WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 15, 2023 – 06:35 PM EDT

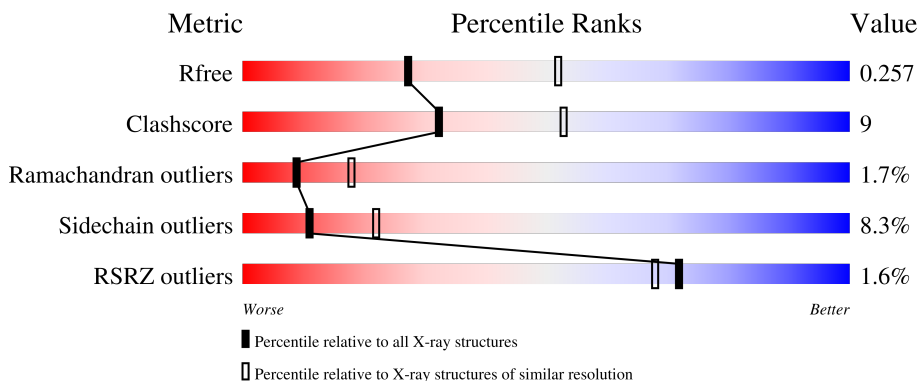|		
PDB ID	:	2A1S
Title	:	Crystal structure of native PARN nuclease domain
Authors	:	Wu, M.; Song, H.
Deposited on	:	2005-06-21
Resolution	:	2.60 Å(reported)

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 2.60 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		0.257
Clashscore		9
Ramachandran outliers		1.7%
Sidechain outliers		8.3%
RSRZ outliers		1.6%

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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	430	3% 67% 20% • 9%
1	B	430	.% 67% 20% • • 9%
1	C	430	65% 22% • 9%
1	D	430	2% 67% 21% • 9%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 13115 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 Poly(A)-specific ribonuclease PARN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	391	Total	C	N	O	S	0	0	0
			3213	2067	529	602	15			
1	B	390	Total	C	N	O	S	0	0	0
			3210	2064	529	603	14			
1	C	393	Total	C	N	O	S	0	0	0
			3238	2083	532	608	15			
1	D	392	Total	C	N	O	S	0	0	0
			3215	2069	530	601	15			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	61	Total	O	0	0
			61	61		
2	B	63	Total	O	0	0
			63	63		
2	C	59	Total	O	0	0
			59	59		
2	D	56	Total	O	0	0
			56	56		

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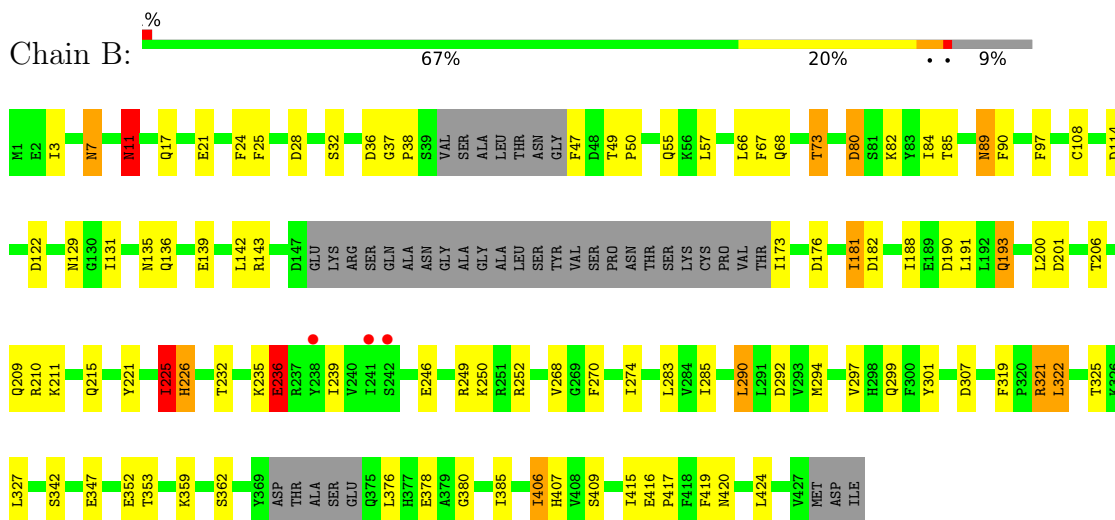 least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
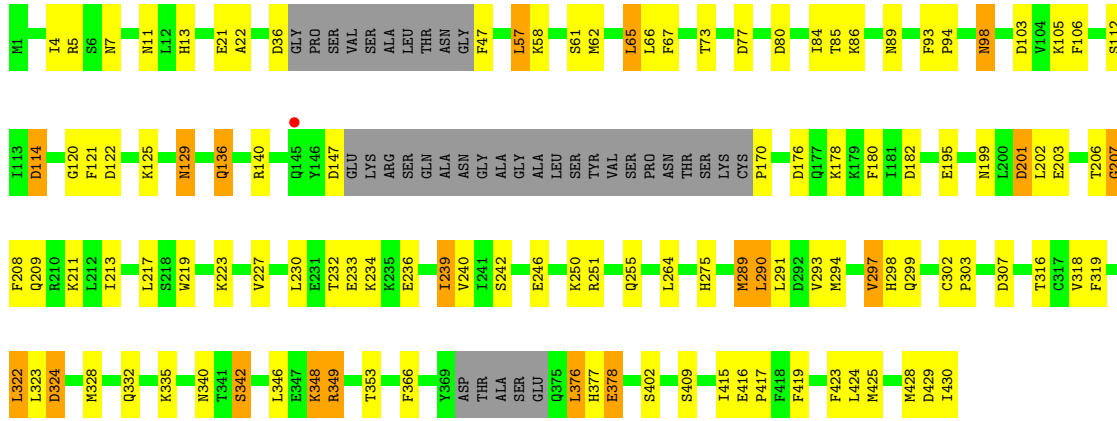
- Molecule 1: Poly(A)-specific ribonuclease PARN

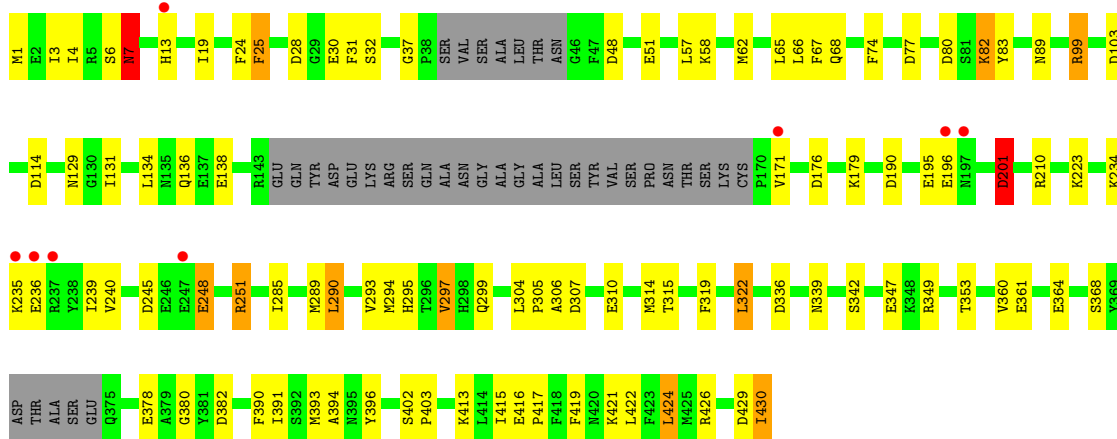- Molecule 1: Poly(A)-specific ribonuclease PARN

- Molecule 1: Poly(A)-specific ribonuclease PARN

M1 I4 R5 S6 N7 N11 L12 H13 E21 A22 D36 GLY PRO SER VAL SER ALA LEU THR ASN GLY F47 L57 K58 S61 M62 L65 L66 F67 T73 D77 D80 I84 T85 K86 N89 F93 P94 N98 D103 V104 K105 F106 S112

I113 D114 G120 F121 D122 K125 N129 Q136 R140 Q145 Y146 D147 GLU LYS ARG SER GLN ALA ASN GLY ALA GLY ALA LEU SER TYR VAL SER PRO ASN THR SER LYS CYS P170 D176 Q177 K178 K179 F180 I181 D182 E195 N199 L200 D201 L202 E203 T206 G207

F208 Q209 R210 K211 L212 I213 L217 S218 W219 K223 V227 L230 E231 T232 E233 K234 K235 E236 I239 V240 I241 S242 E246 K250 R251 Q255 L264 H275 M289 L290 L291 D292 V293 M294 V297 H298 Q299 C302 P303 D307 T316 C317 V318 F319

L322 L323 D324 M328 Q332 K335 N340 T341 S342 L346 E347 K348 R349 T353 F366 Y369 ASP THR ALA SER GLU Q375 L376 H377 E378 S402 S409 I415 E416 P417 F418 F419 F423 L424 M425 M428 D429 I430

• Molecule 1: Poly(A)-specific ribonuclease PARN

Chain D:

2%   67%   21%   •   9%

M1 E2 I3 I4 R5 S6 N7 H13 I19 F24 F25 D28 G29 E30 F31 S32 G37 P38 SER VAL SER ALA LEU THR ASN G46 F47 D48 E51 L57 K58 M62 L65 L66 F67 Q68 F74 D77 D80 S81 K82 Y83 N89 R99 D103

D114 N129 G130 I131 L134 N135 Q136 E137 E138 R143 GLU GLN TYR ASP GLU LYS ARG SER GLN ALA ASN GLY ALA GLY ALA LEU SER TYR VAL SER PRO ASN THR SER LYS CYS P170 V171 D176 K179 D190 E195 E196 N197 D201 R210 K223 K234

K235 E236 R237 Y238 I239 V240 D245 E246 E247 E248 R251 I285 M289 L290 V293 M294 H295 T296 V297 H298 Q299 L304 P305 A306 D307 E310 M314 T315 F319 L322 D336 N339 S342 E347 K348 R349 T353 V360 E361 E364 S368 Y369

ASP THR ALA SER GLU Q375 E378 A379 G380 Y381 D382 F390 I391 S392 M393 A394 N395 Y396 S402 P403 K413 L414 I415 E416 P417 F418 F419 N420 K421 L422 F423 L424 M425 R426 D429 I430

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	205.54Å 123.02Å 82.84Å 90.00° 112.59° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60 75.35 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.4 (20.00-2.60) 99.4 (75.35-2.60)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 2.62Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.219 , 0.254 0.225 , 0.257	Depositor DCC
R_{free} test set	1786 reflections (3.07%)	wwPDB-VP
Wilson B-factor (Å^2)	55.5	Xtriage
Anisotropy	0.283	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.019 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	13115	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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	0.56	0/3288	0.77	13/4425 (0.3%)
1	B	0.57	0/3286	0.80	11/4423 (0.2%)
1	C	0.60	1/3314 (0.0%)	0.79	9/4461 (0.2%)
1	D	0.59	0/3291	0.79	13/4430 (0.3%)
All	All	0.58	1/13179 (0.0%)	0.79	46/17739 (0.3%)

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	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	289	MET	SD-CE	-5.62	1.46	1.77

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	103	ASP	CB-CG-OD2	8.20	125.68	118.30
1	B	182	ASP	CB-CG-OD2	8.16	125.64	118.30
1	C	77	ASP	CB-CG-OD2	7.44	125.00	118.30
1	B	28	ASP	CB-CG-OD2	6.74	124.36	118.30
1	A	63	ASP	CB-CG-OD2	6.64	124.28	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	402	SER	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	3213	0	3167	57	0
1	B	3210	0	3151	43	0
1	C	3238	0	3184	72	0
1	D	3215	0	3171	53	0
2	A	61	0	0	4	0
2	B	63	0	0	1	0
2	C	59	0	0	7	0
2	D	56	0	0	2	0
All	All	13115	0	12673	224	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 9.

The worst 5 of 224 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:4:ILE:H	1:D:7:ASN:HD21	1.12	0.93
1:A:62:MET:HE1	1:A:295:HIS:HA	1.53	0.90
1:B:89:ASN:HD21	1:B:380:GLY:H	1.24	0.85
1:B:66:LEU:H	1:B:299:GLN:HE22	1.24	0.81
1:D:66:LEU:H	1:D:299:GLN:HE22	1.27	0.80

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	383/430 (89%)	355 (93%)	20 (5%)	8 (2%)	7	13
1	B	382/430 (89%)	351 (92%)	22 (6%)	9 (2%)	6	10
1	C	385/430 (90%)	362 (94%)	19 (5%)	4 (1%)	15	32
1	D	384/430 (89%)	358 (93%)	21 (6%)	5 (1%)	12	24
All	All	1534/1720 (89%)	1426 (93%)	82 (5%)	26 (2%)	9	18

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	ASN
1	A	95	LYS
1	A	98	ASN
1	B	225	ILE
1	B	236	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	360/391 (92%)	335 (93%)	25 (7%)	15	31
1	B	359/391 (92%)	326 (91%)	33 (9%)	9	17
1	C	363/391 (93%)	330 (91%)	33 (9%)	9	18
1	D	360/391 (92%)	332 (92%)	28 (8%)	12	25
All	All	1442/1564 (92%)	1323 (92%)	119 (8%)	11	22

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	409	SER
1	D	339	ASN
1	C	203	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	322	LEU
1	D	430	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	89	ASN
1	D	299	GLN
1	D	98	ASN
1	D	183	GLN
1	B	129	ASN

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(Å^2)	Q<0.9
1	A	391/430 (90%)	0.15	13 (3%)	46	39	35, 54, 101, 117	0
1	B	390/430 (90%)	0.07	3 (0%)	86	84	31, 53, 88, 103	0
1	C	393/430 (91%)	-0.05	1 (0%)	94	93	31, 49, 78, 92	0
1	D	392/430 (91%)	0.02	8 (2%)	65	60	28, 50, 81, 98	0
All	All	1566/1720 (91%)	0.05	25 (1%)	72	68	28, 51, 90, 117	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	254	GLN	4.8
1	D	197	ASN	4.6
1	A	97	PHE	4.5
1	D	235	LYS	4.3
1	B	238	TYR	3.6

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.