



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 28, 2024 – 01:08 AM EST

PDB ID : 1G31
Title : GP31 CO-CHAPERONIN FROM BACTERIOPHAGE T4
Authors : Hunt, J.F.; Van Der Vies, S.M.; Henry, L.; Deisenhofer, J.
Deposited on : 1998-03-27
Resolution : 2.30 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 ⓘ 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 ⓘ](#)) 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 : 2.36
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) : 2.36

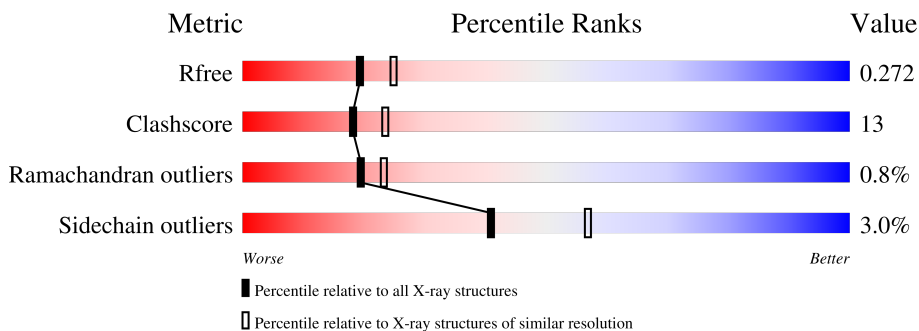
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	111	72% 23% . .
1	B	111	73% 23% . .
1	C	111	74% 20% . .
1	D	111	76% 18% . .
1	E	111	72% 23% . .
1	F	111	68% 27% . .
1	G	111	72% 23% . .

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	PO4	A	1153	-	X	-	-
2	PO4	A	1162	-	X	-	-
2	PO4	B	1163	-	X	-	-
2	PO4	C	1151	-	-	X	-
2	PO4	C	1164	-	X	-	-
2	PO4	C	1171	-	-	X	-
2	PO4	G	1161	-	-	X	-

2 Entry composition [i](#)

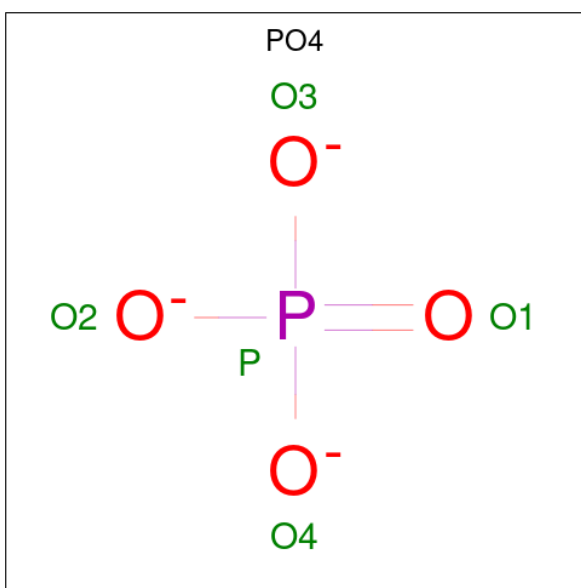
There are 4 unique types of molecules in this entry. The entry contains 6267 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 GP31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	107	Total 814	C 522	N 136	O 152	S 4	0	0	0
1	B	107	Total 814	C 522	N 136	O 152	S 4	0	0	0
1	C	107	Total 814	C 522	N 136	O 152	S 4	0	0	0
1	D	107	Total 814	C 522	N 136	O 152	S 4	0	0	0
1	E	107	Total 814	C 522	N 136	O 152	S 4	0	0	0
1	F	107	Total 814	C 522	N 136	O 152	S 4	0	0	0
1	G	107	Total 814	C 522	N 136	O 152	S 4	0	0	0

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	C	1	Total O P 3 2 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	C	1	Total O P 3 2 1	0	0
2	D	1	Total O P 5 4 1	0	0
2	E	1	Total O P 5 4 1	0	0
2	F	1	Total O P 5 4 1	0	0
2	G	1	Total O P 5 4 1	0	0
2	G	1	Total O P 5 4 1	0	0
2	G	1	Total O P 3 2 1	0	0
2	G	1	Total O P 5 4 1	0	0

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total K 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total K 2 2	0	0
3	C	1	Total K 1 1	0	0
3	D	1	Total K 1 1	0	0
3	E	1	Total K 1 1	0	0
3	F	2	Total K 2 2	0	0
3	G	2	Total K 2 2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	70	Total O 70 70	0	0
4	B	66	Total O 66 66	0	0
4	C	69	Total O 69 69	0	0
4	D	64	Total O 64 64	0	0
4	E	66	Total O 66 66	0	0
4	F	69	Total O 69 69	0	0
4	G	65	Total O 65 65	0	0

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. 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.

- Molecule 1: GP31

Chain A:  72% 23%



- Molecule 1: GP31

Chain B:  73% 23%



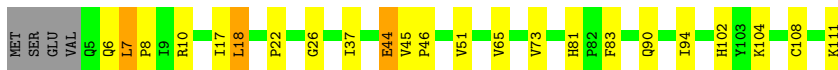
- Molecule 1: GP31

Chain C:  74% 20%



- Molecule 1: GP31

Chain D:  76% 18%



- Molecule 1: GP31

Chain E:  72% 23%



- Molecule 1: GP31

Chain F:  68% 27%



- Molecule 1: GP31

Chain G: 72% 23%



4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	157.68Å 157.68Å 90.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 24.93 – 2.03	Depositor EDS
% Data completeness (in resolution range)	83.5 (20.00-2.30) 89.5 (24.93-2.03)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.53 (at 2.03Å)	Xtrriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.225 , 0.254 0.240 , 0.272	Depositor DCC
R_{free} test set	7110 reflections (10.03%)	wwPDB-VP
Wilson B-factor (Å ²)	24.9	Xtrriage
Anisotropy	0.485	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 77.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6267	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.34% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: K, PO4

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	0.57	0/832	0.71	0/1132
1	B	0.58	0/832	0.75	0/1132
1	C	0.55	0/832	0.72	0/1132
1	D	0.57	0/832	0.71	0/1132
1	E	0.56	0/832	0.75	0/1132
1	F	0.61	0/832	0.79	1/1132 (0.1%)
1	G	0.58	0/832	0.79	1/1132 (0.1%)
All	All	0.57	0/5824	0.75	2/7924 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	G	18	LEU	CA-CB-CG	5.24	127.36	115.30
1	F	18	LEU	CA-CB-CG	5.16	127.17	115.30

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	814	0	824	22	0
1	B	814	0	824	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	814	0	824	27	0
1	D	814	0	824	24	0
1	E	814	0	824	28	0
1	F	814	0	824	36	0
1	G	814	0	824	22	0
2	A	20	0	0	0	1
2	B	20	0	0	1	0
2	C	16	0	0	0	6
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
2	G	18	0	0	0	4
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	2	0	0	0	0
3	G	2	0	0	0	0
4	A	70	0	0	1	0
4	B	66	0	0	2	0
4	C	69	0	0	2	0
4	D	64	0	0	1	0
4	E	66	0	0	2	0
4	F	69	0	0	3	1
4	G	65	0	0	2	0
All	All	6267	0	5768	155	11

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 155 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:F:83:PHE:HB3	1:F:89:LYS:HG3	1.43	0.97
1:F:83:PHE:CG	1:F:89:LYS:HE3	2.01	0.95
1:C:7:LEU:HB3	1:C:8:PRO:HD2	1.59	0.84
1:C:10:ARG:HH11	1:D:111:LYS:HG3	1.41	0.83
1:A:17:ILE:HD11	1:B:108:CYS:HB3	1.61	0.82

The worst 5 of 11 symmetry-related close contacts are listed below. The label for Atom-2 includes

the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:1161:PO4:O1	2:G:1161:PO4:O1[7_556]	0.45	1.75
2:G:1161:PO4:P	2:G:1161:PO4:O2[7_556]	1.46	0.74
2:G:1161:PO4:P	2:G:1161:PO4:O1[7_556]	1.47	0.73
2:C:1151:PO4:P	2:C:1151:PO4:O1[7_556]	1.48	0.72
2:C:1171:PO4:P	2:C:1171:PO4:O2[7_556]	1.48	0.72

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	105/111 (95%)	100 (95%)	4 (4%)	1 (1%)	15	17
1	B	105/111 (95%)	99 (94%)	5 (5%)	1 (1%)	15	17
1	C	105/111 (95%)	98 (93%)	6 (6%)	1 (1%)	15	17
1	D	105/111 (95%)	98 (93%)	6 (6%)	1 (1%)	15	17
1	E	105/111 (95%)	98 (93%)	7 (7%)	0	100	100
1	F	105/111 (95%)	100 (95%)	4 (4%)	1 (1%)	15	17
1	G	105/111 (95%)	98 (93%)	6 (6%)	1 (1%)	15	17
All	All	735/777 (95%)	691 (94%)	38 (5%)	6 (1%)	19	23

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	26	GLY
1	B	25	ALA
1	G	26	GLY
1	A	26	GLY
1	D	26	GLY

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	91/96 (95%)	87 (96%)	4 (4%)	28	39
1	B	91/96 (95%)	88 (97%)	3 (3%)	38	53
1	C	91/96 (95%)	88 (97%)	3 (3%)	38	53
1	D	91/96 (95%)	88 (97%)	3 (3%)	38	53
1	E	91/96 (95%)	90 (99%)	1 (1%)	73	86
1	F	91/96 (95%)	89 (98%)	2 (2%)	52	69
1	G	91/96 (95%)	88 (97%)	3 (3%)	38	53
All	All	637/672 (95%)	618 (97%)	19 (3%)	41	57

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	7	LEU
1	G	31	THR
1	G	90	GLN
1	G	18	LEU
1	C	31	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	81	HIS
1	G	6	GLN
1	G	102	HIS
1	G	52	HIS
1	G	5	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 30 ligands modelled in this entry, 11 are monoatomic - leaving 19 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
2	PO4	C	1164	-	4,4,4	2.90	3 (75%)	6,6,6	1.57	2 (33%)
2	PO4	C	151	3	4,4,4	1.19	0	6,6,6	0.83	0
2	PO4	G	1161	-	0,2,4	-	-	0,1,6	-	-
2	PO4	B	151	3	4,4,4	1.33	0	6,6,6	0.59	0
2	PO4	E	151	3	4,4,4	2.09	3 (75%)	6,6,6	1.39	0
2	PO4	A	1173	3	4,4,4	1.45	0	6,6,6	0.68	0
2	PO4	C	1171	3	0,2,4	-	-	0,1,6	-	-
2	PO4	D	151	3	4,4,4	1.02	0	6,6,6	0.75	0
2	PO4	G	151	3	4,4,4	1.39	0	6,6,6	1.11	0
2	PO4	G	1154	3	4,4,4	1.85	1 (25%)	6,6,6	1.23	0
2	PO4	B	1152	3	4,4,4	2.19	2 (50%)	6,6,6	1.10	1 (16%)
2	PO4	B	1163	-	4,4,4	2.95	3 (75%)	6,6,6	1.97	3 (50%)
2	PO4	G	1174	3	4,4,4	2.74	3 (75%)	6,6,6	0.36	0
2	PO4	A	151	3	4,4,4	1.21	0	6,6,6	1.45	1 (16%)
2	PO4	F	151	3	4,4,4	2.11	2 (50%)	6,6,6	1.33	1 (16%)
2	PO4	A	1153	3	4,4,4	1.95	3 (75%)	6,6,6	1.42	1 (16%)
2	PO4	A	1162	-	4,4,4	2.40	2 (50%)	6,6,6	1.46	2 (33%)
2	PO4	C	1151	3	0,2,4	-	-	0,1,6	-	-
2	PO4	B	1172	3	4,4,4	2.21	2 (50%)	6,6,6	0.73	0

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1163	PO4	P-O3	-4.17	1.42	1.54
2	C	1164	PO4	P-O4	-4.00	1.42	1.54
2	G	1174	PO4	P-O3	-3.56	1.43	1.54
2	A	1162	PO4	P-O3	-3.45	1.44	1.54
2	C	1164	PO4	P-O3	-3.31	1.44	1.54

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1163	PO4	O4-P-O2	3.08	117.86	107.97
2	B	1163	PO4	O4-P-O1	-2.65	101.19	110.89
2	A	1162	PO4	O4-P-O2	-2.42	100.19	107.97
2	F	151	PO4	O4-P-O2	-2.26	100.70	107.97
2	A	151	PO4	O3-P-O2	2.26	115.21	107.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	1161	PO4	0	4
2	B	151	PO4	1	0
2	C	1171	PO4	0	3
2	A	1162	PO4	0	1
2	C	1151	PO4	0	3

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.