



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 12:39 AM GMT

PDB ID : 2B7F  
Title : Crystal structure of human T-cell leukemia virus protease, a novel target for anti-cancer design  
Authors : Li, M.; Laco, G.S.; Jaskolski, M.; Rozycki, J.; Alexandratos, J.; Wlodawer, A.; Gustchina, A.  
Deposited on : 2005-10-04  
Resolution : 2.60 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

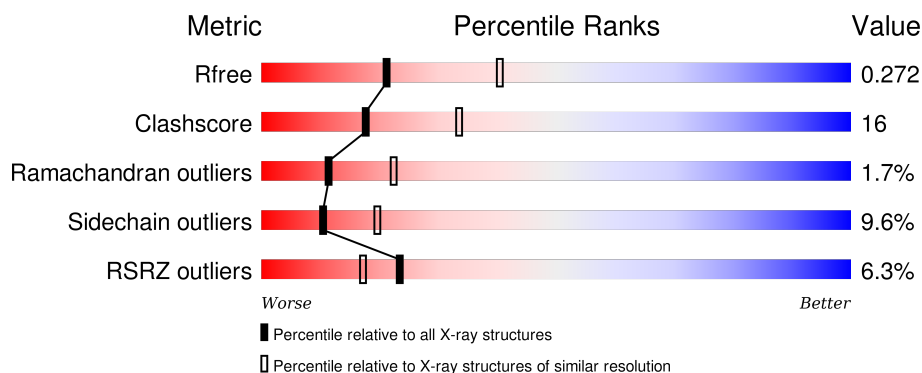
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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. 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>4%</div> <div>78%</div> <div>21%</div> <div>.</div> </div>
1	B	116	<div> <div>2%</div> <div>63%</div> <div>32%</div> <div>..</div> </div>
1	C	116	<div> <div>8%</div> <div>65%</div> <div>32%</div> <div>..</div> </div>
1	D	116	<div> <div>11%</div> <div>64%</div> <div>29%</div> <div>7%</div> </div>
1	E	116	<div> <div>4%</div> <div>67%</div> <div>28%</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>3%</div><div>71%</div><div>26%</div><div></div></div>
2	I	10	<div><div></div><div>20%</div><div>50%</div><div>30%</div><div>10%</div><div>10%</div></div>
2	J	10	<div><div></div><div>30%</div><div>10%</div><div>40%</div><div>20%</div><div>10%</div><div>20%</div></div>
2	K	10	<div><div></div><div>20%</div><div>60%</div><div>20%</div><div>10%</div><div>10%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5764 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 HTLV protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	0	0	0
			883	564	152	164	3			
1	B	116	Total	C	N	O	S	0	0	0
			883	564	152	164	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	0	0
			883	564	152	164	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	ILE	LEU	ENGINEERED	UNP P10274
B	40	ILE	LEU	ENGINEERED	UNP P10274
C	40	ILE	LEU	ENGINEERED	UNP P10274
D	40	ILE	LEU	ENGINEERED	UNP P10274
E	40	ILE	LEU	ENGINEERED	UNP P10274
F	40	ILE	LEU	ENGINEERED	UNP P10274

- Molecule 2 is a protein called (ACE)APQV(STA)VMHP peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	10	Total	C	N	O	S	0	0	0
			75	49	12	13	1			
2	J	8	Total	C	N	O	S	0	8	0
			134	88	22	22	2			
2	K	10	Total	C	N	O	S	0	0	0
			75	49	12	13	1			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		

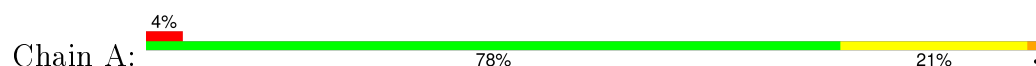
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	28	Total	O	0	0
			28	28		
4	B	28	Total	O	0	0
			28	28		
4	C	23	Total	O	0	0
			23	23		
4	D	21	Total	O	0	0
			21	21		
4	E	27	Total	O	0	0
			27	27		
4	F	42	Total	O	0	0
			42	42		
4	I	2	Total	O	0	0
			2	2		
4	K	1	Total	O	0	0
			1	1		

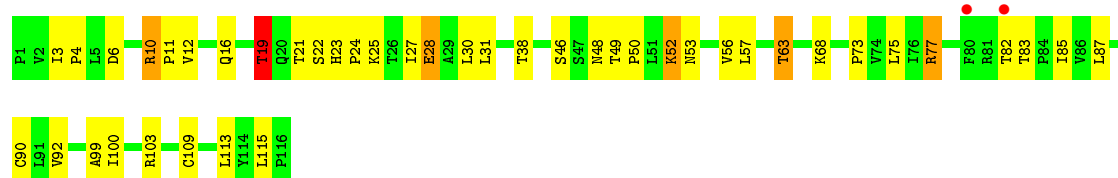
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry 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.

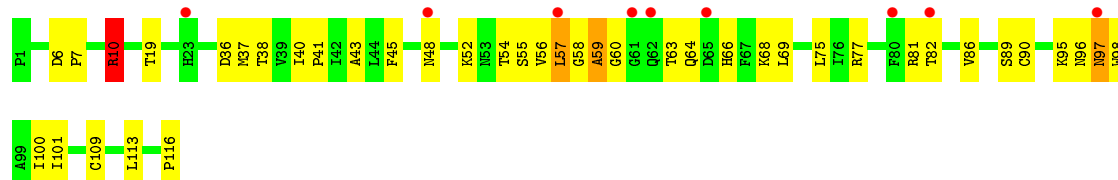
- Molecule 1: HTLV protease



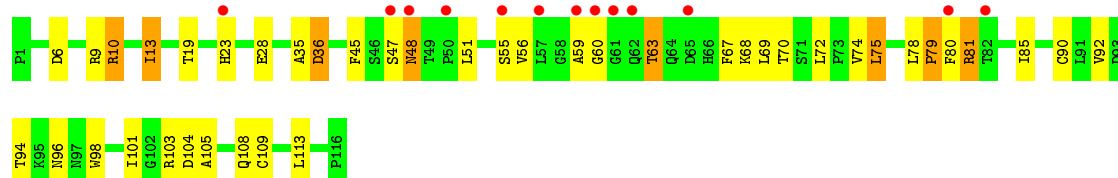
- Molecule 1: HTLV protease



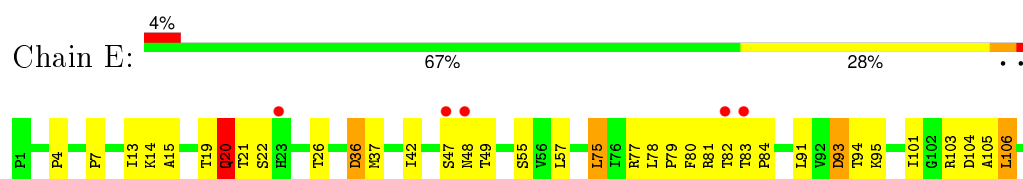
- Molecule 1: HTLV protease



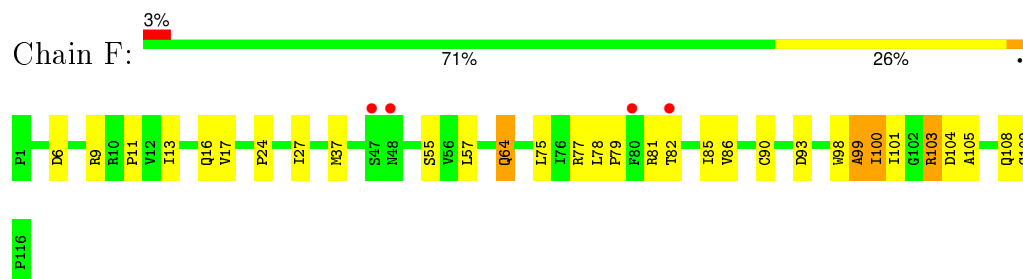
- Molecule 1: HTLV protease



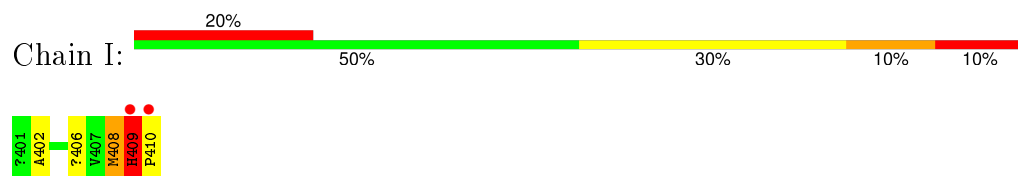
- Molecule 1: HTLV protease



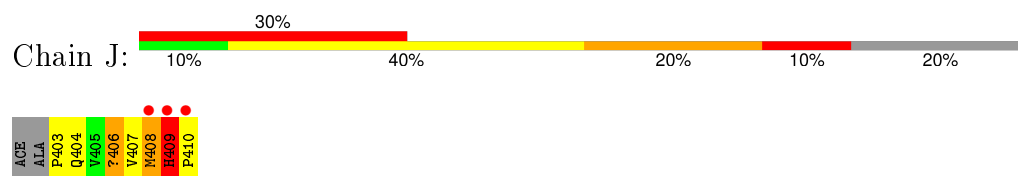
- Molecule 1: HTLV protease



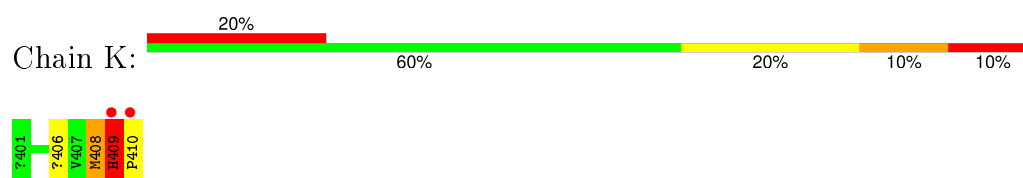
- Molecule 2: (ACE)APQV(STA)VMHP peptide



- Molecule 2: (ACE)APQV(STA)VMHP peptide



- Molecule 2: (ACE)APQV(STA)VMHP peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.32Å 77.79Å 80.38Å 90.00° 99.28° 90.00°	Depositor
Resolution (Å)	10.00 – 2.60 19.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	97.6 (10.00-2.60) 97.6 (19.99-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.10 (at 2.59Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.198 , 0.278 0.196 , 0.272	Depositor DCC
$R_{free}$ test set	1143 reflections (4.96%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.9	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 65.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 24590 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5764	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.11% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: PO4, STA, ACE

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	1.00	1/901 (0.1%)	1.03	3/1233 (0.2%)
1	B	0.95	0/901	1.13	5/1233 (0.4%)
1	C	1.01	1/901 (0.1%)	1.08	4/1233 (0.3%)
1	D	1.07	2/901 (0.2%)	1.09	4/1233 (0.3%)
1	E	0.95	0/901	1.10	4/1233 (0.3%)
1	F	0.92	0/901	1.01	3/1233 (0.2%)
2	I	0.88	0/63	0.97	0/84
2	J	0.64	0/114	0.71	0/148
2	K	0.79	0/63	0.91	0/84
All	All	0.98	4/5646 (0.1%)	1.07	23/7714 (0.3%)

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	C	0	1
2	I	0	3
2	J	0	5
2	K	0	3
All	All	0	12

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	109	CYS	CB-SG	-8.27	1.68	1.82
1	D	28	GLU	CG-CD	5.71	1.60	1.51
1	C	109	CYS	CB-SG	-5.45	1.73	1.81
1	D	90	CYS	CB-SG	-5.00	1.73	1.81

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	103	ARG	NE-CZ-NH2	-10.44	115.08	120.30
1	F	103	ARG	NE-CZ-NH2	-10.37	115.11	120.30
1	B	103	ARG	NE-CZ-NH1	8.49	124.55	120.30
1	D	103	ARG	NE-CZ-NH2	-7.67	116.47	120.30
1	C	6	ASP	CB-CG-OD1	7.59	125.13	118.30
1	B	6	ASP	CB-CG-OD1	7.37	124.94	118.30
1	D	6	ASP	CB-CG-OD1	7.08	124.67	118.30
1	E	36	ASP	CB-CG-OD1	6.60	124.24	118.30
1	D	36	ASP	CB-CG-OD1	6.29	123.96	118.30
1	C	57	LEU	CA-CB-CG	6.16	129.47	115.30
1	E	36	ASP	CB-CG-OD2	-5.99	112.91	118.30
1	B	19	THR	N-CA-CB	-5.93	99.03	110.30
1	B	31	LEU	CA-CB-CG	5.74	128.49	115.30
1	C	10	ARG	NE-CZ-NH1	-5.69	117.45	120.30
1	C	113	LEU	CB-CG-CD2	-5.44	101.76	111.00
1	E	75	LEU	CA-CB-CG	5.43	127.79	115.30
1	A	6	ASP	CB-CG-OD2	-5.25	113.57	118.30
1	A	44	LEU	CA-CB-CG	5.25	127.38	115.30
1	A	109	CYS	CB-CA-C	-5.21	99.99	110.40
1	E	93	ASP	CB-CG-OD1	5.15	122.93	118.30
1	F	6	ASP	CB-CG-OD1	5.06	122.85	118.30
1	D	103	ARG	NE-CZ-NH1	5.05	122.83	120.30
1	F	99	ALA	N-CA-C	-5.03	97.43	111.00

There are no chirality outliers.

All (12) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	97	ASN	Peptide
2	I	406	STA	Mainchain,Peptide
2	I	409	HIS	Peptide
2	J	406[A]	STA	Mainchain,Peptide
2	J	406[B]	STA	Mainchain,Peptide
2	J	409[B]	HIS	Peptide
2	K	406	STA	Mainchain,Peptide
2	K	409	HIS	Peptide

## 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	883	0	926	19	0
1	B	883	0	926	27	0
1	C	883	0	926	32	1
1	D	883	0	926	49	0
1	E	883	0	926	22	0
1	F	883	0	926	28	0
2	I	75	0	78	9	0
2	J	134	0	140	43	0
2	K	75	0	78	5	0
3	A	5	0	0	0	0
3	C	5	0	0	0	1
4	A	28	0	0	2	0
4	B	28	0	0	1	0
4	C	23	0	0	0	0
4	D	21	0	0	3	0
4	E	27	0	0	0	0
4	F	42	0	0	4	0
4	I	2	0	0	0	0
4	K	1	0	0	0	0
All	All	5764	0	5852	182	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 16.

All (182) 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:37:MET:CE	1:A:90:CYS:HA	1.78	1.14
1:A:10:ARG:HH11	1:A:10:ARG:CG	1.74	0.98
1:D:98:TRP:CH2	2:J:408[A]:MET:SD	2.57	0.98
2:J:409[A]:HIS:HB3	2:J:410[A]:PRO:HD2	1.46	0.96
2:J:409[B]:HIS:HB3	2:J:410[B]:PRO:HD2	1.46	0.95
1:D:98:TRP:HH2	2:J:408[A]:MET:SD	1.89	0.95
2:I:409:HIS:HB3	2:I:410:PRO:HD2	1.55	0.89
1:A:37:MET:HE1	1:A:90:CYS:HA	1.57	0.86

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:36:ASP:OD2	2:J:409[B]:HIS:N	2.09	0.85
1:A:37:MET:HE2	1:A:90:CYS:HA	1.59	0.84
1:D:56:VAL:O	1:D:63:THR:HG23	1.78	0.83
1:D:98:TRP:CH2	2:J:408[A]:MET:CE	2.61	0.83
1:D:10:ARG:NH1	2:J:404[B]:GLN:OE1	2.13	0.82
2:K:409:HIS:HB3	2:K:410:PRO:HD2	1.62	0.81
1:A:10:ARG:HH11	1:A:10:ARG:HG2	1.46	0.80
2:J:409[B]:HIS:HB3	2:J:410[B]:PRO:CD	2.11	0.80
1:D:98:TRP:HH2	2:J:408[A]:MET:CG	1.96	0.79
1:E:37:MET:HE3	1:E:91:LEU:HG	1.65	0.78
1:F:75:LEU:HD22	1:F:86:VAL:HG22	1.65	0.78
2:J:409[A]:HIS:HB3	2:J:410[A]:PRO:CD	2.14	0.77
1:C:10:ARG:NH2	2:J:408[B]:MET:HB3	1.99	0.77
1:B:10:ARG:NH1	2:I:408:MET:SD	2.59	0.76
1:A:10:ARG:HH11	1:A:10:ARG:HG3	1.48	0.75
1:D:98:TRP:CH2	2:J:408[A]:MET:HE2	2.21	0.75
1:D:55:SER:O	2:J:403[A]:PRO:HD2	1.87	0.74
1:D:56:VAL:O	1:D:63:THR:CG2	2.36	0.74
1:C:75:LEU:HD22	1:C:86:VAL:HG22	1.69	0.74
1:F:78:LEU:HB3	1:F:79:PRO:HD2	1.70	0.74
1:D:98:TRP:CZ3	2:J:408[A]:MET:SD	2.81	0.73
1:D:98:TRP:HH2	2:J:408[A]:MET:HG3	1.52	0.73
1:A:10:ARG:NH1	1:A:10:ARG:CG	2.45	0.72
1:C:43:ALA:HB2	1:C:96:ASN:HD21	1.55	0.72
2:I:409:HIS:HB3	2:I:410:PRO:CD	2.19	0.72
1:B:85:ILE:HD11	1:B:109:CYS:SG	2.30	0.72
1:B:12:VAL:HG21	1:B:28:GLU:OE2	1.90	0.71
1:A:40:ILE:HD12	1:A:90:CYS:SG	2.30	0.71
1:C:55:SER:HB3	1:C:64:GLN:NE2	2.08	0.69
1:D:98:TRP:CH2	2:J:408[A]:MET:HG3	2.28	0.68
1:C:41:PRO:HB2	1:C:96:ASN:OD1	1.93	0.68
1:D:56:VAL:HG22	2:J:403[A]:PRO:HG2	1.77	0.67
1:E:101:ILE:HG22	1:E:106:LEU:HD13	1.76	0.65
1:D:48:ASN:HD22	1:D:48:ASN:C	2.00	0.65
1:C:98:TRP:CH2	2:J:406[A]:STA:HD11	2.32	0.65
1:C:56:VAL:HG13	2:J:403[B]:PRO:O	1.97	0.64
1:F:85:ILE:HG23	1:F:108:GLN:HG3	1.80	0.64
1:D:13:ILE:HD11	1:D:101:ILE:HD12	1.79	0.64
1:D:75:LEU:N	1:D:75:LEU:HD23	2.13	0.64
1:E:7:PRO:HG3	1:F:104:ASP:HA	1.80	0.64
1:C:55:SER:HB3	1:C:64:GLN:HE22	1.62	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:57:LEU:HD23	2:I:402:ALA:HB3	1.80	0.63
1:F:77:ARG:NH1	1:F:81:ARG:O	2.31	0.63
1:D:85:ILE:HG23	1:D:108:GLN:HG3	1.82	0.61
1:F:77:ARG:HH11	1:F:77:ARG:HG2	1.63	0.61
1:D:10:ARG:NH1	2:J:404[B]:GLN:CD	2.54	0.61
2:J:409[A]:HIS:O	2:J:410[A]:PRO:C	2.39	0.60
1:C:98:TRP:HH2	2:J:408[B]:MET:HG3	1.67	0.60
1:E:37:MET:CE	1:E:91:LEU:HG	2.31	0.60
2:K:409:HIS:HB3	2:K:410:PRO:CD	2.31	0.59
1:C:7:PRO:HG3	1:D:104:ASP:HA	1.84	0.59
1:B:30:LEU:HB3	1:B:100:ILE:HD13	1.83	0.59
1:A:66:HIS:CD2	4:A:216:HOH:O	2.55	0.59
1:C:36:ASP:OD2	2:J:408[A]:MET:HA	2.02	0.58
1:D:98:TRP:CZ3	2:J:408[A]:MET:CE	2.85	0.58
2:J:409[B]:HIS:O	2:J:410[B]:PRO:C	2.41	0.58
1:C:55:SER:CB	1:C:64:GLN:NE2	2.66	0.57
2:J:409[B]:HIS:CB	2:J:410[B]:PRO:CD	2.81	0.57
1:A:2:VAL:HG12	1:A:2:VAL:O	2.03	0.57
2:J:409[A]:HIS:CB	2:J:410[A]:PRO:CD	2.82	0.57
1:A:10:ARG:NH1	1:A:10:ARG:HG3	2.12	0.55
1:C:66:HIS:ND1	1:D:60:GLY:HA2	2.20	0.55
1:D:10:ARG:HH12	2:J:404[B]:GLN:CD	2.09	0.55
1:D:98:TRP:HH2	2:J:408[A]:MET:CE	2.10	0.55
1:F:104:ASP:O	1:F:108:GLN:HG2	2.07	0.55
1:B:23:HIS:O	1:B:25:LYS:HE3	2.07	0.54
1:C:43:ALA:HB2	1:C:96:ASN:ND2	2.21	0.54
1:D:10:ARG:HH12	2:J:408[A]:MET:HB3	1.72	0.54
1:B:19:THR:CG2	1:B:21:THR:HB	2.37	0.54
1:D:68:LYS:HG2	1:D:94:THR:CG2	2.38	0.54
1:C:37:MET:CE	1:C:90:CYS:HA	2.38	0.53
1:C:98:TRP:CH2	2:J:408[B]:MET:HE2	2.43	0.53
1:F:55:SER:HB3	1:F:64:GLN:OE1	2.09	0.53
1:B:68:LYS:O	1:B:92:VAL:N	2.40	0.52
1:F:9:ARG:HD3	4:F:137:HOH:O	2.09	0.52
1:B:10:ARG:NH1	2:I:408:MET:CE	2.72	0.52
1:D:98:TRP:CH2	2:J:408[A]:MET:CG	2.82	0.52
1:C:10:ARG:NH1	2:J:408[B]:MET:HE3	2.26	0.51
1:C:19:THR:HB	1:C:45:PHE:CE2	2.45	0.51
1:E:114:TYR:CZ	1:E:116:PRO:HA	2.45	0.51
1:E:116:PRO:HG3	1:F:110:GLN:HB2	1.92	0.51
2:I:409:HIS:CB	2:I:410:PRO:CD	2.88	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:77:ARG:CZ	1:E:84:PRO:HG3	2.41	0.51
1:B:48:ASN:O	1:B:50:PRO:HD3	2.11	0.51
1:D:23:HIS:HB2	4:D:136:HOH:O	2.11	0.50
1:E:42:ILE:HD12	1:E:94:THR:HG22	1.94	0.50
1:E:103:ARG:HD3	1:F:11:PRO:HD3	1.94	0.50
1:C:38:THR:HG23	1:C:100:ILE:O	2.11	0.50
1:B:38:THR:HG23	1:B:100:ILE:O	2.12	0.49
1:C:37:MET:HE2	1:C:90:CYS:HA	1.93	0.49
2:I:409:HIS:O	2:I:410:PRO:C	2.51	0.49
1:D:51:LEU:HD23	1:D:70:THR:HA	1.95	0.48
1:F:103:ARG:NH2	4:F:135:HOH:O	2.46	0.48
1:C:58:GLY:O	1:C:60:GLY:N	2.47	0.48
1:E:78:LEU:O	1:E:79:PRO:C	2.51	0.48
1:C:57:LEU:HD22	1:D:98:TRP:CZ2	2.48	0.48
1:D:68:LYS:O	1:D:92:VAL:N	2.46	0.48
1:A:36:ASP:OD2	2:I:408:MET:HA	2.14	0.48
1:E:36:ASP:OD2	2:K:408:MET:HA	2.14	0.48
1:D:48:ASN:C	1:D:48:ASN:ND2	2.68	0.47
1:A:95:LYS:HE2	1:A:95:LYS:HB3	1.63	0.47
1:C:77:ARG:HD3	1:C:82:THR:O	2.14	0.47
1:B:52:LYS:NZ	1:B:53:ASN:H	2.13	0.47
2:J:409[A]:HIS:ND1	2:J:410[A]:PRO:HD3	2.29	0.47
1:B:21:THR:HG22	1:B:21:THR:O	2.15	0.46
1:C:36:ASP:OD2	2:J:409[A]:HIS:N	2.47	0.46
1:B:12:VAL:CG2	1:B:28:GLU:OE2	2.61	0.46
1:D:74:VAL:C	1:D:75:LEU:HD23	2.36	0.45
1:C:55:SER:O	2:J:403[B]:PRO:HD2	2.15	0.45
1:E:42:ILE:HB	1:E:93:ASP:O	2.16	0.45
1:B:49:THR:HA	1:B:50:PRO:HD3	1.77	0.45
1:C:57:LEU:HD22	1:D:98:TRP:HZ2	1.82	0.45
1:E:116:PRO:HG2	1:F:109:CYS:O	2.17	0.45
1:F:37:MET:CE	1:F:90:CYS:HA	2.47	0.45
1:D:69:LEU:HD12	1:D:69:LEU:HA	1.72	0.45
1:F:13:ILE:HD12	1:F:78:LEU:HD22	1.99	0.44
1:F:98:TRP:CE3	1:F:100:ILE:HD11	2.52	0.44
1:B:77:ARG:HE	1:B:77:ARG:HB2	1.36	0.44
1:D:104:ASP:OD1	1:D:105:ALA:N	2.49	0.44
1:B:25:LYS:HG2	1:E:20:GLN:HB3	2.00	0.44
1:D:109:CYS:HB2	4:D:124:HOH:O	2.16	0.44
1:B:27:ILE:HD12	1:B:99:ALA:HB1	1.98	0.44
1:A:95:LYS:NZ	4:A:212:HOH:O	2.46	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:15:ALA:O	1:E:26:THR:HA	2.18	0.44
1:B:21:THR:HG22	1:B:46:SER:HB3	1.99	0.44
1:E:19:THR:O	1:E:21:THR:HG23	2.18	0.44
1:B:16:GLN:HG3	4:B:129:HOH:O	2.18	0.44
1:C:98:TRP:HH2	2:J:404[A]:GLN:CD	2.21	0.43
1:F:27:ILE:HD12	1:F:99:ALA:HB1	2.00	0.43
1:E:104:ASP:OD1	1:E:105:ALA:N	2.48	0.43
1:D:75:LEU:CD2	1:D:75:LEU:N	2.81	0.43
1:B:11:PRO:HD2	1:B:30:LEU:CD1	2.48	0.43
1:A:83:THR:HA	1:A:84:PRO:HD3	1.74	0.43
2:J:409[B]:HIS:ND1	2:J:410[B]:PRO:HD3	2.33	0.43
1:B:19:THR:HG22	1:B:21:THR:H	1.82	0.43
1:F:17:VAL:O	1:F:24:PRO:HA	2.19	0.43
1:C:54:THR:HG22	1:C:68:LYS:HA	1.99	0.43
1:D:98:TRP:CH2	2:J:406[B]:STA:HD11	2.53	0.43
2:K:409:HIS:O	2:K:410:PRO:C	2.56	0.43
1:B:11:PRO:HD2	1:B:30:LEU:HD11	2.00	0.43
1:E:4:PRO:HA	1:F:112:VAL:HG12	2.01	0.43
1:F:93:ASP:OD1	1:F:93:ASP:C	2.57	0.43
1:F:9:ARG:NH1	4:F:137:HOH:O	2.45	0.42
1:A:40:ILE:HA	1:A:41:PRO:HD3	1.86	0.42
1:B:73:PRO:HA	1:B:87:LEU:O	2.19	0.42
1:A:101:ILE:HG23	1:A:105:ALA:HB3	2.01	0.42
1:D:48:ASN:O	1:D:48:ASN:ND2	2.52	0.42
1:E:103:ARG:CD	1:F:11:PRO:HD3	2.49	0.42
1:C:59:ALA:HB1	1:D:67:PHE:CZ	2.54	0.42
1:C:69:LEU:HD22	1:C:89:SER:OG	2.19	0.42
1:D:10:ARG:NH1	2:J:408[A]:MET:SD	2.92	0.42
1:E:114:TYR:HB3	1:F:112:VAL:HG23	2.01	0.42
1:D:56:VAL:O	1:D:63:THR:HG22	2.17	0.42
1:F:101:ILE:HG23	1:F:105:ALA:HB3	2.02	0.42
1:B:23:HIS:HA	1:B:24:PRO:HD2	1.80	0.42
1:D:35:ALA:HA	2:J:407[B]:VAL:HB	2.00	0.42
1:D:19:THR:HB	1:D:45:PHE:CE2	2.55	0.42
1:C:95:LYS:HB3	1:C:97:ASN:OD1	2.20	0.42
1:F:77:ARG:NH2	4:F:151:HOH:O	2.52	0.41
1:E:14:LYS:O	1:E:79:PRO:HD3	2.20	0.41
1:D:98:TRP:N	4:D:122:HOH:O	2.53	0.41
1:B:56:VAL:O	1:B:63:THR:HG22	2.20	0.41
1:F:13:ILE:CD1	1:F:78:LEU:HD22	2.50	0.41
1:F:85:ILE:HD13	1:F:105:ALA:HB1	2.01	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:78:LEU:HB3	1:D:79:PRO:HD2	2.02	0.41
1:D:98:TRP:CZ3	2:J:408[A]:MET:HE1	2.56	0.41
1:C:98:TRP:CH2	2:J:408[B]:MET:HG3	2.51	0.41
1:B:3:ILE:HA	1:B:4:PRO:HD3	1.83	0.41
1:D:80:PHE:O	1:D:81:ARG:C	2.59	0.41
1:F:77:ARG:HH11	1:F:77:ARG:CG	2.31	0.40
1:A:62:GLN:NE2	2:I:410:PRO:O	2.54	0.40
2:K:409:HIS:CB	2:K:410:PRO:CD	2.97	0.40
1:A:38:THR:HG23	1:A:100:ILE:O	2.22	0.40
1:E:81:ARG:HD2	1:E:81:ARG:HA	1.95	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:116:PRO:OXT	3:C:202:PO4:O2[2_757]	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	114/116 (98%)	110 (96%)	3 (3%)	1 (1%)	21	42
1	B	114/116 (98%)	110 (96%)	3 (3%)	1 (1%)	21	42
1	C	114/116 (98%)	108 (95%)	5 (4%)	1 (1%)	21	42
1	D	114/116 (98%)	102 (90%)	9 (8%)	3 (3%)	7	11
1	E	114/116 (98%)	106 (93%)	5 (4%)	3 (3%)	7	11
1	F	114/116 (98%)	108 (95%)	6 (5%)	0	100	100
2	I	7/10 (70%)	6 (86%)	0	1 (14%)	0	0
2	J	10/10 (100%)	8 (80%)	0	2 (20%)	0	0

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	K	7/10 (70%)	6 (86%)	0	1 (14%)	0	0
All	All	708/726 (98%)	664 (94%)	31 (4%)	13 (2%)	11	21

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	59	ALA
2	J	409[A]	HIS
2	J	409[B]	HIS
2	I	409	HIS
1	E	20	GLN
1	E	80	PHE
1	A	48	ASN
1	D	81	ARG
1	E	22	SER
2	K	409	HIS
1	B	22	SER
1	D	59	ALA
1	D	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	102/102 (100%)	94 (92%)	8 (8%)	16	30
1	B	102/102 (100%)	90 (88%)	12 (12%)	6	12
1	C	102/102 (100%)	95 (93%)	7 (7%)	19	38
1	D	102/102 (100%)	92 (90%)	10 (10%)	10	19
1	E	102/102 (100%)	90 (88%)	12 (12%)	6	12
1	F	102/102 (100%)	96 (94%)	6 (6%)	24	47
2	I	7/7 (100%)	5 (71%)	2 (29%)	0	1
2	J	14/7 (200%)	10 (71%)	4 (29%)	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	K	7/7 (100%)	5 (71%)	2 (29%)	0	1
All	All	640/633 (101%)	577 (90%)	63 (10%)	10	19

All (63) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	ARG
1	A	22	SER
1	A	55	SER
1	A	69	LEU
1	A	82	THR
1	A	88	THR
1	A	94	THR
1	A	95	LYS
1	B	10	ARG
1	B	19	THR
1	B	28	GLU
1	B	52	LYS
1	B	63	THR
1	B	75	LEU
1	B	77	ARG
1	B	82	THR
1	B	83	THR
1	B	90	CYS
1	B	113	LEU
1	B	115	LEU
2	I	408	MET
2	I	409	HIS
1	C	10	ARG
1	C	40	ILE
1	C	48	ASN
1	C	52	LYS
1	C	63	THR
1	C	81	ARG
1	C	101	ILE
1	D	9	ARG
1	D	10	ARG
1	D	13	ILE
1	D	47	SER
1	D	48	ASN
1	D	63	THR
1	D	72	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	75	LEU
1	D	96	ASN
1	D	113	LEU
2	J	408[A]	MET
2	J	408[B]	MET
2	J	409[A]	HIS
2	J	409[B]	HIS
1	E	13	ILE
1	E	20	GLN
1	E	47	SER
1	E	48	ASN
1	E	49	THR
1	E	55	SER
1	E	57	LEU
1	E	75	LEU
1	E	82	THR
1	E	83	THR
1	E	95	LYS
1	E	106	LEU
1	F	16	GLN
1	F	57	LEU
1	F	64	GLN
1	F	82	THR
1	F	100	ILE
1	F	113	LEU
2	K	408	MET
2	K	409	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	66	HIS
1	B	110	GLN
1	C	62	GLN
1	C	64	GLN
1	D	48	ASN
1	D	110	GLN
1	E	23	HIS
1	E	62	GLN
1	E	64	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 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	I	406	2	10,10,11	2.24	2 (20%)	10,12,14	1.86	4 (40%)
2	STA	J	406[A]	2	10,10,11	1.52	1 (10%)	10,12,14	1.64	2 (20%)
2	STA	J	406[B]	2	10,10,11	2.13	1 (10%)	10,12,14	2.11	3 (30%)
2	STA	K	406	2	10,10,11	2.14	1 (10%)	10,12,14	2.03	4 (40%)

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	I	406	2	-	0/11/11/12	0/0/0/0
2	STA	J	406[A]	2	-	0/11/11/12	0/0/0/0
2	STA	J	406[B]	2	-	0/11/11/12	0/0/0/0
2	STA	K	406	2	-	0/11/11/12	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	406	STA	CM-CH	2.00	1.57	1.53
2	J	406[A]	STA	CH-CA	4.11	1.57	1.53
2	J	406[B]	STA	CH-CA	6.14	1.59	1.53
2	K	406	STA	CH-CA	6.19	1.59	1.53
2	I	406	STA	CH-CA	6.38	1.59	1.53

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	406	STA	CH-CM-C	-2.70	109.20	113.27
2	I	406	STA	CH-CM-C	-2.59	109.36	113.27
2	J	406[B]	STA	CH-CM-C	-2.45	109.57	113.27
2	J	406[A]	STA	CH-CM-C	-2.15	110.02	113.27
2	I	406	STA	CD2-CG-CB	2.31	119.70	111.11
2	K	406	STA	CM-CH-CA	2.48	116.46	112.68
2	K	406	STA	CD2-CG-CB	2.62	120.84	111.11
2	J	406[B]	STA	CM-CH-CA	2.87	117.05	112.68
2	I	406	STA	CG-CB-CA	2.96	122.31	115.77
2	I	406	STA	CM-CH-CA	3.26	117.64	112.68
2	J	406[A]	STA	CG-CB-CA	3.50	123.48	115.77
2	K	406	STA	CG-CB-CA	3.89	124.36	115.77
2	J	406[B]	STA	CG-CB-CA	4.77	126.30	115.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	J	406[A]	STA	1	0
2	J	406[B]	STA	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 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	PO4	A	201	-	4,4,4	0.24	0	6,6,6	0.30	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	PO4	C	202	-	4,4,4	0.46	0	6,6,6	0.28	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	PO4	A	201	-	-	0/0/0/0	0/0/0/0
3	PO4	C	202	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	202	PO4	0	1

## 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.

## 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%)	0.07	5 (4%) 39 31	20, 34, 50, 62	0
1	B	116/116 (100%)	-0.15	2 (1%) 73 68	21, 34, 51, 59	0
1	C	116/116 (100%)	0.19	9 (7%) 16 11	21, 34, 53, 62	0
1	D	116/116 (100%)	0.25	13 (11%) 7 4	23, 35, 54, 61	0
1	E	116/116 (100%)	0.03	5 (4%) 39 31	20, 34, 51, 63	0
1	F	116/116 (100%)	-0.08	4 (3%) 49 41	20, 33, 50, 59	0
2	I	8/10 (80%)	1.50	2 (25%) 1 0	37, 42, 54, 55	0
2	J	7/10 (70%)	2.44	3 (42%) 0 0	37, 42, 54, 55	0
2	K	8/10 (80%)	1.38	2 (25%) 1 0	37, 42, 54, 55	0
All	All	719/726 (99%)	0.11	45 (6%) 23 17	20, 34, 54, 63	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	409[A]	HIS	7.5
2	I	410	PRO	5.7
2	K	410	PRO	5.7
1	E	48	ASN	5.5
1	D	62	GLN	4.9
1	C	82	THR	4.7
1	D	48	ASN	4.6
1	D	60	GLY	4.4
1	C	48	ASN	4.0
1	C	62	GLN	3.9
1	E	47	SER	3.5
1	F	80	PHE	3.2
2	I	409	HIS	3.2
1	B	80	PHE	3.1
1	F	82	THR	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	55	SER	2.9
1	C	97	ASN	2.9
1	F	48	ASN	2.9
1	D	47	SER	2.9
2	J	410[A]	PRO	2.9
1	D	50	PRO	2.8
1	D	61	GLY	2.8
1	D	23	HIS	2.6
1	A	48	ASN	2.6
1	F	47	SER	2.6
1	C	57	LEU	2.5
1	A	46	SER	2.5
1	A	82	THR	2.5
1	C	80	PHE	2.5
1	A	47	SER	2.4
1	D	80	PHE	2.4
1	D	59	ALA	2.4
1	B	82	THR	2.4
1	C	65	ASP	2.3
1	E	23	HIS	2.3
1	D	57	LEU	2.3
1	A	80	PHE	2.3
1	C	23	HIS	2.3
2	J	408[A]	MET	2.2
1	E	82	THR	2.2
1	D	65	ASP	2.2
1	D	82	THR	2.2
1	E	83	THR	2.2
2	K	409	HIS	2.1
1	C	61	GLY	2.1

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	STA	K	406	11/12	0.93	0.29	-	37,38,39,40	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	STA	J	406[B]	11/12	0.88	0.38	-	37,38,39,40	11
2	STA	I	406	11/12	0.92	0.31	-	37,38,39,40	0
2	STA	J	406[A]	11/12	0.88	0.38	-	37,38,39,40	11

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PO4	A	201	5/5	0.97	0.08	-1.95	48,48,51,51	0
3	PO4	C	202	5/5	0.99	0.07	-	45,46,47,48	5

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

There are no such residues in this entry.