WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 25, 2024 – 06:00 PM EST

PDB ID	:	3DRX
Title	:	X-ray crystal structure of human KCTD5 protein crystallized in high-salt buffer
Authors	:	Tereshko, V.; Dementieva, I.; Goldstein, S.A.N.
Deposited on	:	2008-07-11
Resolution	:	3.11 Å(reported)

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (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 3.11 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.
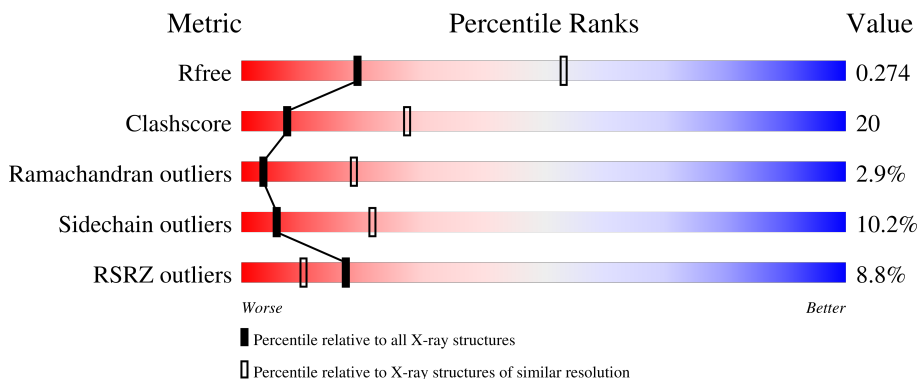
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1668 (3.14-3.10)
Clashscore	180529	1788 (3.14-3.10)
Ramachandran outliers	177936	1696 (3.14-3.10)
Sidechain outliers	177891	1696 (3.14-3.10)
RSRZ outliers	164620	1668 (3.14-3.10)

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	202	
1	B	202	
1	C	202	
1	D	202	
1	E	202	

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 6907 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 BTB/POZ domain-containing protein KCTD5.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	161	Total	C	N	O	S	0	0	0
			1306	831	223	246	6			
1	B	172	Total	C	N	O	S	0	0	0
			1385	874	236	269	6			
1	C	177	Total	C	N	O	S	0	0	0
			1430	902	243	279	6			
1	D	177	Total	C	N	O	S	0	0	0
			1427	901	243	277	6			
1	E	169	Total	C	N	O	S	0	0	0
			1359	860	232	261	6			

There are 5 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	33	SER	-	expression tag	UNP Q9NXV2
B	33	SER	-	expression tag	UNP Q9NXV2
C	33	SER	-	expression tag	UNP Q9NXV2
D	33	SER	-	expression tag	UNP Q9NXV2
E	33	SER	-	expression tag	UNP Q9NXV2

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.
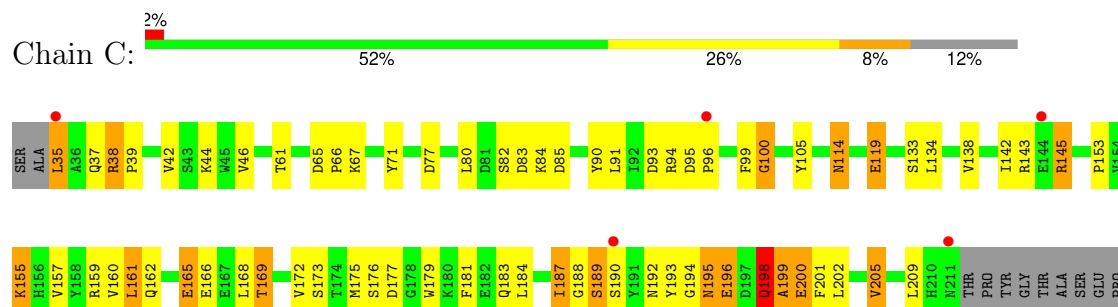
• Molecule 1: BTB/POZ domain-containing protein KCTD5
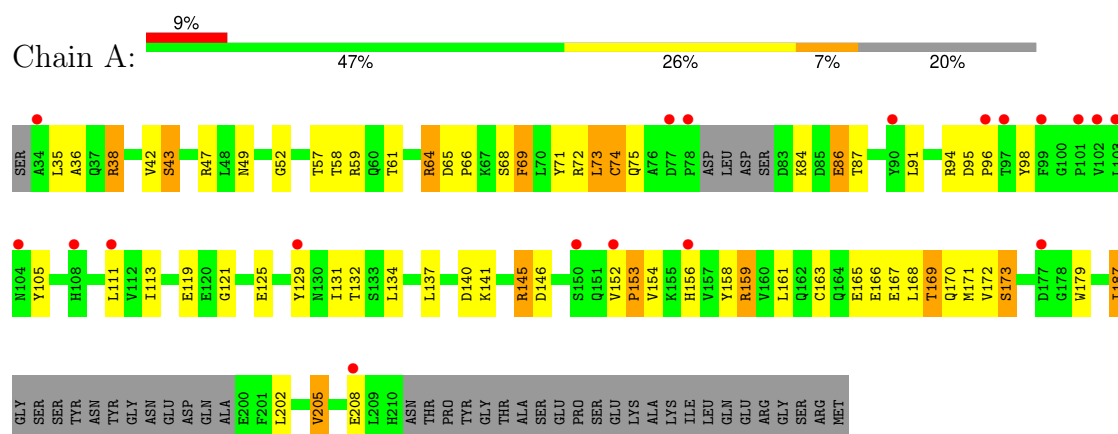
• Molecule 1: BTB/POZ domain-containing protein KCTD5
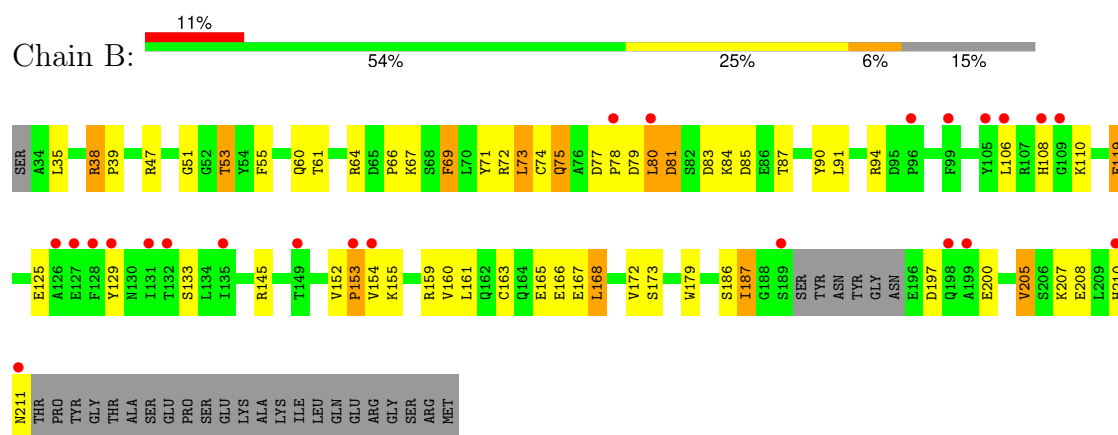
• Molecule 1: BTB/POZ domain-containing protein KCTD5

SER GLU LYS ALA LYS ILE LEU GLN GLU ARG GLY SER ARG MET

• Molecule 1: BTB/POZ domain-containing protein KCTD5

Chain D:

GLY SER ARG MET

• Molecule 1: BTB/POZ domain-containing protein KCTD5

Chain E:

PRO TYR GLY THR ALA SER GLU PRO SER GLU LYS ALA LYS ILE LEU GLU GLN ARG GLY SER ARG MET

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.86Å 128.58Å 152.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.11 20.00 – 3.11	Depositor EDS
% Data completeness (in resolution range)	98.9 (20.00-3.11) 98.5 (20.00-3.11)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 3.09Å)	Xtriage
Refinement program	REFMAC 5.4.0073	Depositor
R, R_{free}	0.227 , 0.275 0.225 , 0.274	Depositor DCC
R_{free} test set	1321 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	62.6	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 66.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	6907	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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 ⓘ

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	0.39	0/1329	0.52	0/1793
1	B	0.44	0/1409	0.53	0/1903
1	C	0.45	0/1457	0.61	1/1970 (0.1%)
1	D	0.42	0/1453	0.53	0/1963
1	E	0.43	0/1383	0.56	0/1868
All	All	0.42	0/7031	0.55	1/9497 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	143	ARG	NE-CZ-NH2	5.36	122.98	120.30

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	1306	0	1309	62	0
1	B	1385	0	1371	46	0
1	C	1430	0	1405	67	0
1	D	1427	0	1405	46	0
1	E	1359	0	1353	65	0
All	All	6907	0	6843	274	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 20.

The worst 5 of 274 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:E:152:VAL:CG1	1:E:153:PRO:HD2	1.77	1.13
1:E:38:ARG:HG2	1:E:38:ARG:HH11	1.09	1.07
1:E:152:VAL:HG12	1:E:153:PRO:HD2	1.37	1.03
1:E:165:GLU:O	1:E:168:LEU:HB2	1.69	0.92
1:A:72:ARG:HG3	1:A:73:LEU:HD13	1.49	0.91

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	155/202 (77%)	128 (83%)	22 (14%)	5 (3%)	3	17
1	B	168/202 (83%)	152 (90%)	12 (7%)	4 (2%)	5	21
1	C	175/202 (87%)	146 (83%)	23 (13%)	6 (3%)	3	15
1	D	173/202 (86%)	148 (86%)	20 (12%)	5 (3%)	3	18
1	E	165/202 (82%)	144 (87%)	17 (10%)	4 (2%)	5	21
All	All	836/1010 (83%)	718 (86%)	94 (11%)	24 (3%)	3	18

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	SER
1	A	74	CYS
1	A	153	PRO
1	B	81	ASP
1	C	198	GLN

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	147/181 (81%)	129 (88%)	18 (12%)	4	16
1	B	156/181 (86%)	134 (86%)	22 (14%)	3	12
1	C	161/181 (89%)	141 (88%)	20 (12%)	4	15
1	D	160/181 (88%)	154 (96%)	6 (4%)	28	57
1	E	153/181 (84%)	140 (92%)	13 (8%)	8	30
All	All	777/905 (86%)	698 (90%)	79 (10%)	6	22

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	198	GLN
1	E	82	SER
1	C	205	VAL
1	D	173	SER
1	E	161	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	151	GLN
1	D	192	ASN
1	C	195	ASN
1	C	164	GLN
1	D	156	HIS

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 ⓘ

> **Warning**: The R factor obtained from EDS is 0.2786, which does not match the depositor's R factor of 0.227. Please interpret the results in this section carefully.

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(Å²)	Q<0.9
1	A	161/202 (79%)	0.91	19 (11%)	10	6	40, 66, 82, 85	2 (1%)
1	B	172/202 (85%)	0.83	23 (13%)	8	5	33, 65, 85, 96	1 (0%)
1	C	177/202 (87%)	0.50	5 (2%)	55	36	52, 66, 83, 85	0
1	D	177/202 (87%)	0.68	19 (10%)	12	8	53, 68, 84, 106	0
1	E	169/202 (83%)	0.48	9 (5%)	33	20	42, 64, 79, 84	0
All	All	856/1010 (84%)	0.67	75 (8%)	17	10	33, 66, 83, 106	3 (0%)

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	78	PRO	12.9
1	B	189	SER	9.5
1	A	77	ASP	8.3
1	D	190	SER	5.5
1	D	191	TYR	4.3

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.