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 2.00 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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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 <=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	465	5% 83% 9% 8%
1	B	465	3% 85% 8% • 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 13421 atoms, of which 6638 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 Reconstructed ancestral glucokinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	438	Total	C	H	N	O	S	0	0	0
			6724	2112	3352	564	666	30			
1	A	429	Total	C	H	N	O	S	0	0	0
			6588	2071	3286	552	649	30			

- Molecule 2 is water.

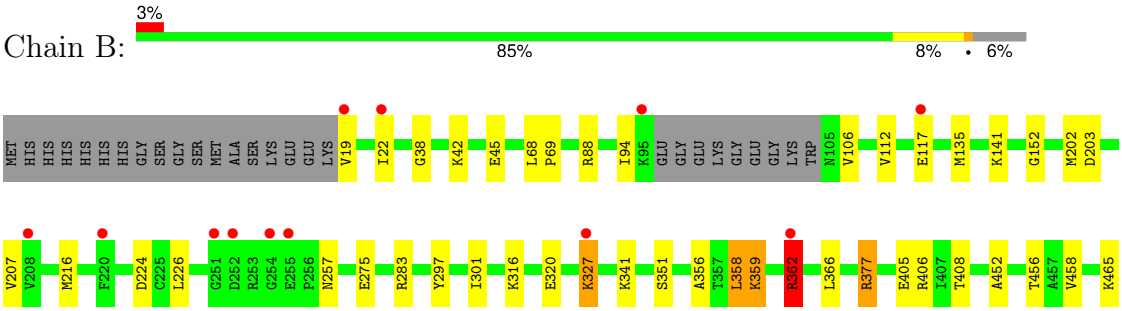
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	73	Total	O	0	0
			73	73		
2	A	36	Total	O	0	0
			36	36		

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.

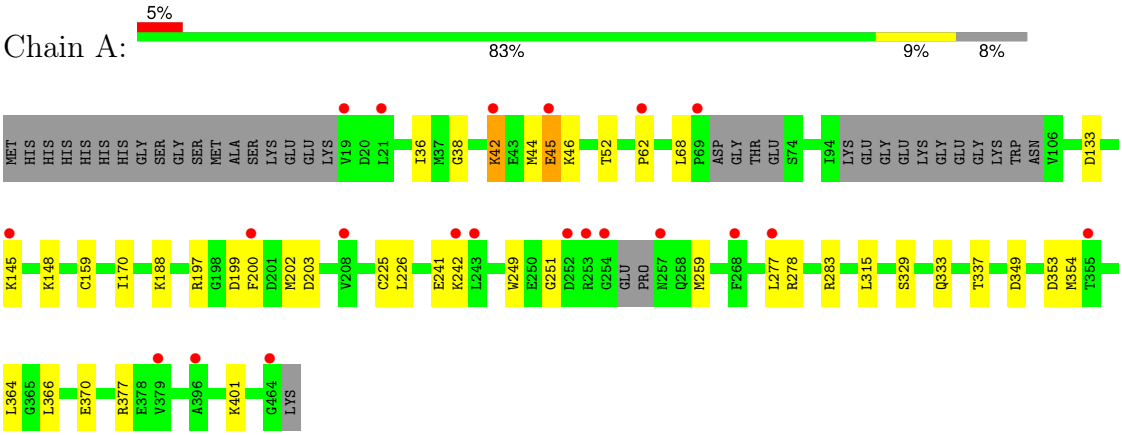
• Molecule 1: Reconstructed ancestral glucokinase

Chain B:



• Molecule 1: Reconstructed ancestral glucokinase

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.30Å 92.83Å 123.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.42 – 2.00 46.42 – 2.00	Depositor EDS
% Data completeness (in resolution range)	86.3 (46.42-2.00) 87.5 (46.42-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.95 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.229 , 0.268 0.230 , 0.269	Depositor DCC
R_{free} test set	2004 reflections (3.01%)	wwPDB-VP
Wilson B-factor (Å^2)	46.7	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 31.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	13421	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.59% 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 (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	0.42	1/3345 (0.0%)	0.71	9/4497 (0.2%)
1	B	0.41	5/3418 (0.1%)	0.69	11/4598 (0.2%)
All	All	0.42	6/6763 (0.1%)	0.70	20/9095 (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	A	0	1
1	B	0	2
All	All	0	3

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	45	GLU	N-CA	8.01	1.56	1.46
1	B	327	LYS	CA-C	6.59	1.60	1.52
1	B	362	ARG	NE-CZ	-6.10	1.26	1.33
1	B	362	ARG	N-CA	-6.09	1.38	1.46
1	B	359	LYS	CE-NZ	5.11	1.64	1.49
1	B	359	LYS	CD-CE	5.05	1.67	1.52

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	45	GLU	CA-CB-CG	-15.91	82.28	114.10
1	B	359	LYS	CB-CG-CD	13.31	141.91	111.30
1	A	45	GLU	CB-CG-CD	11.95	132.92	112.60
1	B	359	LYS	CG-CD-CE	-11.65	84.49	111.30
1	A	45	GLU	N-CA-C	-10.45	99.21	112.90
1	B	359	LYS	N-CA-CB	-8.86	97.17	110.01

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	359	LYS	CA-CB-CG	6.35	126.80	114.10
1	B	117	GLU	CB-CA-C	-6.09	101.07	110.81
1	A	42	LYS	CB-CG-CD	-5.96	97.58	111.30
1	B	341	LYS	CD-CE-NZ	5.93	130.86	111.90
1	A	188	LYS	CB-CG-CD	5.81	124.67	111.30
1	B	117	GLU	N-CA-C	-5.63	104.35	111.11
1	A	44	MET	CA-C-N	-5.62	111.17	121.52
1	A	44	MET	C-N-CA	-5.62	111.17	121.52
1	A	242	LYS	N-CA-CB	-5.59	101.52	109.85
1	B	327	LYS	CA-CB-CG	-5.57	102.96	114.10
1	B	117	GLU	CA-CB-CG	5.54	125.17	114.10
1	B	362	ARG	N-CA-C	-5.26	105.13	112.45
1	A	42	LYS	CA-CB-CG	5.25	124.59	114.10
1	B	377	ARG	CD-NE-CZ	5.21	131.69	124.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	45	GLU	Sidechain
1	B	358	LEU	Mainchain
1	B	362	ARG	Sidechain

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	3302	3286	3283	32	0
1	B	3372	3352	3350	35	0
2	A	36	0	0	0	0
2	B	73	0	0	2	0
All	All	6783	6638	6633	66	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 5.

All (66) 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:B:358:LEU:HB3	1:B:362:ARG:NH1	1.80	0.96
1:A:42:LYS:N	1:A:42:LYS:HD3	1.90	0.85
1:A:133:ASP:OD1	1:A:197:ARG:NH1	2.11	0.83
1:A:145:LYS:NZ	1:A:199:ASP:OD2	2.16	0.77
1:A:62:PRO:HB3	1:A:259:MET:HE3	1.66	0.75
1:A:42:LYS:HD3	1:A:42:LYS:H	1.51	0.74
1:B:45:GLU:OE2	2:B:501:HOH:O	2.05	0.73
1:A:68:LEU:HD11	1:A:226:LEU:HD23	1.72	0.71
1:B:358:LEU:HD22	1:B:362:ARG:NH1	2.06	0.70
1:B:358:LEU:HB3	1:B:362:ARG:HH12	1.57	0.70
1:B:358:LEU:O	1:B:362:ARG:HD3	1.92	0.69
1:B:358:LEU:HD22	1:B:362:ARG:CZ	2.24	0.68
1:B:19:VAL:HA	1:B:22:ILE:HD13	1.76	0.67
1:B:358:LEU:CB	1:B:362:ARG:NH1	2.57	0.67
1:B:358:LEU:HB3	1:B:362:ARG:HH11	1.60	0.66
1:A:62:PRO:CB	1:A:259:MET:HE3	2.29	0.63
1:B:68:LEU:HD11	1:B:226:LEU:HD23	1.84	0.59
1:A:38:GLY:C	1:A:42:LYS:HZ3	2.11	0.58
1:A:249:TRP:NE1	1:A:251:GLY:O	2.37	0.58
1:A:38:GLY:C	1:A:42:LYS:NZ	2.63	0.57
1:A:38:GLY:CA	1:A:42:LYS:HZ3	2.16	0.57
1:A:197:ARG:NH2	1:A:199:ASP:OD2	2.38	0.56
1:A:38:GLY:O	1:A:42:LYS:CE	2.55	0.55
1:A:241:GLU:OE2	1:A:401:LYS:NZ	2.38	0.55
1:A:337:THR:CG2	1:A:364:LEU:HD22	2.37	0.54
1:A:145:LYS:HG3	1:A:200:PHE:HB3	1.89	0.54
1:B:452:ALA:O	1:B:456:THR:HG23	2.08	0.54
1:A:337:THR:HG23	1:A:364:LEU:HD22	1.90	0.53
1:A:38:GLY:O	1:A:42:LYS:HE2	2.08	0.53
1:B:94:ILE:HG22	1:B:106:VAL:HG22	1.90	0.52
1:A:278:ARG:O	1:A:283:ARG:NH2	2.43	0.52
1:B:224:ASP:O	1:B:408:THR:N	2.39	0.51
1:B:22:ILE:HD12	1:B:22:ILE:H	1.76	0.51
1:A:38:GLY:O	1:A:42:LYS:NZ	2.44	0.50
1:A:38:GLY:CA	1:A:42:LYS:NZ	2.75	0.49
1:B:152:GLY:CA	1:B:456:THR:HG21	2.43	0.49
1:A:159:CYS:HA	1:A:170:ILE:HG13	1.95	0.49
1:B:216:MET:HE1	1:B:458:VAL:HG21	1.94	0.49
1:B:362:ARG:N	1:B:362:ARG:HD2	2.28	0.48
1:A:36:ILE:HD11	1:A:277:LEU:CD1	2.44	0.47
1:A:148:LYS:HG2	1:A:202:MET:HA	1.96	0.47
1:B:358:LEU:CD2	1:B:362:ARG:NH1	2.77	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:46:LYS:O	1:A:52:THR:HG23	2.17	0.44
1:A:354:MET:HE3	1:A:370:GLU:HG2	2.00	0.44
1:B:141:LYS:NZ	1:A:353:ASP:OD1	2.48	0.43
1:B:362:ARG:HD2	1:B:366:LEU:O	2.18	0.43
1:A:68:LEU:HD11	1:A:226:LEU:CD2	2.46	0.43
1:B:257:ASN:OD1	1:B:257:ASN:N	2.52	0.43
1:B:69:PRO:HD3	1:B:216:MET:CE	2.49	0.43
1:B:362:ARG:N	1:B:362:ARG:CD	2.81	0.43
1:B:203:ASP:OD2	1:B:465:LYS:NZ	2.52	0.42
1:B:405:GLU:HB3	1:B:406:ARG:NH2	2.33	0.42
1:B:316:LYS:O	1:B:320:GLU:HG2	2.18	0.42
1:B:356:ALA:O	1:B:359:LYS:HB3	2.18	0.42
1:B:297:TYR:O	1:B:301:ILE:HG12	2.20	0.42
1:B:207:VAL:HG13	1:B:207:VAL:O	2.20	0.42
1:A:349:ASP:HB3	1:A:377:ARG:HH21	1.85	0.42
1:B:38:GLY:O	1:B:42:LYS:HD3	2.20	0.41
1:A:36:ILE:HD11	1:A:277:LEU:HD11	2.03	0.41
1:B:135:MET:HB3	1:B:202:MET:HE1	2.02	0.41
1:B:275:GLU:OE1	1:B:283:ARG:NH1	2.54	0.41
1:B:358:LEU:HD22	1:B:362:ARG:NH2	2.35	0.41
1:B:377:ARG:NH2	2:B:511:HOH:O	2.51	0.41
1:A:315:LEU:HD11	1:A:333:GLN:HA	2.03	0.40
1:A:329:SER:HB3	1:A:366:LEU:HD23	2.03	0.40
1:B:88:ARG:HG3	1:B:112:VAL:HG12	2.03	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	421/465 (90%)	409 (97%)	12 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	434/465 (93%)	424 (98%)	10 (2%)	0	100 100
All	All	855/930 (92%)	833 (97%)	22 (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	361/390 (93%)	359 (99%)	2 (1%)	78 85
1	B	369/390 (95%)	367 (100%)	2 (0%)	81 87
All	All	730/780 (94%)	726 (100%)	4 (0%)	81 87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	327	LYS
1	B	351	SER
1	A	203	ASP
1	A	225	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	53	ASN
1	B	111	GLN
1	B	162	ASN
1	A	34	ASN
1	A	162	ASN
1	A	339	GLN
1	A	421	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 oligosaccharides 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	429/465 (92%)	0.70	21 (4%) 35 34	45, 63, 89, 121	0
1	B	438/465 (94%)	0.32	12 (2%) 56 55	36, 52, 78, 114	0
All	All	867/930 (93%)	0.51	33 (3%) 44 43	36, 58, 85, 121	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	464	GLY	5.0
1	A	254	GLY	3.7
1	A	145	LYS	3.7
1	A	277	LEU	3.7
1	B	254	GLY	3.4
1	A	268	PHE	3.3
1	B	252	ASP	3.3
1	A	257	ASN	3.2
1	B	95	LYS	3.1
1	B	208	VAL	3.0
1	A	62	PRO	2.9
1	B	362	ARG	2.8
1	B	255	GLU	2.7
1	A	42	LYS	2.7
1	A	253	ARG	2.7
1	A	242	LYS	2.6
1	B	117	GLU	2.5
1	A	45	GLU	2.5
1	B	19	VAL	2.5
1	A	21	LEU	2.5
1	B	251	GLY	2.3
1	A	69	PRO	2.3
1	A	355	THR	2.3
1	A	200	PHE	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	327	LYS	2.3
1	A	19	VAL	2.1
1	A	243	LEU	2.1
1	B	22	ILE	2.1
1	A	252	ASP	2.1
1	B	220	PHE	2.0
1	A	208	VAL	2.0
1	A	379	VAL	2.0
1	A	396	ALA	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 oligosaccharides 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.