

# Full wwPDB X-ray Structure Validation Report (i)

#### Oct 21, 2024 – 07:29 PM EDT

:	2Q3X
:	The RIM1alpha C2B domain
:	Guan, R.; Dai, H.; Tomchick, D.R.; Machius, M.; Sudhof, T.C.; Rizo, J.
	2007-05-30
:	1.73 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 (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

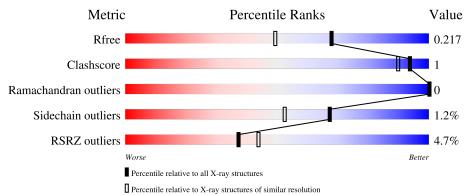
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
$\mathrm{EDS}$	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 1.73 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	164625	1043 (1.74-1.74)
Clashscore	180529	1119 (1.74-1.74)
Ramachandran outliers	177936	1112 (1.74-1.74)
Sidechain outliers	177891	1112 (1.74-1.74)
RSRZ outliers	164620	1043 (1.74-1.74)

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	А	171	79%	•	17%
1	В	171	<b>4%</b> 81%	•	16%



## 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2491 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.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	А	142	Total 1121			O 203		0	3	0
1	В	144	Total 1142	С 733		0 211	 Se 4	7	4	0

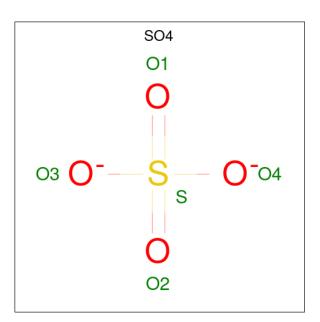
• Molecule 1 is a protein called Regulating synaptic membrane exocytosis protein 1.

Chain	Residue	Modelled	Actual	Comment	Reference
А	1445	SER	-	expression tag	UNP Q9JIR4
А	1446	PRO	-	expression tag	UNP Q9JIR4
А	1462	MSE	MET	modified residue	UNP Q9JIR4
А	1469	MSE	MET	modified residue	UNP Q9JIR4
А	1553	MSE	MET	modified residue	UNP Q9JIR4
А	1559	MSE	MET	modified residue	UNP Q9JIR4
А	1574	MSE	MET	modified residue	UNP Q9JIR4
В	1445	SER	-	expression tag	UNP Q9JIR4
В	1446	PRO	-	expression tag	UNP Q9JIR4
В	1462	MSE	MET	modified residue	UNP Q9JIR4
В	1469	MSE	MET	modified residue	UNP Q9JIR4
В	1553	MSE	MET	modified residue	UNP Q9JIR4
В	1559	MSE	MET	modified residue	UNP Q9JIR4
В	1574	MSE	MET	modified residue	UNP Q9JIR4

There are 14 discrepancies between the modelled and reference sequences:

• Molecule 2 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	В	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	В	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0

• Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	2	Total Cl 2 2	0	0
3	В	1	Total Cl 1 1	0	0

• Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	В	1	Total 1	Na 1	0	0

• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	91	Total         O           91         91	0	0
5	В	98	Total         O           98         98	0	0

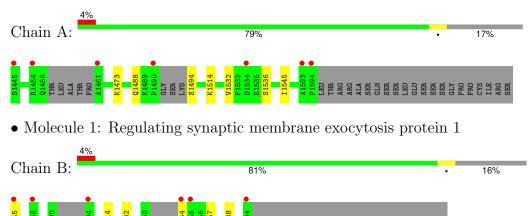


## 3 Residue-property plots (i)

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.

SER GLN SER SER LEU CLU GLU GLU SER SER SER PRO PRO CYS SER ARG

• Molecule 1: Regulating synaptic membrane exocytosis protein 1





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants	62.04Å 62.04Å 145.18Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $120.00^{\circ}$	Depositor
Resolution (Å)	26.40 - 1.73	Depositor
Resolution (A)	26.40 - 1.73	EDS
% Data completeness	99.7 (26.40-1.73)	Depositor
(in resolution range)	99.7 (26.40-1.73)	EDS
R <sub>merge</sub>	(Not available)	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.62 (at 1.73 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
D D.	0.178 , $0.215$	Depositor
$R, R_{free}$	0.179 , $0.217$	DCC
R <sub>free</sub> test set	1400 reflections $(4.05\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	24.1	Xtriage
Anisotropy	0.005	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.37, $38.6$	EDS
L-test for twinning <sup>2</sup>	$<  L  > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.031 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2491	wwPDB-VP
Average B, all atoms $(Å^2)$	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.51% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: NA, SO4, CL

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		
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.81	0/1143	0.89	0/1536	
1	В	0.82	0/1163	0.83	0/1570	
All	All	0.82	0/2306	0.86	0/3106	

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	А	1121	0	1167	4	0
1	В	1142	0	1178	4	0
2	А	25	0	0	0	0
2	В	10	0	0	0	0
3	А	2	0	0	0	0
3	В	1	0	0	0	0
4	В	1	0	0	0	0
5	А	91	0	0	0	0
5	В	98	0	0	0	0
All	All	2491	0	2345	6	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 (6) 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:1554:ASP:HB3	1:B:1557[B]:CYS:SG	2.36	0.65
1:A:1514:LYS:HE3	1:A:1532:VAL:O	2.00	0.61
1:A:1488:GLN:NE2	1:A:1494:SER:O	2.41	0.54
1:B:1514:LYS:HE3	1:B:1532:VAL:O	2.12	0.50
1:A:1545:ILE:CD1	1:B:1588:VAL:HG13	2.44	0.47
1:A:1545:ILE:HD11	1:B:1588:VAL:HG13	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	Perce	ntiles
1	А	139/171~(81%)	133~(96%)	6 (4%)	0	100	100
1	В	142/171~(83%)	140 (99%)	2(1%)	0	100	100
All	All	281/342 ( $82%$ )	273~(97%)	8(3%)	0	100	100

There are no Ramachandran outliers to report.

#### 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	А	125/142~(88%)	123~(98%)	2(2%)	58 38
1	В	128/142~(90%)	127~(99%)	1 (1%)	79 70
All	All	253/284~(89%)	250~(99%)	3~(1%)	67 52

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	1473	LYS
1	А	1536	SER
1	В	1445	SER

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such side chains are listed below:

Mol	Chain	Res	Type
1	А	1507	ASN

#### 5.3.3 RNA (i)

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 oligosaccharides in this entry.

### 5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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).



Mal	Mol Type		Dec	Link	Bond lengths			Bond angles		
IVIOI	Type	Chain	nes	Res Link	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
2	SO4	А	202	-	4,4,4	0.37	0	$6,\!6,\!6$	0.55	0
2	SO4	А	201	-	4,4,4	0.40	0	$6,\!6,\!6$	0.39	0
2	SO4	А	206	-	4,4,4	0.24	0	$6,\!6,\!6$	0.19	0
2	SO4	В	205	-	4,4,4	0.25	0	$6,\!6,\!6$	0.19	0
2	SO4	А	204	-	4,4,4	0.30	0	$6,\!6,\!6$	0.48	0
2	SO4	А	207	-	4,4,4	0.23	0	$6,\!6,\!6$	0.09	0
2	SO4	В	203	-	4,4,4	0.43	0	$6,\!6,\!6$	0.32	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 (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,  $95^{th}$  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	$\langle RSRZ \rangle$	# RSRZ > 2	$OWAB(Å^2)$	Q<0.9
1	А	137/171~(80%)	0.10	7 (5%) 34 41	12, 23, 39, 43	2 (1%)
1	В	$139/171 \ (81\%)$	0.23	6 (4%) 40 48	12, 23, 36, 44	3 (2%)
All	All	276/342~(80%)	0.16	13 (4%) 37 44	12, 23, 37, 44	5 (1%)

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	1554	ASP	4.1
1	А	1534	ASP	3.6
1	В	1445	SER	3.3
1	А	1445	SER	3.3
1	А	1454	ARG	3.2
1	В	1594	PRO	3.2
1	А	1461	ALA	2.7
1	А	1594	PRO	2.5
1	А	1490	PRO	2.3
1	А	1593	ALA	2.3
1	В	1555	HIS	2.2
1	В	1458	ALA	2.1
1	В	1494	SER	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 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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
2	SO4	А	207	5/5	0.85	0.10	$67,\!67,\!68,\!68$	0
3	CL	А	403	1/1	0.86	0.13	$73,\!73,\!73,\!73$	0
2	SO4	А	206	5/5	0.87	0.10	$61,\!61,\!62,\!62$	0
3	CL	А	402	1/1	0.92	0.11	48,48,48,48	0
2	SO4	В	205	5/5	0.95	0.07	44,45,48,48	0
2	SO4	А	204	5/5	0.97	0.13	$34,\!37,\!41,\!41$	0
4	NA	В	301	1/1	0.97	0.05	$22,\!22,\!22,\!22$	0
2	SO4	А	201	5/5	0.98	0.06	$24,\!24,\!26,\!28$	0
2	SO4	В	203	5/5	0.98	0.06	$24,\!25,\!29,\!32$	0
2	SO4	А	202	5/5	0.98	0.08	$26,\!26,\!28,\!29$	0
3	CL	В	401	1/1	0.99	0.10	33,33,33,33	0

### 6.5 Other polymers (i)

There are no such residues in this entry.

