Full wwPDB X-ray Structure Validation Report ⓘ

Oct 10, 2023 – 12:47 AM EDT

PDB ID : 7SR2
Title : Crystal structure of the human SNX25 regulator of G-protein signalling (RGS) domain
Authors : Collins, B.M.; Paul, B.; Weeratunga, S.
Deposited on : 2021-11-07
Resolution : 2.42 Å(reported)

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

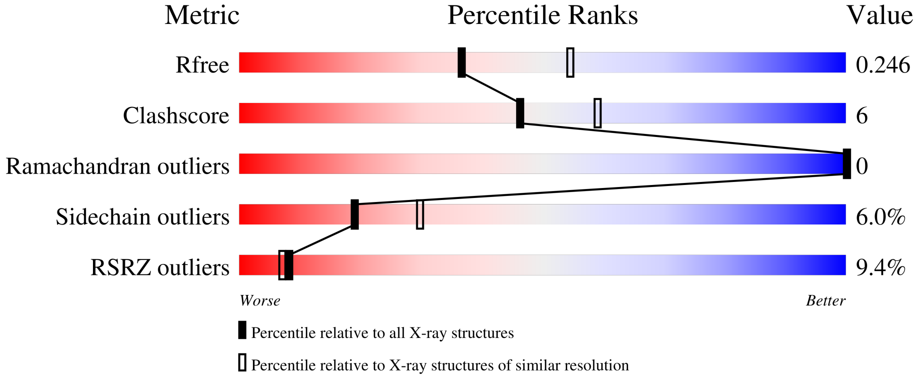
MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

1 Overall quality at a glance ⓘ

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

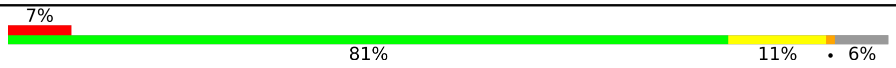
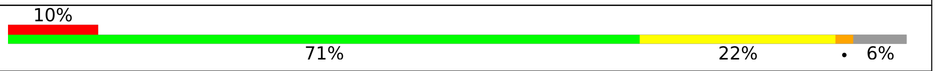
The reported resolution of this entry is 2.42 Å.

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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

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% 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	124	
1	B	124	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	LEU	B	590	-	-	-	X

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 2026 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 Sorting nexin-25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	116	987	643	155	186	3	0	0	0
1	B	117	996	649	157	187	3	0	0	0

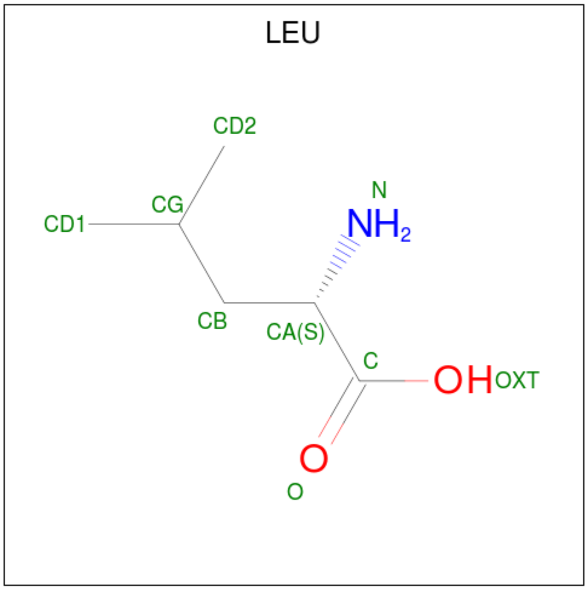
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	526	ALA	CYS	engineered mutation	UNP A0A494C0S0
B	526	ALA	CYS	engineered mutation	UNP A0A494C0S0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

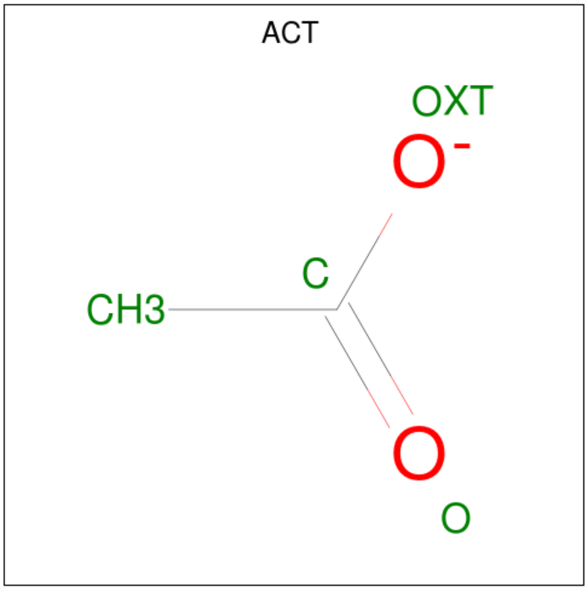
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0

- Molecule 3 is LEUCINE (three-letter code: LEU) (formula: $C_6H_{13}NO_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	B	1	8	6	1	1	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



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

- Molecule 5 is water.

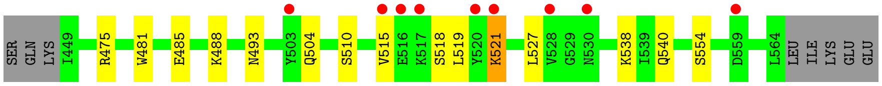
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	19	Total O 19 19	0	0
5	B	11	Total O 11 11	0	0

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 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.

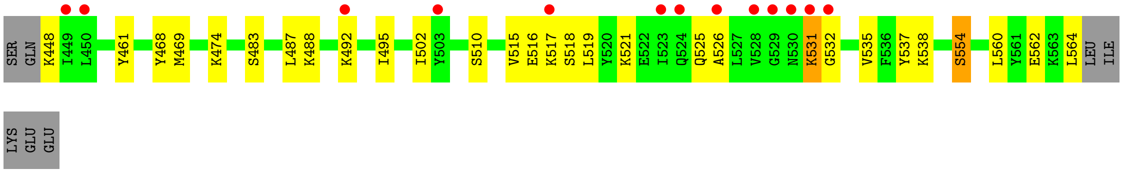
- Molecule 1: Sorting nexin-25

Chain A:



- Molecule 1: Sorting nexin-25

Chain B:



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	44.03Å 75.86Å 51.11Å 90.00° 109.27° 90.00°	Depositor
Resolution (Å)	41.57 – 2.42 41.57 – 2.42	Depositor EDS
% Data completeness (in resolution range)	97.7 (41.57-2.42) 97.7 (41.57-2.42)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 2.42Å)	Xtriage
Refinement program	PHENIX 1.20rc3_4406	Depositor
R, R_{free}	0.211 , 0.247 0.209 , 0.246	Depositor DCC
R_{free} test set	1186 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	43.5	Xtriage
Anisotropy	0.359	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 47.3	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2026	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.26% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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)

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

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	0/1009	0.56	1/1357 (0.1%)
1	B	0.37	0/1018	0.50	0/1368
All	All	0.38	0/2027	0.53	1/2725 (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	A	527	LEU	CB-CG-CD2	-5.49	101.67	111.00

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	987	0	966	9	0
1	B	996	0	979	16	0
2	A	1	0	0	0	0
3	B	8	0	10	0	0
4	B	4	0	3	0	0
5	A	19	0	0	5	1
5	B	11	0	0	0	0
All	All	2026	0	1958	25	1

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 (25) 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:554:SER:OG	5:A:601:HOH:O	2.04	0.75
1:A:515:VAL:CG2	1:A:519:LEU:HB2	2.27	0.64
1:B:488:LYS:HG2	1:B:537:TYR:CE1	2.33	0.63
1:B:526:ALA:HA	1:B:531:LYS:HB3	1.80	0.62
1:B:461:TYR:CD1	1:B:564:LEU:HD22	2.36	0.61
1:A:485:GLU:OE2	5:A:602:HOH:O	2.16	0.60
1:B:521:LYS:O	1:B:525:GLN:HG3	2.02	0.60
1:B:531:LYS:HG3	1:B:532:GLY:H	1.68	0.58
1:A:493:ASN:OD1	5:A:603:HOH:O	2.17	0.57
1:B:487:LEU:HD12	1:B:495:ILE:HD12	1.88	0.54
1:B:492:LYS:HA	1:B:495:ILE:HG12	1.92	0.51
1:B:516:GLU:HG3	1:B:518:SER:H	1.74	0.50
1:A:493:ASN:ND2	5:A:603:HOH:O	2.42	0.48
1:B:469:MET:HE3	1:B:474:LYS:HB3	1.96	0.47
1:A:488:LYS:HD2	1:A:540:GLN:HE22	1.79	0.46
1:B:468:TYR:OH	1:B:554:SER:HB2	2.16	0.46
1:B:483:SER:HB3	1:B:502:ILE:HD11	1.99	0.44
1:B:517:LYS:HA	1:B:517:LYS:HD3	1.68	0.43
1:B:469:MET:CE	1:B:474:LYS:HB3	2.48	0.42
1:B:531:LYS:HG3	1:B:532:GLY:N	2.35	0.42
1:A:521:LYS:H	1:A:521:LYS:HG3	1.61	0.41
1:A:481:TRP:CD1	1:A:540:GLN:HG3	2.56	0.41
1:B:515:VAL:HG12	1:B:516:GLU:O	2.20	0.40
1:B:519:LEU:HD22	1:B:535:VAL:HG23	2.03	0.40
1:A:493:ASN:CG	5:A:603:HOH:O	2.59	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:604:HOH:O	5:A:610:HOH:O[2_656]	1.98	0.22

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	114/124 (92%)	109 (96%)	5 (4%)	0	100	100
1	B	115/124 (93%)	114 (99%)	1 (1%)	0	100	100
All	All	229/248 (92%)	223 (97%)	6 (3%)	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	108/116 (93%)	102 (94%)	6 (6%)	21	32
1	B	109/116 (94%)	102 (94%)	7 (6%)	17	27
All	All	217/232 (94%)	204 (94%)	13 (6%)	19	30

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

Mol	Chain	Res	Type
1	A	475	ARG
1	A	504	GLN
1	A	510	SER
1	A	518	SER
1	A	521	LYS
1	A	538	LYS
1	B	448	LYS
1	B	510	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	531	LYS
1	B	538	LYS
1	B	554	SER
1	B	560	LEU
1	B	562	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	504	GLN
1	A	505	ASN
1	A	540	GLN
1	B	491	ASN
1	B	525	GLN

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 ⓘ

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

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
3	LEU	B	590	-	5,7,8	0.58	0	5,8,10	2.10	1 (20%)
4	ACT	B	591	-	3,3,3	1.61	1 (33%)	3,3,3	1.26	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
3	LEU	B	590	-	-	4/5/6/8	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	591	ACT	CH3-C	2.44	1.59	1.49

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	590	LEU	CD2-CG-CB	-4.22	95.57	111.11

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	590	LEU	O-C-CA-CB
3	B	590	LEU	C-CA-CB-CG
3	B	590	LEU	CA-CB-CG-CD2
3	B	590	LEU	CA-CB-CG-CD1

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

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	116/124 (93%)	0.59	9 (7%)	13	11	33, 47, 87, 106	0
1	B	117/124 (94%)	0.67	13 (11%)	5	4	32, 50, 79, 101	0
All	All	233/248 (93%)	0.63	22 (9%)	8	7	32, 49, 84, 106	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	530	ASN	8.8
1	A	517	LYS	5.3
1	A	515	VAL	4.3
1	A	516	GLU	4.2
1	B	523	ILE	4.0
1	A	530	ASN	3.8
1	B	529	GLY	3.5
1	B	503	TYR	2.9
1	B	450	LEU	2.8
1	B	528	VAL	2.7
1	B	517	LYS	2.6
1	B	531	LYS	2.6
1	A	503	TYR	2.4
1	A	559	ASP	2.3
1	A	520	TYR	2.3
1	A	521	LYS	2.3
1	B	526	ALA	2.2
1	B	524	GLN	2.2
1	B	532	GLY	2.2
1	B	492	LYS	2.1
1	B	449	ILE	2.1
1	A	528	VAL	2.1

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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**	**B-factors(Å^2)**	**Q<0.9**
3	LEU	B	590	8/9	0.79	0.49	65,70,89,91	0
4	ACT	B	591	4/4	0.90	0.20	47,61,74,75	0
2	ZN	A	590	1/1	0.99	0.13	34,34,34,34	0

6.5 Other polymers (i)

There are no such residues in this entry.