W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 8, 2021 – 10:03 am GMT

|		
PDB ID	:	7BGT
Title	:	Mason-Pfizer Monkey Virus Protease mutant C7A/D26N/C106A in complex with peptidomimetic inhibitor
Authors	:	Wosicki, S.; Gilski, M.; Jaskolski, M.; Zabranska, H.; Pichova, I.
Deposited on	:	2021-01-08
Resolution	:	1.93 Å(reported)

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 1.93 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.
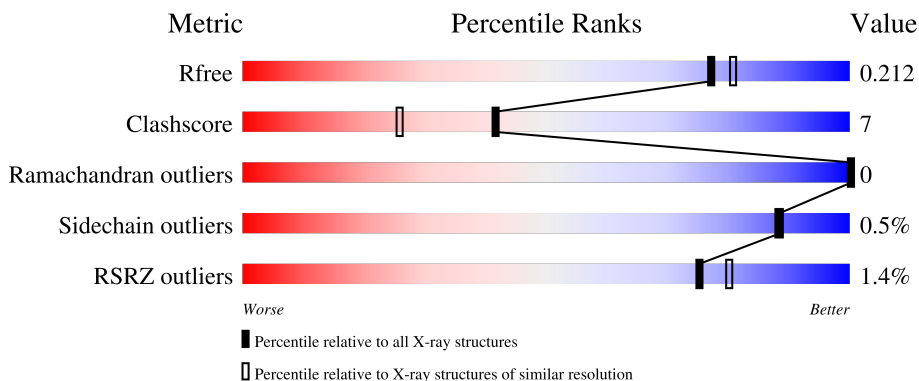
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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% 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	114	84% 10% · 5%
1	B	114	82% 12% 5%
1	C	114	78% 13% 9%
1	D	114	85% 10% 5%
2	F	7	71% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	7	71% 29%

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
3	ACT	A	201	-	-	X	-
3	ACT	C	201	-	-	X	-
3	ACT	F	101	-	-	X	-

2　Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 3916 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 Gag-Pro-Pol polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	108	Total	C	N	O	S	0	1	0
			864	560	143	157	4			
1	B	108	Total	C	N	O	S	0	1	0
			870	563	146	157	4			
1	C	104	Total	C	N	O	S	0	0	0
			831	541	136	150	4			
1	D	108	Total	C	N	O	S	0	1	0
			870	563	146	157	4			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	CYS	engineered mutation	UNP P07572
A	26	ASN	ASP	engineered mutation	UNP P07572
A	106	ALA	CYS	engineered mutation	UNP P07572
B	7	ALA	CYS	engineered mutation	UNP P07572
B	26	ASN	ASP	engineered mutation	UNP P07572
B	106	ALA	CYS	engineered mutation	UNP P07572
C	7	ALA	CYS	engineered mutation	UNP P07572
C	26	ASN	ASP	engineered mutation	UNP P07572
C	106	ALA	CYS	engineered mutation	UNP P07572
D	7	ALA	CYS	engineered mutation	UNP P07572
D	26	ASN	ASP	engineered mutation	UNP P07572
D	106	ALA	CYS	engineered mutation	UNP P07572

- Molecule 2 is a protein called peptidomimetic inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	7	Total	C	N	O	S	0	0	0
			64	44	9	10	1			
2	F	7	Total	C	N	O	S	0	0	0
			64	44	9	10	1			

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 4	C 2	O 2	0	0
3	C	1	Total 4	C 2	O 2	0	0
3	F	1	Total 4	C 2	O 2	0	0

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).
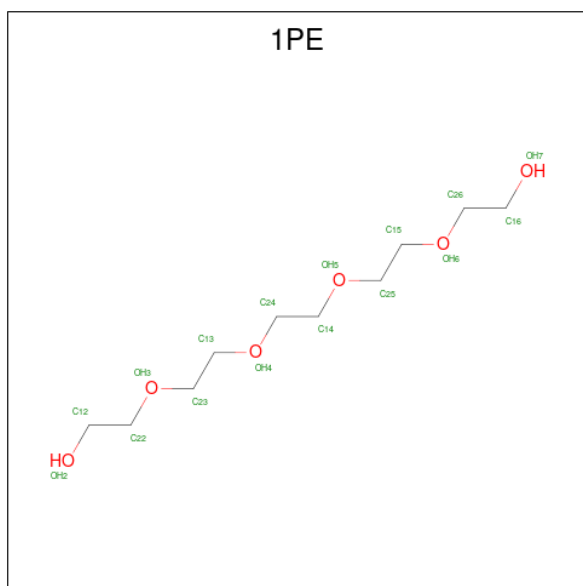
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 16	C 10	O 6	0	0
4	D	1	Total 16	C 10	O 6	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	64	Total 64	O 64	0	0
5	B	84	Total 84	O 84	0	0
5	C	72	Total 73	O 73	0	1
5	D	81	Total 81	O 81	0	0
5	G	5	Total 5	O 5	0	0
5	F	2	Total 2	O 2	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 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: Gag-Pro-Pol polyprotein
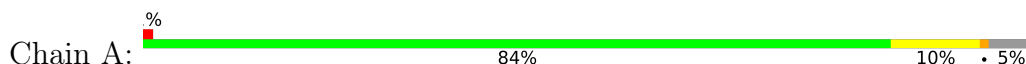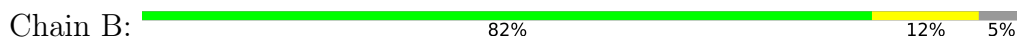
Chain A:

- Molecule 1: Gag-Pro-Pol polyprotein

Chain B:

- Molecule 1: Gag-Pro-Pol polyprotein

Chain C:

- Molecule 1: Gag-Pro-Pol polyprotein

Chain D:

- Molecule 2: peptidomimetic inhibitor

Chain G:

- Molecule 2: peptidomimetic inhibitor

Chain F:
71%   29%
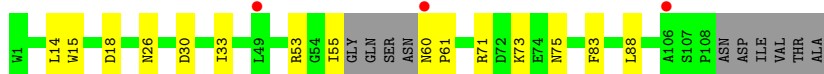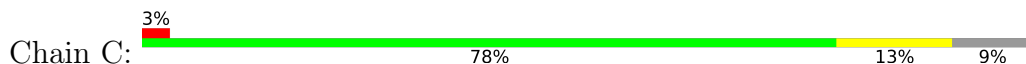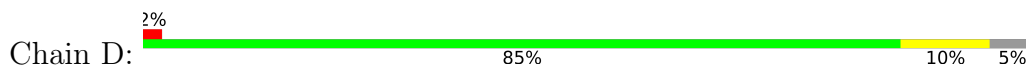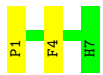
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	29.30Å 67.62Å 69.72Å 76.84° 83.88° 83.64°	Depositor
Resolution (Å)	32.42 – 1.93 42.66 – 1.93	Depositor EDS
% Data completeness (in resolution range)	97.0 (32.42-1.93) 97.0 (42.66-1.93)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 1.92Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.177 , 0.212 0.177 , 0.212	Depositor DCC
R_{free} test set	999 reflections (2.65%)	wwPDB-VP
Wilson B-factor (Å^2)	27.3	Xtriage
Anisotropy	0.107	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.199 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3916	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.31% 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, 0.333 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: 1PE, ACT, PSA

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	0.38	0/891	0.66	0/1218
1	B	0.38	0/897	0.60	0/1225
1	C	0.40	0/854	0.64	0/1168
1	D	0.37	0/897	0.62	0/1225
2	F	0.43	0/51	0.52	0/65
2	G	0.56	0/51	0.70	0/65
All	All	0.38	0/3641	0.63	0/4966

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	864	0	883	12	0
1	B	870	0	894	15	0
1	C	831	0	847	16	0
1	D	870	0	894	15	0
2	F	64	0	60	2	0
2	G	64	0	60	2	0
3	A	4	0	3	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	4	0	3	4	0
3	F	4	0	3	2	0
4	B	16	0	22	3	0
4	D	16	0	22	6	0
5	A	64	0	0	1	0
5	B	84	0	0	3	0
5	C	73	0	0	3	0
5	D	81	0	0	0	0
5	F	2	0	0	0	0
5	G	5	0	0	0	0
All	All	3916	0	3691	51	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 7.

The worst 5 of 51 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:B:88:LEU:HD11	3:F:101:ACT:H2	1.55	0.89
1:D:71:ARG:NE	4:D:201:1PE:H251	1.92	0.85
1:C:30:ASP:OD1	2:G:7:HIS:HB3	1.78	0.83
1:B:71:ARG:NE	4:B:201:1PE:H251	1.96	0.81
1:C:15:TRP:CD2	3:C:201:ACT:H3	2.17	0.80

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	107/114 (94%)	105 (98%)	2 (2%)	0	100	100
1	B	107/114 (94%)	105 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	100/114 (88%)	99 (99%)	1 (1%)	0	100	100
1	D	107/114 (94%)	104 (97%)	3 (3%)	0	100	100
2	F	4/7 (57%)	3 (75%)	1 (25%)	0	100	100
2	G	4/7 (57%)	3 (75%)	1 (25%)	0	100	100
All	All	429/470 (91%)	419 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	99/104 (95%)	98 (99%)	1 (1%)	76	71
1	B	100/104 (96%)	100 (100%)	0	100	100
1	C	95/104 (91%)	94 (99%)	1 (1%)	73	67
1	D	100/104 (96%)	100 (100%)	0	100	100
2	F	5/5 (100%)	5 (100%)	0	100	100
2	G	5/5 (100%)	5 (100%)	0	100	100
All	All	404/426 (95%)	402 (100%)	2 (0%)	88	88

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	ASN
1	C	14	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	3	GLN
1	C	87	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	26	ASN
1	B	60	ASN
1	C	60	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

5 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

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	108/114 (94%)	0.03	1 (0%)	84	87	18, 33, 71, 86	0
1	B	108/114 (94%)	-0.11	0	100	100	19, 31, 56, 79	0
1	C	104/114 (91%)	-0.04	3 (2%)	51	59	19, 32, 70, 100	0
1	D	108/114 (94%)	-0.13	2 (1%)	66	72	19, 32, 60, 71	0
2	F	6/7 (85%)	-0.27	0	100	100	24, 34, 39, 42	0
2	G	6/7 (85%)	-0.01	0	100	100	29, 36, 49, 77	0
All	All	440/470 (93%)	-0.06	6 (1%)	75	80	18, 32, 65, 100	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	49	LEU	4.4
1	A	55	ILE	3.1
1	C	60	ASN	2.7
1	D	1	TRP	2.6
1	D	42	ASN	2.5

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

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. 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	B-factors(Å^2)	Q<0.9
2	PSA	G	4	14/15	0.94	0.10	21,26,31,35	0
2	PSA	F	4	14/15	0.97	0.10	19,23,28,29	0

6.3   Carbohydrates ⓘ

There are no monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
3	ACT	A	201	4/4	0.54	0.29	64,78,78,79	0
3	ACT	C	201	4/4	0.77	0.20	63,67,67,70	0
4	1PE	D	201	16/16	0.79	0.19	36,51,59,59	0
4	1PE	B	201	16/16	0.80	0.19	39,51,67,68	0
3	ACT	F	101	4/4	0.91	0.22	53,57,61,72	0

6.5   Other polymers ⓘ

There are no such residues in this entry.