



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 06:16 am BST

PDB ID : 4ZUZ
Title : SidC 1-871
Authors : Luo, X.; Wasilko, D.J.; Liu, Y.; Sun, J.; Wu, X.; Luo, Z.-Q.; Mao, Y.
Deposited on : 2015-05-18
Resolution : 2.86 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

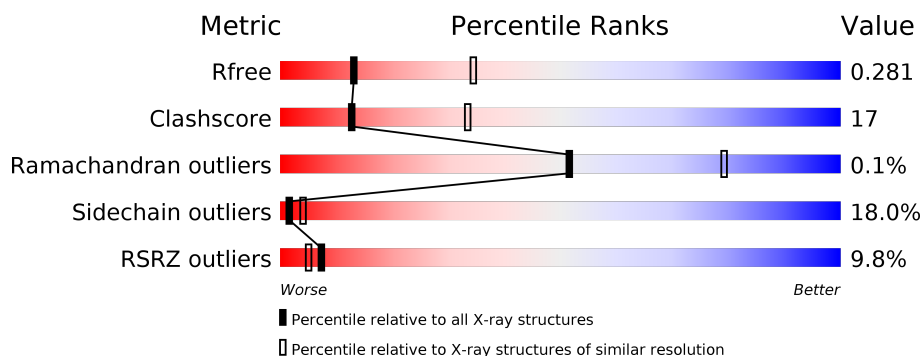
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.86 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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 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 $\leq 5\%$. 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	917	<div> <div>9%</div> <div>62%</div> <div>24%</div> <div>7%</div> <div>7%</div> </div>
1	B	917	<div> <div>9%</div> <div>65%</div> <div>22%</div> <div>7%</div> <div>6%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14094 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 SidC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	857	Total	C	N	O	S	0	0	0
			6979	4405	1190	1374	10			
1	B	858	Total	C	N	O	S	0	0	0
			6986	4410	1191	1375	10			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	325	ALA	VAL	conflict	UNP Q6RCR4
A	326	VAL	ALA	conflict	UNP Q6RCR4
A	334	GLN	ASP	conflict	UNP Q6RCR4
A	646	ALA	LYS	conflict	UNP Q6RCR4
A	731	TYR	ALA	conflict	UNP Q6RCR4
B	325	ALA	VAL	conflict	UNP Q6RCR4
B	326	VAL	ALA	conflict	UNP Q6RCR4
B	334	GLN	ASP	conflict	UNP Q6RCR4
B	646	ALA	LYS	conflict	UNP Q6RCR4
B	731	TYR	ALA	conflict	UNP Q6RCR4

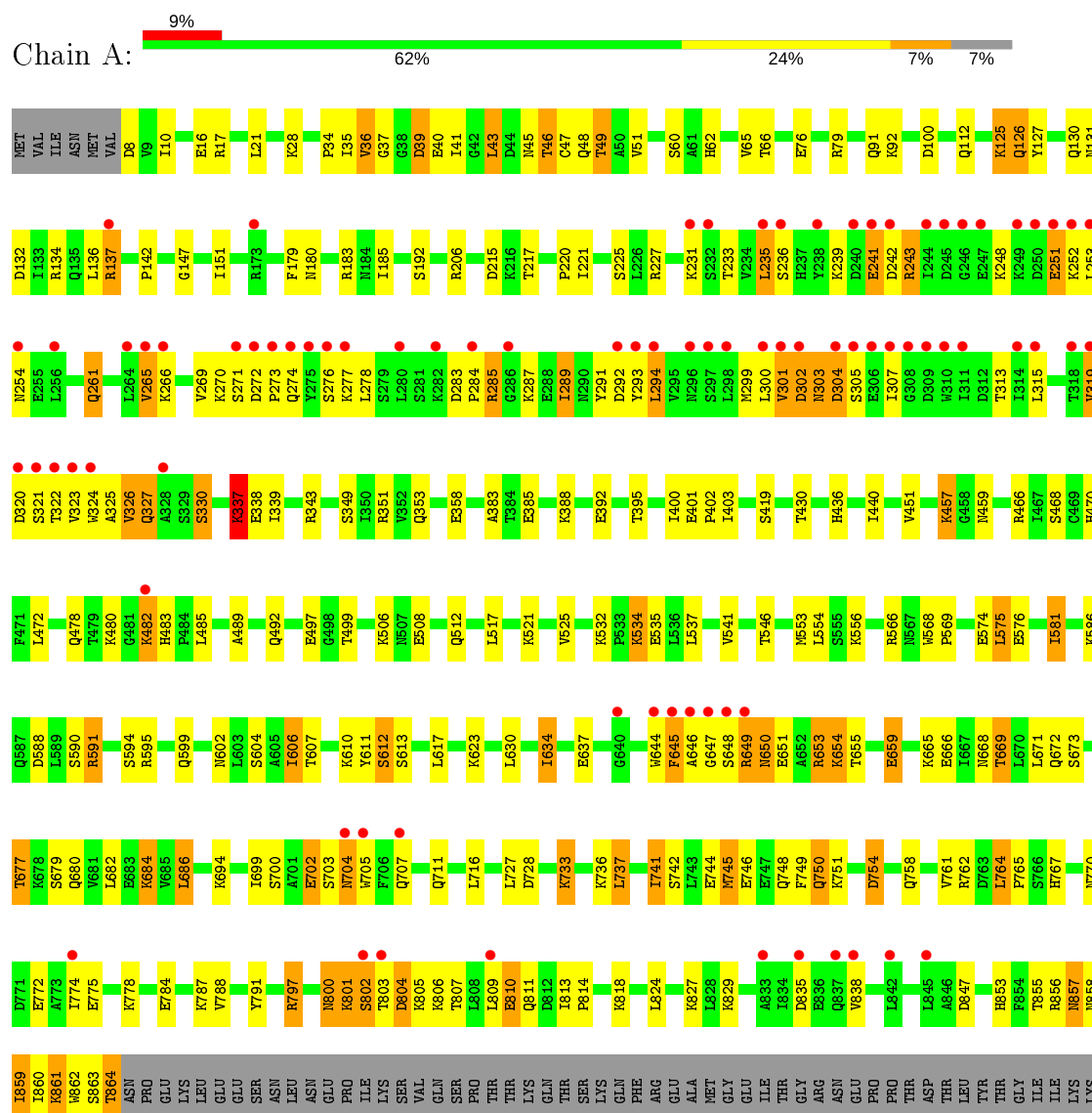
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	65	Total	O	0	0
			65	65		
2	B	64	Total	O	0	0
			64	64		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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: SidC



• Molecule 1: SidC



LEU	ASN	GLU	PRO	ILE	LYS	SER	VAL	GLN	SER	PRO	THR	THR	LYS	GLN	THR	SER	LYS	GLN	PHE	ARG	GLU	ALA	MET	GLY	GLU	ILE	THR	GLY	ARG	ASN	GLU	PRO	THR	ASP	THR	LEU	TYR	THR	GLY	ILE	ILE	LYS	LYS						
F776	F777	K778	E784	K787	V788	Y791	R797	N800	R801	S802	T803	D804	K805	K806	R807	L808	L809	E810	Q811	R812	I813	P814	K818	L824	K827	L828	K829	D835	K846	D847	T855	R856	I859	I860	K861	S863	T864	ASN	PRO	GLU	LYS	LEU	GLU	GLU	SER	ASN			
V681	L682	E683	K684	V685	L686	D693	K694	I699	E702	S703	N704	W705	F706	Q707	Q711	L716	F717	R718	L727	K733	K736	L737	I741	S742	N745	E746	E747	Q748	F749	Q750	K751	I752	Q753	D754	Q758	V761	R762	S763	L764	P765	S766	H767	E772	A773	I774	S775			
R595	D596	N597	I598	Q599	H600	D601	N602	L603	S604	A605	I606	T607	K610	Y611	S612	S613	L617	K623	I634	E637	G640	R641	E642	W643	W644	F645	A646	G647	S648	R649	N650	R653	K654	E659	R662	K665	E666	I667	N668	T669	L670	L671	Q672	S673	T677	K678	S679	Q680	
I460	R466	S321	T322	V323	C469	H470	F471	L472	K480	G481	K482	Q492	T499	E508	Q512	L517	K521	V525	K534	E535	L536	L537	D538	Y539	V541	T546	L554	R566	N567	W568	P569	E574	L575	E576	I581	K586	Q587	D588	L589	S590	R591	S594							
V319	D320	S321	T322	K323	C324	A325	V326	Q327	A328	A336	G337	E338	I339	D344	T348	S349	R351	V352	Q353	G374	A383	T384	E385	L386	A387	T395	I400	E401	P402	I403	I404	Y405	L415	S419	P420	L421	T430	T434	K435	H436	Y437	I440	K457	G458	N459				
D245	G246	E247	K248	R249	D250	E251	K252	L253	N254	E255	L256	L257	L261	L264	V265	K266	V269	K270	S271	D272	P273	Q274	S275	K276	K277	L278	D283	P284	R285	G286	K287	E288	L289	Y291	D292	Y293	L294	K299	L300	Y301	D302	N303	D304	S305	E306	W310	T313	I314	L315
Q121	K125	Q126	Y127	D128	E129	N130	N131	D132	I133	R134	Q135	L136	R137	P142	I151	R173	S192	G204	Y205	K206	L207	L211	F212	E214	D215	K216	T217	P218	T219	P220	I221	S225	L226	R227	V230	K231	S232	T233	V234	L235	S236	H237	Y238	E241	D242	R243	I244		
MET	VAL	ILE	ASN	MET	V7	D8	V9	I10	K13	E16	R17	K28	V29	P34	I36	V36	G37	G38	D39	E40	I41	G42	L43	D44	N45	T46	C47	Q48	T49	A50	V51	S60	A61	H62	S63	G64	V65	T66	E76	R79	K86	Q91	K92	D100	Q112	Y116	I117		

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	228.16Å 83.93Å 129.41Å 90.00° 108.82° 90.00°	Depositor
Resolution (Å)	49.16 – 2.86 47.14 – 2.86	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.16-2.86) 99.2 (47.14-2.86)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.52 (at 2.86Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.225 , 0.281 0.224 , 0.281	Depositor DCC
R_{free} test set	2708 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	73.5	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 60.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	14094	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.40% 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, 0.333 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	0.51	0/7110	0.66	1/9595 (0.0%)
1	B	0.47	0/7117	0.65	1/9605 (0.0%)
All	All	0.49	0/14227	0.66	2/19200 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	337	LYS	CD-CE-NZ	-5.67	98.67	111.70
1	A	337	LYS	CD-CE-NZ	-5.50	99.05	111.70

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	6979	0	6935	236	0
1	B	6986	0	6942	230	0
2	A	65	0	0	27	0
2	B	64	0	0	22	0
All	All	14094	0	13877	461	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 17.

The worst 5 of 461 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:B:703:SER:HB3	1:B:704:ASN:CA	1.46	1.42
1:A:703:SER:HB3	1:A:704:ASN:CA	1.48	1.41
1:B:703:SER:CB	1:B:704:ASN:HA	1.41	1.41
1:A:703:SER:CB	1:A:704:ASN:HA	1.41	1.40
1:B:323:VAL:C	1:B:326:VAL:HG21	1.42	1.36

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	855/917 (93%)	829 (97%)	25 (3%)	1 (0%)	51	79
1	B	856/917 (93%)	823 (96%)	33 (4%)	0	100	100
All	All	1711/1834 (93%)	1652 (97%)	58 (3%)	1 (0%)	51	79

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	497	GLU

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	784/841 (93%)	642 (82%)	142 (18%)	1	4
1	B	785/841 (93%)	645 (82%)	140 (18%)	2	4
All	All	1569/1682 (93%)	1287 (82%)	282 (18%)	1	4

5 of 282 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	802	SER
1	B	131	ASN
1	B	751	LYS
1	A	809	LEU
1	B	36	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	707	GLN
1	B	89	ASN
1	B	661	GLN
1	A	750	GLN
1	A	800	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	857/917 (93%)	0.35	83 (9%) 7 5	31, 65, 134, 199	8 (0%)
1	B	858/917 (93%)	0.44	85 (9%) 7 5	36, 76, 140, 181	8 (0%)
All	All	1715/1834 (93%)	0.40	168 (9%) 7 5	31, 71, 139, 199	16 (0%)

The worst 5 of 168 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	707	GLN	9.4
1	A	244	ILE	8.1
1	A	294	LEU	7.0
1	A	236	SER	6.9
1	A	301	VAL	6.8

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.