



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 04:37 AM GMT

PDB ID : 3KBT
Title : Crystal structure of the ankyrin binding domain of human erythroid beta spectrin (repeats 13-15) in complex with the spectrin binding domain of human erythroid ankyrin (ZU5-ANK)
Authors : Ipsaro, J.J.; Mondragon, A.
Deposited on : 2009-10-20
Resolution : 2.75 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

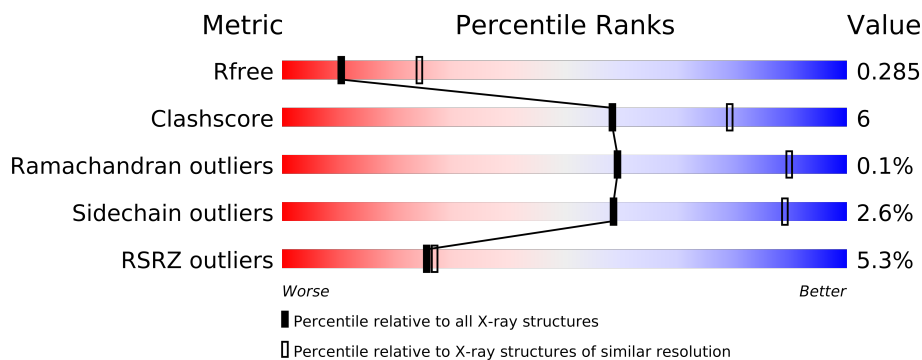
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.75 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}	66092	2406 (2.80-2.72)
Clashscore	79885	2995 (2.80-2.72)
Ramachandran outliers	78287	2941 (2.80-2.72)
Sidechain outliers	78261	2944 (2.80-2.72)
RSRZ outliers	66119	2409 (2.80-2.72)

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	326	
1	B	326	
2	C	161	
2	D	161	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7076 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 Spectrin beta chain, erythrocyte.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	284	Total	C	N	O	S	0	0	0
			2308	1444	407	452	5			
1	B	294	Total	C	N	O	S	0	0	0
			2377	1486	417	468	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1581	SER	-	EXPRESSION TAG	UNP P11277
A	1582	ASN	-	EXPRESSION TAG	UNP P11277
A	1844	ASP	GLU	SEE REMARK 999	UNP P11277
A	1845	VAL	LEU	SEE REMARK 999	UNP P11277
B	1581	SER	-	EXPRESSION TAG	UNP P11277
B	1582	ASN	-	EXPRESSION TAG	UNP P11277
B	1844	ASP	GLU	SEE REMARK 999	UNP P11277
B	1845	VAL	LEU	SEE REMARK 999	UNP P11277

- Molecule 2 is a protein called Ankyrin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	157	Total	C	N	O	S	0	0	0
			1216	765	222	222	7			
2	D	151	Total	C	N	O	S	0	0	0
			1175	743	211	214	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	908	SER	-	EXPRESSION TAG	UNP P16157
C	909	ASN	-	EXPRESSION TAG	UNP P16157
C	910	ALA	-	EXPRESSION TAG	UNP P16157
D	908	SER	-	EXPRESSION TAG	UNP P16157

Continued on next page...

Continued from previous page...

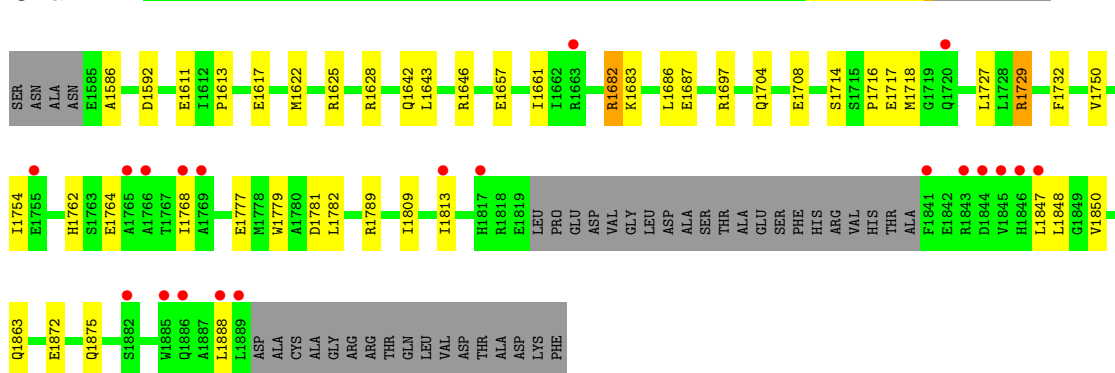
Chain	Residue	Modelled	Actual	Comment	Reference
D	909	ASN	-	EXPRESSION TAG	UNP P16157
D	910	ALA	-	EXPRESSION TAG	UNP P16157

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.

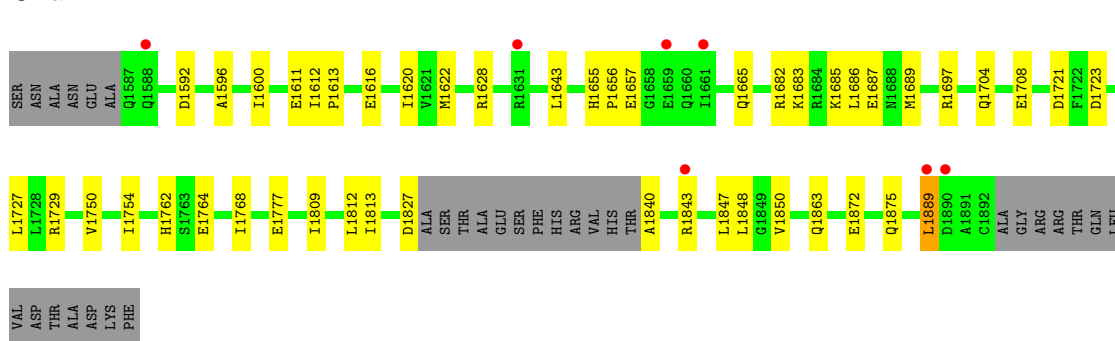
- Molecule 1: Spectrin beta chain, erythrocyte

Chain A:



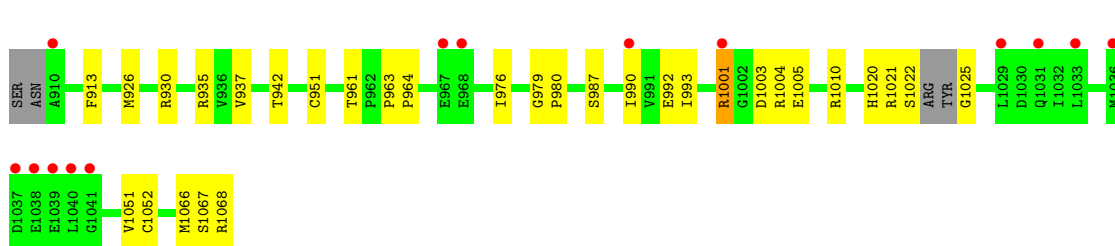
- Molecule 1: Spectrin beta chain, erythrocyte

Chain B:

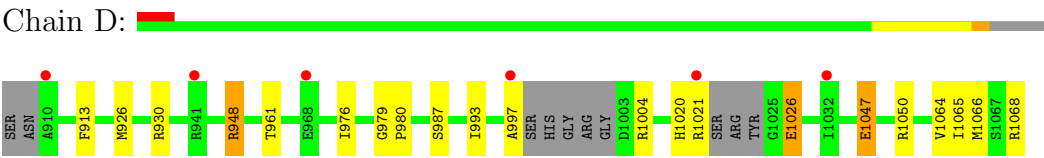


- Molecule 2: Ankyrin-1

Chain C:



- Molecule 2: Ankyrin-1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	90.45Å 95.66Å 137.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.82 – 2.75 37.82 – 2.75	Depositor EDS
% Data completeness (in resolution range)	95.7 (37.82-2.75) 95.7 (37.82-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.17 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.245 , 0.292 0.239 , 0.285	Depositor DCC
R_{free} test set	1576 reflections (5.47%)	DCC
Wilson B-factor (Å ²)	55.9	Xtriage
Anisotropy	0.374	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 24.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 30387 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7076	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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	0.73	1/2346 (0.0%)	0.81	9/3161 (0.3%)
1	B	0.72	0/2416	0.80	6/3258 (0.2%)
2	C	0.67	2/1239 (0.2%)	0.78	0/1673
2	D	0.66	0/1196	0.77	0/1615
All	All	0.71	3/7197 (0.0%)	0.79	15/9707 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1052	CYS	CB-SG	-6.63	1.71	1.82
1	A	1617	GLU	CG-CD	5.28	1.59	1.51
2	C	951	CYS	CB-SG	-5.19	1.73	1.81

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1628	ARG	NE-CZ-NH1	10.51	125.56	120.30
1	A	1628	ARG	NE-CZ-NH1	-9.99	115.31	120.30
1	A	1697	ARG	NE-CZ-NH1	-7.91	116.35	120.30
1	A	1628	ARG	NE-CZ-NH2	7.66	124.13	120.30
1	A	1697	ARG	NE-CZ-NH2	7.56	124.08	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 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	2308	0	2230	30	0
1	B	2377	0	2293	29	1
2	C	1216	0	1238	20	0
2	D	1175	0	1201	12	1
All	All	7076	0	6962	87	1

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 6.

The worst 5 of 87 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:C:1001:ARG:HH11	2:C:1001:ARG:HG3	1.13	1.05
1:A:1754:ILE:HG23	1:A:1768:ILE:CG2	2.00	0.92
1:B:1754:ILE:HG23	1:B:1768:ILE:CG2	2.05	0.85
1:B:1611:GLU:O	1:B:1622:MET:HE1	1.77	0.84
2:C:1001:ARG:HH11	2:C:1001:ARG:CG	1.93	0.80

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:1723:ASP:OD2	2:D:1047:GLU:OE2[4.555]	2.19	0.01

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	280/326 (86%)	277 (99%)	3 (1%)	0	100	100
1	B	290/326 (89%)	287 (99%)	3 (1%)	0	100	100
2	C	153/161 (95%)	137 (90%)	16 (10%)	0	100	100
2	D	145/161 (90%)	134 (92%)	10 (7%)	1 (1%)	30	67
All	All	868/974 (89%)	835 (96%)	32 (4%)	1 (0%)	59	90

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	1026	GLU

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	236/269 (88%)	231 (98%)	5 (2%)	66	92
1	B	244/269 (91%)	238 (98%)	6 (2%)	60	89
2	C	134/138 (97%)	132 (98%)	2 (2%)	76	95
2	D	130/138 (94%)	124 (95%)	6 (5%)	37	71
All	All	744/814 (91%)	725 (97%)	19 (3%)	59	89

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1827	ASP
1	B	1889	LEU
2	D	987	SER
1	B	1812	LEU
2	D	1004	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1856	GLN
1	A	1875	GLN
1	B	1875	GLN
1	A	1694	GLN
1	B	1688	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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, 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	284/326 (87%)	0.32	20 (7%) 16 16	18, 33, 46, 51	0
1	B	294/326 (90%)	0.25	7 (2%) 56 59	18, 33, 47, 51	0
2	C	157/161 (97%)	0.49	14 (8%) 10 9	24, 32, 45, 54	0
2	D	151/161 (93%)	0.31	6 (3%) 36 39	24, 31, 45, 54	0
All	All	886/974 (90%)	0.33	47 (5%) 25 27	18, 32, 46, 54	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1843	ARG	7.3
2	C	967	GLU	6.0
2	C	1029	LEU	5.8
2	C	1031	GLN	5.6
2	D	910	ALA	4.7

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.