wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 9, 2023 – 03:10 PM EDT

PDB ID	:	4HRE
Title	:	Crystal Structure of p11/Annexin A2 Heterotetramer in Complex with SMARCA3 Peptide
Authors	:	Gao, P.; Patel, D.J.
Deposited on	:	2012-10-27
Resolution	:	2.79 Å(reported)

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

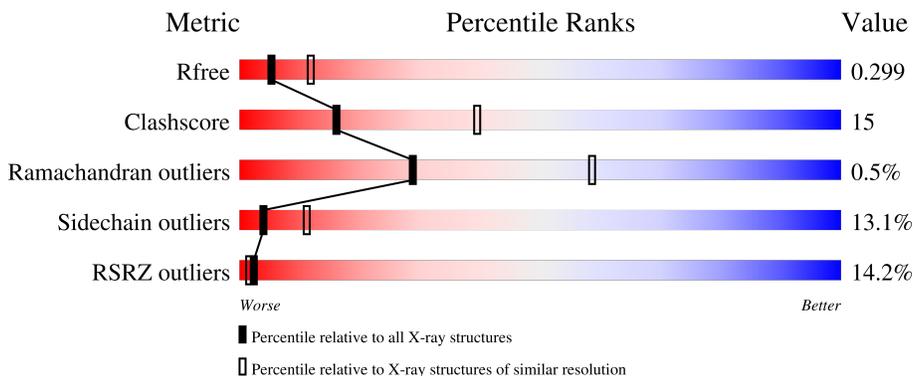|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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		0.299
Clashscore		15
Ramachandran outliers		0.5%
Sidechain outliers		13.1%
RSRZ outliers		14.2%

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}	130704	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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	339	6% / 69% / 27% / •
1	B	339	5% / 70% / 25% / •
1	C	339	6% / 67% / 27% / 6%
1	D	339	5% / 64% / 32% / •
2	E	97	40% / 42% / 45% / 6% / 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	97	43% 49% 34% 10% 6%
2	I	97	33% 48% 37% 8% 6%
2	J	97	41% 47% 35% 10% 6%
3	G	14	36% 14% 36% 14% 36%
3	H	14	43% 21% 29% 14% 36%
3	K	14	50% 50% 14% 36%
3	L	14	36% 29% 21% 14% 36%

2   Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14088 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 Annexin A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total	C	N	O	S	0	0	0
			2705	1700	463	529	13			
1	C	338	Total	C	N	O	S	0	0	0
			2705	1700	463	529	13			
1	B	338	Total	C	N	O	S	0	0	0
			2705	1700	463	529	13			
1	D	338	Total	C	N	O	S	0	0	0
			2701	1699	463	526	13			

- Molecule 2 is a protein called Protein S100-A10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	91	Total	C	N	O	S	0	0	0
			733	472	118	134	9			
2	F	91	Total	C	N	O	S	0	0	0
			733	472	118	134	9			
2	I	91	Total	C	N	O	S	0	0	0
			733	472	118	134	9			
2	J	91	Total	C	N	O	S	0	0	0
			733	472	118	134	9			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	SER	-	expression tag	UNP P60903
F	0	SER	-	expression tag	UNP P60903
I	0	SER	-	expression tag	UNP P60903
J	0	SER	-	expression tag	UNP P60903

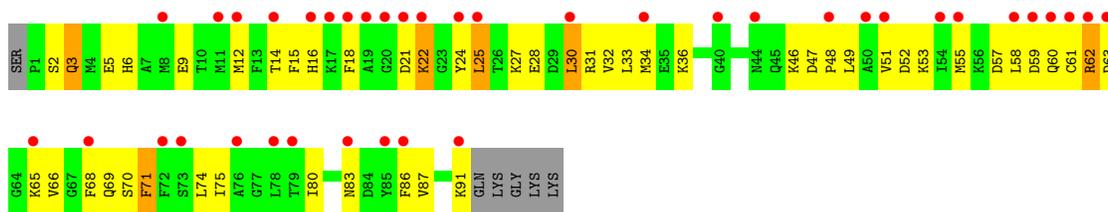- Molecule 3 is a protein called Helicase-like transcription factor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	9	Total	C	N	O	0	0	0
			85	61	12	12			
3	H	9	Total	C	N	O	0	0	0
			85	61	12	12			
3	K	9	Total	C	N	O	0	0	0
			85	61	12	12			
3	L	9	Total	C	N	O	0	0	0
			85	61	12	12			

3 Residue-property plots ⓘ

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: Annexin A2

Chain A:

• Molecule 1: Annexin A2

Chain C:

• Molecule 1: Annexin A2

Chain B:

- Molecule 1: Annexin A2

Chain D:

- Molecule 2: Protein S100-A10

Chain E:

- Molecule 2: Protein S100-A10

Chain F:

- Molecule 2: Protein S100-A10

Chain I:

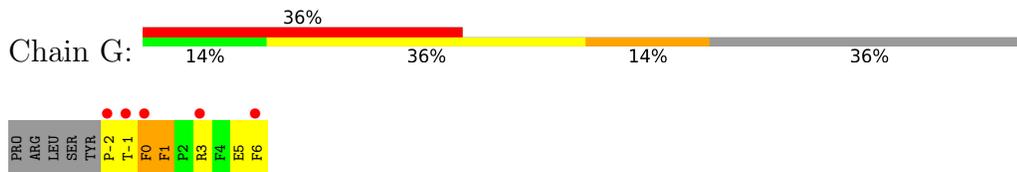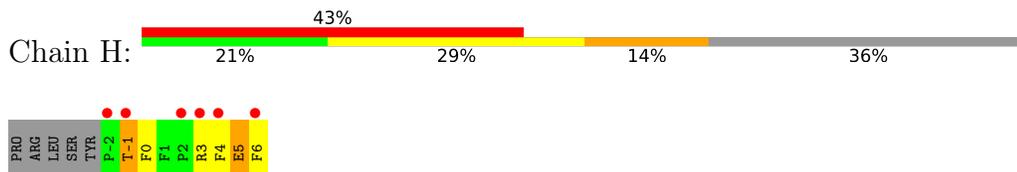
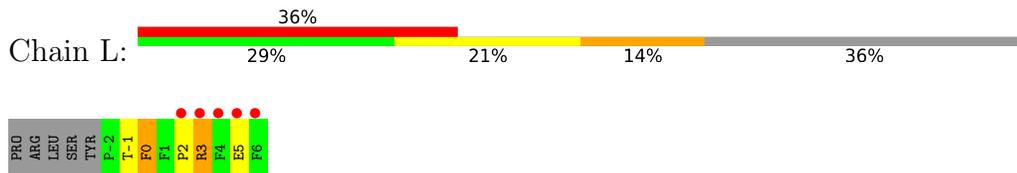• Molecule 2: Protein S100-A10

Chain J:

• Molecule 3: Helicase-like transcription factor

Chain G:

• Molecule 3: Helicase-like transcription factor

Chain H:

• Molecule 3: Helicase-like transcription factor

Chain K:

• Molecule 3: Helicase-like transcription factor

Chain L:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	50.79Å 74.35Å 154.47Å 90.14° 96.56° 108.37°	Depositor
Resolution (Å)	47.34 – 2.79 47.34 – 2.79	Depositor EDS
% Data completeness (in resolution range)	95.2 (47.34-2.79) 95.4 (47.34-2.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.31 (at 2.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.249 , 0.296 0.259 , 0.299	Depositor DCC
R_{free} test set	2583 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	44.1	Xtriage
Anisotropy	0.426	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 41.9	EDS
L-test for twinning[2]	$< |L| > = 0.52, < L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.000 for -h,h+k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	14088	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.79 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.4504e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

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	0.38	0/2744	0.59	0/3687
1	B	0.39	0/2744	0.61	0/3687
1	C	0.38	0/2744	0.61	0/3687
1	D	0.39	0/2740	0.62	1/3682 (0.0%)
2	E	0.32	0/749	0.57	0/1000
2	F	0.33	0/749	0.57	0/1000
2	I	0.31	0/749	0.57	0/1000
2	J	0.41	1/749 (0.1%)	0.70	2/1000 (0.2%)
3	G	0.44	0/90	0.51	0/120
3	H	0.35	0/90	0.62	0/120
3	K	0.43	0/90	0.51	0/120
3	L	0.36	0/90	0.51	0/120
All	All	0.38	1/14328 (0.0%)	0.61	3/19223 (0.0%)

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
1	C	0	1
1	D	0	3
2	J	0	1
3	H	0	1
All	All	0	7

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	48	PRO	N-CD	6.88	1.57	1.47

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	12	SER	CB-CA-C	9.10	127.39	110.10
2	J	47	ASP	C-N-CD	-7.49	104.13	120.60
2	J	47	ASP	CB-CA-C	5.91	122.23	110.40

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	181	GLU	Peptide
1	C	15	GLY	Peptide
1	D	120	GLY	Peptide
1	D	121	LEU	Peptide
3	H	-1	THR	Peptide

5.2 Too-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 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	2705	0	2737	81	0
1	B	2705	0	2737	69	0
1	C	2705	0	2737	73	0
1	D	2701	0	2732	69	0
2	E	733	0	723	39	0
2	F	733	0	723	32	0
2	I	733	0	723	33	0
2	J	733	0	723	36	0
3	G	85	0	78	9	0
3	H	85	0	78	5	0
3	K	85	0	78	0	0
3	L	85	0	78	2	0
All	All	14088	0	14147	410	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 15.

The worst 5 of 410 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:197:GLU:OE1	1:B:212:LYS:CE	1.82	1.27
1:A:118:MET:O	1:A:123:THR:HG21	1.42	1.17
1:A:120:GLY:CA	1:A:123:THR:HG22	1.76	1.16
1:A:120:GLY:HA3	1:A:123:THR:HG22	1.28	1.11
2:I:4:MET:HB2	2:J:37:GLU:HG3	1.14	1.11

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	336/339 (99%)	319 (95%)	16 (5%)	1 (0%)	41	70
1	B	336/339 (99%)	328 (98%)	8 (2%)	0	100	100
1	C	336/339 (99%)	316 (94%)	16 (5%)	4 (1%)	13	36
1	D	336/339 (99%)	318 (95%)	16 (5%)	2 (1%)	25	54
2	E	89/97 (92%)	81 (91%)	8 (9%)	0	100	100
2	F	89/97 (92%)	82 (92%)	7 (8%)	0	100	100
2	I	89/97 (92%)	78 (88%)	10 (11%)	1 (1%)	14	38
2	J	89/97 (92%)	74 (83%)	15 (17%)	0	100	100
3	G	7/14 (50%)	7 (100%)	0	0	100	100
3	H	7/14 (50%)	7 (100%)	0	0	100	100
3	K	7/14 (50%)	7 (100%)	0	0	100	100
3	L	7/14 (50%)	7 (100%)	0	0	100	100
All	All	1728/1800 (96%)	1624 (94%)	96 (6%)	8 (0%)	29	58

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	23	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	21	PRO
1	C	130	GLU
1	D	31	THR
1	D	51	VAL

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	300/301 (100%)	269 (90%)	31 (10%)	7	19
1	B	300/301 (100%)	265 (88%)	35 (12%)	5	15
1	C	300/301 (100%)	268 (89%)	32 (11%)	6	18
1	D	298/301 (99%)	269 (90%)	29 (10%)	8	22
2	E	80/85 (94%)	70 (88%)	10 (12%)	4	13
2	F	80/85 (94%)	61 (76%)	19 (24%)	0	2
2	I	80/85 (94%)	62 (78%)	18 (22%)	1	2
2	J	80/85 (94%)	63 (79%)	17 (21%)	1	3
3	G	9/14 (64%)	5 (56%)	4 (44%)	0	0
3	H	9/14 (64%)	6 (67%)	3 (33%)	0	0
3	K	9/14 (64%)	7 (78%)	2 (22%)	1	2
3	L	9/14 (64%)	5 (56%)	4 (44%)	0	0
All	All	1554/1600 (97%)	1350 (87%)	204 (13%)	4	11

5 of 204 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	162	ASP
1	D	112	SER
3	K	6	PHE
1	B	179	ARG
1	B	231	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	5	HIS
2	J	45	GLN
2	J	69	GLN
2	J	3	GLN
2	E	45	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides 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(Å²)	Q<0.9
1	A	338/339 (99%)	0.40	20 (5%)	22	17	27, 54, 112, 197	0
1	B	338/339 (99%)	0.30	16 (4%)	31	25	26, 55, 124, 168	0
1	C	338/339 (99%)	0.32	20 (5%)	22	17	25, 51, 128, 188	0
1	D	338/339 (99%)	0.29	16 (4%)	31	25	23, 50, 111, 210	0
2	E	91/97 (93%)	1.93	39 (42%)	0	0	88, 146, 209, 277	0
2	F	91/97 (93%)	2.06	42 (46%)	0	0	77, 139, 204, 267	0
2	I	91/97 (93%)	2.04	32 (35%)	0	0	69, 142, 204, 291	0
2	J	91/97 (93%)	2.09	40 (43%)	0	0	85, 139, 182, 217	0
3	G	9/14 (64%)	2.60	5 (55%)	0	0	105, 118, 159, 193	0
3	H	9/14 (64%)	4.26	6 (66%)	0	0	136, 156, 169, 211	0
3	K	9/14 (64%)	2.59	7 (77%)	0	0	90, 122, 150, 166	0
3	L	9/14 (64%)	2.80	5 (55%)	0	0	106, 126, 152, 158	0
All	All	1752/1800 (97%)	0.74	248 (14%)	2	1	23, 60, 168, 291	0

The worst 5 of 248 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	18	SER	19.2
2	F	13	PHE	14.2
2	I	39	PRO	10.7
3	H	-2	PRO	10.1
1	A	16	ASP	9.7

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 monosaccharides 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.