Full wwPDB X-ray Structure Validation Report (i)

Feb 10, 2024 – 08:41 PM EST

PDB ID	:	2QZV
Title	:	Draft Crystal Structure of the Vault Shell at 9 Angstroms Resolution
Authors	:	Anderson, D.H.; Kickhoefer, V.A.; Sievers, S.A.; Rome, L.H.; Eisenberg, D.
Deposited on	:	2007-08-17
Resolution	:	9.00 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 (1)) were used in the production of this report:

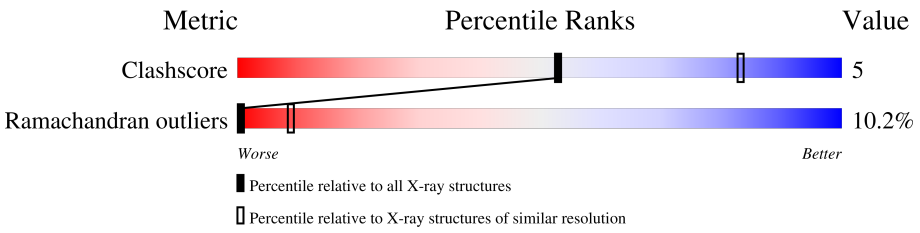
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 9.00 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))
Clashscore	141614	1070 (11.50-3.90)
Ramachandran outliers	138981	1003 (11.50-3.90)

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 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%

Mol	Chain	Length	Quality of chain
1	A	873	71% 14% • 14%
1	B	873	72% 13% • 14%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 7404 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 Major vault protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	749	3702	2204	749	749	0	0	0
1	B	749	3702	2204	749	749	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

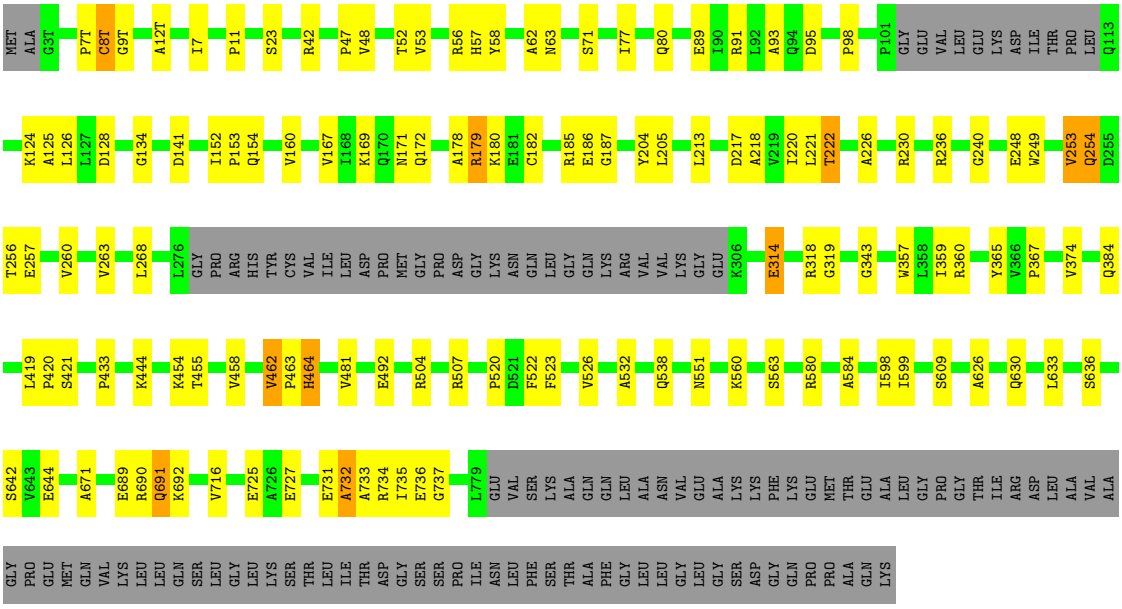
Chain	Residue	Modelled	Actual	Comment	Reference
A	1T	MET	-	expression tag	UNP Q62667
A	2T	ALA	-	expression tag	UNP Q62667
A	3T	GLY	-	expression tag	UNP Q62667
A	4T	CYS	-	expression tag	UNP Q62667
A	5T	GLY	-	expression tag	UNP Q62667
A	6T	CYS	-	expression tag	UNP Q62667
A	7T	PRO	-	expression tag	UNP Q62667
A	8T	CYS	-	expression tag	UNP Q62667
A	9T	GLY	-	expression tag	UNP Q62667
A	10T	CYS	-	expression tag	UNP Q62667
A	11T	GLY	-	expression tag	UNP Q62667
A	12T	ALA	-	expression tag	UNP Q62667
B	1T	MET	-	expression tag	UNP Q62667
B	2T	ALA	-	expression tag	UNP Q62667
B	3T	GLY	-	expression tag	UNP Q62667
B	4T	CYS	-	expression tag	UNP Q62667
B	5T	GLY	-	expression tag	UNP Q62667
B	6T	CYS	-	expression tag	UNP Q62667
B	7T	PRO	-	expression tag	UNP Q62667
B	8T	CYS	-	expression tag	UNP Q62667
B	9T	GLY	-	expression tag	UNP Q62667
B	10T	CYS	-	expression tag	UNP Q62667
B	11T	GLY	-	expression tag	UNP Q62667
B	12T	ALA	-	expression tag	UNP Q62667

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. 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. 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.

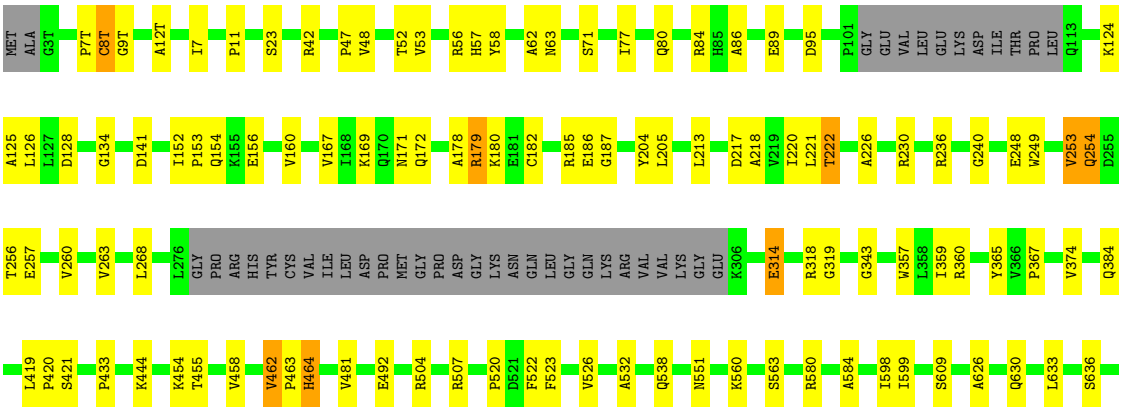
- Molecule 1: Major vault protein

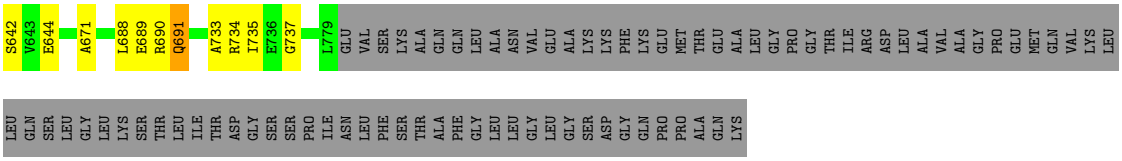
Chain A: 71% 14% • 14%



- Molecule 1: Major vault protein

Chain B: 72% 13% • 14%




S642
V643
E644
A671
L688
E689
R690
Q691
A733
R734
I735
E736
G737
L779
GLU VAL SER LYS ALA GLN GLN LEU ALA ASN VAL GLU ALA LYS LYS PHE LYS GLU MET THR GLU ALA LEU GLY PRO GLY THR ILE ARG ASP LEU ALA VAL ALA GLY PRO GLU MET GLN VAL LYS LEU
LEU GLN SER LEU GLY LEU LYS SER THR LEU ILE THR ASP GLY SER SER PRO ILE ASN LEU PHE SER THR ALA PHE GLY LEU LEU GLY LEU GLY SER ASP GLY GLN PRO PRO ALA GLN LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	631.45Å 464.72Å 584.57Å 90.00° 123.84° 90.00°	Depositor
Resolution (Å)	200.00 – 9.00 188.73 – 9.00	Depositor EDS
% Data completeness (in resolution range)	98.0 (200.00-9.00) 91.2 (188.73-9.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.22	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 8.44Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.615 , (Not available) 0.563 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	-1.2	Xtriage
Anisotropy	4.910	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.37$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.66	EDS
Total number of atoms	7404	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.46% 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	0.31	0/3699	0.70	4/5148 (0.1%)
1	B	0.31	0/3699	0.69	5/5148 (0.1%)
All	All	0.31	0/7398	0.69	9/10296 (0.1%)

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	222	THR	N-CA-C	5.51	125.87	111.00
1	B	222	THR	N-CA-C	5.50	125.86	111.00
1	B	8(T)	CYS	N-CA-C	5.29	125.27	111.00
1	A	8(T)	CYS	N-CA-C	5.28	125.25	111.00
1	A	220	ILE	C-N-CA	5.20	134.71	121.70
1	B	220	ILE	C-N-CA	5.19	134.68	121.70
1	A	9(T)	GLY	N-CA-C	-5.17	100.17	113.10
1	B	9(T)	GLY	N-CA-C	-5.17	100.18	113.10
1	B	220	ILE	CA-C-N	-5.01	106.18	117.20

There are no chirality outliers.

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	3702	0	1718	33	0
1	B	3702	0	1718	32	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7404	0	3436	58	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 (58) 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:98:PRO:CB	1:B:156:GLU:CB	2.42	0.98
1:A:692:LYS:CB	1:B:688:LEU:HA	2.32	0.58
1:A:56:ARG:O	1:A:58:TYR:N	2.36	0.58
1:B:56:ARG:O	1:B:58:TYR:N	2.36	0.58
1:A:172:GLN:HA	1:A:217:ASP:HA	1.86	0.57
1:B:172:GLN:HA	1:B:217:ASP:HA	1.86	0.57
1:B:63:ASN:CB	1:B:77:ILE:HA	2.35	0.56
1:A:63:ASN:CB	1:A:77:ILE:HA	2.35	0.56
1:A:91:ARG:O	1:B:84:ARG:HA	2.06	0.54
1:B:185:ARG:O	1:B:187:GLY:N	2.42	0.53
1:A:185:ARG:O	1:A:187:GLY:N	2.42	0.52
1:A:626:ALA:O	1:A:636:SER:HA	2.10	0.52
1:A:630:GLN:N	1:A:633:LEU:O	2.42	0.51
1:B:630:GLN:N	1:B:633:LEU:O	2.42	0.51
1:A:167:VAL:H	1:A:204:TYR:HA	1.76	0.51
1:A:731:GLU:O	1:A:732:ALA:HB2	2.11	0.50
1:B:626:ALA:O	1:B:636:SER:HA	2.10	0.50
1:B:167:VAL:H	1:B:204:TYR:HA	1.76	0.50
1:B:253:VAL:O	1:B:254:GLN:CB	2.59	0.50
1:B:462:VAL:O	1:B:464:HIS:N	2.45	0.50
1:A:462:VAL:O	1:A:464:HIS:N	2.45	0.49
1:A:253:VAL:O	1:A:254:GLN:CB	2.59	0.49
1:A:230:ARG:HA	1:A:248:GLU:CB	2.43	0.48
1:B:230:ARG:HA	1:B:248:GLU:CB	2.43	0.48
1:B:124:LYS:HA	1:B:141:ASP:O	2.15	0.47
1:B:419:LEU:O	1:B:421:SER:N	2.48	0.47
1:B:52:THR:O	1:B:62:ALA:HA	2.15	0.47
1:A:236:ARG:HA	1:A:240:GLY:O	2.15	0.46
1:B:580:ARG:HA	1:B:584:ALA:HB3	1.97	0.46
1:A:734:ARG:CB	1:B:733:ALA:HA	2.46	0.46
1:A:580:ARG:HA	1:A:584:ALA:HB3	1.97	0.46
1:A:52:THR:O	1:A:62:ALA:HA	2.15	0.46
1:A:124:LYS:HA	1:A:141:ASP:O	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:419:LEU:O	1:A:421:SER:N	2.48	0.46
1:B:236:ARG:HA	1:B:240:GLY:O	2.15	0.46
1:A:314:GLU:HA	1:A:319:GLY:O	2.16	0.46
1:B:152:ILE:O	1:B:154:GLN:N	2.49	0.45
1:B:12(T):ALA:HA	1:B:7:ILE:O	2.17	0.45
1:A:171:ASN:O	1:A:218:ALA:N	2.50	0.45
1:B:314:GLU:HA	1:B:319:GLY:O	2.16	0.45
1:B:690:ARG:O	1:B:691:GLN:C	2.55	0.45
1:A:12(T):ALA:HA	1:A:7:ILE:O	2.17	0.45
1:B:171:ASN:O	1:B:218:ALA:N	2.50	0.45
1:B:532:ALA:CB	1:B:551:ASN:HA	2.47	0.45
1:A:178:ALA:O	1:A:179:ARG:CB	2.65	0.45
1:B:178:ALA:O	1:B:179:ARG:CB	2.65	0.45
1:A:152:ILE:O	1:A:154:GLN:N	2.49	0.44
1:A:690:ARG:O	1:A:691:GLN:C	2.55	0.44
1:A:532:ALA:CB	1:A:551:ASN:HA	2.47	0.44
1:B:222:THR:O	1:B:257:GLU:N	2.51	0.43
1:A:98:PRO:CB	1:B:156:GLU:CA	2.96	0.43
1:A:222:THR:O	1:A:257:GLU:N	2.51	0.43
1:A:256:THR:O	1:A:257:GLU:C	2.57	0.43
1:B:256:THR:O	1:B:257:GLU:C	2.57	0.43
1:A:226:ALA:HB2	1:A:253:VAL:HA	2.02	0.42
1:A:737:GLY:HA3	1:B:737:GLY:O	2.20	0.41
1:B:226:ALA:HB2	1:B:253:VAL:HA	2.02	0.41
1:A:93:ALA:HB3	1:B:86:ALA:HA	2.03	0.41

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	743/873 (85%)	524 (70%)	141 (19%)	78 (10%)	0 8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	743/873 (85%)	528 (71%)	142 (19%)	73 (10%)	0 10
All	All	1486/1746 (85%)	1052 (71%)	283 (19%)	151 (10%)	0 9

All (151) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	8(T)	CYS
1	A	57	HIS
1	A	71	SER
1	A	160	VAL
1	A	179	ARG
1	A	186	GLU
1	A	205	LEU
1	A	213	LEU
1	A	221	LEU
1	A	260	VAL
1	A	268	LEU
1	A	314	GLU
1	A	360	ARG
1	A	454	LYS
1	A	492	GLU
1	A	507	ARG
1	A	563	SER
1	A	598	ILE
1	A	599	ILE
1	A	671	ALA
1	A	689	GLU
1	A	727	GLU
1	A	732	ALA
1	B	8(T)	CYS
1	B	57	HIS
1	B	71	SER
1	B	160	VAL
1	B	179	ARG
1	B	186	GLU
1	B	205	LEU
1	B	213	LEU
1	B	221	LEU
1	B	260	VAL
1	B	268	LEU
1	B	314	GLU
1	B	360	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	454	LYS
1	B	492	GLU
1	B	507	ARG
1	B	563	SER
1	B	598	ILE
1	B	599	ILE
1	B	671	ALA
1	B	689	GLU
1	B	735	ILE
1	A	80	GLN
1	A	126	LEU
1	A	169	LYS
1	A	253	VAL
1	A	254	GLN
1	A	367	PRO
1	A	560	LYS
1	A	609	SER
1	A	691	GLN
1	A	725	GLU
1	B	80	GLN
1	B	126	LEU
1	B	169	LYS
1	B	253	VAL
1	B	254	GLN
1	B	367	PRO
1	B	560	LYS
1	B	609	SER
1	B	691	GLN
1	A	11	PRO
1	A	42	ARG
1	A	48	VAL
1	A	125	ALA
1	A	153	PRO
1	A	180	LYS
1	A	182	CYS
1	A	365	TYR
1	A	384	GLN
1	A	433	PRO
1	A	463	PRO
1	A	464	HIS
1	A	520	PRO
1	A	523	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	526	VAL
1	A	538	GLN
1	A	642	SER
1	A	644	GLU
1	B	11	PRO
1	B	42	ARG
1	B	48	VAL
1	B	125	ALA
1	B	153	PRO
1	B	180	LYS
1	B	182	CYS
1	B	365	TYR
1	B	384	GLN
1	B	433	PRO
1	B	463	PRO
1	B	464	HIS
1	B	520	PRO
1	B	523	PHE
1	B	526	VAL
1	B	538	GLN
1	B	642	SER
1	B	644	GLU
1	B	734	ARG
1	A	23	SER
1	A	89	GLU
1	A	128	ASP
1	A	318	ARG
1	A	359	ILE
1	A	522	PHE
1	A	733	ALA
1	B	23	SER
1	B	89	GLU
1	B	128	ASP
1	B	359	ILE
1	B	522	PHE
1	A	134	GLY
1	A	249	TRP
1	A	357	TRP
1	A	455	THR
1	A	504	ARG
1	A	716	VAL
1	A	736	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	134	GLY
1	B	249	TRP
1	B	318	ARG
1	B	357	TRP
1	B	455	THR
1	B	504	ARG
1	A	7(T)	PRO
1	A	95	ASP
1	A	420	PRO
1	A	462	VAL
1	A	735	ILE
1	B	7(T)	PRO
1	B	95	ASP
1	B	420	PRO
1	B	462	VAL
1	A	481	VAL
1	B	481	VAL
1	A	47	PRO
1	A	53	VAL
1	A	458	VAL
1	B	47	PRO
1	B	53	VAL
1	B	458	VAL
1	A	263	VAL
1	A	343	GLY
1	A	374	VAL
1	A	444	LYS
1	B	263	VAL
1	B	374	VAL
1	B	444	LYS
1	B	343	GLY

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.