



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2021 – 09:15 PM EDT

PDB ID : 3LIY  
Title : crystal structure of HTLV protease complexed with Statine-containing peptide inhibitor  
Authors : Satoh, T.; Li, M.; Nguyen, J.; Kiso, Y.; Wlodawer, A.; Gustchina, A.  
Deposited on : 2010-01-25  
Resolution : 1.86 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 : 2.23.2  
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) : 2.23.2

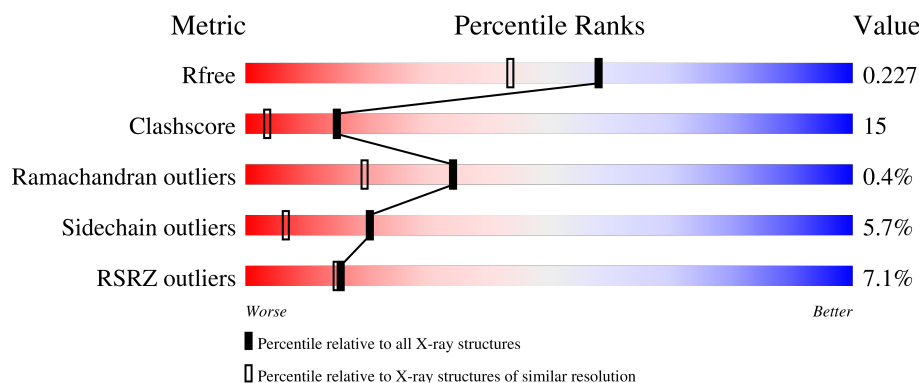
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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  $\geq 3$ , 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	116	<div> <div>3%</div> <div>87%</div> <div>11%</div> <div>.</div> </div>
1	B	116	<div> <div>3%</div> <div>80%</div> <div>17%</div> <div>.</div> </div>
1	C	116	<div> <div>9%</div> <div>79%</div> <div>20%</div> <div>.</div> </div>
1	D	116	<div> <div>15%</div> <div>78%</div> <div>19%</div> <div>.</div> </div>
1	E	116	<div> <div>8%</div> <div>88%</div> <div>9%</div> <div>.</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	116	<div><div></div><div>5%</div><div>80%</div><div>18%</div><div></div></div>
2	I	10	<div><div></div><div>60%</div><div>20%</div><div>10%</div><div>10%</div></div>
2	J	10	<div><div></div><div>10%</div><div>20%</div><div>50%</div><div>10%</div><div>20%</div></div>
2	K	10	<div><div></div><div>40%</div><div>30%</div><div>10%</div><div>20%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5970 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 Protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	0	3	0
			896	576	152	165	3			
1	B	116	Total	C	N	O	S	0	3	0
			901	576	155	167	3			
1	C	116	Total	C	N	O	S	0	0	0
			883	564	152	164	3			
1	D	116	Total	C	N	O	S	0	0	0
			883	564	152	164	3			
1	E	116	Total	C	N	O	S	0	0	0
			883	564	152	164	3			
1	F	116	Total	C	N	O	S	0	1	0
			894	574	153	164	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	ILE	LEU	engineered mutation	UNP Q82134
B	40	ILE	LEU	engineered mutation	UNP Q82134
C	40	ILE	LEU	engineered mutation	UNP Q82134
D	40	ILE	LEU	engineered mutation	UNP Q82134
E	40	ILE	LEU	engineered mutation	UNP Q82134
F	40	ILE	LEU	engineered mutation	UNP Q82134

- Molecule 2 is a protein called statine-containing inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	9	Total	C	N	O	S	0	0	0
			72	47	12	12	1			
2	J	8	Total	C	N	O	S	0	8	0
			128	84	22	20	2			
2	K	8	Total	C	N	O	S	0	8	0
			128	84	22	20	2			

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			7	4	3		
3	C	1	Total	C	O	0	0
			7	4	3		
3	E	1	Total	C	O	0	0
			7	4	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	58	Total	O	0	0
			58	58		
4	B	43	Total	O	0	0
			43	43		
4	I	2	Total	O	0	0
			2	2		
4	C	45	Total	O	0	0
			45	45		
4	D	30	Total	O	0	0
			30	30		
4	J	4	Total	O	0	0
			4	4		
4	E	38	Total	O	0	0
			38	38		
4	F	55	Total	O	0	0
			55	55		

*Continued on next page...*

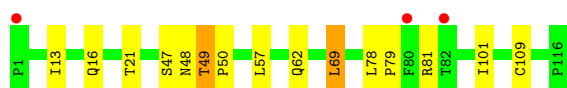
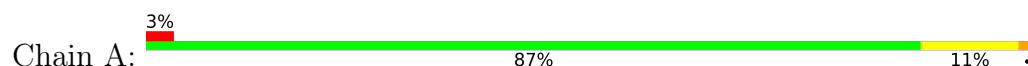
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	K	6	Total	O	0	0
			6	6		

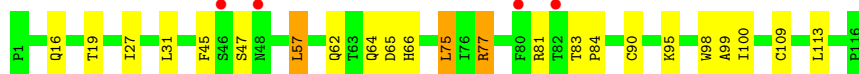
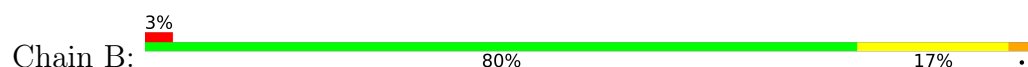
### 3 Residue-property plots [i](#)

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: Protease



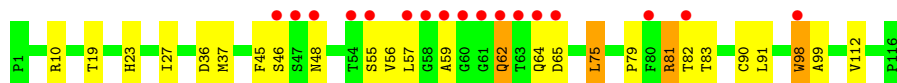
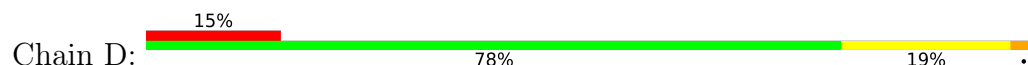
- Molecule 1: Protease



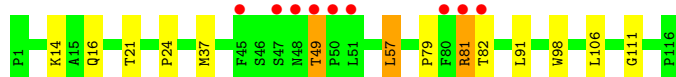
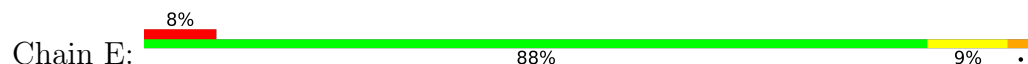
- Molecule 1: Protease



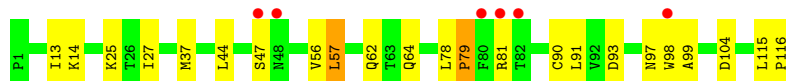
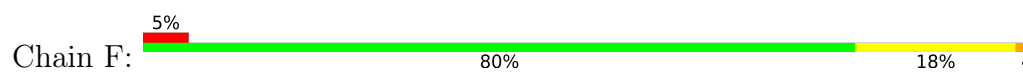
- Molecule 1: Protease



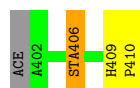
- Molecule 1: Protease



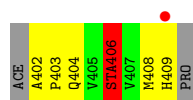
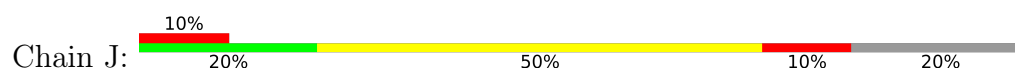
- Molecule 1: Protease



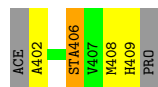
- Molecule 2: statine-containing inhibitor



- Molecule 2: statine-containing inhibitor



- Molecule 2: statine-containing inhibitor





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	133.69Å 77.39Å 80.17Å 90.00° 99.69° 90.00°	Depositor
Resolution (Å)	50.00 – 1.86 38.20 – 1.86	Depositor EDS
% Data completeness (in resolution range)	98.5 (50.00-1.86) 98.5 (38.20-1.86)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.5.0104	Depositor
R, $R_{free}$	0.188 , 0.222 0.193 , 0.227	Depositor DCC
$R_{free}$ test set	2021 reflections (3.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.2	Xtriage
Anisotropy	0.103	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 56.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5970	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.25% 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

### 5.1 Standard geometry

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

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	0.63	1/923 (0.1%)	0.71	1/1263 (0.1%)
1	B	0.54	1/928 (0.1%)	0.70	1/1269 (0.1%)
1	C	0.45	0/901	0.64	0/1233
1	D	0.46	0/901	0.64	0/1233
1	E	0.50	0/901	0.66	0/1233
1	F	0.57	0/917	0.74	1/1256 (0.1%)
2	I	0.67	0/62	0.62	0/82
2	J	0.43	0/106	0.45	0/140
2	K	0.51	0/106	0.64	0/140
All	All	0.53	2/5745 (0.0%)	0.68	3/7849 (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
1	A	0	1
2	I	0	3
2	J	0	4
2	K	0	4
All	All	0	12

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	109	CYS	CB-SG	-8.37	1.68	1.82
1	B	109	CYS	CB-SG	-6.12	1.71	1.82

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	104	ASP	CB-CG-OD1	5.69	123.42	118.30
1	A	69	LEU	CA-CB-CG	5.42	127.77	115.30
1	B	75	LEU	CA-CB-CG	5.21	127.28	115.30

There are no chirality outliers.

All (12) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	48	ASN	Peptide
2	I	406	STA	Peptide,Mainchain
2	I	409	HIS	Peptide
2	J	406[A]	STA	Peptide,Mainchain
2	J	406[B]	STA	Peptide,Mainchain
2	K	406[A]	STA	Peptide,Mainchain
2	K	406[B]	STA	Peptide,Mainchain

## 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	896	0	953	9	0
1	B	901	0	952	15	0
1	C	883	0	926	40	0
1	D	883	0	926	49	4
1	E	883	0	926	27	0
1	F	894	0	936	37	0
2	I	72	0	74	1	0
2	J	128	0	135	54	0
2	K	128	0	131	45	0
3	A	7	0	10	1	0
3	C	7	0	10	0	0
3	E	7	0	10	0	0
4	A	58	0	0	1	0
4	B	43	0	0	1	0
4	C	45	0	0	0	0
4	D	30	0	0	1	0
4	E	38	0	0	0	0
4	F	55	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	I	2	0	0	0	0
4	J	4	0	0	4	0
4	K	6	0	0	2	0
All	All	5970	0	5989	176	4

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 15.

All (176) 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:C:98:TRP:CH2	2:J:408[B]:MET:HE2	1.21	1.69
1:C:98:TRP:HH2	2:J:408[B]:MET:CE	0.98	1.59
1:E:98:TRP:HH2	2:K:408[B]:MET:CE	1.19	1.52
1:C:98:TRP:CH2	2:J:408[B]:MET:CE	1.79	1.45
1:D:37:MET:HB2	2:J:409[B]:HIS:NE2	1.16	1.42
1:E:98:TRP:CH2	2:K:408[B]:MET:CE	2.04	1.39
1:F:91:LEU:HD21	2:K:409[B]:HIS:CE1	1.60	1.36
1:D:37:MET:CB	2:J:409[B]:HIS:NE2	1.87	1.35
1:E:98:TRP:CH2	2:K:408[B]:MET:HE2	1.60	1.31
1:E:37:MET:HE2	2:K:409[A]:HIS:NE2	1.50	1.27
1:D:37:MET:SD	2:J:409[B]:HIS:CE1	2.30	1.24
1:F:91:LEU:HD11	2:K:409[B]:HIS:NE2	1.57	1.20
2:K:402[B]:ALA:N	4:K:192:HOH:O	1.75	1.17
1:C:37:MET:HB3	2:J:409[A]:HIS:NE2	1.59	1.16
1:B:98:TRP:CE3	1:B:100:ILE:HD11	1.81	1.14
1:E:91:LEU:HD21	2:K:409[A]:HIS:CE1	1.82	1.13
1:D:37:MET:CG	2:J:409[B]:HIS:NE2	2.11	1.13
1:E:98:TRP:CH2	2:K:408[B]:MET:HE1	1.77	1.12
1:F:57:LEU:HD13	2:K:408[B]:MET:HE3	1.30	1.12
1:F:98[B]:TRP:CZ3	4:F:281:HOH:O	2.01	1.11
1:C:98:TRP:CH2	2:J:408[B]:MET:HE1	1.84	1.11
1:F:98[B]:TRP:HZ3	4:F:281:HOH:O	1.31	1.11
1:E:91:LEU:HD21	2:K:409[A]:HIS:HE1	0.94	1.11
1:F:91:LEU:CD2	2:K:409[B]:HIS:HE1	1.63	1.10
1:F:57:LEU:HD13	2:K:408[B]:MET:CE	1.89	1.02
1:B:98:TRP:HE3	1:B:100:ILE:HD11	1.17	1.01
1:C:98:TRP:CZ3	2:J:408[B]:MET:HE2	1.95	1.01
1:C:91:LEU:HD11	2:J:409[A]:HIS:CE1	1.96	1.01
1:F:91:LEU:CD2	2:K:409[B]:HIS:CE1	2.43	0.98
1:D:19:THR:HB	1:D:45:PHE:CE2	1.98	0.98

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:97:ASN:HB2	1:F:98[B]:TRP:CE3	1.98	0.98
1:E:37:MET:HE2	2:K:409[A]:HIS:CE1	1.97	0.98
1:F:91:LEU:HD21	2:K:409[B]:HIS:HE1	0.84	0.98
1:C:49:THR:HG22	4:F:136:HOH:O	1.67	0.95
1:D:37:MET:HG3	2:J:409[B]:HIS:CD2	2.01	0.95
1:C:37:MET:CB	2:J:409[A]:HIS:NE2	2.29	0.94
1:D:37:MET:HB2	2:J:409[B]:HIS:CE1	2.01	0.94
1:F:37:MET:HE3	2:K:409[B]:HIS:CD2	2.03	0.93
1:F:91:LEU:HD11	2:K:409[B]:HIS:CE1	2.04	0.93
1:E:98:TRP:HH2	2:K:408[B]:MET:HE2	0.75	0.91
1:C:98:TRP:HH2	2:J:408[B]:MET:HE3	1.34	0.91
1:D:37:MET:CG	2:J:409[B]:HIS:CE1	2.54	0.89
1:D:37:MET:CB	2:J:409[B]:HIS:CE1	2.56	0.88
1:E:91:LEU:CD2	2:K:409[A]:HIS:HE1	1.84	0.87
1:F:57:LEU:CD1	2:K:408[B]:MET:HE3	2.04	0.87
1:E:16:GLN:HE21	1:E:24:PRO:HB2	1.38	0.86
1:F:91:LEU:CD1	2:K:409[B]:HIS:NE2	2.37	0.86
1:C:37:MET:SD	2:J:409[A]:HIS:CE1	2.69	0.85
1:D:19:THR:HG21	1:D:45:PHE:CZ	2.12	0.84
1:D:37:MET:HG3	2:J:409[B]:HIS:NE2	1.89	0.83
2:K:409[A]:HIS:O	4:K:285:HOH:O	1.97	0.83
1:D:19:THR:CG2	1:D:45:PHE:CZ	2.62	0.82
1:C:98:TRP:CZ3	2:J:408[B]:MET:CE	2.55	0.81
1:E:91:LEU:HD11	2:K:409[A]:HIS:NE2	1.96	0.81
1:D:81:ARG:CG	1:D:81:ARG:HH11	1.95	0.79
1:C:91:LEU:CD1	2:J:409[A]:HIS:CE1	2.66	0.79
1:F:98[A]:TRP:HE1	2:K:408[A]:MET:CE	1.97	0.78
1:D:10:ARG:HH12	2:J:404[B]:GLN:HB2	1.47	0.78
1:F:37:MET:CE	2:K:409[B]:HIS:CD2	2.66	0.78
1:E:91:LEU:HD11	2:K:409[A]:HIS:CE1	2.18	0.78
1:C:91:LEU:CG	2:J:409[A]:HIS:HE1	1.97	0.78
1:D:81:ARG:HH11	1:D:81:ARG:HG3	1.51	0.76
1:D:19:THR:CB	1:D:45:PHE:CE2	2.69	0.75
1:D:19:THR:HB	1:D:45:PHE:HE2	1.50	0.74
1:F:93:ASP:OD2	1:F:98[B]:TRP:HE3	1.70	0.74
1:D:57:LEU:HD23	2:J:408[B]:MET:HE3	1.70	0.73
1:B:57:LEU:CD2	1:B:62:GLN:HE21	2.03	0.71
1:D:36:ASP:HB2	2:J:409[B]:HIS:HD2	1.56	0.71
1:F:98[A]:TRP:HE1	2:K:408[A]:MET:HE2	1.55	0.71
1:C:21:THR:HB	1:C:49:THR:HG21	1.72	0.71
1:F:91:LEU:CD1	2:K:409[B]:HIS:CE1	2.73	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:37:MET:CE	1:D:90:CYS:HA	2.21	0.70
1:C:91:LEU:CG	2:J:409[A]:HIS:CE1	2.74	0.70
1:E:98:TRP:CZ3	2:K:408[B]:MET:CE	2.72	0.69
1:C:91:LEU:HG	2:J:409[A]:HIS:HE1	1.58	0.69
1:B:77[A]:ARG:HD3	1:B:84:PRO:HB3	1.76	0.68
1:E:37:MET:HG2	2:K:409[A]:HIS:NE2	2.09	0.68
1:C:91:LEU:HD21	2:J:409[A]:HIS:CE1	2.28	0.67
1:D:91:LEU:HD11	2:J:409[B]:HIS:CE1	2.31	0.66
1:E:57:LEU:HD13	2:K:408[A]:MET:HE3	1.77	0.66
1:F:27:ILE:HD12	1:F:99:ALA:HB1	1.76	0.66
1:D:37:MET:HE2	1:D:90:CYS:HA	1.79	0.65
2:J:406[B]:STA:CM	4:J:284:HOH:O	2.43	0.65
2:J:406[A]:STA:O	4:J:105:HOH:O	2.09	0.65
1:E:37:MET:HE2	2:K:409[A]:HIS:CD2	2.30	0.64
1:E:98:TRP:CZ3	2:K:408[B]:MET:HE1	2.29	0.64
1:E:37:MET:CE	2:K:409[A]:HIS:CE1	2.80	0.64
1:B:27:ILE:HD12	1:B:99:ALA:HB1	1.79	0.63
1:B:57:LEU:HD23	1:B:62:GLN:HE21	1.63	0.62
1:D:81:ARG:HG3	1:D:81:ARG:NH1	2.12	0.62
1:D:27:ILE:HD12	1:D:99:ALA:HB1	1.81	0.61
1:B:19[A]:THR:HG21	1:B:45:PHE:CD2	2.36	0.61
1:C:37:MET:SD	2:J:409[A]:HIS:NE2	2.75	0.60
1:F:91:LEU:CG	2:K:409[B]:HIS:CE1	2.84	0.60
1:C:98:TRP:HH2	2:J:408[B]:MET:HE2	0.55	0.60
1:D:10:ARG:NH1	2:J:404[B]:GLN:HB2	2.16	0.59
1:C:59:ALA:HA	1:D:98:TRP:HZ2	1.67	0.59
1:B:19[A]:THR:HG23	1:B:45:PHE:CE2	2.37	0.59
1:F:98[A]:TRP:HE1	2:K:408[A]:MET:HE1	1.67	0.59
1:C:98:TRP:CH2	2:J:408[B]:MET:HE3	2.18	0.59
1:F:97:ASN:HB2	1:F:98[B]:TRP:CZ3	2.37	0.59
1:D:98:TRP:CH2	2:J:406[B]:STA:HD21	2.37	0.59
1:C:91:LEU:CD2	2:J:409[A]:HIS:HE1	2.15	0.58
1:D:56:VAL:N	1:D:62:GLN:HE22	2.02	0.57
1:E:91:LEU:CD2	2:K:409[A]:HIS:CE1	2.70	0.57
1:C:59:ALA:HA	1:D:98:TRP:CZ2	2.39	0.57
1:E:21:THR:HG21	1:E:49:THR:HG21	1.85	0.57
1:D:56:VAL:C	1:D:62:GLN:NE2	2.59	0.56
1:D:62:GLN:NE2	2:J:402[A]:ALA:HB1	2.20	0.56
1:C:91:LEU:HD11	2:J:409[A]:HIS:NE2	2.20	0.55
1:B:16:GLN:HG3	4:B:128:HOH:O	2.06	0.55
1:E:21:THR:CG2	1:E:49:THR:HG21	2.37	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:J:406[B]:STA:HM2	4:J:284:HOH:O	2.07	0.54
1:E:16:GLN:HE21	1:E:24:PRO:CB	2.15	0.54
1:D:36:ASP:HB2	2:J:409[B]:HIS:CD2	2.41	0.53
1:E:14:LYS:O	1:E:79:PRO:HD3	2.09	0.53
1:A:57[B]:LEU:HD22	2:I:410:PRO:HD2	1.90	0.53
1:D:57:LEU:CD2	2:J:408[B]:MET:HE3	2.37	0.53
1:B:62:GLN:HG2	1:B:64:GLN:HE21	1.74	0.52
1:C:91:LEU:CD2	2:J:409[A]:HIS:CE1	2.91	0.52
1:A:16:GLN:HB2	3:A:117:PEG:H11	1.91	0.52
1:A:49:THR:HG23	4:D:121:HOH:O	2.10	0.52
1:F:78:LEU:HB3	1:F:79:PRO:HD2	1.92	0.52
1:F:57:LEU:HB3	2:K:408[B]:MET:HB2	1.92	0.51
1:C:75:LEU:HD22	1:C:86:VAL:HG22	1.91	0.51
1:F:91:LEU:CG	2:K:409[B]:HIS:NE2	2.74	0.51
1:D:37:MET:SD	2:J:409[B]:HIS:ND1	2.81	0.51
1:C:91:LEU:HG	2:J:409[A]:HIS:CE1	2.40	0.49
1:D:19:THR:CG2	1:D:45:PHE:CE2	2.96	0.49
1:B:66:HIS:CE1	1:B:95:LYS:HD2	2.48	0.49
1:C:58:GLY:HA3	1:D:59:ALA:HB3	1.95	0.49
1:B:19[A]:THR:CG2	1:B:45:PHE:CD2	2.96	0.48
1:F:14:LYS:O	1:F:79:PRO:HD3	2.13	0.48
1:F:93:ASP:OD2	1:F:98[B]:TRP:CE3	2.60	0.47
2:J:406[B]:STA:C	4:J:284:HOH:O	2.63	0.47
1:D:36:ASP:CB	2:J:409[B]:HIS:HD2	2.25	0.47
1:D:55:SER:HB3	1:D:62:GLN:OE1	2.14	0.47
1:D:37:MET:SD	2:J:409[B]:HIS:HE1	2.24	0.47
1:C:91:LEU:HD21	2:J:409[A]:HIS:HE1	1.72	0.46
1:B:19[A]:THR:CG2	1:B:45:PHE:CE2	2.99	0.46
1:D:75:LEU:CD2	1:D:75:LEU:N	2.78	0.46
1:F:25:LYS:HD2	1:F:44:LEU:HD23	1.97	0.46
1:C:36:ASP:OD1	1:D:10:ARG:NH1	2.49	0.46
1:D:75:LEU:N	1:D:75:LEU:HD23	2.31	0.45
1:C:4:PRO:HA	1:D:112:VAL:HG12	1.99	0.44
1:A:78:LEU:HB3	1:A:79:PRO:HD2	1.99	0.44
1:C:98:TRP:CZ3	2:J:408[B]:MET:HE1	2.36	0.44
1:C:98:TRP:CZ2	2:J:406[A]:STA:HD21	2.52	0.44
1:F:37:MET:HE3	1:F:37:MET:HB2	1.65	0.44
1:D:82:THR:HG23	1:D:83:THR:HG23	1.99	0.44
1:E:81:ARG:NH2	1:F:116:PRO:O	2.51	0.44
1:F:98[A]:TRP:NE1	2:K:408[A]:MET:HE2	2.29	0.44
1:A:49:THR:HA	1:A:50:PRO:HD3	1.79	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:63:THR:HG21	1:C:66:HIS:HB2	2.00	0.43
1:B:98:TRP:HE3	1:B:100:ILE:CD1	2.07	0.43
1:B:65:ASP:OD1	1:B:66:HIS:HD2	2.01	0.43
1:C:38:THR:HG22	1:C:40:ILE:HD12	2.00	0.43
1:F:37:MET:HE1	2:K:409[B]:HIS:CD2	2.53	0.43
1:D:19:THR:HG21	1:D:45:PHE:CE1	2.52	0.43
4:A:168:HOH:O	1:D:23:HIS:HE1	2.01	0.42
1:F:37:MET:HE3	2:K:409[B]:HIS:HD2	1.73	0.42
1:D:91:LEU:HG	2:J:409[B]:HIS:HE1	1.85	0.42
1:A:21:THR:HG21	1:A:49:THR:HG21	2.02	0.42
1:E:91:LEU:CD1	2:K:409[A]:HIS:CE1	2.97	0.42
1:C:17:VAL:HG11	1:C:40:ILE:HG12	2.01	0.41
1:A:57[B]:LEU:HD12	1:A:62:GLN:HG3	2.03	0.41
1:D:57:LEU:HD23	2:J:408[B]:MET:CE	2.45	0.41
1:F:56:VAL:HG22	2:K:409[B]:HIS:CE1	2.56	0.41
1:A:13:ILE:HG21	1:A:101[A]:ILE:HD12	2.03	0.41
1:C:58:GLY:O	1:D:98:TRP:HH2	2.03	0.41
1:A:21:THR:CG2	1:A:49:THR:HG21	2.51	0.40
1:C:79:PRO:O	1:C:80:PHE:HB2	2.21	0.40
1:E:111:GLY:HA3	1:F:115:LEU:HD23	2.04	0.40
2:J:402[A]:ALA:HA	2:J:403[A]:PRO:HD2	1.90	0.40
1:F:13:ILE:HD12	1:F:78:LEU:HD22	2.04	0.40
1:C:40:ILE:HG23	1:C:41:PRO:HD2	2.03	0.40

All (4) 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:D:64:GLN:N	1:D:64:GLN:OE1[2_756]	1.60	0.60
1:D:64:GLN:N	1:D:64:GLN:CD[2_756]	1.93	0.27
1:D:64:GLN:N	1:D:64:GLN:NE2[2_756]	2.10	0.10
1:D:64:GLN:CG	1:D:64:GLN:NE2[2_756]	2.16	0.04

## 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	117/116 (101%)	117 (100%)	0	0	100	100
1	B	117/116 (101%)	115 (98%)	1 (1%)	1 (1%)	17	6
1	C	114/116 (98%)	113 (99%)	1 (1%)	0	100	100
1	D	114/116 (98%)	108 (95%)	5 (4%)	1 (1%)	17	6
1	E	114/116 (98%)	113 (99%)	1 (1%)	0	100	100
1	F	115/116 (99%)	112 (97%)	2 (2%)	1 (1%)	17	6
2	I	6/10 (60%)	6 (100%)	0	0	100	100
2	J	10/10 (100%)	10 (100%)	0	0	100	100
2	K	10/10 (100%)	10 (100%)	0	0	100	100
All	All	717/726 (99%)	704 (98%)	10 (1%)	3 (0%)	34	19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	47	SER
1	D	79	PRO
1	F	79	PRO

### 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	105/102 (103%)	101 (96%)	4 (4%)	33	16
1	B	105/102 (103%)	96 (91%)	9 (9%)	10	2
1	C	102/102 (100%)	96 (94%)	6 (6%)	19	6
1	D	102/102 (100%)	95 (93%)	7 (7%)	15	4
1	E	102/102 (100%)	97 (95%)	5 (5%)	25	9
1	F	103/102 (101%)	97 (94%)	6 (6%)	20	6
2	I	7/7 (100%)	7 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	J	12/7 (171%)	12 (100%)	0	100	100
2	K	12/7 (171%)	12 (100%)	0	100	100
All	All	650/633 (103%)	613 (94%)	37 (6%)	20	6

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	47	SER
1	A	49	THR
1	A	69	LEU
1	A	81	ARG
1	B	31	LEU
1	B	57	LEU
1	B	75	LEU
1	B	77[A]	ARG
1	B	77[B]	ARG
1	B	81	ARG
1	B	83	THR
1	B	90	CYS
1	B	113	LEU
1	C	40	ILE
1	C	48	ASN
1	C	52	LYS
1	C	57	LEU
1	C	81	ARG
1	C	82	THR
1	D	46	SER
1	D	48	ASN
1	D	62	GLN
1	D	65	ASP
1	D	75	LEU
1	D	81	ARG
1	D	98	TRP
1	E	49	THR
1	E	57	LEU
1	E	81	ARG
1	E	82	THR
1	E	106	LEU
1	F	47	SER
1	F	57	LEU
1	F	62	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	64	GLN
1	F	81	ARG
1	F	90	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	62	GLN
1	B	23	HIS
1	B	62	GLN
1	B	64	GLN
1	B	66	HIS
1	C	48	ASN
1	C	97	ASN
1	D	62	GLN
1	D	66	HIS
1	E	16	GLN
1	E	64	GLN
1	E	107	GLN
1	F	97	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

5 non-standard protein/DNA/RNA residues are 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	$\# Z  > 2$
2	STA	J	406[A]	2	10,10,11	0.85	0	9,12,14	1.25	2 (22%)
2	STA	K	406[A]	2	10,10,11	0.85	0	9,12,14	1.27	1 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	STA	I	406	2	10,10,11	2.24	2 (20%)	9,12,14	2.86	3 (33%)
2	STA	J	406[B]	2	10,10,11	0.96	0	9,12,14	1.00	0
2	STA	K	406[B]	2	10,10,11	0.88	0	9,12,14	0.99	1 (11%)

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
2	STA	J	406[A]	2	-	7/11/11/12	-
2	STA	K	406[A]	2	-	3/11/11/12	-
2	STA	I	406	2	-	3/11/11/12	-
2	STA	J	406[B]	2	-	8/11/11/12	-
2	STA	K	406[B]	2	-	3/11/11/12	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	406	STA	CM-C	-6.05	1.34	1.49
2	I	406	STA	CH-CA	3.15	1.56	1.53

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	406	STA	O-C-CM	-7.36	103.97	125.43
2	I	406	STA	OH-CH-CA	2.63	114.15	109.40
2	K	406[A]	STA	CG-CB-CA	-2.60	110.22	115.82
2	J	406[A]	STA	CG-CB-CA	2.51	121.22	115.82
2	I	406	STA	CG-CB-CA	-2.48	110.49	115.82
2	J	406[A]	STA	CH-CM-C	-2.29	109.09	113.11
2	K	406[B]	STA	CH-CM-C	-2.02	109.56	113.11

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	406	STA	CA-CH-CM-C
2	I	406	STA	OH-CH-CM-C
2	J	406[A]	STA	CA-CH-CM-C

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	J	406[A]	STA	OH-CH-CM-C
2	J	406[B]	STA	CB-CA-CH-OH
2	J	406[B]	STA	CB-CA-CH-CM
2	J	406[B]	STA	CA-CH-CM-C
2	J	406[B]	STA	OH-CH-CM-C
2	K	406[A]	STA	CA-CH-CM-C
2	K	406[A]	STA	OH-CH-CM-C
2	K	406[B]	STA	CA-CH-CM-C
2	K	406[B]	STA	OH-CH-CM-C
2	J	406[B]	STA	CA-CB-CG-CD2
2	J	406[B]	STA	CA-CB-CG-CD1
2	J	406[A]	STA	CB-CA-CH-OH
2	J	406[B]	STA	N-CA-CH-CM
2	I	406	STA	CB-CA-CH-CM
2	J	406[A]	STA	CB-CA-CH-CM
2	J	406[A]	STA	CA-CB-CG-CD2
2	J	406[A]	STA	O-C-CM-CH
2	J	406[B]	STA	O-C-CM-CH
2	K	406[A]	STA	O-C-CM-CH
2	K	406[B]	STA	O-C-CM-CH
2	J	406[A]	STA	CA-CB-CG-CD1

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	J	406[A]	STA	2	0
2	J	406[B]	STA	4	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

3 ligands are 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	$\# Z  > 2$
3	PEG	C	117	-	6,6,6	0.42	0	5,5,5	0.34	0
3	PEG	E	117	-	6,6,6	0.44	0	5,5,5	0.35	0
3	PEG	A	117	-	6,6,6	0.39	0	5,5,5	0.36	0

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
3	PEG	C	117	-	-	2/4/4/4	-
3	PEG	E	117	-	-	3/4/4/4	-
3	PEG	A	117	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	117	PEG	O1-C1-C2-O2
3	E	117	PEG	O1-C1-C2-O2
3	E	117	PEG	O2-C3-C4-O4
3	C	117	PEG	C4-C3-O2-C2
3	A	117	PEG	O2-C3-C4-O4
3	A	117	PEG	C4-C3-O2-C2
3	E	117	PEG	C4-C3-O2-C2
3	A	117	PEG	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	117	PEG	1	0

## 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	116/116 (100%)	-0.02	3 (2%) 56 54	19, 33, 55, 72	0
1	B	116/116 (100%)	0.18	4 (3%) 45 42	22, 40, 70, 87	0
1	C	116/116 (100%)	0.23	11 (9%) 8 7	26, 40, 70, 85	0
1	D	116/116 (100%)	0.66	17 (14%) 2 2	26, 41, 90, 113	0
1	E	116/116 (100%)	0.38	9 (7%) 13 13	25, 45, 72, 91	0
1	F	116/116 (100%)	0.04	6 (5%) 27 26	20, 34, 57, 75	0
2	I	8/10 (80%)	0.33	0 100 100	28, 39, 52, 67	0
2	J	7/10 (70%)	1.29	1 (14%) 2 2	43, 54, 60, 94	0
2	K	7/10 (70%)	0.55	0 100 100	27, 36, 42, 74	0
All	All	718/726 (98%)	0.26	51 (7%) 16 15	19, 39, 72, 113	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	80	PHE	9.3
1	D	60	GLY	8.2
1	D	80	PHE	7.6
1	E	48	ASN	6.6
1	D	48	ASN	6.2
1	F	80	PHE	6.2
1	D	64	GLN	5.6
1	D	61	GLY	5.5
1	F	82	THR	5.4
1	D	62	GLN	5.2
1	D	47	SER	5.2
1	F	48	ASN	5.0
1	D	46	SER	4.8
2	J	409[A]	HIS	4.7
1	A	80	PHE	4.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	47	SER	3.9
1	A	82	THR	3.9
1	E	82	THR	3.6
1	C	59	ALA	3.5
1	B	48	ASN	3.4
1	E	50	PRO	3.2
1	D	63	THR	3.2
1	C	80	PHE	3.2
1	D	55	SER	3.0
1	B	82	THR	3.0
1	C	60	GLY	2.9
1	E	49	THR	2.8
1	F	47	SER	2.8
1	D	65	ASP	2.7
1	C	66	HIS	2.7
1	C	61	GLY	2.7
1	A	1	PRO	2.7
1	D	82	THR	2.6
1	D	57	LEU	2.5
1	D	54	THR	2.5
1	C	62	GLN	2.5
1	C	64	GLN	2.5
1	D	59	ALA	2.4
1	E	45	PHE	2.4
1	C	48	ASN	2.3
1	B	46	SER	2.3
1	F	81	ARG	2.3
1	E	80	PHE	2.3
1	C	82	THR	2.3
1	D	58	GLY	2.2
1	D	98	TRP	2.2
1	C	1	PRO	2.2
1	E	81	ARG	2.1
1	C	65	ASP	2.1
1	F	98[A]	TRP	2.0
1	E	51	LEU	2.0

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

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
2	STA	I	406	11/12	0.88	0.13	23,27,32,32	0
2	STA	J	406[A]	11/12	0.92	0.14	28,34,46,55	11
2	STA	J	406[B]	11/12	0.92	0.14	39,47,61,62	11
2	STA	K	406[A]	11/12	0.95	0.20	23,27,32,34	11
2	STA	K	406[B]	11/12	0.95	0.20	26,28,33,37	11

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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
3	PEG	E	117	7/7	0.85	0.12	51,56,65,66	0
3	PEG	C	117	7/7	0.89	0.11	37,53,56,62	0
3	PEG	A	117	7/7	0.90	0.10	41,46,66,68	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.