



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 19, 2017 – 04:30 PM EDT

PDB ID : 4DRW
Title : Crystal Structure of the Ternary Complex between S100A10, an Annexin A2 N-terminal Peptide and an AHNAK Peptide
Authors : Rezvanpour, A.; Lee, T.-W.; Junop, M.S.; Shaw, G.S.
Deposited on : unknown
Resolution : 3.50 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20029824

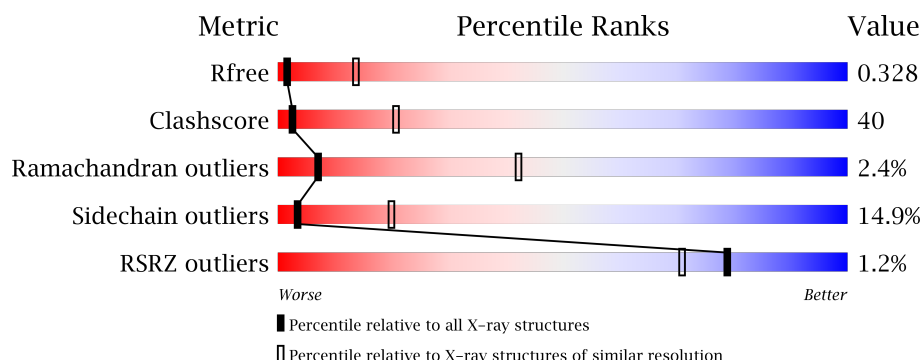
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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}	100719	1195 (3.60-3.40)
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)
RSRZ outliers	101464	1226 (3.60-3.40)

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 $\leq 5\%$. 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	121	<div> <div>2%</div> <div>28%</div> <div>50%</div> <div>6%</div> <div>15%</div> </div>
1	B	121	<div> <div>0%</div> <div>30%</div> <div>46%</div> <div>9%</div> <div>14%</div> </div>
1	C	121	<div> <div>21%</div> <div>47%</div> <div>17%</div> <div>15%</div> </div>
1	D	121	<div> <div>2%</div> <div>26%</div> <div>51%</div> <div>7%</div> <div>14%</div> </div>
2	E	20	<div> <div>15%</div> <div>30%</div> <div>5%</div> <div>50%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	20	 A horizontal bar chart representing the quality of chain F. The bar is divided into four segments: a small green segment at the beginning labeled '5%', followed by a yellow segment labeled '35%', then an orange segment labeled '10%', and finally a long grey segment at the end labeled '50%'.

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3490 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 Protein S100-A10/Annexin A2 chimeric protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	103	Total	C	N	O	S	0	0	0
			825	531	133	152	9			
1	B	104	Total	C	N	O	S	0	0	0
			834	536	134	155	9			
1	C	103	Total	C	N	O	S	0	0	0
			825	531	133	152	9			
1	D	104	Total	C	N	O	S	0	0	0
			834	536	134	155	9			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	EXPRESSION TAG	UNP P60903
A	-4	PRO	-	EXPRESSION TAG	UNP P60903
A	-3	LEU	-	EXPRESSION TAG	UNP P60903
A	-2	GLY	-	EXPRESSION TAG	UNP P60903
A	-1	SER	-	EXPRESSION TAG	UNP P60903
A	93	GLU	-	LINKER	UNP P07355
A	94	ASN	-	LINKER	UNP P07355
A	95	LEU	-	LINKER	UNP P07355
A	96	TYR	-	LINKER	UNP P07355
A	97	PHE	-	LINKER	UNP P07355
A	98	GLN	-	LINKER	UNP P07355
A	99	GLY	-	LINKER	UNP P07355
A	100	ASP	-	LINKER	UNP P07355
A	108	SER	CYS	ENGINEERED MUTATION	UNP P07355
B	-5	GLY	-	EXPRESSION TAG	UNP P60903
B	-4	PRO	-	EXPRESSION TAG	UNP P60903
B	-3	LEU	-	EXPRESSION TAG	UNP P60903
B	-2	GLY	-	EXPRESSION TAG	UNP P60903
B	-1	SER	-	EXPRESSION TAG	UNP P60903
B	93	GLU	-	LINKER	UNP P07355
B	94	ASN	-	LINKER	UNP P07355

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	95	LEU	-	LINKER	UNP P07355
B	96	TYR	-	LINKER	UNP P07355
B	97	PHE	-	LINKER	UNP P07355
B	98	GLN	-	LINKER	UNP P07355
B	99	GLY	-	LINKER	UNP P07355
B	100	ASP	-	LINKER	UNP P07355
B	108	SER	CYS	ENGINEERED MUTATION	UNP P07355
C	-5	GLY	-	EXPRESSION TAG	UNP P60903
C	-4	PRO	-	EXPRESSION TAG	UNP P60903
C	-3	LEU	-	EXPRESSION TAG	UNP P60903
C	-2	GLY	-	EXPRESSION TAG	UNP P60903
C	-1	SER	-	EXPRESSION TAG	UNP P60903
C	93	GLU	-	LINKER	UNP P07355
C	94	ASN	-	LINKER	UNP P07355
C	95	LEU	-	LINKER	UNP P07355
C	96	TYR	-	LINKER	UNP P07355
C	97	PHE	-	LINKER	UNP P07355
C	98	GLN	-	LINKER	UNP P07355
C	99	GLY	-	LINKER	UNP P07355
C	100	ASP	-	LINKER	UNP P07355
C	108	SER	CYS	ENGINEERED MUTATION	UNP P07355
D	-5	GLY	-	EXPRESSION TAG	UNP P60903
D	-4	PRO	-	EXPRESSION TAG	UNP P60903
D	-3	LEU	-	EXPRESSION TAG	UNP P60903
D	-2	GLY	-	EXPRESSION TAG	UNP P60903
D	-1	SER	-	EXPRESSION TAG	UNP P60903
D	93	GLU	-	LINKER	UNP P07355
D	94	ASN	-	LINKER	UNP P07355
D	95	LEU	-	LINKER	UNP P07355
D	96	TYR	-	LINKER	UNP P07355
D	97	PHE	-	LINKER	UNP P07355
D	98	GLN	-	LINKER	UNP P07355
D	99	GLY	-	LINKER	UNP P07355
D	100	ASP	-	LINKER	UNP P07355
D	108	SER	CYS	ENGINEERED MUTATION	UNP P07355

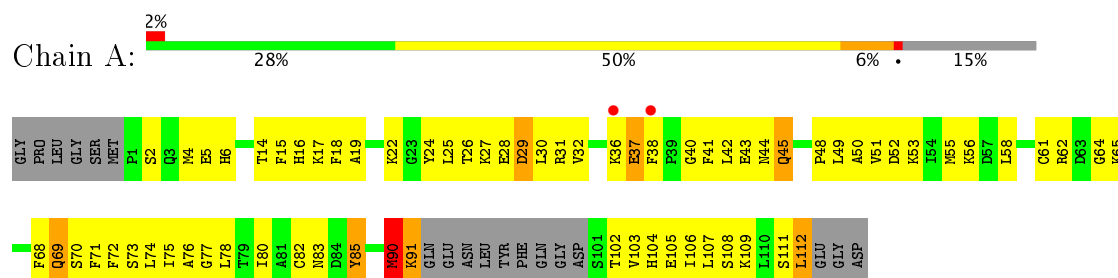
- Molecule 2 is a protein called Neuroblast differentiation-associated protein AHNAK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	10	Total 86	C 61	N 13	O 11	S 1	0	0	0
2	F	10	Total 86	C 61	N 13	O 11	S 1	0	0	0

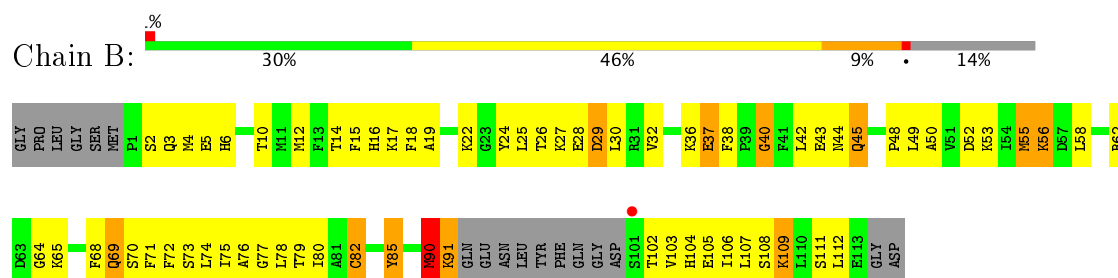
3 Residue-property plots [i](#)

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.

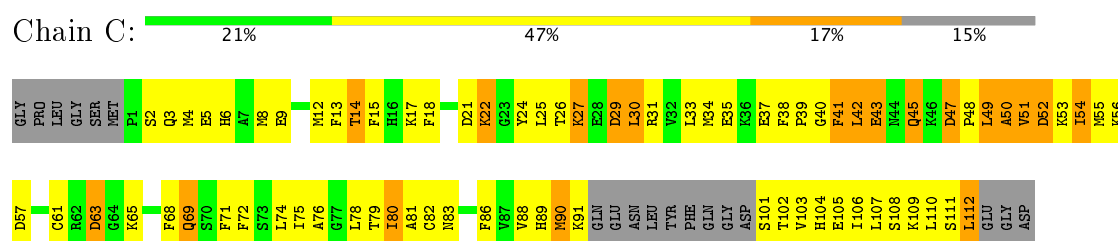
- Molecule 1: Protein S100-A10/Annexin A2 chimeric protein



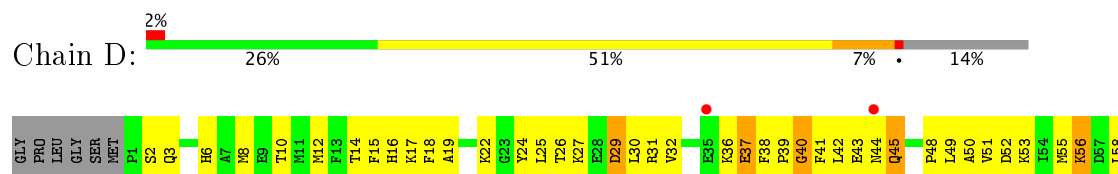
- Molecule 1: Protein S100-A10/Annexin A2 chimeric protein

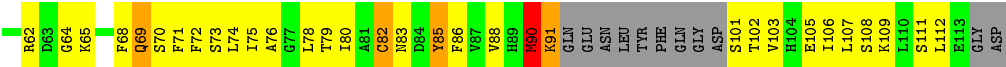


- Molecule 1: Protein S100-A10/Annexin A2 chimeric protein



- Molecule 1: Protein S100-A10/Annexin A2 chimeric protein

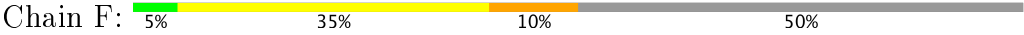




- Molecule 2: Neuroblast differentiation-associated protein AHNAK



- Molecule 2: Neuroblast differentiation-associated protein AHNAK



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	47.76Å 47.79Å 62.82Å 71.62° 71.63° 68.62°	Depositor
Resolution (Å)	43.34 – 3.50 43.34 – 3.27	Depositor EDS
% Data completeness (in resolution range)	(Not available) (43.34-3.50) 87.3 (43.34-3.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.37 (at 3.25Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.296 , 0.327 0.296 , 0.328	Depositor DCC
R_{free} test set	633 reflections (12.17%)	DCC
Wilson B-factor (Å ²)	121.2	Xtriage
Anisotropy	0.426	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 156.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.458 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	3490	wwPDB-VP
Average B, all atoms (Å ²)	154.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.27	0/841	0.57	1/1123 (0.1%)
1	B	0.27	0/850	0.58	1/1135 (0.1%)
1	C	0.27	0/841	0.66	3/1123 (0.3%)
1	D	0.32	0/850	0.57	0/1135
2	E	0.38	0/89	0.68	0/116
2	F	0.36	0/89	0.89	0/116
All	All	0.29	0/3560	0.61	5/4748 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	50	ALA	CB-CA-C	-5.76	101.46	110.10
1	A	41	PHE	N-CA-CB	-5.37	100.93	110.60
1	B	109	LYS	CB-CA-C	-5.27	99.86	110.40
1	C	63	ASP	N-CA-CB	-5.07	101.48	110.60
1	C	45	GLN	CB-CA-C	5.06	120.51	110.40

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	825	0	823	69	0
1	B	834	0	829	65	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	825	0	823	97	0
1	D	834	0	829	74	0
2	E	86	0	98	8	0
2	F	86	0	98	9	0
All	All	3490	0	3500	280	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 40.

The worst 5 of 280 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:D:90:MET:O	1:D:91:LYS:HG3	1.20	1.28
1:D:40:GLY:O	1:D:108:SER:HB2	1.35	1.26
1:B:40:GLY:O	1:B:108:SER:HB2	1.26	1.25
1:A:40:GLY:O	1:A:108:SER:HB2	1.23	1.25
1:B:90:MET:O	1:B:91:LYS:HG3	1.36	1.21

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	99/121 (82%)	78 (79%)	20 (20%)	1 (1%)	18	61
1	B	100/121 (83%)	78 (78%)	20 (20%)	2 (2%)	9	46
1	C	99/121 (82%)	75 (76%)	21 (21%)	3 (3%)	5	37
1	D	100/121 (83%)	79 (79%)	18 (18%)	3 (3%)	5	37
2	E	8/20 (40%)	4 (50%)	4 (50%)	0	100	100
2	F	8/20 (40%)	3 (38%)	4 (50%)	1 (12%)	0	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	414/524 (79%)	317 (77%)	87 (21%)	10 (2%)	7	42

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	41	PHE
1	C	90	MET
2	F	11	PRO
1	A	90	MET
1	B	90	MET

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	92/106 (87%)	81 (88%)	11 (12%)	6	28
1	B	93/106 (88%)	81 (87%)	12 (13%)	5	25
1	C	92/106 (87%)	74 (80%)	18 (20%)	1	8
1	D	93/106 (88%)	82 (88%)	11 (12%)	6	29
2	E	10/18 (56%)	7 (70%)	3 (30%)	0	3
2	F	10/18 (56%)	7 (70%)	3 (30%)	0	3
All	All	390/460 (85%)	332 (85%)	58 (15%)	3	20

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	27	LYS
1	C	49	LEU
2	E	10	ILE
1	C	29	ASP
1	C	42	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	69	GLN
1	D	69	GLN
1	C	16	HIS
1	B	16	HIS
1	C	3	GLN

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 ⓘ

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	103/121 (85%)	-0.18	2 (1%) 67 59	84, 161, 210, 253	0
1	B	104/121 (85%)	-0.18	1 (0%) 82 75	82, 161, 209, 253	0
1	C	103/121 (85%)	-0.38	0 100 100	63, 138, 203, 232	0
1	D	104/121 (85%)	-0.28	2 (1%) 67 59	82, 160, 210, 253	0
2	E	10/20 (50%)	-0.66	0 100 100	54, 101, 174, 185	0
2	F	10/20 (50%)	-0.05	0 100 100	92, 121, 160, 162	0
All	All	434/524 (82%)	-0.26	5 (1%) 79 71	54, 152, 210, 253	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	35	GLU	2.7
1	B	101	SER	2.7
1	A	38	PHE	2.3
1	D	44	ASN	2.3
1	A	36	LYS	2.2

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.