



Full wwPDB X-ray Structure Validation Report ⓘ

May 1, 2018 – 12:11 PM EDT

PDB ID : 6F8D
Title : Crystal structure of Human ARS2 residues 171-270 + 408-763 (P65 form)
Authors : Cusack, S.; Schulze, W.M.
Deposited on : 2017-12-13
Resolution : 3.48 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

<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 : rb-20031172
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031172

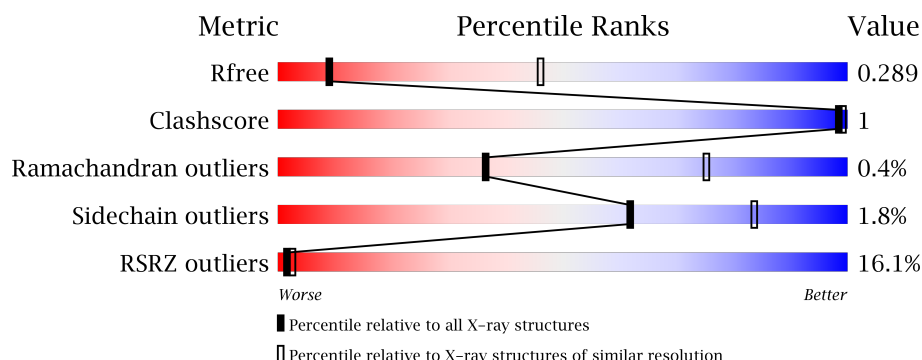
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.48 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}	111664	1159 (3.56-3.40)
Clashscore	122126	1235 (3.56-3.40)
Ramachandran outliers	120053	1202 (3.56-3.40)
Sidechain outliers	120020	1203 (3.56-3.40)
RSRZ outliers	108989	1080 (3.56-3.40)

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 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	100	
1	B	100	
2	C	356	
2	D	356	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6504 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 Serrate RNA effector molecule homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	96	Total	C	N	O	S	0	0	0
			805	509	144	148	4			
1	B	94	Total	C	N	O	S	0	0	0
			785	496	142	143	4			

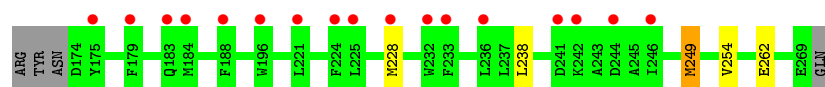
- Molecule 2 is a protein called Serrate RNA effector molecule homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	296	Total	C	N	O	S	0	0	0
			2457	1553	447	445	12			
2	D	296	Total	C	N	O	S	0	0	0
			2457	1553	447	445	12			

3 Residue-property plots

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.

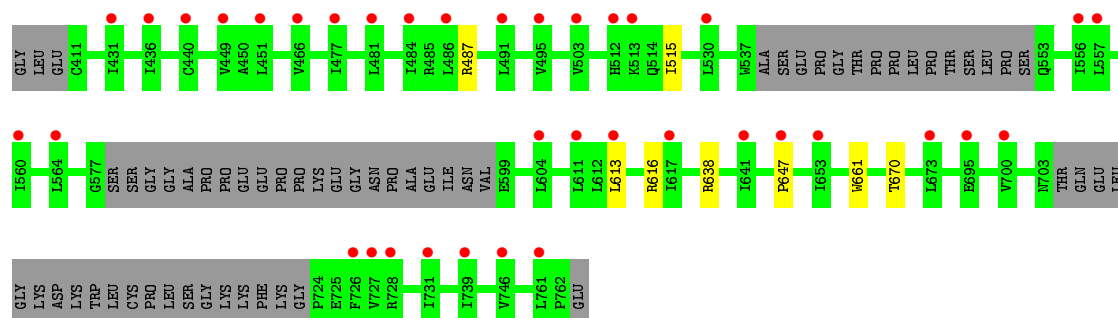
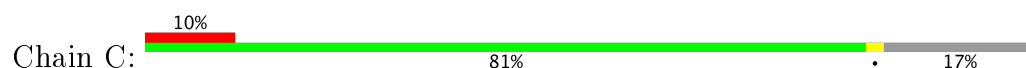
- Molecule 1: Serrate RNA effector molecule homolog



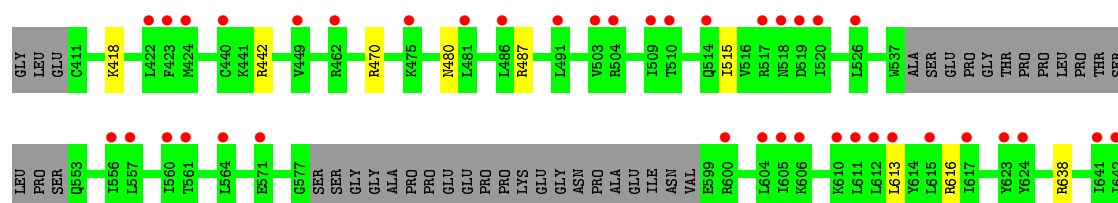
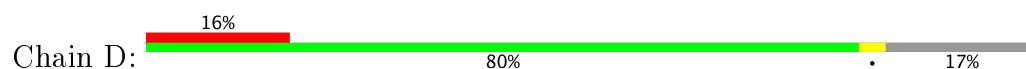
- Molecule 1: Serrate RNA effector molecule homolog

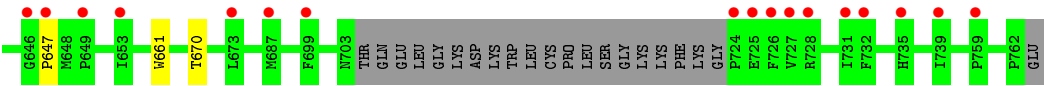


- Molecule 2: Serrate RNA effector molecule homolog



- Molecule 2: Serrate RNA effector molecule homolog





4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	105.21Å 105.21Å 267.04Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	91.11 – 3.48 48.94 – 3.48	Depositor EDS
% Data completeness (in resolution range)	97.8 (91.11-3.48) 97.9 (48.94-3.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.30 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.264 , 0.292 0.265 , 0.289	Depositor DCC
R_{free} test set	922 reflections (4.40%)	wwPDB-VP
Wilson B-factor (Å ²)	123.2	Xtriage
Anisotropy	0.342	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 105.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.097 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6504	wwPDB-VP
Average B, all atoms (Å ²)	164.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.51% 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	0.36	0/819	0.51	0/1096
1	B	0.36	0/798	0.52	0/1067
2	C	0.35	0/2505	0.53	0/3374
2	D	0.36	0/2505	0.54	0/3374
All	All	0.36	0/6627	0.53	0/8911

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	805	0	792	3	0
1	B	785	0	779	2	0
2	C	2457	0	2486	4	0
2	D	2457	0	2486	4	0
All	All	6504	0	6543	8	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 1.

All (8) 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)
2:D:515:ILE:HD11	2:D:661:TRP:CD1	2.34	0.63
2:C:515:ILE:HD11	2:C:661:TRP:CD1	2.34	0.62
2:D:418:LYS:HB2	2:D:470:ARG:HG3	1.97	0.46
1:A:238:LEU:HD21	2:C:613:LEU:HD21	2.02	0.41
1:A:254:VAL:HG22	2:C:638:ARG:HB2	2.03	0.41
1:B:254:VAL:HG22	2:D:638:ARG:HB2	2.03	0.41
1:A:249:MET:HE1	2:C:613:LEU:HD22	2.04	0.40
1:B:238:LEU:HD21	2:D:613:LEU:HD21	2.02	0.40

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	94/100 (94%)	90 (96%)	4 (4%)	0	100	100
1	B	92/100 (92%)	88 (96%)	4 (4%)	0	100	100
2	C	288/356 (81%)	269 (93%)	18 (6%)	1 (0%)	43	78
2	D	288/356 (81%)	268 (93%)	18 (6%)	2 (1%)	24	65
All	All	762/912 (84%)	715 (94%)	44 (6%)	3 (0%)	36	75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	442	ARG
2	D	647	PRO
2	C	647	PRO

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	84/88 (96%)	81 (96%)	3 (4%)	38	70
1	B	82/88 (93%)	79 (96%)	3 (4%)	37	69
2	C	276/325 (85%)	273 (99%)	3 (1%)	76	89
2	D	276/325 (85%)	272 (99%)	4 (1%)	69	87
All	All	718/826 (87%)	705 (98%)	13 (2%)	62	83

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

Mol	Chain	Res	Type
1	A	228	MET
1	A	249	MET
1	A	262	GLU
2	C	487	ARG
2	C	616	ARG
2	C	670	THR
1	B	228	MET
1	B	249	MET
1	B	262	GLU
2	D	480	ASN
2	D	487	ARG
2	D	616	ARG
2	D	670	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	625	ASN
2	D	480	ASN
2	D	625	ASN

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](#)

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 ⓘ

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	96/100 (96%)	1.06	17 (17%) 1 2	109, 151, 193, 229	0
1	B	94/100 (94%)	0.99	15 (15%) 1 3	121, 157, 187, 204	0
2	C	296/356 (83%)	0.97	37 (12%) 4 5	103, 155, 223, 253	0
2	D	296/356 (83%)	1.24	57 (19%) 1 1	114, 167, 240, 273	0
All	All	782/912 (85%)	1.08	126 (16%) 1 2	103, 159, 223, 273	0

All (126) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	724	PRO	8.9
2	D	731	ILE	6.1
2	D	739	ILE	6.0
2	D	604	LEU	5.7
1	A	175	TYR	5.4
2	D	673	LEU	4.9
2	D	653	ILE	4.8
2	C	604	LEU	4.7
2	D	725	GLU	4.7
2	D	509	ILE	4.6
2	D	514	GLN	4.6
2	D	732	PHE	4.5
2	D	641	ILE	4.3
2	D	610	LYS	4.1
2	D	560	ILE	4.0
2	D	624	TYR	4.0
2	D	510	THR	3.9
2	D	728	ARG	3.9
2	C	440	CYS	3.8
1	A	242	LYS	3.6
1	B	178	ASP	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	611	LEU	3.6
1	A	228	MET	3.6
2	D	649	PRO	3.5
1	B	236	LEU	3.5
2	D	611	LEU	3.5
2	D	699	PHE	3.5
2	D	600	ARG	3.5
1	A	241	ASP	3.5
1	B	224	PHE	3.3
2	D	440	CYS	3.3
2	D	557	LEU	3.2
2	C	731	ILE	3.2
2	D	642	ILE	3.2
2	C	491	LEU	3.1
2	C	726	PHE	3.1
2	C	431	ILE	3.1
1	B	233	PHE	3.0
2	D	422	LEU	3.0
2	D	727	VAL	3.0
1	B	249	MET	3.0
2	D	617	ILE	3.0
2	D	503	VAL	2.9
1	A	233	PHE	2.9
2	C	739	ILE	2.9
2	C	647	PRO	2.9
2	D	605	ILE	2.8
1	A	179	PHE	2.8
2	D	571	GLU	2.8
2	D	613	LEU	2.8
2	C	641	ILE	2.8
1	A	232	TRP	2.8
2	D	526	LEU	2.7
2	C	746	VAL	2.7
2	D	520	ILE	2.7
2	D	504	ARG	2.7
2	D	687	MET	2.6
1	B	238	LEU	2.6
1	B	196	TRP	2.6
1	A	224	PHE	2.6
2	D	759	PRO	2.6
1	A	184	MET	2.6
1	A	244	ASP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	612	LEU	2.6
2	C	695	GLU	2.5
2	C	451	LEU	2.5
1	A	236	LEU	2.5
2	D	491	LEU	2.5
1	B	179	PHE	2.5
2	D	486	LEU	2.5
2	C	436	ILE	2.5
2	D	564	LEU	2.5
1	B	237	LEU	2.5
2	C	530	LEU	2.4
2	C	728	ARG	2.4
2	C	486	LEU	2.4
2	C	700	VAL	2.4
2	D	517	ARG	2.4
2	D	475	LYS	2.4
2	C	557	LEU	2.4
2	D	518	ASN	2.4
2	C	477	ILE	2.4
2	C	613	LEU	2.4
2	C	761	LEU	2.4
2	D	615	LEU	2.4
2	C	481	LEU	2.3
2	D	735	HIS	2.3
2	D	646	GLY	2.3
1	A	246	ILE	2.3
2	C	560	ILE	2.3
2	C	673	LEU	2.3
2	D	481	LEU	2.3
2	D	449	VAL	2.3
2	C	495	VAL	2.3
2	C	503	VAL	2.3
2	C	617	ILE	2.3
1	A	196	TRP	2.3
1	A	221	LEU	2.3
2	D	561	THR	2.3
2	C	484	ILE	2.3
2	D	462	ARG	2.2
1	A	225	LEU	2.2
2	D	647	PRO	2.2
2	D	623	TYR	2.2
1	B	265	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	183	GLN	2.2
1	B	182	GLN	2.2
2	C	449	VAL	2.2
2	D	519	ASP	2.2
2	C	564	LEU	2.1
2	D	726	PHE	2.1
1	B	246	ILE	2.1
2	D	556	ILE	2.1
2	C	512	HIS	2.1
2	C	466	VAL	2.1
1	B	228	MET	2.1
2	C	727	VAL	2.1
2	D	424	MET	2.1
1	B	232	TRP	2.1
2	C	556	ILE	2.1
2	D	606	LYS	2.1
2	C	653	ILE	2.1
2	C	513	LYS	2.1
1	A	188	PHE	2.1
1	B	235	ASN	2.0
2	D	423	PHE	2.0

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 carbohydrates 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.