



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:14 PM GMT

PDB ID : 4BY2
Title : SAS-4 (dCPAP) TCP domain in complex with a Proline Rich Motif of Ana2 (dSTIL) of Drosophila Melanogaster
Authors : Cottee, M.A.; Muschalik, N.; Wong, Y.L.; Johnson, C.M.; Johnson, S.; Andreeva, A.; Oegema, K.; Lea, S.M.; Raff, J.W.; vanBreugel, M.
Deposited on : 2013-07-17
Resolution : 2.57 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

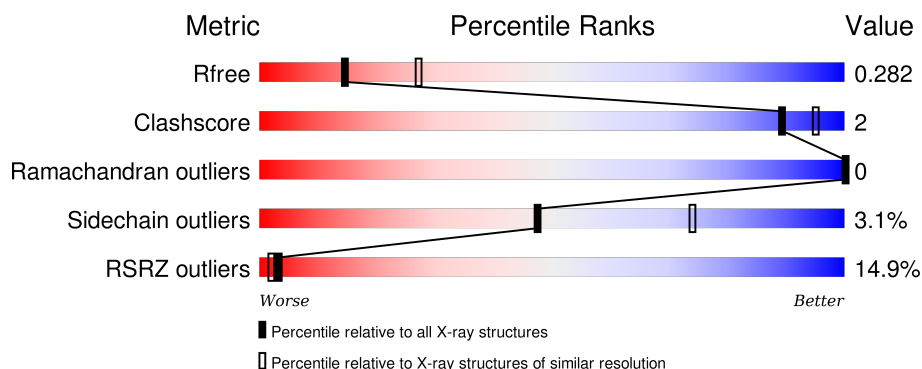
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.57 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}	91344	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	252	<div> <div>10%</div> <div> <div></div> <div>54%</div> <div>7%</div> <div>39%</div> </div> </div>
1	B	252	<div> <div>4%</div> <div> <div></div> <div>57%</div> <div>5%</div> <div>38%</div> </div> </div>
1	C	252	<div> <div>13%</div> <div> <div></div> <div>54%</div> <div>8%</div> <div>38%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3859 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 ANASTRAL SPINDLE 2, SAS 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total	C	N	O	S	0	0	0
			1258	792	223	240	3			
1	B	155	Total	C	N	O	S	0	0	0
			1265	797	224	241	3			
1	C	155	Total	C	N	O	S	0	0	0
			1267	797	224	243	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP Q9XZ31
A	2	PRO	-	EXPRESSION TAG	UNP Q9XZ31
A	3	MET	-	EXPRESSION TAG	UNP Q9XZ31
A	4	GLY	-	EXPRESSION TAG	UNP Q9XZ31
B	1	GLY	-	EXPRESSION TAG	UNP Q9XZ31
B	2	PRO	-	EXPRESSION TAG	UNP Q9XZ31
B	3	MET	-	EXPRESSION TAG	UNP Q9XZ31
B	4	GLY	-	EXPRESSION TAG	UNP Q9XZ31
C	1	GLY	-	EXPRESSION TAG	UNP Q9XZ31
C	2	PRO	-	EXPRESSION TAG	UNP Q9XZ31
C	3	MET	-	EXPRESSION TAG	UNP Q9XZ31
C	4	GLY	-	EXPRESSION TAG	UNP Q9XZ31

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	20	Total	O	0	0
			20	20		
3	B	28	Total	O	0	0
			28	28		
3	C	17	Total	O	0	0
			17	17		

T203	
K204	
LEU	ASN
LYS	ARG
ARG	ARG
GLU	TYR
PRO	ASP
GLY	THR
THR	VAL
LYS	LYS
LEU	VAL
TYR	TYR
PRO	ASP
ASP	GLY
GLY	SER
SER	GLN
GLU	GLU
THR	THR
ARG	ARG
TYR	TYR
SER	SER
ASN	ASN
GLY	GLY
ARG	ARG
VAL	VAL
ARG	ARG
LEU	LEU
LYS	LYS
ASP	ASP
LYS	LYS
ASP	ASP
GLY	GLY
LYS	LYS
LEU	LEU
ILE	ILE
MET	MET
ASP	ASP
THR	THR
ASP	ASP
TYR	TYR
ALA	ALA
LYS	LYS
TYR	TYR

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	58.64Å 69.91Å 69.98Å 86.96° 88.64° 67.69°	Depositor
Resolution (Å)	64.60 – 2.57 64.60 – 2.57	Depositor EDS
% Data completeness (in resolution range)	97.6 (64.60-2.57) 96.1 (64.60-2.57)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.95 (at 2.58Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, R_{free}	0.245 , 0.263 0.260 , 0.282	Depositor DCC
R_{free} test set	1611 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	53.7	Xtriage
Anisotropy	0.791	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 46.4	EDS
Estimated twinning fraction	0.027 for h,h-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 31909 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3859	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.14% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EDO

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	0.38	0/1286	0.57	0/1738
1	B	0.35	0/1294	0.54	0/1750
1	C	0.37	0/1295	0.55	0/1750
All	All	0.37	0/3875	0.56	0/5238

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	1258	0	1230	7	0
1	B	1265	0	1237	5	0
1	C	1267	0	1236	7	0
2	B	4	0	6	0	0
3	A	20	0	0	0	0
3	B	28	0	0	0	0
3	C	17	0	0	0	0
All	All	3859	0	3709	18	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 2.

All (18) 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:68:GLU:HG3	1:A:78:ILE:HG12	1.82	0.60
1:C:192:LEU:HB3	1:C:200:GLU:HB2	1.84	0.60
1:C:68:GLU:HG3	1:C:78:ILE:HG12	1.84	0.58
1:B:143:GLU:HG2	1:B:153:ILE:HG12	1.90	0.53
1:B:67:ARG:HB2	1:B:79:TRP:HB2	1.92	0.52
1:C:180:THR:HB	1:C:195:PRO:HD3	1.93	0.50
1:A:151:VAL:HB	1:A:163:VAL:HG13	1.93	0.49
1:A:170:LYS:HA	1:A:185:LEU:HA	1.97	0.47
1:C:144:HIS:HB2	1:C:152:GLU:HB3	1.97	0.46
1:B:190:LYS:HB2	1:B:202:HIS:HB2	1.98	0.45
1:C:181:HIS:HB3	1:C:193:ASN:HB2	1.98	0.44
1:A:69:ILE:HG23	1:A:77:ASP:HB2	2.01	0.42
1:A:135:LEU:O	1:A:142:THR:HA	2.20	0.42
1:B:138:PRO:HA	1:C:16:LEU:HB2	2.00	0.42
1:A:192:LEU:HB3	1:A:200:GLU:HB2	2.03	0.41
1:A:190:LYS:HB2	1:A:202:HIS:HB2	2.01	0.41
1:C:69:ILE:HG23	1:C:77:ASP:HB2	2.02	0.41
1:B:162:ILE:HB	1:B:173:GLU:HB2	2.02	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	150/252 (60%)	147 (98%)	3 (2%)	0	100	100
1	B	151/252 (60%)	150 (99%)	1 (1%)	0	100	100
1	C	151/252 (60%)	148 (98%)	3 (2%)	0	100	100
All	All	452/756 (60%)	445 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	137/221 (62%)	133 (97%)	4 (3%)	50	76
1	B	138/221 (62%)	135 (98%)	3 (2%)	60	82
1	C	138/221 (62%)	132 (96%)	6 (4%)	35	62
All	All	413/663 (62%)	400 (97%)	13 (3%)	47	74

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	91	ASP
1	A	129	LEU
1	A	168	THR
1	A	203	THR
1	B	12	MET
1	B	65	PHE
1	B	147	LYS
1	C	79	TRP
1	C	91	ASP
1	C	158	ASN
1	C	169	GLU
1	C	185	LEU
1	C	203	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	EDO	B	1205	-	3,3,3	0.55	0	2,2,2	0.45	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	B	1205	-	-	0/1/1/1	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	154/252 (61%)	1.11	24 (15%) 3 2	44, 72, 98, 121	0
1	B	155/252 (61%)	0.92	11 (7%) 19 15	40, 65, 88, 109	0
1	C	155/252 (61%)	1.25	34 (21%) 1 0	41, 71, 102, 127	0
All	All	464/756 (61%)	1.09	69 (14%) 3 2	40, 69, 99, 127	0

All (69) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	24	VAL	6.4
1	A	23	ALA	4.8
1	A	186	ARG	4.8
1	A	72	ALA	4.3
1	C	192	LEU	4.3
1	C	168	THR	4.3
1	C	72	ALA	4.1
1	A	191	ILE	4.1
1	C	177	ALA	3.6
1	B	191	ILE	3.3
1	C	186	ARG	3.3
1	A	69	ILE	3.2
1	C	182	LEU	3.2
1	C	73	ASP	3.2
1	C	151	VAL	3.2
1	C	166	SER	3.1
1	A	192	LEU	3.0
1	C	23	ALA	2.9
1	C	69	ILE	2.9
1	C	171	LEU	2.8
1	C	162	ILE	2.8
1	A	183	VAL	2.7
1	B	12	MET	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	104	ILE	2.7
1	A	174	TRP	2.7
1	A	182	LEU	2.6
1	A	162	ILE	2.5
1	A	126	THR	2.5
1	A	163	VAL	2.5
1	B	182	LEU	2.5
1	A	24	VAL	2.5
1	B	13	LEU	2.5
1	B	98	LEU	2.4
1	C	195	PRO	2.4
1	C	201	ILE	2.4
1	C	95	LEU	2.4
1	B	160	ILE	2.4
1	B	199	LYS	2.4
1	C	160	ILE	2.4
1	A	98	LEU	2.3
1	B	126	THR	2.3
1	C	191	ILE	2.3
1	A	177	ALA	2.3
1	C	98	LEU	2.3
1	C	123	THR	2.3
1	C	104	ILE	2.2
1	C	67	ARG	2.2
1	C	190	LYS	2.2
1	A	201	ILE	2.2
1	C	183	VAL	2.2
1	C	78	ILE	2.2
1	A	168	THR	2.2
1	A	194	LEU	2.1
1	C	167	ASP	2.1
1	C	173	GLU	2.1
1	C	194	LEU	2.1
1	A	153	ILE	2.1
1	B	78	ILE	2.1
1	A	104	ILE	2.1
1	A	198	GLN	2.1
1	C	184	GLN	2.1
1	A	12	MET	2.1
1	B	174	TRP	2.1
1	C	176	TYR	2.1
1	A	155	PHE	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	170	LYS	2.1
1	C	204	LYS	2.0
1	A	187	ASN	2.0
1	C	153	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	EDO	B	1205	4/4	0.94	0.23	0.50	59,59,59,59	0

6.5 Other polymers [i](#)

There are no such residues in this entry.