Full wwPDB X-ray Structure Validation Report (i)

Nov 24, 2021 – 03:07 pm GMT

PDB ID	:	7NFT
Title	:	Fujian capbinding domain in complex with Nb8208
Authors	:	Keown, J.R.; Grimes, J.M.; Fodor, E.
Deposited on	:	2021-02-07
Resolution	:	3.14 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references (1)) were used in the production of this report:

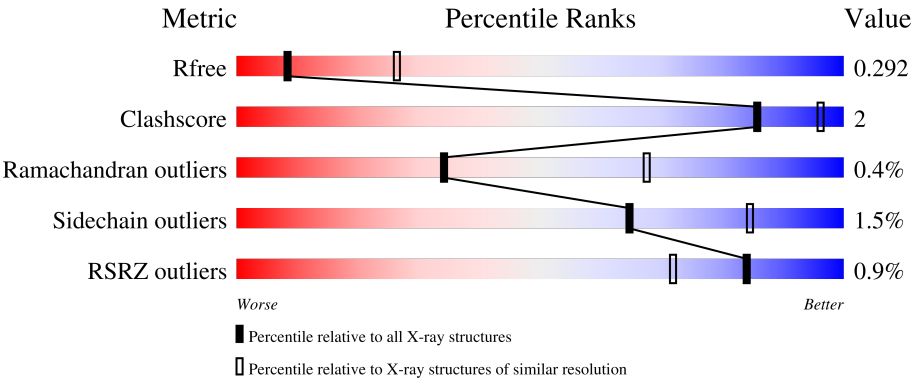
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.14 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	1626 (3.18-3.10)
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)
RSRZ outliers	127900	1588 (3.18-3.10)

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	304	47% 6% 47%
1	B	304	47% 5% 48%
2	C	132	2% 89% 9%
2	D	132	2% 86% 5% 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8839 atoms, of which 4402 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 Polymerase basic protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	161	Total	C	H	N	O	S	0	0	0
			2587	802	1311	234	231	9			
1	B	159	Total	C	H	N	O	S	0	0	0
			2556	790	1297	232	228	9			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	233	HIS	-	expression tag	UNP Q6E3N3
A	234	HIS	-	expression tag	UNP Q6E3N3
A	235	HIS	-	expression tag	UNP Q6E3N3
A	236	HIS	-	expression tag	UNP Q6E3N3
A	237	HIS	-	expression tag	UNP Q6E3N3
A	238	HIS	-	expression tag	UNP Q6E3N3
A	239	GLY	-	expression tag	UNP Q6E3N3
A	240	GLU	-	expression tag	UNP Q6E3N3
A	241	ASN	-	expression tag	UNP Q6E3N3
A	242	LEU	-	expression tag	UNP Q6E3N3
A	243	TYR	-	expression tag	UNP Q6E3N3
A	244	PHE	-	expression tag	UNP Q6E3N3
A	245	GLN	-	expression tag	UNP Q6E3N3
A	246	GLY	-	expression tag	UNP Q6E3N3
B	233	HIS	-	expression tag	UNP Q6E3N3
B	234	HIS	-	expression tag	UNP Q6E3N3
B	235	HIS	-	expression tag	UNP Q6E3N3
B	236	HIS	-	expression tag	UNP Q6E3N3
B	237	HIS	-	expression tag	UNP Q6E3N3
B	238	HIS	-	expression tag	UNP Q6E3N3
B	239	GLY	-	expression tag	UNP Q6E3N3
B	240	GLU	-	expression tag	UNP Q6E3N3
B	241	ASN	-	expression tag	UNP Q6E3N3
B	242	LEU	-	expression tag	UNP Q6E3N3
B	243	TYR	-	expression tag	UNP Q6E3N3

Continued on next page...

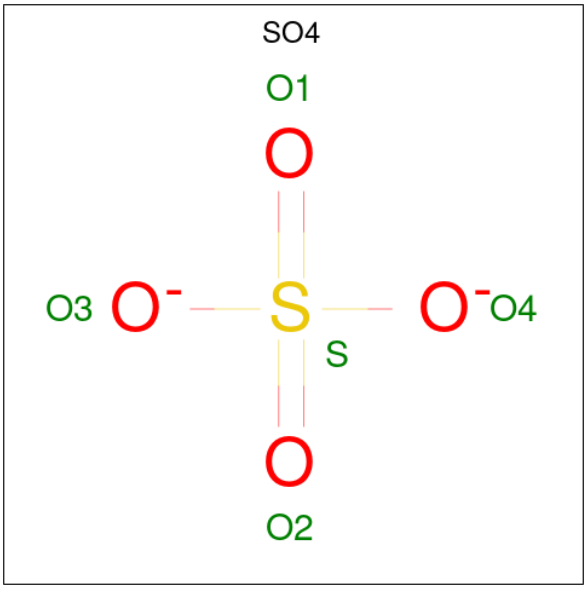
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	244	PHE	-	expression tag	UNP Q6E3N3
B	245	GLN	-	expression tag	UNP Q6E3N3
B	246	GLY	-	expression tag	UNP Q6E3N3

- Molecule 2 is a protein called Nb8208.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	120	Total	C	H	N	O	S	0	0	0
			1823	580	897	165	176	5			
2	D	120	Total	C	H	N	O	S	0	0	0
			1823	580	897	165	176	5			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

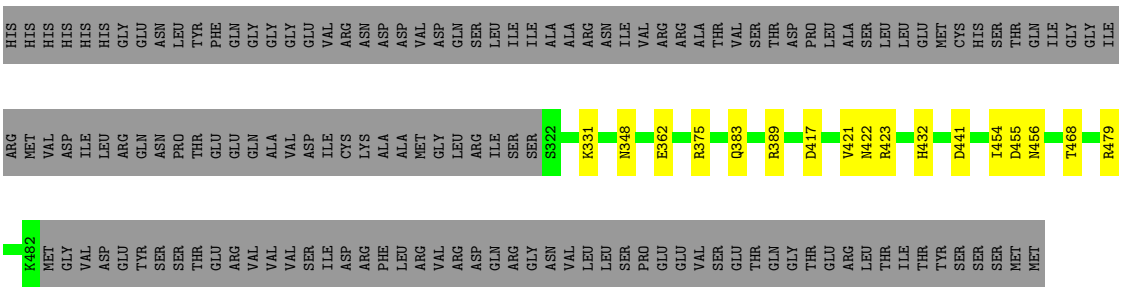
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0

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.

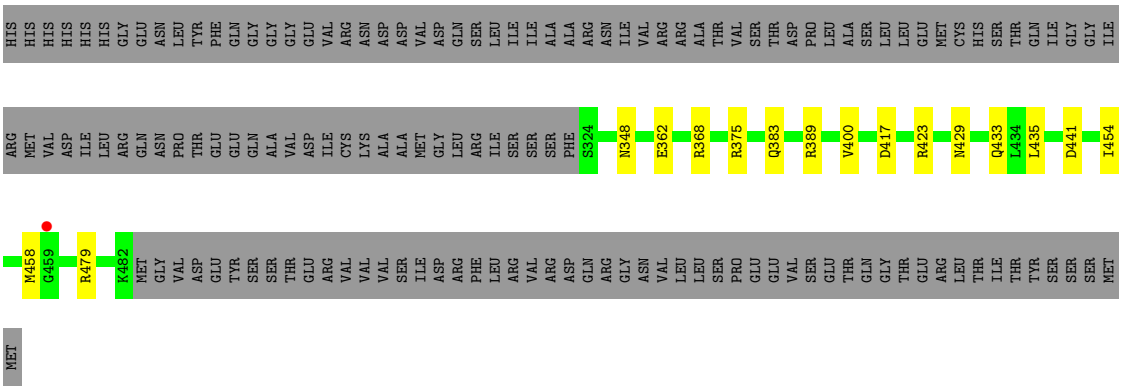
• Molecule 1: Polymerase basic protein 2

Chain A:



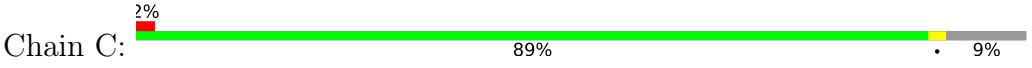
• Molecule 1: Polymerase basic protein 2

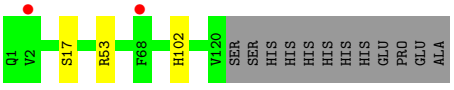
Chain B:



• Molecule 2: Nb8208

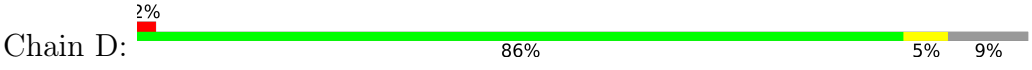
Chain C:

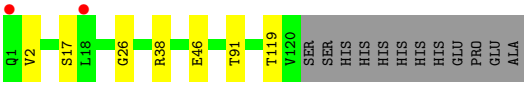




• Molecule 2: Nb8208

Chain D:




Q1
V2
S17
L18
G26
R38
E46
T91
T119
V120
SER
HIS
HIS
HIS
HIS
HIS
GLU
PRO
GLU
ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 6	Depositor
Cell constants a, b, c, α, β, γ	128.19Å 128.19Å 70.27Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	70.27 – 3.14 111.02 – 3.14	Depositor EDS
% Data completeness (in resolution range)	93.0 (70.27-3.14) 87.9 (111.02-3.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.60 (at 3.13Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.244 , 0.291 0.244 , 0.292	Depositor DCC
R_{free} test set	582 reflections (5.38%)	wwPDB-VP
Wilson B-factor (Å^2)	60.8	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.077 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	8839	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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: SO4

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	0.25	0/1294	0.53	0/1735
1	B	0.25	0/1276	0.53	0/1711
2	C	0.26	0/945	0.52	0/1275
2	D	0.26	0/945	0.53	0/1275
All	All	0.26	0/4460	0.53	0/5996

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	1276	1311	1311	8	2
1	B	1259	1297	1297	7	0
2	C	926	897	897	2	0
2	D	926	897	897	3	0
3	A	15	0	0	2	0
3	B	15	0	0	1	0
3	C	10	0	0	0	0
3	D	10	0	0	0	0
All	All	4437	4402	4402	19	2

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 2.

All (19) 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:348:ASN:ND2	1:A:417:ASP:OD1	2.26	0.68
1:A:423:ARG:NH2	3:A:603:SO4:O2	2.28	0.66
1:B:423:ARG:NE	3:B:603:SO4:O3	2.30	0.63
1:B:348:ASN:ND2	1:B:417:ASP:OD1	2.39	0.54
1:B:454:ILE:HG22	1:B:479:ARG:CZ	2.38	0.54
2:D:38:ARG:O	2:D:46:GLU:N	2.39	0.54
1:A:375:ARG:NE	1:A:383:GLN:OE1	2.41	0.53
1:B:362:GLU:OE2	1:B:375:ARG:NH1	2.46	0.48
2:C:102:HIS:O	2:C:102:HIS:ND1	2.48	0.47
1:A:454:ILE:HG22	1:A:479:ARG:CZ	2.46	0.46
1:A:468:THR:HG23	2:C:53:ARG:HG3	1.97	0.46
2:D:91:THR:HG23	2:D:119:THR:HA	2.00	0.44
1:B:429:ASN:O	1:B:433:GLN:N	2.44	0.43
1:B:400:VAL:HG22	1:B:435:LEU:HD11	2.01	0.43
1:A:362:GLU:OE2	1:A:375:ARG:NH1	2.51	0.43
2:D:2:VAL:HG22	2:D:26:GLY:HA3	2.01	0.43
1:B:375:ARG:NE	1:B:383:GLN:OE1	2.52	0.42
1:A:432:HIS:NE2	3:A:601:SO4:O4	2.52	0.42
1:A:421:VAL:HG12	1:A:422:ASN:O	2.21	0.41

All (2) 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)
1:A:331:LYS:HZ1	1:A:455:ASP:OD2[2_545]	1.53	0.07
1:A:331:LYS:NZ	1:A:455:ASP:OD2[2_545]	2.14	0.06

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	159/304 (52%)	152 (96%)	6 (4%)	1 (1%)	25 59
1	B	157/304 (52%)	148 (94%)	8 (5%)	1 (1%)	25 59
2	C	118/132 (89%)	115 (98%)	3 (2%)	0	100 100
2	D	118/132 (89%)	115 (98%)	3 (2%)	0	100 100
All	All	552/872 (63%)	530 (96%)	20 (4%)	2 (0%)	34 67

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	456	ASN
1	B	458	MET

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	140/266 (53%)	138 (99%)	2 (1%)	67 85
1	B	138/266 (52%)	135 (98%)	3 (2%)	52 77
2	C	94/105 (90%)	93 (99%)	1 (1%)	73 88
2	D	94/105 (90%)	93 (99%)	1 (1%)	73 88
All	All	466/742 (63%)	459 (98%)	7 (2%)	65 84

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	389	ARG
1	A	441	ASP
2	C	17	SER
1	B	368	ARG
1	B	389	ARG
1	B	441	ASP
2	D	17	SER

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

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

10 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, 95^{th} 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	161/304 (52%)	0.28	0 100 100	53, 65, 82, 92	0
1	B	159/304 (52%)	0.39	1 (0%) 89 80	56, 74, 89, 114	0
2	C	120/132 (90%)	0.56	2 (1%) 70 51	58, 70, 86, 103	0
2	D	120/132 (90%)	0.59	2 (1%) 70 51	59, 72, 101, 124	0
All	All	560/872 (64%)	0.44	5 (0%) 84 72	53, 71, 90, 124	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	1	GLN	2.6
2	C	2	VAL	2.4
2	D	18	LEU	2.3
1	B	459	GLY	2.1
2	C	68	PHE	2.1

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

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

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, 95^{th} 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	SO4	C	202	5/5	0.90	0.18	68,72,78,80	0
3	SO4	B	601	5/5	0.92	0.15	58,61,65,66	0
3	SO4	D	201	5/5	0.93	0.22	56,60,61,65	0
3	SO4	C	201	5/5	0.94	0.16	49,49,52,59	0
3	SO4	B	602	5/5	0.94	0.13	60,64,67,72	0
3	SO4	A	603	5/5	0.94	0.19	55,61,65,66	0
3	SO4	D	202	5/5	0.94	0.14	72,73,80,81	0
3	SO4	A	601	5/5	0.95	0.11	55,56,59,65	0
3	SO4	A	602	5/5	0.96	0.18	55,61,66,68	0
3	SO4	B	603	5/5	0.98	0.10	47,47,49,53	0

6.5 Other polymers (i)

There are no such residues in this entry.