wwPDB X-ray Structure Validation Summary Report (i)

Oct 15, 2023 – 12:50 PM EDT

PDB ID	:	1PFG
Title	:	Strategy to design inhibitors: Structure of a complex of Proteinase K with a designed octapeptide inhibitor N-Ac-Pro-Ala-Pro-Phe-DAla-Ala-Ala-Ala-NH 2 at 2.5A resolution
Authors	:	Saxena, A.K.; Singh, T.P.; Peters, K.; Fittkau, S.; Betzel, C.
Deposited on	:	2003-05-27
Resolution	:	2.50 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 (i) 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 (i)) were used in the production of this report:

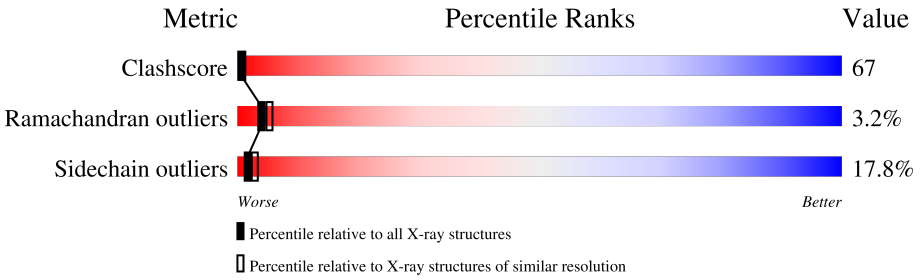
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al.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

1 Overall quality at a glance (i)

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	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	279	35% 47% 15% .
2	B	10	30% 30% 40%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2277 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 Proteinase K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	279	Total	C	N	O	S	0	0	0
			2017	1242	352	413	10			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	VAL	ALA	conflict	UNP P06873

- Molecule 2 is a protein called N-Ac-PAPFAAAA-NH2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	10	Total	C	N	O	0	0	1
			54	36	9	9			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	196	Total	O	0	0
			196	196		
3	B	10	Total	O	0	0
			10	10		

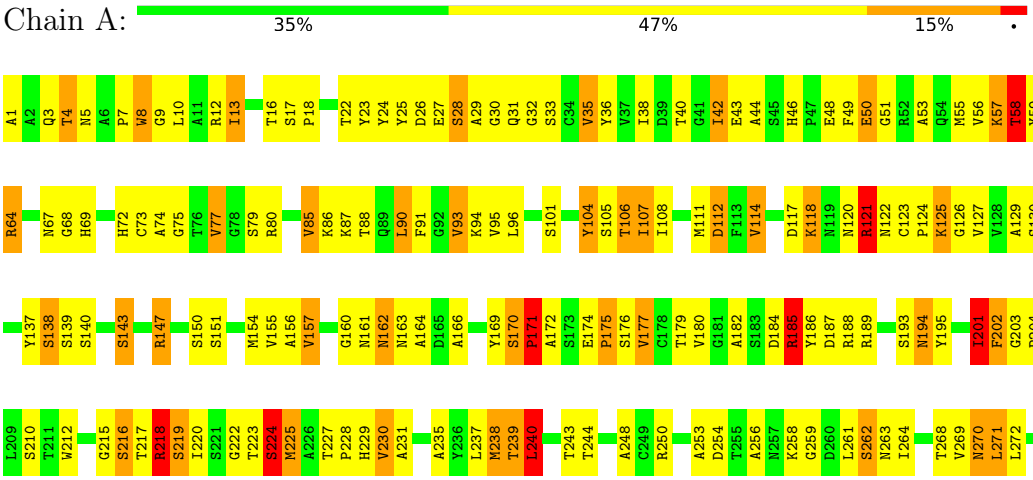
3 Residue-property plots (i)

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.

Note EDS was not executed.

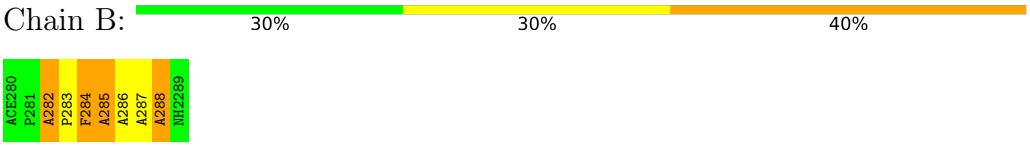
• Molecule 1: Proteinase K

Chain A:



• Molecule 2: N-Ac-PAPFAAAA-NH2

Chain B:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	68.00Å 68.00Å 107.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.50)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.167 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2277	wwPDB-VP
Average B, all atoms (Å^2)	13.0	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: ACE, NH2

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	1.26	6/2056 (0.3%)	1.87	46/2796 (1.6%)
2	B	2.71	1/52 (1.9%)	2.48	4/71 (5.6%)
All	All	1.32	7/2108 (0.3%)	1.89	50/2867 (1.7%)

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	285	ALA	N-CA	-17.11	1.12	1.46
1	A	238	MET	CG-SD	-6.13	1.65	1.81
1	A	125	LYS	CB-CG	-5.83	1.36	1.52
1	A	123	CYS	CB-SG	-5.49	1.72	1.81
1	A	202	PHE	CB-CG	-5.36	1.42	1.51

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	80	ARG	NE-CZ-NH1	-19.01	110.80	120.30
1	A	121	ARG	NE-CZ-NH1	17.91	129.25	120.30
1	A	171	PRO	CA-N-CD	-17.36	87.19	111.50
1	A	170	SER	C-N-CD	-16.43	84.45	120.60
1	A	121	ARG	NE-CZ-NH2	-16.28	112.16	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	2017	0	1916	261	3
2	B	54	0	50	31	0
3	A	196	0	0	35	5
3	B	10	0	0	1	0
All	All	2277	0	1966	271	5

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 67.

The worst 5 of 271 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:224:SER:CB	1:A:225:MET:HE3	1.61	1.30
1:A:126:GLY:HA3	1:A:238:MET:CE	1.60	1.28
1:A:85:VAL:O	1:A:85:VAL:CG1	1.73	1.23
1:A:220:ILE:HD12	2:B:286:ALA:HB2	1.23	1.18
1:A:269:VAL:CG1	1:A:271:LEU:HD12	1.75	1.16

All (5) 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)
3:A:457:HOH:O	3:A:457:HOH:O[7_465]	0.69	1.51
3:A:476:HOH:O	3:A:543:HOH:O[6_455]	0.81	1.39
1:A:112:ASP:OD1	3:A:339:HOH:O[7_465]	1.77	0.43
1:A:216:SER:CB	3:A:511:HOH:O[6_555]	1.80	0.40
1:A:150:SER:OG	3:A:423:HOH:O[7_465]	1.93	0.27

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	277/279 (99%)	244 (88%)	25 (9%)	8 (3%)	4 6
2	B	6/10 (60%)	4 (67%)	1 (17%)	1 (17%)	0 0
All	All	283/289 (98%)	248 (88%)	26 (9%)	9 (3%)	4 5

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	171	PRO
1	A	50	GLU
2	B	282	ALA
1	A	42	ILE
1	A	43	GLU

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	211/214 (99%)	173 (82%)	38 (18%)	1 3
2	B	3/3 (100%)	3 (100%)	0	100 100
All	All	214/217 (99%)	176 (82%)	38 (18%)	2 3

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	216	SER
1	A	250	ARG
1	A	218	ARG
1	A	239	THR
1	A	270	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	194	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	229	HIS
1	A	276	ASN
1	A	270	ASN
1	A	162	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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	284:PHE	C	285:ALA	N	2.83

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.