WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

May 23, 2022 – 01:18 pm BST

PDB ID	:	7R24
Title	:	Crystal structure of rat Arc CTD in complex with two anti-Arc nanobodies
Authors	:	Markusson, S.; Kursula, P.
Deposited on	:	2022-02-04
Resolution	:	2.70 Å(reported)

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

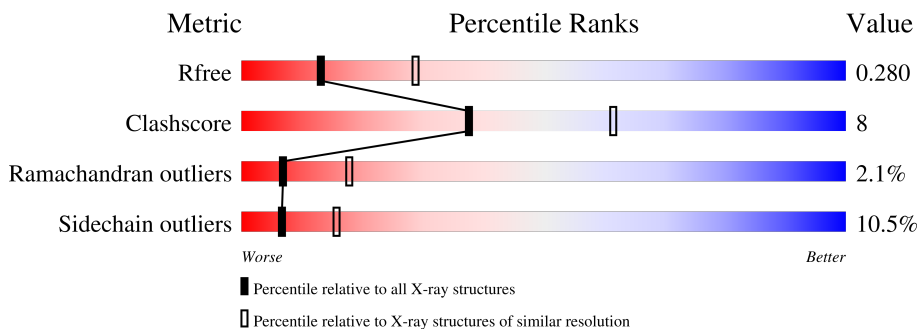|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

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 2.70 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		0.280
Clashscore		8
Ramachandran outliers		2.1%
Sidechain outliers		10.5%

■ 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)

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%

Mol	Chain	Length	Quality of chain
1	A	397	29% 7% • 62%
2	C	120	79% 18% • •
3	E	128	73% 20% • • •

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6236 atoms, of which 3068 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 Activity-regulated cytoskeleton-associated protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	150	Total 2526	C 823	H 1245	N 218	O 238	S 2	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP Q63053
A	114	ALA	MET	engineered mutation	UNP Q63053
A	115	ALA	HIS	engineered mutation	UNP Q63053
A	116	ALA	VAL	engineered mutation	UNP Q63053
A	117	ALA	TRP	engineered mutation	UNP Q63053
A	118	ALA	ARG	engineered mutation	UNP Q63053
A	119	ALA	GLU	engineered mutation	UNP Q63053
A	120	ALA	VAL	engineered mutation	UNP Q63053

- Molecule 2 is a protein called anti-Arc nanobody.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	118	Total 1803	C 570	H 889	N 162	O 179	S 3	0	2	0

- Molecule 3 is a protein called anti-Arc nanobody.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	E	125	Total 1907	C 610	H 934	N 170	O 188	S 5	0	0	0

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. 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. 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: Activity-regulated cytoskeleton-associated protein

Chain A:

29% 7% 62%

SER MET GLU LEU ASP HIS THR MET THR GLY GLY LEU LEU HIS ALA TYR PRO ALA PRO GLY ARG GLY GLY PRO ALA ALA LYS PRO ASN VAL ILE LEU GLN ILE GLY LYS CYS ARG ALA GLU MET LEU GLU HIS VAL ARG ARG THR HIS ARG HIS LEU LEU THR GLU VAL SER LYS GLN VAL GLU ARG

GLU LEU LYS GLY LEU HIS ARG SER VAL GLY LYS LEU GLU ASN ASN LEU ASP GLY TYR VAL PRO THR GLY ASP SER GLN GLN ARG TRP LYS LYS SER ILE LYS ALA CYS LEU CYS ARG CYS GLN GLU THR ILE ALA ASN LEU GLU ARG TRP VAL LYS ARG GLU ALA ALA ALA ALA ALA ALA

PHE TYR ARG LEU GLU GLU ARG TRP ALA ASP ARG LEU GLU SER MET GLY LYS LYS TYR PRO VAL GLY SER GLU PRO ALA ARG HIS THR VAL SER VAL GLY VAL GLY GLY PRO GLU PRO TYR CYS GLN GLU ALA ASP GLY TYR ASP TYR THR VAL SER PRO TYR ALA ILE THR PRO PRO PRO ALA

ALA GLY GLU LEU PRO GLU GLN GLU GLU SER VAL GLY ALA GLN GLN GLN TYR GLN SER TRP VAL PRO GLY GLU ASP GLY GLN PRO SER PRO GLY VAL **D211** **I214** **F215** **E216** **L222** **Q231** **L240** **I243** **Q244** **K252** **F257** **K258** **Q259** **V262** **K269** **L280** **P292** **Q293** **K294**

Q301 **R307** **Q311** **E320** **E321** **I322** **I323** **Q324** **V327** **L330** **R336** **F337** **H340** **P341** **T345** **L349** **R352** **V356** **Q357** **D358** **G359** **L360** GLU GLN ALA ALA GLU PRO SER VAL THR PRO LEU PRO THR GLU GLU ASP GLU THR GLU GLY ALA LEU THR GLU PRO ALA LEU THR SER

GLU SER VAL ALA SER ASP ARG THR GLN PRO GLU

• Molecule 2: anti-Arc nanobody

Chain C:

79% 18%

GLY SER **E3** **V4** **Q5** **E8** **L13** **V14** **L20** **T30** **S31** **V36** **R47** **E48** **Y49** **S74** **D75** **N76** **A77** **V81** **M85** **L88** **T93** **V100** **R101** **T114** **Q115** **V116** **T117** **V118** **S119** **S120**

• Molecule 3: anti-Arc nanobody

Chain E:

73% 20%

GLY SER **E3** **S9** **V14** **Q15** **L20** **R21** **L22** **S23** **C24** **S27** **S32** **Q41** **K45** **S59** **V81** **Y82** **L83** **L88** **T93** **A100** **M104** **Y105** **R106** **I109** **S110** **K111** **G112** **Q113** **N114** **Y117** **T122** **Q123** **V124** **T125** **V126** **S127** SER

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	40.88Å 92.81Å 114.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.51 – 2.70 38.51 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.2 (38.51-2.70) 83.6 (38.51-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	0.62 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.257 , 0.274 0.264 , 0.280	Depositor DCC
R_{free} test set	1248 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	37.6	Xtriage
Anisotropy	0.822	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	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	0.83	EDS
Total number of atoms	6236	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.41% 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	0.25	0/1314	0.45	0/1772
2	C	0.25	0/949	0.53	0/1286
3	E	0.26	0/994	0.51	0/1343
All	All	0.25	0/3257	0.49	0/4401

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	1281	1245	1245	17	0
2	C	914	889	870	15	0
3	E	973	934	934	20	0
All	All	3168	3068	3049	48	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 8.

All (48) 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)
3:E:14:VAL:HG11	3:E:20:LEU:HD11	1.42	0.98

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:14:VAL:HG23	3:E:126:VAL:HG13	1.68	0.75
2:C:36:VAL:HG11	2:C:81:VAL:HG21	1.69	0.73
3:E:93:THR:HG23	3:E:125:THR:HA	1.77	0.67
3:E:14:VAL:HG11	3:E:20:LEU:CD1	2.23	0.66
1:A:352:ARG:O	1:A:356:VAL:HG22	1.97	0.64
3:E:124:VAL:HG12	3:E:124:VAL:O	2.01	0.60
3:E:20:LEU:HD13	3:E:88:LEU:HD21	1.87	0.57
2:C:13:LEU:HD23	2:C:14:VAL:N	2.20	0.56
1:A:321:GLU:OE2	2:C:101:ARG:NE	2.38	0.56
1:A:320:GLU:OE1	1:A:324:GLN:NE2	2.38	0.56
3:E:22:LEU:HD12	3:E:83:LEU:HD23	1.89	0.55
3:E:20:LEU:HD13	3:E:88:LEU:CD2	2.37	0.54
1:A:337:PHE:CG	1:A:356:VAL:HG11	2.43	0.54
1:A:216:GLU:HG2	3:E:109:ILE:HD11	1.90	0.53
2:C:8:GLU:CB	2:C:114:THR:HG23	2.39	0.52
3:E:126:VAL:HG12	3:E:127:SER:H	1.74	0.52
1:A:327:VAL:HA	1:A:330:LEU:HD22	1.92	0.51
1:A:307:ARG:HG2	1:A:322:ILE:HG21	1.93	0.50
1:A:292:PRO:O	1:A:294:LYS:N	2.45	0.50
2:C:13:LEU:HD21	2:C:119:SER:HB3	1.94	0.49
2:C:114:THR:O	2:C:115:GLN:CB	2.60	0.49
2:C:30:THR:HG23	2:C:30:THR:O	2.14	0.48
2:C:114:THR:O	2:C:115:GLN:HB3	2.13	0.48
3:E:124:VAL:O	3:E:124:VAL:CG1	2.62	0.47
1:A:280:LEU:HD21	3:E:106:ARG:CZ	2.44	0.47
3:E:100:ALA:HB3	3:E:117:TYR:HB2	1.96	0.46
3:E:24:CYS:HB3	3:E:81:VAL:HG12	1.98	0.46
1:A:240:LEU:HD23	1:A:243:ILE:HD11	1.98	0.46
1:A:340:HIS:CB	1:A:341:PRO:CD	2.94	0.46
2:C:8:GLU:OE1	2:C:8:GLU:N	2.49	0.45
3:E:20:LEU:HD22	3:E:124:VAL:HG13	1.98	0.45
1:A:214:ILE:HG22	3:E:109:ILE:HD12	1.98	0.44
1:A:340:HIS:HB2	1:A:341:PRO:CD	2.48	0.44
2:C:93:THR:HG23	2:C:117:THR:HA	1.99	0.43
1:A:337:PHE:CD2	1:A:356:VAL:HG11	2.53	0.43
2:C:75:ASP:O	2:C:77:ALA:N	2.48	0.42
3:E:124:VAL:O	3:E:125:THR:C	2.57	0.42
2:C:8:GLU:HB3	2:C:114:THR:HG23	2.01	0.42
1:A:293:GLN:O	1:A:294:LYS:C	2.58	0.42
2:C:8:GLU:HB2	2:C:114:THR:HG23	2.01	0.41
2:C:85:MET:CE	2:C:88:LEU:HD21	2.50	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:41:GLN:NE2	3:E:45:LYS:O	2.53	0.41
3:E:14:VAL:CG1	3:E:20:LEU:HD11	2.30	0.41
1:A:222:LEU:HD21	1:A:269:LYS:HG2	2.02	0.41
1:A:340:HIS:HB2	1:A:341:PRO:HD2	2.04	0.40
3:E:9:SER:N	3:E:23:SER:O	2.53	0.40

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	148/397 (37%)	138 (93%)	5 (3%)	5 (3%)	3	8
2	C	118/120 (98%)	110 (93%)	7 (6%)	1 (1%)	19	43
3	E	123/128 (96%)	110 (89%)	11 (9%)	2 (2%)	9	24
All	All	389/645 (60%)	358 (92%)	23 (6%)	8 (2%)	7	18

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	293	GLN
1	A	294	LYS
3	E	113	GLN
2	C	115	GLN
1	A	340	HIS
1	A	345	THR
1	A	358	ASP
3	E	125	THR

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	138/338 (41%)	126 (91%)	12 (9%)	10	23
2	C	98/97 (101%)	91 (93%)	7 (7%)	14	34
3	E	100/102 (98%)	84 (84%)	16 (16%)	2	6
All	All	336/537 (63%)	301 (90%)	35 (10%)	7	16

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	231	GLN
1	A	244	GLN
1	A	252	LYS
1	A	257	PHE
1	A	259	GLN
1	A	262	VAL
1	A	301	GLN
1	A	311	GLN
1	A	330	LEU
1	A	336	ARG
1	A	349	LEU
1	A	356	VAL
2	C	5	GLN
2	C	20	LEU
2	C	30	THR
2	C	31	SER
2	C	47	ARG
2	C	49	TYR
2	C	74	SER
3	E	3	GLU
3	E	15	GLN
3	E	23	SER
3	E	27	SER
3	E	32	SER
3	E	45	LYS
3	E	59	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	88	LEU
3	E	104	MET
3	E	106	ARG
3	E	111	LYS
3	E	114	ASN
3	E	122	THR
3	E	123	GLN
3	E	125	THR
3	E	127	SER

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

Mol	Chain	Res	Type
1	A	351	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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.