Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2022 – 01:14 pm BST

PDB ID	:	7R23
Title	:	Crystal structure of human Arc CTD in complex with two anti-Arc nanobodies
Authors	:	Markusson, S.; Kursula, P.
Deposited on	:	2022-02-04
Resolution	:	2.77 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*
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:

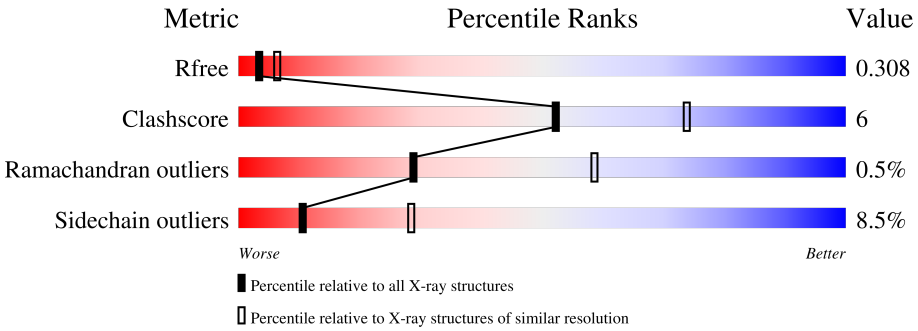
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.77 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	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)

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	181	61% 18% • 20%
2	B	128	78% 17% • •
3	C	120	82% 16% ••

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6159 atoms, of which 3033 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.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	145	2459	799	1217	213	228	2	0	1	0

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	181	MET	-	initiating methionine	UNP A0A3L7I2Q1
A	182	HIS	-	expression tag	UNP A0A3L7I2Q1
A	183	HIS	-	expression tag	UNP A0A3L7I2Q1
A	184	HIS	-	expression tag	UNP A0A3L7I2Q1
A	185	HIS	-	expression tag	UNP A0A3L7I2Q1
A	186	HIS	-	expression tag	UNP A0A3L7I2Q1
A	187	HIS	-	expression tag	UNP A0A3L7I2Q1
A	188	LEU	-	expression tag	UNP A0A3L7I2Q1
A	189	GLU	-	expression tag	UNP A0A3L7I2Q1
A	190	SER	-	expression tag	UNP A0A3L7I2Q1
A	191	THR	-	expression tag	UNP A0A3L7I2Q1
A	192	SER	-	expression tag	UNP A0A3L7I2Q1
A	193	LEU	-	expression tag	UNP A0A3L7I2Q1
A	194	TYR	-	expression tag	UNP A0A3L7I2Q1
A	195	LYS	-	expression tag	UNP A0A3L7I2Q1
A	196	LYS	-	expression tag	UNP A0A3L7I2Q1
A	197	ALA	-	expression tag	UNP A0A3L7I2Q1
A	198	GLY	-	expression tag	UNP A0A3L7I2Q1
A	199	SER	-	expression tag	UNP A0A3L7I2Q1
A	200	GLU	-	expression tag	UNP A0A3L7I2Q1
A	201	ASN	-	expression tag	UNP A0A3L7I2Q1
A	202	LEU	-	expression tag	UNP A0A3L7I2Q1
A	203	TYR	-	expression tag	UNP A0A3L7I2Q1
A	204	PHE	-	expression tag	UNP A0A3L7I2Q1
A	205	GLN	-	expression tag	UNP A0A3L7I2Q1
A	317	ASP	GLU	conflict	UNP A0A3L7I2Q1
A	358	ASP	GLY	conflict	UNP A0A3L7I2Q1

- Molecule 2 is a protein called Chains: B.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	124	Total	C	H	N	O	S	0	1	0
			1888	605	924	166	188	5			

- Molecule 3 is a protein called Chains: C.

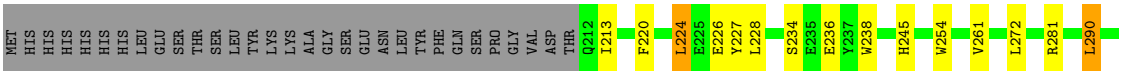
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	119	Total	C	H	N	O	S	0	0	0
			1812	573	892	163	181	3			

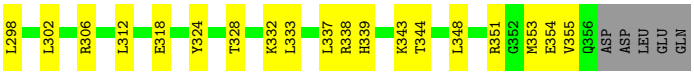
3 Residue-property plots (i)

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. 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: ARC

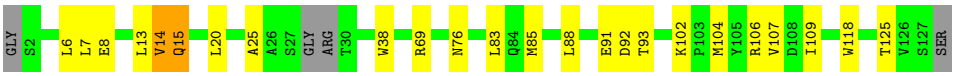
Chain A:





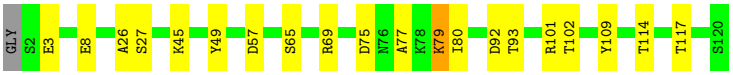
- Molecule 2: Chains: B

Chain B:



- Molecule 3: Chains: C

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	40.80Å 61.64Å 171.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.99 – 2.77 42.99 – 2.77	Depositor EDS
% Data completeness (in resolution range)	98.5 (42.99-2.77) 85.0 (42.99-2.77)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.27 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.252 , 0.302 0.256 , 0.308	Depositor DCC
R_{free} test set	1163 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	50.4	Xtriage
Anisotropy	0.626	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	6159	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.65% 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	0.24	0/1285	0.43	0/1732
2	B	0.26	0/989	0.50	0/1336
3	C	0.25	0/938	0.51	0/1271
All	All	0.25	0/3212	0.47	0/4339

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

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	1242	1217	1203	15	0
2	B	964	924	917	14	0
3	C	920	892	892	9	0
All	All	3126	3033	3012	34	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 6.

All (34) 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:281:ARG:NH1	1:A:312:LEU:O	2.14	0.81

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:93:THR:HG23	2:B:125:THR:HA	1.73	0.71
1:A:324:TYR:O	1:A:328:THR:HG23	1.96	0.66
3:C:69:ARG:NH2	3:C:92:ASP:OD2	2.31	0.63
2:B:8:GLU:N	2:B:8:GLU:OE1	2.33	0.60
2:B:14:VAL:HG22	2:B:15:GLN:H	1.66	0.60
1:A:338:ARG:NH2	3:C:57:ASP:OD2	2.34	0.59
3:C:93:THR:HG23	3:C:117:THR:HA	1.86	0.58
3:C:8:GLU:HB2	3:C:114:THR:HG23	1.89	0.55
1:A:228:LEU:HD23	1:A:238:TRP:HB3	1.90	0.54
2:B:14:VAL:HG11	2:B:88:LEU:HD13	1.89	0.54
2:B:104:MET:HB3	2:B:107:VAL:HG11	1.90	0.54
2:B:7:LEU:HB3	2:B:25:ALA:HB3	1.89	0.53
3:C:75:ASP:O	3:C:77:ALA:N	2.38	0.53
3:C:102:THR:HG21	3:C:109:TYR:CE2	2.44	0.52
3:C:8:GLU:N	3:C:8:GLU:OE1	2.43	0.51
1:A:213:ILE:HG21	2:B:109:ILE:HB	1.92	0.50
1:A:254:TRP:CH2	1:A:261:VAL:HG21	2.46	0.50
3:C:102:THR:HG21	3:C:109:TYR:HE2	1.77	0.50
3:C:26:ALA:HB3	3:C:79:LYS:HA	1.95	0.48
2:B:14:VAL:HG13	2:B:15:GLN:N	2.28	0.48
1:A:298:LEU:HD21	1:A:302:LEU:HD12	1.96	0.47
2:B:6:LEU:HD12	2:B:118:TRP:C	2.36	0.46
2:B:69:ARG:NH2	2:B:92:ASP:OD2	2.49	0.45
1:A:337:LEU:HD13	1:A:348:LEU:HG	1.99	0.45
1:A:298:LEU:HD23	1:A:298:LEU:C	2.37	0.45
2:B:38:TRP:CG	2:B:83:LEU:HD22	2.52	0.45
1:A:298:LEU:HD22	1:A:353:MET:CG	2.48	0.44
1:A:227:TYR:OH	2:B:102:LYS:NZ	2.52	0.43
1:A:245:HIS:O	2:B:104:MET:HA	2.19	0.42
1:A:337:LEU:HD21	1:A:351:ARG:HB3	2.02	0.42
2:B:13:LEU:HD12	2:B:125:THR:O	2.20	0.41
1:A:220:PHE:CZ	1:A:224:LEU:HD12	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	144/181 (80%)	138 (96%)	5 (4%)	1 (1%)	22 50
2	B	121/128 (94%)	110 (91%)	10 (8%)	1 (1%)	19 47
3	C	117/120 (98%)	109 (93%)	8 (7%)	0	100 100
All	All	382/429 (89%)	357 (94%)	23 (6%)	2 (0%)	29 58

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	355	VAL
2	B	14	VAL

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	135/167 (81%)	121 (90%)	14 (10%)	7 19
2	B	100/102 (98%)	94 (94%)	6 (6%)	19 45
3	C	97/97 (100%)	89 (92%)	8 (8%)	11 30
All	All	332/366 (91%)	304 (92%)	28 (8%)	10 29

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	224	LEU
1	A	226	GLU
1	A	234	SER
1	A	236	GLU
1	A	272	LEU
1	A	290	LEU
1	A	306	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	318	GLU
1	A	332	LYS
1	A	333	LEU
1	A	339	HIS
1	A	343	LYS
1	A	344	THR
1	A	354	GLU
2	B	15	GLN
2	B	20	LEU
2	B	76	ASN
2	B	85	MET
2	B	91	GLU
2	B	106	ARG
3	C	3	GLU
3	C	27	SER
3	C	45	LYS
3	C	49	TYR
3	C	65	SER
3	C	79	LYS
3	C	80	ILE
3	C	101	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	A	243	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 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.