wwPDB X-ray Structure Validation Summary Report (i)

Nov 16, 2023 – 10:26 AM JST

PDB ID	:	6LKN
Title	:	Crystal structure of ATP11C-CDC50A in PtdSer-bound E2P state
Authors	:	Abe, K.; Irie, K.; Nakanishi, H.; Hasegawa, K.
Deposited on	:	2019-12-19
Resolution	:	3.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

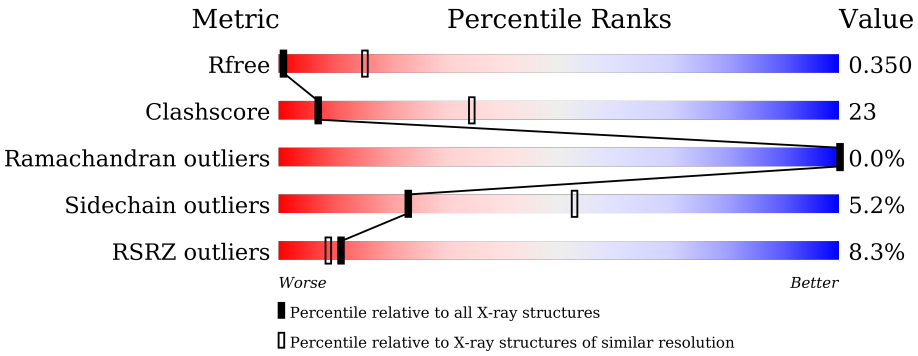
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



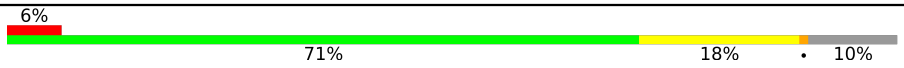
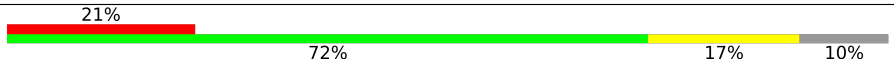
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1002 (4.14-3.66)
Clashscore	141614	1004 (4.12-3.68)
Ramachandran outliers	138981	1021 (4.14-3.66)
Sidechain outliers	138945	1014 (4.14-3.66)
RSRZ outliers	127900	1275 (4.20-3.60)

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 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	1129	4% 56% 35% · 6%
1	E	1129	8% 57% 34% · 6%
1	I	1129	9% 65% 27% · 6%
1	M	1129	11% 65% 27% · 6%
2	C	361	59% 29% · 10%
2	F	361	% 58% 29% · 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	361	
2	N	361	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	BFD	E	409	-	-	X	-
5	NAG	F	405	-	-	X	-
5	NAG	J	404	-	-	X	-
5	NAG	J	405	-	-	X	-
5	NAG	N	401	-	-	-	X
5	NAG	N	405	-	-	X	-
6	AH2	N	403	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 45252 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 Phospholipid-transporting ATPase IG.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	1064	Total	Be	C	F	N	O	S	0	0	0
			8601	1	5556	3	1401	1594	46			
1	E	1064	Total	Be	C	F	N	O	S	0	0	0
			8601	1	5556	3	1401	1594	46			
1	I	1064	Total	Be	C	F	N	O	S	0	0	0
			8601	1	5556	3	1401	1594	46			
1	M	1064	Total	Be	C	F	N	O	S	0	0	0
			8601	1	5556	3	1401	1594	46			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q8NB49
A	2	PHE	-	expression tag	UNP Q8NB49
A	3	ARG	-	expression tag	UNP Q8NB49
A	4	ARG	-	expression tag	UNP Q8NB49
A	5	SER	-	expression tag	UNP Q8NB49
A	6	LEU	-	expression tag	UNP Q8NB49
A	7	ASN	-	expression tag	UNP Q8NB49
A	8	ARG	-	expression tag	UNP Q8NB49
A	9	PHE	-	expression tag	UNP Q8NB49
A	111	TRP	CYS	engineered mutation	UNP Q8NB49
E	1	MET	-	initiating methionine	UNP Q8NB49
E	2	PHE	-	expression tag	UNP Q8NB49
E	3	ARG	-	expression tag	UNP Q8NB49
E	4	ARG	-	expression tag	UNP Q8NB49
E	5	SER	-	expression tag	UNP Q8NB49
E	6	LEU	-	expression tag	UNP Q8NB49
E	7	ASN	-	expression tag	UNP Q8NB49
E	8	ARG	-	expression tag	UNP Q8NB49
E	9	PHE	-	expression tag	UNP Q8NB49
E	111	TRP	CYS	engineered mutation	UNP Q8NB49
I	1	MET	-	initiating methionine	UNP Q8NB49

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	2	PHE	-	expression tag	UNP Q8NB49
I	3	ARG	-	expression tag	UNP Q8NB49
I	4	ARG	-	expression tag	UNP Q8NB49
I	5	SER	-	expression tag	UNP Q8NB49
I	6	LEU	-	expression tag	UNP Q8NB49
I	7	ASN	-	expression tag	UNP Q8NB49
I	8	ARG	-	expression tag	UNP Q8NB49
I	9	PHE	-	expression tag	UNP Q8NB49
I	111	TRP	CYS	engineered mutation	UNP Q8NB49
M	1	MET	-	initiating methionine	UNP Q8NB49
M	2	PHE	-	expression tag	UNP Q8NB49
M	3	ARG	-	expression tag	UNP Q8NB49
M	4	ARG	-	expression tag	UNP Q8NB49
M	5	SER	-	expression tag	UNP Q8NB49
M	6	LEU	-	expression tag	UNP Q8NB49
M	7	ASN	-	expression tag	UNP Q8NB49
M	8	ARG	-	expression tag	UNP Q8NB49
M	9	PHE	-	expression tag	UNP Q8NB49
M	111	TRP	CYS	engineered mutation	UNP Q8NB49

- Molecule 2 is a protein called Cell cycle control protein 50A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	324	Total	C	N	O	S	0	0	0
			2613	1696	439	466	12			
2	F	324	Total	C	N	O	S	0	0	0
			2613	1696	439	466	12			
2	J	324	Total	C	N	O	S	0	0	0
			2613	1696	439	466	12			
2	N	324	Total	C	N	O	S	0	0	0
			2613	1696	439	466	12			

There are 8 discrepancies between the modelled and reference sequences:

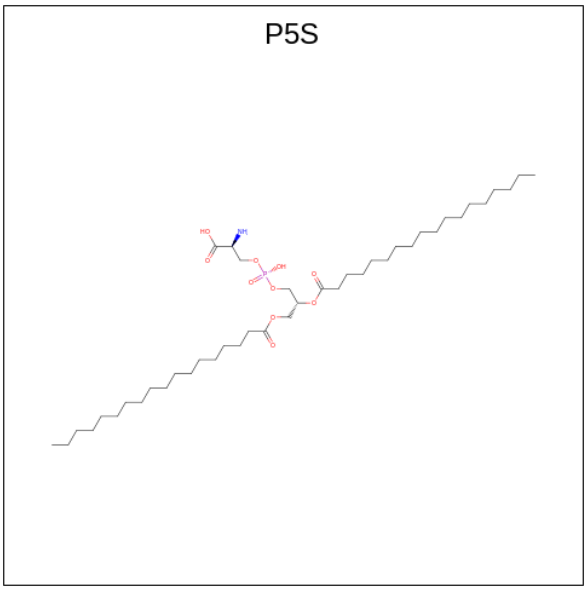
Chain	Residue	Modelled	Actual	Comment	Reference
C	190	GLN	ASN	engineered mutation	UNP Q9NV96
C	292	TRP	SER	engineered mutation	UNP Q9NV96
F	190	GLN	ASN	engineered mutation	UNP Q9NV96
F	292	TRP	SER	engineered mutation	UNP Q9NV96
J	190	GLN	ASN	engineered mutation	UNP Q9NV96
J	292	TRP	SER	engineered mutation	UNP Q9NV96
N	190	GLN	ASN	engineered mutation	UNP Q9NV96

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	292	TRP	SER	engineered mutation	UNP Q9NV96

- Molecule 3 is O-[(R)-{[(2R)-2,3-bis(octadecanoyloxy)propyl]oxy}(hydroxy)phosphoryl]-L-serine (three-letter code: P5S) (formula: $C_{42}H_{82}NO_{10}P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O P 11 3 1 6 1	0	0
3	C	1	Total C N O P 54 42 1 10 1	0	0
3	E	1	Total C N O P 11 3 1 6 1	0	0
3	E	1	Total C N O P 11 3 1 6 1	0	0
3	I	1	Total C N O P 11 3 1 6 1	0	0
3	I	1	Total C N O P 11 3 1 6 1	0	0
3	M	1	Total C N O P 11 3 1 6 1	0	0

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

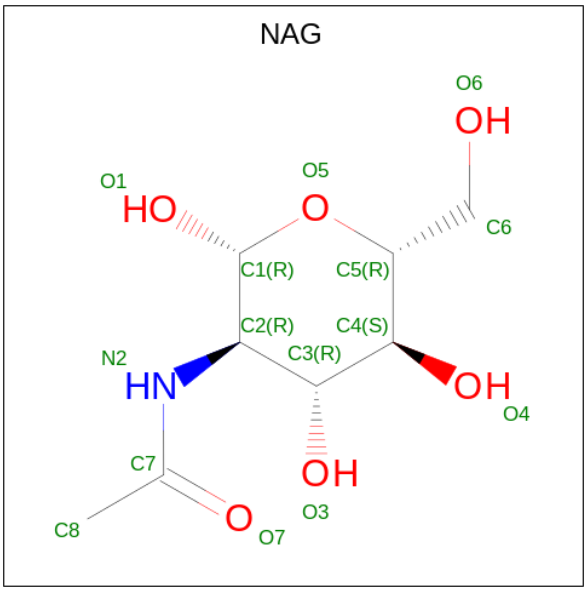
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Mg 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	1	Total Mg 1 1	0	0
4	I	1	Total Mg 1 1	0	0
4	M	1	Total Mg 1 1	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



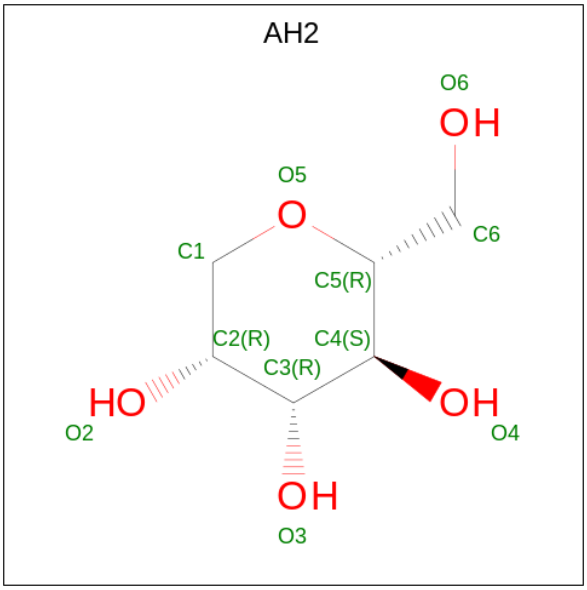
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total C N O 14 8 1 5	0	0
5	C	1	Total C N O 14 8 1 5	0	0
5	C	1	Total C N O 14 8 1 5	0	0
5	C	1	Total C N O 13 8 1 4	0	0
5	F	1	Total C N O 14 8 1 5	0	0
5	F	1	Total C N O 14 8 1 5	0	0
5	F	1	Total C N O 14 8 1 5	0	0
5	F	1	Total C N O 13 8 1 4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	J	1	Total C N O 14 8 1 5	0	0
5	J	1	Total C N O 14 8 1 5	0	0
5	J	1	Total C N O 14 8 1 5	0	0
5	J	1	Total C N O 13 8 1 4	0	0
5	N	1	Total C N O 14 8 1 5	0	0
5	N	1	Total C N O 14 8 1 5	0	0
5	N	1	Total C N O 14 8 1 5	0	0
5	N	1	Total C N O 13 8 1 4	0	0

- Molecule 6 is 1-deoxy-alpha-D-mannopyranose (three-letter code: AH2) (formula: $C_6H_{12}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total C O 11 6 5	0	0
6	F	1	Total C O 11 6 5	0	0
6	J	1	Total C O 11 6 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	N	1	Total C O 11 6 5	0	0

- Molecule 7 is water.

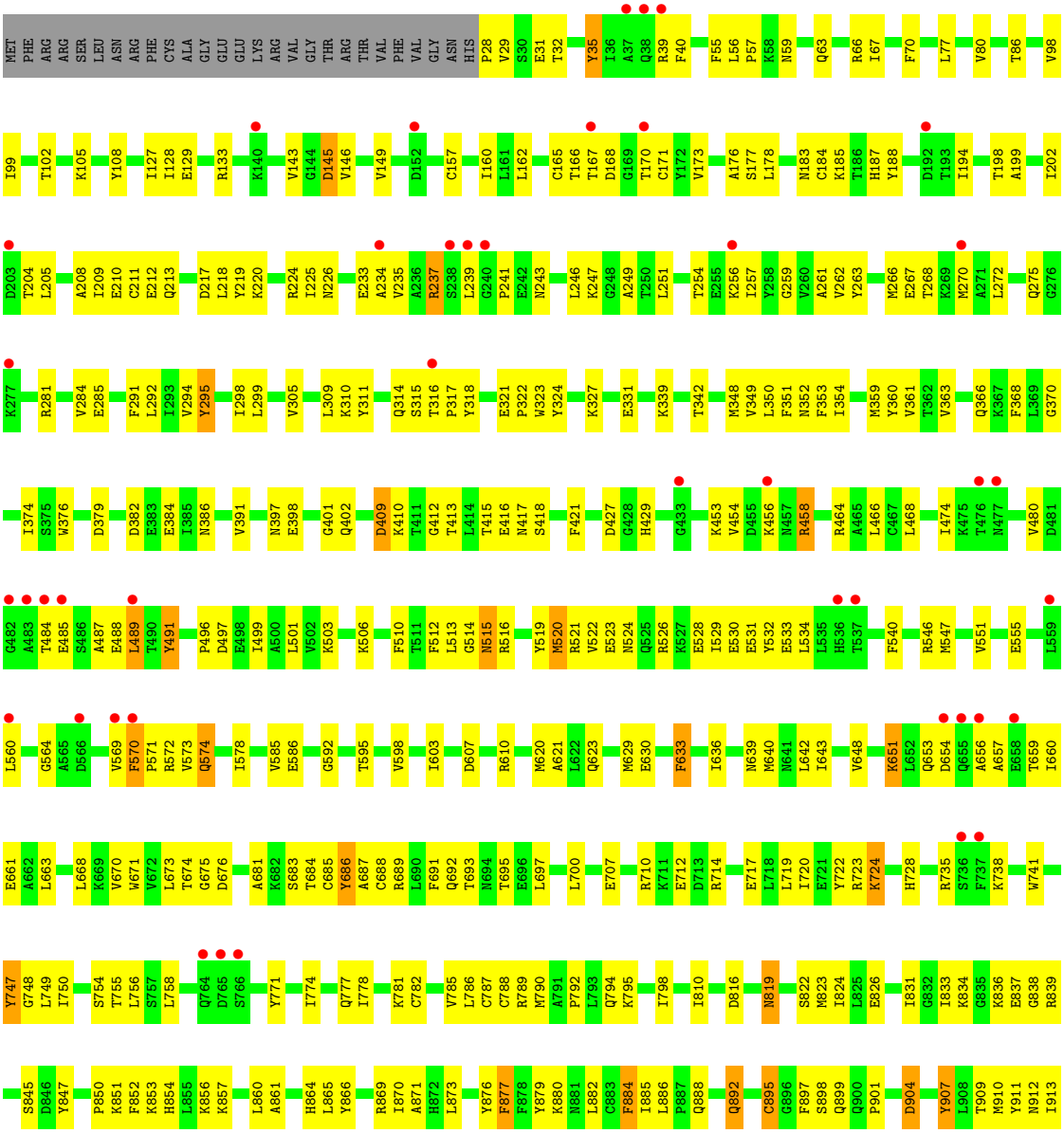
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	2	Total O 2 2	0	0
7	E	2	Total O 2 2	0	0
7	I	2	Total O 2 2	0	0
7	M	2	Total O 2 2	0	0

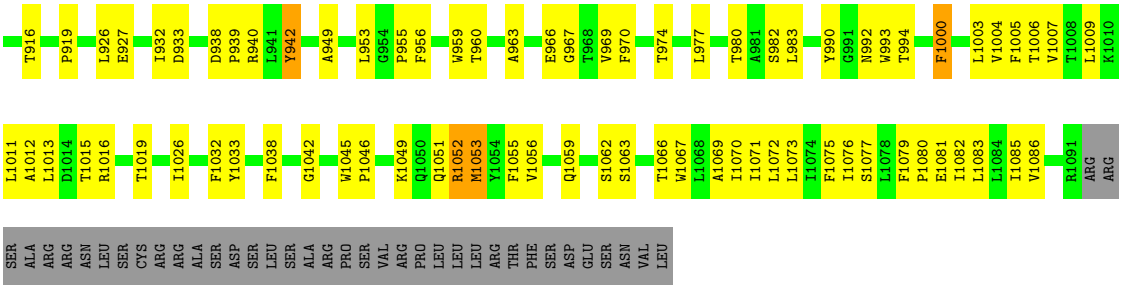
3 Residue-property plots (i)

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: Phospholipid-transporting ATPase IG

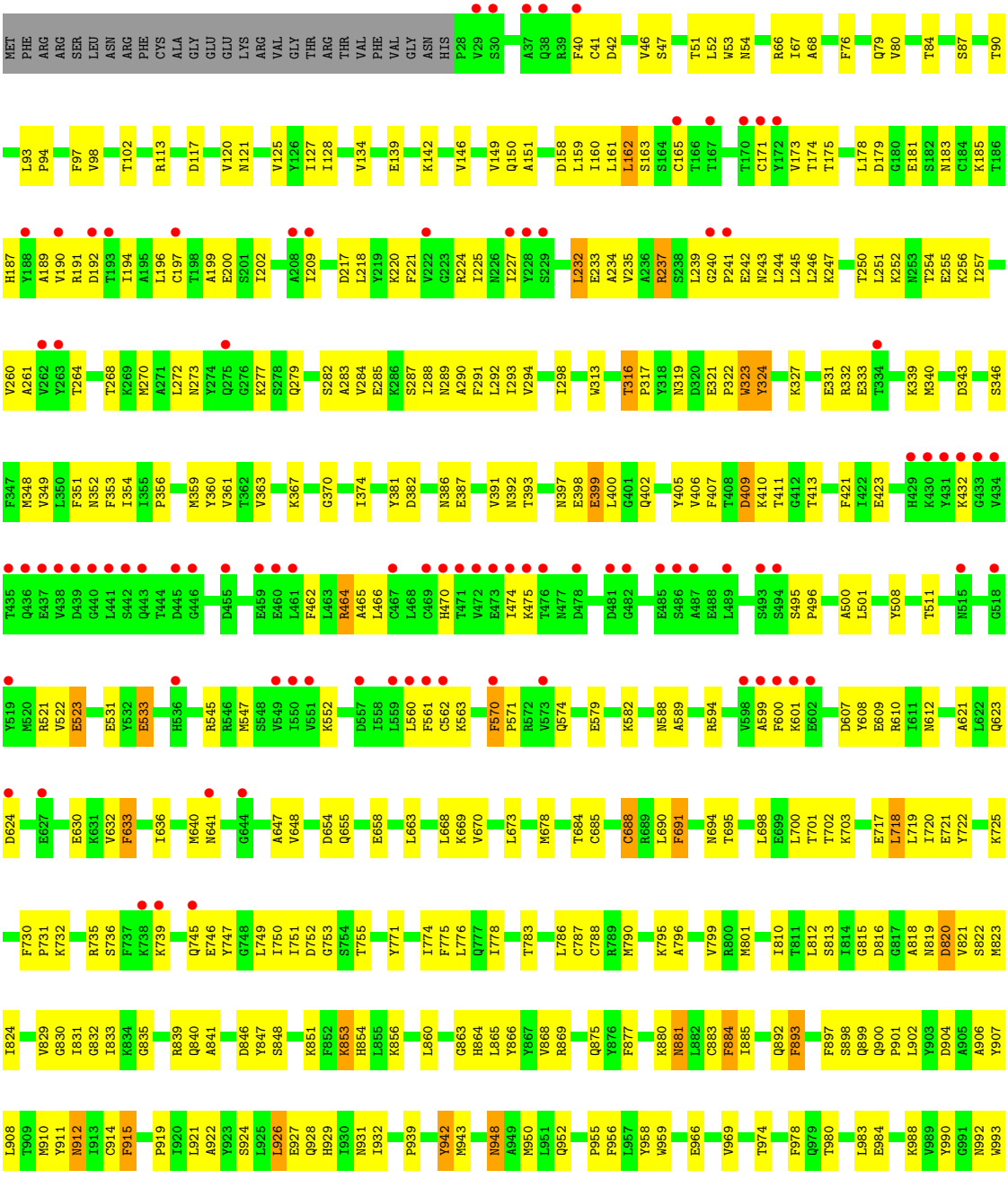
Chain A:

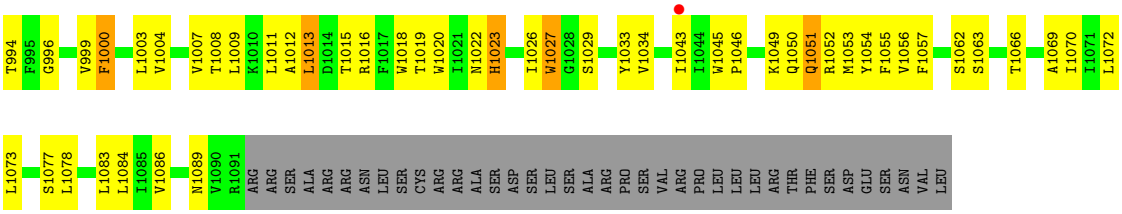




• Molecule 1: Phospholipid-transporting ATPase IG

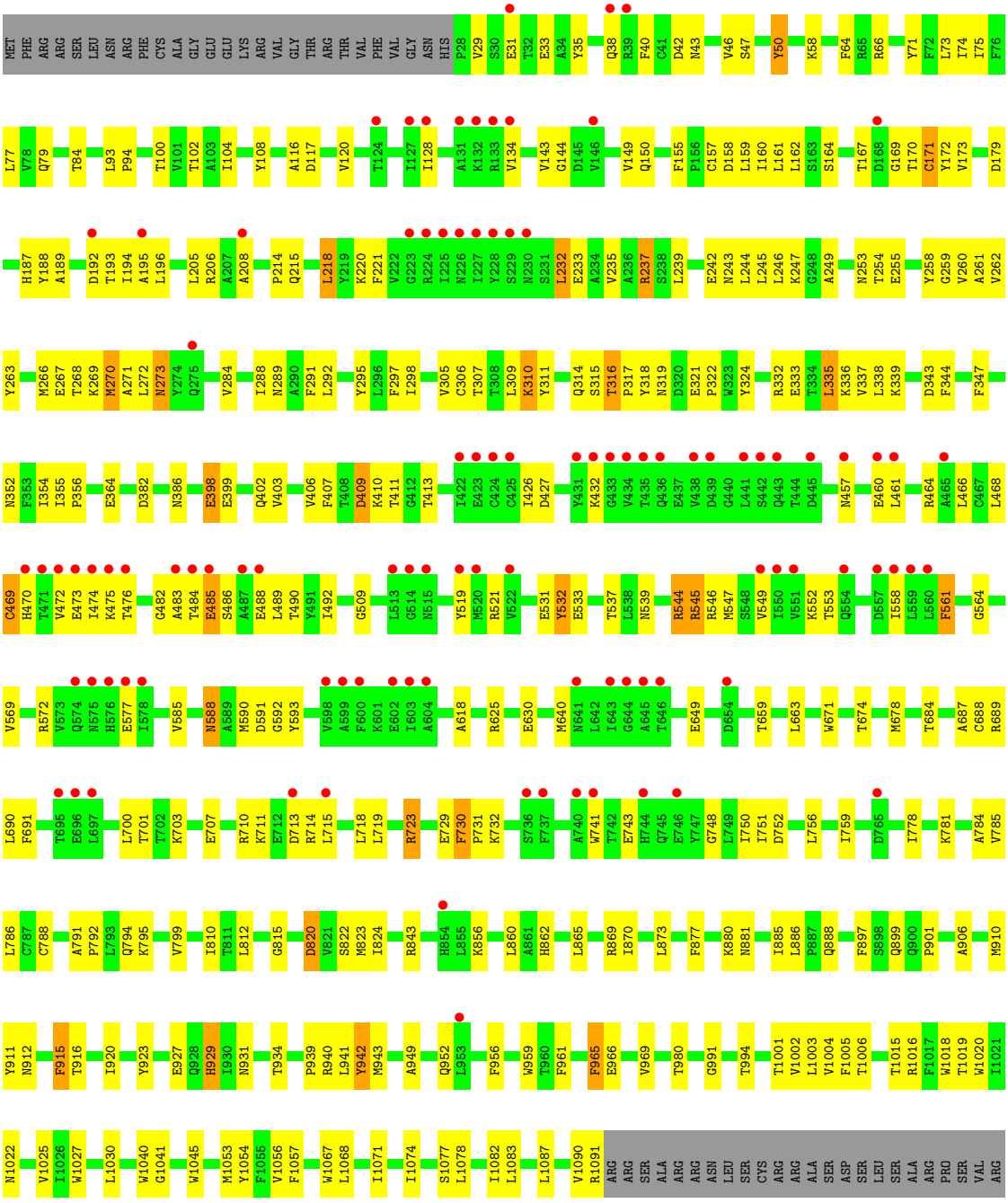
Chain E:





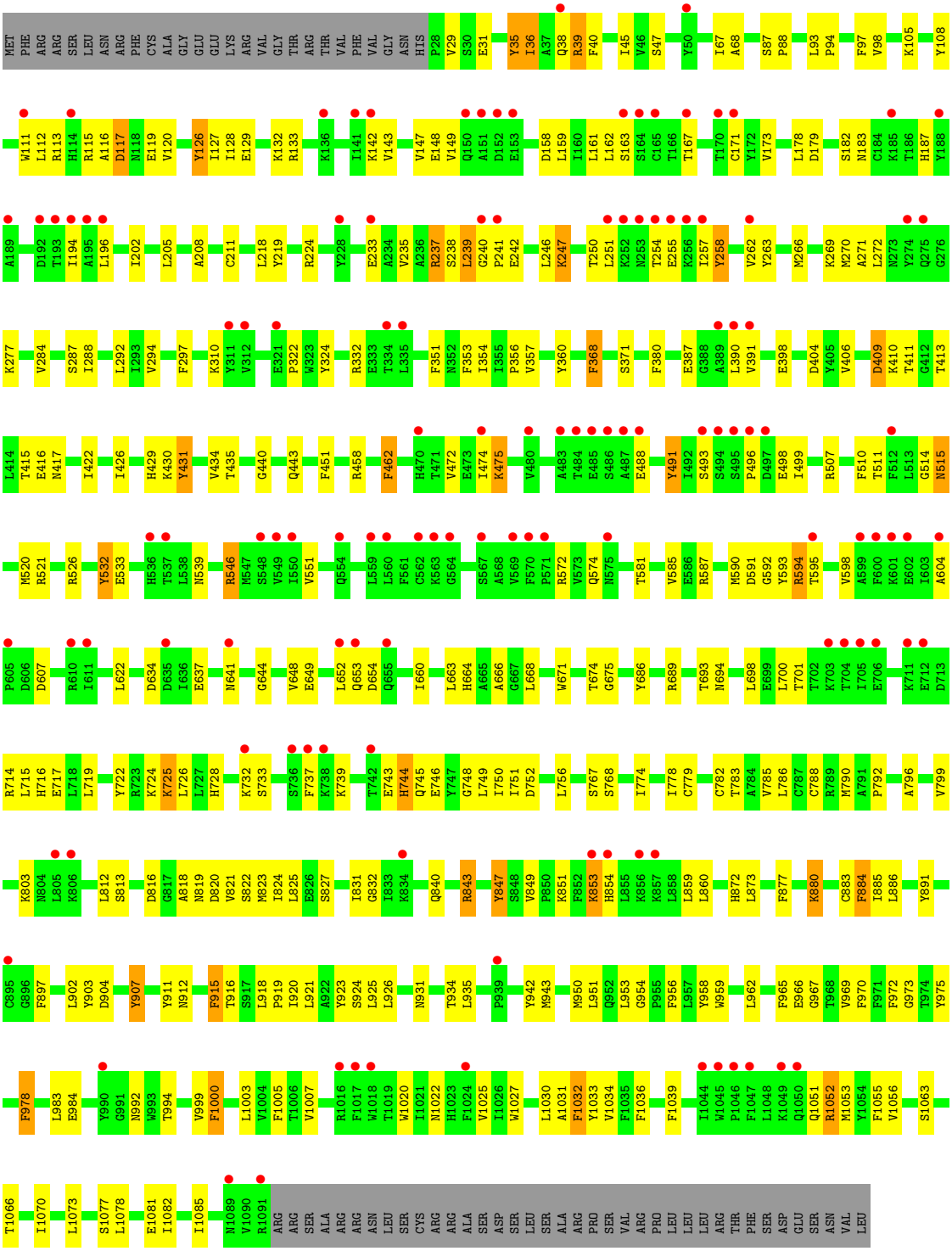
• Molecule 1: Phospholipid-transporting ATPase IG

Chain I:



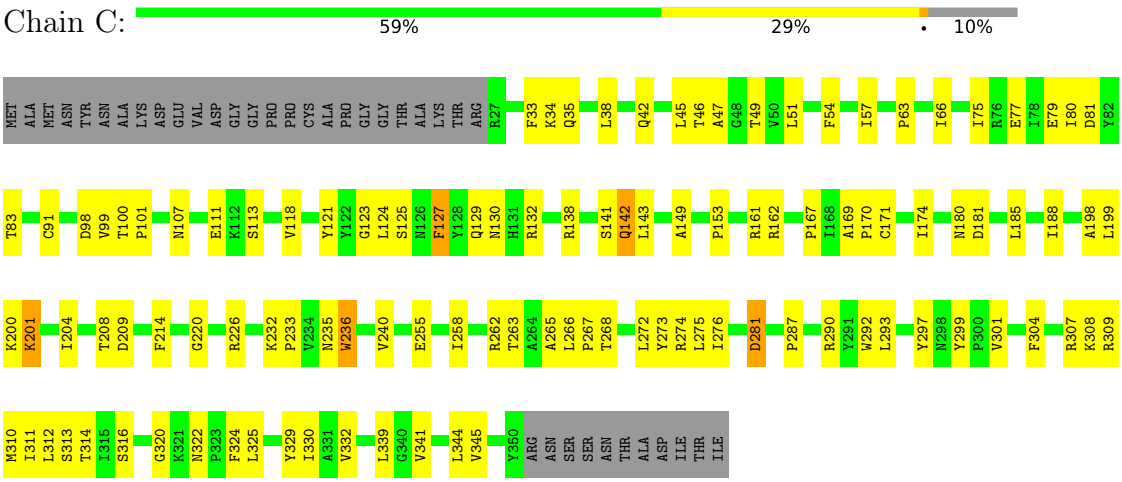
• Molecule 1: Phospholipid-transporting ATPase IG

Chain M:



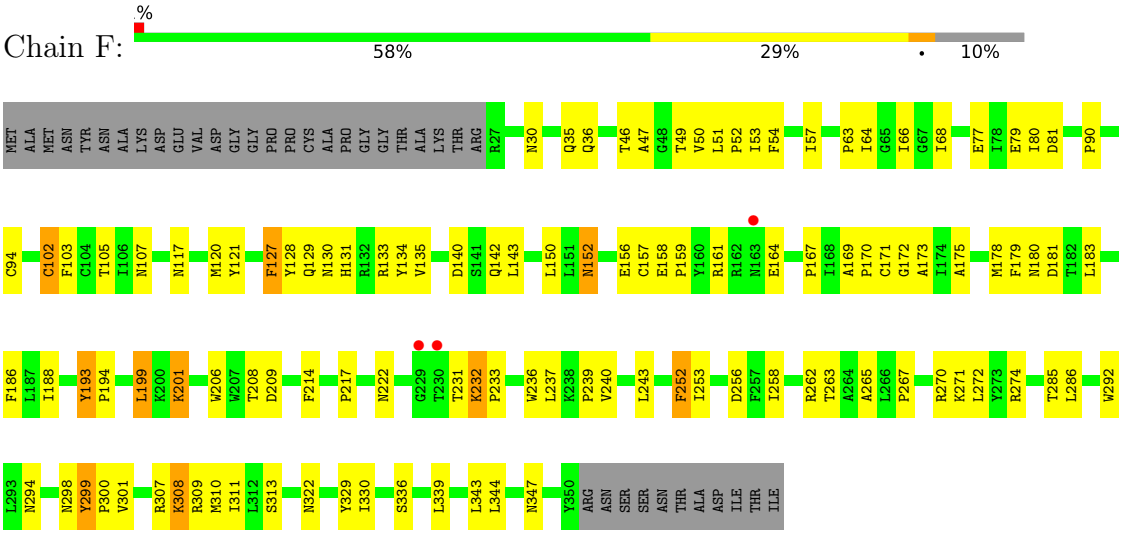
• Molecule 2: Cell cycle control protein 50A

Chain C:



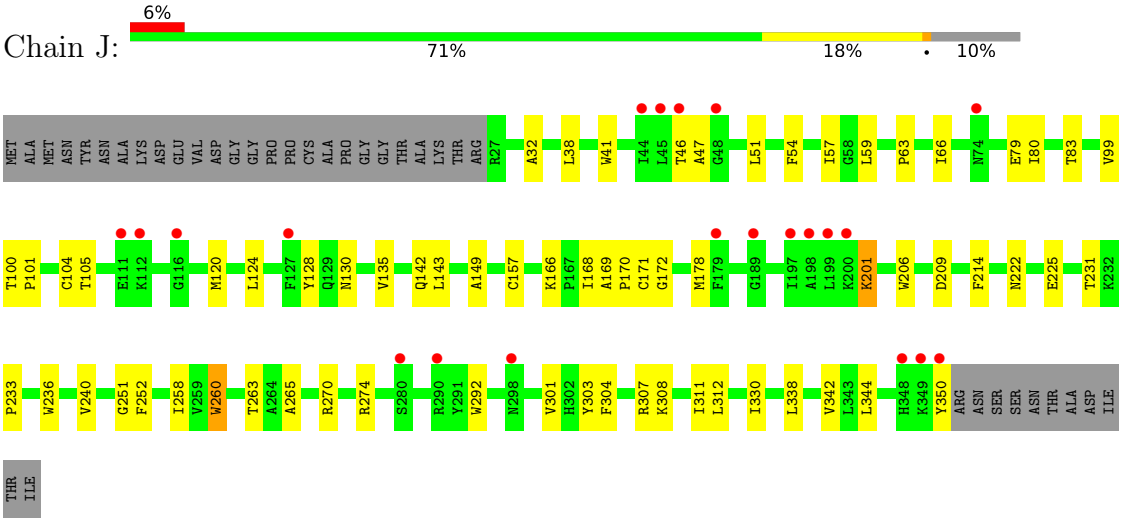
• Molecule 2: Cell cycle control protein 50A

Chain F:



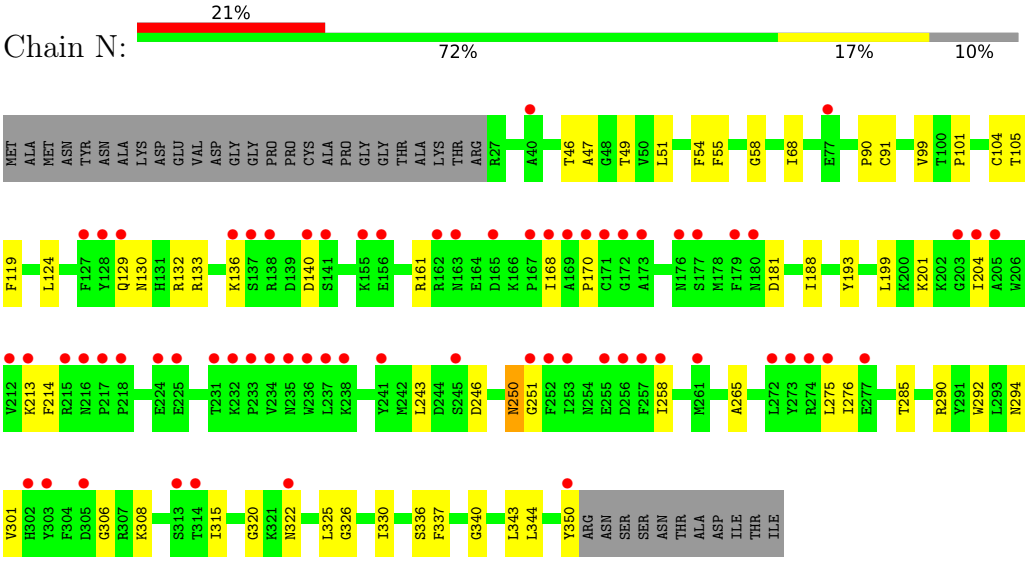
• Molecule 2: Cell cycle control protein 50A

Chain J:



• Molecule 2: Cell cycle control protein 50A

Chain N:


21%
72%
17%
10%
MET ALA MET ASN TYR ASN ALA LYS ASP GLU VAL ASP GLY GLY PRO PRO CYS ALA PRO GLY GLY THR ALA LYS THR ARG R27 A40 T46 A47 G48 T49 V50 L51 F54 F55 G58 I68 E77 P90 C91 V99 T100 P101 C104 T105 I106 N107 N117 V118
F119 L124 F127 Y128 Q129 N130 H131 R132 R133 K136 S137 R138 D139 D140 S141 K155 E156 R161 R162 N163 E164 D165 K166 P167 I168 A169 P170 C171 G172 A173 N176 S177 M178 F179 N180 D181 I188 Y193 L199 K200 K201 K202 G203 I204 A205 W206 W207 T208 D209 K210 N211
V212 K213 F214 R215 N216 P217 P218 E224 E225 T231 K232 P233 V234 N235 W236 L237 K238 Y241 M242 L243 D244 S245 D246 N250 G251 F252 I253 N254 E255 D256 F257 I258 M261 A265 L272 Y273 R274 L275 I276 E277 T285 R290 Y291 W292 L293 N294 Y297 N298 Y299 P300
V301 H302 Y303 F304 D305 G306 R307 K308 S313 T314 I315 G320 K321 N322 L325 G326 I330 S336 F337 G340 L343 L344 Y350 ARG ASN SER SER ASN THR ALA ASP ILE THR ILE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	100.46Å 232.83Å 492.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.97 – 3.90 49.97 – 3.78	Depositor EDS
% Data completeness (in resolution range)	71.4 (49.97-3.90) 65.4 (49.97-3.78)	Depositor EDS
R_{merge}	1.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.95 (at 3.77Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.281 , 0.350 0.281 , 0.350	Depositor DCC
R_{free} test set	3779 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	98.3	Xtriage
Anisotropy	0.458	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.22 , 92.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.21$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	45252	wwPDB-VP
Average B, all atoms (Å^2)	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.21% 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, BFD, P5S, NAG, AH2

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	0.34	0/8777	0.64	0/11883
1	E	0.34	0/8777	0.64	0/11883
1	I	0.29	0/8777	0.59	0/11883
1	M	0.29	0/8777	0.58	0/11883
2	C	0.33	0/2690	0.64	0/3659
2	F	0.33	0/2690	0.64	0/3659
2	J	0.27	0/2690	0.55	0/3659
2	N	0.28	0/2690	0.53	0/3659
All	All	0.31	0/45868	0.61	0/62168

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	8601	0	8587	591	0
1	E	8601	0	8588	436	0
1	I	8601	0	8587	461	0
1	M	8601	0	8588	267	0
2	C	2613	0	2592	101	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	2613	0	2594	107	0
2	J	2613	0	2592	68	0
2	N	2613	0	2592	84	0
3	A	11	0	5	0	0
3	C	54	0	80	2	0
3	E	22	0	10	1	0
3	I	22	0	10	9	0
3	M	11	0	5	1	0
4	A	1	0	0	0	0
4	E	1	0	0	0	0
4	I	1	0	0	0	0
4	M	1	0	0	0	0
5	C	55	0	50	3	0
5	F	55	0	50	23	0
5	J	55	0	50	18	0
5	N	55	0	49	15	0
6	C	11	0	0	0	0
6	F	11	0	0	0	0
6	J	11	0	0	0	0
6	N	11	0	0	0	0
7	A	2	0	0	0	0
7	E	2	0	0	4	0
7	I	2	0	0	1	0
7	M	2	0	0	1	0
All	All	45252	0	45029	2057	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 23.

The worst 5 of 2057 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:I:214:PRO:HB3	1:I:268:THR:CG2	1.23	1.60
1:I:678:MET:CE	1:I:700:LEU:HD13	1.27	1.60
1:A:28:PRO:CD	1:A:212:GLU:HA	1.15	1.58
1:A:519:TYR:HD2	1:A:532:TYR:CB	0.97	1.58
1:A:519:TYR:CE1	1:A:551:VAL:HG11	1.35	1.58

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	1061/1129 (94%)	917 (86%)	143 (14%)	1 (0%)	51 84
1	E	1061/1129 (94%)	875 (82%)	186 (18%)	0	100 100
1	I	1061/1129 (94%)	878 (83%)	182 (17%)	1 (0%)	51 84
1	M	1061/1129 (94%)	885 (83%)	176 (17%)	0	100 100
2	C	322/361 (89%)	266 (83%)	56 (17%)	0	100 100
2	F	322/361 (89%)	275 (85%)	47 (15%)	0	100 100
2	J	322/361 (89%)	281 (87%)	41 (13%)	0	100 100
2	N	322/361 (89%)	297 (92%)	25 (8%)	0	100 100
All	All	5532/5960 (93%)	4674 (84%)	856 (16%)	2 (0%)	100 100

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	316	THR
1	A	570	PHE

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	946/1004 (94%)	890 (94%)	56 (6%)	19 49
1	E	946/1004 (94%)	895 (95%)	51 (5%)	22 52
1	I	946/1004 (94%)	897 (95%)	49 (5%)	23 53

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	M	946/1004 (94%)	888 (94%)	58 (6%)	18 48
2	C	288/316 (91%)	276 (96%)	12 (4%)	30 57
2	F	288/316 (91%)	273 (95%)	15 (5%)	23 53
2	J	288/316 (91%)	283 (98%)	5 (2%)	60 78
2	N	288/316 (91%)	275 (96%)	13 (4%)	27 56
All	All	4936/5280 (94%)	4677 (95%)	259 (5%)	23 53

5 of 259 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	884	PHE
1	M	1005	PHE
1	E	820	ASP
1	E	691	PHE
1	M	1039	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	314	GLN
1	M	694	ASN
1	M	402	GLN
1	M	840	GLN
1	E	243	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
1	BFD	A	409	1,4	8,11,12	5.58	3 (37%)	3,15,17	1.72	1 (33%)
1	BFD	I	409	1	8,11,12	5.60	3 (37%)	3,15,17	2.12	1 (33%)
1	BFD	E	409	1	8,11,12	5.57	3 (37%)	3,15,17	1.14	0
1	BFD	M	409	1	8,11,12	5.62	3 (37%)	3,15,17	1.77	1 (33%)

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
1	BFD	A	409	1,4	-	2/5/11/13	-
1	BFD	I	409	1	-	2/5/11/13	-
1	BFD	E	409	1	-	4/5/11/13	-
1	BFD	M	409	1	-	4/5/11/13	-

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	M	409	BFD	F3-BE	9.39	1.77	1.54
1	I	409	BFD	F3-BE	9.38	1.77	1.54
1	A	409	BFD	F3-BE	9.33	1.76	1.54
1	E	409	BFD	F3-BE	9.32	1.76	1.54
1	M	409	BFD	F2-BE	9.01	1.76	1.54

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	I	409	BFD	CA-CB-CG	-3.26	106.02	112.86
1	M	409	BFD	CA-CB-CG	-2.65	107.30	112.86
1	A	409	BFD	OD2-CG-CB	-2.62	118.94	124.73

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	I	409	BFD	N-CA-CB-CG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	I	409	BFD	C-CA-CB-CG
1	M	409	BFD	C-CA-CB-CG
1	M	409	BFD	N-CA-CB-CG
1	A	409	BFD	CA-CB-CG-OD2

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	409	BFD	2	0
1	I	409	BFD	1	0
1	E	409	BFD	5	0
1	M	409	BFD	2	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 31 ligands modelled in this entry, 4 are monoatomic - leaving 27 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	$\#\|Z\| > 2$
5	NAG	F	401	-	14,14,15	1.12	1 (7%)	17,19,21	1.00	2 (11%)
5	NAG	C	405	-	14,14,15	0.63	1 (7%)	17,19,21	0.82	1 (5%)
5	NAG	C	402	-	14,14,15	1.24	1 (7%)	17,19,21	0.63	0
6	AH2	F	403	-	11,11,11	0.99	0	15,15,15	1.14	1 (6%)
6	AH2	N	403	-	11,11,11	0.81	0	15,15,15	1.39	3 (20%)
5	NAG	J	402	-	14,14,15	0.21	0	17,19,21	0.37	0
5	NAG	F	402	-	14,14,15	0.51	0	17,19,21	0.45	0
5	NAG	J	401	-	14,14,15	0.18	0	17,19,21	0.49	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	NAG	F	405	-	13,13,15	0.34	0	14,17,21	0.59	0
5	NAG	N	405	-	13,13,15	0.35	0	14,17,21	0.53	0
5	NAG	N	404	-	14,14,15	0.29	0	17,19,21	0.65	1 (5%)
3	P5S	I	1202	-	9,10,53	0.92	0	12,14,60	1.12	0
5	NAG	J	405	-	13,13,15	0.56	0	14,17,21	0.71	0
6	AH2	C	404	-	11,11,11	0.72	0	15,15,15	1.16	1 (6%)
5	NAG	C	403	-	14,14,15	0.56	0	17,19,21	0.97	2 (11%)
5	NAG	J	404	-	14,14,15	0.38	0	17,19,21	0.72	1 (5%)
5	NAG	N	401	-	14,14,15	0.40	0	17,19,21	0.38	0
3	P5S	A	1201	-	9,10,53	0.93	0	12,14,60	1.18	1 (8%)
3	P5S	E	1202	-	9,10,53	0.92	0	12,14,60	1.35	2 (16%)
5	NAG	C	406	-	13,13,15	0.49	0	14,17,21	0.52	0
3	P5S	E	1201	-	9,10,53	0.92	0	12,14,60	1.09	1 (8%)
3	P5S	I	1201	-	9,10,53	0.91	0	12,14,60	1.44	3 (25%)
3	P5S	C	401	-	52,53,53	0.98	2 (3%)	56,60,60	1.09	2 (3%)
6	AH2	J	403	-	11,11,11	0.83	0	15,15,15	1.28	1 (6%)
5	NAG	F	404	-	14,14,15	0.25	0	17,19,21	0.61	1 (5%)
5	NAG	N	402	-	14,14,15	0.22	0	17,19,21	0.42	0
3	P5S	M	1201	-	9,10,53	0.91	0	12,14,60	1.28	3 (25%)

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
5	NAG	F	401	-	-	4/6/23/26	0/1/1/1
5	NAG	C	405	-	-	1/6/23/26	0/1/1/1
5	NAG	C	402	-	-	4/6/23/26	0/1/1/1
6	AH2	F	403	-	-	1/2/19/19	0/1/1/1
6	AH2	N	403	-	-	1/2/19/19	1/1/1/1
5	NAG	J	402	-	-	1/6/23/26	0/1/1/1
5	NAG	F	402	-	-	1/6/23/26	0/1/1/1
5	NAG	J	401	-	-	3/6/23/26	0/1/1/1
5	NAG	F	405	-	-	3/6/19/26	1/1/1/1
5	NAG	N	405	-	-	1/6/19/26	0/1/1/1
5	NAG	N	404	-	-	2/6/23/26	0/1/1/1
3	P5S	I	1202	-	-	2/10/10/59	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	J	405	-	-	3/6/19/26	1/1/1/1
6	AH2	C	404	-	-	1/2/19/19	1/1/1/1
5	NAG	C	403	-	-	2/6/23/26	0/1/1/1
5	NAG	J	404	-	-	2/6/23/26	0/1/1/1
5	NAG	N	401	-	-	2/6/23/26	0/1/1/1
3	P5S	A	1201	-	-	3/10/10/59	-
3	P5S	E	1202	-	-	3/10/10/59	-
5	NAG	C	406	-	-	2/6/19/26	0/1/1/1
3	P5S	E	1201	-	-	9/10/10/59	-
3	P5S	I	1201	-	-	5/10/10/59	-
3	P5S	C	401	-	-	19/59/59/59	-
6	AH2	J	403	-	-	2/2/19/19	1/1/1/1
5	NAG	F	404	-	-	1/6/23/26	0/1/1/1
5	NAG	N	402	-	-	0/6/23/26	0/1/1/1
3	P5S	M	1201	-	-	4/10/10/59	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	401	P5S	O37-C38	4.53	1.47	1.34
3	C	401	P5S	O19-C17	4.43	1.46	1.33
5	C	402	NAG	O5-C1	-4.43	1.36	1.43
5	F	401	NAG	O5-C1	-3.72	1.37	1.43
5	C	405	NAG	O5-C1	2.07	1.47	1.43

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	401	P5S	O37-C38-C39	4.42	121.02	111.50
6	J	403	AH2	C1-O5-C5	3.87	117.44	112.19
6	N	403	AH2	C1-O5-C5	3.75	117.27	112.19
6	C	404	AH2	C1-O5-C5	3.38	116.78	112.19
3	I	1201	P5S	OG-CB-CA	3.25	110.89	108.06

There are no chirality outliers.

5 of 82 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1201	P5S	C-CA-CB-OG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	1201	P5S	N-CA-CB-OG
3	C	401	P5S	O-C-CA-N
3	C	401	P5S	N-CA-CB-OG
3	C	401	P5S	C39-C38-O37-C2

All (5) ring outliers are listed below:

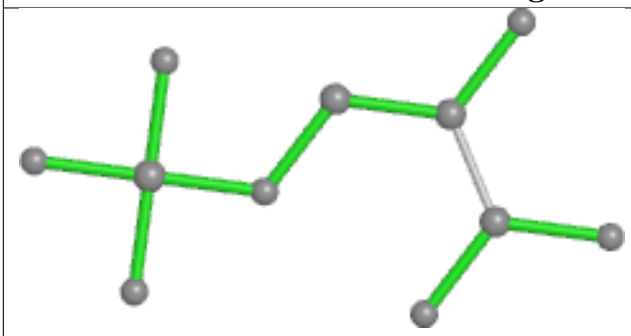
Mol	Chain	Res	Type	Atoms
5	F	405	NAG	C1-C2-C3-C4-C5-O5
6	N	403	AH2	C1-C2-C3-C4-C5-O5
6	J	403	AH2	C1-C2-C3-C4-C5-O5
6	C	404	AH2	C1-C2-C3-C4-C5-O5
5	J	405	NAG	C1-C2-C3-C4-C5-O5

12 monomers are involved in 72 short contacts:

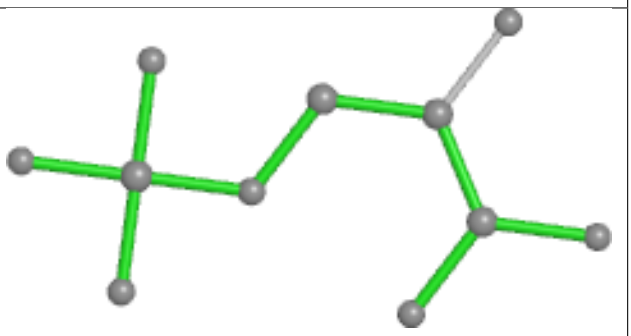
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	F	401	NAG	1	0
5	C	402	NAG	2	0
5	F	405	NAG	18	0
5	N	405	NAG	15	0
5	J	405	NAG	7	0
5	J	404	NAG	11	0
3	E	1202	P5S	1	0
5	C	406	NAG	1	0
3	I	1201	P5S	9	0
3	C	401	P5S	2	0
5	F	404	NAG	4	0
3	M	1201	P5S	1	0

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

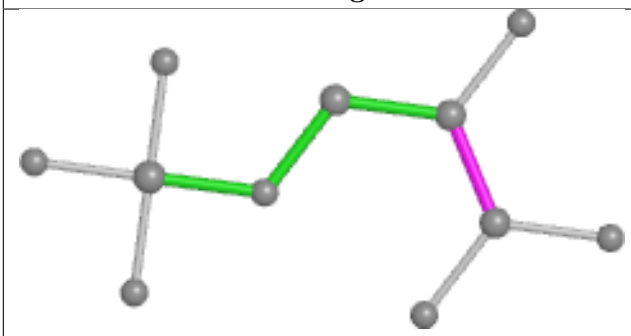
Ligand P5S I 1202



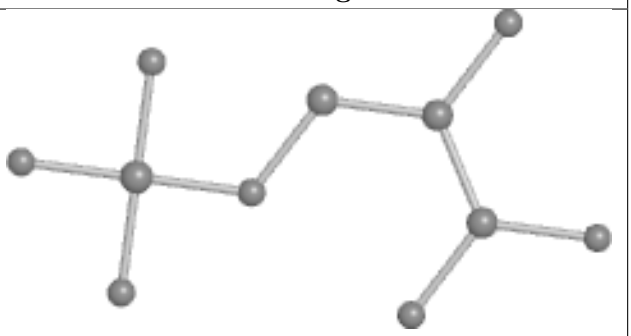
Bond lengths



Bond angles

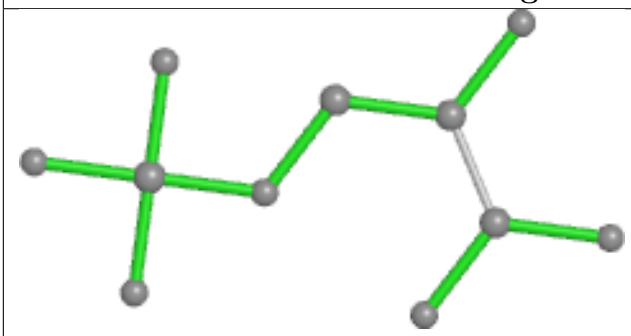


Torsions

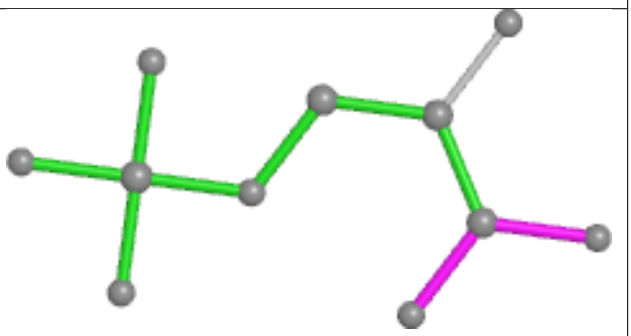


Rings

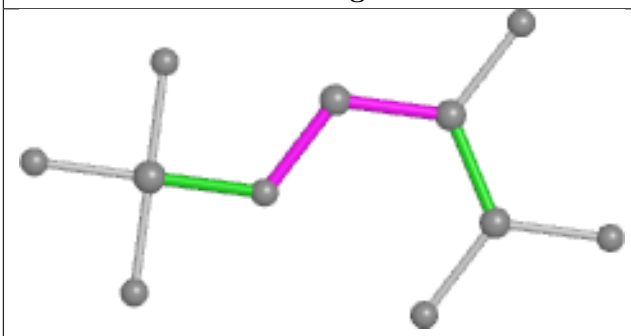
Ligand P5S A 1201



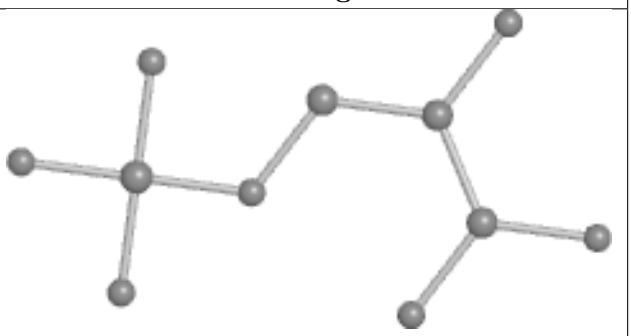
Bond lengths



Bond angles

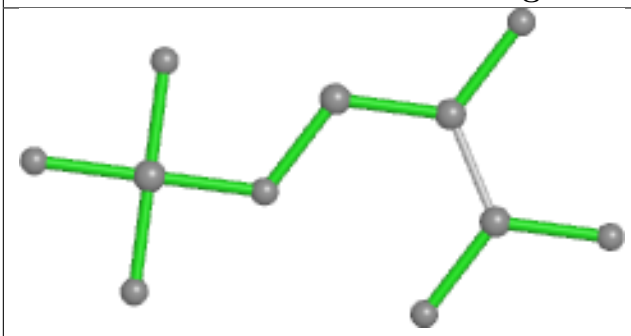


Torsions

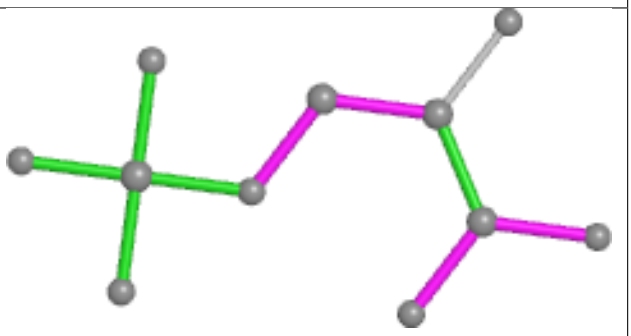


Rings

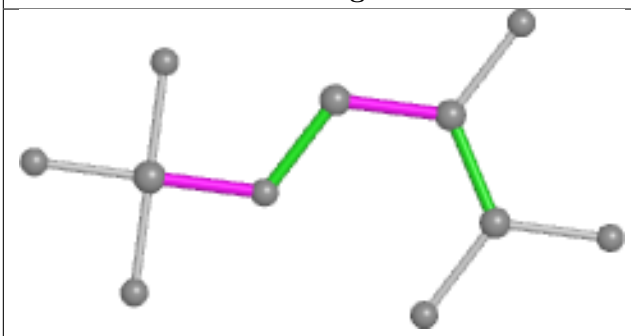
Ligand P5S E 1202



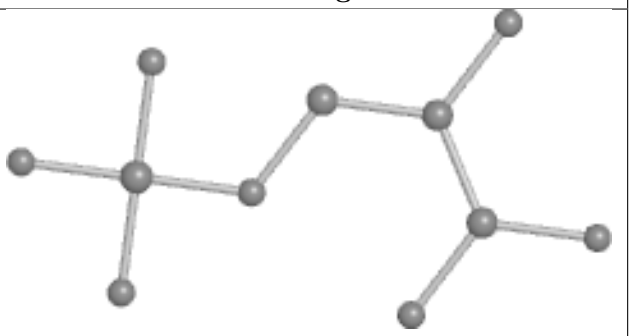
Bond lengths



Bond angles

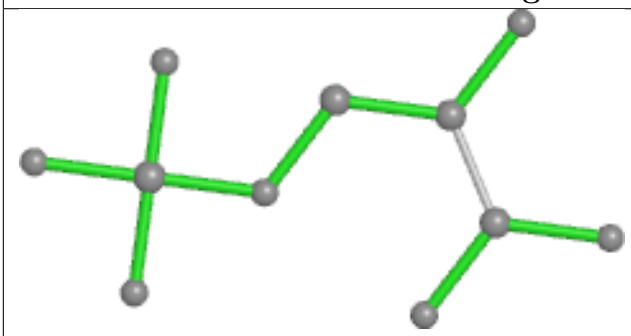


Torsions

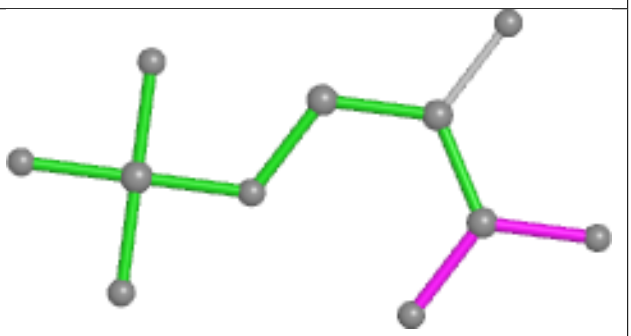


Rings

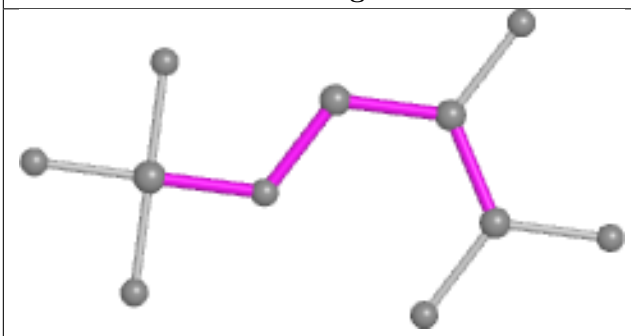
Ligand P5S E 1201



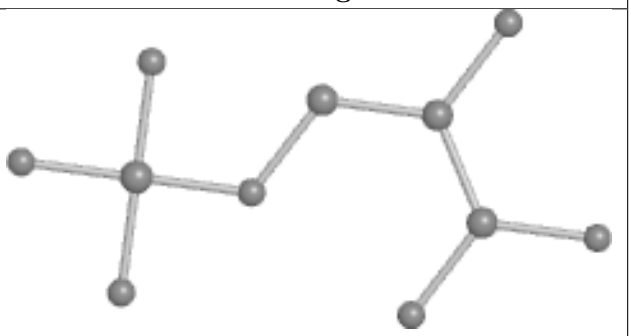
Bond lengths



Bond angles

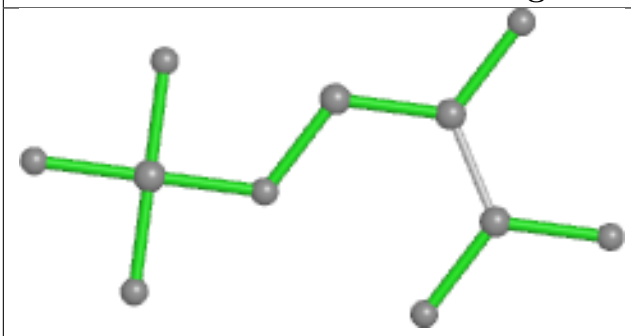


Torsions

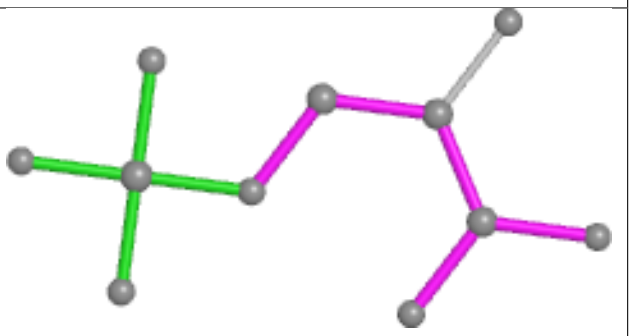


Rings

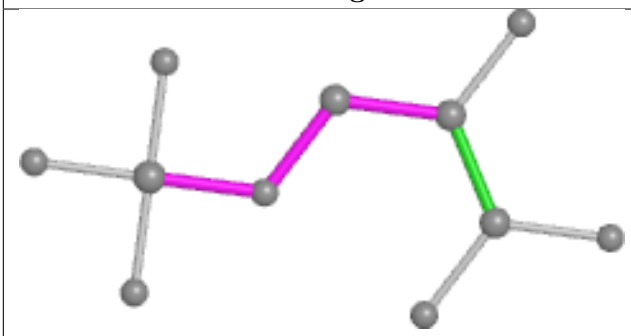
Ligand P5S I 1201



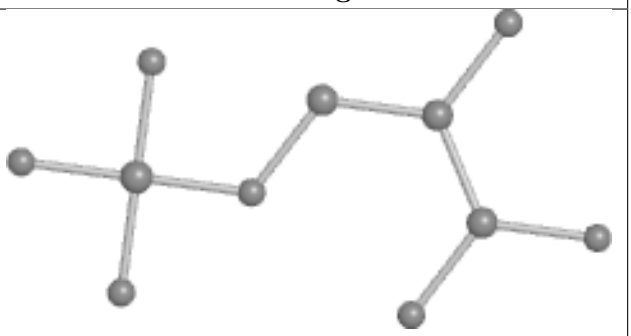
Bond lengths



Bond angles

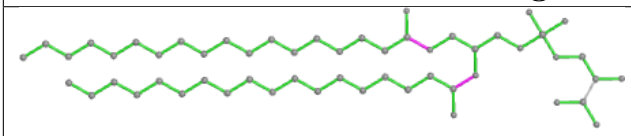


Torsions

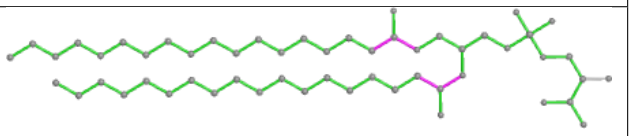


Rings

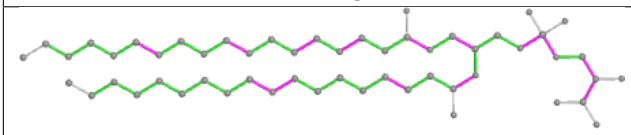
Ligand P5S C 401



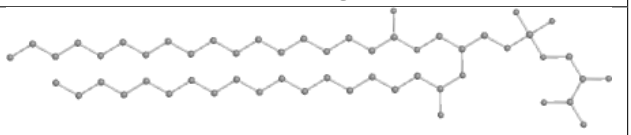
Bond lengths



Bond angles

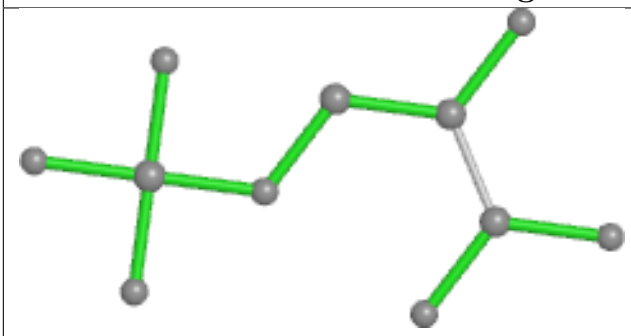


Torsions

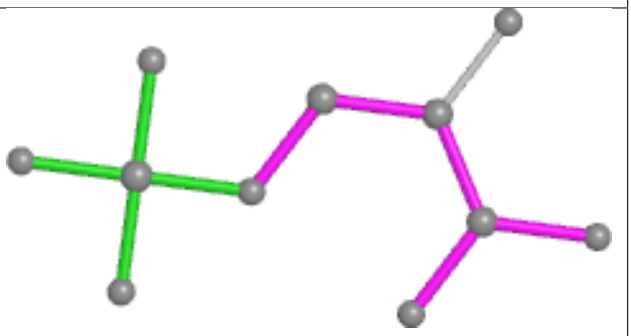


Rings

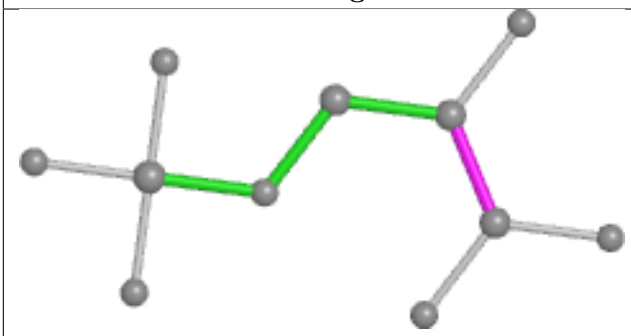
Ligand P5S M 1201



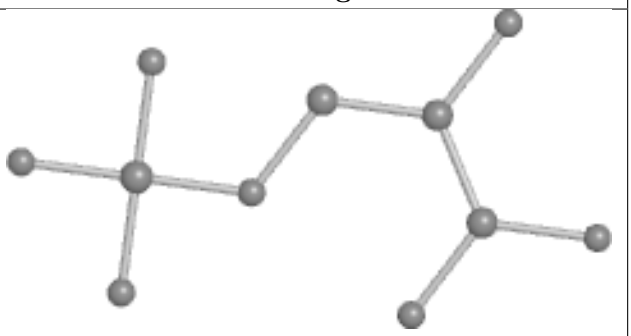
Bond lengths



Bond angles



Torsions



Rings

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^{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	1063/1129 (94%)	-0.16	42 (3%) 38 30	24, 120, 194, 235	0
1	E	1063/1129 (94%)	0.05	93 (8%) 10 8	31, 115, 211, 246	0
1	I	1063/1129 (94%)	0.18	101 (9%) 8 6	96, 173, 226, 261	0
1	M	1063/1129 (94%)	0.33	125 (11%) 4 4	124, 214, 315, 367	0
2	C	324/361 (89%)	-0.60	0 100 100	21, 79, 153, 183	0
2	F	324/361 (89%)	-0.58	3 (0%) 84 77	25, 71, 136, 166	0
2	J	324/361 (89%)	-0.05	21 (6%) 18 13	118, 189, 255, 281	0
2	N	324/361 (89%)	0.86	75 (23%) 0 0	204, 340, 392, 416	0
All	All	5548/5960 (93%)	0.05	460 (8%) 11 9	21, 164, 320, 416	0

The worst 5 of 460 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	N	172	GLY	13.3
1	E	37	ALA	10.7
2	N	176	ASN	10.2
1	E	431	TYR	9.9
1	E	432	LYS	9.1

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. 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
1	BFD	M	409	12/13	0.81	0.32	184,209,238,250	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	BFD	I	409	12/13	0.87	0.29	148,153,162,162	0
1	BFD	A	409	12/13	0.95	0.27	70,84,106,109	0
1	BFD	E	409	12/13	0.95	0.23	59,78,88,89	0

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, 95th 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
6	AH2	N	403	11/11	-0.02	0.90	347,350,356,359	0
5	NAG	N	404	14/15	0.55	0.25	398,409,413,415	0
3	P5S	M	1201	11/54	0.60	0.21	293,299,310,312	0
5	NAG	J	405	13/15	0.63	0.21	255,269,274,276	0
5	NAG	N	401	14/15	0.68	0.47	335,342,344,344	0
6	AH2	F	403	11/11	0.71	0.24	143,165,168,170	0
6	AH2	J	403	11/11	0.78	0.22	155,161,168,172	0
5	NAG	J	404	14/15	0.78	0.20	238,242,246,249	0
6	AH2	C	404	11/11	0.79	0.27	117,126,144,145	0
3	P5S	I	1201	11/54	0.80	0.26	220,231,238,239	0
5	NAG	J	401	14/15	0.80	0.22	145,176,183,185	0
3	P5S	I	1202	11/54	0.81	0.34	151,162,189,189	0
5	NAG	F	405	13/15	0.83	0.20	125,133,149,153	0
3	P5S	E	1202	11/54	0.85	0.25	78,97,159,160	0
3	P5S	C	401	54/54	0.86	0.33	92,120,135,141	0
5	NAG	N	402	14/15	0.87	0.53	309,312,317,319	0
5	NAG	N	405	13/15	0.87	0.16	356,362,366,367	0
3	P5S	A	1201	11/54	0.90	0.37	107,117,167,172	0
4	MG	I	1203	1/1	0.90	0.27	139,139,139,139	0
5	NAG	C	406	13/15	0.90	0.19	70,93,122,127	0
5	NAG	F	404	14/15	0.90	0.23	61,100,139,145	0
5	NAG	C	405	14/15	0.92	0.19	99,111,127,137	0
5	NAG	F	402	14/15	0.93	0.16	45,82,110,113	0
5	NAG	J	402	14/15	0.93	0.13	156,172,186,190	0
5	NAG	F	401	14/15	0.94	0.18	66,80,102,109	0

Continued on next page...

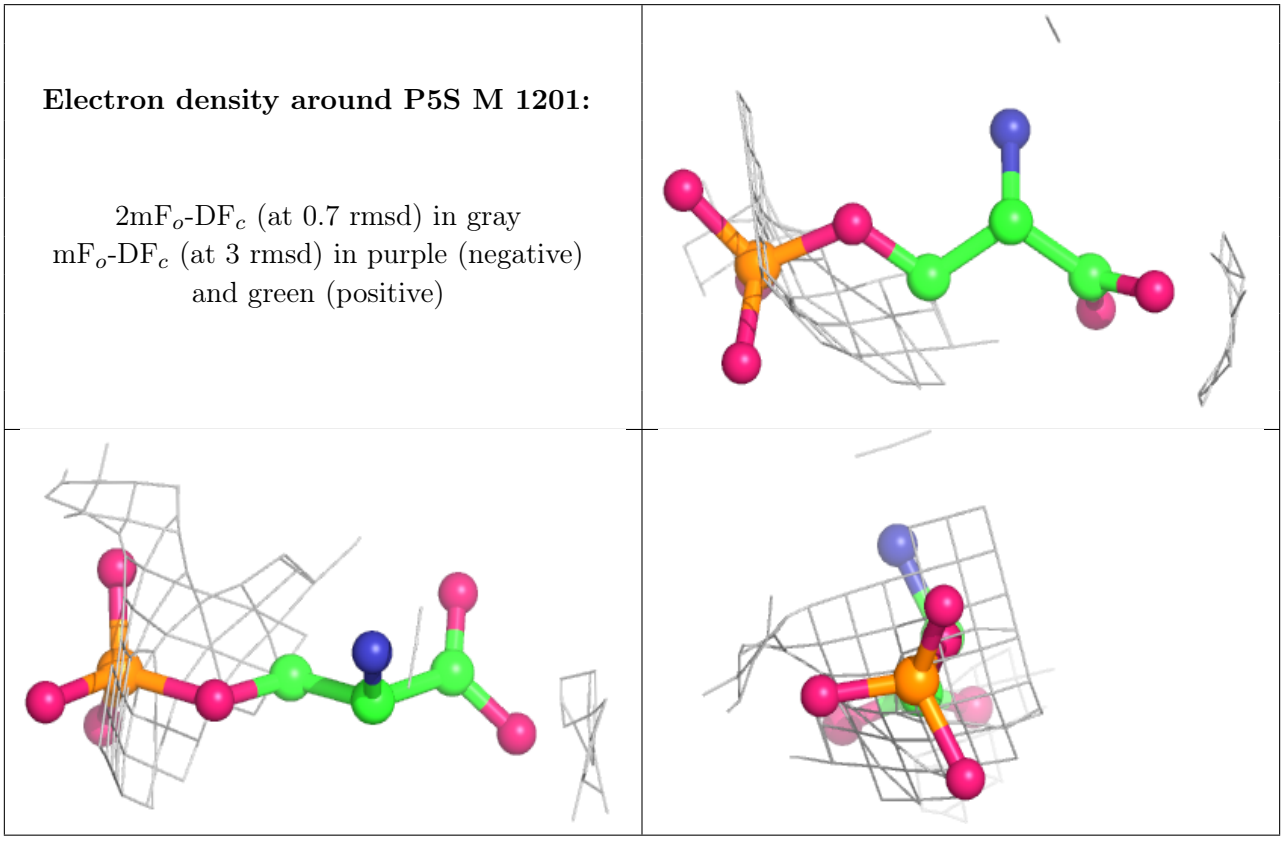
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	NAG	C	402	14/15	0.95	0.18	41,71,97,98	0
5	NAG	C	403	14/15	0.95	0.16	75,107,121,124	0
4	MG	M	1202	1/1	0.97	0.40	185,185,185,185	0
4	MG	E	1203	1/1	0.97	0.33	91,91,91,91	0
3	P5S	E	1201	11/54	0.97	0.18	70,85,102,110	0
4	MG	A	1202	1/1	0.99	0.28	172,172,172,172	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.

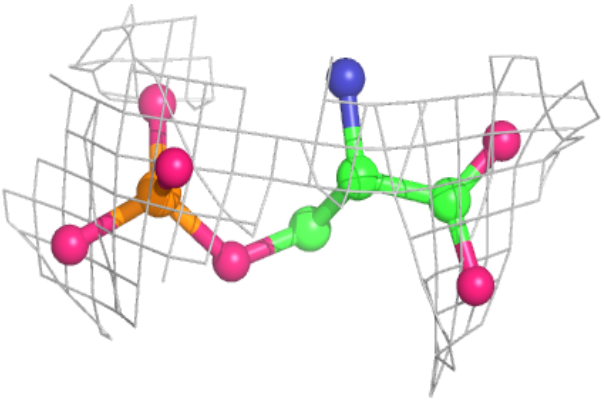
Electron density around P5S M 1201:

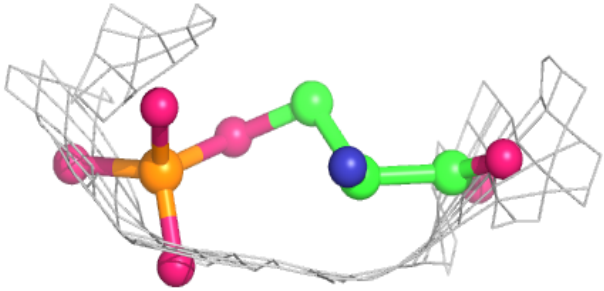
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

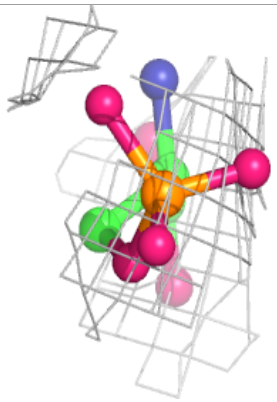


Electron density around P5S I 1201:

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

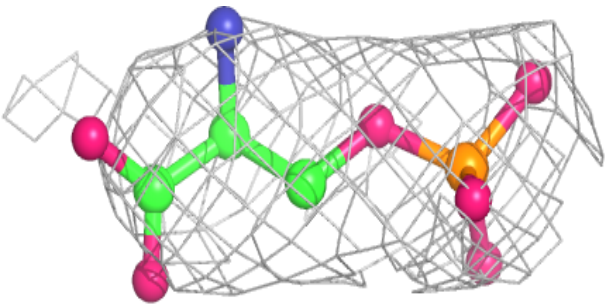


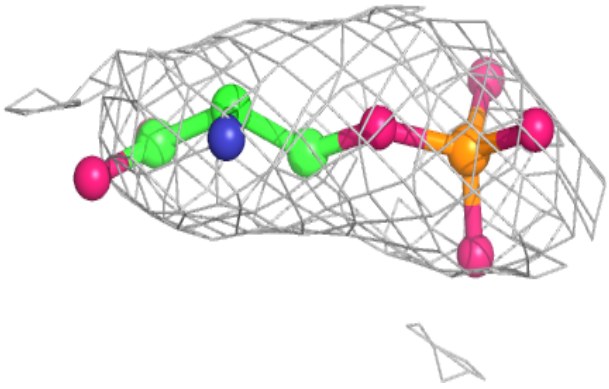


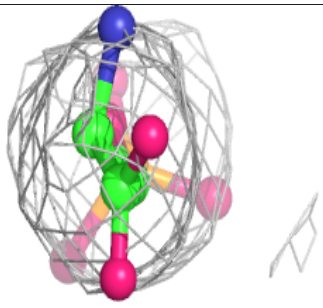


Electron density around P5S I 1202:

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

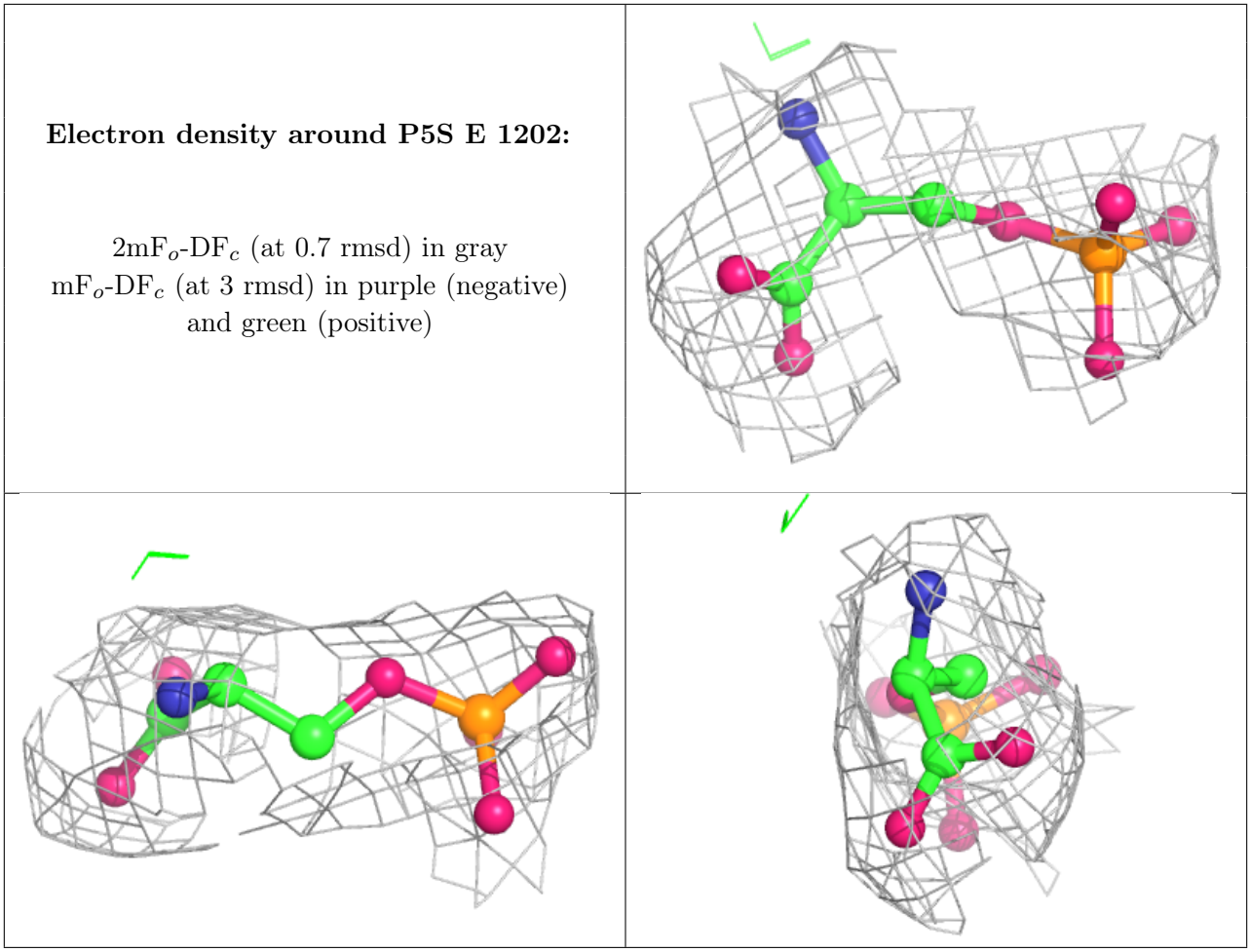






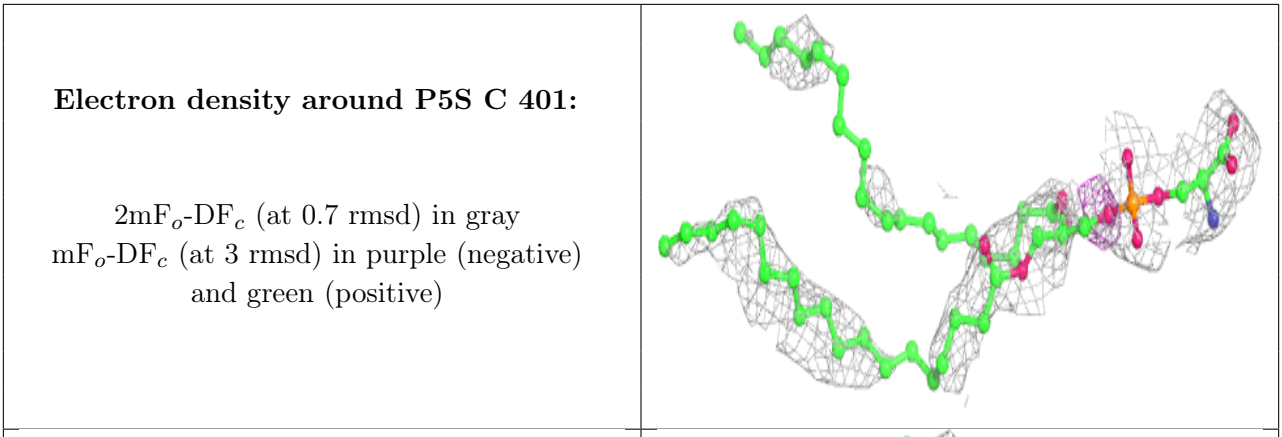
Electron density around P5S E 1202:

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



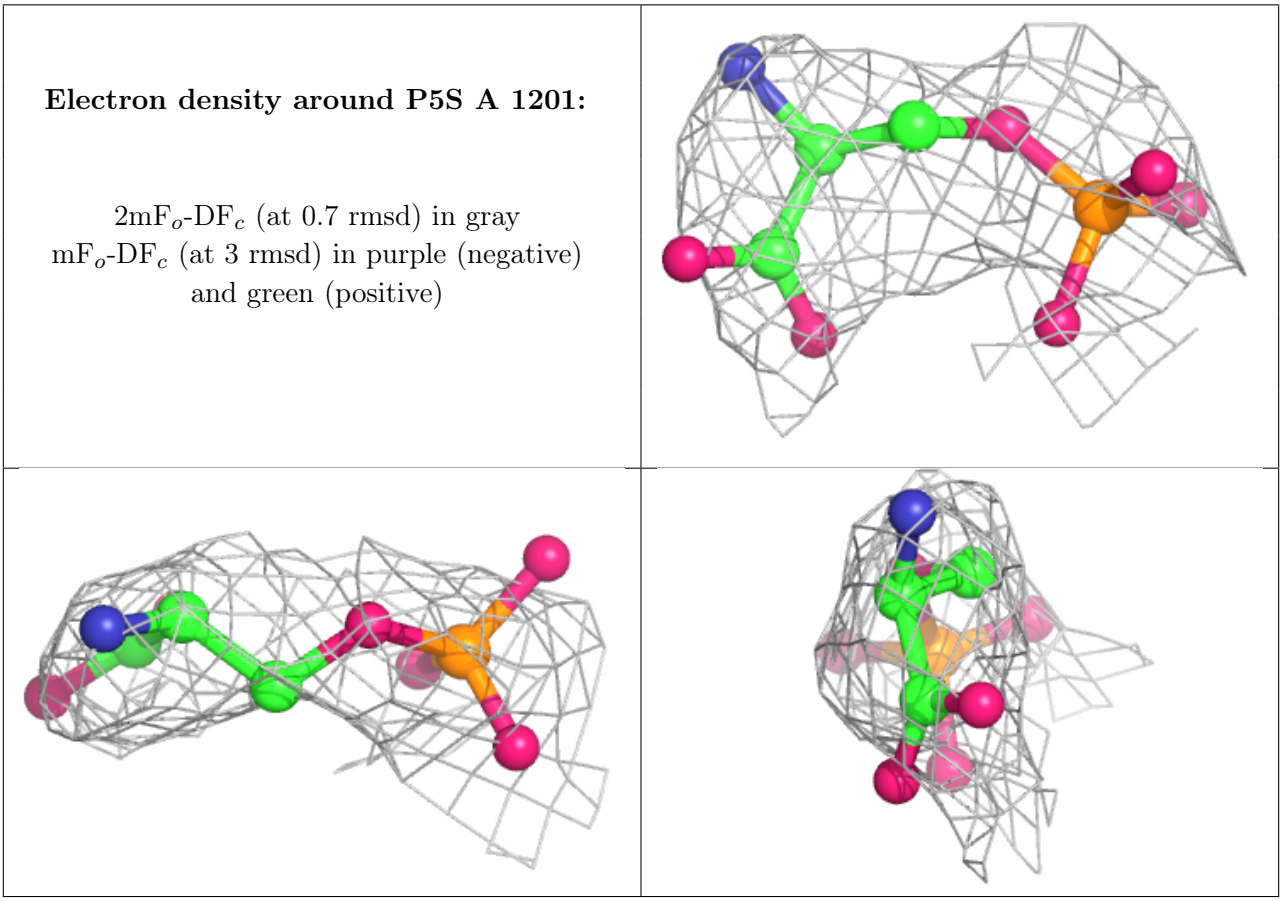
Electron density around P5S C 401:

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



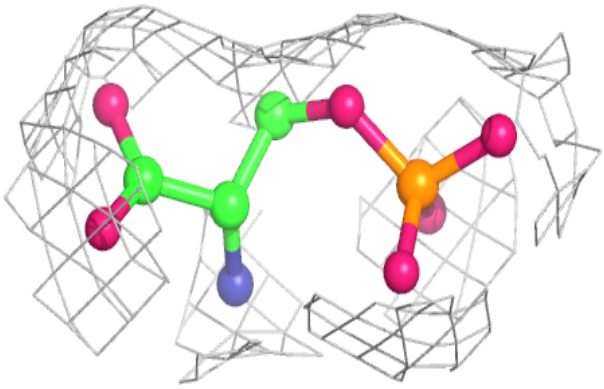
Electron density around P5S A 1201:

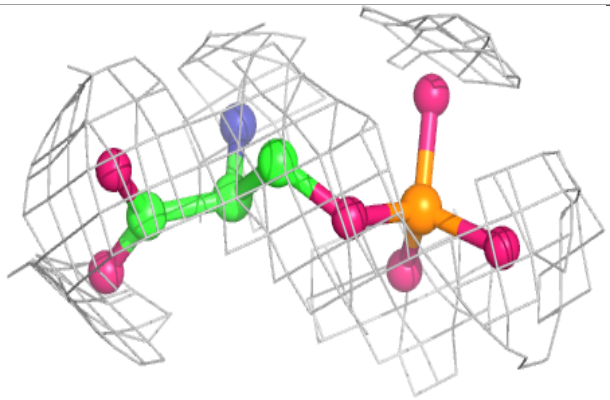
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

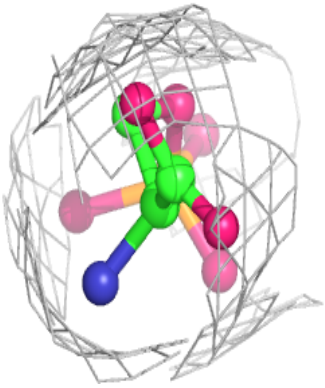


Electron density around P5S E 1201:

$2mF_o$-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

There are no such residues in this entry.