



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 14, 2020 – 10:45 AM BST

PDB ID : 6UUJ
Title : Structure of PE5-PPE4-EspG3 complex from the type VII (ESX-3) secretion system, space group P212121
Authors : Williamson, Z.A.; Korotkov, K.V.
Deposited on : 2019-10-30
Resolution : 3.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.14.4.dev1
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)	:	2.14.4.dev1

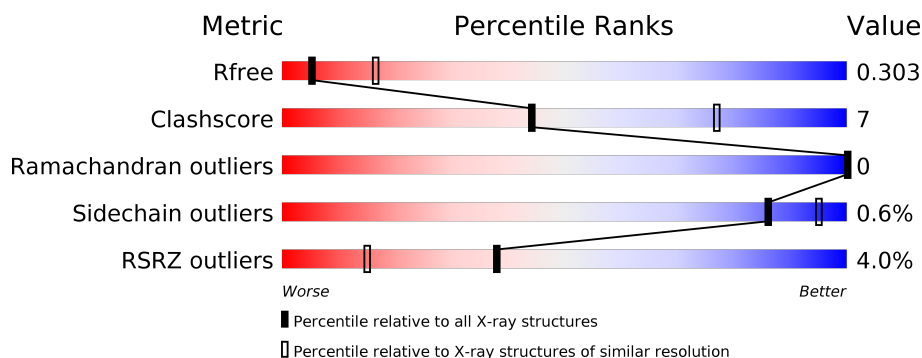
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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 $\leq 5\%$. 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	100	<div> <div></div> <div>67% 6% 27%</div> </div>
1	D	100	<div> <div>%</div> <div>64% 7% 29%</div> </div>
1	G	100	<div> <div>3%</div> <div>66% 8% 26%</div> </div>
1	J	100	<div> <div></div> <div>66% 5% 29%</div> </div>
2	B	178	<div> <div>%</div> <div>81% 11% 8%</div> </div>
2	E	178	<div> <div></div> <div>84% 7% 8%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	178	<div><div>%</div><div><div></div><div>83%</div><div>8%</div><div>9%</div></div></div>
2	K	178	<div><div>%</div><div><div></div><div>84%</div><div>7%</div><div>8%</div></div></div>
3	C	288	<div><div>4%</div><div><div></div><div>72%</div><div>18%</div><div>10%</div></div></div>
3	F	288	<div><div>5%</div><div><div></div><div>77%</div><div>13%</div><div>10%</div></div></div>
3	I	288	<div><div>7%</div><div><div></div><div>74%</div><div>15%</div><div>10%</div></div></div>
3	L	288	<div><div>9%</div><div><div></div><div>71%</div><div>19%</div><div>10%</div></div></div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 28770 atoms, of which 14231 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 PE family immunomodulator PE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	73	Total	C	H	N	O	0	0	0
			966	301	483	85	97			
1	D	71	Total	C	H	N	O	0	0	0
			944	294	472	83	95			
1	G	74	Total	C	H	N	O	0	0	0
			982	306	492	86	98			
1	J	71	Total	C	H	N	O	0	0	0
			944	294	472	83	95			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	GLY	-	expression tag	UNP L7N695
A	4	ALA	-	expression tag	UNP L7N695
A	5	MET	-	expression tag	UNP L7N695
D	3	GLY	-	expression tag	UNP L7N695
D	4	ALA	-	expression tag	UNP L7N695
D	5	MET	-	expression tag	UNP L7N695
G	3	GLY	-	expression tag	UNP L7N695
G	4	ALA	-	expression tag	UNP L7N695
G	5	MET	-	expression tag	UNP L7N695
J	3	GLY	-	expression tag	UNP L7N695
J	4	ALA	-	expression tag	UNP L7N695
J	5	MET	-	expression tag	UNP L7N695

- Molecule 2 is a protein called PPE family protein PPE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	164	Total	C	H	N	O	0	0	0
			2309	757	1134	194	220			
2	E	163	Total	C	H	N	O	0	0	0
			2299	754	1129	193	219			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	H	162	Total	C	H	N	O	S	0	0	0
			2288	751	1124	192	217	4			
2	K	163	Total	C	H	N	O	S	0	0	0
			2298	754	1129	193	218	4			

- Molecule 3 is a protein called ESX-3 secretion-associated protein EspG3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	259	Total	C	H	N	O	S	0	0	0
			3934	1264	1949	340	375	6			
3	F	259	Total	C	H	N	O	S	0	0	0
			3934	1264	1949	340	375	6			
3	I	259	Total	C	H	N	O	S	0	0	0
			3934	1264	1949	340	375	6			
3	L	259	Total	C	H	N	O	S	0	0	0
			3934	1264	1949	340	375	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	MET	-	initiating methionine	UNP B2HNX0
C	4	ALA	-	expression tag	UNP B2HNX0
F	3	MET	-	initiating methionine	UNP B2HNX0
F	4	ALA	-	expression tag	UNP B2HNX0
I	3	MET	-	initiating methionine	UNP B2HNX0
I	4	ALA	-	expression tag	UNP B2HNX0
L	3	MET	-	initiating methionine	UNP B2HNX0
L	4	ALA	-	expression tag	UNP B2HNX0

- Molecule 4 is water.

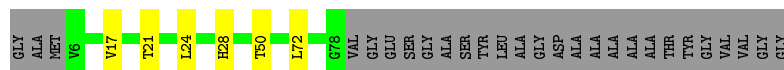
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	1	Total	O	0	0
			1	1		
4	H	1	Total	O	0	0
			1	1		
4	I	1	Total	O	0	0
			1	1		
4	J	1	Total	O	0	0
			1	1		

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.

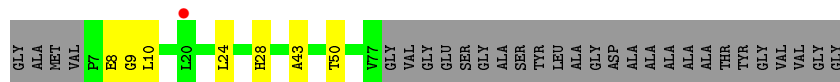
- Molecule 1: PE family immunomodulator PE5

Chain A: 



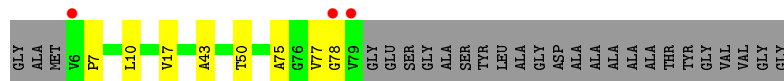
- Molecule 1: PE family immunomodulator PE5

Chain D: 



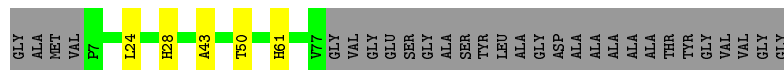
- Molecule 1: PE family immunomodulator PE5

Chain G: 




- Molecule 1: PE family immunomodulator PE5

Chain J: 




- Molecule 2: PPE family protein PPE4

Chain B: 




- Molecule 2: PPE family protein PPE4

Chain E:  84% 7% 8%




- Molecule 2: PPE family protein PPE4

Chain H:  83% 8% 9%




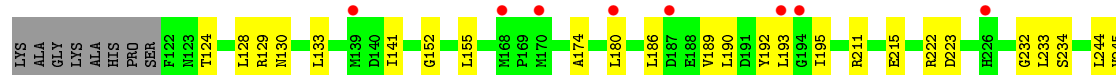
- Molecule 2: PPE family protein PPE4

Chain K:  84% 7% 8%




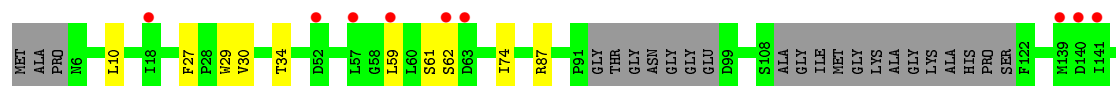
- Molecule 3: ESX-3 secretion-associated protein EspG3

Chain C:  72% 18% 10%



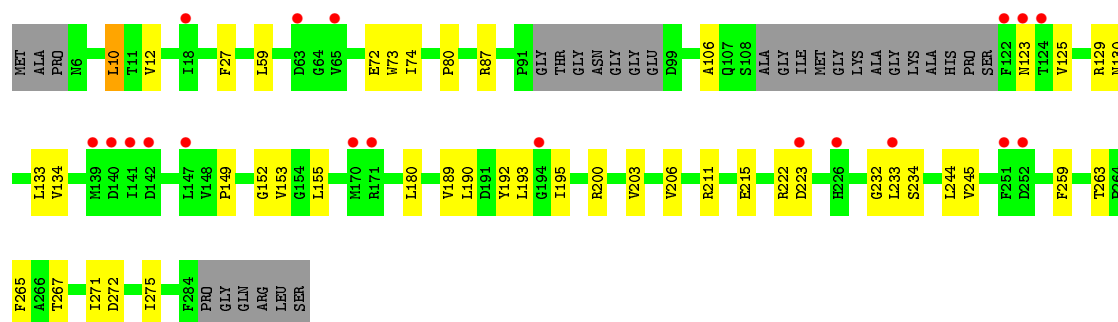
- Molecule 3: ESX-3 secretion-associated protein EspG3

Chain F:  77% 13% 10%

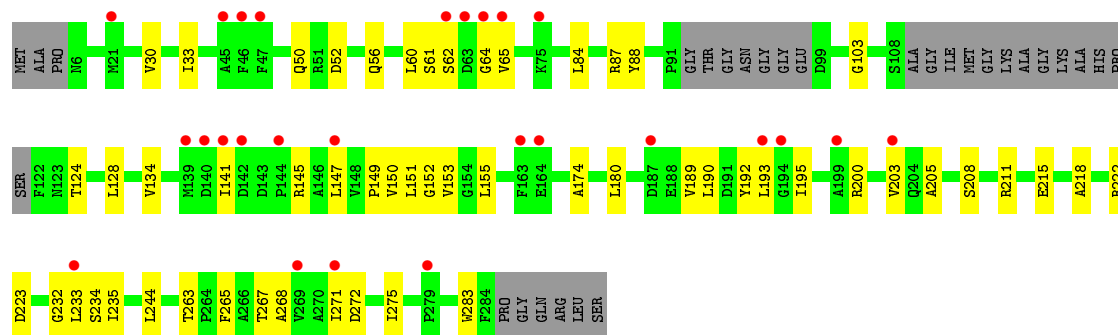


- Molecule 3: ESX-3 secretion-associated protein EspG3

Chain I:  74% 15% 10%



• Molecule 3: ESX-3 secretion-associated protein EspG3



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	72.27Å 158.63Å 209.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.52 – 3.00 39.51 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (39.52-3.00) 99.1 (39.51-3.00)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.79 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.16 _3549	Depositor
R, R_{free}	0.266 , 0.301 0.266 , 0.303	Depositor DCC
R_{free} test set	2475 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	85.8	Xtriage
Anisotropy	0.604	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 60.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	28770	wwPDB-VP
Average B, all atoms (Å ²)	111.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.75% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.24	0/490	0.42	0/675
1	D	0.25	0/479	0.42	0/659
1	G	0.23	0/497	0.41	0/685
1	J	0.24	0/479	0.41	0/659
2	B	0.23	0/1209	0.36	0/1670
2	E	0.23	0/1204	0.35	0/1663
2	H	0.24	0/1198	0.36	0/1655
2	K	0.23	0/1203	0.36	0/1662
3	C	0.24	0/2029	0.45	0/2772
3	F	0.25	0/2029	0.45	0/2772
3	I	0.24	0/2029	0.45	0/2772
3	L	0.24	0/2029	0.46	0/2772
All	All	0.24	0/14875	0.42	0/20416

There are no bond length outliers.

There are no bond angle outliers.

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	483	483	483	6	0
1	D	472	472	472	9	0
1	G	490	492	492	9	0
1	J	472	472	472	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1175	1134	1134	19	0
2	E	1170	1129	1129	13	0
2	H	1164	1124	1124	16	0
2	K	1169	1129	1129	10	0
3	C	1985	1949	1949	39	0
3	F	1985	1949	1949	30	0
3	I	1985	1949	1949	31	0
3	L	1985	1949	1949	41	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
4	I	1	0	0	0	0
4	J	1	0	0	1	0
All	All	14539	14231	14231	190	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 7.

The worst 5 of 190 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)
3:C:62:SER:HB3	3:C:65:VAL:HG12	1.69	0.74
1:A:50:THR:HG21	2:B:161:SER:HB3	1.69	0.74
1:D:50:THR:HG21	2:E:161:SER:HB3	1.69	0.73
1:G:50:THR:HG21	2:H:161:SER:HB3	1.73	0.70
3:L:60:LEU:HD23	3:L:61:SER:O	1.91	0.70

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	71/100 (71%)	69 (97%)	2 (3%)	0	100	100
1	D	69/100 (69%)	68 (99%)	1 (1%)	0	100	100
1	G	72/100 (72%)	70 (97%)	2 (3%)	0	100	100
1	J	69/100 (69%)	68 (99%)	1 (1%)	0	100	100
2	B	160/178 (90%)	159 (99%)	1 (1%)	0	100	100
2	E	159/178 (89%)	158 (99%)	1 (1%)	0	100	100
2	H	158/178 (89%)	156 (99%)	2 (1%)	0	100	100
2	K	159/178 (89%)	157 (99%)	2 (1%)	0	100	100
3	C	253/288 (88%)	238 (94%)	15 (6%)	0	100	100
3	F	253/288 (88%)	241 (95%)	12 (5%)	0	100	100
3	I	253/288 (88%)	238 (94%)	15 (6%)	0	100	100
3	L	253/288 (88%)	238 (94%)	15 (6%)	0	100	100
All	All	1929/2264 (85%)	1860 (96%)	69 (4%)	0	100	100

There are no Ramachandran outliers to report.

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 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	44/56 (79%)	44 (100%)	0	100	100
1	D	43/56 (77%)	43 (100%)	0	100	100
1	G	45/56 (80%)	45 (100%)	0	100	100
1	J	43/56 (77%)	43 (100%)	0	100	100
2	B	107/119 (90%)	107 (100%)	0	100	100
2	E	107/119 (90%)	106 (99%)	1 (1%)	78	92
2	H	106/119 (89%)	106 (100%)	0	100	100
2	K	106/119 (89%)	105 (99%)	1 (1%)	78	92
3	C	211/228 (92%)	209 (99%)	2 (1%)	78	92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	F	211/228 (92%)	210 (100%)	1 (0%)	88	96
3	I	211/228 (92%)	209 (99%)	2 (1%)	78	92
3	L	211/228 (92%)	210 (100%)	1 (0%)	88	96
All	All	1445/1612 (90%)	1437 (99%)	8 (1%)	86	95

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	211	ARG
3	L	211	ARG
3	I	211	ARG
2	E	66	GLU
3	I	10	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	73/100 (73%)	-0.05	0 100 100	66, 93, 121, 134	0
1	D	71/100 (71%)	0.09	1 (1%) 75 49	65, 91, 110, 134	0
1	G	74/100 (74%)	0.13	3 (4%) 37 14	64, 84, 123, 139	0
1	J	71/100 (71%)	0.09	0 100 100	76, 98, 118, 126	0
2	B	164/178 (92%)	0.05	1 (0%) 89 72	59, 79, 105, 127	0
2	E	163/178 (91%)	0.00	0 100 100	59, 78, 105, 126	0
2	H	162/178 (91%)	0.02	1 (0%) 89 72	59, 77, 100, 121	0
2	K	163/178 (91%)	0.05	2 (1%) 79 54	57, 79, 108, 136	0
3	C	259/288 (89%)	0.33	11 (4%) 36 14	68, 107, 146, 162	0
3	F	259/288 (89%)	0.36	15 (5%) 23 7	72, 112, 155, 168	0
3	I	259/288 (89%)	0.43	19 (7%) 15 4	69, 112, 161, 208	0
3	L	259/288 (89%)	0.54	26 (10%) 7 2	73, 118, 168, 214	0
All	All	1977/2264 (87%)	0.24	79 (3%) 38 15	57, 95, 149, 214	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	63	ASP	9.5
3	L	194	GLY	5.6
3	F	141	ILE	4.9
3	C	194	GLY	4.9
3	L	21	MET	4.7

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 monosaccharides 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.