wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 04:54 AM GMT

PDB ID	:	2ONG
Title	:	Crystal Structure of of limonene synthase with 2-fluorogeranyl diphosphate (FGPP).
Authors	:	Hyatt, D.C.; Youn, B.; Croteau, R.; Kang, C.
Deposited on	:	2007-01-23
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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

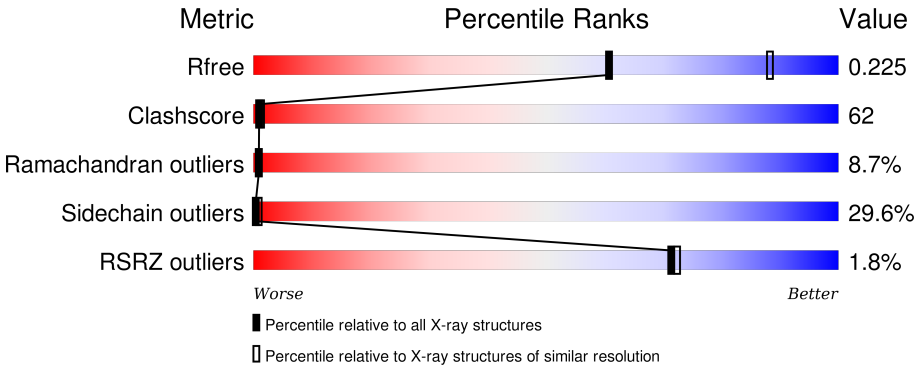
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (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 on the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	543	3% 19% 47% 29% 6%
1	B	543	% 20% 45% 30% 5%

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
2	MN	A	601	-	-	-	X
3	FPG	A	600	-	-	X	-
3	FPG	B	1600	-	-	X	-
4	BTB	A	604	-	-	-	X
4	BTB	A	605	-	-	X	-
4	BTB	B	1604	-	-	X	X
4	BTB	B	1605	-	-	X	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9181 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 4S-limonene synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	543	Total	C	N	O	S	0	0	0
			4495	2871	761	843	20			
1	B	543	Total	C	N	O	S	0	0	0
			4491	2870	758	843	20			

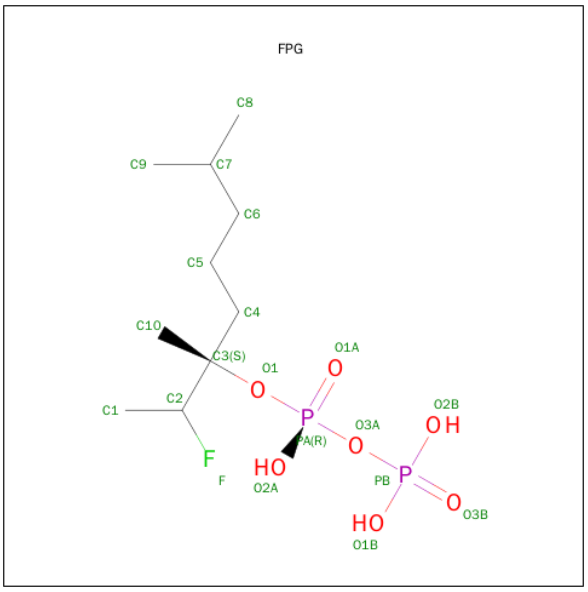
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	57	MET	GLU	ENGINEERED	UNP Q40322
B	57	MET	GLU	ENGINEERED	UNP Q40322

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

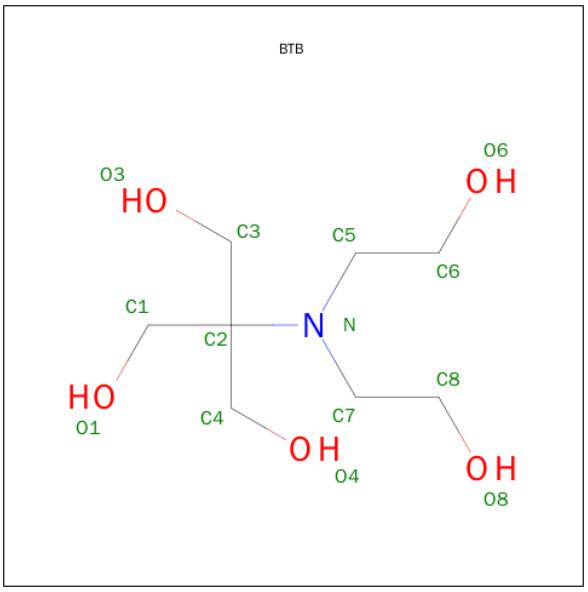
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total	Mn	0	0
			3	3		
2	A	3	Total	Mn	0	0
			3	3		

- Molecule 3 is (1S)-1-[(1R)-1-FLUOROETHYL]-1,5-DIMETHYLHEXYL TRIHYDROGEN DIPHOSPHATE (three-letter code: FPG) (formula: $C_{10}H_{23}FO_7P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	F	O	P	0	0
			20	10	1	7	2		
3	B	1	Total	C	F	O	P	0	0
			20	10	1	7	2		

- Molecule 4 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: BTB) (formula: $C_8H_{19}NO_5$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is water.

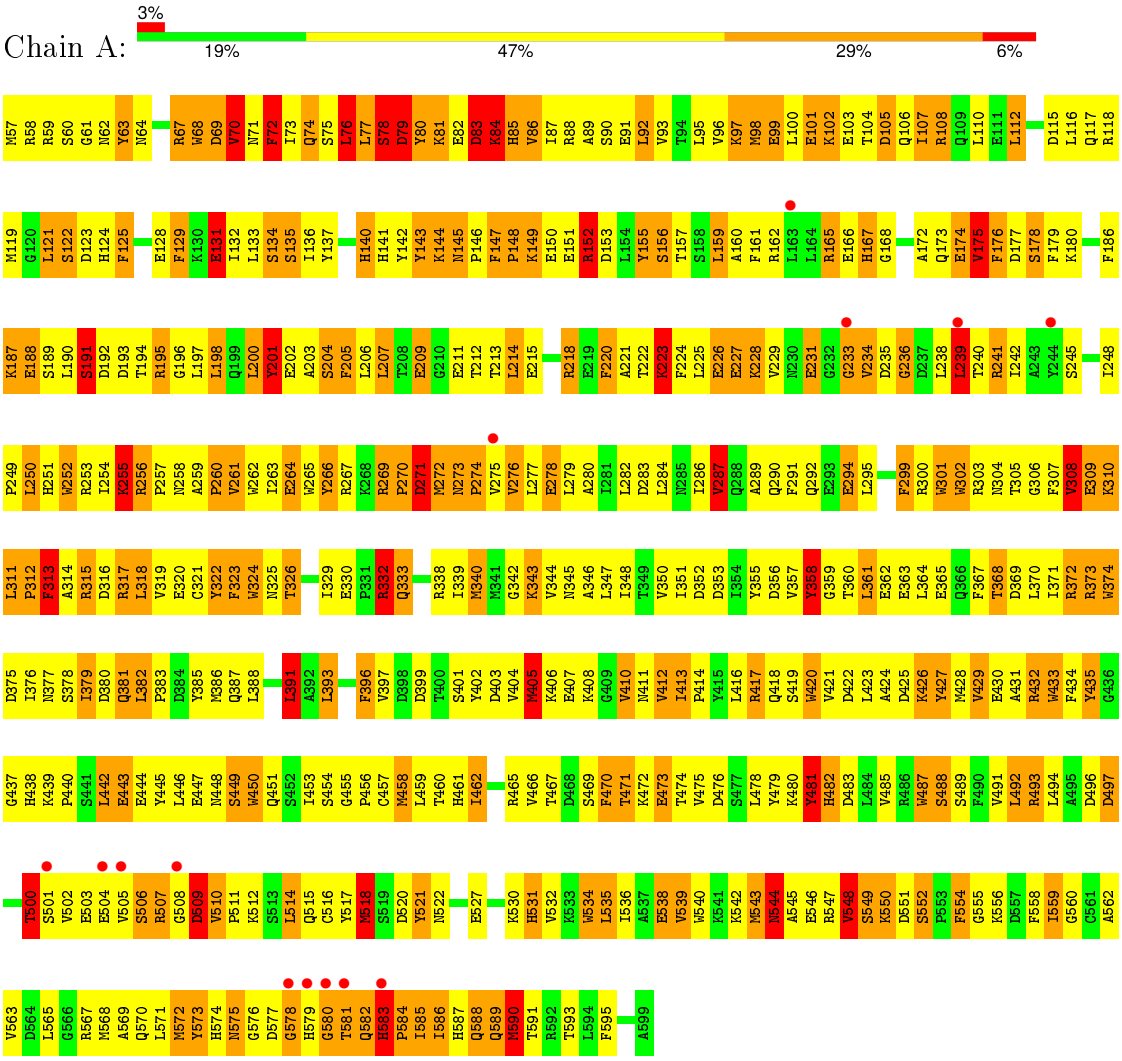
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	39	Total O 39 39	0	0
5	B	54	Total O 54 54	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

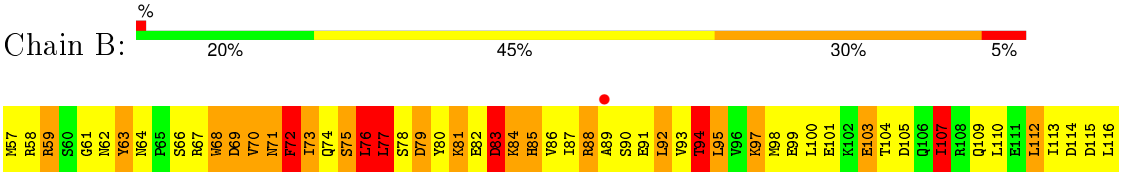
- Molecule 1: 4S-limonene synthase

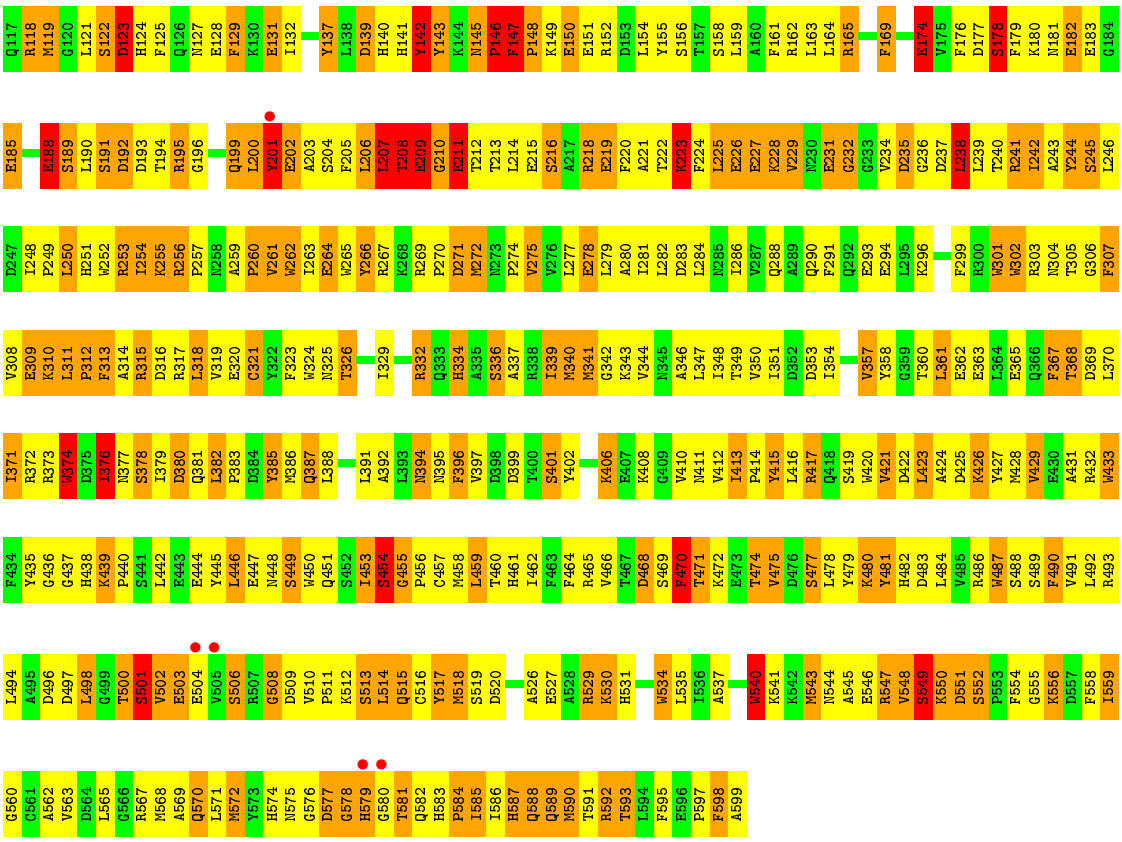
Chain A:



- Molecule 1: 4S-limonene synthase

Chain B:




Q117 R118 M119 G120 L121 S122 D123 H124 F125 Q126 N127 E128 F129 K130 E131 I132 Y137 L138 D139 H140 H141 Y142 Y143 K144 N145 P146 F147 P148 K149 E150 E151 R152 D153 L154 Y155 S156 T157 S158 L159 A160 F161 R162 L163 L164 R165 F169 E174 V175 F176 D177 S178 F179 K180 N181 E182 E183 G184
E185 E188 S189 L190 S191 D192 D193 T194 R195 G196 Q199 L200 Y201 E202 A203 S204 F205 L206 L207 T208 E209 G210 E211 T212 T213 L214 E215 S216 A217 R218 E219 F220 A221 T222 K223 F224 L225 E226 E227 K228 V229 N230 E231 G232 G233 V234 D235 G236 D237 L238 L239 T240 R241 I242 A243 Y244 S245 L246
D247 I248 P249 L250 H251 W252 R253 I254 K255 R256 P257 N258 A259 P260 V261 W262 I263 E264 W265 Y266 R267 K268 R269 P270 D271 M272 N273 P274 V275 Y276 L277 E278 L279 A280 I281 L282 D283 L284 N285 I286 V287 Q288 K289 Q290 F291 Q292 E293 E294 L295 K296 F299 R300 W301 W302 R303 N304 T305 G306 F307
V308 E309 K310 L311 P312 F313 A314 R315 D316 R317 L318 V319 E320 G321 Y322 F323 W324 N325 T326 I329 R332 Q333 H334 A335 S336 A337 R338 I339 M340 M341 G342 K343 V344 N345 A346 L347 I348 T349 V350 I351 D352 D353 I354 V357 Y358 G359 T360 L361 E362 E363 L364 E365 Q366 F367 T368 D369 L370
I371 R372 R373 W374 D375 I376 N377 S378 I379 D380 Q381 L382 P383 D384 Y385 M386 Q387 L388 L391 A392 L393 N394 N395 F396 V397 D398 D399 T400 S401 Y402 K406 E407 K408 G409 V410 N411 V412 I413 P414 Y415 L416 R417 Q418 S419 W420 V421 D422 L423 A424 D425 K426 Y427 M428 V429 E430 A431 R432 W433
F434 Y435 G436 G437 H438 K439 P440 S441 L442 E443 E444 Y445 L446 E447 N448 S449 W450 Q451 S452 I453 S454 G455 P456 C457 M458 L459 T460 H461 I462 F463 F464 R465 V466 Y467 D468 S469 F470 T471 K472 E473 T474 V475 D476 S477 L478 Y479 K480 Y481 H482 D483 L484 V485 R486 W487 S488 S489 F490 V491 L492 R493
L494 A495 D496 D497 L498 G499 T500 S501 V502 E503 E504 V505 S506 R507 G508 D509 V510 P511 K512 S513 L514 Q515 C516 Y517 M518 S519 D520 A526 E527 A528 R529 K530 H531 W534 L535 I536 A537 W540 K541 K542 M543 N544 A545 E546 R547 V548 S549 K550 D551 S552 P553 F554 G555 K556 D557 F558 I559
G560 C561 A562 V563 D564 L565 G566 R567 M568 A569 Q570 L571 M572 Y573 H574 N575 G576 D577 G578 H579 G580 T581 Q582 H583 P584 I585 I586 H587 Q588 Q589 M590 T591 R592 T593 L594 F595 E596 P597 F598 A599

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	200.48Å 200.48Å 123.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.70 27.90 – 2.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.70) 93.0 (27.90-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 2.61Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.208 , 0.241 0.217 , 0.225	Depositor DCC
R_{free} test set	3061 reflections (5.27%)	DCC
Wilson B-factor (Å^2)	48.9	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 120.2	EDS
Estimated twinning fraction	0.478 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 69724 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9181	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: FPG, MN, BTB

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	0.91	9/4607 (0.2%)	1.67	129/6234 (2.1%)
1	B	0.89	4/4603 (0.1%)	1.62	108/6230 (1.7%)
All	All	0.90	13/9210 (0.1%)	1.65	237/12464 (1.9%)

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	368	THR	C-O	8.06	1.38	1.23
1	A	131	GLU	CG-CD	7.48	1.63	1.51
1	B	368	THR	C-O	6.90	1.36	1.23
1	A	487	TRP	CG-CD2	-6.30	1.32	1.43
1	A	433	TRP	CG-CD2	-6.29	1.32	1.43

The worst 5 of 237 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	481	TYR	N-CA-C	-11.62	79.63	111.00
1	B	266	TYR	CB-CG-CD1	-11.53	114.08	121.00
1	A	266	TYR	CB-CG-CD1	-10.92	114.45	121.00
1	A	433	TRP	CD1-CG-CD2	10.65	114.82	106.30
1	A	143	TYR	CB-CG-CD1	-10.62	114.62	121.00

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	4495	0	4346	544	0
1	B	4491	0	4342	495	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	20	0	18	9	0
3	B	20	0	18	13	0
4	A	28	0	38	22	0
4	B	28	0	38	40	0
5	A	39	0	0	1	0
5	B	54	0	0	2	0
All	All	9181	0	8800	1101	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 62.

The worst 5 of 1101 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)
4:B:1604:BTB:N	4:B:1604:BTB:C2	1.68	1.56
4:B:1605:BTB:C2	4:B:1605:BTB:N	1.69	1.51
4:A:605:BTB:C2	4:A:605:BTB:N	1.71	1.49
1:B:579:HIS:CD2	3:B:1600:FPG:H92	1.67	1.28
4:B:1605:BTB:C7	4:B:1605:BTB:H32	1.72	1.19

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	541/543 (100%)	400 (74%)	93 (17%)	48 (9%)	1 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	541/543 (100%)	409 (76%)	86 (16%)	46 (8%)	1 1
All	All	1082/1086 (100%)	809 (75%)	179 (16%)	94 (9%)	1 1

5 of 94 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	76	LEU
1	A	78	SER
1	A	83	ASP
1	A	125	PHE
1	A	152	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	490/492 (100%)	352 (72%)	138 (28%)	0 1
1	B	490/492 (100%)	338 (69%)	152 (31%)	0 0
All	All	980/984 (100%)	690 (70%)	290 (30%)	0 1

5 of 290 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	572	MET
1	B	118	ARG
1	B	513	SER
1	A	589	GLN
1	B	73	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	587	HIS
1	B	145	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	575	ASN
1	B	140	HIS
1	B	290	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	FPG	A	600	2	15,19,19	1.75	4 (26%)	22,29,29	1.58	3 (13%)
4	BTB	A	604	-	12,13,13	1.46	3 (25%)	8,16,16	1.13	1 (12%)
4	BTB	A	605	-	12,13,13	3.08	5 (41%)	8,16,16	1.96	2 (25%)
3	FPG	B	1600	2	15,19,19	1.67	3 (20%)	22,29,29	1.94	4 (18%)
4	BTB	B	1604	-	12,13,13	3.39	5 (41%)	8,16,16	1.38	2 (25%)
4	BTB	B	1605	-	12,13,13	3.04	5 (41%)	8,16,16	1.79	2 (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
3	FPG	A	600	2	-	0/19/25/25	0/0/0/0
4	BTB	A	604	-	-	0/21/21/21	0/0/0/0
4	BTB	A	605	-	-	0/21/21/21	0/0/0/0
3	FPG	B	1600	2	-	0/19/25/25	0/0/0/0
4	BTB	B	1604	-	-	0/21/21/21	0/0/0/0
4	BTB	B	1605	-	-	0/21/21/21	0/0/0/0

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	600	FPG	F-C2	-3.01	1.33	1.41
3	B	1600	FPG	F-C2	-2.38	1.35	1.41
3	A	600	FPG	C4-C3	2.21	1.58	1.51
4	A	605	BTB	C3-C2	2.28	1.56	1.53
4	A	604	BTB	C3-C2	2.35	1.56	1.53

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1600	FPG	O1-C3-C10	-5.92	86.95	107.51
4	A	605	BTB	C7-N-C5	-4.17	99.23	112.52
3	B	1600	FPG	PA-O3A-PB	-3.94	119.46	132.67
4	B	1605	BTB	C7-N-C5	-3.64	100.92	112.52
3	A	600	FPG	PA-O3A-PB	-3.41	121.22	132.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 84 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	600	FPG	9	0
4	A	604	BTB	3	0
4	A	605	BTB	19	0
3	B	1600	FPG	13	0
4	B	1604	BTB	19	0
4	B	1605	BTB	21	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

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	543/543 (100%)	0.10	14 (2%) 59 59	18, 46, 94, 100	0
1	B	543/543 (100%)	0.07	6 (1%) 82 83	16, 46, 96, 100	0
All	All	1086/1086 (100%)	0.09	20 (1%) 71 72	16, 46, 95, 100	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	579	HIS	6.6
1	B	579	HIS	6.4
1	B	505	VAL	5.0
1	B	580	GLY	4.4
1	B	504	GLU	4.2

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 carbohydrates in this entry.

6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
4	BTB	A	604	14/14	0.82	0.50	10.89	83,90,96,100	0
4	BTB	B	1604	14/14	0.66	0.33	8.38	67,89,100,100	0
4	BTB	B	1605	14/14	0.87	0.20	2.86	93,100,100,100	0
2	MN	A	601	1/1	0.98	0.26	2.57	52,52,52,52	0
2	MN	A	602	1/1	0.97	0.27	1.70	41,41,41,41	0
2	MN	A	603	1/1	1.00	0.26	0.75	43,43,43,43	0
3	FPG	B	1600	20/20	0.97	0.27	0.63	44,71,100,100	0
3	FPG	A	600	20/20	0.97	0.26	0.57	54,69,76,82	0
2	MN	B	1603	1/1	0.99	0.25	-	44,44,44,44	0
2	MN	B	1602	1/1	0.98	0.24	-	49,49,49,49	0
2	MN	B	1601	1/1	0.99	0.25	-	52,52,52,52	0
4	BTB	A	605	14/14	0.83	0.23	-	98,100,100,100	0

6.5 Other polymers (i)

There are no such residues in this entry.