



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2017 – 03:09 PM EDT

PDB ID : 1VG9  
Title : The crystal structures of the REP-1 protein in complex with C-terminally truncated Rab7 protein  
Authors : Rak, A.; Pylypenko, O.; Niculae, A.; Pyatkov, K.; Goody, R.S.; Alexandrov, K.  
Deposited on : unknown  
Resolution : 2.50 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

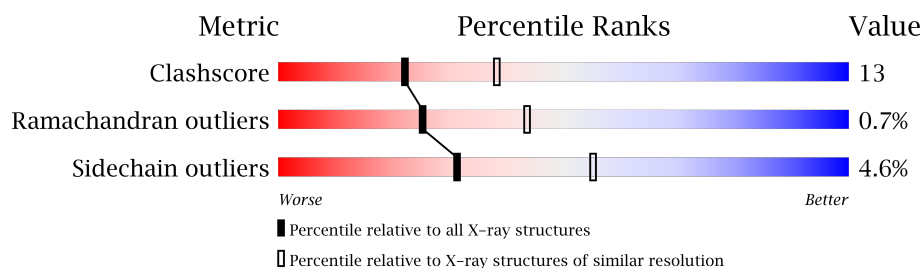
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)


The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	650	
1	C	650	
1	E	650	
1	G	650	
2	B	185	
2	D	185	
2	F	185	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	185	 A horizontal bar chart showing the quality of chain H. The bar is divided into three segments: green (59%), yellow (31%), and red (6%). The red segment is the smallest and is located at the far right of the bar. The percentages are labeled below the corresponding segments.

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 22604 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 Rab proteins geranylgeranyltransferase component A 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	501	Total	C	N	O	S	0	0	0
			3993	2524	672	767	30			
1	C	502	Total	C	N	O	S	0	0	0
			3980	2516	670	764	30			
1	E	500	Total	C	N	O	S	0	0	0
			3988	2521	671	766	30			
1	G	502	Total	C	N	O	S	0	0	0
			3984	2518	670	766	30			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	231	LYS	GLN	ENGINEERED	UNP P37727
A	462	ARG	LYS	ENGINEERED	UNP P37727
A	473	THR	ALA	ENGINEERED	UNP P37727
A	483	ALA	GLY	ENGINEERED	UNP P37727
C	231	LYS	GLN	ENGINEERED	UNP P37727
C	462	ARG	LYS	ENGINEERED	UNP P37727
C	473	THR	ALA	ENGINEERED	UNP P37727
C	483	ALA	GLY	ENGINEERED	UNP P37727
E	231	LYS	GLN	ENGINEERED	UNP P37727
E	462	ARG	LYS	ENGINEERED	UNP P37727
E	473	THR	ALA	ENGINEERED	UNP P37727
E	483	ALA	GLY	ENGINEERED	UNP P37727
G	231	LYS	GLN	ENGINEERED	UNP P37727
G	462	ARG	LYS	ENGINEERED	UNP P37727
G	473	THR	ALA	ENGINEERED	UNP P37727
G	483	ALA	GLY	ENGINEERED	UNP P37727

- Molecule 2 is a protein called Ras-related protein Rab-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total	C	N	O	S	0	0	0
			1335	849	225	255	6			
2	D	173	Total	C	N	O	S	0	0	0
			1371	870	233	262	6			
2	F	168	Total	C	N	O	S	0	0	0
			1332	847	227	252	6			
2	H	173	Total	C	N	O	S	0	0	0
			1371	870	233	262	6			

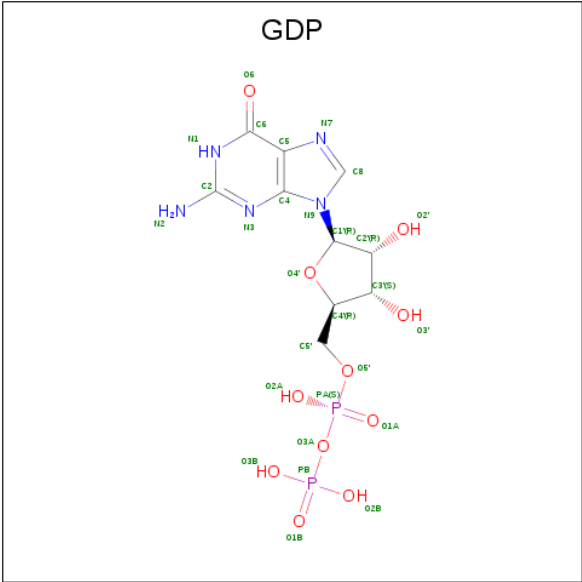
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	H	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	D	1	Total	Mg	0	0
			1	1		
3	F	1	Total	Mg	0	0
			1	1		

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

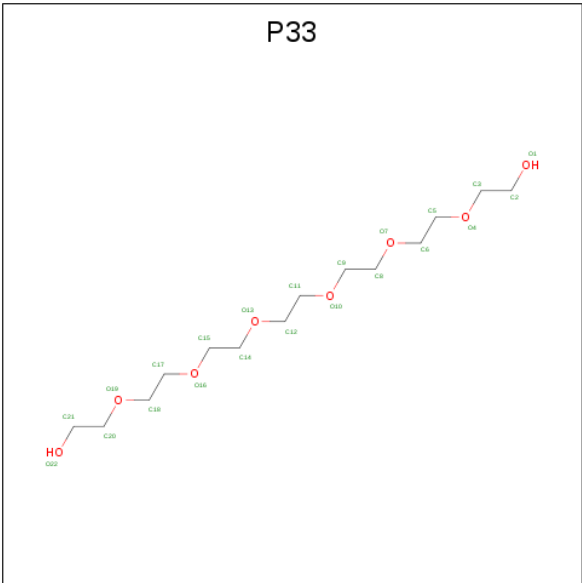
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	1	Total	K	0	0
			1	1		
4	B	1	Total	K	0	0
			1	1		
4	D	1	Total	K	0	0
			1	1		
4	F	1	Total	K	0	0
			1	1		

- Molecule 5 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>11</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
5	D	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
5	F	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
5	H	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 6 is 3,6,9,12,15,18-HEXAOSAICOSANE-1,20-DIOL (three-letter code: P33) (formula: C<sub>14</sub>H<sub>30</sub>O<sub>8</sub>).



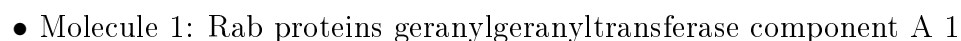
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			22	14	8		
6	D	1	Total	C	O	0	0
			22	14	8		
6	F	1	Total	C	O	0	0
			22	14	8		
6	H	1	Total	C	O	0	0
			22	14	8		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	223	Total	O	0	0
			223	223		
7	B	62	Total	O	0	0
			62	62		
7	C	164	Total	O	0	0
			164	164		
7	D	63	Total	O	0	0
			63	63		
7	E	222	Total	O	0	0
			222	222		
7	F	69	Total	O	0	0
			69	69		
7	G	171	Total	O	0	0
			171	171		
7	H	68	Total	O	0	0
			68	68		

These plots are drawn for all protein, RNA and DNA 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 ( $\text{RSRZ} > 2$ ). 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: Rab proteins geranylgeranyltransferase component A 1

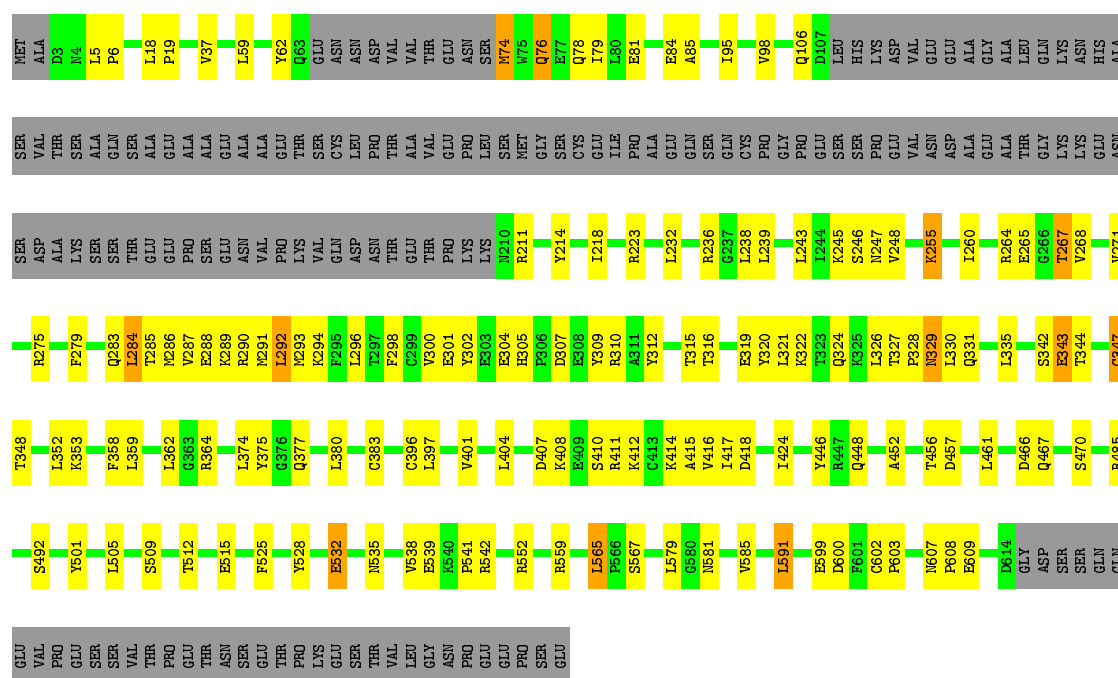






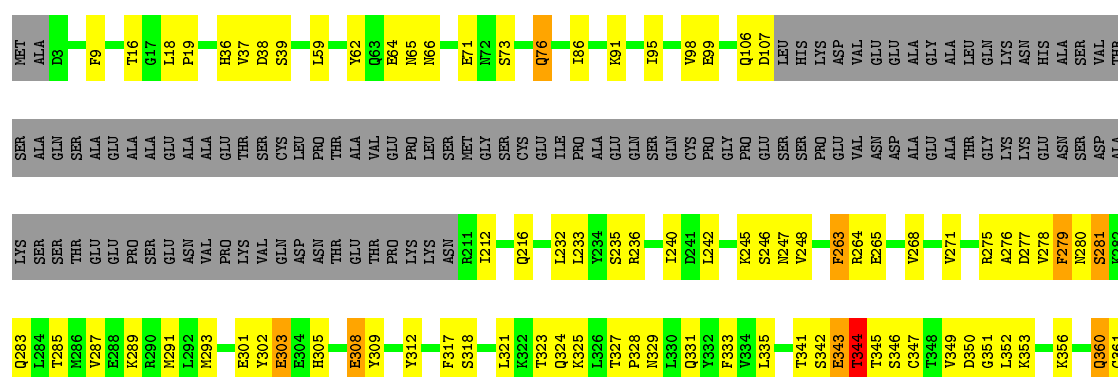
• Molecule 1: Rab proteins geranylgeranyltransferase component A 1

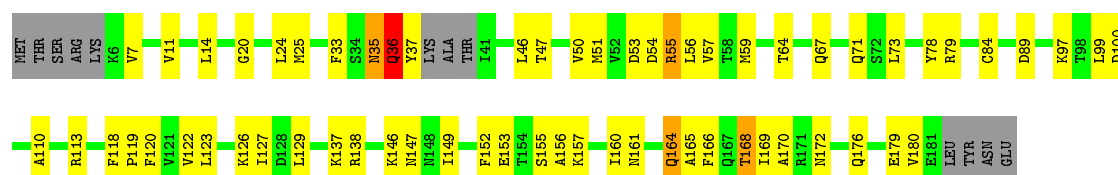
Chain E: 54% 21% 23%



• Molecule 1: Rab proteins geranylgeranyltransferase component A 1

Chain G: 54% 21% 23%





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.39 Å   144.69 Å   200.09 Å 90.00°   90.13°   90.00°	Depositor
Resolution (Å)	19.42 – 2.50	Depositor
% Data completeness (in resolution range)	99.7 (19.42-2.50)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.206 , 0.250	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	22604	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GDP, P33, K, MG

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	0.47	0/4071	0.67	0/5508
1	C	0.46	0/4058	0.66	0/5493
1	E	0.46	0/4066	0.66	0/5501
1	G	0.45	0/4062	0.66	0/5498
2	B	0.46	0/1359	0.64	0/1840
2	D	0.47	0/1395	0.63	0/1888
2	F	0.44	0/1356	0.61	0/1835
2	H	0.47	0/1395	0.65	0/1888
All	All	0.46	0/21762	0.66	0/29451

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 [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	3993	0	3934	116	0
1	C	3980	0	3912	90	0
1	E	3988	0	3929	125	0
1	G	3984	0	3916	101	0
2	B	1335	0	1306	23	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1371	0	1342	32	0
2	F	1332	0	1311	29	0
2	H	1371	0	1342	50	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	H	1	0	0	0	0
5	B	28	0	12	1	0
5	D	28	0	12	0	0
5	F	28	0	12	0	0
5	H	28	0	12	2	0
6	B	22	0	30	5	0
6	D	22	0	30	1	0
6	F	22	0	30	2	0
6	H	22	0	30	1	0
7	A	223	0	0	3	0
7	B	62	0	0	1	0
7	C	164	0	0	0	0
7	D	63	0	0	0	0
7	E	222	0	0	3	0
7	F	69	0	0	2	0
7	G	171	0	0	2	0
7	H	68	0	0	2	0
All	All	22604	0	21160	562	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 13.

The worst 5 of 562 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:316:THR:HA	1:A:347:CYS:HA	1.29	1.05
1:A:600:ASP:HB3	1:A:603:PRO:HG3	1.43	1.00
1:E:600:ASP:HB3	1:E:603:PRO:HG3	1.54	0.89
1:E:316:THR:HA	1:E:347:CYS:HA	1.56	0.87
1:A:327:THR:HB	1:A:328:PRO:HD2	1.56	0.85

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	495/650 (76%)	456 (92%)	32 (6%)	7 (1%)	13	23
1	C	496/650 (76%)	454 (92%)	38 (8%)	4 (1%)	22	39
1	E	494/650 (76%)	461 (93%)	30 (6%)	3 (1%)	28	48
1	G	496/650 (76%)	462 (93%)	32 (6%)	2 (0%)	38	59
2	B	165/185 (89%)	160 (97%)	5 (3%)	0	100	100
2	D	169/185 (91%)	160 (95%)	9 (5%)	0	100	100
2	F	164/185 (89%)	156 (95%)	8 (5%)	0	100	100
2	H	169/185 (91%)	159 (94%)	8 (5%)	2 (1%)	15	27
All	All	2648/3340 (79%)	2468 (93%)	162 (6%)	18 (1%)	25	43

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	310	ARG
1	A	467	GLN
1	A	3	ASP
1	C	309	TYR
1	E	310	ARG

### 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	452/581 (78%)	434 (96%)	18 (4%)	36	62
1	C	450/581 (78%)	439 (98%)	11 (2%)	54	80
1	E	452/581 (78%)	430 (95%)	22 (5%)	29	52
1	G	451/581 (78%)	429 (95%)	22 (5%)	29	52
2	B	144/162 (89%)	135 (94%)	9 (6%)	21	38
2	D	148/162 (91%)	136 (92%)	12 (8%)	14	26
2	F	144/162 (89%)	138 (96%)	6 (4%)	34	59
2	H	148/162 (91%)	139 (94%)	9 (6%)	22	40
All	All	2389/2972 (80%)	2280 (95%)	109 (5%)	31	55

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	179	GLU
1	E	329	ASN
2	H	35	ASN
1	E	74	MET
1	E	279	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	148	ASN
1	E	329	ASN
2	H	26	ASN
1	E	210	ASN
1	E	467	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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	GDP	B	2302	3	25,30,30	1.90	3 (12%)	26,47,47	2.76	11 (42%)
6	P33	B	3302	4	21,21,21	0.51	0	20,20,20	0.57	0
5	GDP	D	2300	3	25,30,30	1.95	3 (12%)	26,47,47	2.70	8 (30%)
6	P33	D	3304	4	21,21,21	0.51	0	20,20,20	0.53	0
5	GDP	F	2305	3	25,30,30	2.01	3 (12%)	26,47,47	2.74	11 (42%)
6	P33	F	3306	4	21,21,21	0.51	0	20,20,20	0.57	0
5	GDP	H	2303	3	25,30,30	1.92	3 (12%)	26,47,47	2.81	10 (38%)
6	P33	H	3308	4	21,21,21	0.54	0	20,20,20	0.53	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
5	GDP	B	2302	3	-	0/12/32/32	0/3/3/3
6	P33	B	3302	4	-	0/19/19/19	0/0/0/0
5	GDP	D	2300	3	-	0/12/32/32	0/3/3/3
6	P33	D	3304	4	-	0/19/19/19	0/0/0/0
5	GDP	F	2305	3	-	0/12/32/32	0/3/3/3
6	P33	F	3306	4	-	0/19/19/19	0/0/0/0
5	GDP	H	2303	3	-	0/12/32/32	0/3/3/3
6	P33	H	3308	4	-	0/19/19/19	0/0/0/0

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



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	2303	GDP	C2'-C1'	-2.52	1.49	1.53
5	F	2305	GDP	C2'-C1'	-2.31	1.50	1.53
5	D	2300	GDP	C2'-C1'	-2.24	1.50	1.53
5	B	2302	GDP	C2-N1	2.08	1.39	1.35
5	H	2303	GDP	C6-N1	5.03	1.42	1.33

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	H	2303	GDP	C5-C6-N1	-8.11	111.94	123.48
5	F	2305	GDP	C5-C6-N1	-8.00	112.09	123.48
5	D	2300	GDP	C5-C6-N1	-7.83	112.33	123.48
5	B	2302	GDP	C5-C6-N1	-7.81	112.36	123.48
5	F	2305	GDP	C1'-N9-C4	-3.75	120.16	126.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	2302	GDP	1	0
6	B	3302	P33	5	0
6	D	3304	P33	1	0
6	F	3306	P33	2	0
5	H	2303	GDP	2	0
6	H	3308	P33	1	0

## 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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.