W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Dec 18, 2024 – 12:07 pm GMT

PDB ID	:	9ERU
Title	:	Mouse CNPase catalytic domain with nanobody 7E
Authors	:	Markusson, S.; Raasakka, A.; Opazo, F.; Kursula, P.
Deposited on	:	2024-03-25
Resolution	:	2.50 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

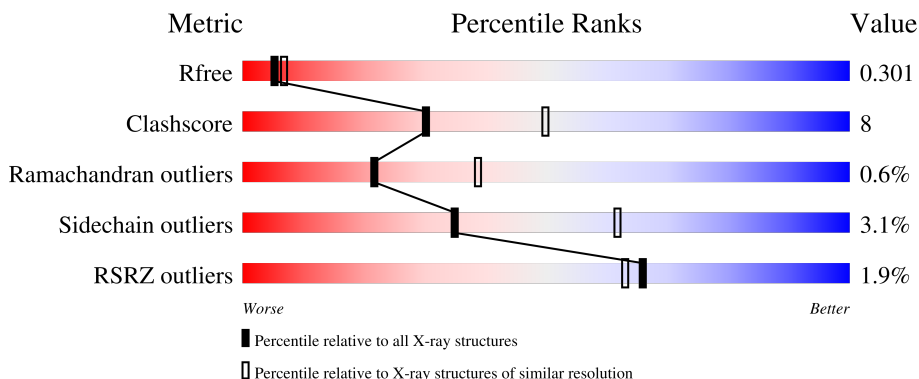|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.50 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		0.301
Clashscore		8
Ramachandran outliers		0.6%
Sidechain outliers		3.1%
RSRZ outliers		1.9%

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}	164625	5504 (2.50-2.50)
Clashscore	180529	6282 (2.50-2.50)
Ramachandran outliers	177936	6191 (2.50-2.50)
Sidechain outliers	177891	6193 (2.50-2.50)
RSRZ outliers	164620	5504 (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 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	220	79% 14% 6%
1	B	220	3% 76% 18% · ·
1	D	220	3% 78% 14% · 7%
1	F	220	2% 74% 19% · 5%
2	C	127	2% 75% 24% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	127	2% 80% 18%
2	G	127	% 76% 21%
2	H	127	2% 79% 20%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 20570 atoms, of which 10234 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 2',3'-cyclic-nucleotide 3'-phosphodiesterase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	213	Total	C	H	N	O	S	0	0	0
			3351	1074	1689	279	304	5			
1	A	206	Total	C	H	N	O	S	0	0	0
			3228	1036	1626	270	291	5			
1	D	204	Total	C	H	N	O	S	0	0	0
			3199	1027	1613	267	287	5			
1	F	208	Total	C	H	N	O	S	0	0	0
			3269	1048	1650	273	293	5			

- Molecule 2 is a protein called Chains: G,C,E,H.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	G	126	Total	C	H	N	O	S	0	0	0
			1880	602	914	162	196	6			
2	C	126	Total	C	H	N	O	S	0	0	0
			1880	602	914	162	196	6			
2	E	126	Total	C	H	N	O	S	0	0	0
			1880	602	914	162	196	6			
2	H	126	Total	C	H	N	O	S	0	0	0
			1880	602	914	162	196	6			

- Molecule 3 is water.

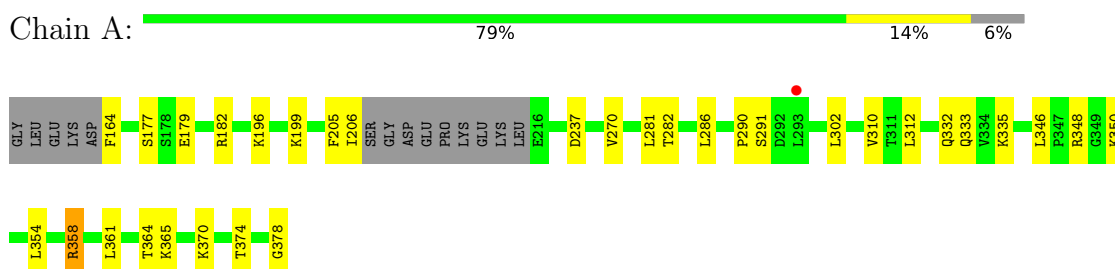| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total	O	0	0
			1	1		
3	F	2	Total	O	0	0
			2	2		

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.
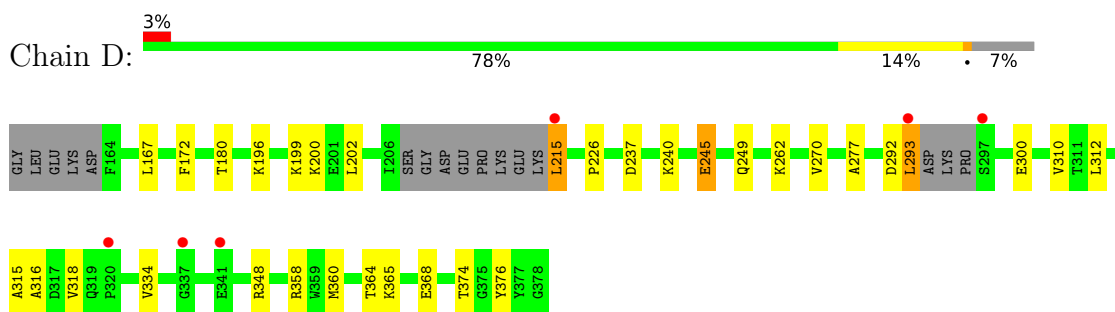
• Molecule 1: 2',3'-cyclic-nucleotide 3'-phosphodiesterase
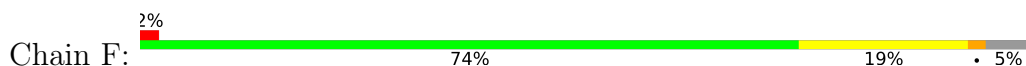
Chain B:

• Molecule 1: 2',3'-cyclic-nucleotide 3'-phosphodiesterase

Chain A:

• Molecule 1: 2',3'-cyclic-nucleotide 3'-phosphodiesterase

Chain D:

• Molecule 1: 2',3'-cyclic-nucleotide 3'-phosphodiesterase

Chain F:

GLY LEU GLU LYS ASP F164 K175 K176 T180 L181 R182 G185 L189 H195 K196 K199 R203 I206 SER GLY ASP GLU PRO LYS GLU K214 L215 K223 P226 L229 T233 K234 K240 Q249 Q250 E251 V252 V253 K254 K259 A260 F261 A267

V270 T271 P272 K273 T282 L286 Q287 L288 W289 P290 L293 D294 K295 P296 A315 V318 D326 Q333 S338 A342 R348 G349 K350 R358 T364 K365 K370 T374 G378

• Molecule 2: Chains: G,C,E,H

Chain G:

E1 E6 V12 S17 L18 R19 L20 N32 T33 M34 K43 Q44 R45 E46 Q47 V51 I54 T55 D56 Y57 S65 R69 N86 K89 P90 T93 M100 D101 L102 T103 A107 L111 G118 T121 V125 S126 SER

• Molecule 2: Chains: G,C,E,H

Chain C:

E1 E6 V12 S17 L18 R19 L20 A23 N27 N32 Q39 V51 I54 T55 D56 Y57 T60 A63 K67 A77 T80 P90 T93 M100 D101 L102 T103 L111 Y112 A113 Y114 D115 G118 T121 V125 S126 SER

• Molecule 2: Chains: G,C,E,H

Chain E:

E1 S17 L18 R19 L20 V31 Q39 K43 Q44 R45 R48 E49 L50 V51 I54 T55 D56 T60 Q84 M85 N86 T93 C98 N99 M100 D101 L102 T106 L111 Y112 T121 V125 S126 SER

• Molecule 2: Chains: G,C,E,H

Chain H:

E1 W11 V12 G16 S17 L18 R19 L20 S21 C22 F29 N32 T33 R38 Q44 I54 T55 D56 T60 A77 M85 L88 T93 M100 D101 L102 T103 L111 T121 V125 S126 SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	37.15Å 79.00Å 117.59Å 77.67° 87.71° 84.29°	Depositor
Resolution (Å)	41.96 – 2.50 41.96 – 2.50	Depositor EDS
% Data completeness (in resolution range)	91.5 (41.96-2.50) 78.6 (41.96-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.60 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.21rc1_5127	Depositor
R, R_{free}	0.255 , 0.300 0.254 , 0.301	Depositor DCC
R_{free} test set	39229 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	43.4	Xtriage
Anisotropy	0.686	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 40.7	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	20570	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.31% 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 (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	0.30	0/1639	0.50	0/2207
1	B	0.30	0/1699	0.48	0/2286
1	D	0.31	0/1621	0.50	0/2181
1	F	0.31	0/1656	0.50	0/2229
2	C	0.31	0/986	0.54	0/1340
2	E	0.31	0/986	0.55	0/1340
2	G	0.30	0/986	0.53	0/1340
2	H	0.32	0/986	0.54	0/1340
All	All	0.31	0/10559	0.51	0/14263

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	1602	1626	1625	20	0
1	B	1662	1689	1689	35	0
1	D	1586	1613	1611	17	1
1	F	1619	1650	1649	32	1
2	C	966	914	914	28	0
2	E	966	914	914	15	0
2	G	966	914	914	23	0
2	H	966	914	914	20	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	1	0	0	0	0
3	F	2	0	0	0	0
All	All	10336	10234	10230	171	2

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 8.

The worst 5 of 171 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:179:GLU:OE1	2:C:77:ALA:HB2	1.75	0.86
1:F:250:GLN:O	1:F:254:LYS:HG3	1.76	0.85
2:H:12:VAL:HG21	2:H:18:LEU:HD13	1.62	0.81
1:F:182:ARG:HG3	1:F:229:LEU:HD11	1.64	0.79
1:B:170:GLY:HA2	1:B:233:THR:HG23	1.65	0.77

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
1:D:180:THR:OG1	1:D:240:LYS:O[1_655]	1.97	0.23
1:F:180:THR:OG1	1:F:240:LYS:O[1_655]	2.17	0.03

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	202/220 (92%)	194 (96%)	6 (3%)	2 (1%)	13	25
1	B	209/220 (95%)	195 (93%)	11 (5%)	3 (1%)	9	17
1	D	198/220 (90%)	191 (96%)	5 (2%)	2 (1%)	13	25

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	204/220 (93%)	197 (97%)	7 (3%)	0	100	100
2	C	124/127 (98%)	117 (94%)	7 (6%)	0	100	100
2	E	124/127 (98%)	118 (95%)	6 (5%)	0	100	100
2	G	124/127 (98%)	117 (94%)	6 (5%)	1 (1%)	16	31
2	H	124/127 (98%)	118 (95%)	6 (5%)	0	100	100
All	All	1309/1388 (94%)	1247 (95%)	54 (4%)	8 (1%)	22	39

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	162	LYS
1	B	163	ASP
1	B	296	PRO
1	A	290	PRO
1	D	300	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	168/180 (93%)	164 (98%)	4 (2%)	44	70
1	B	175/180 (97%)	170 (97%)	5 (3%)	37	64
1	D	166/180 (92%)	161 (97%)	5 (3%)	36	63
1	F	170/180 (94%)	163 (96%)	7 (4%)	26	50
2	C	102/103 (99%)	99 (97%)	3 (3%)	37	64
2	E	102/103 (99%)	98 (96%)	4 (4%)	27	52
2	G	102/103 (99%)	99 (97%)	3 (3%)	37	64
2	H	102/103 (99%)	99 (97%)	3 (3%)	37	64
All	All	1087/1132 (96%)	1053 (97%)	34 (3%)	35	62

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	273	LYS
1	F	338	SER
2	H	38	ARG
2	C	27	ASN
1	A	358	ARG

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

Mol	Chain	Res	Type
2	C	32	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 oligosaccharides 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	206/220 (93%)	0.50	1 (0%)	87	85	48, 75, 112, 144	0
1	B	213/220 (96%)	0.50	6 (2%)	55	51	50, 77, 119, 172	0
1	D	204/220 (92%)	0.35	6 (2%)	54	50	44, 71, 117, 161	0
1	F	208/220 (94%)	0.42	4 (1%)	66	63	45, 69, 122, 141	0
2	C	126/127 (99%)	0.61	2 (1%)	70	67	57, 78, 112, 127	0
2	E	126/127 (99%)	0.51	3 (2%)	59	56	46, 72, 108, 120	0
2	G	126/127 (99%)	0.53	1 (0%)	82	79	63, 79, 99, 119	0
2	H	126/127 (99%)	0.32	2 (1%)	70	67	47, 68, 95, 118	0
All	All	1335/1388 (96%)	0.46	25 (1%)	66	63	44, 74, 112, 172	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	293	LEU	3.1
1	D	297	SER	3.1
2	C	126	SER	3.0
1	F	164	PHE	2.9
1	B	160	LEU	2.8

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.4　Ligands ⓘ

There are no ligands in this entry.

6.5　Other polymers ⓘ

There are no such residues in this entry.