WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Aug 22, 2022 – 07:06 pm BST

|		
PDB ID	:	7ZK1
Title	:	Crystal structure of cystinosin from Arabidopsis thaliana bound to sybody and nanobody
Authors	:	Loebel, M.; Newstead, S.; Omari, K.E.
Deposited on	:	2022-04-12
Resolution	:	2.65 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 2.65 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		0.303
Clashscore		10
Ramachandran outliers		2.5%
Sidechain outliers		12.7%

Worse   Better

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
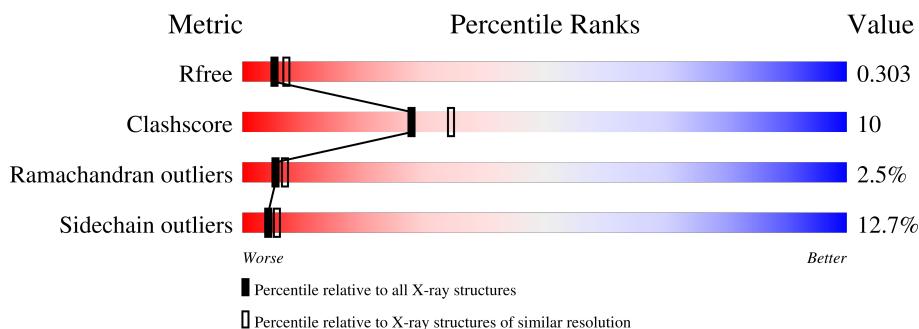
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)

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	277	62% 21% • 12%
2	B	128	71% 23% • ••
3	C	121	71% 24% • •

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 3957 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 Cystinosin homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	243	Total	C	N	O	S	0	0	0
			1987	1330	310	335	12			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	271	GLY	-	expression tag	UNP P57758
A	272	GLU	-	expression tag	UNP P57758
A	273	ASN	-	expression tag	UNP P57758
A	274	LEU	-	expression tag	UNP P57758
A	275	TYR	-	expression tag	UNP P57758
A	276	PHE	-	expression tag	UNP P57758
A	277	GLN	-	expression tag	UNP P57758

- Molecule 2 is a protein called Llama derived nanobody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	127	Total	C	N	O	S	0	1	0
			976	609	175	188	4			

- Molecule 3 is a protein called Synthetic nanobody (Sybody).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	121	Total	C	N	O	S	0	0	0
			962	604	170	184	4			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	14	Total	O	0	0
			14	14		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	11	Total 11	O 11	0	0
4	C	7	Total 7	O 7	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: Cystinosin homolog

Chain A:
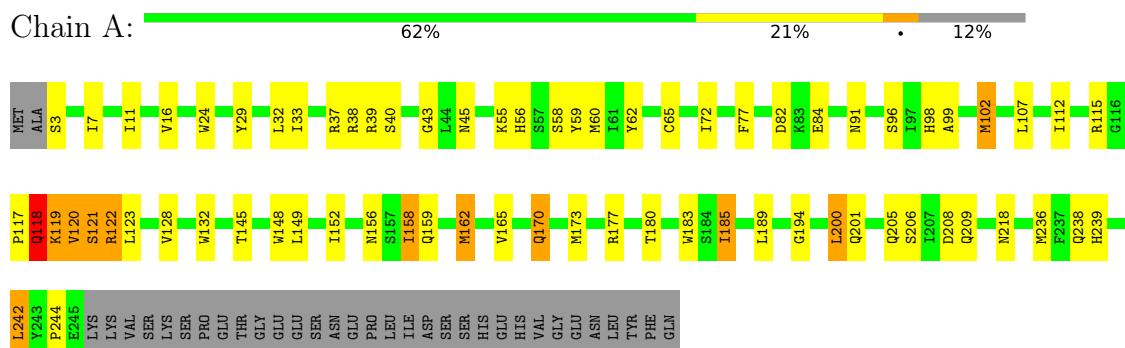
• Molecule 2: Llama derived nanobody

Chain B:
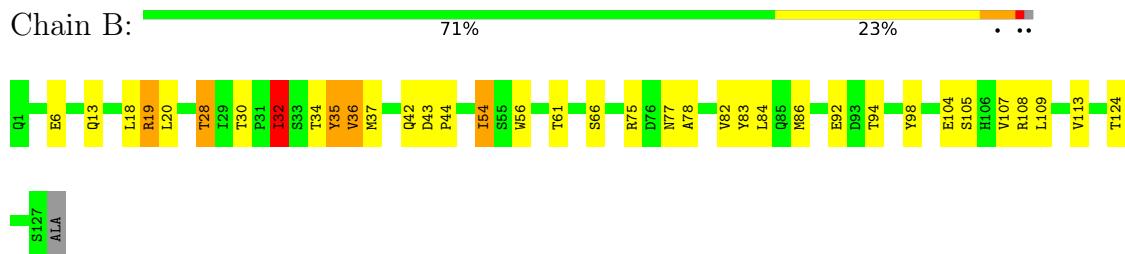
• Molecule 3: Synthetic nanobody (Sybody)

Chain C:
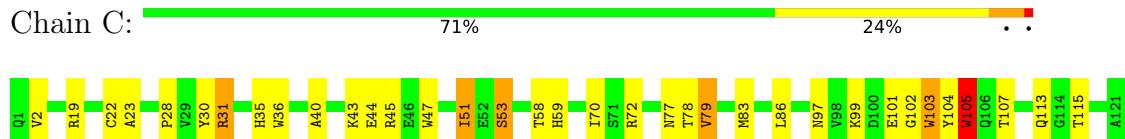
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	62.79Å 319.96Å 45.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	79.99 – 2.65 79.99 – 2.65	Depositor EDS
% Data completeness (in resolution range)	93.4 (79.99-2.65) 93.4 (79.99-2.65)	Depositor EDS
R_{merge}	0.54	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.38 (at 2.65Å)	Xtriage
Refinement program	BUSTER	Depositor
R, R_{free}	0.230 , 0.270 0.257 , 0.303	Depositor DCC
R_{free} test set	1331 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	66.3	Xtriage
Anisotropy	0.564	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	3957	wwPDB-VP
Average B, all atoms (Å^2)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.17% 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	0.44	0/2049	0.64	0/2789
2	B	0.45	0/996	0.66	0/1351
3	C	0.42	0/987	0.69	0/1340
All	All	0.44	0/4032	0.66	0/5480

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	1987	0	1986	43	0
2	B	976	0	949	19	0
3	C	962	0	904	21	0
4	A	14	0	0	1	0
4	B	11	0	0	1	0
4	C	7	0	0	0	0
All	All	3957	0	3839	77	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 10.

All (77) 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:62:TYR:HA	1:A:152:ILE:HG12	1.59	0.84
1:A:65:CYS:HB2	1:A:152:ILE:HD11	1.63	0.79
3:C:36:TRP:HD1	3:C:70:ILE:HD12	1.48	0.79
1:A:183:TRP:NE1	1:A:185:ILE:HG13	1.99	0.77
2:B:54:ILE:HG23	2:B:61:THR:HG22	1.68	0.75
1:A:180:THR:HG21	1:A:239:HIS:ND1	2.06	0.69
1:A:118:GLN:HG2	1:A:120:VAL:HG22	1.74	0.68
1:A:183:TRP:HE1	1:A:185:ILE:HG13	1.59	0.67
3:C:22:CYS:HB3	3:C:79:VAL:HG12	1.77	0.67
1:A:208:ASP:OD2	1:A:209:GLN:NE2	2.28	0.66
3:C:36:TRP:HD1	3:C:70:ILE:CD1	2.09	0.65
1:A:32:LEU:HD11	1:A:102:MET:HG3	1.79	0.64
3:C:102:GLY:H	3:C:107:THR:HG22	1.62	0.64
2:B:28:THR:HG21	2:B:35:TYR:OH	1.97	0.63
2:B:32:ILE:O	2:B:35:TYR:HB2	1.98	0.63
1:A:65:CYS:HB2	1:A:152:ILE:CD1	2.29	0.63
2:B:37:MET:HB2	2:B:82:VAL:HG11	1.80	0.63
1:A:149:LEU:O	1:A:152:ILE:HG22	2.01	0.61
1:A:59:TYR:HB2	1:A:159:GLN:HG3	1.83	0.61
1:A:115:ARG:HA	1:A:115:ARG:HE	1.66	0.60
1:A:238:GLN:HA	1:A:242:LEU:HB2	1.86	0.58
1:A:7:ILE:HG23	2:B:107:VAL:HG22	1.87	0.57
1:A:183:TRP:HE1	1:A:185:ILE:CG1	2.18	0.56
1:A:117:PRO:HG2	3:C:47:TRP:CH2	2.40	0.56
2:B:19:ARG:NH1	2:B:83:TYR:CD2	2.73	0.56
1:A:43:GLY:O	3:C:107:THR:HG21	2.06	0.55
2:B:75:ARG:NH2	4:B:202:HOH:O	2.40	0.55
1:A:173:MET:O	1:A:177:ARG:HB2	2.07	0.54
4:A:314:HOH:O	3:C:105:TRP:HB2	2.07	0.54
2:B:105:SER:HB3	2:B:108:ARG:HB2	1.90	0.54
2:B:32:ILE:HA	2:B:35:TYR:CD1	2.44	0.53
1:A:55:LYS:HB2	1:A:162:MET:HB3	1.91	0.53
1:A:29:TYR:O	1:A:33:ILE:HG12	2.10	0.52
1:A:60:MET:HA	1:A:96:SER:HB3	1.92	0.52
3:C:23:ALA:HA	3:C:78:THR:HG22	1.92	0.51
1:A:206:SER:O	1:A:209:GLN:O	2.29	0.51
2:B:75:ARG:HD3	2:B:77:ASN:OD1	2.11	0.51
1:A:40:SER:HB2	1:A:115:ARG:HB3	1.93	0.51
2:B:94:THR:HG23	2:B:124:THR:HA	1.93	0.51
1:A:77:PHE:HE1	1:A:84:GLU:O	1.94	0.51
1:A:183:TRP:HH2	1:A:236:MET:HG2	1.77	0.50
2:B:113:VAL:HG12	2:B:113:VAL:O	2.11	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:29:TYR:OH	1:A:98:HIS:HE1	1.96	0.49
2:B:35:TYR:O	2:B:36:VAL:O	2.29	0.49
3:C:35:HIS:HB2	3:C:97:ASN:HB3	1.95	0.49
1:A:91:ASN:HD21	1:A:201:GLN:HE22	1.61	0.48
1:A:183:TRP:CD1	1:A:239:HIS:CE1	3.00	0.48
3:C:30:TYR:O	3:C:53:SER:O	2.31	0.48
1:A:128:VAL:HG13	1:A:132:TRP:CD1	2.49	0.48
3:C:51:ILE:HD12	3:C:58:THR:HG22	1.95	0.48
2:B:37:MET:CB	2:B:82:VAL:HG11	2.43	0.47
3:C:40:ALA:HB3	3:C:43:LYS:HD3	1.96	0.47
2:B:34:THR:HA	2:B:56:TRP:HB2	1.96	0.47
3:C:105:TRP:C	3:C:105:TRP:CD1	2.89	0.47
1:A:148:TRP:O	1:A:152:ILE:HB	2.16	0.46
3:C:28:PRO:HA	3:C:77:ASN:HD21	1.79	0.46
1:A:119:LYS:C	1:A:121:SER:H	2.19	0.45
1:A:58:SER:OG	1:A:158:ILE:HD11	2.16	0.45
1:A:16:VAL:HG11	1:A:200:LEU:HB3	1.99	0.44
2:B:84:LEU:HG	2:B:86:MET:HG2	1.99	0.44
1:A:56:HIS:CE1	1:A:99:ALA:HB1	2.53	0.44
1:A:128:VAL:HG13	1:A:132:TRP:HD1	1.82	0.44
1:A:180:THR:CG2	1:A:239:HIS:ND1	2.79	0.44
1:A:117:PRO:HB3	3:C:59:HIS:CG	2.53	0.43
1:A:117:PRO:O	1:A:119:LYS:N	2.51	0.43
2:B:42:GLN:HB2	2:B:98:TYR:HE1	1.83	0.43
1:A:72:ILE:HG12	1:A:148:TRP:HB2	2.01	0.43
3:C:83:MET:HB2	3:C:86:LEU:HD21	1.99	0.43
3:C:28:PRO:HG2	3:C:31:ARG:HB2	1.99	0.43
1:A:24:TRP:CD1	1:A:194:GLY:HA3	2.54	0.42
1:A:170:GLN:OE1	3:C:105:TRP:CD2	2.73	0.42
2:B:18:LEU:HB3	2:B:86:MET:CE	2.50	0.42
1:A:40:SER:HA	1:A:115:ARG:HB2	2.03	0.41
2:B:20:LEU:HD12	2:B:84:LEU:HD23	2.01	0.41
3:C:36:TRP:CD1	3:C:70:ILE:CD1	2.98	0.41
1:A:170:GLN:HG2	3:C:103:TRP:HB3	2.03	0.40
3:C:105:TRP:C	3:C:105:TRP:HD1	2.25	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	241/277 (87%)	224 (93%)	12 (5%)	5 (2%)	7	10
2	B	126/128 (98%)	111 (88%)	10 (8%)	5 (4%)	3	3
3	C	119/121 (98%)	111 (93%)	6 (5%)	2 (2%)	9	13
All	All	486/526 (92%)	446 (92%)	28 (6%)	12 (2%)	5	7

All (12) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	242	LEU
2	B	78	ALA
3	C	103	TRP
1	A	118	GLN
1	A	244	PRO
2	B	36	VAL
1	A	121	SER
3	C	105	TRP
1	A	122	ARG
2	B	32	ILE
2	B	43	ASP
2	B	44	PRO

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	218/249 (88%)	192 (88%)	26 (12%)	5	7
2	B	102/102 (100%)	90 (88%)	12 (12%)	5	7
3	C	98/98 (100%)	83 (85%)	15 (15%)	2	3
All	All	418/449 (93%)	365 (87%)	53 (13%)	4	6

All (53) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	SER
1	A	11	ILE
1	A	37	ARG
1	A	38	ARG
1	A	39	ARG
1	A	45	ASN
1	A	82	ASP
1	A	102	MET
1	A	107	LEU
1	A	112	ILE
1	A	118	GLN
1	A	119	LYS
1	A	120	VAL
1	A	122	ARG
1	A	123	LEU
1	A	145	THR
1	A	156	ASN
1	A	158	ILE
1	A	162	MET
1	A	165	VAL
1	A	170	GLN
1	A	185	ILE
1	A	189	LEU
1	A	200	LEU
1	A	205	GLN
1	A	218	ASN
2	B	6	GLU
2	B	13	GLN
2	B	19	ARG
2	B	28	THR
2	B	30	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	32	ILE
2	B	35	TYR
2	B	54	ILE
2	B	66	SER
2	B	92	GLU
2	B	104	GLU
2	B	109	LEU
3	C	2	VAL
3	C	19	ARG
3	C	31	ARG
3	C	44	GLU
3	C	45	ARG
3	C	51	ILE
3	C	53	SER
3	C	72	ARG
3	C	79	VAL
3	C	99	LYS
3	C	101	GLU
3	C	104	TYR
3	C	105	TRP
3	C	113	GLN
3	C	115	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	ASN
1	A	75	GLN
1	A	91	ASN
1	A	98	HIS
1	A	156	ASN
1	A	159	GLN
1	A	170	GLN
1	A	198	ASN
1	A	218	ASN
1	A	238	GLN
3	C	74	ASN
3	C	77	ASN
3	C	97	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 ⓘ

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.