wwPDB X-ray Structure Validation Summary Report (i)

Jan 20, 2024 – 04:29 pm GMT

PDB ID	:	7A8W
Title	:	Complex of rice blast (Magnaporthe oryzae) effector protein AVR-PikC with an engineered HMA domain of Pikp-1 from rice (Oryza sativa)
Authors	:	Maidment, J.H.R.; De la Concepcion, J.C.; Franceschetti, M.; Banfield, M.J.
Deposited on	:	2020-08-31
Resolution	:	2.15 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 (1)) were used in the production of this report:

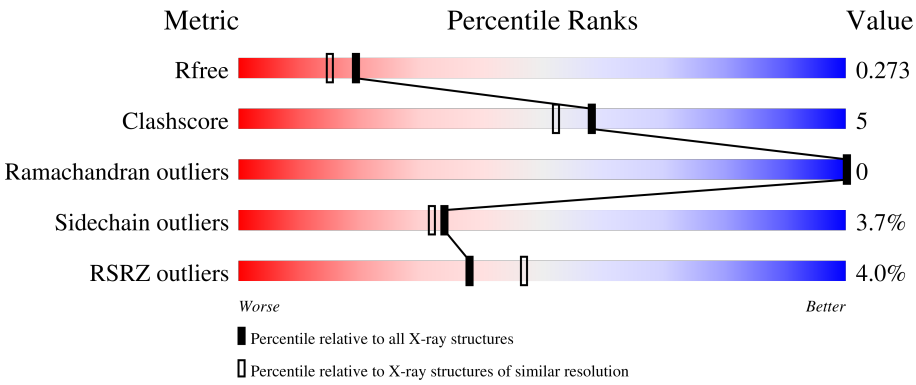
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	AAA	81	%, 84%, 5%, 11%
1	BBB	81	%, 78%, 12%, •, 9%
1	DDD	81	%, 72%, 15%, 14%
1	EEE	81	4%, 78%, 11%, •, 10%
2	CCC	93	13%, 65%, 20%, 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	FFF	93	81% 6% • 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7129 atoms, of which 3595 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 NBS-LRR class disease resistance protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	72	1112	335	582	93	99	3	8	0	0
1	BBB	74	1144	344	600	96	101	3	8	0	0
1	DDD	70	1091	329	573	90	96	3	8	0	0
1	EEE	73	1158	348	609	97	101	3	8	2	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	262	GLU	-	expression tag	UNP D5L9G5
AAA	263	ASP	-	expression tag	UNP D5L9G5
BBB	262	GLU	-	expression tag	UNP D5L9G5
BBB	263	ASP	-	expression tag	UNP D5L9G5
DDD	262	GLU	-	expression tag	UNP D5L9G5
DDD	263	ASP	-	expression tag	UNP D5L9G5
EEE	262	GLU	-	expression tag	UNP D5L9G5
EEE	263	ASP	-	expression tag	UNP D5L9G5

- Molecule 2 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	CCC	79	1254	414	607	111	116	6	10	0	0
2	FFF	82	1292	428	624	114	120	6	10	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	21	MET	-	initiating methionine	UNP G4MXW3
FFF	21	MET	-	initiating methionine	UNP G4MXW3

- Molecule 3 is water.

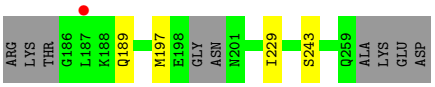
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	AAA	4	4	4	0	0
			Total	O		
3	BBB	13	13	13	0	0
			Total	O		
3	CCC	9	9	9	0	0
			Total	O		
3	DDD	15	15	15	0	0
			Total	O		
3	EEE	14	14	14	0	0
			Total	O		
3	FFF	23	23	23	0	0

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 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.

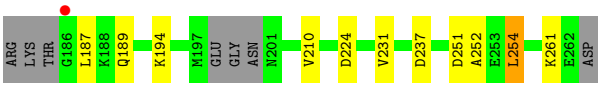
• Molecule 1: NBS-LRR class disease resistance protein

Chain AAA:



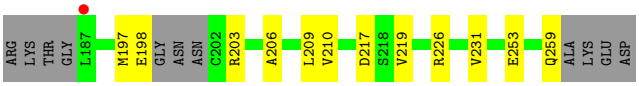
• Molecule 1: NBS-LRR class disease resistance protein

Chain BBB:



• Molecule 1: NBS-LRR class disease resistance protein

Chain DDD:

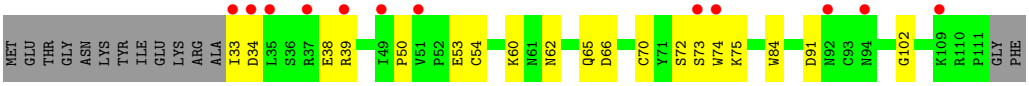


• Molecule 1: NBS-LRR class disease resistance protein

Chain EEE:

• Molecule 2: Uncharacterized protein

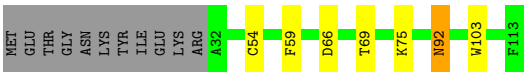
Chain CCC:



• Molecule 2: Uncharacterized protein

Chain FFF:

81%
6%
12%


MET
GLU
THR
GLY
ASN
LYS
TYR
ILE
GLU
LYS
ARG
A32
C54
F59
D66
T69
K75
N92
W103
F113

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.78Å 80.20Å 105.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.16 – 2.15 44.13 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.16-2.15) 99.9 (44.13-2.15)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0266	Depositor
R, R_{free}	0.224 , 0.272 0.226 , 0.273	Depositor DCC
R_{free} test set	1510 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	47.9	Xtriage
Anisotropy	0.676	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7129	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.34 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.6816e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

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	AAA	0.71	0/530	0.83	0/708
1	BBB	0.73	0/544	0.88	0/726
1	DDD	0.70	0/518	0.85	0/692
1	EEE	0.70	0/555	0.86	0/740
2	CCC	0.65	0/670	0.81	0/909
2	FFF	0.67	0/692	0.79	0/937
All	All	0.69	0/3509	0.83	0/4712

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 (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	AAA	530	582	580	1	0
1	BBB	544	600	598	9	0
1	DDD	518	573	571	8	0
1	EEE	549	609	609	7	0
2	CCC	647	607	605	13	0
2	FFF	668	624	622	3	0
3	AAA	4	0	0	0	0
3	BBB	13	0	0	1	0
3	CCC	9	0	0	0	0
3	DDD	15	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	EEE	14	0	0	1	0
3	FFF	23	0	0	0	0
All	All	3534	3595	3585	34	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 5.

The worst 5 of 34 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:EEE:190:LYS:HE3	3:EEE:309:HOH:O	1.95	0.65
1:AAA:197:MET:CE	1:AAA:229:ILE:HD11	2.31	0.61
2:CCC:65:GLN:HE21	2:CCC:65:GLN:HA	1.67	0.60
1:BBB:224:ASP:OD2	2:CCC:39:ARG:NH2	2.37	0.58
1:BBB:261:LYS:HB3	2:CCC:74:TRP:CZ2	2.40	0.56

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	AAA	68/81 (84%)	64 (94%)	4 (6%)	0	100	100
1	BBB	70/81 (86%)	70 (100%)	0	0	100	100
1	DDD	66/81 (82%)	64 (97%)	2 (3%)	0	100	100
1	EEE	71/81 (88%)	70 (99%)	1 (1%)	0	100	100
2	CCC	77/93 (83%)	72 (94%)	5 (6%)	0	100	100
2	FFF	80/93 (86%)	79 (99%)	1 (1%)	0	100	100
All	All	432/510 (85%)	419 (97%)	13 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	58/65 (89%)	56 (97%)	2 (3%)	37	35
1	BBB	59/65 (91%)	56 (95%)	3 (5%)	24	20
1	DDD	57/65 (88%)	54 (95%)	3 (5%)	22	19
1	EEE	60/65 (92%)	59 (98%)	1 (2%)	60	65
2	CCC	73/84 (87%)	71 (97%)	2 (3%)	44	46
2	FFF	74/84 (88%)	71 (96%)	3 (4%)	30	29
All	All	381/428 (89%)	367 (96%)	14 (4%)	34	32

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	DDD	203	ARG
1	DDD	226	ARG
2	FFF	92	ASN
2	FFF	66	ASP
2	FFF	75	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

There are no ligands in this entry.

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	AAA	72/81 (88%)	0.18	1 (1%) 75 80	46, 61, 89, 96	0
1	BBB	74/81 (91%)	0.37	1 (1%) 75 80	42, 53, 81, 111	0
1	DDD	70/81 (86%)	0.31	1 (1%) 75 80	44, 58, 93, 111	0
1	EEE	73/81 (90%)	0.49	3 (4%) 37 46	42, 58, 89, 101	0
2	CCC	79/93 (84%)	0.78	12 (15%) 2 2	51, 75, 100, 116	0
2	FFF	82/93 (88%)	0.11	0 100 100	42, 51, 69, 82	0
All	All	450/510 (88%)	0.37	18 (4%) 38 47	42, 59, 92, 116	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	CCC	35	LEU	5.4
2	CCC	34	ASP	4.1
1	AAA	187	LEU	3.9
1	BBB	186	GLY	3.7
1	EEE	249	VAL	3.1

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 monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.