wwPDB X-ray Structure Validation Summary Report (i)

Apr 28, 2024 – 02:26 am BST

PDB ID	:	4B8B
Title	:	N-Terminal domain of the yeast Not1
Authors	:	Basquin, J.; Conti, E.
Deposited on	:	2012-08-26
Resolution	:	2.80 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:

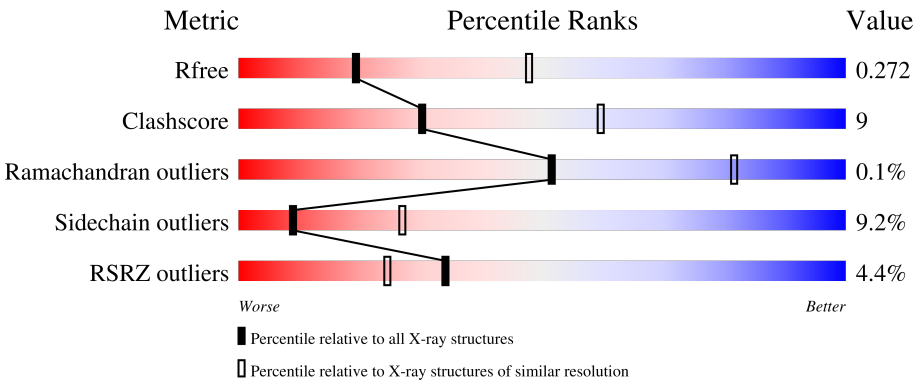
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	603	3% 63% 23% • 12%
1	B	603	4% 63% 23% • 12%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	AU	A	1750	-	-	-	X
2	AU	B	1751	-	-	-	X
2	AU	B	1800	-	-	-	X

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8410 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 GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	528	4179	2705	680	776	18	0	0	0
1	B	529	4202	2716	687	781	18	0	0	0

- Molecule 2 is GOLD ION (three-letter code: AU) (formula: Au).

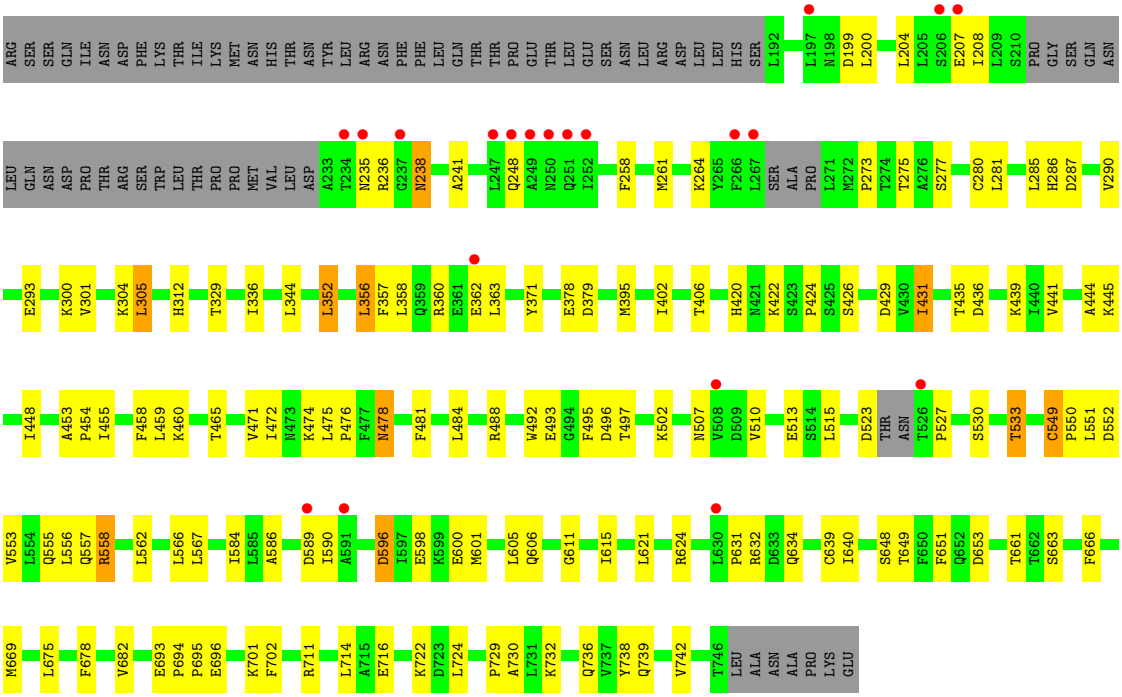
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Au		
2	A	15	15	15	0	0
2	B	14	14	14	0	0

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.

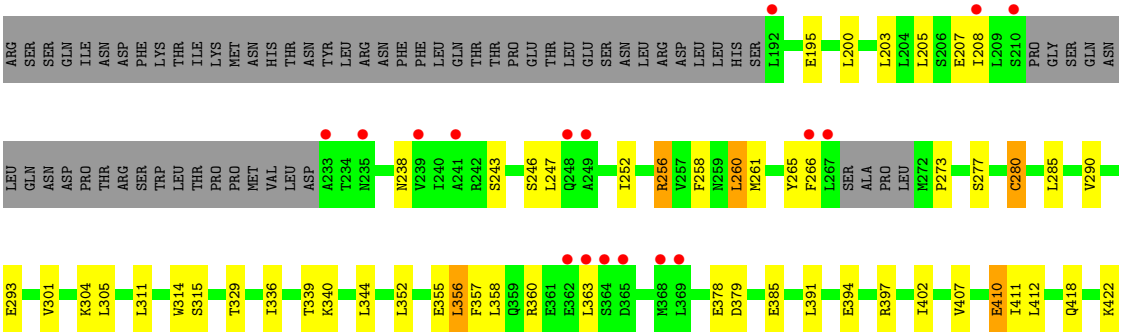
• Molecule 1: GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1

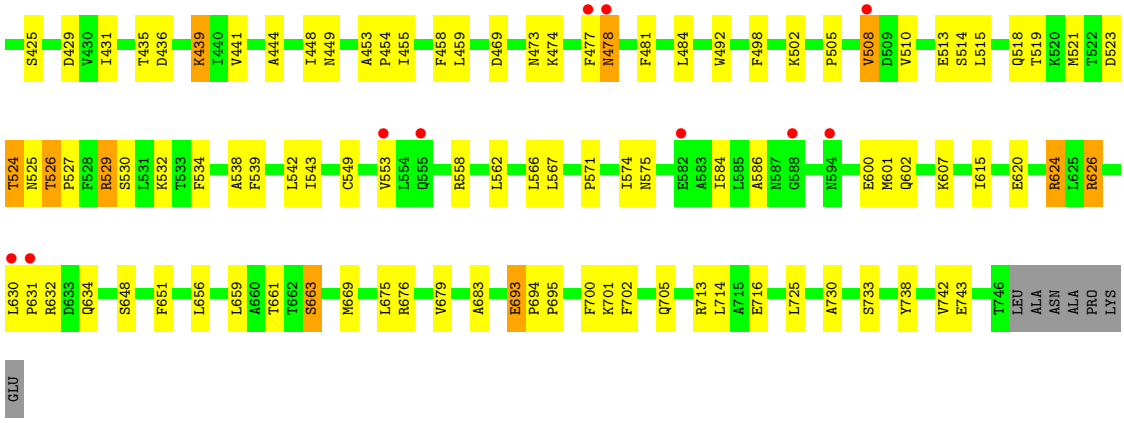
Chain A:



• Molecule 1: GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1

Chain B:




S425
D429
V430
I431
T435
D436
K439
I440
V441
A444
I448
N449
A453
P454
I455
F458
L459
D469
N473
K474
F477
N478
F481
L484
W492
F498
K502
P505
V508
D509
V510
E513
S514
L515
Q518
T519
K520
M521
T522
D523
T524
N525
T526
P527
F528
R529
S530
L531
K532
T533
F534
A538
F539
L542
I543
C549
V553
L554
Q555
R558
L562
L566
L567
P571
I574
N575
E582
A583
I584
L585
A586
N587
G588
N594
E600
M601
Q602
K607
I615
E620
R624
L625
R626
L630
P631
R632
D633
Q634
S648
F651
L656
L659
A660
T661
T662
S663
M669
L675
R676
V679
A683
E693
P694
P695
F700
K701
F702
Q705
R713
L714
A715
E716
L725
A730
S733
Y738
V742
E743
T746
LEU
ALA
ASN
ALA
PRO
LYS
GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.10Å 164.51Å 86.14Å 90.00° 100.05° 90.00°	Depositor
Resolution (Å)	46.05 – 2.80 46.05 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.05-2.80) 99.8 (46.05-2.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.36 (at 2.81Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.238 , 0.276 0.235 , 0.272	Depositor DCC
R_{free} test set	1855 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	61.4	Xtriage
Anisotropy	0.607	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 64.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	8410	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: AU

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	0.47	0/4257	0.64	0/5768
1	B	0.45	0/4281	0.63	0/5800
All	All	0.46	0/8538	0.64	0/11568

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	264	LYS	Peptide

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	A	4179	0	4197	75	0
1	B	4202	0	4235	85	0
2	A	15	0	0	0	0
2	B	14	0	0	0	0
All	All	8410	0	8432	157	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 9.

The worst 5 of 157 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:B:575:ASN:ND2	1:B:634:GLN:OE1	2.10	0.83
1:A:533:THR:HG21	1:A:590:ILE:HD13	1.60	0.83
1:A:455:ILE:HG23	1:A:458:PHE:HB3	1.66	0.76
1:B:478:ASN:HB2	1:B:510:VAL:HG11	1.71	0.71
1:B:336:ILE:HD11	1:B:429:ASP:HB3	1.73	0.70

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	520/603 (86%)	500 (96%)	19 (4%)	1 (0%)	47 78
1	B	523/603 (87%)	501 (96%)	22 (4%)	0	100 100
All	All	1043/1206 (86%)	1001 (96%)	41 (4%)	1 (0%)	51 81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	611	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	461/544 (85%)	415 (90%)	46 (10%)	7 22
1	B	466/544 (86%)	427 (92%)	39 (8%)	11 31
All	All	927/1088 (85%)	842 (91%)	85 (9%)	9 27

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	356	LEU
1	B	529	ARG
1	B	425	SER
1	B	478	ASN
1	B	601	MET

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)

Of 29 ligands modelled in this entry, 29 are monoatomic - leaving 0 for Mogul analysis.

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 (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	A	528/603 (87%)	0.19	20 (3%)	40	30	28, 54, 98, 145	0
1	B	529/603 (87%)	0.24	27 (5%)	28	19	35, 54, 102, 169	0
All	All	1057/1206 (87%)	0.21	47 (4%)	34	24	28, 54, 101, 169	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	266	PHE	5.4
1	B	553	VAL	4.8
1	A	267	LEU	4.3
1	A	248	GLN	4.3
1	B	362	GLU	3.9

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)

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, 95th 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
2	AU	B	1800	1/1	0.15	0.73	543,543,543,543	0
2	AU	B	1756	1/1	0.27	0.28	154,154,154,154	1
2	AU	A	1802	1/1	0.31	0.22	147,147,147,147	1
2	AU	B	1750	1/1	0.56	0.20	127,127,127,127	1
2	AU	A	1750	1/1	0.63	0.49	100,100,100,100	1
2	AU	A	1803	1/1	0.65	0.11	138,138,138,138	1
2	AU	A	1751	1/1	0.67	0.38	217,217,217,217	1
2	AU	B	1801	1/1	0.73	0.28	180,180,180,180	1
2	AU	A	1757	1/1	0.74	0.20	120,120,120,120	1
2	AU	B	1803	1/1	0.76	0.21	118,118,118,118	1
2	AU	B	1751	1/1	0.79	0.49	156,156,156,156	1
2	AU	B	1754	1/1	0.80	0.25	120,120,120,120	1
2	AU	B	1748	1/1	0.81	0.08	67,67,67,67	1
2	AU	A	1756	1/1	0.83	0.27	115,115,115,115	1
2	AU	A	1755	1/1	0.83	0.26	103,103,103,103	1
2	AU	B	1753	1/1	0.85	0.18	93,93,93,93	1
2	AU	A	1754	1/1	0.90	0.08	101,101,101,101	1
2	AU	A	1801	1/1	0.91	0.07	127,127,127,127	1
2	AU	A	1748	1/1	0.93	0.06	77,77,77,77	1
2	AU	A	1753	1/1	0.93	0.14	71,71,71,71	1
2	AU	B	1802	1/1	0.93	0.10	82,82,82,82	1
2	AU	A	1747	1/1	0.93	0.11	66,66,66,66	1
2	AU	A	1800	1/1	0.95	0.04	76,76,76,76	0
2	AU	A	1749	1/1	0.95	0.12	5,5,5,5	1
2	AU	A	1752	1/1	0.96	0.20	71,71,71,71	1
2	AU	B	1755	1/1	0.97	0.13	139,139,139,139	1
2	AU	B	1747	1/1	0.98	0.12	63,63,63,63	1
2	AU	B	1749	1/1	0.99	0.14	1,1,1,1	1
2	AU	B	1752	1/1	0.99	0.90	38,38,38,38	1

6.5 Other polymers (i)

There are no such residues in this entry.