Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 06:03 PM GMT

PDB ID	:	4KF8
Title	:	Nup188(aa1445-1827) from Myceliophthora thermophila
Authors	:	Schwartz, T.U.; Andersen, K.R.
Deposited on	:	2013-04-26
Resolution	:	3.00 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
http://wwpdb.org/validation/2016/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:

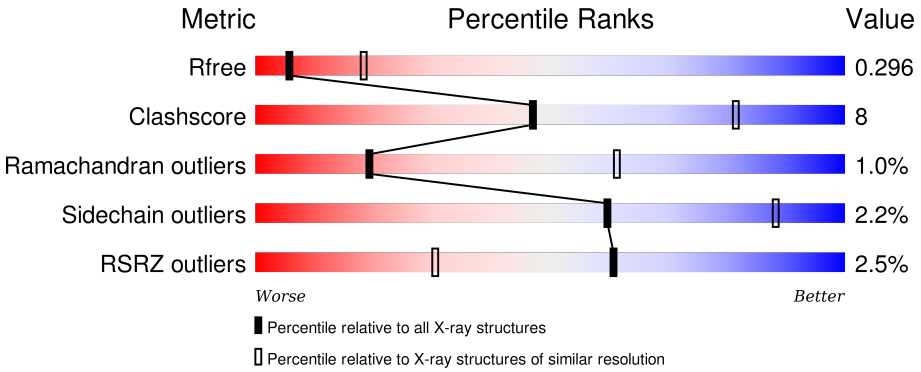
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	383	2% 70% 18% • 10%
1	B	383	2% 56% 11% 33%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4366 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 Nup188.

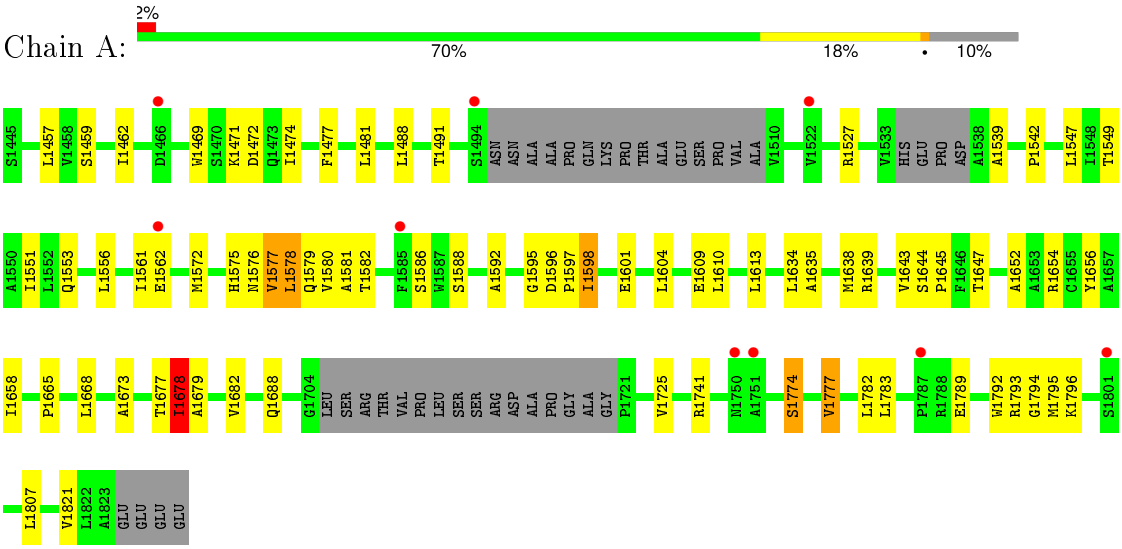
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	344	Total	C	N	O	S	Se	0	0	0
			2526	1626	427	466	4	3			
1	B	258	Total	C	N	O	S	Se	0	0	0
			1840	1181	316	337	4	2			

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

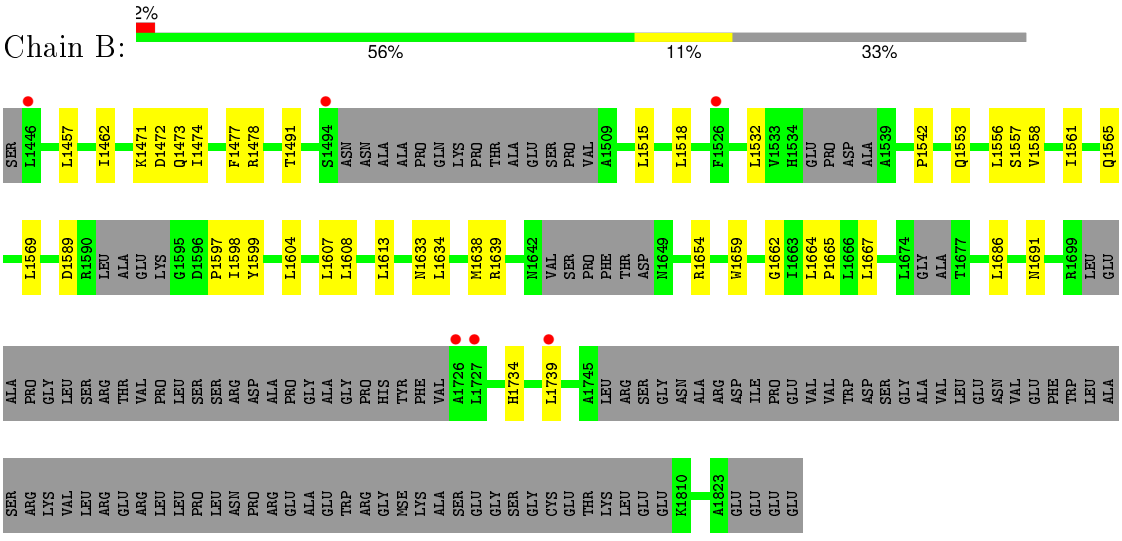
- Molecule 1: Nup188

Chain A:



- Molecule 1: Nup188

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	64.30Å 66.65Å 162.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	81.13 – 3.00 81.13 – 2.74	Depositor EDS
% Data completeness (in resolution range)	99.8 (81.13-3.00) 94.5 (81.13-2.74)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 2.73Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.275 , 0.294 0.283 , 0.296	Depositor DCC
R_{free} test set	1424 reflections (11.12%)	DCC
Wilson B-factor (Å^2)	80.8	Xtriage
Anisotropy	0.670	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 69.1	EDS
Estimated twinning fraction	0.007 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Outliers	1 of 17936 reflections (0.006%)	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4366	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.92% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.22	0/2562	0.44	0/3497
1	B	0.22	0/1856	0.41	0/2528
All	All	0.22	0/4418	0.43	0/6025

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	2526	0	2563	44	0
1	B	1840	0	1858	24	0
All	All	4366	0	4421	66	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 8.

All (66) 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:1601:GLU:OE1	1:A:1654:ARG:NH1	2.20	0.74
1:A:1639:ARG:NH1	1:A:1688:GLN:O	2.27	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1665:PRO:HA	1:A:1668:LEU:HD12	1.77	0.66
1:B:1639:ARG:O	1:B:1691:ASN:ND2	2.33	0.62
1:B:1553:GLN:O	1:B:1557:SER:OG	2.18	0.60
1:A:1794:GLY:O	1:B:1654:ARG:NH2	2.34	0.60
1:B:1558:VAL:HG22	1:B:1561:ILE:HB	1.84	0.59
1:A:1596:ASP:HB2	1:A:1654:ARG:HG3	1.85	0.58
1:A:1634:LEU:HB3	1:A:1638:MSE:HE2	1.84	0.58
1:A:1679:ALA:HA	1:A:1682:VAL:HG22	1.86	0.57
1:B:1598:ILE:HB	1:B:1654:ARG:HH11	1.70	0.57
1:A:1741:ARG:HE	1:A:1821:VAL:HG13	1.70	0.56
1:A:1572:MSE:HE2	1:A:1610:LEU:HD11	1.88	0.56
1:A:1577:VAL:HG23	1:A:1578:LEU:H	1.71	0.56
1:A:1561:ILE:HG22	1:A:1562:GLU:H	1.71	0.55
1:A:1459:SER:O	1:A:1462:ILE:HG22	2.06	0.54
1:A:1597:PRO:HG2	1:A:1634:LEU:HD21	1.89	0.54
1:A:1457:LEU:HD21	1:B:1457:LEU:HD21	1.90	0.54
1:A:1673:ALA:HB3	1:A:1678:ILE:HD11	1.90	0.54
1:A:1645:PRO:HG3	1:A:1656:TYR:CE2	2.42	0.54
1:B:1608:LEU:HD11	1:B:1665:PRO:HB2	1.89	0.53
1:B:1471:LYS:HD3	1:B:1471:LYS:H	1.73	0.53
1:B:1686:LEU:HD11	1:B:1739:LEU:HD23	1.91	0.53
1:A:1635:ALA:HA	1:A:1638:MSE:HE3	1.93	0.51
1:B:1589:ASP:OD1	1:B:1633:ASN:ND2	2.41	0.51
1:B:1638:MSE:HE3	1:B:1659:TRP:HB2	1.92	0.51
1:A:1488:LEU:O	1:A:1491:THR:OG1	2.26	0.50
1:B:1604:LEU:HD23	1:B:1607:LEU:HD12	1.93	0.50
1:A:1604:LEU:HD12	1:A:1658:ILE:HG23	1.93	0.50
1:A:1527:ARG:HD3	1:A:1575:HIS:HB3	1.93	0.50
1:B:1542:PRO:HG3	1:B:1599:TYR:CE2	2.47	0.49
1:A:1576:ASN:HB3	1:A:1580:VAL:HG23	1.94	0.49
1:A:1639:ARG:HH12	1:A:1688:GLN:HB3	1.78	0.49
1:A:1774:SER:HB2	1:A:1777:VAL:HG12	1.95	0.49
1:A:1782:LEU:HD21	1:A:1807:LEU:HG	1.93	0.49
1:A:1549:THR:O	1:A:1553:GLN:HG3	2.14	0.48
1:A:1592:ALA:HB3	1:A:1597:PRO:HA	1.96	0.47
1:A:1556:LEU:HD13	1:A:1613:LEU:HD11	1.95	0.47
1:B:1515:LEU:HD23	1:B:1518:LEU:HD12	1.97	0.46
1:B:1556:LEU:HD13	1:B:1613:LEU:HD11	1.95	0.46
1:B:1556:LEU:HA	1:B:1561:ILE:HG21	1.97	0.46
1:A:1579:GLN:O	1:A:1582:THR:OG1	2.30	0.46
1:B:1462:ILE:HG23	1:B:1477:PHE:HE1	1.80	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1677:THR:C	1:A:1679:ALA:H	2.20	0.45
1:B:1472:ASP:OD1	1:B:1473:GLN:N	2.43	0.45
1:A:1793:ARG:HA	1:A:1807:LEU:HB2	1.99	0.45
1:A:1795:MSE:HG2	1:A:1796:LYS:H	1.83	0.44
1:A:1471:LYS:HD2	1:A:1539:ALA:HB2	1.99	0.44
1:B:1474:ILE:HG12	1:B:1478:ARG:HH21	1.82	0.44
1:A:1644:SER:HB3	1:A:1647:THR:HG23	1.99	0.44
1:B:1664:LEU:HD23	1:B:1667:LEU:HD12	2.01	0.43
1:A:1677:THR:O	1:A:1679:ALA:N	2.47	0.43
1:A:1469:TRP:HA	1:A:1474:ILE:HG13	2.01	0.43
1:A:1578:LEU:HA	1:A:1581:ALA:HB3	2.01	0.43
1:A:1789:GLU:HA	1:A:1792:TRP:CE2	2.54	0.43
1:A:1527:ARG:HD3	1:A:1576:ASN:H	1.84	0.42
1:A:1542:PRO:HG3	1:A:1598:ILE:HD13	2.01	0.42
1:B:1597:PRO:HG3	1:B:1634:LEU:HD23	2.02	0.42
1:B:1471:LYS:HE2	1:B:1471:LYS:HB2	1.78	0.42
1:A:1643:VAL:HA	1:A:1652:ALA:HB1	2.01	0.42
1:A:1477:PHE:O	1:A:1481:LEU:HG	2.20	0.42
1:B:1565:GLN:O	1:B:1569:LEU:HG	2.19	0.41
1:A:1547:LEU:O	1:A:1551:ILE:HG13	2.21	0.41
1:A:1596:ASP:HA	1:A:1597:PRO:HD3	1.91	0.40
1:A:1553:GLN:HG2	1:A:1609:GLU:CD	2.41	0.40
1:B:1532:LEU:H	1:B:1532:LEU:HG	1.74	0.40

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	A	336/383 (88%)	308 (92%)	24 (7%)	4 (1%)	16	56
1	B	242/383 (63%)	225 (93%)	15 (6%)	2 (1%)	24	66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	578/766 (76%)	533 (92%)	39 (7%)	6 (1%)	19 61

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1678	ILE
1	A	1588	SER
1	A	1577	VAL
1	B	1491	THR
1	A	1595	GLY
1	B	1662	GLY

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	262/311 (84%)	253 (97%)	9 (3%)	44 81
1	B	186/311 (60%)	185 (100%)	1 (0%)	92 98
All	All	448/622 (72%)	438 (98%)	10 (2%)	60 88

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1472	ASP
1	A	1578	LEU
1	A	1586	SER
1	A	1598	ILE
1	A	1678	ILE
1	A	1725	VAL
1	A	1774	SER
1	A	1777	VAL
1	A	1783	LEU
1	B	1734	HIS

Some 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 carbohydrates 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

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	341/383 (89%)	-0.17	9 (2%)	59	29	47, 90, 130, 161	0
1	B	256/383 (66%)	-0.13	6 (2%)	64	33	44, 85, 158, 189	0
All	All	597/766 (77%)	-0.16	15 (2%)	61	30	44, 89, 148, 189	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1494	SER	4.7
1	B	1739	LEU	4.4
1	A	1562	GLU	4.0
1	A	1494	SER	3.3
1	A	1751	ALA	2.8
1	A	1466	ASP	2.8
1	A	1801	SER	2.7
1	A	1585	PHE	2.7
1	B	1526	PHE	2.5
1	B	1726	ALA	2.5
1	B	1446	LEU	2.2
1	B	1727	LEU	2.2
1	A	1750	ASN	2.2
1	A	1787	PRO	2.0
1	A	1522	VAL	2.0

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.