



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 06:04 am BST

PDB ID : 4F7U  
Title : The 6S snRNP assembly intermediate  
Authors : Grimm, C.; Pelz, J.P.  
Deposited on : 2012-05-16  
Resolution : 1.90 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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
buster-report : 1.1.7 (2018)  
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) : 2.11

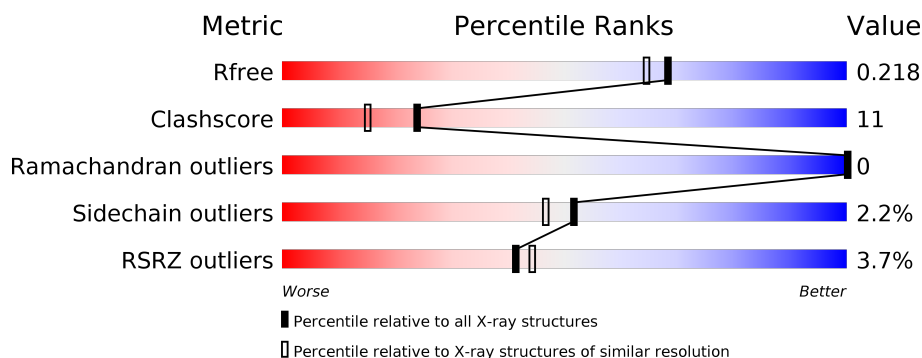
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	119	<div> <div>0%</div> <div> <div></div> <div>57%</div> <div>12%</div> <div>31%</div> </div> </div>
1	C	119	<div> <div>0%</div> <div> <div></div> <div>51%</div> <div>15%</div> <div>32%</div> </div> </div>
2	B	118	<div> <div>3%</div> <div> <div></div> <div>48%</div> <div>19%</div> <div>32%</div> </div> </div>
2	D	118	<div> <div>7%</div> <div> <div></div> <div>58%</div> <div>13%</div> <div>28%</div> </div> </div>
3	E	92	<div> <div>2%</div> <div> <div></div> <div>65%</div> <div>20%</div> <div>15%</div> </div> </div>
3	H	92	<div> <div></div> <div> <div></div> <div>70%</div> <div>15%</div> <div>15%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	86	<div><div></div><div>67%</div><div>17%</div><div>15%</div></div>
4	I	86	<div><div></div><div>66%</div><div>17%</div><div>•</div><div>14%</div></div>
5	G	76	<div><div></div><div>70%</div><div>20%</div><div>11%</div></div>
5	J	76	<div><div></div><div>74%</div><div>14%</div><div>12%</div></div>
6	P	129	<div><div>3%</div><div></div><div>77%</div><div>13%</div><div>10%</div></div>
6	Q	129	<div><div>14%</div><div></div><div>67%</div><div>25%</div><div>•</div><div>7%</div></div>

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 8542 atoms, of which 17 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 Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	82	Total	C	N	O	S	0	1	0
			655	418	114	119	4			
1	C	81	Total	C	N	O	S	0	1	0
			647	412	112	120	3			

- Molecule 2 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	80	Total	C	N	O	S	0	1	0
			653	410	120	117	6			
2	D	85	Total	C	N	O	S	0	2	0
			706	441	131	128	6			

- Molecule 3 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	78	Total	C	N	O	S	0	2	0
			653	417	114	116	6			
3	H	78	Total	C	N	O	S	0	3	0
			658	419	115	118	6			

- Molecule 4 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	73	Total	C	N	O	S	0	0	0
			571	369	94	103	5			
4	I	74	Total	C	N	O	S	0	2	0
			590	382	96	107	5			

- Molecule 5 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	G	68	Total	C	N	O	S	0	4	0
			551	350	96	99	6			
5	J	67	Total	C	N	O	S	0	1	0
			527	333	93	95	6			

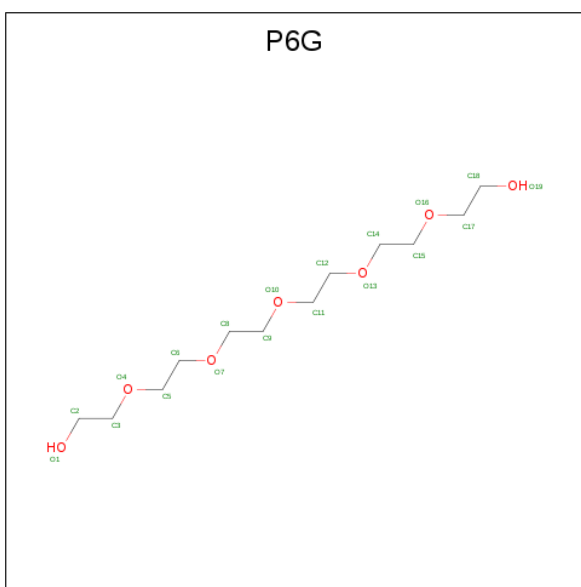
- Molecule 6 is a protein called Methylosome subunit pICln.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	Q	120	Total	C	N	O	S	0	1	0
			948	606	156	177	9			
6	P	116	Total	C	N	O	S	0	9	0
			981	627	161	184	9			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	6006	HIS	ARG	SEE REMARK 999	UNP A1ZAW5
Q	6144	ALA	HIS	ENGINEERED MUTATION	UNP A1ZAW5
Q	6161	HIS	-	EXPRESSION TAG	UNP A1ZAW5
Q	6162	HIS	-	EXPRESSION TAG	UNP A1ZAW5
Q	6163	HIS	-	EXPRESSION TAG	UNP A1ZAW5
Q	6164	HIS	-	EXPRESSION TAG	UNP A1ZAW5
Q	6165	HIS	-	EXPRESSION TAG	UNP A1ZAW5
P	6006	HIS	ARG	SEE REMARK 999	UNP A1ZAW5
P	6144	ALA	HIS	ENGINEERED MUTATION	UNP A1ZAW5
P	6161	HIS	-	EXPRESSION TAG	UNP A1ZAW5
P	6162	HIS	-	EXPRESSION TAG	UNP A1ZAW5
P	6163	HIS	-	EXPRESSION TAG	UNP A1ZAW5
P	6164	HIS	-	EXPRESSION TAG	UNP A1ZAW5
P	6165	HIS	-	EXPRESSION TAG	UNP A1ZAW5

- Molecule 7 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: C<sub>12</sub>H<sub>26</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	H	1	Total	C	H	O	0	0
			30	8	17	5		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	39	Total O 39 39	0	0
8	B	16	Total O 16 16	0	0
8	C	34	Total O 34 34	0	0
8	D	21	Total O 21 21	0	0
8	E	45	Total O 45 45	0	0
8	H	39	Total O 39 39	0	0
8	F	33	Total O 33 33	0	0
8	I	31	Total O 31 31	0	0
8	G	19	Total O 19 19	0	0
8	J	21	Total O 21 21	0	0
8	Q	26	Total O 26 26	0	0

*Continued on next page...*

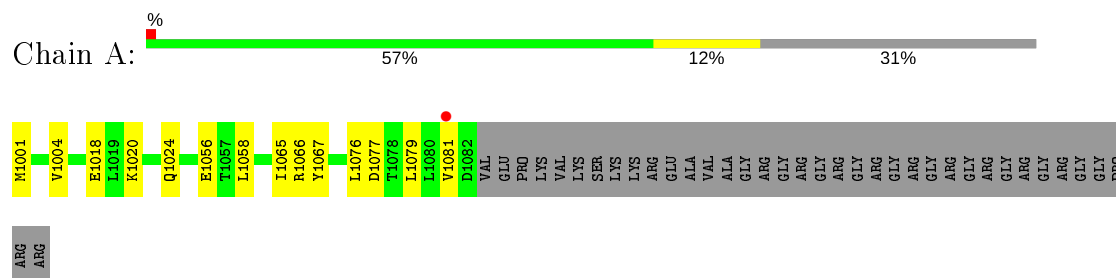
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	P	48	Total	O	0	0
			48	48		

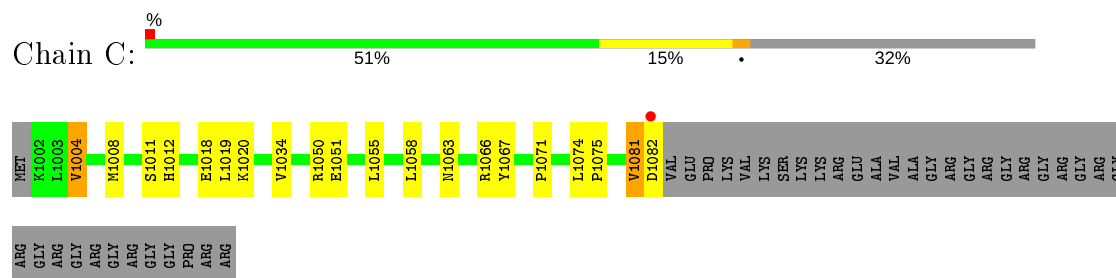
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

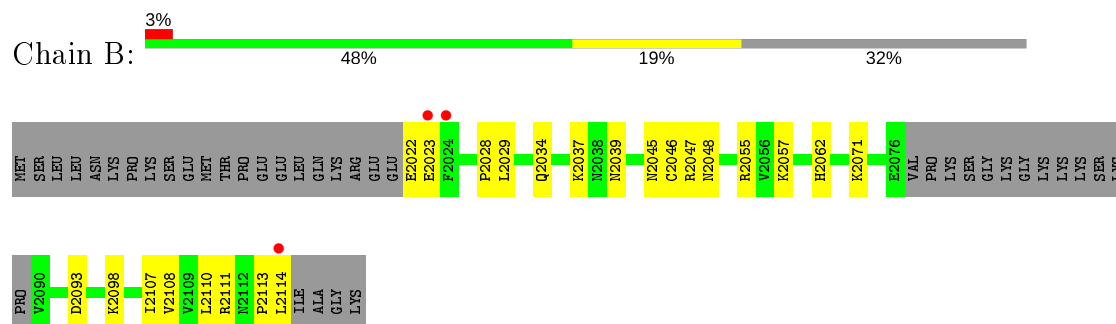
- Molecule 1: Small nuclear ribonucleoprotein Sm D1



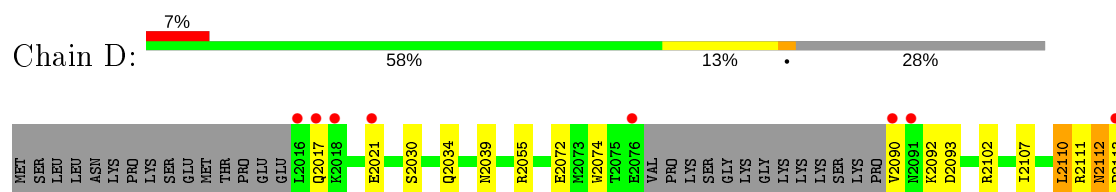
- Molecule 1: Small nuclear ribonucleoprotein Sm D1



- Molecule 2: Small nuclear ribonucleoprotein Sm D2



- Molecule 2: Small nuclear ribonucleoprotein Sm D2





LEU  
ILE  
ALA  
GLY  
LYS

• Molecule 3: Small nuclear ribonucleoprotein E



MET ALA TYR ARG GLY GLN GLY GLN LYS VAL GLN LYS V3013 T3018 T3019 T3020 T3021 T3022 T3023 R3028 R3029 R3030 R3031 T3036 T3040 E3044 E3045 C3046 E3052 D3059 T3077 T3087 V3090 SER ASN

• Molecule 3: Small nuclear ribonucleoprotein E



MET ALA TYR ARG GLY GLN GLY GLN LYS VAL GLN LYS VAL K3014 V3015 Q3016 T3020 R3023 T3031 Q3032 V3033 T3040 F3050 D3059 T3077 T3085 T3086 L3087 S3091 ASN

• Molecule 4: Small nuclear ribonucleoprotein F



MET SER LEU P4004 K4008 P4009 K4022 M4027 E4028 L4033 V4036 M4042 Q4043 L4059 C4066 R4073 G4074 E4075 E4076 GLU GLU GLU ASP GLY GLU MET ARG GLU

• Molecule 4: Small nuclear ribonucleoprotein F



MET SER L4003 N4006 P4010 L4011 K4022 M4025 G4026 M4027 V4036 D4037 E4048 L4059 V4062 R4065 C4066 N4067 V4068 L4070 R4073 E4076 GLU GLU GLU ASP GLY GLU MET ARG GLU

• Molecule 5: Small nuclear ribonucleoprotein G



MET SER LYS ALA HIS PRO P5007 E5008 L5009 M5013 D5014 K5020 L5021 N5022 H5026 R5032 G5033 F5034 D5035 P5036 F5037 M5038 N5039 I5042 I5062 N5065 E5074 ARG VAL

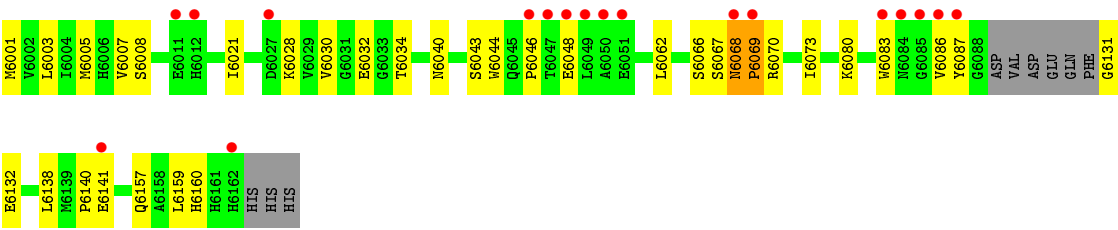
• Molecule 5: Small nuclear ribonucleoprotein G



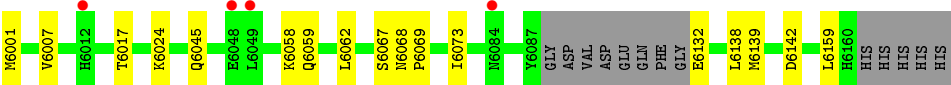
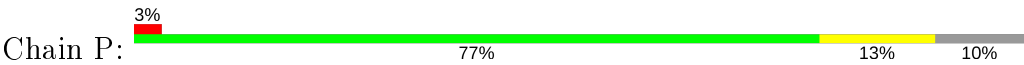
MET SER LYS ALA HIS PRO E5008 K5011 K5016 I5030 L5031 R5032 D5035 P5036 S5051 M5059 S5066 L5073 E5074 ARG VAL

• Molecule 6: Methylosome subunit pICln





● Molecule 6: Methylosome subunit pICln



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	180.72Å 65.30Å 99.35Å 90.00° 92.47° 90.00°	Depositor
Resolution (Å)	49.63 – 1.90 49.63 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.1 (49.63-1.90) 96.1 (49.63-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 1.90Å)	Xtriage
Refinement program	PHENIX phenix.refine: 1.7.1_743	Depositor
R, $R_{free}$	0.181 , 0.221 0.177 , 0.218	Depositor DCC
$R_{free}$ test set	4426 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.9	Xtriage
Anisotropy	0.420	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 53.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8542	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: P6G

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	0.64	0/666	0.76	0/899
1	C	0.64	0/658	0.71	0/890
2	B	0.63	0/663	0.74	0/890
2	D	0.62	0/719	0.71	1/963 (0.1%)
3	E	0.79	0/667	0.83	0/896
3	H	0.74	0/675	0.82	0/906
4	F	0.83	1/583 (0.2%)	0.79	0/787
4	I	0.73	0/608	0.74	0/822
5	G	0.62	0/569	0.74	0/758
5	J	0.54	0/535	0.69	0/713
6	P	0.73	0/1010	0.78	0/1373
6	Q	0.57	0/975	0.71	0/1326
All	All	0.68	1/8328 (0.0%)	0.75	1/11223 (0.0%)

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
6	Q	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	4066	CYS	CB-SG	5.80	1.92	1.82

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	D	2102	ARG	NE-CZ-NH2	-6.39	117.11	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	Q	6068	ASN	Peptide
6	Q	6069	PRO	Peptide

## 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	655	0	703	16	0
1	C	647	0	687	17	1
2	B	653	0	678	25	0
2	D	706	0	732	23	1
3	E	653	0	682	16	0
3	H	658	0	686	17	0
4	F	571	0	580	11	0
4	I	590	0	602	15	0
5	G	551	0	582	15	0
5	J	527	0	549	14	0
6	P	981	0	974	15	0
6	Q	948	0	927	30	0
7	H	13	17	17	1	0
8	A	39	0	0	0	0
8	B	16	0	0	2	0
8	C	34	0	0	0	0
8	D	21	0	0	2	0
8	E	45	0	0	0	0
8	F	33	0	0	2	0
8	G	19	0	0	2	0
8	H	39	0	0	2	0
8	I	31	0	0	2	0
8	J	21	0	0	2	0
8	P	48	0	0	1	0
8	Q	26	0	0	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	8525	17	8399	183	1

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 11.

All (183) 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)
2:D:2110:LEU:HD23	4:I:4062:VAL:HG22	1.45	0.96
2:D:2112:ASN:HB3	2:D:2113:PRO:HB3	1.51	0.92
1:C:1034:VAL:HG13	6:P:6017:THR:HG22	1.50	0.92
5:J:5008:GLU:O	5:J:5011:LYS:NZ	2.08	0.86
6:P:6132:GLU:OE1	6:P:6132:GLU:N	2.10	0.85
5:G:5039:ASN:OD1	8:G:5113:HOH:O	1.95	0.84
5:G:5035:ASP:HB2	5:G:5036:PRO:CD	2.06	0.84
1:C:1034:VAL:CG1	6:P:6017:THR:HG22	2.10	0.82
1:A:1056:GLU:OE1	6:P:6067[B]:SER:HB3	1.79	0.81
4:I:4006:ASN:OD1	8:I:4117:HOH:O	1.98	0.81
6:Q:6157:GLN:NE2	8:Q:6213:HOH:O	2.15	0.79
4:F:4022:LYS:HD2	4:F:4028:GLU:HG2	1.65	0.77
3:H:3031:ILE:HD13	3:H:3087[B]:LEU:HD11	1.65	0.77
6:P:6142:ASP:OD2	8:P:6244:HOH:O	2.01	0.77
6:Q:6068:ASN:HB2	6:Q:6069:PRO:HD2	1.67	0.76
2:D:2112:ASN:HB3	2:D:2113:PRO:CB	2.15	0.76
1:C:1058:LEU:HD11	6:Q:6062:LEU:HD13	1.68	0.75
3:E:3028:ARG:HD2	3:E:3046[B]:CYS:SG	2.27	0.74
6:Q:6068:ASN:HB2	6:Q:6069:PRO:CD	2.18	0.74
5:J:5035:ASP:HB2	5:J:5036:PRO:CD	2.19	0.73
3:H:3014:MET:CE	3:H:3016:GLN:HE22	2.00	0.73
5:G:5020[B]:LYS:HG2	5:G:5026[B]:HIS:ND1	2.04	0.72
2:D:2055:ARG:NH2	8:D:2221:HOH:O	2.08	0.71
6:Q:6068:ASN:CB	6:Q:6069:PRO:CD	2.68	0.71
3:E:3013:VAL:HG23	5:G:5034:PHE:O	1.91	0.70
1:A:1001:MET:O	1:A:1004:VAL:HG23	1.92	0.69
4:F:4004:PRO:HD3	8:F:4118:HOH:O	1.92	0.69
2:B:2022:GLU:N	2:B:2022:GLU:OE1	2.25	0.69
3:H:3023:ARG:NH2	8:H:3237:HOH:O	2.25	0.69
2:D:2110:LEU:HB3	8:I:4124:HOH:O	1.92	0.69
6:P:6024:LYS:HB2	6:P:6139:MET:CE	2.24	0.68
2:B:2045:ASN:HB2	2:B:2108:VAL:HG22	1.76	0.67
3:H:3031:ILE:HD13	3:H:3087[B]:LEU:CD1	2.24	0.67

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1018:GLU:HG3	1:A:1067:TYR:CE1	2.31	0.66
2:D:2030:SER:O	2:D:2034[A]:GLN:HG3	1.96	0.66
2:B:2022:GLU:CG	2:B:2023:GLU:H	2.09	0.66
2:B:2022:GLU:HG2	2:B:2023:GLU:H	1.62	0.64
6:P:6024:LYS:HB2	6:P:6139:MET:HE1	1.80	0.64
4:F:4075:VAL:O	4:F:4076:GLU:HB2	1.97	0.64
1:A:1001:MET:HG2	2:B:2062:HIS:ND1	2.13	0.64
3:H:3014:MET:HE2	3:H:3016:GLN:HE22	1.62	0.64
6:Q:6048:GLU:H	6:Q:6048:GLU:CD	2.02	0.63
6:Q:6067:SER:HA	6:Q:6070:ARG:O	1.97	0.63
2:D:2112:ASN:CB	2:D:2113:PRO:HB3	2.27	0.63
2:B:2022:GLU:HG2	2:B:2023:GLU:N	2.14	0.62
3:E:3018:ILE:HG22	3:E:3022:PHE:CD1	2.35	0.61
3:E:3052:GLU:N	3:E:3052:GLU:OE1	2.31	0.61
5:J:5016:LYS:NZ	8:J:5107:HOH:O	2.34	0.61
5:G:5035:ASP:HB2	5:G:5036:PRO:HD3	1.82	0.61
5:G:5035:ASP:HB2	5:G:5036:PRO:HD2	1.81	0.60
6:Q:6062:LEU:HA	8:Q:6213:HOH:O	2.01	0.60
2:B:2110:LEU:HD21	4:F:4059:LEU:HG	1.84	0.59
3:E:3077:ILE:HG22	4:F:4073:ARG:HB3	1.84	0.59
2:B:2113:PRO:O	2:B:2114:LEU:HB3	2.02	0.59
4:F:4008:LYS:NZ	8:F:4130:HOH:O	2.24	0.59
1:A:1081:VAL:O	1:A:1081:VAL:HG13	2.04	0.58
6:Q:6062:LEU:HD12	6:Q:6062:LEU:C	2.24	0.57
1:C:1018[B]:GLU:HG3	1:C:1067:TYR:CE1	2.38	0.57
1:A:1018:GLU:OE1	1:A:1024:GLN:HG2	2.04	0.57
6:Q:6048:GLU:OE1	6:Q:6048:GLU:N	2.37	0.57
2:B:2022:GLU:CG	2:B:2023:GLU:N	2.68	0.57
2:B:2113:PRO:O	2:B:2114:LEU:CB	2.52	0.57
3:H:3020:LEU:HD11	5:J:5059:MET:HG2	1.87	0.56
2:D:2055:ARG:NE	8:D:2221:HOH:O	2.35	0.56
5:J:5035:ASP:HB2	5:J:5036:PRO:HD2	1.88	0.56
3:H:3014:MET:CE	3:H:3016:GLN:NE2	2.67	0.56
3:E:3052:GLU:CD	3:E:3052:GLU:H	2.09	0.56
2:D:2110:LEU:HD22	4:I:4059:LEU:HG	1.87	0.55
6:Q:6070:ARG:NH2	6:Q:6141:GLU:O	2.40	0.55
3:E:3018:ILE:HG22	3:E:3018:ILE:O	2.05	0.55
6:Q:6083:TRP:N	6:Q:6087:TYR:O	2.29	0.55
3:H:3014:MET:HE2	3:H:3016:GLN:NE2	2.22	0.54
4:F:4008:LYS:HB3	4:F:4009:PRO:HD3	1.90	0.54
5:G:5009:LEU:HD11	5:G:5038:MET:HG3	1.90	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:J:5011:LYS:HE2	6:Q:6005:MET:CE	2.38	0.54
4:F:4075:VAL:O	4:F:4076:GLU:CB	2.51	0.54
1:C:1011:SER:O	1:C:1012:HIS:HB2	2.08	0.53
6:Q:6080:LYS:HD2	6:Q:6131:GLY:O	2.08	0.53
2:B:2047:ARG:O	2:B:2048:ASN:HB2	2.09	0.53
1:C:1050:ARG:HH11	1:C:1050:ARG:HG3	1.73	0.53
6:Q:6044:TRP:CZ3	6:Q:6046:PRO:HG3	2.44	0.53
2:D:2112:ASN:CB	2:D:2113:PRO:CA	2.87	0.52
5:J:5051:SER:HB2	8:J:5117:HOH:O	2.08	0.52
2:B:2071:LYS:HE3	2:B:2093:ASP:OD2	2.10	0.52
6:Q:6001:MET:HG3	6:Q:6003:LEU:CD1	2.38	0.52
6:Q:6068:ASN:CB	6:Q:6069:PRO:HD3	2.38	0.51
6:P:6058:LYS:NZ	6:P:6059[A]:GLN:OE1	2.41	0.51
6:P:6024:LYS:HB2	6:P:6139:MET:HE2	1.92	0.51
2:B:2114:LEU:HD23	2:B:2114:LEU:C	2.31	0.51
1:C:1034:VAL:CG1	6:P:6017:THR:CG2	2.87	0.51
2:D:2072:GLU:O	2:D:2093:ASP:HA	2.10	0.51
3:E:3018:ILE:HG22	3:E:3022:PHE:HD1	1.74	0.51
5:J:5016:LYS:HE3	5:J:5073:LEU:HD12	1.92	0.50
4:I:4036:VAL:HG12	4:I:4037:ASP:N	2.27	0.50
1:A:1077:ASP:OD1	2:B:2057:LYS:HE3	2.11	0.50
1:A:1020[A]:LYS:NZ	1:A:1065:ILE:O	2.45	0.50
5:J:5030:ILE:HG21	5:J:5032:ARG:CZ	2.42	0.50
2:D:2039:ASN:HA	2:D:2055:ARG:HD2	1.94	0.50
5:G:5020[B]:LYS:HG2	5:G:5026[B]:HIS:CE1	2.47	0.50
3:E:3013:VAL:HG22	3:E:3013:VAL:O	2.11	0.50
1:A:1066:ARG:NH1	1:A:1067:TYR:CE2	2.81	0.49
2:D:2090:VAL:HG13	2:D:2090:VAL:O	2.13	0.49
4:I:4025:TRP:HB2	4:I:4027:MET:SD	2.52	0.49
6:Q:6086:VAL:HG12	6:Q:6087:TYR:N	2.27	0.48
3:H:3050:PHE:O	7:H:3101:P6G:H32	2.13	0.48
4:I:4011:LEU:HD23	4:I:4036:VAL:HG11	1.95	0.48
2:B:2022:GLU:CD	2:B:2023:GLU:H	2.17	0.48
3:E:3031:ILE:HD12	3:E:3031:ILE:C	2.34	0.48
4:I:4011:LEU:CD2	4:I:4036:VAL:HG11	2.44	0.48
2:D:2074:TRP:CE2	2:D:2092:LYS:HB2	2.49	0.48
1:C:1055:LEU:CD2	6:Q:6066:SER:HB2	2.44	0.48
4:F:4036:VAL:HG22	4:F:4042:MET:HG2	1.96	0.48
3:H:3077:ILE:HG22	4:I:4073:ARG:HB3	1.95	0.47
5:J:5008:GLU:OE2	6:Q:6040:ASN:HB2	2.14	0.47
2:D:2112:ASN:HA	2:D:2112:ASN:HD22	1.54	0.47

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:1020:LYS:HE2	1:C:1063:ASN:O	2.15	0.47
5:J:5035:ASP:HB2	5:J:5036:PRO:HD3	1.97	0.46
6:Q:6073:ILE:HB	6:Q:6138:LEU:HB2	1.97	0.46
3:H:3040:ASN:HB2	8:H:3224:HOH:O	2.15	0.46
2:D:2112:ASN:CB	2:D:2113:PRO:CB	2.91	0.46
1:C:1018[A]:GLU:HG2	1:C:1066:ARG:HD3	1.97	0.46
3:E:3013:VAL:N	5:G:5013:MET:SD	2.89	0.46
4:I:4022:LYS:HG2	4:I:4070:LEU:HD23	1.98	0.46
5:G:5009:LEU:CD1	5:G:5038:MET:HG3	2.46	0.45
2:B:2057:LYS:NZ	8:B:2209:HOH:O	2.48	0.45
3:H:3014:MET:HE3	3:H:3016:GLN:NE2	2.31	0.45
3:E:3036:TYR:O	5:G:5022[B]:ASN:OD1	2.35	0.45
3:H:3085:THR:HG23	5:J:5066:SER:HB3	1.98	0.45
5:G:5042:ILE:HG13	5:G:5062:ILE:HD12	1.99	0.45
6:Q:6066:SER:OG	6:Q:6067:SER:N	2.50	0.44
3:H:3014:MET:HE3	3:H:3016:GLN:HE22	1.82	0.44
5:J:5011:LYS:HE2	6:Q:6005:MET:HE1	1.98	0.44
5:G:5014:ASP:OD2	5:G:5032:ARG:NH1	2.46	0.44
3:E:3019:ASN:O	3:E:3023:ARG:HB2	2.18	0.44
1:A:1077:ASP:OD1	2:B:2057:LYS:CE	2.66	0.44
2:D:2107:ILE:HA	4:I:4068:ASN:ND2	2.33	0.44
6:Q:6068:ASN:HB3	6:Q:6069:PRO:HD3	2.00	0.44
1:A:1001:MET:HG2	2:B:2062:HIS:CE1	2.53	0.43
3:E:3020:LEU:HA	3:E:3020:LEU:HD23	1.77	0.43
5:G:5035:ASP:CB	5:G:5036:PRO:CD	2.82	0.43
2:B:2034:GLN:OE1	2:B:2111:ARG:CZ	2.67	0.43
3:H:3078:MET:HB2	4:I:4010:PHE:CD2	2.53	0.43
2:D:2111:ARG:O	2:D:2112:ASN:ND2	2.51	0.43
2:D:2110:LEU:CD2	4:I:4059:LEU:HG	2.49	0.43
3:H:3033:VAL:HG22	3:H:3087[A]:LEU:CD2	2.48	0.43
6:P:6001:MET:C	6:P:6059[B]:GLN:OE1	2.58	0.42
6:Q:6030:VAL:HG13	6:Q:6086:VAL:HG13	2.01	0.42
1:A:1079:LEU:HD23	1:A:1079:LEU:HA	1.82	0.42
1:C:1004:VAL:O	1:C:1008:MET:HG3	2.18	0.42
1:C:1074:LEU:HD12	1:C:1075:PRO:HD2	2.02	0.42
6:P:6073:ILE:HB	6:P:6138:LEU:HB2	1.99	0.42
1:C:1071:PRO:HD2	1:C:1074:LEU:HD22	2.00	0.42
6:Q:6007:VAL:HG11	6:Q:6043[B]:SER:HB3	2.00	0.42
1:C:1081:VAL:O	1:C:1082:ASP:HB3	2.20	0.42
3:E:3031:ILE:HD13	3:E:3087[B]:LEU:CD1	2.50	0.42
6:Q:6021:ILE:CD1	6:Q:6140:PRO:HB3	2.50	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:2039:ASN:OD1	2:B:2055:ARG:HD3	2.19	0.42
2:B:2028:PRO:HA	4:F:4043:GLN:HB2	2.01	0.42
4:F:4033:LEU:HD11	4:F:4036:VAL:CG2	2.50	0.42
2:D:2017:GLN:O	2:D:2021:GLU:HG2	2.20	0.42
6:P:6007:VAL:HB	6:P:6045:GLN:HB2	2.01	0.42
1:A:1081:VAL:CG1	1:A:1081:VAL:O	2.67	0.42
1:A:1018:GLU:HG3	1:A:1067:TYR:CZ	2.55	0.41
1:C:1019:LEU:N	1:C:1019:LEU:HD23	2.35	0.41
3:E:3030:ARG:NE	3:E:3044:GLU:OE2	2.36	0.41
4:I:4065:ARG:HD3	4:I:4067:ASN:OD1	2.19	0.41
5:J:5035:ASP:CB	5:J:5036:PRO:CD	2.91	0.41
1:A:1058:LEU:HD11	6:P:6062:LEU:HD13	2.03	0.41
2:B:2107:ILE:HG22	2:B:2108:VAL:HG13	2.03	0.41
2:D:2110:LEU:CD2	4:I:4062:VAL:HG22	2.32	0.41
1:C:1055:LEU:CD1	1:C:1058:LEU:HD13	2.51	0.41
6:Q:6028:LYS:HD2	6:Q:6086:VAL:O	2.20	0.41
2:B:2029:LEU:HB2	8:B:2206:HOH:O	2.21	0.41
5:G:5065:ASN:N	8:G:5106:HOH:O	2.34	0.41
1:C:1011:SER:O	1:C:1012:HIS:CB	2.69	0.41
6:Q:6160:HIS:CE1	8:Q:6226:HOH:O	2.74	0.41
6:Q:6008:SER:HA	8:Q:6217:HOH:O	2.20	0.40
2:B:2034:GLN:HA	2:B:2037:LYS:HG2	2.03	0.40
2:D:2112:ASN:HB3	2:D:2113:PRO:CA	2.49	0.40
3:H:3078:MET:HB2	4:I:4010:PHE:CE2	2.57	0.40
6:P:6068[B]:ASN:HA	6:P:6069[B]:PRO:C	2.40	0.40
1:A:1076:LEU:HD11	2:B:2098:LYS:HG2	2.04	0.40
2:D:2021:GLU:OE1	2:D:2021:GLU:HA	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:1051:GLU:OE2	2:D:2055:ARG:NH1[4_556]	1.96	0.24

## 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	81/119 (68%)	79 (98%)	2 (2%)	0	100	100
1	C	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
2	B	77/118 (65%)	76 (99%)	1 (1%)	0	100	100
2	D	83/118 (70%)	82 (99%)	1 (1%)	0	100	100
3	E	78/92 (85%)	77 (99%)	1 (1%)	0	100	100
3	H	79/92 (86%)	79 (100%)	0	0	100	100
4	F	71/86 (83%)	70 (99%)	1 (1%)	0	100	100
4	I	74/86 (86%)	73 (99%)	1 (1%)	0	100	100
5	G	70/76 (92%)	69 (99%)	1 (1%)	0	100	100
5	J	66/76 (87%)	65 (98%)	1 (2%)	0	100	100
6	P	121/129 (94%)	119 (98%)	2 (2%)	0	100	100
6	Q	117/129 (91%)	110 (94%)	7 (6%)	0	100	100
All	All	997/1240 (80%)	976 (98%)	21 (2%)	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	78/101 (77%)	78 (100%)	0	100	100
1	C	77/101 (76%)	75 (97%)	2 (3%)	46	39
2	B	77/110 (70%)	75 (97%)	2 (3%)	46	39
2	D	83/110 (76%)	81 (98%)	2 (2%)	49	43
3	E	75/84 (89%)	72 (96%)	3 (4%)	31	22
3	H	76/84 (90%)	74 (97%)	2 (3%)	46	39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	F	62/74 (84%)	61 (98%)	1 (2%)	62	60
4	I	65/74 (88%)	60 (92%)	5 (8%)	13	5
5	G	63/66 (96%)	63 (100%)	0	100	100
5	J	59/66 (89%)	59 (100%)	0	100	100
6	P	112/114 (98%)	111 (99%)	1 (1%)	78	79
6	Q	106/114 (93%)	102 (96%)	4 (4%)	33	24
All	All	933/1098 (85%)	911 (98%)	22 (2%)	52	43

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	2046[A]	CYS
2	B	2046[B]	CYS
1	C	1004	VAL
1	C	1081	VAL
2	D	2110	LEU
2	D	2112	ASN
3	E	3029	SER
3	E	3040	ASN
3	E	3059	ASP
3	H	3059	ASP
3	H	3091	SER
4	F	4027	MET
4	I	4003	LEU
4	I	4027	MET
4	I	4048[A]	GLU
4	I	4048[B]	GLU
4	I	4059	LEU
6	Q	6032	GLU
6	Q	6034	THR
6	Q	6132	GLU
6	Q	6159	LEU
6	P	6159	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1024	GLN
2	D	2112	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	H	3016	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	P6G	H	3101	-	12,12,18	0.56	0	11,11,17	1.43	2 (18%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	P6G	H	3101	-	-	3/10/10/16	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	H	3101	P6G	O1-C2-C3	-2.83	95.37	111.81
7	H	3101	P6G	O10-C9-C8	-2.45	99.33	110.39

There are no chirality outliers.

All (3) torsion outliers are listed below:

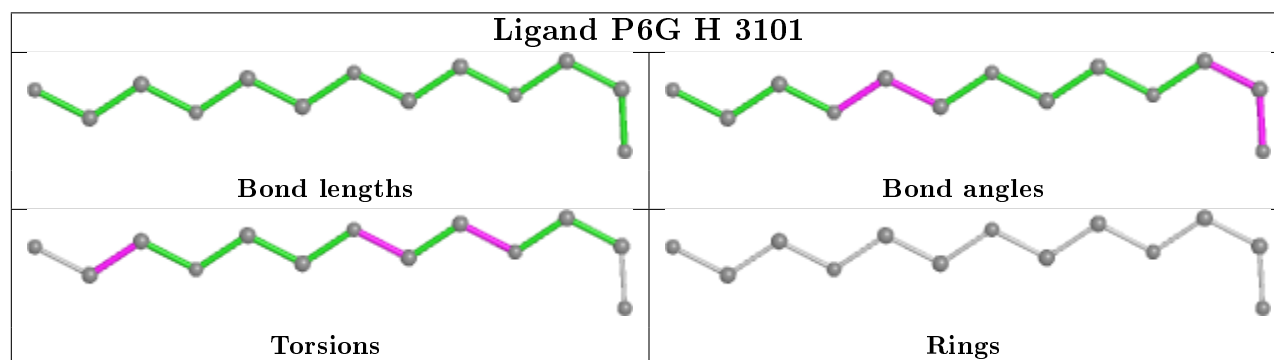
Mol	Chain	Res	Type	Atoms
7	H	3101	P6G	O10-C11-C12-O13
7	H	3101	P6G	C6-C5-O4-C3
7	H	3101	P6G	C5-C6-O7-C8

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	H	3101	P6G	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	82/119 (68%)	0.03	1 (1%) 79 81	20, 36, 59, 89	0
1	C	81/119 (68%)	0.03	1 (1%) 79 81	26, 38, 61, 106	0
2	B	80/118 (67%)	0.13	3 (3%) 40 43	23, 42, 75, 101	0
2	D	85/118 (72%)	0.41	8 (9%) 8 9	26, 41, 92, 104	0
3	E	78/92 (84%)	0.03	2 (2%) 56 58	22, 35, 65, 77	0
3	H	78/92 (84%)	-0.19	0 100 100	25, 33, 56, 69	0
4	F	73/86 (84%)	-0.23	0 100 100	23, 36, 49, 86	0
4	I	74/86 (86%)	-0.15	0 100 100	25, 38, 57, 82	0
5	G	68/76 (89%)	0.01	0 100 100	25, 39, 77, 95	0
5	J	67/76 (88%)	0.15	0 100 100	30, 46, 81, 94	0
6	P	116/129 (89%)	0.32	4 (3%) 45 48	20, 32, 70, 98	0
6	Q	120/129 (93%)	0.78	18 (15%) 2 2	27, 54, 103, 141	0
All	All	1002/1240 (80%)	0.15	37 (3%) 41 44	20, 39, 81, 141	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	Q	6049	LEU	6.9
6	Q	6011	GLU	6.6
3	E	3013	VAL	6.5
6	Q	6050	ALA	6.2
2	D	2090	VAL	6.1
2	D	2016	LEU	4.9
6	P	6084	ASN	4.9
2	D	2018	LYS	4.7
6	Q	6069	PRO	4.4
2	D	2017	GLN	4.2
2	B	2114	LEU	4.1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
6	Q	6048	GLU	3.8
6	Q	6086	VAL	3.8
6	Q	6083	TRP	3.5
6	Q	6012	HIS	3.5
6	P	6012	HIS	3.5
6	Q	6046	PRO	3.4
6	Q	6085	GLY	3.4
6	P	6048	GLU	3.3
6	Q	6087	TYR	3.3
6	Q	6068	ASN	3.2
2	B	2024	PHE	3.2
2	D	2091	ASN	3.2
2	D	2076	GLU	3.2
3	E	3018	ILE	3.1
6	Q	6141	GLU	3.0
6	P	6049	LEU	2.8
1	A	1081	VAL	2.7
6	Q	6051	GLU	2.7
6	Q	6084	ASN	2.6
1	C	1082	ASP	2.6
6	Q	6027	ASP	2.6
6	Q	6047	THR	2.6
2	D	2021	GLU	2.5
6	Q	6162	HIS	2.5
2	D	2113	PRO	2.4
2	B	2023	GLU	2.1

## 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](#)

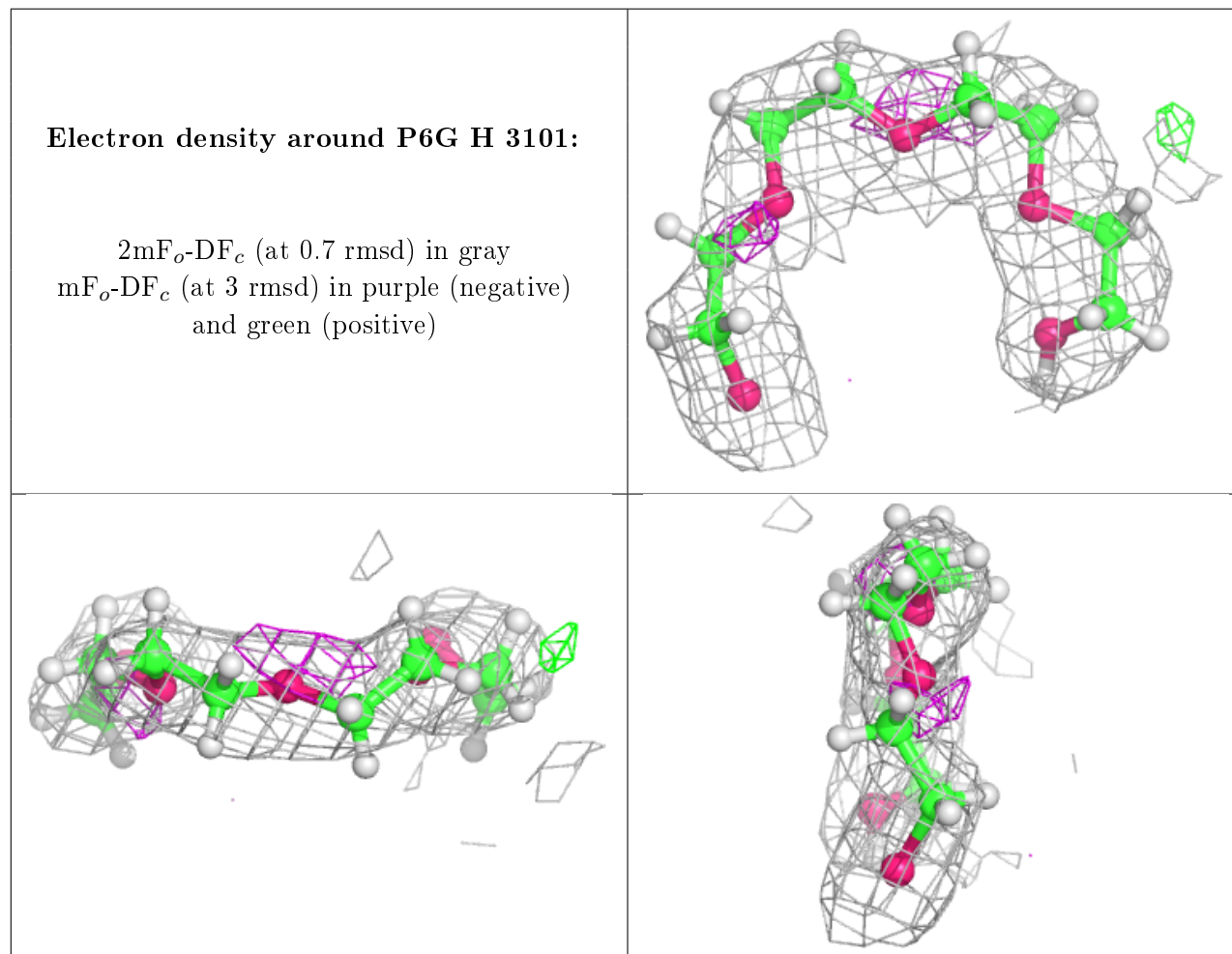
There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	P6G	H	3101	13/19	0.89	0.15	38,54,66,66	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.