



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:53 PM BST

PDB ID : 3J7W  
EMDB ID: : EMD-6035  
Title : Capsid Expansion Mechanism Of Bacteriophage T7 Revealed By Multi-State Atomic Models Derived From Cryo-EM Reconstructions  
Authors : Guo, F.; Liu, Z.; Fang, P.A.; Zhang, Q.; Wright, E.T.; Wu, W.; Zhang, C.; Vago, F.; Ren, Y.; Jakata, J.; Chiu, W.; Serwer, P.; Jiang, W.  
Deposited on : 2014-08-12  
Resolution : 3.50 Å(reported)

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](mailto: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

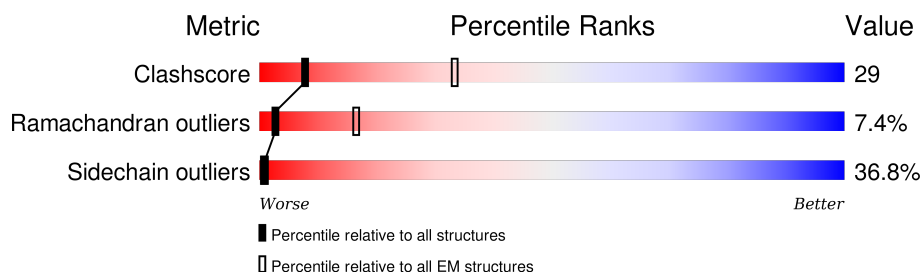
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.50 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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	345	41% 38% 18% ..
1	B	345	42% 37% 16% . .
1	C	345	42% 38% 18% ..
1	D	345	43% 40% 14% ..
1	E	345	44% 41% 11% . .
1	F	345	43% 36% 18% ..
1	G	345	44% 37% 16% ..

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 17829 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.

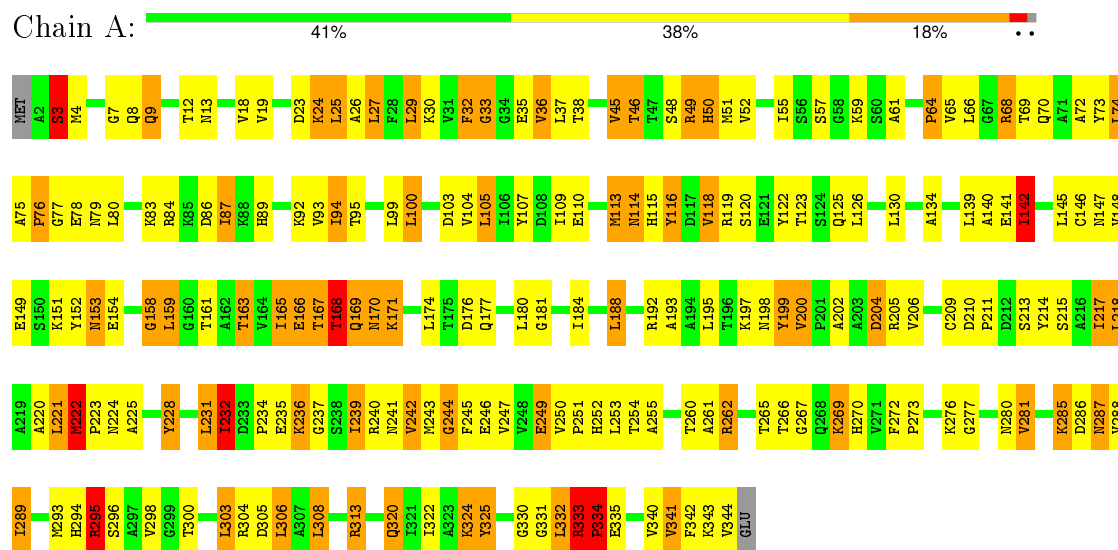
- Molecule 1 is a protein called Major capsid protein 10A.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		
1	B	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		
1	C	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		
1	D	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		
1	E	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		
1	F	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		
1	G	343	Total	C	N	O	S	0	0
			2547	1593	452	491	11		

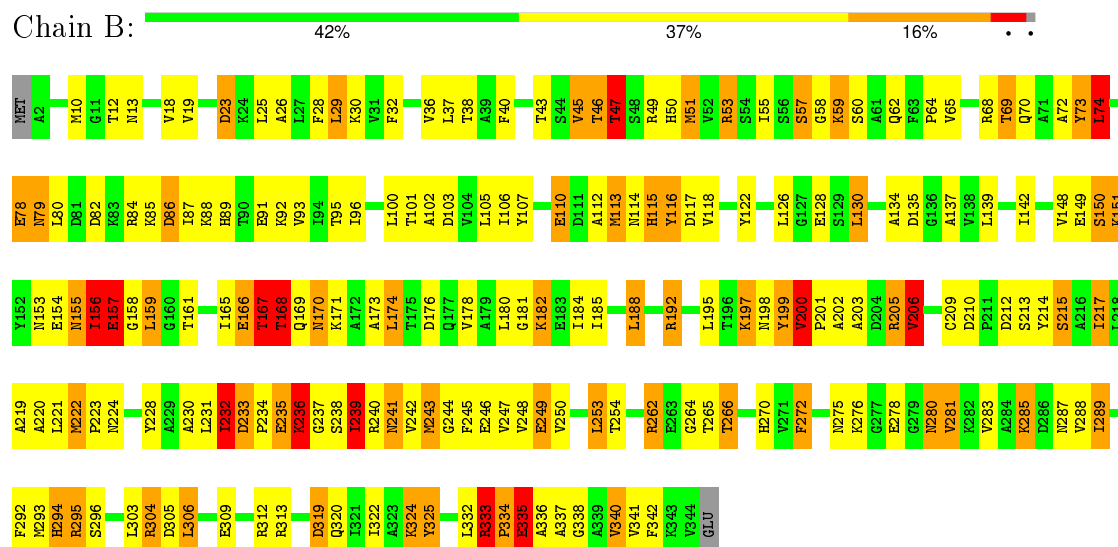
### 3 Residue-property plots

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. 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.

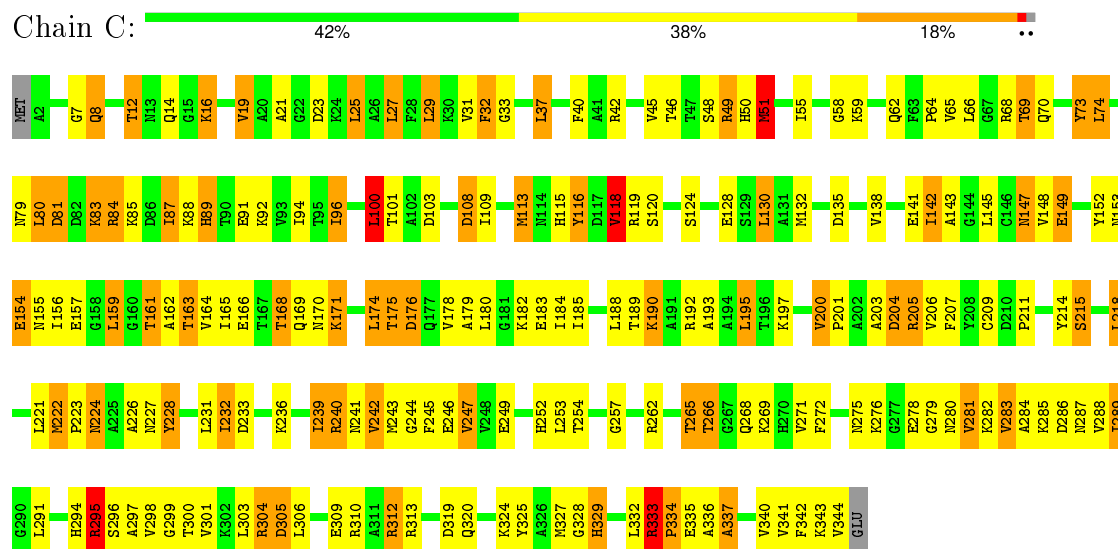
#### • Molecule 1: Major capsid protein 10A



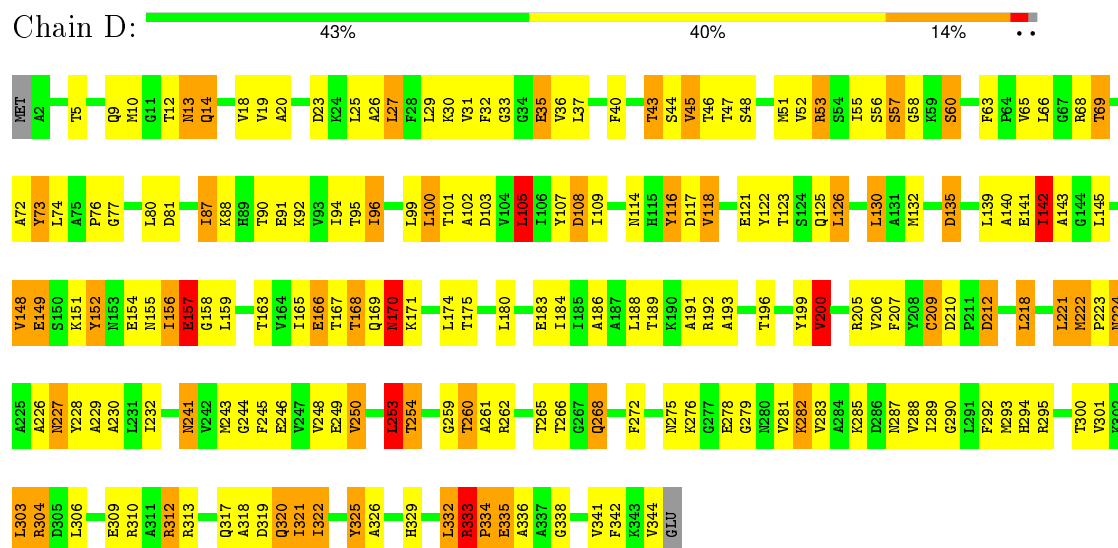
#### • Molecule 1: Major capsid protein 10A



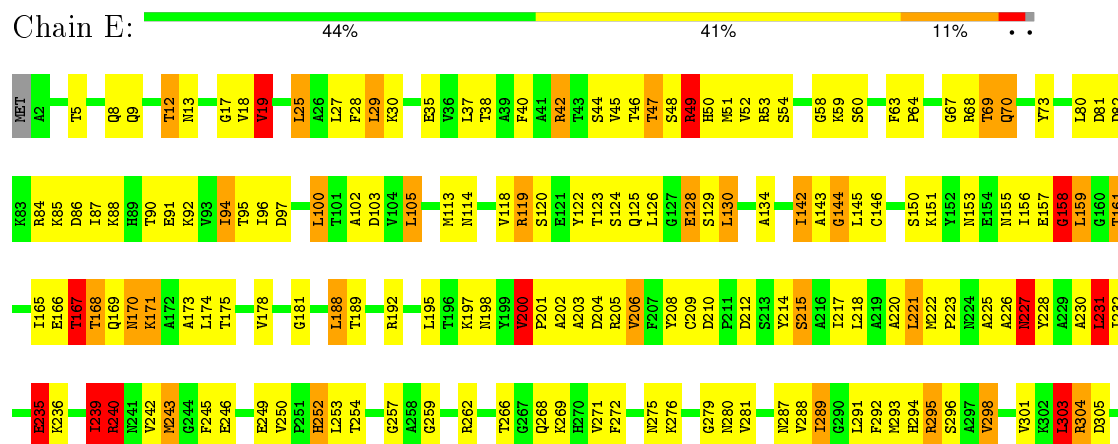
#### • Molecule 1: Major capsid protein 10A



• Molecule 1: Major capsid protein 10A



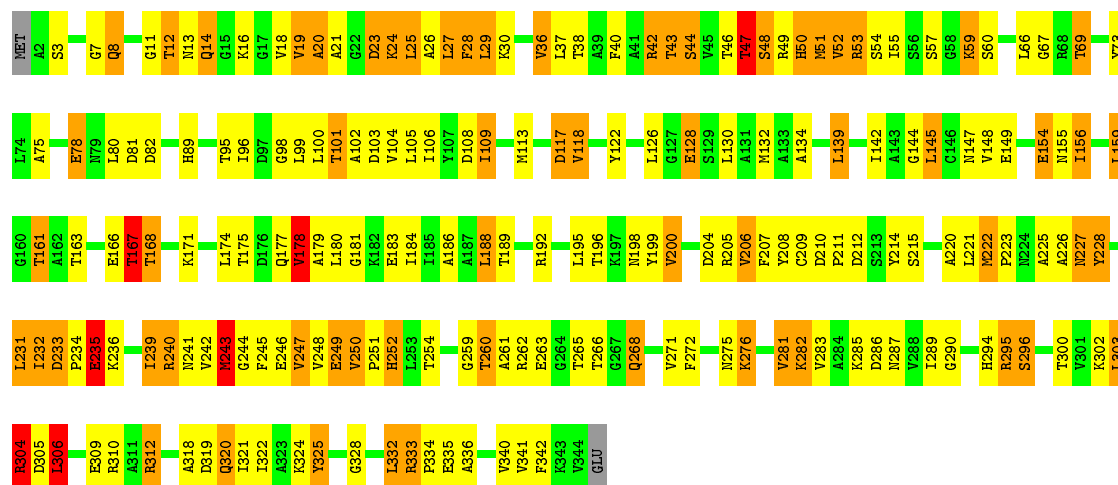
• Molecule 1: Major capsid protein 10A





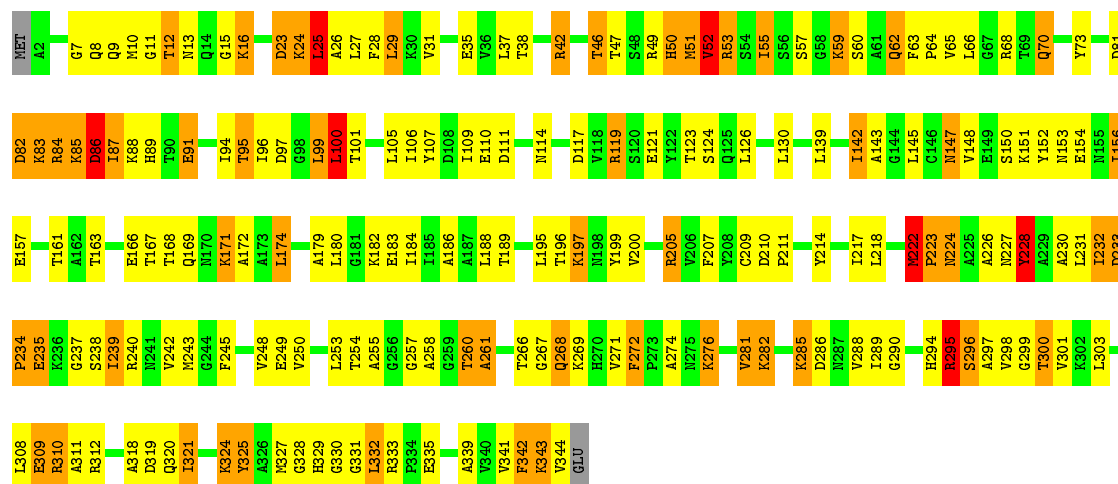
• Molecule 1: Major capsid protein 10A

Chain F: 43% 36% 18% ..



• Molecule 1: Major capsid protein 10A

Chain G: 44% 37% 16% ..



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	43417	Depositor
Resolution determination method	FSC 0.143	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	25	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3300	Depositor
Magnification	59000	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	0.33	1/2582 (0.0%)	0.69	5/3491 (0.1%)
1	B	0.34	1/2582 (0.0%)	0.66	2/3491 (0.1%)
1	C	0.33	1/2582 (0.0%)	0.66	3/3491 (0.1%)
1	D	0.34	0/2582	0.65	4/3491 (0.1%)
1	E	0.32	0/2582	0.62	3/3491 (0.1%)
1	F	0.31	0/2582	0.65	2/3491 (0.1%)
1	G	0.33	1/2582 (0.0%)	0.68	4/3491 (0.1%)
All	All	0.33	4/18074 (0.0%)	0.66	23/24437 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	334	PRO	N-CD	5.34	1.55	1.47
1	A	334	PRO	N-CD	5.30	1.55	1.47
1	B	334	PRO	N-CD	5.16	1.55	1.47
1	G	223	PRO	N-CD	5.00	1.54	1.47

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	74	LEU	CA-CB-CG	6.88	131.12	115.30
1	D	27	LEU	CA-CB-CG	6.70	130.71	115.30
1	G	333	ARG	NE-CZ-NH1	6.52	123.56	120.30
1	G	100	LEU	CA-CB-CG	6.33	129.87	115.30
1	G	222	MET	C-N-CD	5.95	140.89	128.40

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	2547	0	2571	189	0
1	B	2547	0	2573	180	0
1	C	2547	0	2573	177	0
1	D	2547	0	2573	134	0
1	E	2547	0	2573	124	0
1	F	2547	0	2573	171	0
1	G	2547	0	2573	183	0
All	All	17829	0	18009	1051	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 29.

The worst 5 of 1051 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:159:LEU:CB	1:A:335:GLU:HG2	1.27	1.63
1:B:74:LEU:HD22	1:B:80:LEU:CD2	1.35	1.53
1:C:159:LEU:CD1	1:C:335:GLU:HA	1.41	1.48
1:A:159:LEU:HB2	1:A:335:GLU:CD	1.33	1.47
1:A:159:LEU:CB	1:A:335:GLU:CG	1.94	1.44

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	341/345 (99%)	262 (77%)	51 (15%)	28 (8%)	1	13
1	B	341/345 (99%)	275 (81%)	40 (12%)	26 (8%)	1	15
1	C	341/345 (99%)	267 (78%)	48 (14%)	26 (8%)	1	15
1	D	341/345 (99%)	271 (80%)	46 (14%)	24 (7%)	1	18
1	E	341/345 (99%)	266 (78%)	50 (15%)	25 (7%)	1	16
1	F	341/345 (99%)	280 (82%)	38 (11%)	23 (7%)	1	19
1	G	341/345 (99%)	281 (82%)	36 (11%)	24 (7%)	1	18
All	All	2387/2415 (99%)	1902 (80%)	309 (13%)	176 (7%)	3	16

5 of 176 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	GLN
1	A	19	VAL
1	A	76	PRO
1	A	222	MET
1	A	236	LYS

### 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 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	259/261 (99%)	160 (62%)	99 (38%)	0	1
1	B	259/261 (99%)	150 (58%)	109 (42%)	0	1
1	C	259/261 (99%)	171 (66%)	88 (34%)	0	2
1	D	259/261 (99%)	167 (64%)	92 (36%)	0	1
1	E	259/261 (99%)	166 (64%)	93 (36%)	0	1
1	F	259/261 (99%)	165 (64%)	94 (36%)	0	1
1	G	259/261 (99%)	166 (64%)	93 (36%)	0	1
All	All	1813/1827 (99%)	1145 (63%)	668 (37%)	1	1

5 of 668 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	19	VAL
1	D	295	ARG
1	G	121	GLU
1	D	53	ARG
1	D	156	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	14	GLN
1	D	315	ASN
1	G	169	GLN
1	D	50	HIS
1	D	153	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 [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 [i](#)

There are no chain breaks in this entry.