wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 01:03 am BST

PDB ID	:	5N48
Title	:	Structure of Anticalin N9B in complex with extra-domain B of human oncofetal fibronectin
Authors	:	Schiefner, A.; Skerra, A.
Deposited on	:	2017-02-10
Resolution	:	1.60 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 following versions of software and data (see references (1)) were used in the production of this report:

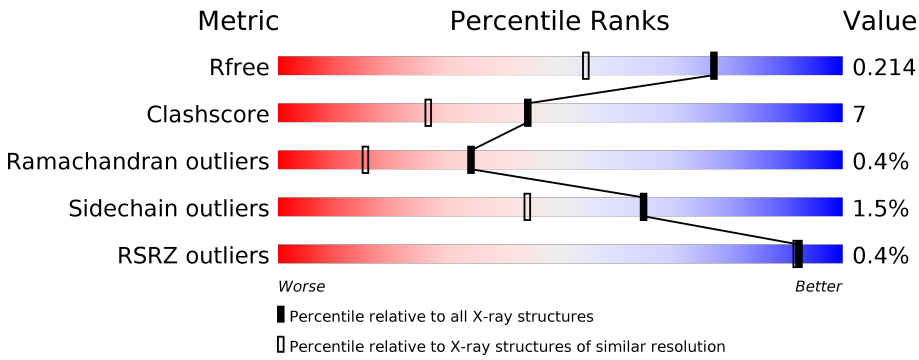
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 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	188	% 78% 14% • 7%
1	C	188	79% 12% • 7%
2	B	106	75% 10% • 13%
2	D	106	% 75% 13% 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4513 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 Neutrophil gelatinase-associated lipocalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	175	1419	913	244	255	7	0	0	0
			Total	C	N	O	S			
1	C	174	1413	910	243	253	7	0	0	0

There are 62 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	HIS	GLN	engineered mutation	UNP P80188
A	36	ARG	LEU	engineered mutation	UNP P80188
A	40	MET	ALA	engineered mutation	UNP P80188
A	41	ARG	ILE	engineered mutation	UNP P80188
A	49	ALA	GLN	engineered mutation	UNP P80188
A	52	VAL	TYR	engineered mutation	UNP P80188
A	68	LYS	SER	engineered mutation	UNP P80188
A	70	MET	LEU	engineered mutation	UNP P80188
A	72	GLN	ARG	engineered mutation	UNP P80188
A	73	ARG	LYS	engineered mutation	UNP P80188
A	77	LYS	ASP	engineered mutation	UNP P80188
A	79	MET	TRP	engineered mutation	UNP P80188
A	81	ASN	ARG	engineered mutation	UNP P80188
A	87	SER	CYS	engineered mutation	UNP P80188
A	96	ALA	ASN	engineered mutation	UNP P80188
A	100	PRO	TYR	engineered mutation	UNP P80188
A	103	PRO	LEU	engineered mutation	UNP P80188
A	106	THR	TYR	engineered mutation	UNP P80188
A	125	HIS	LYS	engineered mutation	UNP P80188
A	127	PHE	SER	engineered mutation	UNP P80188
A	134	HIS	LYS	engineered mutation	UNP P80188
A	179	SER	-	expression tag	UNP P80188
A	180	ALA	-	expression tag	UNP P80188
A	181	TRP	-	expression tag	UNP P80188
A	182	SER	-	expression tag	UNP P80188

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	183	HIS	-	expression tag	UNP P80188
A	184	PRO	-	expression tag	UNP P80188
A	185	GLN	-	expression tag	UNP P80188
A	186	PHE	-	expression tag	UNP P80188
A	187	GLU	-	expression tag	UNP P80188
A	188	LYS	-	expression tag	UNP P80188
C	28	HIS	GLN	engineered mutation	UNP P80188
C	36	ARG	LEU	engineered mutation	UNP P80188
C	40	MET	ALA	engineered mutation	UNP P80188
C	41	ARG	ILE	engineered mutation	UNP P80188
C	49	ALA	GLN	engineered mutation	UNP P80188
C	52	VAL	TYR	engineered mutation	UNP P80188
C	68	LYS	SER	engineered mutation	UNP P80188
C	70	MET	LEU	engineered mutation	UNP P80188
C	72	GLN	ARG	engineered mutation	UNP P80188
C	73	ARG	LYS	engineered mutation	UNP P80188
C	77	LYS	ASP	engineered mutation	UNP P80188
C	79	MET	TRP	engineered mutation	UNP P80188
C	81	ASN	ARG	engineered mutation	UNP P80188
C	87	SER	CYS	engineered mutation	UNP P80188
C	96	ALA	ASN	engineered mutation	UNP P80188
C	100	PRO	TYR	engineered mutation	UNP P80188
C	103	PRO	LEU	engineered mutation	UNP P80188
C	106	THR	TYR	engineered mutation	UNP P80188
C	125	HIS	LYS	engineered mutation	UNP P80188
C	127	PHE	SER	engineered mutation	UNP P80188
C	134	HIS	LYS	engineered mutation	UNP P80188
C	179	SER	-	expression tag	UNP P80188
C	180	ALA	-	expression tag	UNP P80188
C	181	TRP	-	expression tag	UNP P80188
C	182	SER	-	expression tag	UNP P80188
C	183	HIS	-	expression tag	UNP P80188
C	184	PRO	-	expression tag	UNP P80188
C	185	GLN	-	expression tag	UNP P80188
C	186	PHE	-	expression tag	UNP P80188
C	187	GLU	-	expression tag	UNP P80188
C	188	LYS	-	expression tag	UNP P80188

- Molecule 2 is a protein called Fibronectin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	92	690	440	104	146	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	D	93	695	443	105	147	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1252	MET	-	initiating methionine	UNP P02751
B	1253	ALA	-	expression tag	UNP P02751
B	1254	SER	-	expression tag	UNP P02751
B	1255	ARG	-	expression tag	UNP P02751
B	1256	GLY	-	expression tag	UNP P02751
B	1257	SER	-	expression tag	UNP P02751
B	1258	HIS	-	expression tag	UNP P02751
B	1259	HIS	-	expression tag	UNP P02751
B	1260	HIS	-	expression tag	UNP P02751
B	1261	HIS	-	expression tag	UNP P02751
B	1262	HIS	-	expression tag	UNP P02751
B	1263	HIS	-	expression tag	UNP P02751
B	1264	GLY	-	expression tag	UNP P02751
B	1265	ALA	-	expression tag	UNP P02751
D	1252	MET	-	initiating methionine	UNP P02751
D	1253	ALA	-	expression tag	UNP P02751
D	1254	SER	-	expression tag	UNP P02751
D	1255	ARG	-	expression tag	UNP P02751
D	1256	GLY	-	expression tag	UNP P02751
D	1257	SER	-	expression tag	UNP P02751
D	1258	HIS	-	expression tag	UNP P02751
D	1259	HIS	-	expression tag	UNP P02751
D	1260	HIS	-	expression tag	UNP P02751
D	1261	HIS	-	expression tag	UNP P02751
D	1262	HIS	-	expression tag	UNP P02751
D	1263	HIS	-	expression tag	UNP P02751
D	1264	GLY	-	expression tag	UNP P02751
D	1265	ALA	-	expression tag	UNP P02751

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	102	102	102	0	0
3	B	40	40	40	0	0

Continued on next page...

Continued from previous page...

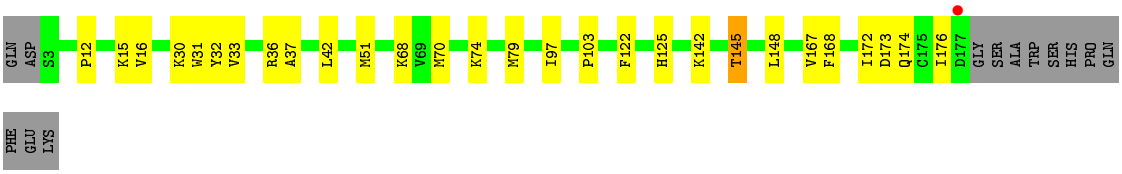
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	C	103	103	103	0	0
			Total	O		
3	D	51	51	51	0	0

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 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.

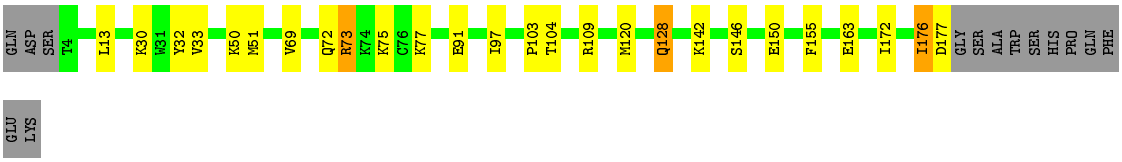
- Molecule 1: Neutrophil gelatinase-associated lipocalin

Chain A:



- Molecule 1: Neutrophil gelatinase-associated lipocalin

Chain C:

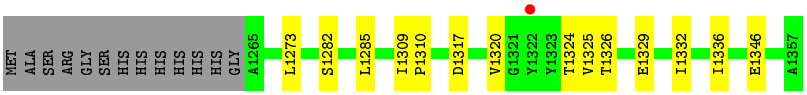


- Molecule 2: Fibronectin

Chain B:

- Molecule 2: Fibronectin

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	52.46Å 67.80Å 76.81Å 90.00° 90.06° 90.00°	Depositor
Resolution (Å)	28.47 – 1.60 28.47 – 1.60	Depositor EDS
% Data completeness (in resolution range)	99.4 (28.47-1.60) 99.4 (28.47-1.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.81 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.169 , 0.208 0.175 , 0.214	Depositor DCC
R_{free} test set	3579 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	17.2	Xtriage
Anisotropy	0.685	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 34.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.449 for h,-k,-l	Xtriage
Reported twinning fraction	0.510 for H, K, L 0.490 for h,-k,-l	Depositor
Outliers	0 of 70827 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4513	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.65% 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)

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	1.15	2/1457 (0.1%)	1.11	2/1971 (0.1%)
1	C	0.85	0/1451	0.92	0/1963
2	B	0.81	0/703	0.93	2/966 (0.2%)
2	D	0.80	1/708 (0.1%)	0.87	0/973
All	All	0.95	3/4319 (0.1%)	0.98	4/5873 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	TRP	CG-CD1	6.18	1.45	1.36
1	A	37	ALA	C-O	-5.66	1.12	1.23
2	D	1346	GLU	CD-OE2	5.28	1.31	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1299	ARG	NE-CZ-NH1	7.37	123.99	120.30
2	B	1299	ARG	NE-CZ-NH2	-6.24	117.18	120.30
1	A	36	ARG	NE-CZ-NH2	-5.60	117.50	120.30
1	A	145	THR	CA-CB-CG2	5.33	119.86	112.40

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	A	1419	0	1418	16	0
1	C	1413	0	1413	23	0
2	B	690	0	673	12	0
2	D	695	0	678	9	0
3	A	102	0	0	0	0
3	B	40	0	0	1	0
3	C	103	0	0	1	0
3	D	51	0	0	0	0
All	All	4513	0	4182	55	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 7.

The worst 5 of 55 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:C:72:GLN:OE1	1:C:77:LYS:HE3	1.77	0.83
1:C:72:GLN:OE1	1:C:77:LYS:CE	2.37	0.71
1:A:51:MET:SD	1:A:70:MET:SD	2.92	0.68
1:A:145:THR:HG22	1:A:148:LEU:H	1.61	0.66
2:D:1317:ASP:OD1	2:D:1320:VAL:HG23	1.96	0.65

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	173/188 (92%)	165 (95%)	8 (5%)	0	100 100
1	C	172/188 (92%)	162 (94%)	8 (5%)	2 (1%)	13 2
2	B	90/106 (85%)	88 (98%)	2 (2%)	0	100 100
2	D	91/106 (86%)	89 (98%)	2 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	526/588 (90%)	504 (96%)	20 (4%)	2 (0%)	34 15

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	128	GLN
1	C	73	ARG

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	160/171 (94%)	158 (99%)	2 (1%)	69 50
1	C	159/171 (93%)	154 (97%)	5 (3%)	40 15
2	B	79/89 (89%)	79 (100%)	0	100 100
2	D	79/89 (89%)	79 (100%)	0	100 100
All	All	477/520 (92%)	470 (98%)	7 (2%)	65 44

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	104	THR
1	C	176	ILE
1	C	128	GLN
1	A	173	ASP
1	C	163	GLU

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

Mol	Chain	Res	Type
2	D	1291	ASN

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 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 (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	175/188 (93%)	-0.37	1 (0%)	89	89	11, 17, 32, 46	0
1	C	174/188 (92%)	-0.29	0	100	100	13, 23, 38, 50	0
2	B	92/106 (86%)	-0.34	0	100	100	13, 24, 34, 37	0
2	D	93/106 (87%)	-0.25	1 (1%)	80	80	15, 24, 39, 42	0
All	All	534/588 (90%)	-0.32	2 (0%)	92	92	11, 22, 37, 50	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	1322	TYR	2.1
1	A	177	ASP	2.1

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 carbohydrates 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.