Full wwPDB X-ray Structure Validation Report ⓘ

Dec 13, 2023 – 11:21 am GMT

PDB ID	:	4CZ2
Title	:	Complex of human VARP-ANKRD1 with Rab32-GppCp. Selenomet derivative.
Authors	:	Perez-Dorado, I.; Schaefer, I.B.; McCoy, A.J.; Owen, D.J.; Evans, P.R.
Deposited on	:	2014-04-16
Resolution	:	2.97 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

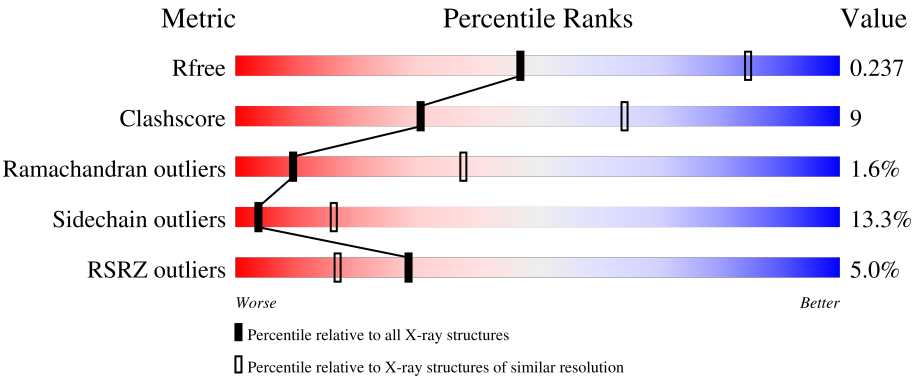
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.97 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.



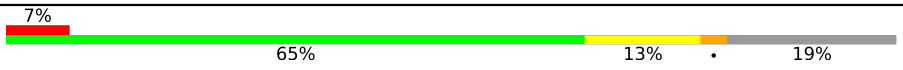
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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 >=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 <=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	230	55% 16% · 25%
1	B	230	3% 59% 14% · 22%
1	C	230	2% 57% 14% 5% 25%
2	D	203	5% 63% 16% · 17%
2	E	203	6% 64% 15% · · 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	203	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8249 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 RAS-RELATED PROTEIN RAB-32.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	172	Total	C	N	O	S	Se	0	0	0
			1388	892	242	247	2	5			
1	B	179	Total	C	N	O	S	Se	0	0	0
			1441	924	250	260	2	5			
1	C	173	Total	C	N	O	S	Se	0	0	0
			1394	895	243	249	2	5			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP Q13637
A	-3	PRO	-	expression tag	UNP Q13637
A	-2	LEU	-	expression tag	UNP Q13637
A	-1	GLY	-	expression tag	UNP Q13637
A	0	SER	-	expression tag	UNP Q13637
A	85	LEU	GLN	engineered mutation	UNP Q13637
A	100	MSE	VAL	engineered mutation	UNP Q13637
A	153	MSE	GLN	engineered mutation	UNP Q13637
A	158	MSE	VAL	engineered mutation	UNP Q13637
A	192	MSE	ILE	engineered mutation	UNP Q13637
B	-4	GLY	-	expression tag	UNP Q13637
B	-3	PRO	-	expression tag	UNP Q13637
B	-2	LEU	-	expression tag	UNP Q13637
B	-1	GLY	-	expression tag	UNP Q13637
B	0	SER	-	expression tag	UNP Q13637
B	85	LEU	GLN	engineered mutation	UNP Q13637
B	100	MSE	VAL	engineered mutation	UNP Q13637
B	153	MSE	GLN	engineered mutation	UNP Q13637
B	158	MSE	VAL	engineered mutation	UNP Q13637
B	192	MSE	ILE	engineered mutation	UNP Q13637
C	-4	GLY	-	expression tag	UNP Q13637
C	-3	PRO	-	expression tag	UNP Q13637
C	-2	LEU	-	expression tag	UNP Q13637

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	expression tag	UNP Q13637
C	0	SER	-	expression tag	UNP Q13637
C	85	LEU	GLN	engineered mutation	UNP Q13637
C	100	MSE	VAL	engineered mutation	UNP Q13637
C	153	MSE	GLN	engineered mutation	UNP Q13637
C	158	MSE	VAL	engineered mutation	UNP Q13637
C	192	MSE	ILE	engineered mutation	UNP Q13637

- Molecule 2 is a protein called ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	168	Total	C	N	O	S	0	0	0
			1283	804	225	246	8			
2	E	166	Total	C	N	O	S	0	0	0
			1269	795	223	243	8			
2	F	165	Total	C	N	O	S	0	0	0
			1262	790	222	242	8			

There are 36 discrepancies between the modelled and reference sequences:

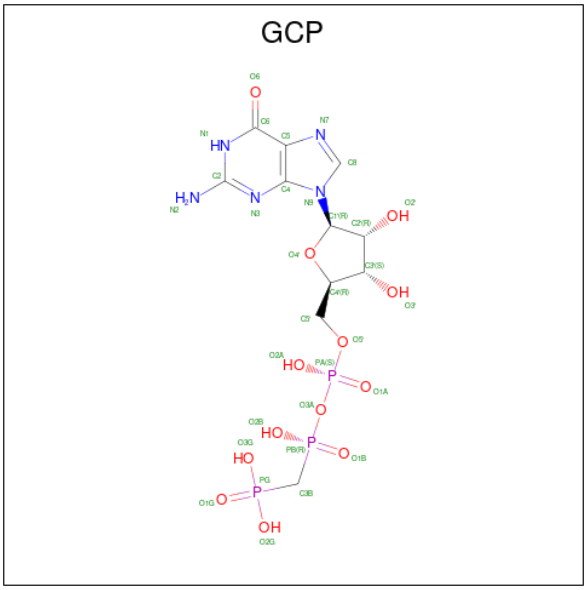
Chain	Residue	Modelled	Actual	Comment	Reference
D	444	GLY	-	expression tag	UNP Q96NW4
D	445	PRO	-	expression tag	UNP Q96NW4
D	446	LEU	-	expression tag	UNP Q96NW4
D	447	GLY	-	expression tag	UNP Q96NW4
D	448	SER	-	expression tag	UNP Q96NW4
D	449	MET	-	expression tag	UNP Q96NW4
D	641	HIS	-	expression tag	UNP Q96NW4
D	642	HIS	-	expression tag	UNP Q96NW4
D	643	HIS	-	expression tag	UNP Q96NW4
D	644	HIS	-	expression tag	UNP Q96NW4
D	645	HIS	-	expression tag	UNP Q96NW4
D	646	HIS	-	expression tag	UNP Q96NW4
E	444	GLY	-	expression tag	UNP Q96NW4
E	445	PRO	-	expression tag	UNP Q96NW4
E	446	LEU	-	expression tag	UNP Q96NW4
E	447	GLY	-	expression tag	UNP Q96NW4
E	448	SER	-	expression tag	UNP Q96NW4
E	449	MET	-	expression tag	UNP Q96NW4
E	641	HIS	-	expression tag	UNP Q96NW4
E	642	HIS	-	expression tag	UNP Q96NW4
E	643	HIS	-	expression tag	UNP Q96NW4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	644	HIS	-	expression tag	UNP Q96NW4
E	645	HIS	-	expression tag	UNP Q96NW4
E	646	HIS	-	expression tag	UNP Q96NW4
F	444	GLY	-	expression tag	UNP Q96NW4
F	445	PRO	-	expression tag	UNP Q96NW4
F	446	LEU	-	expression tag	UNP Q96NW4
F	447	GLY	-	expression tag	UNP Q96NW4
F	448	SER	-	expression tag	UNP Q96NW4
F	449	MET	-	expression tag	UNP Q96NW4
F	641	HIS	-	expression tag	UNP Q96NW4
F	642	HIS	-	expression tag	UNP Q96NW4
F	643	HIS	-	expression tag	UNP Q96NW4
F	644	HIS	-	expression tag	UNP Q96NW4
F	645	HIS	-	expression tag	UNP Q96NW4
F	646	HIS	-	expression tag	UNP Q96NW4

- Molecule 3 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: GCP) (formula: $C_{11}H_{18}N_5O_{13}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
3	B	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
3	C	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Mg 1 1	0	0
4	B	1	Total Mg 1 1	0	0
4	C	1	Total Mg 1 1	0	0

- Molecule 5 is water.

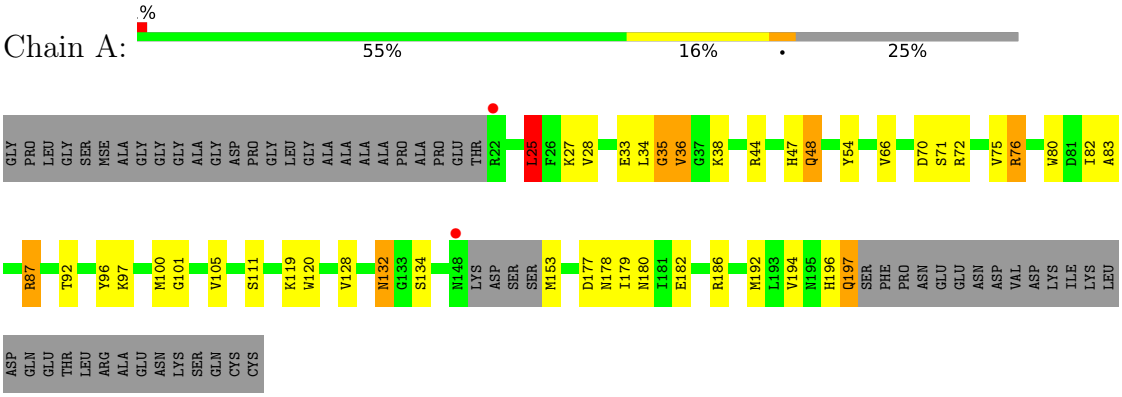
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	24	Total O 24 24	0	0
5	B	24	Total O 24 24	0	0
5	C	16	Total O 16 16	0	0
5	D	18	Total O 18 18	0	0
5	E	17	Total O 17 17	0	0
5	F	14	Total O 14 14	0	0

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.

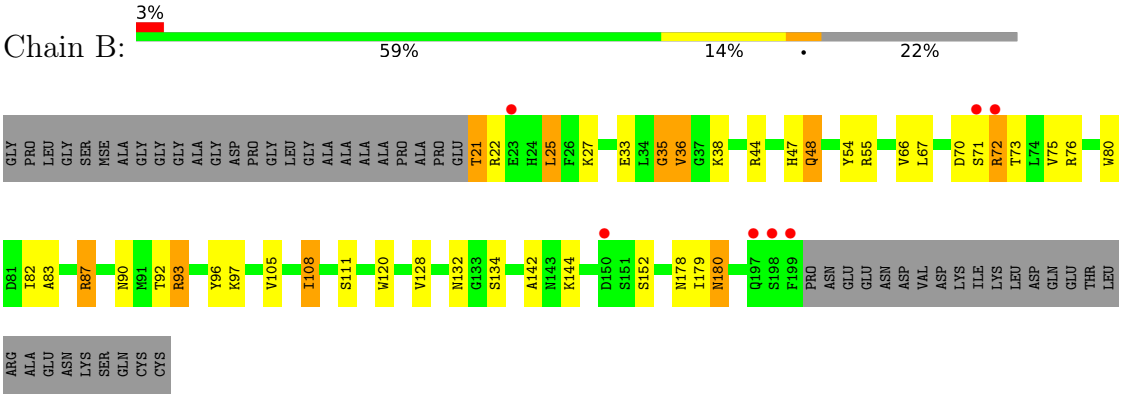
• Molecule 1: RAS-RELATED PROTEIN RAB-32

Chain A:



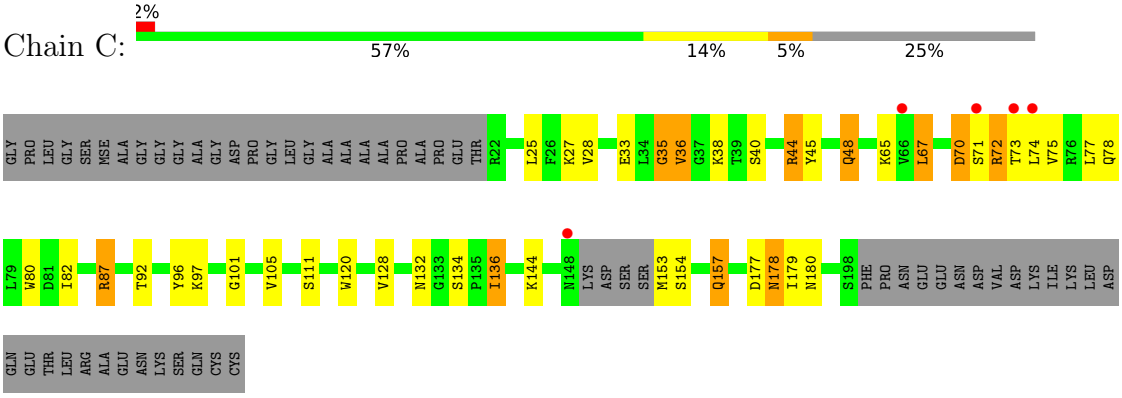
• Molecule 1: RAS-RELATED PROTEIN RAB-32

Chain B:



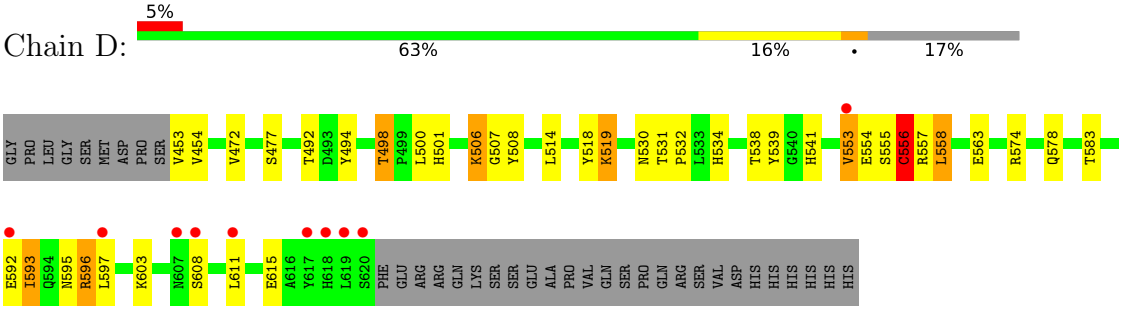
• Molecule 1: RAS-RELATED PROTEIN RAB-32

Chain C:



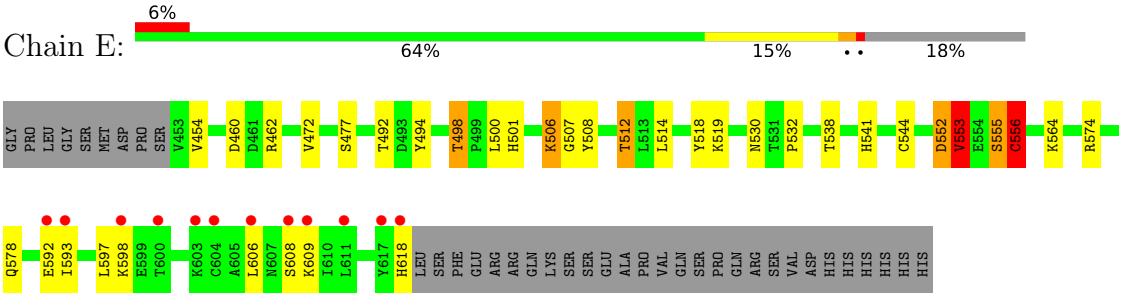
• Molecule 2: ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 27

Chain D:



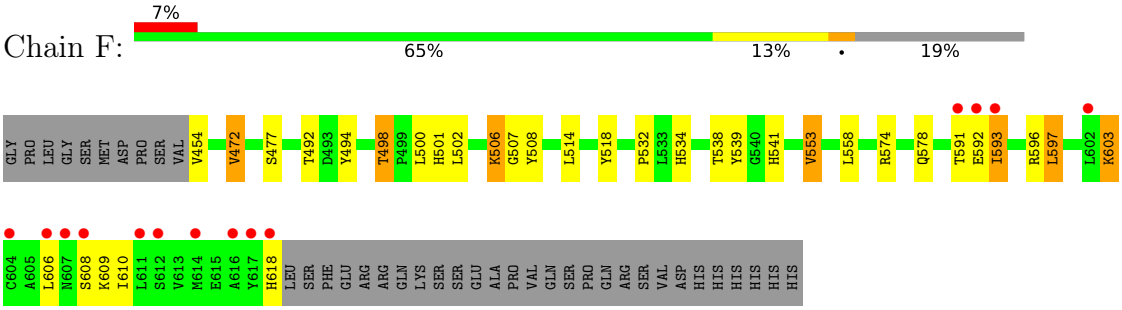
• Molecule 2: ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 27

Chain E:



• Molecule 2: ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 27

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	113.05Å 122.49Å 131.58Å 90.00° 113.25° 90.00°	Depositor
Resolution (Å)	120.89 – 2.97 57.05 – 2.97	Depositor EDS
% Data completeness (in resolution range)	97.1 (120.89-2.97) 97.1 (57.05-2.97)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.195 , 0.232 0.203 , 0.237	Depositor DCC
R_{free} test set	1679 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	45.7	Xtriage
Anisotropy	0.746	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 54.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.036 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8249	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.05% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: MG, GCP

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	0.59	0/1415	0.78	1/1901 (0.1%)
1	B	0.57	0/1469	0.76	1/1973 (0.1%)
1	C	0.54	0/1421	0.77	0/1909
2	D	0.57	0/1308	0.77	0/1780
2	E	0.58	0/1294	0.75	0/1761
2	F	0.56	0/1287	0.75	0/1751
All	All	0.57	0/8194	0.76	2/11075 (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	C	0	1
2	D	0	1
2	E	0	1
2	F	0	1
All	All	0	4

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	25	LEU	CB-CG-CD1	5.40	120.18	111.00
1	B	55	ARG	NE-CZ-NH1	5.17	122.89	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	70	ASP	Peptide
2	D	593	ILE	Peptide
2	E	555	SER	Peptide
2	F	593	ILE	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	1388	0	1382	34	0
1	B	1441	0	1431	27	0
1	C	1394	0	1387	24	0
2	D	1283	0	1268	34	0
2	E	1269	0	1252	28	0
2	F	1262	0	1243	14	0
3	A	32	0	14	3	0
3	B	32	0	14	4	0
3	C	32	0	14	3	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	24	0	0	0	0
5	B	24	0	0	2	0
5	C	16	0	0	0	0
5	D	18	0	0	2	0
5	E	17	0	0	2	0
5	F	14	0	0	1	0
All	All	8249	0	8005	149	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

All (149) 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:25:LEU:HG	2:D:553:VAL:HG12	1.28	1.13
1:A:25:LEU:HG	2:D:553:VAL:CG1	1.88	1.04

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:25:LEU:HD13	1:A:76:ARG:HB3	1.42	0.99
1:A:101:GLY:N	1:A:192:MSE:HE2	1.86	0.90
2:D:507:GLY:HA3	2:D:541:HIS:ND1	1.86	0.90
2:D:453:VAL:HG12	2:D:454:VAL:H	1.38	0.89
1:A:101:GLY:CA	1:A:192:MSE:HE2	2.11	0.80
1:B:25:LEU:CD2	2:E:553:VAL:HG12	2.13	0.78
1:A:25:LEU:CG	2:D:553:VAL:HG12	2.12	0.77
1:A:28:VAL:HG22	1:A:192:MSE:HE1	1.70	0.71
2:F:507:GLY:HA3	2:F:541:HIS:CD2	2.24	0.71
2:E:507:GLY:HA3	2:E:541:HIS:CD2	2.26	0.71
1:B:25:LEU:HD23	2:E:553:VAL:HG12	1.72	0.71
1:A:101:GLY:HA3	1:A:192:MSE:HE2	1.72	0.70
1:A:35:GLY:N	3:A:1198:GCP:H3B1	2.07	0.69
2:E:498:THR:HB	2:E:501:HIS:HD2	1.57	0.69
2:E:555:SER:HB3	2:E:556:CYS:O	1.93	0.69
2:E:506:LYS:HB3	2:E:508:TYR:CE1	2.29	0.67
2:D:498:THR:HB	2:D:501:HIS:HD2	1.57	0.67
2:D:531:THR:OG1	2:D:534:HIS:HD2	1.78	0.67
1:C:96:TYR:HB3	1:C:128:VAL:HG21	1.76	0.67
2:F:506:LYS:HB3	2:F:508:TYR:CE1	2.30	0.66
2:E:508:TYR:O	2:E:512:THR:HG23	1.97	0.66
1:B:96:TYR:HB3	1:B:128:VAL:HG21	1.78	0.65
2:E:498:THR:HB	2:E:501:HIS:CD2	2.31	0.65
2:D:498:THR:HB	2:D:501:HIS:CD2	2.31	0.65
1:C:35:GLY:N	3:C:1198:GCP:H3B1	2.13	0.64
1:A:96:TYR:HB3	1:A:128:VAL:HG21	1.79	0.64
1:B:48:GLN:OE1	1:B:48:GLN:N	2.31	0.63
2:D:453:VAL:CG1	2:D:454:VAL:H	2.12	0.63
1:A:101:GLY:HA3	1:A:192:MSE:CE	2.28	0.63
2:D:557:ARG:O	2:D:558:LEU:HB2	1.99	0.63
1:B:35:GLY:N	3:B:1198:GCP:H3B1	2.15	0.62
2:E:555:SER:O	2:E:556:CYS:HB3	2.00	0.62
1:B:25:LEU:HD22	2:E:553:VAL:HG12	1.83	0.60
2:D:507:GLY:HA3	2:D:541:HIS:CE1	2.37	0.59
1:C:96:TYR:HB3	1:C:128:VAL:CG2	2.32	0.59
1:C:48:GLN:N	1:C:48:GLN:OE1	2.36	0.59
1:A:44:ARG:HD2	1:A:178:ASN:ND2	2.17	0.59
1:A:70:ASP:HB3	1:A:72:ARG:H	1.68	0.58
1:B:96:TYR:HB3	1:B:128:VAL:CG2	2.34	0.57
1:A:96:TYR:HB3	1:A:128:VAL:CG2	2.34	0.57
1:C:179:ILE:HG22	1:C:180:ASN:HB2	1.87	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:87:ARG:HB3	1:A:120:TRP:CH2	2.41	0.56
1:A:101:GLY:CA	1:A:192:MSE:CE	2.82	0.56
1:A:179:ILE:HG22	1:A:180:ASN:HB2	1.87	0.55
1:C:154:SER:OG	1:C:157:GLN:HG2	2.05	0.55
1:C:65:LYS:HG3	1:C:67:LEU:HD12	1.87	0.55
1:C:101:GLY:HA2	1:C:136:ILE:HG13	1.89	0.55
2:D:453:VAL:HG12	2:D:454:VAL:N	2.16	0.54
2:D:498:THR:HG22	2:D:501:HIS:H	1.72	0.54
1:B:25:LEU:HD23	2:E:553:VAL:CG1	2.38	0.54
1:B:179:ILE:HG22	1:B:180:ASN:HB2	1.90	0.53
2:E:498:THR:HG22	2:E:501:HIS:H	1.73	0.53
1:C:28:VAL:HG21	1:C:77:LEU:HD22	1.91	0.53
2:E:552:ASP:C	2:E:553:VAL:HG23	2.29	0.53
1:A:177:ASP:HB2	1:A:179:ILE:HD12	1.91	0.52
1:A:100:MSE:C	1:A:192:MSE:HE2	2.28	0.52
2:D:498:THR:H	2:D:501:HIS:CD2	2.28	0.52
2:D:498:THR:CG2	2:D:500:LEU:HB3	2.40	0.52
2:E:498:THR:H	2:E:501:HIS:CD2	2.28	0.52
2:F:508:TYR:N	2:F:508:TYR:CD1	2.75	0.52
2:F:498:THR:HG22	2:F:501:HIS:H	1.75	0.52
2:E:508:TYR:CD1	2:E:508:TYR:N	2.77	0.52
2:D:498:THR:HG23	5:D:2003:HOH:O	2.09	0.51
1:C:70:ASP:HB3	1:C:72:ARG:H	1.74	0.51
1:C:177:ASP:HB2	1:C:179:ILE:HD12	1.92	0.51
2:F:500:LEU:HD23	2:F:532:PRO:HG2	1.93	0.51
1:B:87:ARG:HB3	1:B:120:TRP:CZ2	2.46	0.51
1:B:54:TYR:CE1	3:B:1198:GCP:H3B2	2.46	0.51
2:D:506:LYS:HB3	2:D:508:TYR:CE2	2.45	0.51
2:E:498:THR:CB	2:E:501:HIS:HD2	2.24	0.50
1:B:93:ARG:NH2	5:B:2014:HOH:O	2.36	0.50
2:E:498:THR:CG2	2:E:500:LEU:HB3	2.41	0.49
1:B:70:ASP:OD1	1:B:73:THR:N	2.45	0.49
1:B:87:ARG:HB3	1:B:120:TRP:CH2	2.47	0.49
2:E:498:THR:HG23	5:E:2008:HOH:O	2.12	0.49
1:B:25:LEU:CD2	2:E:553:VAL:CG1	2.87	0.49
2:F:472:VAL:HG13	2:F:502:LEU:HB3	1.94	0.49
2:D:583:THR:HG23	5:D:2015:HOH:O	2.13	0.49
2:F:498:THR:CG2	2:F:500:LEU:HB3	2.43	0.49
2:D:563:GLU:O	2:D:596:ARG:NH1	2.46	0.49
1:B:25:LEU:HD22	2:E:553:VAL:O	2.13	0.48
2:E:501:HIS:HE1	2:E:530:ASN:O	1.96	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:21:THR:HA	1:B:72:ARG:O	2.13	0.48
1:B:90:ASN:HB2	5:B:2012:HOH:O	2.12	0.48
1:B:35:GLY:H	3:B:1198:GCP:H3B1	1.78	0.48
1:C:33:GLU:OE2	1:C:87:ARG:NH2	2.42	0.48
1:A:35:GLY:H	3:A:1198:GCP:H3B1	1.76	0.48
1:C:87:ARG:HB3	1:C:120:TRP:CZ2	2.48	0.48
1:B:108:ILE:HD12	1:B:142:ALA:HB1	1.96	0.47
2:D:501:HIS:HE1	2:D:530:ASN:O	1.98	0.47
2:D:531:THR:OG1	2:D:534:HIS:CD2	2.62	0.47
1:A:87:ARG:HB3	1:A:120:TRP:CZ2	2.50	0.47
1:C:87:ARG:HB3	1:C:120:TRP:CH2	2.49	0.47
2:E:500:LEU:HD23	2:E:532:PRO:HG2	1.96	0.47
1:A:25:LEU:HG	2:D:553:VAL:HG11	1.87	0.47
1:C:70:ASP:HB2	1:C:73:THR:H	1.80	0.46
1:A:33:GLU:OE2	1:A:87:ARG:NH2	2.40	0.46
2:D:500:LEU:HD23	2:D:532:PRO:HG2	1.97	0.46
2:D:498:THR:CB	2:D:501:HIS:HD2	2.26	0.45
2:F:514:LEU:HD11	2:F:518:TYR:CE2	2.51	0.45
1:A:27:LYS:HE3	1:A:80:TRP:CE2	2.51	0.45
1:A:38:LYS:HE3	1:A:82:ILE:O	2.17	0.45
1:A:197:GLN:HE21	1:A:197:GLN:N	2.14	0.45
1:C:40:SER:OG	3:C:1198:GCP:O1A	2.33	0.45
2:E:454:VAL:HA	5:E:2003:HOH:O	2.17	0.44
1:A:54:TYR:CE1	3:A:1198:GCP:H3B2	2.52	0.44
1:C:70:ASP:HB2	1:C:73:THR:N	2.33	0.44
1:B:38:LYS:HE3	1:B:82:ILE:O	2.18	0.44
1:C:78:GLN:NE2	2:F:553:VAL:HG22	2.32	0.44
1:A:48:GLN:N	1:A:48:GLN:CD	2.71	0.44
1:B:27:LYS:HE3	1:B:80:TRP:CE2	2.52	0.43
1:A:27:LYS:HE3	1:A:80:TRP:CZ2	2.54	0.43
1:B:47:HIS:O	1:B:48:GLN:C	2.56	0.43
2:E:598:LYS:HE3	2:E:598:LYS:HB3	1.79	0.43
1:B:36:VAL:HG13	1:B:105:VAL:HG12	2.01	0.43
1:C:44:ARG:HD3	1:C:178:ASN:CG	2.39	0.43
2:F:534:HIS:O	2:F:538:THR:HG23	2.19	0.43
1:C:27:LYS:HE3	1:C:80:TRP:CE2	2.54	0.42
1:B:144:LYS:HG2	3:B:1198:GCP:C6	2.50	0.42
1:C:38:LYS:HE3	1:C:82:ILE:O	2.20	0.42
2:D:507:GLY:CA	2:D:541:HIS:ND1	2.70	0.42
1:A:194:VAL:O	1:A:197:GLN:HG2	2.20	0.42
2:D:453:VAL:HG12	2:D:454:VAL:HG23	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:514:LEU:HD11	2:E:518:TYR:CE2	2.55	0.42
2:F:558:LEU:N	5:F:2014:HOH:O	2.35	0.42
1:A:36:VAL:HG13	1:A:105:VAL:HG12	2.01	0.42
2:D:514:LEU:HD11	2:D:518:TYR:CE2	2.54	0.42
2:D:595:ASN:OD1	2:D:595:ASN:C	2.58	0.42
1:A:132:ASN:N	1:A:132:ASN:ND2	2.67	0.41
2:F:508:TYR:N	2:F:508:TYR:HD1	2.18	0.41
2:F:538:THR:OG1	2:F:539:TYR:CD1	2.71	0.41
2:D:534:HIS:O	2:D:538:THR:HG23	2.19	0.41
2:D:519:LYS:HA	2:D:519:LYS:HE2	2.02	0.41
2:D:555:SER:O	2:D:556:CYS:O	2.39	0.41
2:F:603:LYS:HD2	2:F:603:LYS:O	2.19	0.41
1:B:33:GLU:OE2	1:B:87:ARG:NH2	2.44	0.41
1:C:45:TYR:OH	1:C:67:LEU:HD11	2.21	0.41
1:B:27:LYS:HE3	1:B:80:TRP:CZ2	2.56	0.41
1:A:47:HIS:O	1:A:48:GLN:C	2.59	0.41
2:E:512:THR:HG21	2:E:544:CYS:SG	2.61	0.41
2:D:538:THR:OG1	2:D:539:TYR:CD1	2.72	0.41
2:E:460:ASP:C	2:E:460:ASP:OD1	2.59	0.41
1:C:36:VAL:HG13	1:C:105:VAL:HG12	2.02	0.40
1:C:144:LYS:HG2	3:C:1198:GCP:C6	2.51	0.40
1:A:182:GLU:O	1:A:186:ARG:HG3	2.22	0.40
2:D:611:LEU:O	2:D:615:GLU:HG3	2.21	0.40
2:D:508:TYR:CE1	2:E:508:TYR:CE1	3.10	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	168/230 (73%)	163 (97%)	3 (2%)	2 (1%)	13 45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	177/230 (77%)	166 (94%)	8 (4%)	3 (2%)	9 36
1	C	169/230 (74%)	161 (95%)	6 (4%)	2 (1%)	13 45
2	D	166/203 (82%)	153 (92%)	10 (6%)	3 (2%)	8 35
2	E	164/203 (81%)	150 (92%)	10 (6%)	4 (2%)	6 27
2	F	163/203 (80%)	151 (93%)	10 (6%)	2 (1%)	13 45
All	All	1007/1299 (78%)	944 (94%)	47 (5%)	16 (2%)	9 38

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	553	VAL
2	D	556	CYS
2	D	593	ILE
2	E	556	CYS
2	E	593	ILE
2	F	593	ILE
1	C	35	GLY
2	D	597	LEU
2	E	597	LEU
2	F	597	LEU
1	A	35	GLY
1	B	35	GLY
1	B	48	GLN
1	B	83	ALA
1	C	48	GLN
1	A	83	ALA

5.3.2 Protein sidechains (i)

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	150/186 (81%)	132 (88%)	18 (12%)	5 20
1	B	157/186 (84%)	135 (86%)	22 (14%)	3 15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	C	151/186 (81%)	133 (88%)	18 (12%)	5 21
2	D	140/172 (81%)	123 (88%)	17 (12%)	5 20
2	E	138/172 (80%)	117 (85%)	21 (15%)	3 12
2	F	137/172 (80%)	117 (85%)	20 (15%)	3 13
All	All	873/1074 (81%)	757 (87%)	116 (13%)	4 16

All (116) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LEU
1	A	34	LEU
1	A	36	VAL
1	A	48	GLN
1	A	66	VAL
1	A	71	SER
1	A	75	VAL
1	A	76	ARG
1	A	87	ARG
1	A	92	THR
1	A	97	LYS
1	A	111	SER
1	A	119	LYS
1	A	132	ASN
1	A	134	SER
1	A	153	MSE
1	A	196	HIS
1	A	197	GLN
1	B	21	THR
1	B	22	ARG
1	B	25	LEU
1	B	36	VAL
1	B	44	ARG
1	B	66	VAL
1	B	67	LEU
1	B	71	SER
1	B	72	ARG
1	B	75	VAL
1	B	76	ARG
1	B	87	ARG
1	B	92	THR
1	B	93	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	97	LYS
1	B	108	ILE
1	B	111	SER
1	B	132	ASN
1	B	134	SER
1	B	152	SER
1	B	178	ASN
1	B	180	ASN
1	C	25	LEU
1	C	36	VAL
1	C	44	ARG
1	C	67	LEU
1	C	71	SER
1	C	72	ARG
1	C	74	LEU
1	C	75	VAL
1	C	87	ARG
1	C	92	THR
1	C	97	LYS
1	C	111	SER
1	C	132	ASN
1	C	134	SER
1	C	136	ILE
1	C	153	MSE
1	C	157	GLN
1	C	178	ASN
2	D	472	VAL
2	D	477	SER
2	D	492	THR
2	D	494	TYR
2	D	498	THR
2	D	506	LYS
2	D	519	LYS
2	D	553	VAL
2	D	554	GLU
2	D	556	CYS
2	D	558	LEU
2	D	574	ARG
2	D	578	GLN
2	D	592	GLU
2	D	596	ARG
2	D	603	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	608	SER
2	E	462	ARG
2	E	472	VAL
2	E	477	SER
2	E	492	THR
2	E	494	TYR
2	E	498	THR
2	E	506	LYS
2	E	512	THR
2	E	519	LYS
2	E	538	THR
2	E	552	ASP
2	E	553	VAL
2	E	556	CYS
2	E	564	LYS
2	E	574	ARG
2	E	578	GLN
2	E	592	GLU
2	E	606	LEU
2	E	608	SER
2	E	609	LYS
2	E	618	HIS
2	F	454	VAL
2	F	472	VAL
2	F	477	SER
2	F	492	THR
2	F	494	TYR
2	F	498	THR
2	F	506	LYS
2	F	553	VAL
2	F	574	ARG
2	F	578	GLN
2	F	591	THR
2	F	592	GLU
2	F	596	ARG
2	F	597	LEU
2	F	603	LYS
2	F	606	LEU
2	F	608	SER
2	F	609	LYS
2	F	610	ILE
2	F	618	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	52	GLN
1	A	68	ASN
1	A	132	ASN
1	A	178	ASN
1	A	197	GLN
1	B	178	ASN
1	C	78	GLN
1	C	132	ASN
1	C	178	ASN
2	D	501	HIS
2	D	534	HIS
2	D	618	HIS
2	E	464	HIS
2	E	501	HIS
2	E	541	HIS
2	F	541	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 for Mogul analysis.

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	GCP	A	1198	4	27,34,34	1.30	3 (11%)	34,54,54	2.04	10 (29%)
3	GCP	C	1198	4	27,34,34	1.38	4 (14%)	34,54,54	1.96	8 (23%)
3	GCP	B	1198	4	27,34,34	1.26	4 (14%)	34,54,54	1.79	8 (23%)

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	GCP	A	1198	4	-	1/15/38/38	0/3/3/3
3	GCP	C	1198	4	-	2/15/38/38	0/3/3/3
3	GCP	B	1198	4	-	0/15/38/38	0/3/3/3

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1198	GCP	C5-C6	3.37	1.47	1.41
3	C	1198	GCP	PB-O3A	3.10	1.61	1.58
3	B	1198	GCP	O4'-C1'	2.96	1.45	1.41
3	C	1198	GCP	PG-O2G	2.86	1.61	1.54
3	A	1198	GCP	PB-O3A	2.86	1.61	1.58
3	B	1198	GCP	C5-C6	2.76	1.46	1.41
3	C	1198	GCP	C5-C6	2.66	1.45	1.41
3	C	1198	GCP	PG-O3G	2.50	1.60	1.54
3	B	1198	GCP	C5-C4	2.34	1.47	1.40
3	A	1198	GCP	C5-C4	2.07	1.46	1.40
3	B	1198	GCP	C4-N3	-2.06	1.32	1.35

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1198	GCP	C5-C6-N1	-5.46	115.97	123.43
3	C	1198	GCP	C5-C6-N1	-4.85	116.80	123.43
3	A	1198	GCP	C5-C6-N1	-4.53	117.23	123.43
3	C	1198	GCP	C2-N1-C6	4.26	122.70	115.93
3	C	1198	GCP	PB-O3A-PA	-4.14	119.44	132.56
3	A	1198	GCP	O2G-PG-C3B	4.04	116.21	106.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	1198	GCP	C4-C5-C6	-3.92	117.06	120.80
3	A	1198	GCP	C2-N3-C4	3.90	119.81	115.36
3	A	1198	GCP	C2-N1-C6	3.84	122.04	115.93
3	B	1198	GCP	C2-N1-C6	3.70	121.80	115.93
3	A	1198	GCP	C3'-C2'-C1'	3.49	106.23	100.98
3	C	1198	GCP	C2-N3-C4	3.49	119.34	115.36
3	B	1198	GCP	PB-O3A-PA	-3.26	122.23	132.56
3	B	1198	GCP	O3G-PG-O1G	-3.21	103.90	112.39
3	A	1198	GCP	O1G-PG-C3B	-3.02	104.74	111.24
3	C	1198	GCP	N3-C2-N1	-2.85	123.42	127.22
3	A	1198	GCP	C4-C5-C6	-2.68	118.23	120.80
3	A	1198	GCP	N3-C2-N1	-2.65	123.68	127.22
3	B	1198	GCP	O3G-PG-O2G	2.58	115.61	108.08
3	A	1198	GCP	O2B-PB-C3B	2.44	116.56	106.58
3	A	1198	GCP	PB-O3A-PA	-2.22	125.51	132.56
3	B	1198	GCP	O3G-PG-C3B	-2.11	101.27	106.40
3	B	1198	GCP	O2B-PB-C3B	2.06	115.01	106.58
3	C	1198	GCP	O2G-PG-C3B	2.05	111.37	106.40
3	C	1198	GCP	N2-C2-N1	2.05	120.43	117.25
3	B	1198	GCP	C2-N3-C4	2.01	117.66	115.36

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1198	GCP	PG-C3B-PB-O1B
3	C	1198	GCP	O4'-C4'-C5'-O5'
3	C	1198	GCP	C3'-C4'-C5'-O5'

There are no ring outliers.

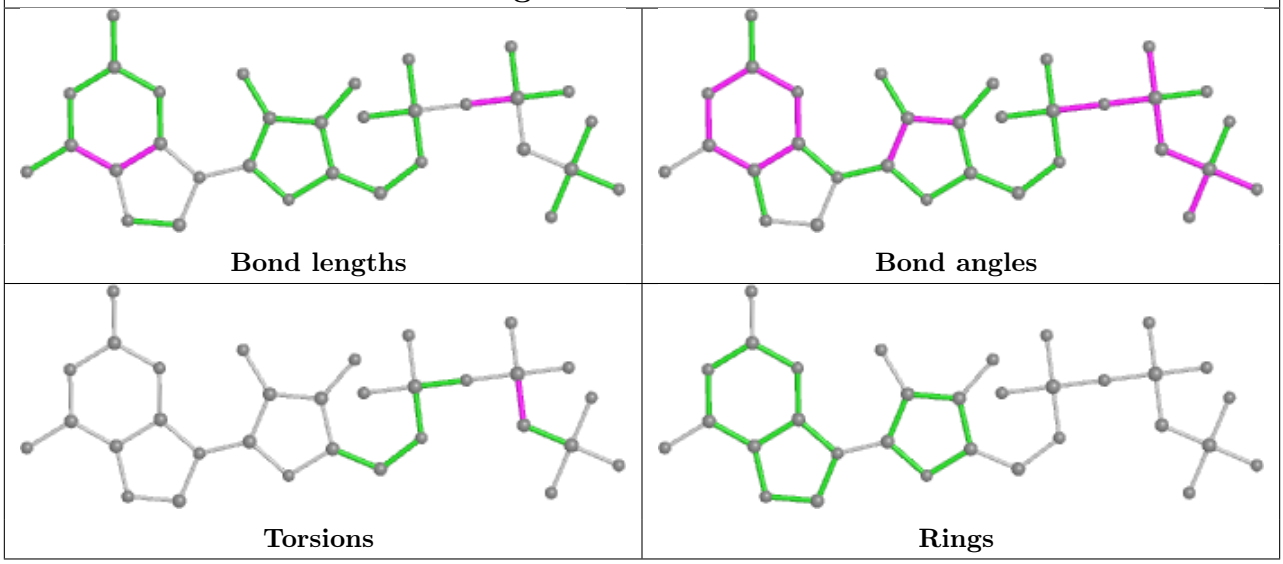
3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1198	GCP	3	0
3	C	1198	GCP	3	0
3	B	1198	GCP	4	0

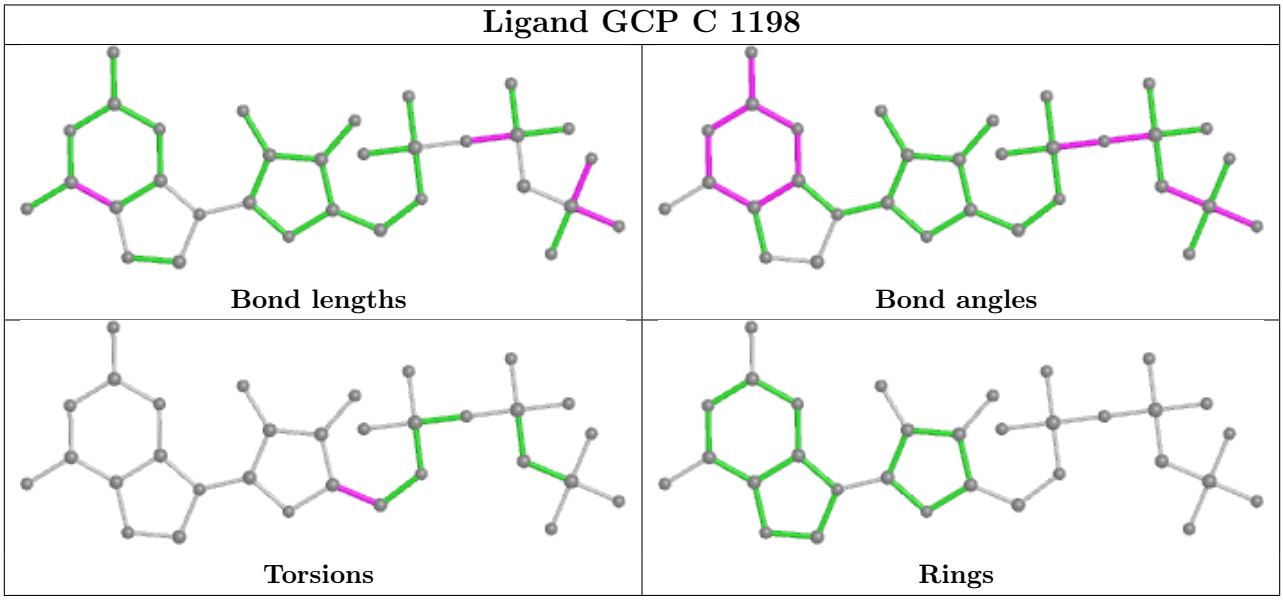
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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.

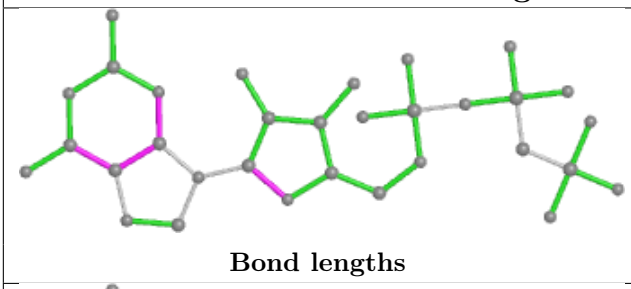
Ligand GCP A 1198



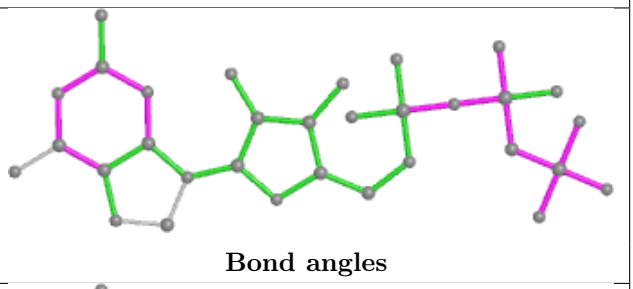
Ligand GCP C 1198



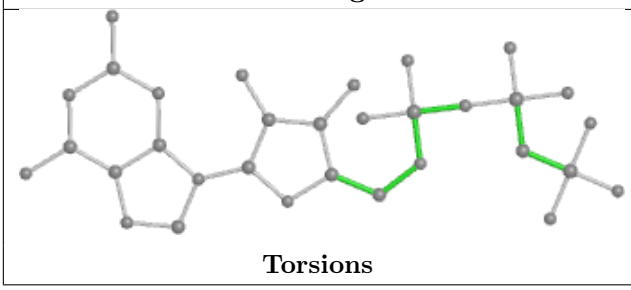
Ligand GCP B 1198



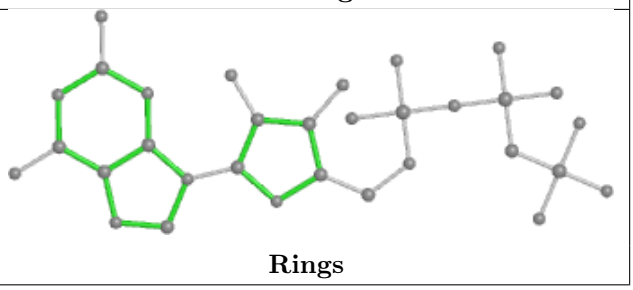
Bond lengths



Bond angles



Torsions



Rings

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^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	167/230 (72%)	-0.11	2 (1%) 79 61	15, 30, 58, 89	0
1	B	174/230 (75%)	0.11	7 (4%) 38 23	21, 37, 81, 105	0
1	C	168/230 (73%)	0.16	5 (2%) 50 31	25, 49, 75, 86	0
2	D	168/203 (82%)	0.14	10 (5%) 21 12	15, 33, 67, 95	0
2	E	166/203 (81%)	0.24	12 (7%) 15 8	18, 39, 77, 105	0
2	F	165/203 (81%)	0.24	14 (8%) 10 5	21, 36, 74, 91	0
All	All	1008/1299 (77%)	0.13	50 (4%) 28 17	15, 37, 75, 105	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	618	HIS	4.8
1	B	199	PHE	4.3
1	B	198	SER	3.7
2	F	592	GLU	3.4
2	E	592	GLU	3.4
2	D	608	SER	3.3
2	E	617	TYR	3.2
2	F	616	ALA	3.2
2	D	619	LEU	3.2
1	C	74	LEU	3.1
2	F	591	THR	2.9
2	F	618	HIS	2.9
2	F	617	TYR	2.9
2	D	553	VAL	2.9
1	B	197	GLN	2.8
2	F	604	CYS	2.8
2	E	608	SER	2.7
2	F	607	ASN	2.7
2	D	618	HIS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	603	LYS	2.6
2	E	593	ILE	2.5
1	C	71	SER	2.5
2	D	620	SER	2.5
2	F	612	SER	2.5
1	B	150	ASP	2.5
2	E	611	LEU	2.5
1	A	22	ARG	2.4
1	A	148	ASN	2.4
1	C	66	VAL	2.4
2	E	598	LYS	2.4
1	C	73	THR	2.4
1	C	148	ASN	2.3
2	F	606	LEU	2.3
2	E	600	THR	2.3
2	F	608	SER	2.3
1	B	23	GLU	2.3
2	E	609	LYS	2.2
2	E	604	CYS	2.2
1	B	71	SER	2.2
2	D	597	LEU	2.2
2	E	606	LEU	2.2
2	D	607	ASN	2.2
2	F	614	MET	2.1
1	B	72	ARG	2.1
2	D	617	TYR	2.1
2	F	593	ILE	2.1
2	D	592	GLU	2.1
2	D	611	LEU	2.0
2	F	602	LEU	2.0
2	F	611	LEU	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

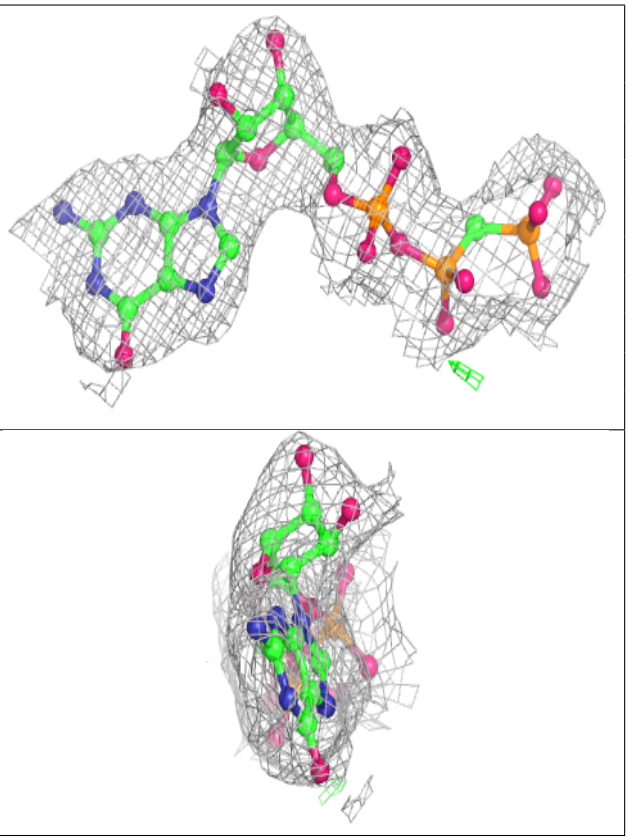
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

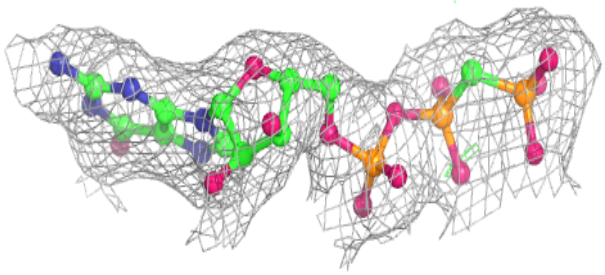
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MG	B	1199	1/1	0.95	0.09	18,18,18,18	0
4	MG	C	1199	1/1	0.96	0.10	26,26,26,26	0
4	MG	A	1199	1/1	0.97	0.10	15,15,15,15	0
3	GCP	A	1198	32/32	0.98	0.15	13,16,20,21	0
3	GCP	B	1198	32/32	0.98	0.15	15,19,27,29	0
3	GCP	C	1198	32/32	0.98	0.14	23,33,42,45	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.

Electron density around GCP A 1198:

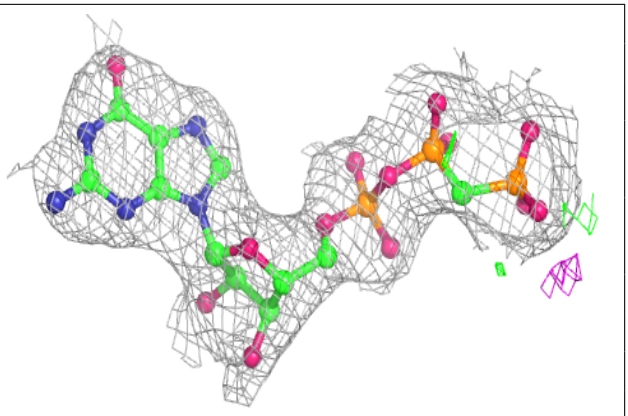
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

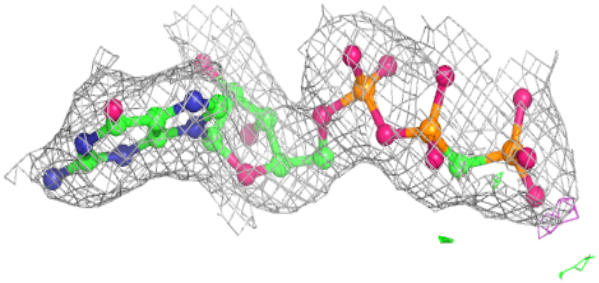


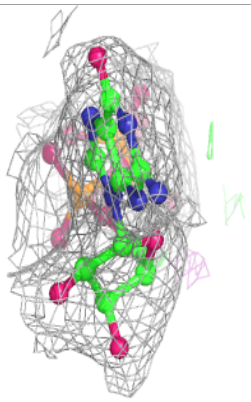


Electron density around GCP B 1198:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

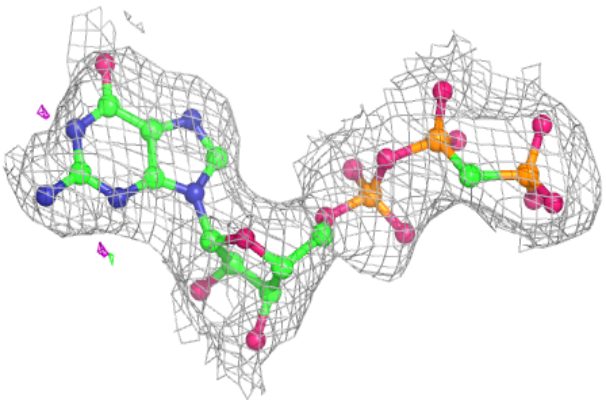


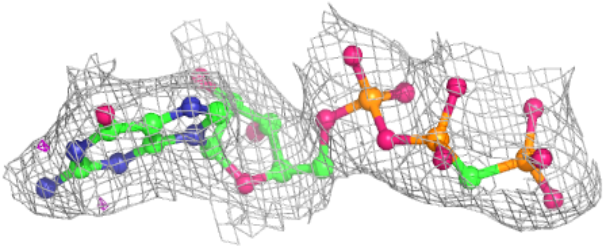


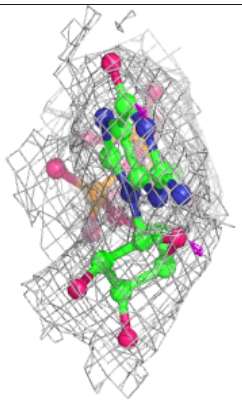


Electron density around GCP C 1198:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.