WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2023 – 10:57 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	:	2004-04-23
Resolution	:	2.50 Å(reported)

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

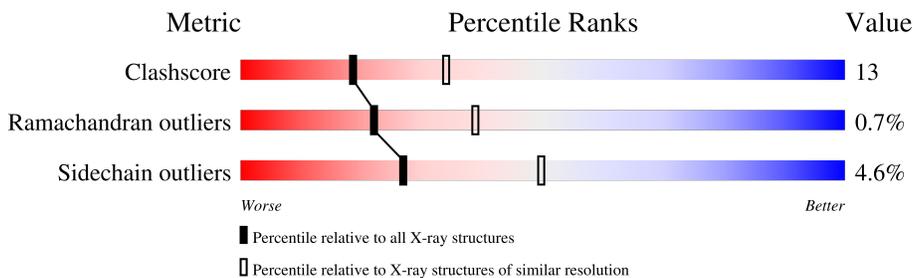|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1　Overall quality at a glance ⓘ

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	Percentile Ranks	Value
Clashscore		13
Ramachandran outliers		0.7%
Sidechain outliers		4.6%

Worse　　　　　　　　　　　　　　　　*Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	650	54%　　22%　•　23%
1	C	650	57%　　19%　•　23%
1	E	650	54%　　21%　•　23%
1	G	650	54%　　21%　•　23%
2	B	185	72%　　18%　•　9%
2	D	185	69%　　22%　•　6%
2	F	185	68%　　22%　•　9%
2	H	185	59%　　31%　••　6%

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 3993	C 2524	N 672	O 767	S 30	0	0	0
1	C	502	Total 3980	C 2516	N 670	O 764	S 30	0	0	0
1	E	500	Total 3988	C 2521	N 671	O 766	S 30	0	0	0
1	G	502	Total 3984	C 2518	N 670	O 766	S 30	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

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

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

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

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

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

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

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

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

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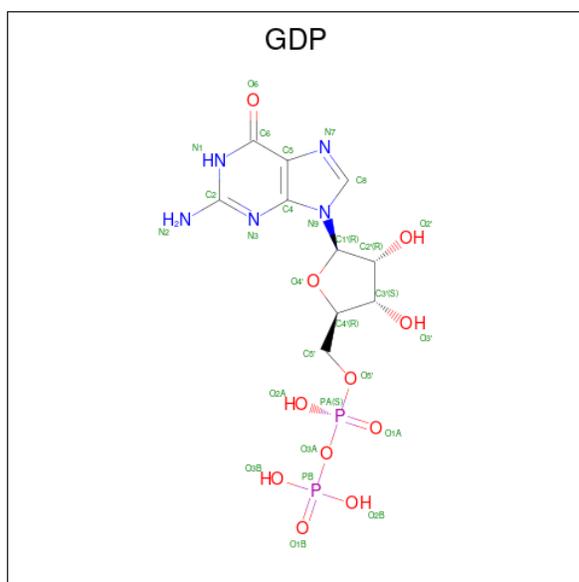
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
5	D	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
5	F	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
5	H	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 6 is 3,6,9,12,15,18-HEXAOXAICOSANE-1,20-DIOL (three-letter code: P33) (formula: $C_{14}H_{30}O_8$).

P33
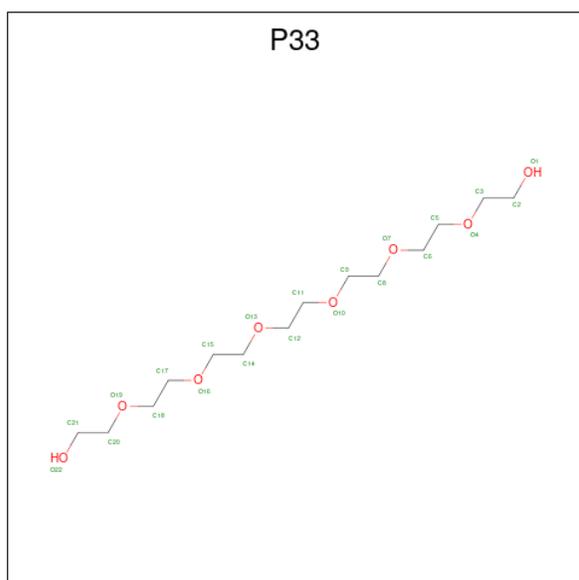
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total 22	C 14	O 8	0	0
6	D	1	Total 22	C 14	O 8	0	0
6	F	1	Total 22	C 14	O 8	0	0
6	H	1	Total 22	C 14	O 8	0	0

• Molecule 7 is water.

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

3 Residue-property plots ⓘ

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. 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. 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.

Note EDS was not executed.

• Molecule 1: Rab proteins geranylgeranyltransferase component A 1

• Molecule 1: Rab proteins geranylgeranyltransferase component A 1

N280 S281 K282 Q283 L284 T285 M286 V287 E288 K289 R290 M291 L292 M293 L296 T297 F298 Y302 H305 P306 D307 E308 Y309 R310 F317 S318 K322 K325 L326 T327 Q331 Y332 F333 V334 L335 H336 S337 C347 D350 G351 K356 Q360 C361 L362 G363 R364 T368 P369

L374 Y375 L380 C383 F391 C396 V401 L404 K408 E409 S410 R411 A415 D418 I424 I425 S426 N439 V444 Q445 Y446 A452 T456 D457 R462 T463 E477 E478 S481 V486 L505 T512 E515 R519 K523 L524

F525 E530 ILE ALA GLU ASN GLU GLN VAL E539 R542 L547 D560 L579 K586 Q587 A588 E589 T590 L591 C596 E599 D600 F601 C602 P603 A604 P605 P606 G615 ASP SER SER GLN GLU VAL PRO GLU SER VAL THR PRO GLU THR ASN SER GLU THR

PRO LYS GLU SER THR VAL LEU GLY ASN PRO GLU GLU PRO SER GLU

• Molecule 1: Rab proteins geranylgeranyltransferase component A 1

Chain E: 54% 21% • 23%

MET ALA D3 N4 L5 P6 L18 P19 V37 L59 Y62 Q63 GLU ASN ASN ASP VAL VAL THR GLU ASN SER M74 W75 Q76 E77 Q78 I79 L80 E81 E84 A85 I95 V98 Q106 D107 LEU HIS LYS ASP VAL GLU GLU ALA GLY ALA LEU GLN LYS ASN HIS ALA

SER VAL THR SER ALA GLN SER ALA GLU ALA GLU ALA ALA GLU ALA GLU THR SER CYS LEU PRO THR ALA VAL GLU PRO LEU SER MET GLY SER CYS GLU ILE PRO ALA GLU GLN GLN CYS PRO GLY PRO GLU SER SER PRO GLY VAL ASN ASP ALA GLU ALA THR GLY LYS LYS LYS GLU ASN

SER ASP ALA LYS SER SER THR GLU GLU PRO SER GLU ASN VAL PRO LYS VAL GLN ASP ASN THR GLU THR PRO LYS N210 R211 Y214 I218 R223 L232 R236 G237 L238 L239 L243 I244 K245 S246 N247 V248 K255 I260 R264 E265 G266 T267 V268 V271

R275 F279 Q283 L284 T285 M286 V287 E288 K289 R290 M291 L292 M293 K294 F295 L296 T297 F298 V300 E301 Y302 E303 E304 H305 P306 D307 E308 Y309 R310 A311 Y312 T315 T316 E319 Y320 L321 K322 T323 Q324 K325 L326 T327 P328 N329 L330 Q331 L335 S342 E343 T344 C347

T348 L352 K353 F358 L359 L362 G363 R364 L374 Y375 G376 Q377 L380 C383 C396 L397 V401 L404 D407 K408 E409 S410 R411 K412 C413 K414 A415 V416 I417 D418 I424 Y446 R447 Q448 A452 T456 D457 L461 D466 Q467 S470 R485

S492 Y501 L505 S509 T512 E515 F525 Y528 E532 N535 V538 E539 K540 P541 R542 R552 R559 L565 P566 S567 L579 G580 N581 V585 L591 E599 D600 F601 C602 P603 N607 P608 E609 D614 GLY ASP SER SER GLN GLN

GLU VAL PRO GLU SER SER VAL THR PRO GLU GLU THR ASN SER GLU GLU THR PRO LYS GLU SER THR VAL LEU GLY ASN PRO GLU GLU PRO SER GLU

• Molecule 1: Rab proteins geranylgeranyltransferase component A 1

Chain G: 54% 21% • 23%

MET ALA D3 F9 T16 G17 L18 P19 H36 V37 D38 S39 L59 Y62 Q63 E64 N65 N66 E71 N72 S73 Q76 I86 K91 I95 V98 E99 Q106 D107 LEU HIS LYS ASP VAL GLU GLU ALA GLY ALA LEU GLN LYS ASN HIS ALA SER VAL THR

SER ALA GLN SER ALA GLU ALA ALA GLU ALA ALA GLU ALA GLU THR SER CYS LEU PRO THR ALA VAL GLU PRO LEU LEU MET GLY SER CYS GLU ILE PRO ALA GLU GLN SER GLN PRO GLY PRO GLU SER SER PRO GLU VAL ASN ASP ALA GLU ALA THR GLY LYS LYS LYS GLU ASN HIS ALA SER VAL ASP ALA

LYS SER SER THR GLU GLU PRO SER GLU ASN VAL PRO LYS VAL GLN ASP ASN THR GLU THR PRO LYS LYS ASN R211 I212 Q216 L232 L233 Y234 S235 R236 I240 D241 L242 K245 S246 N247 V248 F263 R264 E265 V268 V271 R275 A276 D277 V278 F279 N280 S281 K282

Q283 L284 T285 M286 V287 E288 K289 R290 M291 L292 M293 E301 Y302 E303 E304 H305 E308 Y309 Y312 F317 S318 L321 K322 T323 Q324 K325 L326 T327 P328 N329 L330 Q331 Y332 F333 V334 L335 T341 S342 E343 T344 T345 S346 C347 T348 V349 D350 G351 L352 K353 K356 Q360 C361

L362 G363 R364 Y365 T368 P369 L374 Y375 G376 Q377 C383 F391 Y395 C396 S400 V401 L404 S410 R411 A415 D418 I424 I425 S426 K427 H428 F429 N439 V444 Q445 Y446 A452 I455 T456 D457 R462 T463 E477 E478 S481 F482

R485 L505 M508 T512 E515 K523 L524 F525 E530 ILE GLU ALA GLU ASN GLU GLN VAL E539 R542 W545 R552 R559 L565 L579 K586 Q587 A588 E589 T590 L591 F592 Q593 C596 E599 C602 P603 A604 P605 P606 N607 P608 E609 G615

ASP SER SER GLN GLN GLU VAL PRO GLU SER SER VAL THR PRO GLU THR ASN SER GLU GLU THR PRO LYS GLU SER THR VAL LEU GLY ASN PRO GLU GLU PRO SER GLU

• Molecule 2: Ras-related protein Rab-7

Chain B:

72%   18% • 9%

MET THR SER ARG LYS K6 L14 L24 N30 K31 K32 N35 GLN TYR LYS ALA THR I41 E49 V52 T64 R69 F70 Q71 Y78 R79 L99 D100 S101 D104 I108 Q109 A110 S111 F118 P119 F120 L123 K126 I127 D128 L129 I149

E153 K157 E158 N161 Q164 R171 N172 K175 Q176 E179 VAL GLU LEU TYR ASN GLU
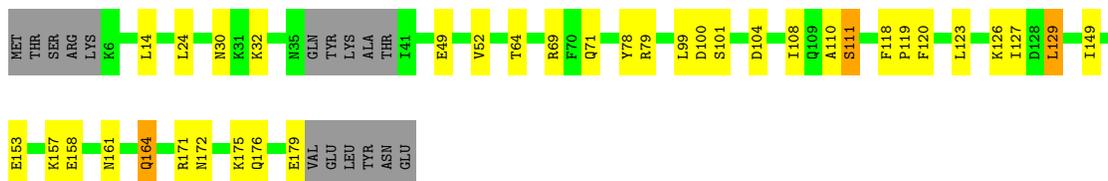
• Molecule 2: Ras-related protein Rab-7

Chain D:

69%   22% • 6%

MET THR SER ARG LYS K6 V7 D16 V19 M25 F33 S34 N35 Q36 Y37 LYS ALA THR I41 L46 V50 D53 D54 R55 L56 V57 G66 R69 F70 Q71 Y78 R79 L99 D100 A110 S111 P112 R113 F118 P119 F120 V121 V122 I127 D128 L129

T136 Q140 K146 I149 F152 E153 T154 S155 A156 K157 I160 N161 Q164 T168 N172 Q176 E177 T178 E179 V180 E181 LEU TYR ASN GLU
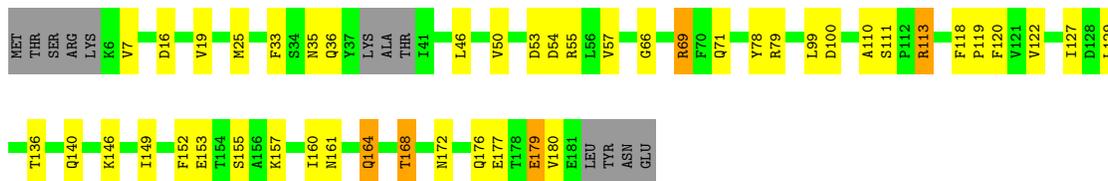
• Molecule 2: Ras-related protein Rab-7

Chain F:

68%   22% • 9%

MET THR SER ARG LYS K6 L14 T22 N30 K31 K32 F33 S34 N35 GLN TYR LYS ALA THR I41 K48 E49 V52 D53 D54 R55 M59 Q60 I61 W62 D63 Q67 S72 L73 C84 V85 L86 V87 F88 D89 A92 P93 T98 L99 D100 S101 P119 F120

L123 I127 D128 L129 K137 K146 N147 N148 I149 E153 N161 Q164 I169 A170 R171 N172 A173 L174 K175 Q176 E177 T178 GLU VAL GLU LEU TYR ASN GLU
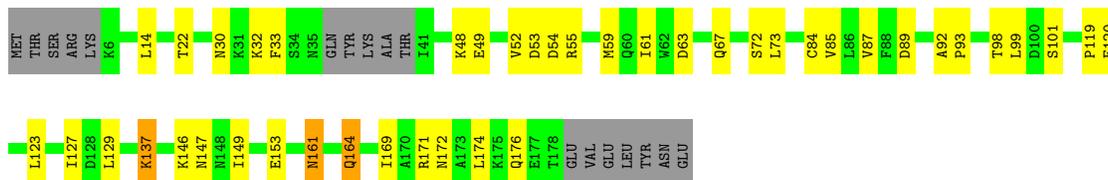
• Molecule 2: Ras-related protein Rab-7

Chain H:

59%   31% •• 6%

MET THR SER ARG LYS K6 V7 V11 L14 G20 L24 M25 F33 S34 N35 Q36 Y37 LYS ALA THR I41 L46 T47 V50 M51 V52 D53 D54 R55 L56 V57 T58 M59 T64 Q67 Q71 S72 L73 Y78 R79 C84 D89 K97 T98 L99 D100

A110 R113 F118 P119 F120 V121 V122 L123 K126 I127 D128 L129 K137 R138 K146 N147 N148 I149 F152 E153 T154 S155 A156 K157 I160 N161 Q164 A165 F166 Q167 T168 I169 A170 R171 N172 Q176 E179 V180 E181 LEU TYR ASN GLU

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, α, β, γ	66.39Å 144.69Å 200.09Å 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^2)	35.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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 ⓘ

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 ⓘ

5.3.1 Protein backbone ⓘ

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%)	11	20
1	C	496/650 (76%)	454 (92%)	38 (8%)	4 (1%)	19	35
1	E	494/650 (76%)	461 (93%)	30 (6%)	3 (1%)	25	43
1	G	496/650 (76%)	462 (93%)	32 (6%)	2 (0%)	34	54
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%)	13	24
All	All	2648/3340 (79%)	2468 (93%)	162 (6%)	18 (1%)	22	39

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 ⓘ

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%)	31	56
1	C	450/581 (78%)	439 (98%)	11 (2%)	49	74
1	E	452/581 (78%)	430 (95%)	22 (5%)	25	47
1	G	451/581 (78%)	429 (95%)	22 (5%)	25	47
2	B	144/162 (89%)	135 (94%)	9 (6%)	18	34
2	D	148/162 (91%)	136 (92%)	12 (8%)	11	23
2	F	144/162 (89%)	138 (96%)	6 (4%)	30	54
2	H	148/162 (91%)	139 (94%)	9 (6%)	18	36
All	All	2389/2972 (80%)	2280 (95%)	109 (5%)	27	50

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

Mol	Chain	Res	Type
1	E	284	LEU
1	E	599	GLU
1	G	607	ASN
1	E	290	ARG
1	E	470	SER

Sometimes 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	172	ASN
2	H	26	ASN
1	E	467	GLN
1	G	607	ASN
2	H	172	ASN

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 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	H	2303	3	24,30,30	1.26	2 (8%)	30,47,47	1.38	6 (20%)
6	P33	F	3306	4	21,21,21	0.51	0	20,20,20	0.58	0
6	P33	H	3308	4	21,21,21	0.54	0	20,20,20	0.54	0
5	GDP	F	2305	3	24,30,30	1.42	3 (12%)	30,47,47	1.42	6 (20%)
5	GDP	B	2302	3	24,30,30	1.37	4 (16%)	30,47,47	1.41	6 (20%)
6	P33	D	3304	4	21,21,21	0.51	0	20,20,20	0.54	0
5	GDP	D	2300	3	24,30,30	1.33	3 (12%)	30,47,47	1.36	3 (10%)
6	P33	B	3302	4	21,21,21	0.51	0	20,20,20	0.59	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	H	2303	3	-	1/12/32/32	0/3/3/3
6	P33	F	3306	4	-	6/19/19/19	-
6	P33	H	3308	4	-	7/19/19/19	-
5	GDP	F	2305	3	-	1/12/32/32	0/3/3/3
5	GDP	B	2302	3	-	1/12/32/32	0/3/3/3
6	P33	D	3304	4	-	7/19/19/19	-
5	GDP	D	2300	3	-	1/12/32/32	0/3/3/3
6	P33	B	3302	4	-	8/19/19/19	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	2302	GDP	C6-N1	3.59	1.43	1.37
5	F	2305	GDP	C6-N1	3.52	1.43	1.37
5	D	2300	GDP	C6-N1	3.02	1.42	1.37
5	H	2303	GDP	C6-N1	2.87	1.42	1.37
5	H	2303	GDP	C2'-C1'	-2.72	1.49	1.53

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	2305	GDP	PA-O3A-PB	-3.19	121.88	132.83
5	B	2302	GDP	PA-O3A-PB	-3.11	122.16	132.83
5	D	2300	GDP	PA-O3A-PB	-2.85	123.03	132.83
5	F	2305	GDP	O3B-PB-O2B	-2.51	98.05	107.64
5	H	2303	GDP	PA-O3A-PB	-2.50	124.23	132.83

There are no chirality outliers.

5 of 32 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	3302	P33	O13-C14-C15-O16
6	F	3306	P33	O7-C8-C9-O10
6	H	3308	P33	O13-C14-C15-O16
6	H	3308	P33	O7-C8-C9-O10
6	D	3304	P33	O13-C14-C15-O16

There are no ring outliers.

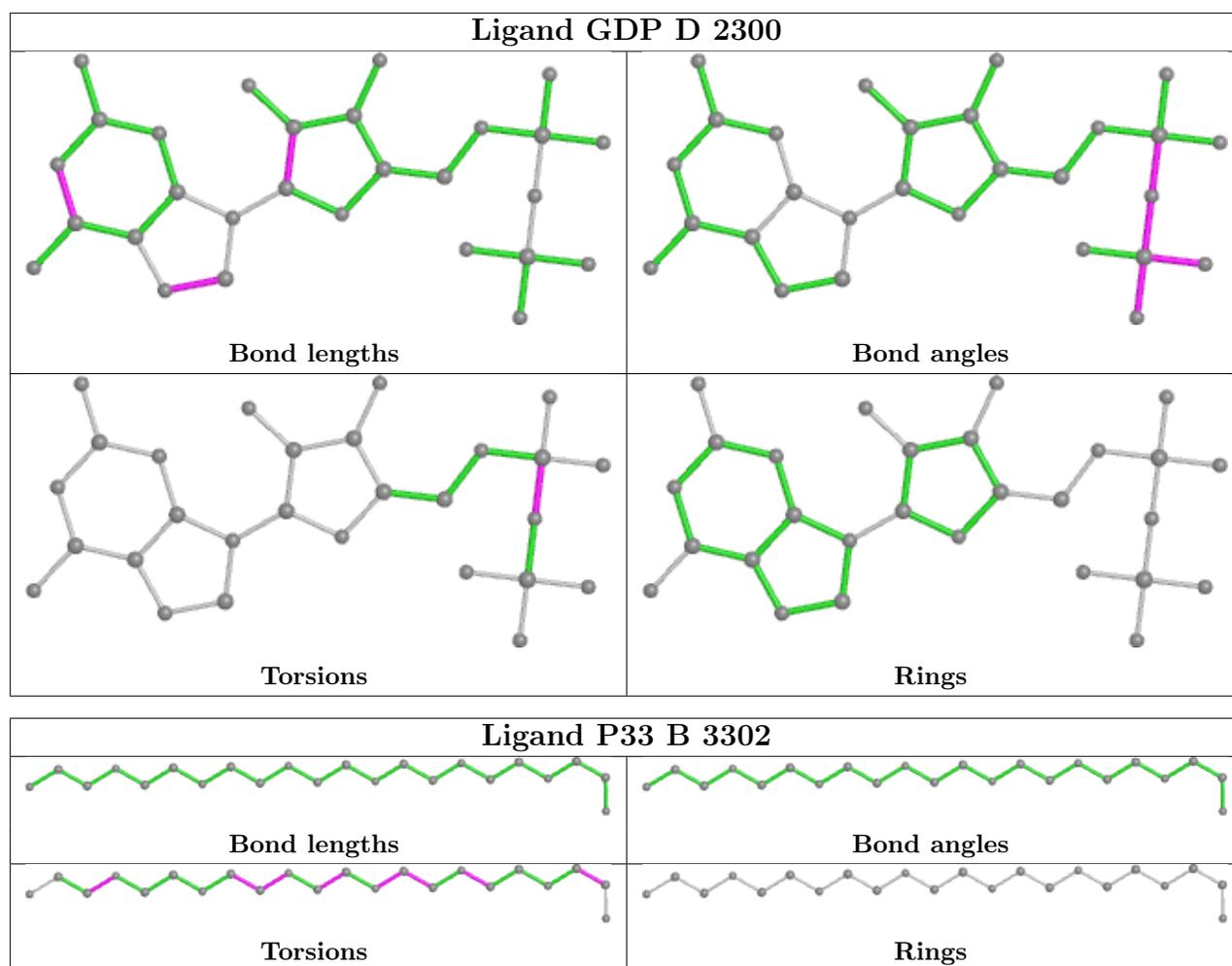6 monomers are involved in 12 short contacts:

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

Bond lengths

Bond angles

Torsions

Rings

Ligand P33 F 3306

Bond lengths

Bond angles

Torsions

Rings

Ligand P33 H 3308

Bond lengths

Bond angles

Torsions

Rings

Ligand GDP F 2305

Bond lengths

Bond angles

Torsions

Rings

Ligand GDP B 2302

Bond lengths

Bond angles

Torsions

Rings

Ligand P33 D 3304

Bond lengths

Bond angles

Torsions

Rings

Ligand GDP D 2300

Bond lengths

Bond angles

Torsions

Rings

Ligand P33 B 3302

Bond lengths

Bond angles

Torsions

Rings

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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.