wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:39 PM BST

PDB ID	:	1HB7
EMDB ID:	:	EMD-1013
Title	:	quasi-atomic resolution model of bacteriophage PRD1 sus1 mutant, obtained by combined cryo-EM and X-ray crystallography.
Authors	:	San Martin, C.; Burnett, R.M.; De Haas, F.; Heinkel, R.; Rutten, T.; Fuller, S.D.; Butcher, S.J.; Bamford, D.H.
Deposited on	:	2001-04-12
Resolution	:	14.00 Å(reported)
Based on PDB ID	:	1HX6

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

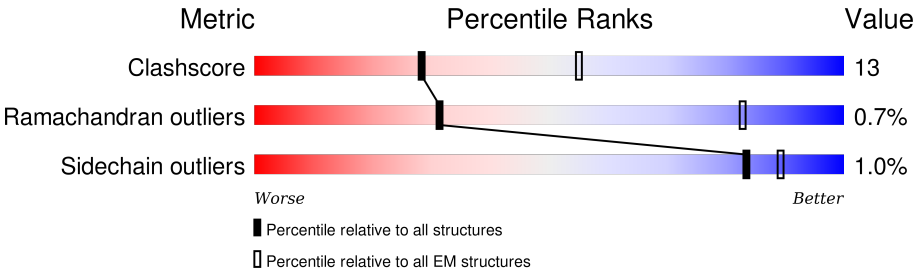
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 14.00 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	394	80% 13% • 6%
1	B	394	79% 14% • 5%
1	C	394	82% 12% • 6%
1	D	394	76% 17% • 6%
1	E	394	80% 14% • 5%
1	F	394	83% 10% • 6%
1	G	394	80% 13% • 6%
1	H	394	84% 10% • 5%
1	I	394	78% 15% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	394	
1	K	394	
1	L	394	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 34180 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID.

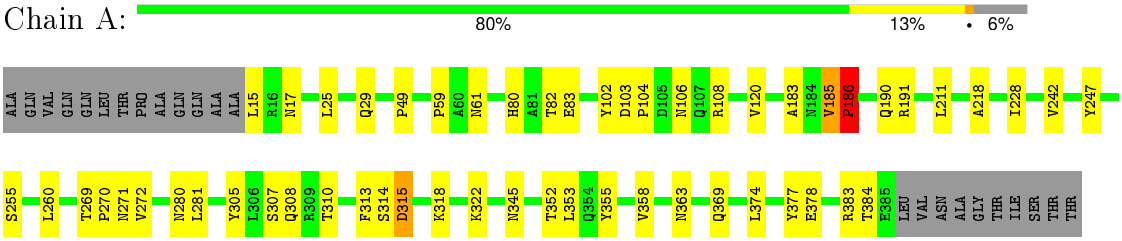
Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	371	Total	C	N	O	S	0	1
			2837	1803	479	548	7		
1	B	373	Total	C	N	O	S	0	1
			2847	1804	485	551	7		
1	C	372	Total	C	N	O	S	0	1
			2861	1817	484	553	7		
1	D	371	Total	C	N	O	S	0	1
			2837	1803	479	548	7		
1	E	373	Total	C	N	O	S	0	1
			2847	1804	485	551	7		
1	F	372	Total	C	N	O	S	0	1
			2861	1817	484	553	7		
1	G	371	Total	C	N	O	S	0	1
			2837	1803	479	548	7		
1	H	373	Total	C	N	O	S	0	1
			2847	1804	485	551	7		
1	I	372	Total	C	N	O	S	0	1
			2861	1817	484	553	7		
1	J	371	Total	C	N	O	S	0	1
			2837	1803	479	548	7		
1	K	373	Total	C	N	O	S	0	1
			2847	1804	485	551	7		
1	L	372	Total	C	N	O	S	0	1
			2861	1817	484	553	7		

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 errors 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.

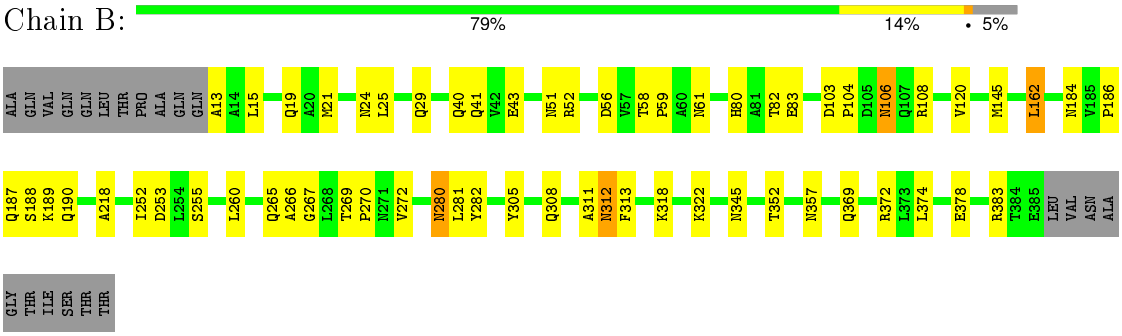
- Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain A:



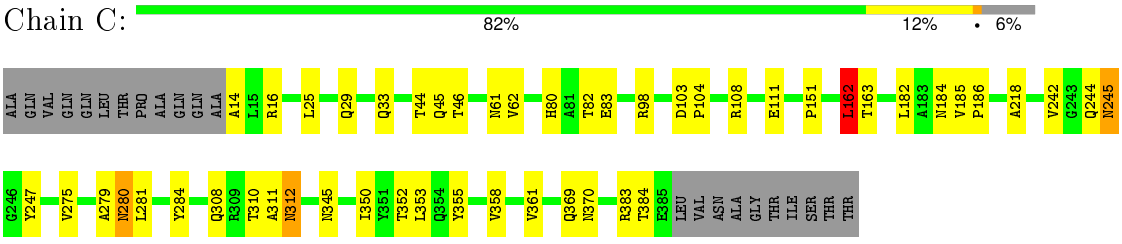
- Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain B:



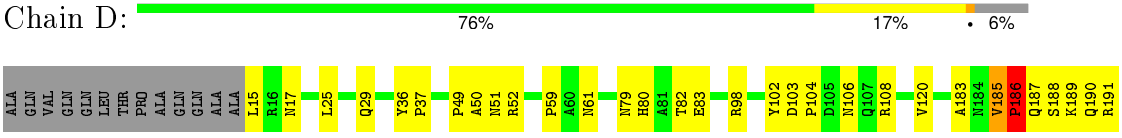
- Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

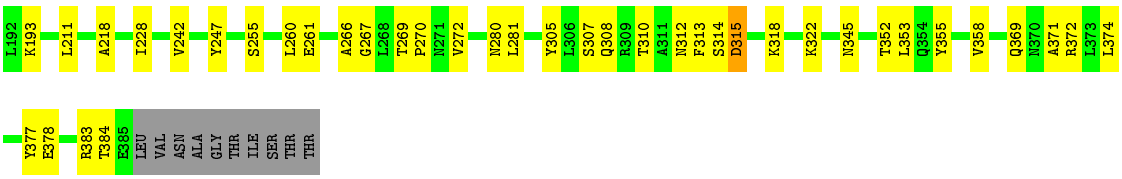
Chain C:



- Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

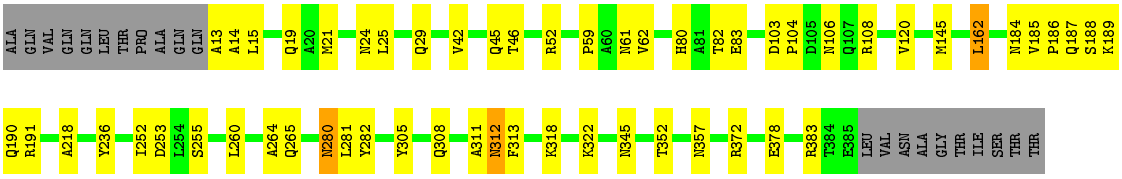
Chain D:





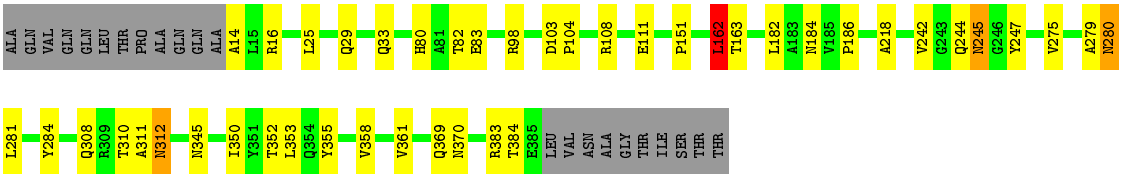
• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain E:



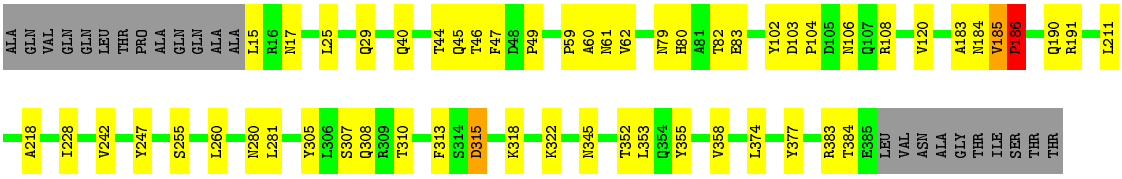
• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain F:



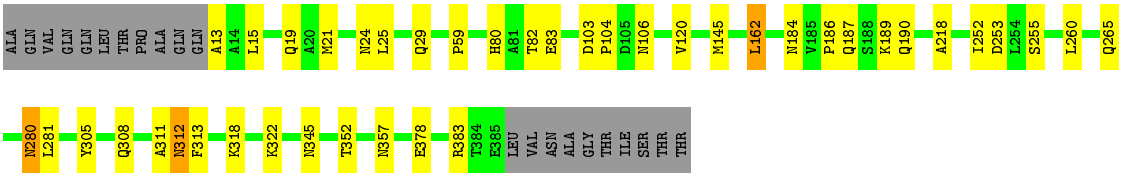
• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain G:



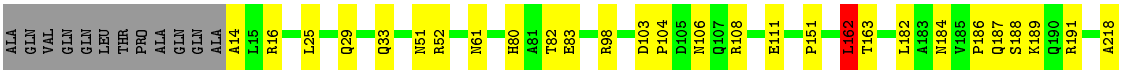
• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

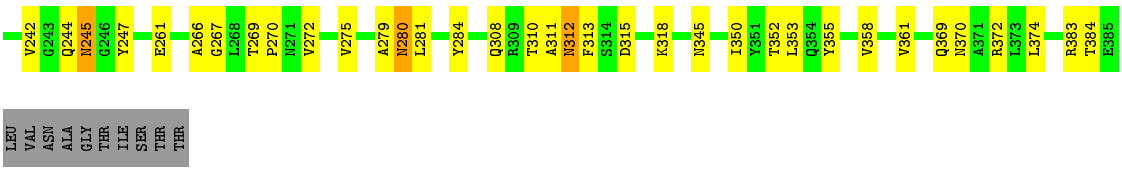
Chain H:



• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

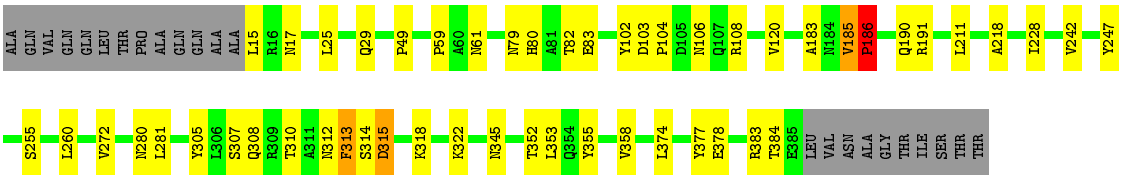
Chain I:





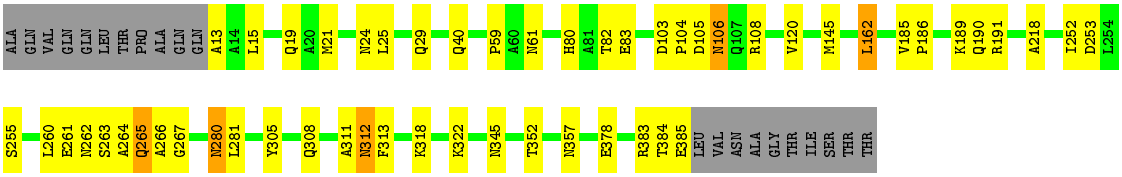
• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain J:



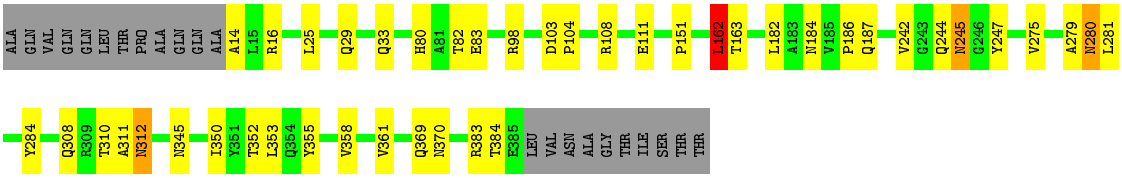
• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain K:



• Molecule 1: BACTERIOPHAGE PRD1 SUS1 MUTANT CAPSID

Chain L:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	PHILIPS CM200 FEG	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	1000	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	4100	Depositor
Magnification	36000	Depositor
Image detector	KODAK SO-163 FILM	Depositor

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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.44	0/2902	0.72	0/3971
1	B	0.46	0/2910	0.73	1/3982 (0.0%)
1	C	0.43	0/2926	0.72	1/4003 (0.0%)
1	D	0.44	0/2902	0.72	0/3971
1	E	0.45	0/2910	0.73	1/3982 (0.0%)
1	F	0.43	0/2926	0.72	1/4003 (0.0%)
1	G	0.44	0/2902	0.72	0/3971
1	H	0.45	0/2910	0.73	1/3982 (0.0%)
1	I	0.43	0/2926	0.72	1/4003 (0.0%)
1	J	0.44	0/2902	0.72	0/3971
1	K	0.45	0/2910	0.73	1/3982 (0.0%)
1	L	0.43	0/2926	0.72	1/4003 (0.0%)
All	All	0.44	0/34952	0.72	8/47824 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	K	162	LEU	CA-CB-CG	6.45	130.14	115.30
1	B	162	LEU	CA-CB-CG	6.43	130.10	115.30
1	E	162	LEU	CA-CB-CG	6.42	130.08	115.30
1	H	162	LEU	CA-CB-CG	6.42	130.08	115.30
1	I	162	LEU	CA-CB-CG	5.96	129.00	115.30

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	2837	0	2709	80	0
1	B	2847	0	2723	202	0
1	C	2861	0	2751	92	0
1	D	2837	0	2705	163	0
1	E	2847	0	2723	165	0
1	F	2861	0	2752	35	0
1	G	2837	0	2710	82	0
1	H	2847	0	2729	31	0
1	I	2861	0	2742	203	0
1	J	2837	0	2710	68	0
1	K	2847	0	2729	128	0
1	L	2861	0	2752	33	0
All	All	34180	0	32735	837	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 13.

The worst 5 of 837 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:269:THR:CG2	1:K:263:SER:HB2	1.19	1.65
1:B:189:LYS:CG	1:E:61:ASN:HB2	1.19	1.62
1:A:269:THR:HG22	1:K:263:SER:CB	1.16	1.61
1:B:372:ARG:CD	1:D:269:THR:HG21	1.21	1.57
1:B:372:ARG:HD3	1:D:269:THR:CG2	1.12	1.56

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 PDB entries followed by that with respect to all EM entries.

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	369/394 (94%)	349 (95%)	16 (4%)	4 (1%)	17 63
1	B	371/394 (94%)	349 (94%)	21 (6%)	1 (0%)	46 83
1	C	370/394 (94%)	348 (94%)	19 (5%)	3 (1%)	24 69
1	D	369/394 (94%)	349 (95%)	16 (4%)	4 (1%)	17 63
1	E	371/394 (94%)	349 (94%)	21 (6%)	1 (0%)	46 83
1	F	370/394 (94%)	348 (94%)	19 (5%)	3 (1%)	24 69
1	G	369/394 (94%)	349 (95%)	16 (4%)	4 (1%)	17 63
1	H	371/394 (94%)	349 (94%)	21 (6%)	1 (0%)	46 83
1	I	370/394 (94%)	348 (94%)	19 (5%)	3 (1%)	24 69
1	J	369/394 (94%)	349 (95%)	16 (4%)	4 (1%)	17 63
1	K	371/394 (94%)	349 (94%)	21 (6%)	1 (0%)	46 83
1	L	370/394 (94%)	348 (94%)	19 (5%)	3 (1%)	24 69
All	All	4440/4728 (94%)	4184 (94%)	224 (5%)	32 (1%)	31 71

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	185	VAL
1	C	245	ASN
1	C	280	ASN
1	C	312	ASN
1	D	185	VAL

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 PDB entries followed by that with respect to all EM entries.

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	292/325 (90%)	290 (99%)	2 (1%)	88 94
1	B	294/325 (90%)	289 (98%)	5 (2%)	68 87
1	C	298/325 (92%)	296 (99%)	2 (1%)	88 94
1	D	292/325 (90%)	290 (99%)	2 (1%)	88 94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	E	294/325 (90%)	289 (98%)	5 (2%)	68 87
1	F	298/325 (92%)	296 (99%)	2 (1%)	88 94
1	G	292/325 (90%)	290 (99%)	2 (1%)	88 94
1	H	294/325 (90%)	289 (98%)	5 (2%)	68 87
1	I	298/325 (92%)	296 (99%)	2 (1%)	88 94
1	J	292/325 (90%)	290 (99%)	2 (1%)	88 94
1	K	294/325 (90%)	289 (98%)	5 (2%)	68 87
1	L	298/325 (92%)	296 (99%)	2 (1%)	88 94
All	All	3536/3900 (91%)	3500 (99%)	36 (1%)	83 92

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	162	LEU
1	H	106	ASN
1	K	357	ASN
1	G	186	PRO
1	H	162	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 107 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	271	ASN
1	G	200	ASN
1	K	200	ASN
1	F	345	ASN
1	G	61	ASN

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

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.