



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 05:23 am GMT

PDB ID : 2Q4C  
Title : Ensemble refinement of the protein crystal structure of annexin from *Arabidopsis thaliana* gene At1g35720  
Authors : Levin, E.J.; Kondrashov, D.A.; Wesenberg, G.E.; Phillips Jr., G.N.; Center for Eukaryotic Structural Genomics (CESG)  
Deposited on : 2007-05-31  
Resolution : 2.51 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](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 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	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

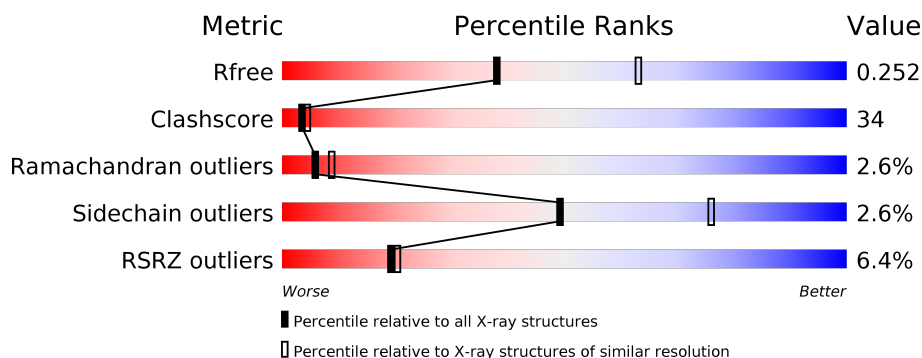
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.51 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))
$R_{free}$	111664	4155 (2.50-2.50)
Clashscore	122126	4827 (2.50-2.50)
Ramachandran outliers	120053	4735 (2.50-2.50)
Sidechain outliers	120020	4737 (2.50-2.50)
RSRZ outliers	108989	4058 (2.50-2.50)

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  $\geq 3$ , 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	1-A	317	<div> <div>4%</div> <div>47%</div> <div>48%</div> <div>• •</div> </div>
1	1-B	317	<div> <div>10%</div> <div>41%</div> <div>53%</div> <div>• •</div> </div>
1	2-A	317	<div> <div>4%</div> <div>44%</div> <div>52%</div> <div>• •</div> </div>
1	2-B	317	<div> <div>10%</div> <div>49%</div> <div>47%</div> <div>• •</div> </div>
1	3-A	317	<div> <div>4%</div> <div>41%</div> <div>53%</div> <div>• • •</div> </div>
1	3-B	317	<div> <div>10%</div> <div>42%</div> <div>52%</div> <div>• •</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	4-A	317	
1	4-B	317	
1	5-A	317	
1	5-B	317	
1	6-A	317	
1	6-B	317	
1	7-A	317	
1	7-B	317	
1	8-A	317	
1	8-B	317	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 42008 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 D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	2-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	3-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	4-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	5-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	6-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	7-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	8-A	312	Total	C	N	O	S	0	0	0
			2508	1557	442	504	5			
1	1-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	2-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	3-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	4-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	5-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	6-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	7-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			
1	8-B	307	Total	C	N	O	S	0	0	0
			2462	1526	436	495	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	EXPRESSION TAG	UNP Q9SYT0
B	1	SER	-	EXPRESSION TAG	UNP Q9SYT0

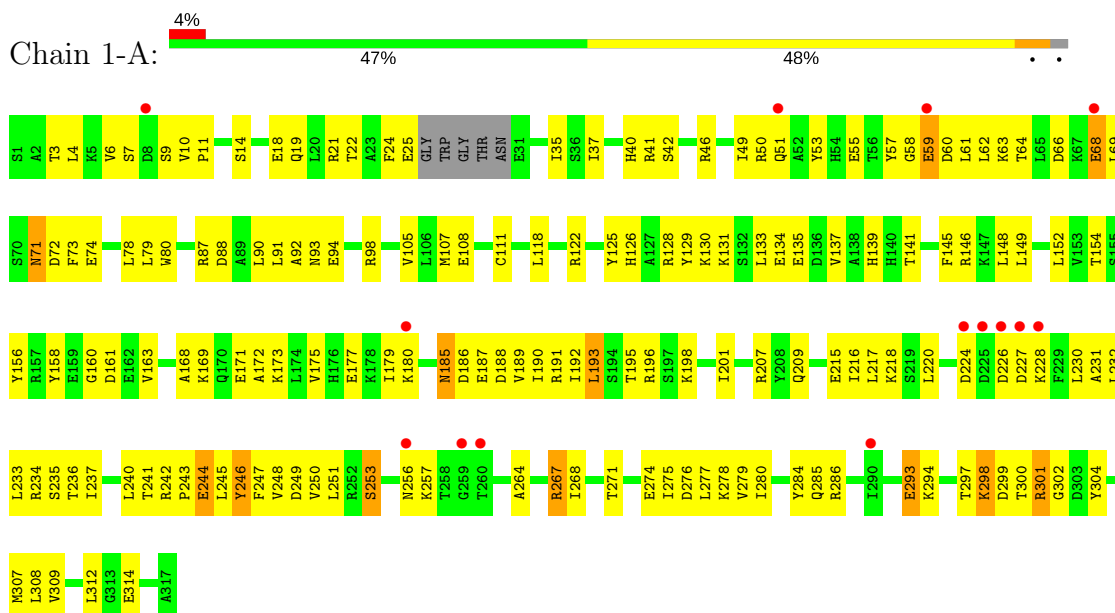
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	1-A	170	Total	O	0	0
			170	170		
2	2-A	167	Total	O	0	0
			167	167		
2	3-A	164	Total	O	0	0
			164	164		
2	4-A	167	Total	O	0	0
			167	167		
2	5-A	171	Total	O	0	0
			171	171		
2	6-A	165	Total	O	0	0
			165	165		
2	7-A	165	Total	O	0	0
			165	165		
2	8-A	165	Total	O	0	0
			165	165		
2	1-B	111	Total	O	0	0
			111	111		
2	2-B	114	Total	O	0	0
			114	114		
2	3-B	117	Total	O	0	0
			117	117		
2	4-B	114	Total	O	0	0
			114	114		
2	5-B	110	Total	O	0	0
			110	110		
2	6-B	116	Total	O	0	0
			116	116		
2	7-B	116	Total	O	0	0
			116	116		
2	8-B	116	Total	O	0	0
			116	116		

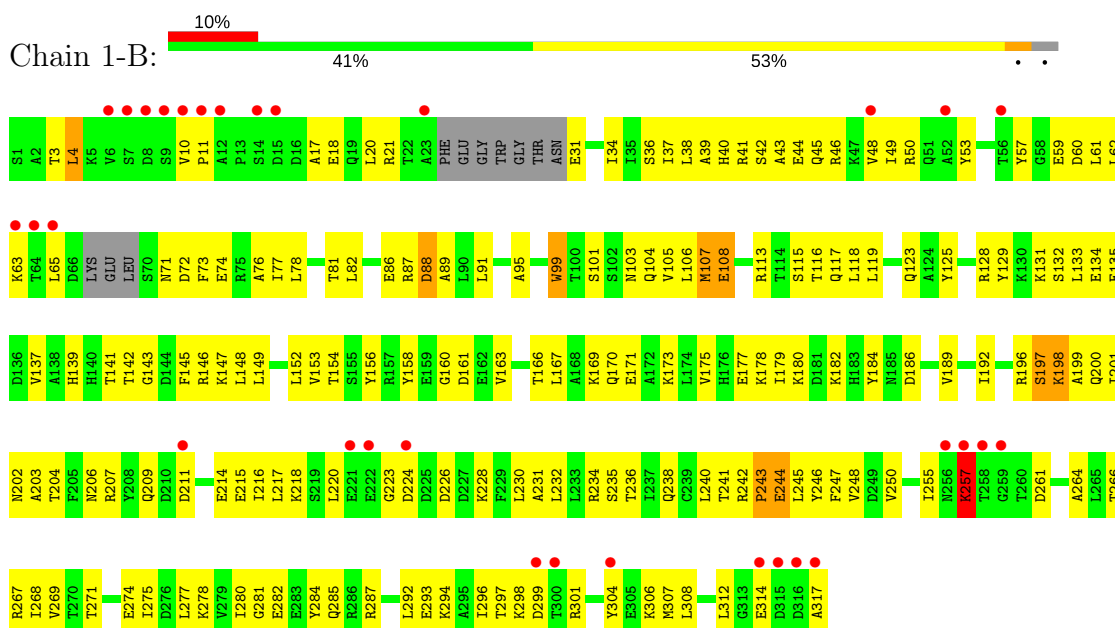
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

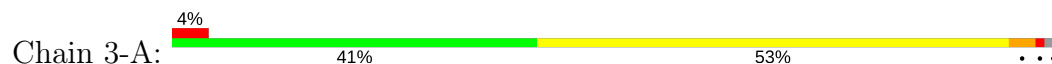
#### • Molecule 1: Annexin D1

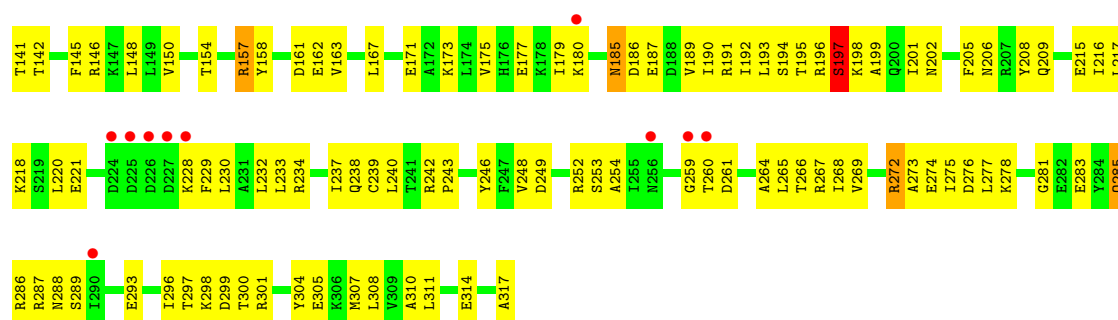


#### • Molecule 1: Annexin D1

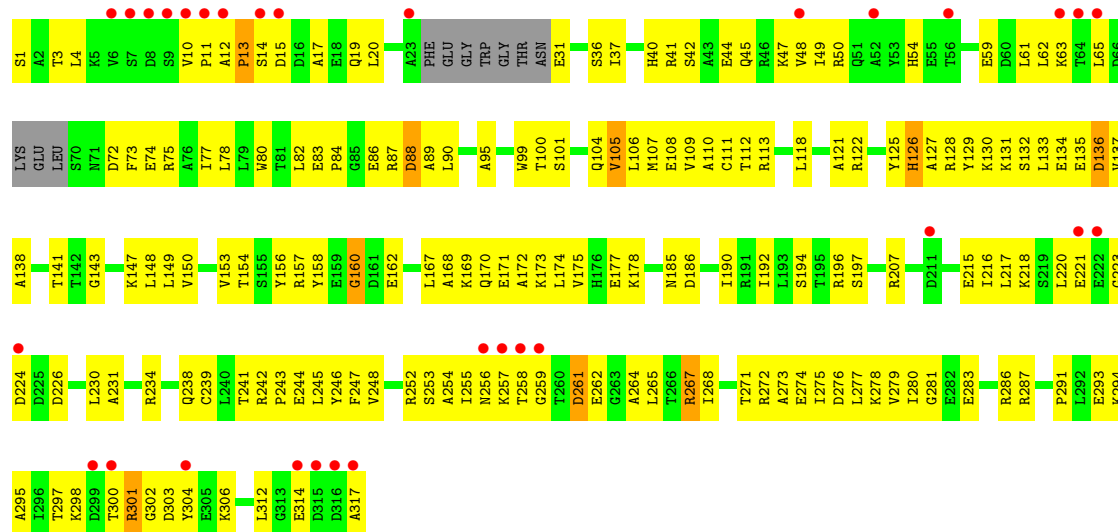
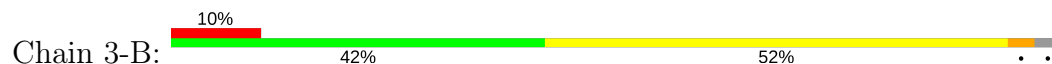


Chain 2-A:  4% 44% 52%

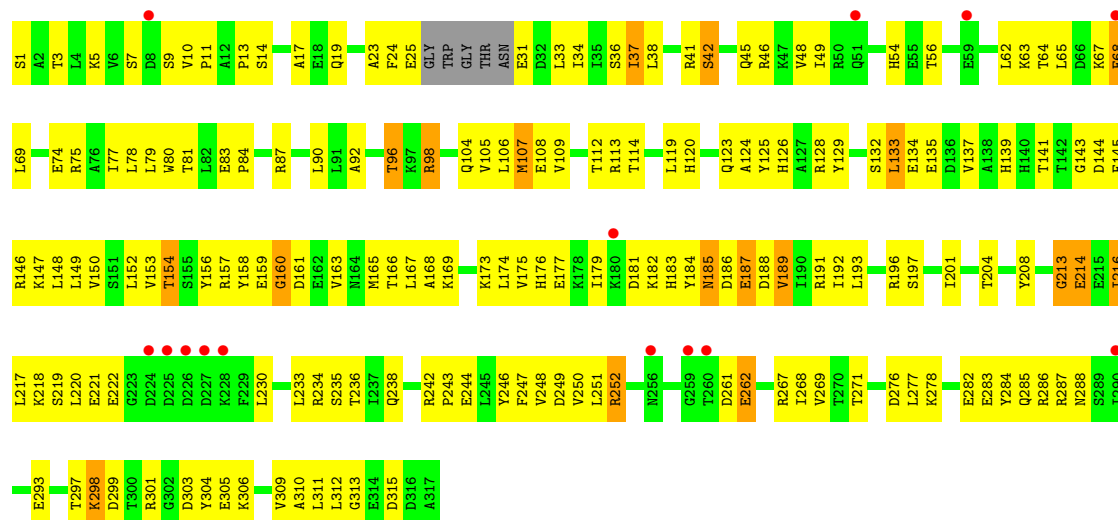
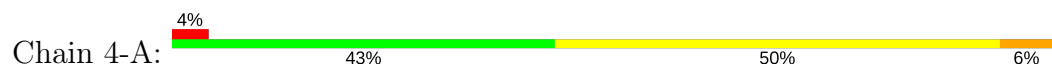




• Molecule 1: Annexin D1

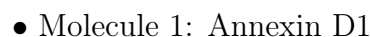
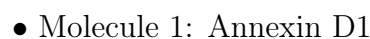


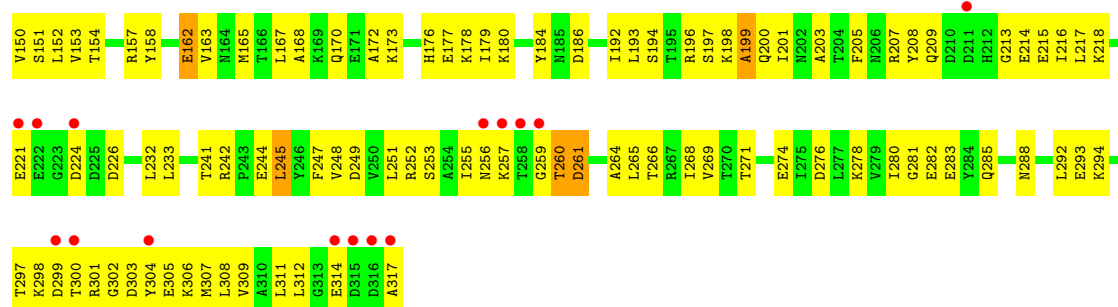
• Molecule 1: Annexin D1



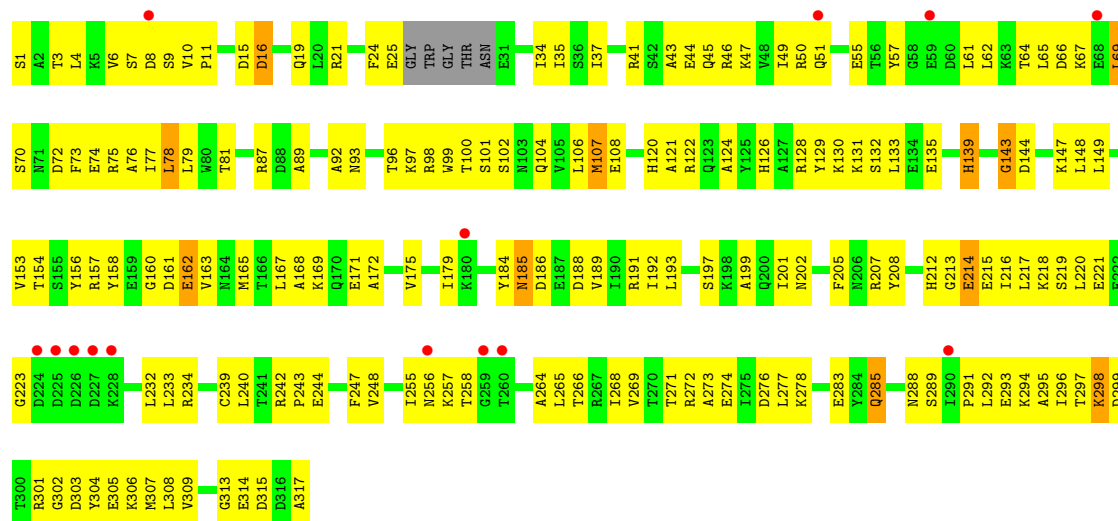
• Molecule 1: Annexin D1



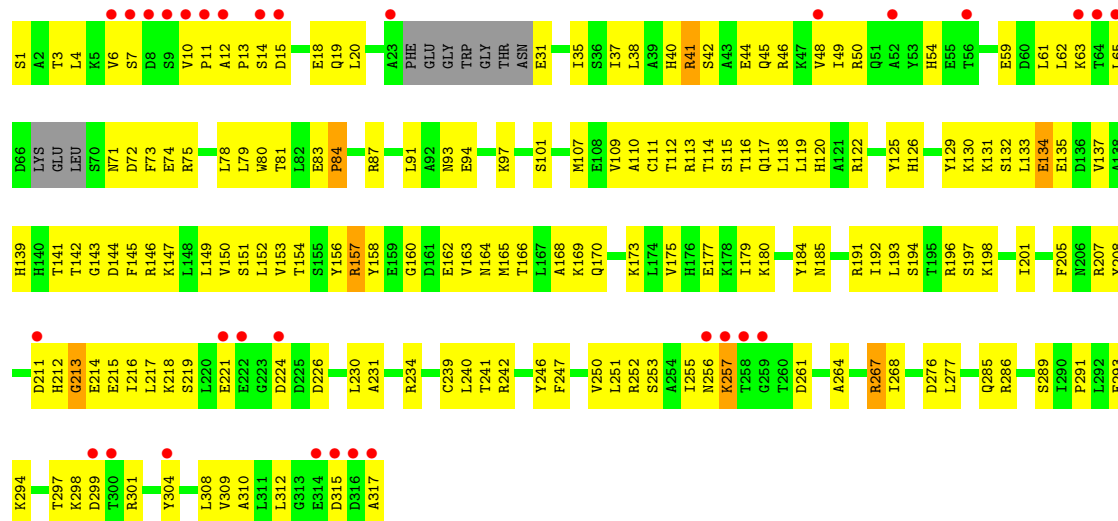




• Molecule 1: Annexin D1

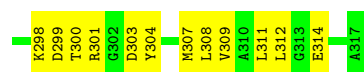


• Molecule 1: Annexin D1

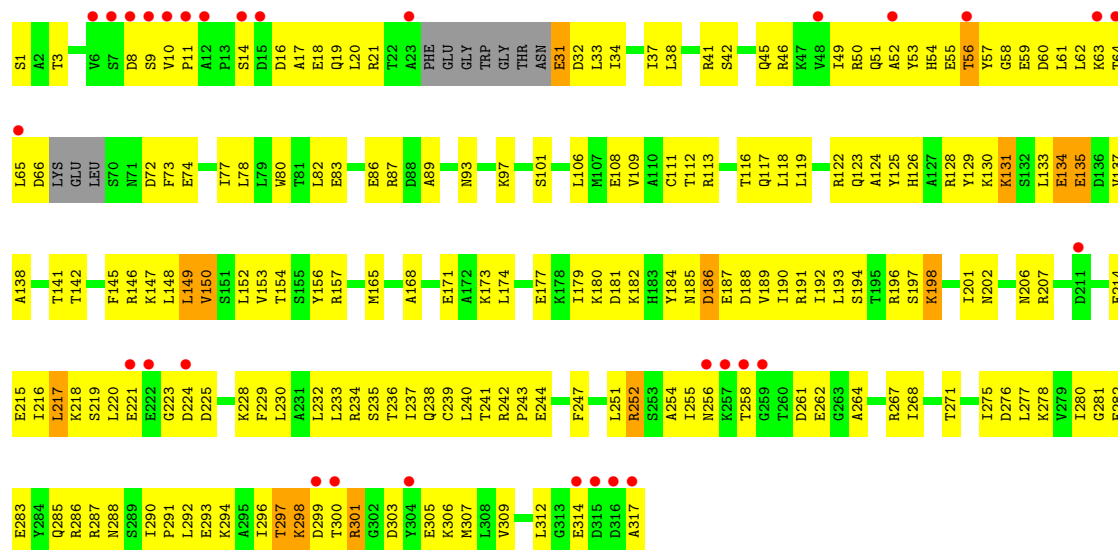


• Molecule 1: Annexin D1





● Molecule 1: Annexin D1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.54Å 96.94Å 226.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.81 – 2.51 33.12 – 2.51	Depositor EDS
% Data completeness (in resolution range)	93.6 (29.81-2.51) 93.5 (33.12-2.51)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.79 (at 2.51Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.177 , 0.243 0.187 , 0.252	Depositor DCC
$R_{free}$ test set	1320 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.5	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.21 , 78.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	42008	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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	1-A	0.32	0/2541	0.54	0/3428
1	1-B	0.29	0/2493	0.53	0/3363
1	2-A	0.32	0/2541	0.55	0/3428
1	2-B	0.29	0/2493	0.53	0/3363
1	3-A	0.31	0/2541	0.54	0/3428
1	3-B	0.29	0/2493	0.53	0/3363
1	4-A	0.33	0/2541	0.54	0/3428
1	4-B	0.29	0/2493	0.52	0/3363
1	5-A	0.32	0/2541	0.57	0/3428
1	5-B	0.30	0/2493	0.54	0/3363
1	6-A	0.33	0/2541	0.57	0/3428
1	6-B	0.30	0/2493	0.55	0/3363
1	7-A	0.34	0/2541	0.56	0/3428
1	7-B	0.31	0/2493	0.54	0/3363
1	8-A	0.34	0/2541	0.56	0/3428
1	8-B	0.31	0/2493	0.56	0/3363
All	All	0.31	0/40272	0.55	0/54328

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	1-A	2508	0	2493	171	0
1	1-B	2462	0	2447	185	0
1	2-A	2508	0	2493	186	0
1	2-B	2462	0	2447	135	0
1	3-A	2508	0	2493	185	0
1	3-B	2462	0	2447	184	0
1	4-A	2508	0	2493	188	0
1	4-B	2462	0	2447	151	0
1	5-A	2508	0	2493	189	0
1	5-B	2462	0	2447	150	0
1	6-A	2508	0	2493	173	0
1	6-B	2462	0	2447	162	0
1	7-A	2508	0	2493	155	0
1	7-B	2462	0	2447	163	0
1	8-A	2508	0	2493	176	0
1	8-B	2462	0	2447	205	0
2	1-A	170	0	0	14	0
2	1-B	111	0	0	7	0
2	2-A	167	0	0	14	0
2	2-B	114	0	0	4	0
2	3-A	164	0	0	13	0
2	3-B	117	0	0	5	0
2	4-A	167	0	0	14	0
2	4-B	114	0	0	12	0
2	5-A	171	0	0	19	0
2	5-B	110	0	0	3	0
2	6-A	165	0	0	12	0
2	6-B	116	0	0	15	0
2	7-A	165	0	0	11	0
2	7-B	116	0	0	10	0
2	8-A	165	0	0	17	0
2	8-B	116	0	0	14	0
All	All	42008	0	39520	2690	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 34.

The worst 5 of 2690 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:134:GLU:HG2	1:B:154:THR:HG22	1.28	1.15
1:B:90:LEU:HD13	1:B:128:ARG:HH21	1.12	1.10

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:134:GLU:HG2	1:A:154:THR:HG22	1.33	1.09
1:A:272:ARG:HH11	1:A:272:ARG:HA	1.14	1.08
1:B:267:ARG:HB3	1:B:267:ARG:HH11	1.11	1.06

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	1-A	308/317 (97%)	268 (87%)	35 (11%)	5 (2%)	11	19
1	1-B	301/317 (95%)	249 (83%)	42 (14%)	10 (3%)	4	5
1	2-A	308/317 (97%)	266 (86%)	38 (12%)	4 (1%)	13	23
1	2-B	301/317 (95%)	248 (82%)	46 (15%)	7 (2%)	7	11
1	3-A	308/317 (97%)	267 (87%)	32 (10%)	9 (3%)	5	7
1	3-B	301/317 (95%)	241 (80%)	48 (16%)	12 (4%)	3	4
1	4-A	308/317 (97%)	254 (82%)	40 (13%)	14 (4%)	3	3
1	4-B	301/317 (95%)	258 (86%)	40 (13%)	3 (1%)	17	31
1	5-A	308/317 (97%)	262 (85%)	40 (13%)	6 (2%)	9	15
1	5-B	301/317 (95%)	253 (84%)	38 (13%)	10 (3%)	4	5
1	6-A	308/317 (97%)	271 (88%)	30 (10%)	7 (2%)	7	11
1	6-B	301/317 (95%)	266 (88%)	29 (10%)	6 (2%)	8	13
1	7-A	308/317 (97%)	256 (83%)	46 (15%)	6 (2%)	9	15
1	7-B	301/317 (95%)	257 (85%)	39 (13%)	5 (2%)	10	17
1	8-A	308/317 (97%)	263 (85%)	36 (12%)	9 (3%)	5	7
1	8-B	301/317 (95%)	249 (83%)	39 (13%)	13 (4%)	3	3
All	All	4872/5072 (96%)	4128 (85%)	618 (13%)	126 (3%)	6	9



5 of 126 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-A	298	LYS
1	1-B	88	ASP
1	1-B	108	GLU
1	2-A	13	PRO
1	2-B	97	LYS

### 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	1-A	274/277 (99%)	265 (97%)	9 (3%)	41	68
1	1-B	269/277 (97%)	264 (98%)	5 (2%)	60	83
1	2-A	274/277 (99%)	267 (97%)	7 (3%)	49	76
1	2-B	269/277 (97%)	265 (98%)	4 (2%)	67	87
1	3-A	274/277 (99%)	264 (96%)	10 (4%)	38	65
1	3-B	269/277 (97%)	265 (98%)	4 (2%)	67	87
1	4-A	274/277 (99%)	263 (96%)	11 (4%)	34	60
1	4-B	269/277 (97%)	261 (97%)	8 (3%)	44	71
1	5-A	274/277 (99%)	264 (96%)	10 (4%)	38	65
1	5-B	269/277 (97%)	266 (99%)	3 (1%)	76	91
1	6-A	274/277 (99%)	264 (96%)	10 (4%)	38	65
1	6-B	269/277 (97%)	263 (98%)	6 (2%)	55	80
1	7-A	274/277 (99%)	269 (98%)	5 (2%)	62	84
1	7-B	269/277 (97%)	267 (99%)	2 (1%)	85	95
1	8-A	274/277 (99%)	266 (97%)	8 (3%)	45	72
1	8-B	269/277 (97%)	259 (96%)	10 (4%)	37	64
All	All	4344/4432 (98%)	4232 (97%)	112 (3%)	49	76

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	4-B	16	ASP
1	5-A	136	ASP
1	8-B	32	ASP
1	4-B	32	ASP
1	4-B	183	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 92 such sidechains are listed below:

Mol	Chain	Res	Type
1	4-A	212	HIS
1	5-A	200	GLN
1	8-A	206	ASN
1	4-A	285	GLN
1	4-B	170	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, 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	1-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	1-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	2-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	2-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	3-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	3-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	4-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	4-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	5-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	5-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	6-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	6-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	7-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	7-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
1	8-A	312/317 (98%)	-0.01	14 (4%) 33 36	11, 35, 68, 92	312 (100%)
1	8-B	307/317 (96%)	0.30	31 (10%) 7 6	11, 42, 76, 92	307 (100%)
All	All	4952/5072 (97%)	0.15	360 (7%) 19 16	11, 37, 75, 92	4952 (100%)

The worst 5 of 360 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-B	64	THR	6.9
1	2-B	64	THR	6.9
1	3-B	64	THR	6.9
1	4-B	64	THR	6.9
1	5-B	64	THR	6.9

## 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.