



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 08:53 pm GMT

PDB ID : 2F3T
Title : Crystal Structure Of E.coli Guanylate Kinase In Complex With Ganciclovir monophosphate
Authors : Hible, G.; Cherfils, J.
Deposited on : 2005-11-22
Resolution : 3.16 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 ⓘ 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) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

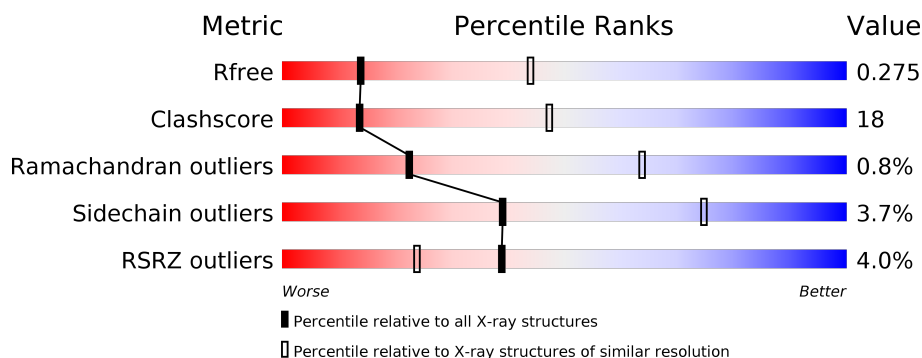
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.16 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}	100719	1259 (3.20-3.12)
Clashscore	112137	1397 (3.20-3.12)
Ramachandran outliers	110173	1368 (3.20-3.12)
Sidechain outliers	110143	1367 (3.20-3.12)
RSRZ outliers	101464	1264 (3.20-3.12)

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 $\leq 5\%$. 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.

Mol	Chain	Length	Quality of chain
1	A	207	<div> <div>7%</div> <div>65%</div> <div>33%</div> <div>..</div> </div>
1	B	207	<div> <div>%</div> <div>67%</div> <div>29%</div> <div>..</div> </div>
1	C	207	<div> <div>6%</div> <div>63%</div> <div>32%</div> <div>..</div> </div>
1	D	207	<div> <div>2%</div> <div>69%</div> <div>29%</div> <div>..</div> </div>
1	E	207	<div> <div>%</div> <div>68%</div> <div>30%</div> <div>.</div> </div>
1	F	207	<div> <div>7%</div> <div>61%</div> <div>37%</div> <div>..</div> </div>

2 Entry composition [i](#)

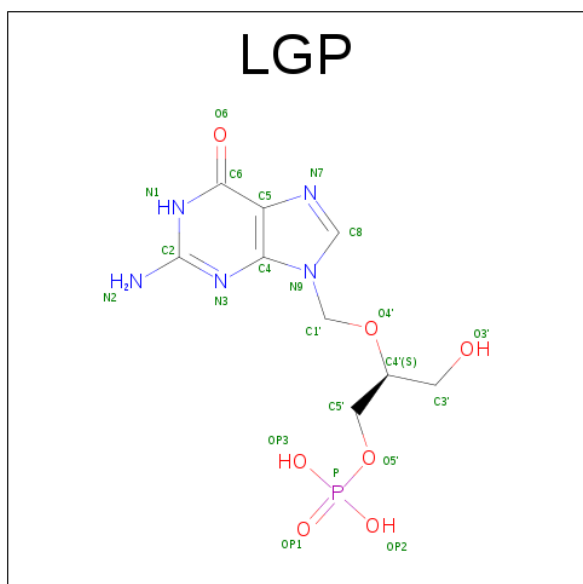
There are 3 unique types of molecules in this entry. The entry contains 9925 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Guanylate kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total	C	N	O	S	0	0	0
			1645	1034	296	310	5			
1	B	203	Total	C	N	O	S	0	0	0
			1632	1027	293	307	5			
1	C	201	Total	C	N	O	S	0	0	0
			1617	1019	288	305	5			
1	D	205	Total	C	N	O	S	0	0	0
			1640	1031	294	310	5			
1	E	204	Total	C	N	O	S	0	0	0
			1640	1031	295	309	5			
1	F	205	Total	C	N	O	S	0	0	0
			1638	1030	295	308	5			

- Molecule 2 is N9-1-HYDROXY-PROP-2-OXYMETHYL-GUANINE-3'-MONOPHOSPHATE (three-letter code: LGP) (formula: $C_9H_{14}N_5O_7P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	4	Total	C	N	O	P	0	0
			88	36	20	28	4		

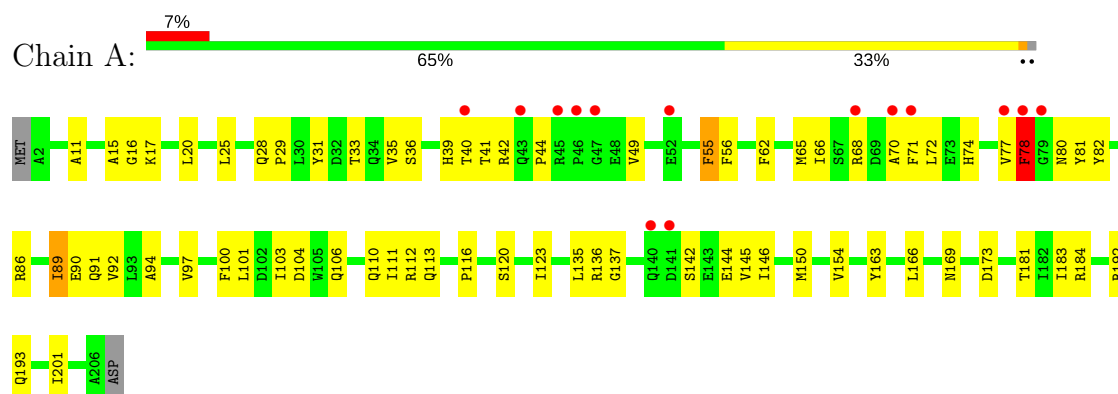
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	O	0	0
			3	3		
3	B	2	Total	O	0	0
			2	2		
3	C	5	Total	O	0	0
			5	5		
3	D	4	Total	O	0	0
			4	4		
3	E	6	Total	O	0	0
			6	6		
3	F	5	Total	O	0	0
			5	5		

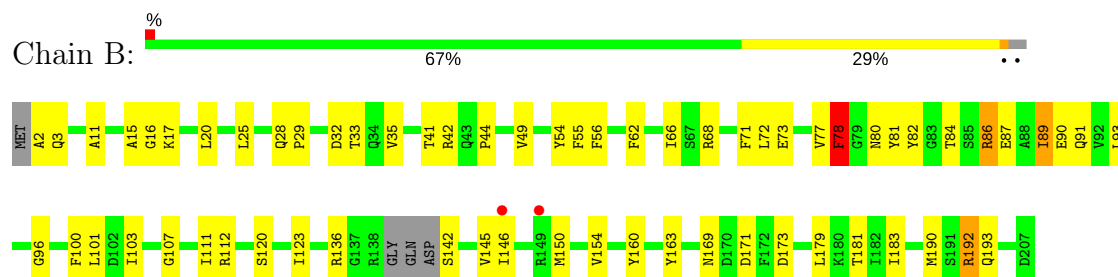
3 Residue-property plots [i](#)

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

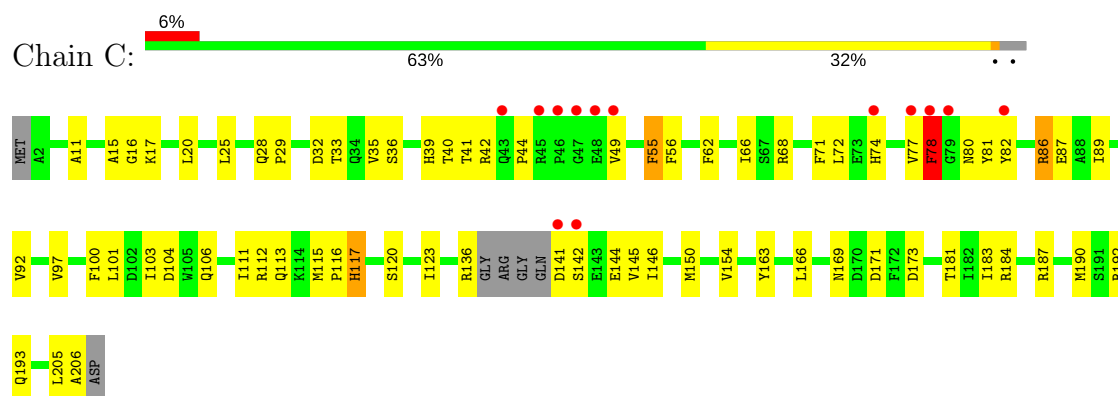
• Molecule 1: Guanylate kinase



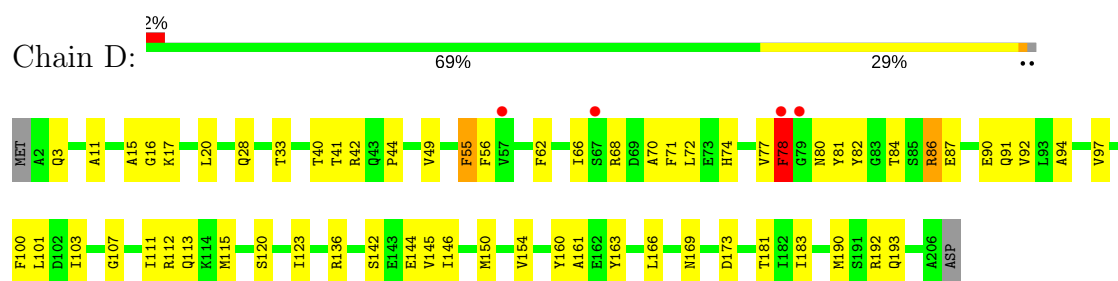
• Molecule 1: Guanylate kinase



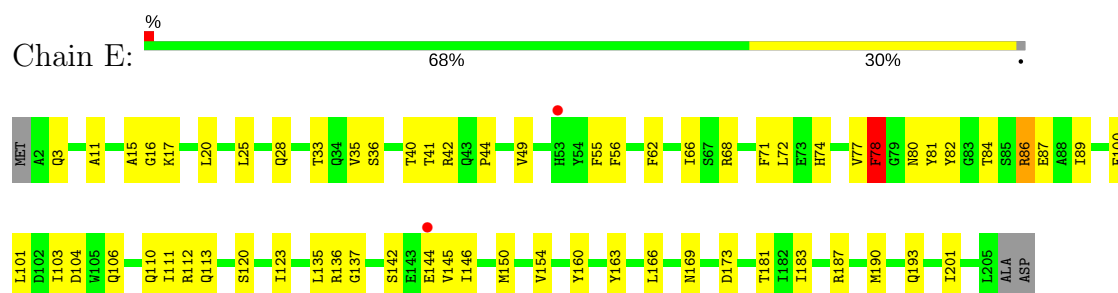
• Molecule 1: Guanylate kinase



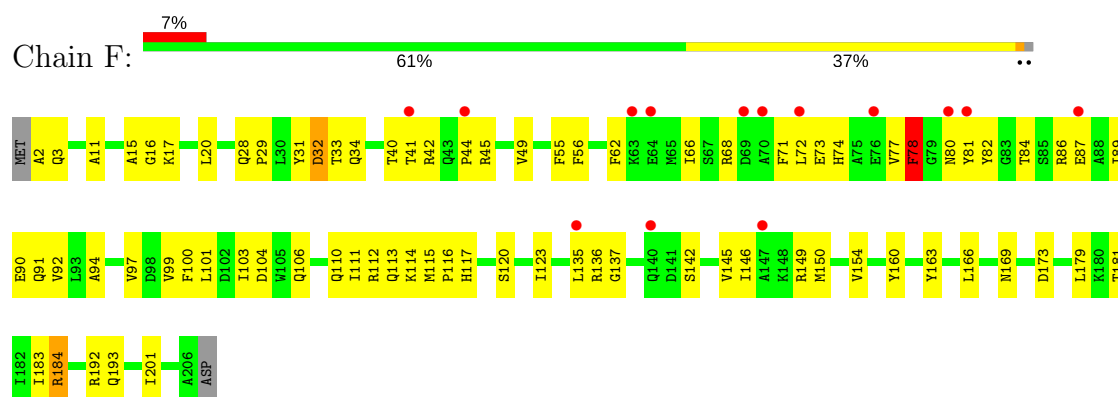
• Molecule 1: Guanylate kinase



● Molecule 1: Guanylate kinase



● Molecule 1: Guanylate kinase



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	107.75Å 107.75Å 272.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 3.16 29.89 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.96-3.16) 92.4 (29.89-2.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 2.80Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.268 , 0.290 0.256 , 0.275	Depositor DCC
R_{free} test set	1394 reflections (4.93%)	DCC
Wilson B-factor (Å ²)	63.5	Xtriage
Anisotropy	0.318	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 40.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	9925	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.71% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: LGP

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	0.45	0/1677	0.76	7/2264 (0.3%)
1	B	0.46	0/1663	0.97	6/2244 (0.3%)
1	C	0.46	0/1648	0.79	7/2225 (0.3%)
1	D	0.43	0/1671	0.76	6/2256 (0.3%)
1	E	0.47	0/1672	1.23	6/2257 (0.3%)
1	F	0.47	1/1670 (0.1%)	1.34	12/2255 (0.5%)
All	All	0.46	1/10001 (0.0%)	1.00	44/13501 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	192	ARG	CD-NE	-5.37	1.37	1.46

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	184	ARG	NE-CZ-NH1	-28.61	105.99	120.30
1	F	184	ARG	NE-CZ-NH2	26.48	133.54	120.30
1	E	112	ARG	NE-CZ-NH1	-25.30	107.65	120.30
1	E	112	ARG	NE-CZ-NH2	24.82	132.71	120.30
1	F	192	ARG	NE-CZ-NH1	-23.11	108.75	120.30

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	1645	0	1626	57	0
1	B	1632	0	1614	52	0
1	C	1617	0	1598	55	0
1	D	1640	0	1621	53	0
1	E	1640	0	1621	56	0
1	F	1638	0	1610	86	0
2	A	88	0	48	15	0
3	A	3	0	0	0	0
3	B	2	0	0	0	0
3	C	5	0	0	1	0
3	D	4	0	0	0	0
3	E	6	0	0	0	0
3	F	5	0	0	0	0
All	All	9925	0	9738	349	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 18.

The worst 5 of 349 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:C:28:GLN:HG2	1:C:33:THR:HB	1.43	0.99
1:F:72:LEU:HD21	1:F:115:MET:SD	2.03	0.98
2:A:302:LGP:C5'	2:A:302:LGP:H8	1.93	0.97
2:A:302:LGP:C8	2:A:302:LGP:H5''	1.94	0.97
2:A:302:LGP:H8	2:A:302:LGP:H5''	1.00	0.97

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	203/207 (98%)	182 (90%)	19 (9%)	2 (1%)	18	60
1	B	199/207 (96%)	184 (92%)	14 (7%)	1 (0%)	32	73
1	C	197/207 (95%)	180 (91%)	15 (8%)	2 (1%)	18	60
1	D	203/207 (98%)	187 (92%)	15 (7%)	1 (0%)	32	73
1	E	202/207 (98%)	184 (91%)	16 (8%)	2 (1%)	18	60
1	F	203/207 (98%)	186 (92%)	15 (7%)	2 (1%)	18	60
All	All	1207/1242 (97%)	1103 (91%)	94 (8%)	10 (1%)	22	64

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	78	PHE
1	B	78	PHE
1	C	78	PHE
1	D	78	PHE
1	E	78	PHE

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	175/177 (99%)	170 (97%)	5 (3%)	48	79
1	B	174/177 (98%)	166 (95%)	8 (5%)	31	69
1	C	173/177 (98%)	164 (95%)	9 (5%)	27	64
1	D	174/177 (98%)	168 (97%)	6 (3%)	42	76
1	E	175/177 (99%)	170 (97%)	5 (3%)	48	79
1	F	172/177 (97%)	166 (96%)	6 (4%)	41	76
All	All	1043/1062 (98%)	1004 (96%)	39 (4%)	39	75

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	86	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	173	ASP
1	F	78	PHE
1	C	113	GLN
1	C	115	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	169	ASN
1	D	80	ASN
1	F	152	GLN
1	C	197	HIS
1	D	28	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](#)

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	LGP	A	300	-	19,23,23	1.57	3 (15%)	19,33,33	2.28	8 (42%)
2	LGP	A	301	-	19,23,23	1.58	3 (15%)	19,33,33	2.31	5 (26%)
2	LGP	A	302	-	19,23,23	1.57	3 (15%)	19,33,33	2.39	6 (31%)
2	LGP	A	303	-	19,23,23	1.59	3 (15%)	19,33,33	2.38	6 (31%)

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
2	LGP	A	300	-	-	0/11/13/13	0/2/2/2
2	LGP	A	301	-	-	0/11/13/13	0/2/2/2
2	LGP	A	302	-	-	0/11/13/13	0/2/2/2
2	LGP	A	303	-	-	0/11/13/13	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	302	LGP	C2-N1	2.68	1.40	1.35
2	A	300	LGP	C2-N1	2.86	1.40	1.35
2	A	303	LGP	C2-N1	2.92	1.40	1.35
2	A	301	LGP	C2-N1	2.97	1.40	1.35
2	A	300	LGP	P-OP1	3.17	1.61	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	302	LGP	N3-C2-N1	-5.38	119.61	127.46
2	A	301	LGP	N3-C2-N1	-5.35	119.64	127.46
2	A	300	LGP	N3-C2-N1	-5.07	120.06	127.46
2	A	303	LGP	N3-C2-N1	-4.90	120.31	127.46
2	A	303	LGP	C5'-C4'-C3'	-4.46	101.92	111.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	300	LGP	4	0
2	A	301	LGP	6	0
2	A	302	LGP	4	0
2	A	303	LGP	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	205/207 (99%)	0.21	14 (6%) 18 10	22, 55, 78, 84	0
1	B	203/207 (98%)	-0.11	2 (0%) 82 71	23, 51, 76, 82	0
1	C	201/207 (97%)	0.21	13 (6%) 20 11	23, 53, 79, 85	0
1	D	205/207 (99%)	-0.00	4 (1%) 65 49	24, 57, 79, 88	0
1	E	204/207 (98%)	-0.05	2 (0%) 82 71	23, 50, 76, 82	0
1	F	205/207 (99%)	0.24	14 (6%) 18 10	23, 56, 80, 88	0
All	All	1223/1242 (98%)	0.08	49 (4%) 39 24	22, 54, 79, 88	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	70	ALA	6.0
1	C	48	GLU	5.4
1	F	81	TYR	5.3
1	A	78	PHE	4.4
1	F	87	GLU	4.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	LGP	A	303	22/22	0.82	0.22	-0.88	54,57,70,72	0
2	LGP	A	300	22/22	0.89	0.18	-1.07	55,57,63,64	0
2	LGP	A	301	22/22	0.89	0.20	-1.48	26,33,45,46	0
2	LGP	A	302	22/22	0.95	0.13	-4.95	29,33,49,51	0

6.5 Other polymers [i](#)

There are no such residues in this entry.