wwPDB X-ray Structure Validation Summary Report (i)

Sep 9, 2023 – 11:43 PM EDT

PDB ID	:	4I1O
Title	:	Crystal structure of the Legionella pneumophila GAP domain of LepB in complex with Rab1b bound to GDP and BeF3
Authors	:	Gazdag, E.M.; Streller, A.; Vetter, I.R.; Goody, R.S.; Itzen, A.
Deposited on	:	2012-11-21
Resolution	:	2.70 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:

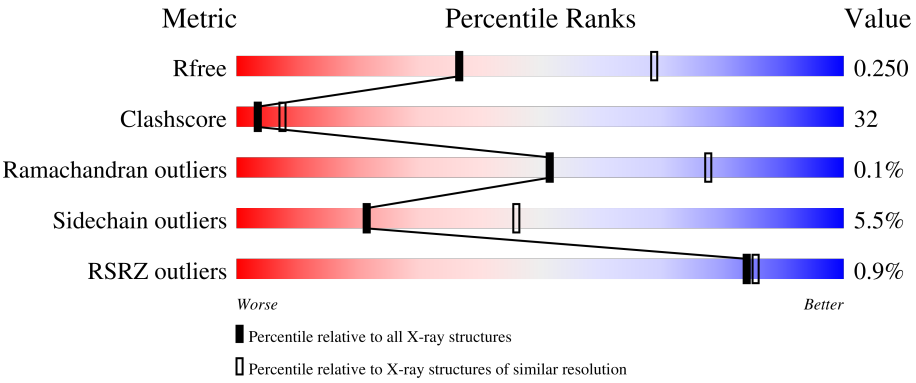
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	181	67% 28%
1	C	181	64% 32%
1	E	181	62% 34%
1	G	181	65% 30%
2	B	302	66% 24% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	302	
2	F	302	
2	H	302	

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
6	PEG	A	204	-	-	X	-
6	PEG	C	204	-	-	X	-
6	PEG	C	205	-	-	X	-
6	PEG	C	206	-	-	X	-
6	PEG	G	204	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 14370 atoms, of which 60 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-1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	177	Total	C	N	O	S	3	0	0
			1360	869	219	268	4			
1	C	177	Total	C	N	O	S	0	0	0
			1381	881	225	271	4			
1	E	177	Total	C	N	O	S	0	0	0
			1378	880	225	269	4			
1	G	177	Total	C	N	O	S	0	0	0
			1373	877	225	267	4			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q9H0U4
A	1	HIS	-	expression tag	UNP Q9H0U4
A	2	MET	-	expression tag	UNP Q9H0U4
A	175	LEU	-	expression tag	UNP Q9H0U4
A	176	GLU	-	expression tag	UNP Q9H0U4
A	177	VAL	-	expression tag	UNP Q9H0U4
A	178	LEU	-	expression tag	UNP Q9H0U4
A	179	PHE	-	expression tag	UNP Q9H0U4
A	180	GLN	-	expression tag	UNP Q9H0U4
C	0	GLY	-	expression tag	UNP Q9H0U4
C	1	HIS	-	expression tag	UNP Q9H0U4
C	2	MET	-	expression tag	UNP Q9H0U4
C	175	LEU	-	expression tag	UNP Q9H0U4
C	176	GLU	-	expression tag	UNP Q9H0U4
C	177	VAL	-	expression tag	UNP Q9H0U4
C	178	LEU	-	expression tag	UNP Q9H0U4
C	179	PHE	-	expression tag	UNP Q9H0U4
C	180	GLN	-	expression tag	UNP Q9H0U4
E	0	GLY	-	expression tag	UNP Q9H0U4
E	1	HIS	-	expression tag	UNP Q9H0U4
E	2	MET	-	expression tag	UNP Q9H0U4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	175	LEU	-	expression tag	UNP Q9H0U4
E	176	GLU	-	expression tag	UNP Q9H0U4
E	177	VAL	-	expression tag	UNP Q9H0U4
E	178	LEU	-	expression tag	UNP Q9H0U4
E	179	PHE	-	expression tag	UNP Q9H0U4
E	180	GLN	-	expression tag	UNP Q9H0U4
G	0	GLY	-	expression tag	UNP Q9H0U4
G	1	HIS	-	expression tag	UNP Q9H0U4
G	2	MET	-	expression tag	UNP Q9H0U4
G	175	LEU	-	expression tag	UNP Q9H0U4
G	176	GLU	-	expression tag	UNP Q9H0U4
G	177	VAL	-	expression tag	UNP Q9H0U4
G	178	LEU	-	expression tag	UNP Q9H0U4
G	179	PHE	-	expression tag	UNP Q9H0U4
G	180	GLN	-	expression tag	UNP Q9H0U4

- Molecule 2 is a protein called LepB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	280	Total	C	N	O	S	27	0	0
			2121	1370	353	393	5			
2	D	281	Total	C	N	O	S	19	0	0
			2128	1377	355	390	6			
2	F	281	Total	C	N	O	S	35	0	0
			2122	1371	353	393	5			
2	H	281	Total	C	N	O	S	41	0	0
			2127	1377	354	390	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	457	ALA	LYS	engineered mutation	UNP Q5ZSM7
B	458	ALA	GLU	engineered mutation	UNP Q5ZSM7
B	460	ALA	LYS	engineered mutation	UNP Q5ZSM7
D	457	ALA	LYS	engineered mutation	UNP Q5ZSM7
D	458	ALA	GLU	engineered mutation	UNP Q5ZSM7
D	460	ALA	LYS	engineered mutation	UNP Q5ZSM7
F	457	ALA	LYS	engineered mutation	UNP Q5ZSM7
F	458	ALA	GLU	engineered mutation	UNP Q5ZSM7
F	460	ALA	LYS	engineered mutation	UNP Q5ZSM7
H	457	ALA	LYS	engineered mutation	UNP Q5ZSM7
H	458	ALA	GLU	engineered mutation	UNP Q5ZSM7

Continued on next page...

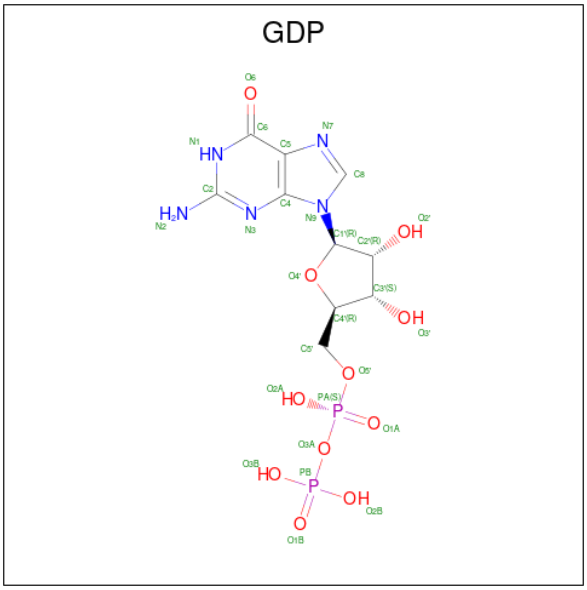
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	460	ALA	LYS	engineered mutation	UNP Q5ZSM7

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

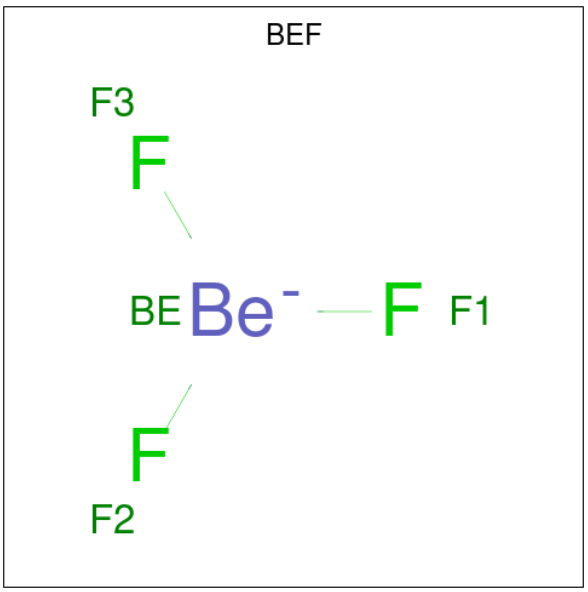
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0
3	E	1	Total Mg 1 1	0	0
3	G	1	Total Mg 1 1	0	0

- Molecule 4 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



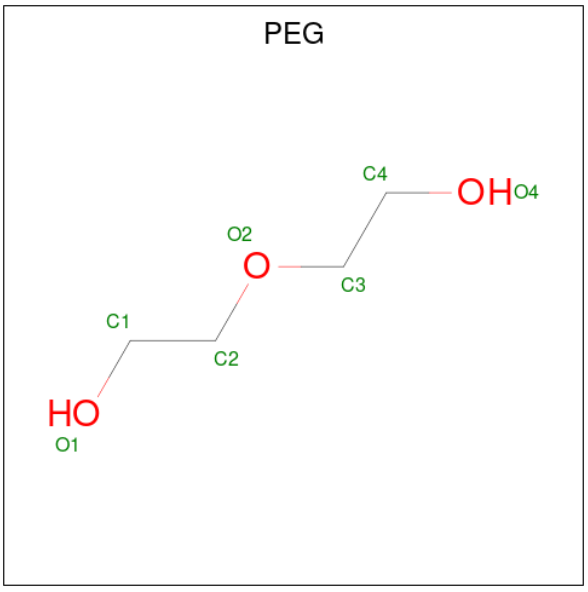
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O P 28 10 5 11 2	0	0
4	C	1	Total C N O P 28 10 5 11 2	0	0
4	E	1	Total C N O P 28 10 5 11 2	0	0
4	G	1	Total C N O P 28 10 5 11 2	0	0

- Molecule 5 is BERYLLIUM TRIFLUORIDE ION (three-letter code: BEF) (formula: BeF_3).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Be F 4 1 3	0	0
5	C	1	Total Be F 4 1 3	0	0
5	E	1	Total Be F 4 1 3	0	0
5	G	1	Total Be F 4 1 3	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	H	O	0	0
			17	4	10	3		
6	C	1	Total	C	H	O	0	0
			17	4	10	3		
6	C	1	Total	C	H	O	0	0
			17	4	10	3		
6	C	1	Total	C	H	O	0	0
			17	4	10	3		
6	D	1	Total	C	H	O	0	0
			17	4	10	3		
6	G	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 7 is water.

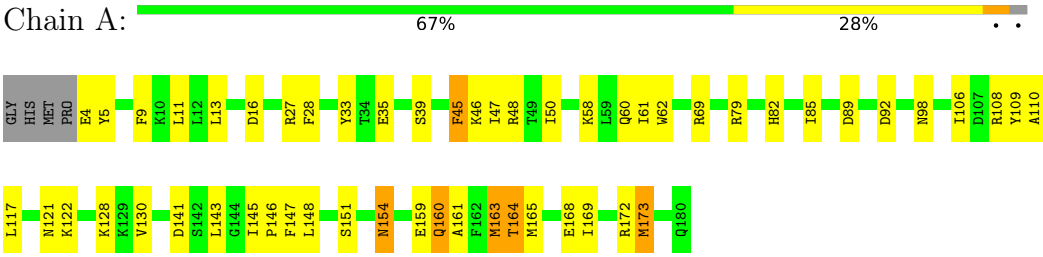
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	16	Total	O	0	0
			16	16		
7	B	23	Total	O	0	0
			23	23		
7	C	14	Total	O	0	0
			14	14		
7	D	16	Total	O	0	0
			16	16		
7	E	18	Total	O	0	0
			18	18		
7	F	26	Total	O	0	0
			26	26		
7	G	15	Total	O	0	0
			15	15		
7	H	18	Total	O	0	0
			18	18		

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.

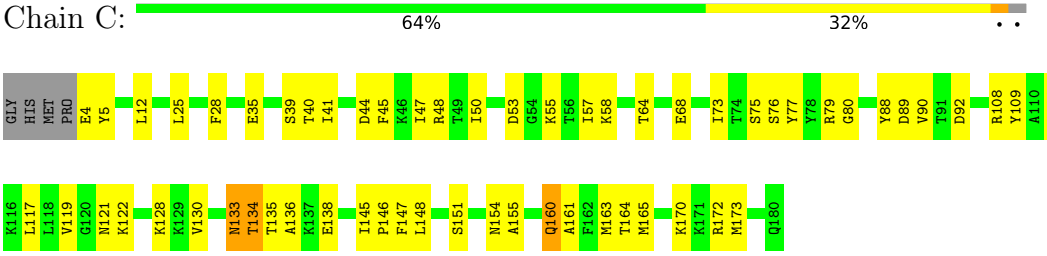
• Molecule 1: Ras-related protein Rab-1B

Chain A:



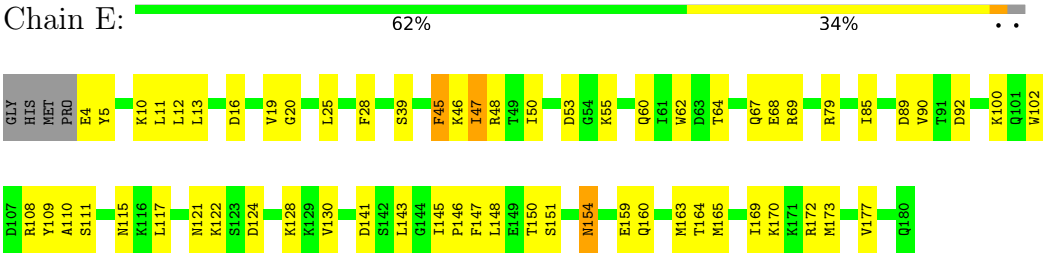
• Molecule 1: Ras-related protein Rab-1B

Chain C:



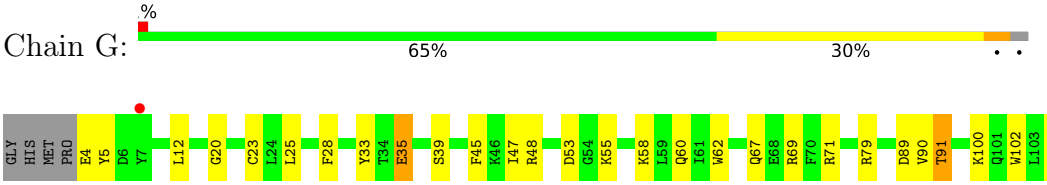
• Molecule 1: Ras-related protein Rab-1B

Chain E:



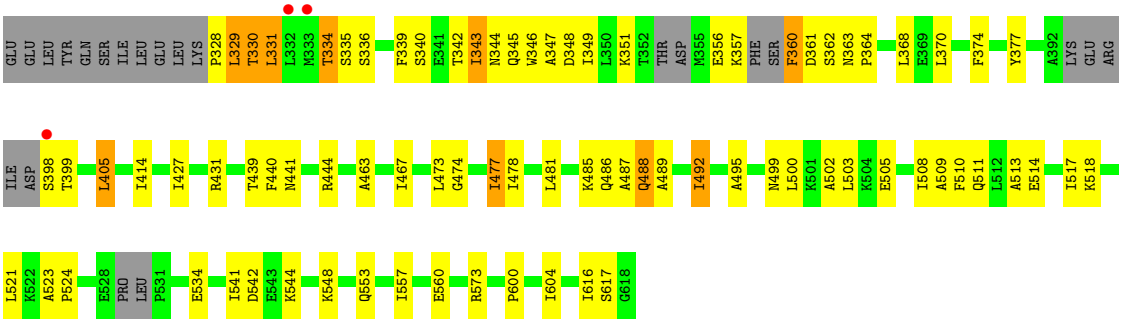
• Molecule 1: Ras-related protein Rab-1B

Chain G:



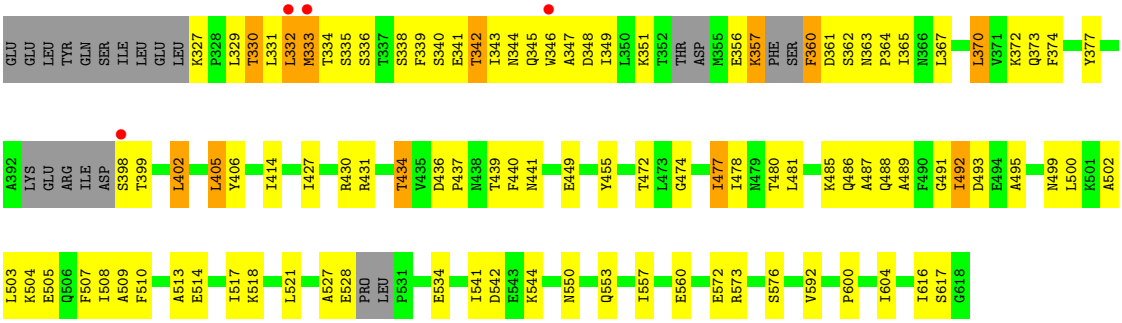
• Molecule 2: LepB

Chain B:



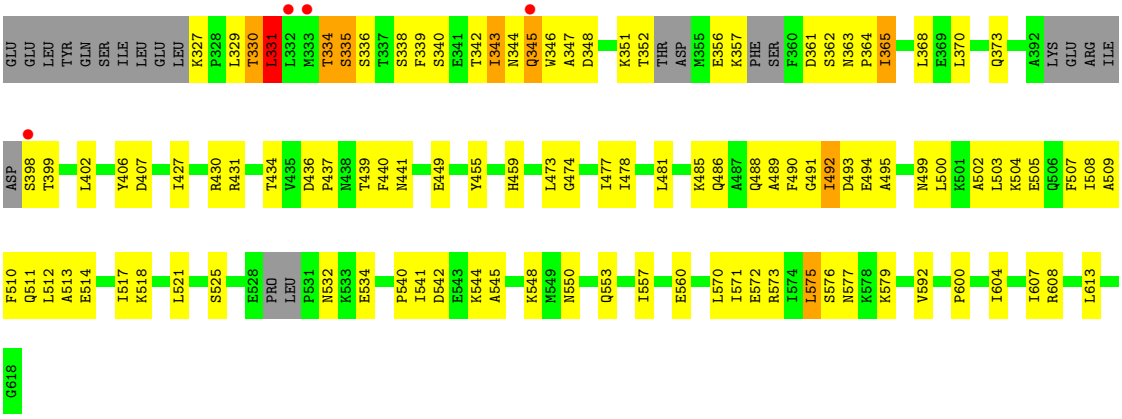
• Molecule 2: LepB

Chain D:



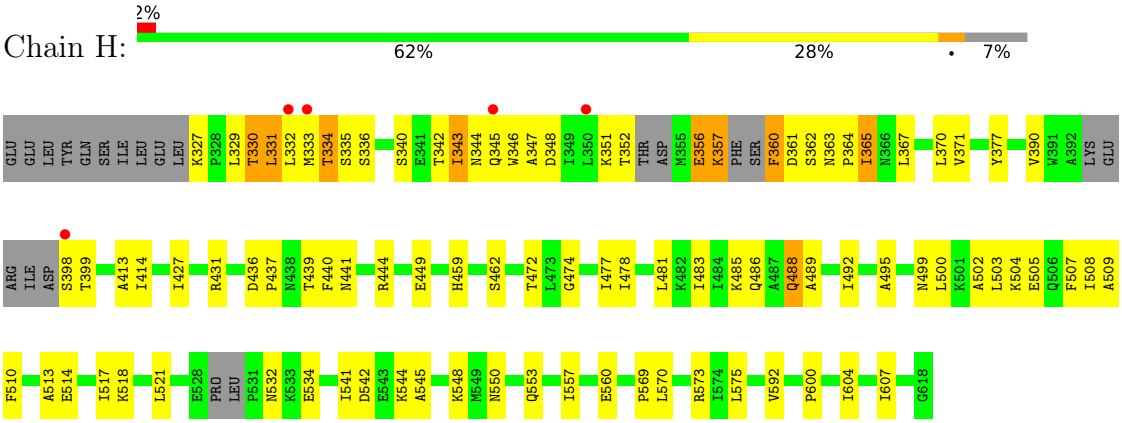
• Molecule 2: LepB

Chain F:



• Molecule 2: LepB

Chain H:


2%
62%
28%
7%
GLU GLU LEU TYR GLN SER ILE LEU LEU GLU LEU K327 P328 L329 T330 L331 L332 M333 T334 S335 S336 S340 E341 T342 I343 N344 Q345 W346 A347 D348 I349 L350 K351 T352 THR ASP M355 E356 K357 PHE SER F360 D361 S362 N363 P364 I365 N366 L367 L370 V371 Y377 V390 W391 A392 LYS GLU
ARG ILE ASP S398 T399 A413 I414 I427 R431 D436 P437 N438 T439 F440 N441 R444 E449 H459 S462 T472 L473 G474 I477 I478 L481 K482 I483 I484 K485 Q486 A487 Q488 A489 I492 A495 N499 L500 K501 A502 L503 K504 E505 Q506 F507 I508 A509
F510 A513 E514 I517 K518 L521 E528 PRO LEU P531 N532 K533 E534 I541 D542 E543 K544 A545 K548 M549 N550 Q553 I557 E560 P569 L570 R573 I574 L575 V592 P600 I604 I607 G618

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	124.41Å 124.41Å 147.47Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.16 – 2.70 49.16 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.16-2.70) 100.0 (49.16-2.70)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.28 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.217 , 0.249 0.220 , 0.250	Depositor DCC
R_{free} test set	3502 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	51.5	Xtriage
Anisotropy	0.520	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 28.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.487 for -h,-k,l 0.487 for h,-h-k,-l 0.487 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14370	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% 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: PEG, MG, GDP, BEF

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	0.65	0/1382	0.54	0/1875
1	C	0.65	0/1403	0.54	0/1900
1	E	0.63	0/1400	0.52	0/1896
1	G	0.65	0/1395	0.53	0/1890
2	B	0.68	0/2167	0.56	0/2947
2	D	0.66	0/2174	0.57	0/2956
2	F	0.67	0/2168	0.56	1/2951 (0.0%)
2	H	0.68	0/2173	0.57	0/2956
All	All	0.66	0/14262	0.55	1/19371 (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
2	B	0	1
2	F	0	1
2	H	0	1
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	331	LEU	CA-CB-CG	-5.18	103.38	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	334	THR	Peptide
2	F	334	THR	Peptide
2	H	334	THR	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	1360	0	1302	79	0
1	C	1381	0	1349	81	0
1	E	1378	0	1347	71	0
1	G	1373	0	1341	65	0
2	B	2121	0	2008	132	0
2	D	2128	0	2028	162	0
2	F	2122	0	2006	159	0
2	H	2127	0	2026	133	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
4	A	28	0	12	2	0
4	C	28	0	12	2	0
4	E	28	0	12	4	0
4	G	28	0	12	7	0
5	A	4	0	0	0	0
5	C	4	0	0	0	0
5	E	4	0	0	1	0
5	G	4	0	0	0	0
6	A	7	10	9	11	0
6	C	21	30	27	23	0
6	D	7	10	9	0	0
6	G	7	10	9	4	0
7	A	16	0	0	3	0
7	B	23	0	0	1	0
7	C	14	0	0	0	0
7	D	16	0	0	1	0
7	E	18	0	0	1	0
7	F	26	0	0	2	0
7	G	15	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	H	18	0	0	2	0
All	All	14310	60	13509	879	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 32.

The worst 5 of 879 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:329:LEU:CD2	2:B:357:LYS:HG2	1.79	1.12
2:D:348:ASP:HA	2:D:351:LYS:HE2	1.29	1.09
2:B:508:ILE:HD12	2:B:509:ALA:N	1.68	1.09
2:D:341:GLU:OE1	2:D:345:GLN:NE2	1.84	1.08
2:H:360:PHE:CE1	2:H:362:SER:HB3	1.89	1.07

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	175/181 (97%)	165 (94%)	10 (6%)	0	100	100
1	C	175/181 (97%)	167 (95%)	8 (5%)	0	100	100
1	E	175/181 (97%)	165 (94%)	10 (6%)	0	100	100
1	G	175/181 (97%)	162 (93%)	13 (7%)	0	100	100
2	B	270/302 (89%)	255 (94%)	14 (5%)	1 (0%)	34	60
2	D	271/302 (90%)	251 (93%)	20 (7%)	0	100	100
2	F	271/302 (90%)	246 (91%)	24 (9%)	1 (0%)	34	60
2	H	271/302 (90%)	248 (92%)	23 (8%)	0	100	100
All	All	1783/1932 (92%)	1659 (93%)	122 (7%)	2 (0%)	51	78

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	493	ASP
2	B	492	ILE

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	140/157 (89%)	134 (96%)	6 (4%)	29	57
1	C	146/157 (93%)	143 (98%)	3 (2%)	53	80
1	E	145/157 (92%)	138 (95%)	7 (5%)	25	53
1	G	144/157 (92%)	138 (96%)	6 (4%)	30	58
2	B	208/262 (79%)	197 (95%)	11 (5%)	22	48
2	D	209/262 (80%)	193 (92%)	16 (8%)	13	30
2	F	208/262 (79%)	192 (92%)	16 (8%)	13	30
2	H	209/262 (80%)	196 (94%)	13 (6%)	18	40
All	All	1409/1676 (84%)	1331 (94%)	78 (6%)	21	46

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	575	LEU
2	H	352	THR
1	G	35	GLU
2	H	330	THR
2	H	365	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	488	GLN
2	D	488	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 18 ligands modelled in this entry, 4 are monoatomic - leaving 14 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	$\#\|Z\| > 2$
4	GDP	E	202	5,3	24,30,30	1.27	2 (8%)	30,47,47	1.12	3 (10%)
5	BEF	A	203	4	0,3,3	-	-	-		
5	BEF	C	203	4	0,3,3	-	-	-		
5	BEF	E	203	4	0,3,3	-	-	-		
6	PEG	A	204	-	6,6,6	0.99	0	5,5,5	0.64	0
6	PEG	C	206	-	6,6,6	0.96	0	5,5,5	0.40	0
4	GDP	G	202	5,3	24,30,30	1.22	1 (4%)	30,47,47	1.18	4 (13%)
6	PEG	D	701	-	6,6,6	0.93	0	5,5,5	0.52	0
6	PEG	C	205	-	6,6,6	1.03	0	5,5,5	0.57	0
6	PEG	G	204	-	6,6,6	1.04	1 (16%)	5,5,5	0.40	0
4	GDP	A	202	5,3	24,30,30	1.21	1 (4%)	30,47,47	1.16	4 (13%)
4	GDP	C	202	5,3	24,30,30	1.22	1 (4%)	30,47,47	1.13	4 (13%)
5	BEF	G	203	4	0,3,3	-	-	-		
6	PEG	C	204	-	6,6,6	0.97	0	5,5,5	0.73	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
4	GDP	E	202	5,3	-	4/12/32/32	0/3/3/3
6	PEG	A	204	-	-	2/4/4/4	-
6	PEG	C	206	-	-	1/4/4/4	-
4	GDP	G	202	5,3	-	5/12/32/32	0/3/3/3
6	PEG	D	701	-	-	4/4/4/4	-
6	PEG	C	205	-	-	2/4/4/4	-
6	PEG	G	204	-	-	2/4/4/4	-
4	GDP	A	202	5,3	-	2/12/32/32	0/3/3/3
4	GDP	C	202	5,3	-	0/12/32/32	0/3/3/3
6	PEG	C	204	-	-	2/4/4/4	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	202	GDP	C6-N1	-3.72	1.32	1.37
4	E	202	GDP	C6-N1	-3.70	1.32	1.37
4	A	202	GDP	C6-N1	-3.67	1.32	1.37
4	G	202	GDP	C6-N1	-3.66	1.32	1.37
4	E	202	GDP	C2'-C1'	-2.07	1.50	1.53

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	202	GDP	PA-O3A-PB	-2.83	123.11	132.83
4	A	202	GDP	PA-O3A-PB	-2.63	123.81	132.83
4	E	202	GDP	C5-C6-N1	2.45	118.27	113.95
4	C	202	GDP	C5-C6-N1	2.39	118.17	113.95
4	G	202	GDP	C5-C6-N1	2.34	118.09	113.95

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	202	GDP	PA-O3A-PB-O3B
4	G	202	GDP	C5'-O5'-PA-O3A
4	G	202	GDP	C5'-O5'-PA-O1A
4	G	202	GDP	C5'-O5'-PA-O2A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	G	202	GDP	O4'-C4'-C5'-O5'

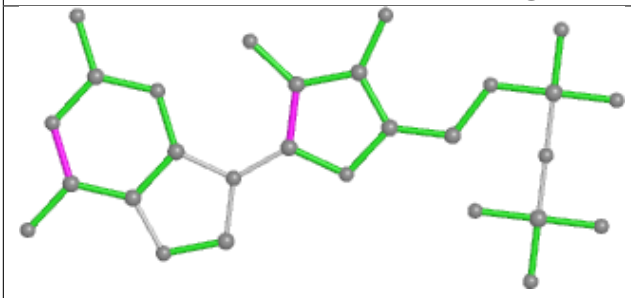
There are no ring outliers.

10 monomers are involved in 49 short contacts:

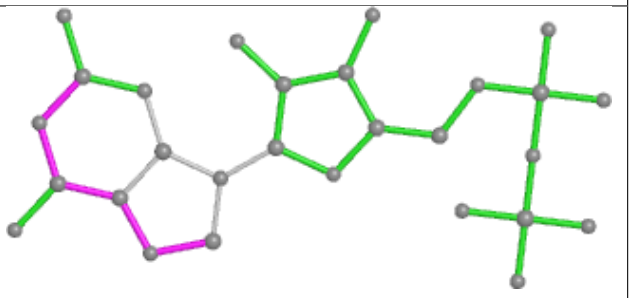
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	202	GDP	4	0
5	E	203	BEF	1	0
6	A	204	PEG	11	0
6	C	206	PEG	7	0
4	G	202	GDP	7	0
6	C	205	PEG	5	0
6	G	204	PEG	4	0
4	A	202	GDP	2	0
4	C	202	GDP	2	0
6	C	204	PEG	11	0

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.

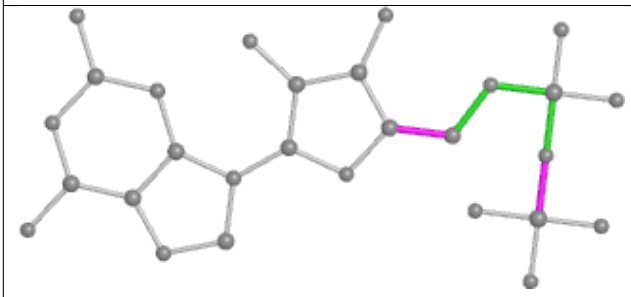
Ligand GDP E 202



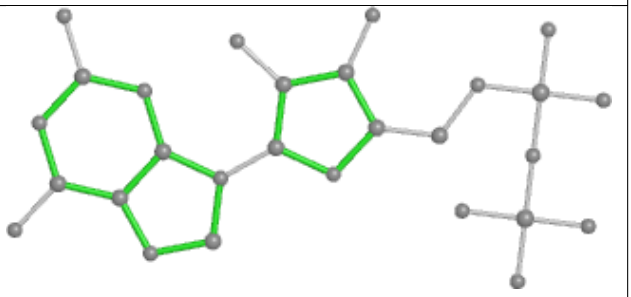
Bond lengths



Bond angles

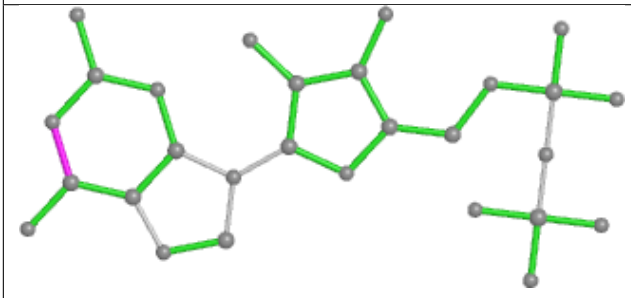


Torsions

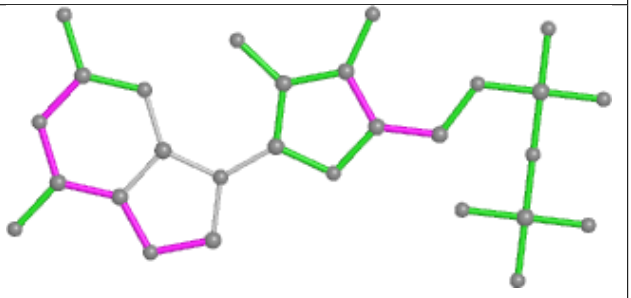


Rings

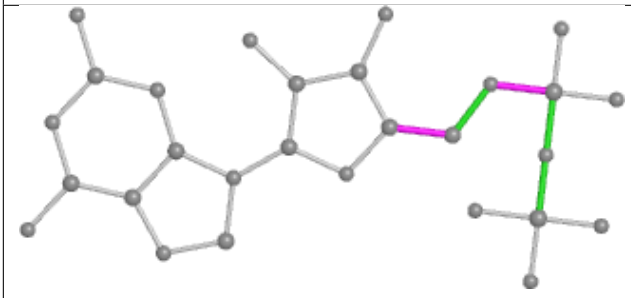
Ligand GDP G 202



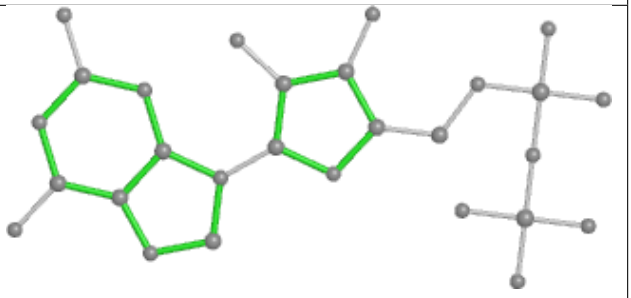
Bond lengths



Bond angles

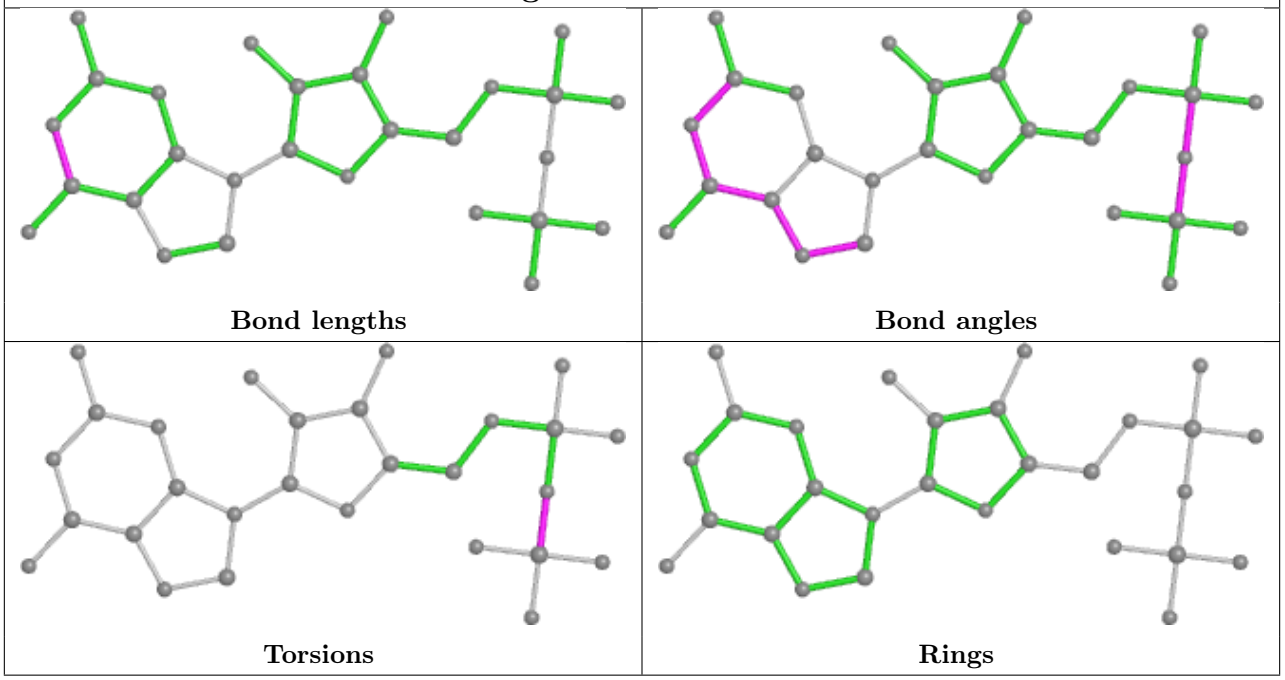


Torsions

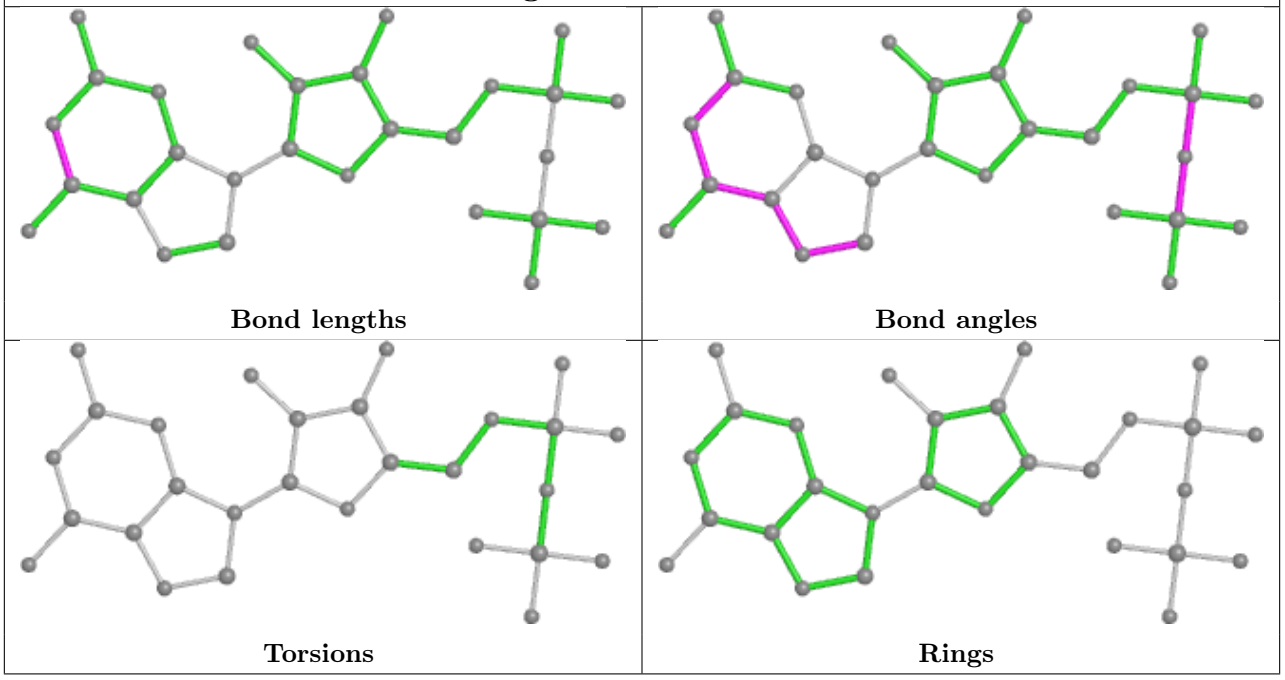


Rings

Ligand GDP A 202



Ligand GDP C 202



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	177/181 (97%)	-0.34	0	100	100	29, 48, 69, 91	1 (0%)
1	C	177/181 (97%)	-0.30	0	100	100	27, 48, 73, 91	0
1	E	177/181 (97%)	-0.29	0	100	100	28, 48, 73, 96	0
1	G	177/181 (97%)	-0.31	1 (0%)	89	91	27, 47, 72, 94	0
2	B	280/302 (92%)	-0.23	3 (1%)	80	82	23, 46, 91, 138	11 (3%)
2	D	281/302 (93%)	-0.22	4 (1%)	75	77	23, 47, 95, 141	7 (2%)
2	F	281/302 (93%)	-0.23	4 (1%)	75	77	23, 47, 95, 142	13 (4%)
2	H	281/302 (93%)	-0.22	5 (1%)	68	70	22, 46, 92, 140	17 (6%)
All	All	1831/1932 (94%)	-0.26	17 (0%)	84	85	22, 47, 88, 142	49 (2%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	333	MET	3.1
2	F	398	SER	3.0
2	H	332	LEU	2.7
2	H	350	LEU	2.6
2	D	398	SER	2.6

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands (i)

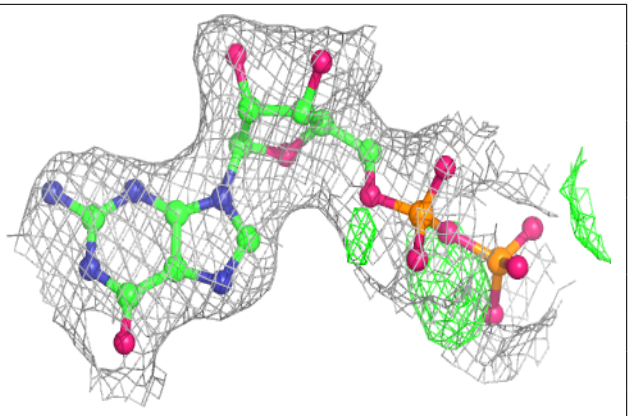
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

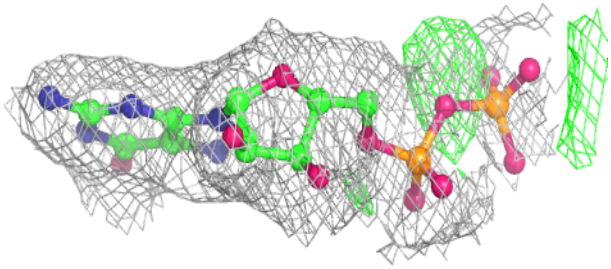
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	PEG	C	204	7/7	0.84	0.23	57,79,99,99	0
6	PEG	D	701	7/7	0.84	0.17	70,84,99,99	0
6	PEG	G	204	7/7	0.89	0.18	63,76,84,89	0
6	PEG	A	204	7/7	0.90	0.16	57,76,95,95	0
6	PEG	C	205	7/7	0.90	0.22	58,74,85,88	0
6	PEG	C	206	7/7	0.91	0.12	75,100,117,120	0
3	MG	C	201	1/1	0.92	0.10	34,34,34,34	0
5	BEF	C	203	4/4	0.92	0.20	27,45,47,59	0
5	BEF	G	203	4/4	0.92	0.24	36,38,57,137	0
5	BEF	E	203	4/4	0.94	0.22	36,55,61,66	0
3	MG	G	201	1/1	0.94	0.12	35,35,35,35	0
5	BEF	A	203	4/4	0.95	0.24	26,32,38,51	0
3	MG	E	201	1/1	0.95	0.10	32,32,32,32	0
4	GDP	A	202	28/28	0.97	0.14	15,33,58,61	0
4	GDP	E	202	28/28	0.97	0.13	18,37,52,55	0
3	MG	A	201	1/1	0.97	0.11	32,32,32,32	0
4	GDP	C	202	28/28	0.98	0.14	17,37,54,55	0
4	GDP	G	202	28/28	0.98	0.13	12,40,56,59	0

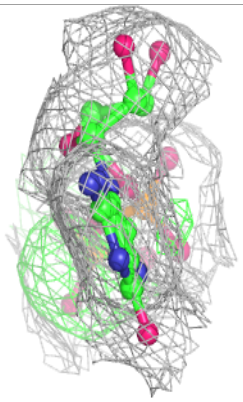
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around GDP A 202:

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

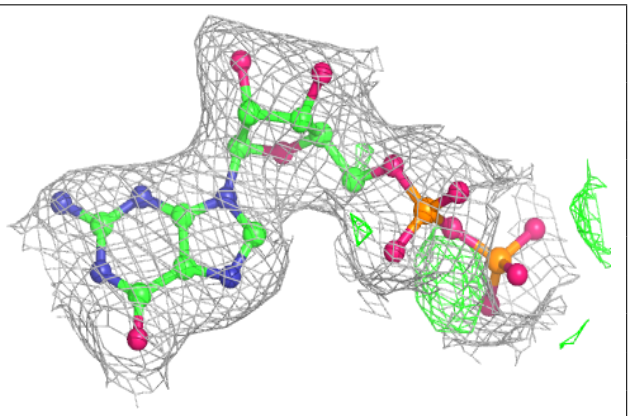


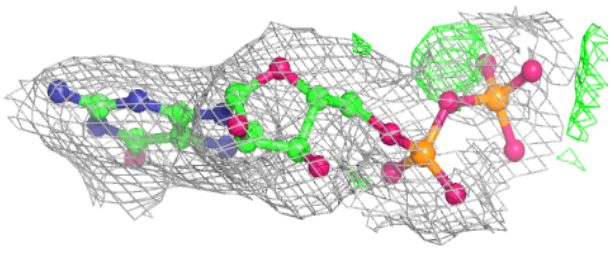


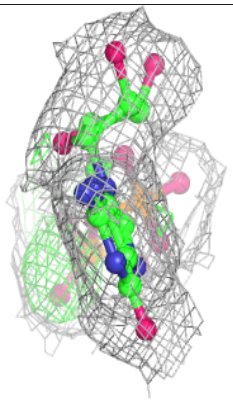


Electron density around GDP E 202:

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

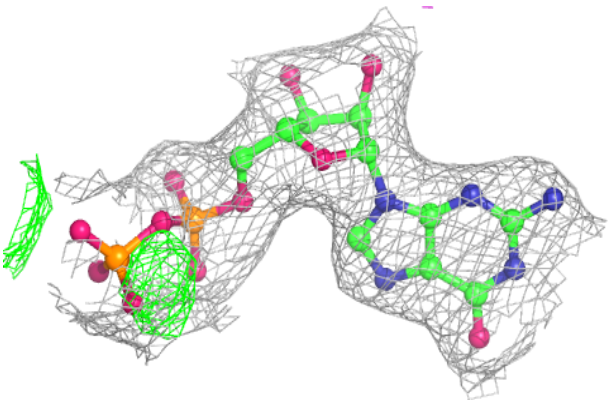


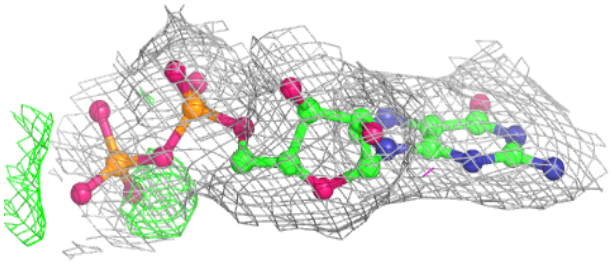


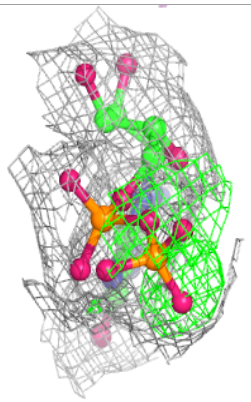


Electron density around GDP C 202:

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

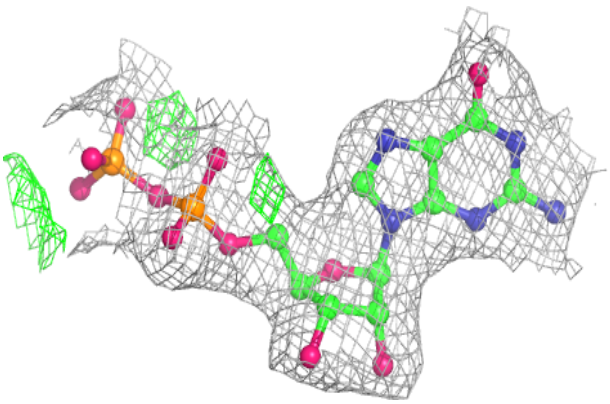


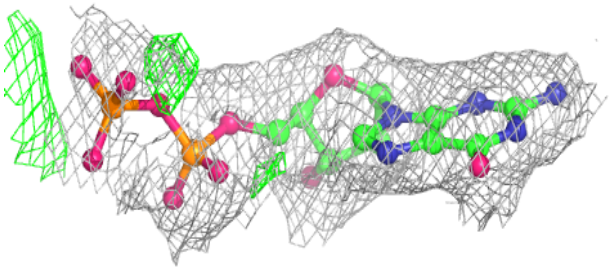


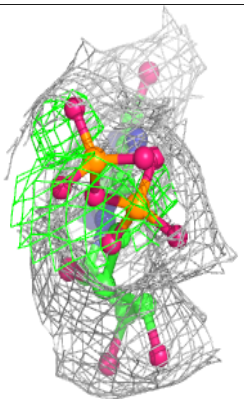


Electron density around GDP G 202:

$2mF_o$-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.