



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 06:41 AM EDT

PDB ID : 6TNO  
Title : Crystal structure of the human Arc N-lobe bound to stargazin  
Authors : Hallin, E.I.; Bramham, C.R.; Kursula, P.  
Deposited on : 2019-12-09  
Resolution : 1.90 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

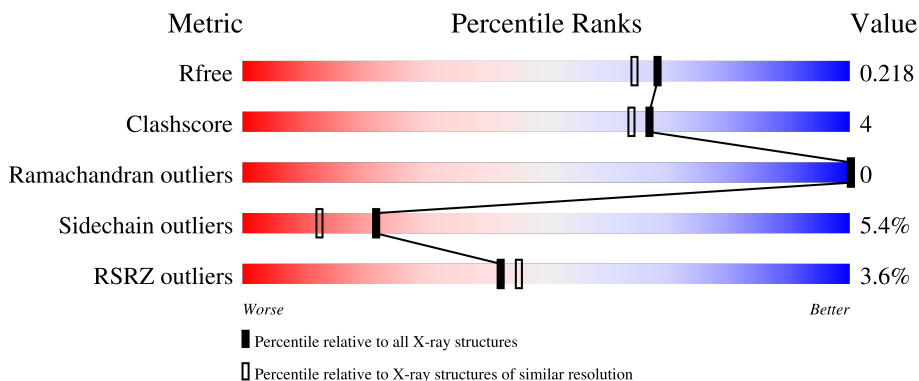
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	75	
1	C	75	
1	E	75	
2	B	8	
2	D	8	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	8	 88% 12%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3928 atoms, of which 1845 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 Arc\_C domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	67	1119	376	538	96	108	1	0	0	0
1	C	67	1158	391	558	98	110	1	0	2	0
1	E	67	1119	376	538	96	108	1	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	GLY	-	expression tag	UNP A0A2I2UQ80
A	204	ALA	-	expression tag	UNP A0A2I2UQ80
A	205	MET	-	expression tag	UNP A0A2I2UQ80
C	203	GLY	-	expression tag	UNP A0A2I2UQ80
C	204	ALA	-	expression tag	UNP A0A2I2UQ80
C	205	MET	-	expression tag	UNP A0A2I2UQ80
E	203	GLY	-	expression tag	UNP A0A2I2UQ80
E	204	ALA	-	expression tag	UNP A0A2I2UQ80
E	205	MET	-	expression tag	UNP A0A2I2UQ80

- Molecule 2 is a protein called Chains: B,D,F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	B	7	134	44	66	13	11	0	0	0
2	D	7	134	44	66	13	11	0	0	0
2	F	8	157	50	79	17	11	0	0	0

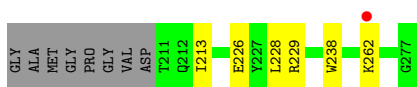
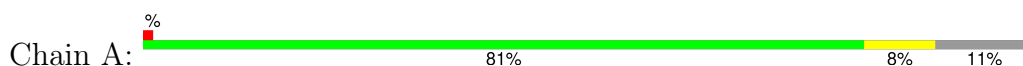
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	19	Total O 19 19	0	0
3	B	5	Total O 5 5	0	0
3	C	55	Total O 55 55	0	0
3	D	6	Total O 6 6	0	0
3	E	21	Total O 21 21	0	0
3	F	1	Total O 1 1	0	0

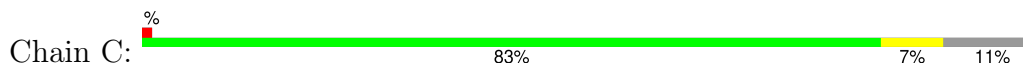
### 3 Residue-property plots [i](#)

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.

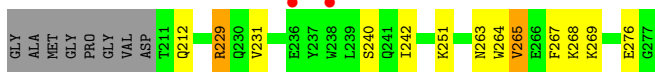
- Molecule 1: Arc\_C domain-containing protein



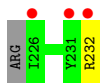
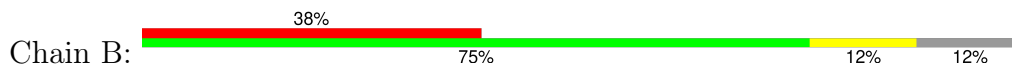
- Molecule 1: Arc\_C domain-containing protein



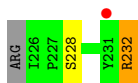
- Molecule 1: Arc\_C domain-containing protein



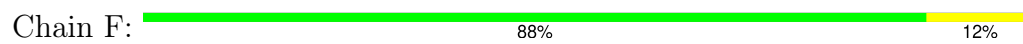
- Molecule 2: Chains: B,D,F



- Molecule 2: Chains: B,D,F



- Molecule 2: Chains: B,D,F



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.39Å 29.77Å 71.79Å 90.00° 92.51° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 42.64 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.8 (50.00-1.90) 97.9 (42.64-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.16 (at 1.91Å)	Xtrriage
Refinement program	PHENIX 1.16-3549	Depositor
R, $R_{free}$	0.182 , 0.216 0.185 , 0.218	Depositor DCC
$R_{free}$ test set	1726 reflections (10.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.1	Xtrriage
Anisotropy	0.528	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 59.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.031 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3928	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.55% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.36	0/600	0.52	0/806
1	C	0.34	0/620	0.52	0/833
1	E	0.34	0/600	0.55	0/806
2	B	0.29	0/70	0.50	0/92
2	D	0.34	0/70	0.52	0/92
2	F	0.26	0/80	0.53	0/106
All	All	0.34	0/2040	0.53	0/2735

There are no bond length outliers.

There are no bond angle outliers.

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	581	538	538	5	0
1	C	600	558	556	3	0
1	E	581	538	538	9	0
2	B	68	66	66	0	0
2	D	68	66	66	2	0
2	F	78	79	79	0	0
3	A	19	0	0	0	0
3	B	5	0	0	0	0
3	C	55	0	0	1	0
3	D	6	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	21	0	0	3	0
3	F	1	0	0	0	0
All	All	2083	1845	1843	15	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 4.

All (15) 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:A:262:LYS:HD2	1:E:229:ARG:HD2	1.61	0.83
1:A:262:LYS:CD	1:E:229:ARG:HD2	2.20	0.71
1:E:276:GLU:O	3:E:301:HOH:O	2.11	0.68
2:D:232:ARG:NH2	3:D:301:HOH:O	2.30	0.62
1:A:226:GLU:OE1	1:A:229:ARG:NH2	2.37	0.57
1:A:262:LYS:HD2	1:E:229:ARG:CD	2.34	0.56
1:E:263:ASN:OD1	1:E:265:VAL:HG22	2.06	0.56
1:C:226:GLU:OE2	3:C:301:HOH:O	2.18	0.54
1:E:242:ILE:HD13	1:E:267:PHE:CE1	2.48	0.48
1:E:265:VAL:HG13	3:E:309:HOH:O	2.13	0.48
1:E:251:LYS:NZ	3:E:303:HOH:O	2.45	0.46
1:C:213:ILE:HD13	2:D:228:SER:HB3	2.01	0.43
1:A:228:LEU:HB3	1:A:238:TRP:CE2	2.54	0.42
1:E:264:TRP:CD1	1:E:268:LYS:HE2	2.56	0.41
1:C:228:LEU:HB3	1:C:238:TRP:CE2	2.56	0.41

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	65/75 (87%)	65 (100%)	0	0	100	100
1	C	67/75 (89%)	67 (100%)	0	0	100	100
1	E	65/75 (87%)	65 (100%)	0	0	100	100
2	B	5/8 (62%)	5 (100%)	0	0	100	100
2	D	5/8 (62%)	5 (100%)	0	0	100	100
2	F	6/8 (75%)	6 (100%)	0	0	100	100
All	All	213/249 (86%)	213 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

### 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	61/65 (94%)	60 (98%)	1 (2%)	62	60
1	C	63/65 (97%)	62 (98%)	1 (2%)	62	60
1	E	61/65 (94%)	55 (90%)	6 (10%)	8	3
2	B	7/8 (88%)	6 (86%)	1 (14%)	3	1
2	D	7/8 (88%)	6 (86%)	1 (14%)	3	1
2	F	8/8 (100%)	7 (88%)	1 (12%)	4	1
All	All	207/219 (94%)	196 (95%)	11 (5%)	22	13

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	213	ILE
2	B	232	ARG
1	C	263	ASN
2	D	232	ARG
1	E	212	GLN
1	E	229	ARG
1	E	231	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	240	SER
1	E	265	VAL
1	E	269	LYS
2	F	225	ARG

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

Mol	Chain	Res	Type
1	E	223	HIS

### 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 monosaccharides 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	67/75 (89%)	0.02	1 (1%) 73 76	27, 42, 70, 84	0
1	C	67/75 (89%)	-0.21	1 (1%) 73 76	18, 26, 53, 59	0
1	E	67/75 (89%)	0.33	2 (2%) 50 53	30, 55, 82, 97	0
2	B	7/8 (87%)	1.48	3 (42%) 0 0	59, 65, 88, 109	0
2	D	7/8 (87%)	0.79	1 (14%) 2 2	27, 32, 66, 82	0
2	F	8/8 (100%)	0.50	0 100 100	39, 45, 68, 91	0
All	All	223/249 (89%)	0.13	8 (3%) 42 45	18, 42, 80, 109	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	226	ILE	5.1
2	D	231	TYR	3.2
1	E	238	TRP	2.8
1	A	262	LYS	2.7
1	E	236	GLU	2.3
2	B	232	ARG	2.2
1	C	256[A]	PHE	2.2
2	B	231	TYR	2.2

### 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 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.