Full wwPDB X-ray Structure Validation Report (i)

Oct 2, 2023 – 06:21 AM EDT

PDB ID	:	6NAG
Title	:	X-ray structure of a secreted C11 cysteine protease from Bacteroides thetaiotaomicron "iotapain
Authors	:	Wolan, D.W.; Gonzalez-Paez, G.E.; Roncase, E.J.
Deposited on	:	2018-12-05
Resolution	:	2.68 Å(reported)

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

MolProbity	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
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)	:	2.35.1

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 2.68 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6181 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 Clostripain-related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	351	Total 2775	C 1774	N 435	O 548	S 18	0	0	0
1	B	350	Total 2775	C 1776	N 435	O 546	S 18	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	GLY	-	expression tag	UNP Q8A9T8
A	154	ALA	ARG	engineered mutation	UNP Q8A9T8
B	29	GLY	-	expression tag	UNP Q8A9T8
B	152B	ALA	ARG	engineered mutation	UNP Q8A9T8

- Molecule 2 is PROLINE (three-letter code: PRO) (formula: $C_5H_9NO_2$).
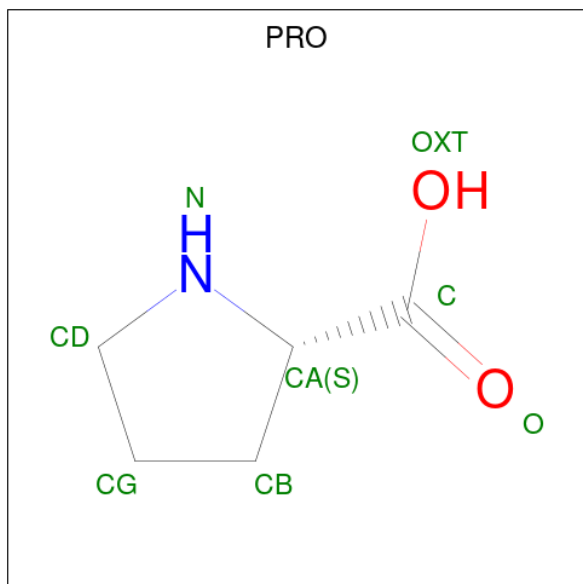
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 8	C 5	N 1	O 2	0	0
2	A	1	Total 8	C 5	N 1	O 2	0	0
2	A	1	Total 8	C 5	N 1	O 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	326	Total 326	O 326	0	0
3	B	281	Total 281	O 281	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	157.68Å 157.68Å 119.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.52 – 2.68	Depositor
% Data completeness (in resolution range)	98.8 (45.52-2.68)	Depositor
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.176 , 0.211	Depositor
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.290	Xtriage
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.039 for -h,-k,l	Xtriage
Total number of atoms	6181	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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.

4 Model quality ⓘ

4.1 Standard geometry ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles ⓘ

4.3.1 Protein backbone ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA ⓘ

MolProbity failed to run properly - this section is therefore empty.

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

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

4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

4.6 Ligand geometry ⓘ

3 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	PRO	A	401	-	8,8,8	0.85	1 (12%)	10,10,10	1.40	2 (20%)
2	PRO	A	403	-	8,8,8	0.86	1 (12%)	10,10,10	1.39	2 (20%)
2	PRO	A	402	-	8,8,8	0.87	1 (12%)	10,10,10	1.36	2 (20%)

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	PRO	A	401	-	-	2/4/11/11	0/1/1/1
2	PRO	A	403	-	-	4/4/11/11	0/1/1/1
2	PRO	A	402	-	-	2/4/11/11	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	402	PRO	OXT-C	-2.14	1.23	1.30
2	A	403	PRO	OXT-C	-2.14	1.23	1.30
2	A	401	PRO	OXT-C	-2.13	1.23	1.30

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	PRO	OXT-C-O	-2.67	118.03	124.09
2	A	402	PRO	OXT-C-O	-2.66	118.06	124.09
2	A	403	PRO	OXT-C-O	-2.65	118.08	124.09
2	A	401	PRO	OXT-C-CA	2.27	120.94	113.40
2	A	403	PRO	OXT-C-CA	2.27	120.93	113.40
2	A	402	PRO	OXT-C-CA	2.21	120.75	113.40

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	403	PRO	O-C-CA-N

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	403	PRO	OXT-C-CA-N
2	A	403	PRO	O-C-CA-CB
2	A	403	PRO	OXT-C-CA-CB
2	A	402	PRO	OXT-C-CA-CB
2	A	401	PRO	O-C-CA-N
2	A	402	PRO	O-C-CA-CB
2	A	401	PRO	OXT-C-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	171:ARG	C	173:TYR	N	33.01
1	B	171:ARG	C	173:TYR	N	32.91

5 Fit of model and data ⓘ

5.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

5.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.