WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 27, 2014 – 01:52 PM GMT

PDB ID	:	2V71
Title	:	COILED-COIL REGION OF NUDEL
Authors	:	Derewenda, U.; Cooper, D.R.; Kim, M.H.; Derewenda, Z.S.
Deposited on	:	2007-07-25
Resolution	:	2.24 Å(reported)

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

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.24 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	Percentile Ranks	Value
Rfree		0.292
Clashscore		2
Ramachandran outliers		0.6%
Sidechain outliers		4.9%
RSRZ outliers		15.9%

Worse ← → Better

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
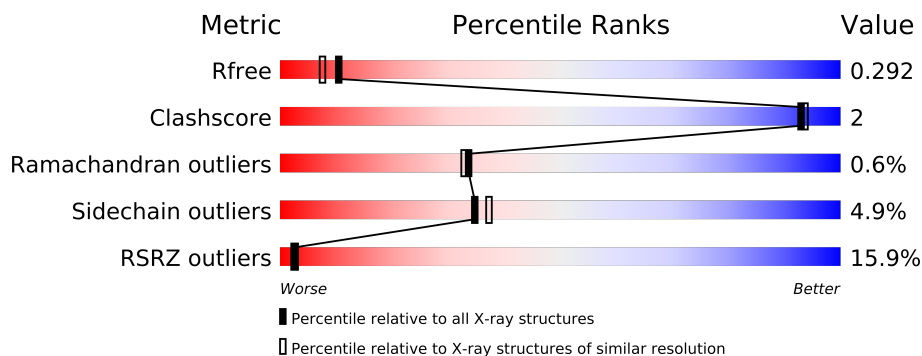
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	66092	1112 (2.26-2.22)
Clashscore	79885	1317 (2.26-2.22)
Ramachandran outliers	78287	1282 (2.26-2.22)
Sidechain outliers	78261	1282 (2.26-2.22)
RSRZ outliers	66119	1112 (2.26-2.22)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	189	
1	B	189	

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 2671 atoms, of which 0 are hydrogen and 0 are deuterium.

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 NUCLEAR DISTRIBUTION PROTEIN NUDE-LIKE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	160	Total	C	N	O	Se	0	0	0
			1334	824	231	277	2			
1	B	160	Total	C	N	O	Se	0	0	0
			1337	826	231	278	2			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	MSE	LEU	ENGINEERED MUTATION	UNP Q78PB6
A	134	MSE	VAL	ENGINEERED MUTATION	UNP Q78PB6
B	24	MSE	LEU	ENGINEERED MUTATION	UNP Q78PB6
B	134	MSE	VAL	ENGINEERED MUTATION	UNP Q78PB6

3 Residue-property plots ⓘ

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.
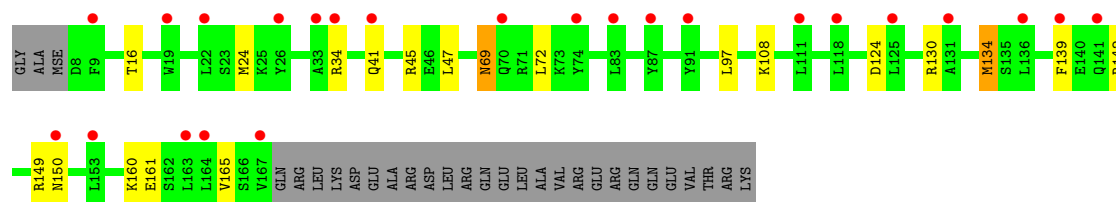
• Molecule 1: NUCLEAR DISTRIBUTION PROTEIN NUDE-LIKE 1

Chain A:
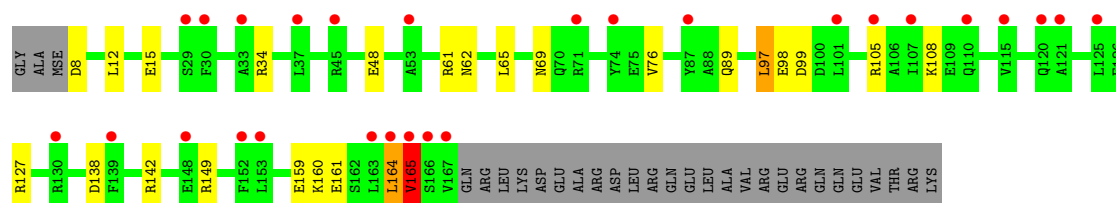
• Molecule 1: NUCLEAR DISTRIBUTION PROTEIN NUDE-LIKE 1

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	89.87Å 44.98Å 124.38Å 90.00° 97.35° 90.00°	Depositor
Resolution (Å)	123.09 – 2.24 43.74 – 2.24	Depositor EDS
% Data completeness (in resolution range)	95.3 (123.09-2.24) 95.3 (43.74-2.24)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.74 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.253 , 0.290 0.257 , 0.292	Depositor DCC
R_{free} test set	1172 reflections (5.36%)	DCC
Wilson B-factor (Å^2)	54.6	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 34.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Outliers	1 of 23035 reflections (0.004%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2671	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.35% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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	0.56	1/1346 (0.1%)	1.27	14/1800 (0.8%)
1	B	0.53	0/1349	1.27	12/1804 (0.7%)
All	All	0.54	1/2695 (0.0%)	1.27	26/3604 (0.7%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	62	ASN	CG-OD1	5.28	1.35	1.24

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	149	ARG	NE-CZ-NH2	-10.26	115.17	120.30
1	B	134	MSE	CA-CB-CG	8.32	127.45	113.30
1	B	45	ARG	NE-CZ-NH1	8.24	124.42	120.30
1	B	149	ARG	NE-CZ-NH1	7.75	124.17	120.30
1	B	24	MSE	CB-CG-SE	-7.65	89.74	112.70
1	B	34	ARG	NE-CZ-NH2	-7.39	116.61	120.30
1	A	8	ASP	CB-CG-OD1	7.37	124.93	118.30
1	A	127	ARG	NE-CZ-NH2	-7.28	116.66	120.30
1	B	124	ASP	CB-CG-OD1	7.26	124.83	118.30
1	B	34	ARG	NE-CZ-NH1	6.56	123.58	120.30
1	B	149	ARG	NE-CZ-NH2	-6.52	117.04	120.30
1	A	142	ARG	NE-CZ-NH2	-6.49	117.06	120.30
1	A	138	ASP	CB-CG-OD1	6.27	123.94	118.30
1	A	149	ARG	NE-CZ-NH1	6.25	123.42	120.30
1	A	99	ASP	CB-CG-OD1	6.23	123.91	118.30
1	A	97	LEU	CB-CG-CD1	-6.19	100.48	111.00
1	A	34	ARG	NE-CZ-NH1	5.95	123.28	120.30
1	A	164	LEU	CA-CB-CG	5.73	128.47	115.30
1	B	41	GLN	CB-CA-C	-5.68	99.04	110.40
1	B	130	ARG	NE-CZ-NH1	5.63	123.12	120.30
1	B	130	ARG	NE-CZ-NH2	-5.61	117.50	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	108	LYS	CD-CE-NZ	5.45	124.24	111.70
1	A	61	ARG	NE-CZ-NH1	5.45	123.02	120.30
1	A	34	ARG	NE-CZ-NH2	-5.27	117.67	120.30
1	A	105	ARG	NE-CZ-NH1	5.25	122.93	120.30
1	B	45	ARG	NE-CZ-NH2	-5.06	117.77	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1334	0	1285	6	0
1	B	1337	0	1295	8	0
All	All	2671	0	2580	9	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (9) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:15:GLU:HG2	1:B:16:THR:CG2	2.41	0.51
1:B:161:GLU:O	1:B:165:VAL:HG23	2.11	0.51
1:A:69:ASN:OD1	1:B:69:ASN:ND2	2.47	0.48
1:A:161:GLU:O	1:A:165:VAL:HG22	2.14	0.47
1:A:97:LEU:HD23	1:B:97:LEU:HD23	1.96	0.47
1:B:47:LEU:HD23	1:B:47:LEU:O	2.15	0.46
1:A:76:VAL:HG21	1:B:72:LEU:HD13	2.02	0.41
1:A:48:GLU:HG2	1:B:47:LEU:HD22	2.01	0.41
1:B:47:LEU:C	1:B:47:LEU:HD23	2.42	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	158/189 (84%)	155 (98%)	1 (1%)	2 (1%)	18	11
1	B	158/189 (84%)	158 (100%)	0	0	100	100
All	All	316/378 (84%)	313 (99%)	1 (0%)	2 (1%)	33	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	165	VAL
1	A	164	LEU

5.3.2 Protein sidechains ⓘ

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	143/167 (86%)	136 (95%)	7 (5%)	35	37
1	B	145/167 (87%)	138 (95%)	7 (5%)	35	38
All	All	288/334 (86%)	274 (95%)	14 (5%)	35	37

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	LEU
1	A	65	LEU
1	A	89	GLN
1	A	98	GLU
1	A	159	GLU
1	A	160	LYS
1	A	165	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	69	ASN
1	B	108	LYS
1	B	134	MSE
1	B	139	PHE
1	B	142	ARG
1	B	150	ASN
1	B	160	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	69	ASN

5.3.3 RNA (i)

There are no RNA chains 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 carbohydrates 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

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, 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	160/189 (84%)	1.34	27 (16%)	2	2	52, 67, 71, 74	0
1	B	160/189 (84%)	1.15	24 (15%)	3	3	53, 67, 71, 73	0
All	All	320/378 (84%)	1.24	51 (15%)	3	2	52, 67, 71, 74	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	164	LEU	10.8
1	A	163	LEU	9.2
1	A	167	VAL	6.7
1	B	164	LEU	4.2
1	B	139	PHE	4.1
1	A	165	VAL	4.0
1	A	74	TYR	3.9
1	B	74	TYR	3.7
1	A	101	LEU	3.6
1	A	115	VAL	3.3
1	A	166	SER	3.3
1	B	136	LEU	3.1
1	B	125	LEU	3.0
1	A	139	PHE	2.9
1	A	130	ARG	2.9
1	B	34	ARG	2.8
1	B	22	LEU	2.8
1	B	41	GLN	2.7
1	B	141	GLN	2.7
1	B	33	ALA	2.7
1	A	125	LEU	2.7
1	B	87	TYR	2.7
1	B	19	TRP	2.7
1	B	163	LEU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	105	ARG	2.7
1	B	131	ALA	2.7
1	A	71	ARG	2.7
1	B	70	GLN	2.5
1	B	83	LEU	2.5
1	B	153	LEU	2.5
1	A	121	ALA	2.5
1	B	150	ASN	2.4
1	B	9	PHE	2.4
1	A	30	PHE	2.3
1	A	33	ALA	2.3
1	A	153	LEU	2.3
1	A	110	GLN	2.3
1	B	111	LEU	2.3
1	B	91	TYR	2.2
1	B	167	VAL	2.2
1	A	37	LEU	2.2
1	B	26	TYR	2.2
1	A	152	PHE	2.2
1	A	148	GLU	2.1
1	A	87	TYR	2.1
1	A	45	ARG	2.1
1	A	53	ALA	2.1
1	A	120	GLN	2.1
1	A	29	SER	2.1
1	A	107	ILE	2.1
1	B	118	LEU	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 (i)

There are no such residues in this entry.