Full wwPDB X-ray Structure Validation Report (i)

Dec 17, 2023 – 01:01 am GMT

PDB ID	:	3ZMF
Title	:	Salmonella enterica SadA 303-358 fused to GCN4 adaptors (SadAK2)
Authors	:	Hartmann, M.D.; Hernandez Alvarez, B.; Albrecht, R.; Lupas, A.N.
Deposited on	:	2013-02-08
Resolution	:	1.85 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:

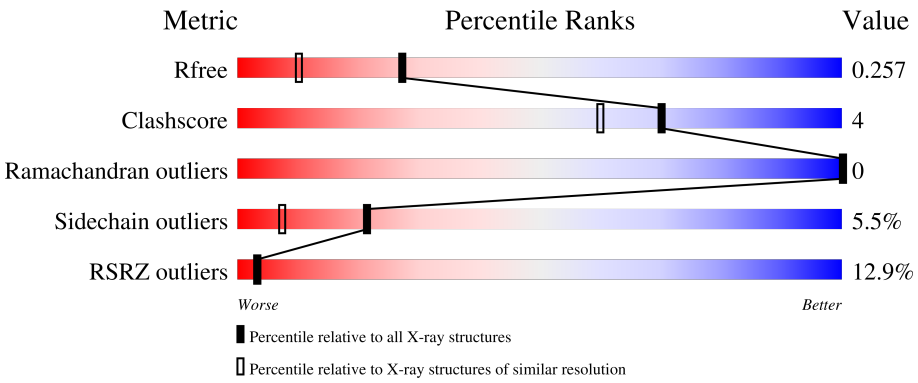
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	113	15% 85% 13% •
1	B	113	12% 81% 15% • •
1	C	113	12% 88% 8% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2550 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 GENERAL CONTROL PROTEIN GCN4, PUTATIVE INNER MEMBRANE PROTEIN, GENERAL CONTROL PROTEIN GCN4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	111	807	508	134	164	1	0	0	0
1	B	111	812	513	131	167	1	0	0	0
1	C	111	808	509	131	166	2	0	1	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	278	ILE	LEU	engineered mutation	UNP P03069
A	282	ILE	VAL	engineered mutation	UNP P03069
A	285	ILE	LEU	engineered mutation	UNP P03069
A	289	ILE	ASN	engineered mutation	UNP P03069
A	292	ILE	LEU	engineered mutation	UNP P03069
A	296	ILE	VAL	engineered mutation	UNP P03069
A	299	ILE	LEU	engineered mutation	UNP P03069
A	303	ILE	VAL	engineered mutation	UNP P03069
A	362	ILE	LEU	engineered mutation	UNP P03069
A	366	ILE	VAL	engineered mutation	UNP P03069
A	369	ILE	LEU	engineered mutation	UNP P03069
A	373	ILE	ASN	engineered mutation	UNP P03069
A	376	ILE	LEU	engineered mutation	UNP P03069
A	380	ILE	VAL	engineered mutation	UNP P03069
A	383	ILE	LEU	engineered mutation	UNP P03069
A	387	ILE	VAL	engineered mutation	UNP P03069
B	278	ILE	LEU	engineered mutation	UNP P03069
B	282	ILE	VAL	engineered mutation	UNP P03069
B	285	ILE	LEU	engineered mutation	UNP P03069
B	289	ILE	ASN	engineered mutation	UNP P03069
B	292	ILE	LEU	engineered mutation	UNP P03069
B	296	ILE	VAL	engineered mutation	UNP P03069

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	299	ILE	LEU	engineered mutation	UNP P03069
B	303	ILE	VAL	engineered mutation	UNP P03069
B	362	ILE	LEU	engineered mutation	UNP P03069
B	366	ILE	VAL	engineered mutation	UNP P03069
B	369	ILE	LEU	engineered mutation	UNP P03069
B	373	ILE	ASN	engineered mutation	UNP P03069
B	376	ILE	LEU	engineered mutation	UNP P03069
B	380	ILE	VAL	engineered mutation	UNP P03069
B	383	ILE	LEU	engineered mutation	UNP P03069
B	387	ILE	VAL	engineered mutation	UNP P03069
C	278	ILE	LEU	engineered mutation	UNP P03069
C	282	ILE	VAL	engineered mutation	UNP P03069
C	285	ILE	LEU	engineered mutation	UNP P03069
C	289	ILE	ASN	engineered mutation	UNP P03069
C	292	ILE	LEU	engineered mutation	UNP P03069
C	296	ILE	VAL	engineered mutation	UNP P03069
C	299	ILE	LEU	engineered mutation	UNP P03069
C	303	ILE	VAL	engineered mutation	UNP P03069
C	362	ILE	LEU	engineered mutation	UNP P03069
C	366	ILE	VAL	engineered mutation	UNP P03069
C	369	ILE	LEU	engineered mutation	UNP P03069
C	373	ILE	ASN	engineered mutation	UNP P03069
C	376	ILE	LEU	engineered mutation	UNP P03069
C	380	ILE	VAL	engineered mutation	UNP P03069
C	383	ILE	LEU	engineered mutation	UNP P03069
C	387	ILE	VAL	engineered mutation	UNP P03069

- Molecule 2 is water.

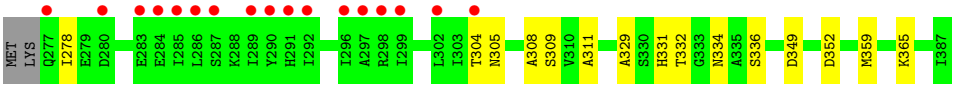
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	39	39	39	0	0
			Total	O		
2	B	45	45	45	0	0
			Total	O		
2	C	39	39	39	0	0

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.

• Molecule 1: GENERAL CONTROL PROTEIN GCN4, PUTATIVE INNER MEMBRANE PROTEIN, GENERAL CONTROL PROTEIN GCN4

Chain A:

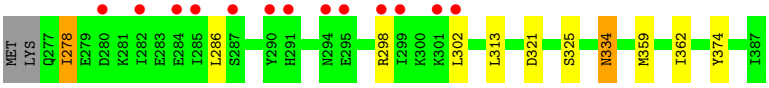


• Molecule 1: GENERAL CONTROL PROTEIN GCN4, PUTATIVE INNER MEMBRANE PROTEIN, GENERAL CONTROL PROTEIN GCN4

Chain B:

• Molecule 1: GENERAL CONTROL PROTEIN GCN4, PUTATIVE INNER MEMBRANE PROTEIN, GENERAL CONTROL PROTEIN GCN4

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	46.59Å 46.59Å 128.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.51 – 1.85 38.51 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.1 (38.51-1.85) 99.8 (38.51-1.85)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.36 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.181 , 0.221 0.211 , 0.257	Depositor DCC
R_{free} test set	1325 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	23.7	Xtriage
Anisotropy	0.677	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 78.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.54$, $< L^2 > = 0.39$	Xtriage
Estimated twinning fraction	0.001 for -h,-k,l 0.080 for h,-h-k,-l 0.022 for -k,-h,-l	Xtriage
Reported twinning fraction	0.726 for H, K, L 0.274 for K, H, -L	Depositor
Outliers	0 of 26768 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2550	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.70% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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)

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	1.16	1/814 (0.1%)	1.03	3/1104 (0.3%)
1	B	1.30	8/819 (1.0%)	1.12	2/1110 (0.2%)
1	C	1.12	0/817	1.09	5/1108 (0.5%)
All	All	1.20	9/2450 (0.4%)	1.08	10/3322 (0.3%)

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	322	GLU	CD-OE2	6.16	1.32	1.25
1	B	322	GLU	CG-CD	6.10	1.61	1.51
1	B	345	THR	CB-CG2	-5.67	1.33	1.52
1	B	339	THR	CB-CG2	-5.55	1.34	1.52
1	A	352	ASP	CB-CG	5.38	1.63	1.51
1	B	309	SER	CB-OG	-5.38	1.35	1.42
1	B	336	SER	CB-OG	-5.30	1.35	1.42
1	B	354	VAL	CB-CG2	-5.13	1.42	1.52
1	B	322	GLU	CD-OE1	5.12	1.31	1.25

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	352	ASP	CB-CG-OD1	9.57	126.91	118.30
1	A	359	MET	CG-SD-CE	-7.47	88.25	100.20
1	A	352	ASP	CB-CG-OD2	-6.55	112.40	118.30
1	C	359[A]	MET	CG-SD-CE	-5.94	90.69	100.20
1	C	359[B]	MET	CG-SD-CE	-5.94	90.69	100.20
1	B	319	LEU	CB-CG-CD1	5.72	120.73	111.00
1	A	352	ASP	N-CA-CB	5.65	120.77	110.60
1	C	321	ASP	CB-CG-OD2	-5.64	113.22	118.30
1	C	359[A]	MET	CB-CG-SD	-5.41	96.16	112.40
1	C	359[B]	MET	CB-CG-SD	-5.41	96.16	112.40

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	807	0	780	7	0
1	B	812	0	791	6	0
1	C	808	0	779	4	0
2	A	39	0	0	1	0
2	B	45	0	0	2	0
2	C	39	0	0	2	0
All	All	2550	0	2350	17	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 4.

All (17) 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:C:374:TYR:OH	2:C:2037:HOH:O	2.11	0.66
1:A:331:HIS:O	1:A:332:THR:HG22	1.98	0.62
1:A:332:THR:HG23	1:A:332:THR:O	2.02	0.58
1:B:365:LYS:NZ	2:B:2045:HOH:O	2.39	0.56
1:B:358:GLN:HA	1:B:361:GLN:NE2	2.28	0.48
1:A:332:THR:HG23	1:A:334:ASN:HD22	1.80	0.46
1:B:278:ILE:O	1:B:282:ILE:HG12	2.16	0.45
1:A:349:ASP:OD1	1:A:349:ASP:N	2.50	0.44
1:C:334:ASN:ND2	2:C:2018:HOH:O	2.49	0.44
1:C:298:ARG:O	1:C:302:LEU:HB2	2.18	0.43
1:B:332:THR:OG1	2:B:2027:HOH:O	2.04	0.43
1:A:365:LYS:NZ	2:A:2031:HOH:O	2.51	0.42
1:C:278:ILE:C	1:C:278:ILE:HD12	2.40	0.42
1:A:308:ALA:O	1:A:311:ALA:N	2.54	0.41
1:B:281:LYS:HA	1:B:284:GLU:OE2	2.21	0.41
1:A:329:ALA:O	1:A:336:SER:HB2	2.21	0.41
1:B:339:THR:O	1:B:340:ASN:HB2	2.20	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	109/113 (96%)	107 (98%)	2 (2%)	0	100	100
1	B	109/113 (96%)	108 (99%)	1 (1%)	0	100	100
1	C	110/113 (97%)	110 (100%)	0	0	100	100
All	All	328/339 (97%)	325 (99%)	3 (1%)	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	A	79/95 (83%)	75 (95%)	4 (5%)	24	9
1	B	81/95 (85%)	78 (96%)	3 (4%)	34	17
1	C	79/95 (83%)	73 (92%)	6 (8%)	13	3
All	All	239/285 (84%)	226 (95%)	13 (5%)	21	8

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

Mol	Chain	Res	Type
1	A	278	ILE
1	A	304	THR
1	A	305	ASN
1	A	309	SER
1	B	302	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	365	LYS
1	B	373	ILE
1	C	278	ILE
1	C	286	LEU
1	C	313	LEU
1	C	325	SER
1	C	334	ASN
1	C	362	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	334	ASN
1	B	361	GLN
1	C	334	ASN
1	C	361	GLN

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 monosaccharides 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 (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	111/113 (98%)	0.75	17 (15%)	2	2	18, 38, 103, 111	0
1	B	111/113 (98%)	0.50	13 (11%)	4	4	14, 33, 82, 102	0
1	C	111/113 (98%)	0.48	13 (11%)	4	4	18, 36, 81, 93	0
All	All	333/339 (98%)	0.58	43 (12%)	3	3	14, 36, 94, 111	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	286	LEU	7.3
1	A	280	ASP	6.3
1	A	302	LEU	6.1
1	A	287	SER	5.6
1	B	387	ILE	5.3
1	A	290	TYR	5.2
1	B	297	ALA	4.3
1	C	302	LEU	4.0
1	C	285	ILE	3.9
1	A	298	ARG	3.9
1	B	290	TYR	3.8
1	C	298	ARG	3.8
1	B	294	ASN	3.8
1	A	299	ILE	3.8
1	A	285	ILE	3.4
1	C	290	TYR	3.1
1	A	304	THR	3.1
1	C	291	HIS	3.1
1	A	277	GLN	3.1
1	A	292	ILE	3.1
1	B	277	GLN	3.0
1	A	284	GLU	3.0
1	B	300	LYS	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	278	ILE	2.9
1	A	297	ALA	2.7
1	C	284	GLU	2.7
1	C	301	LYS	2.6
1	A	296	ILE	2.6
1	C	299	ILE	2.5
1	A	291	HIS	2.5
1	B	385	LYS	2.5
1	C	287	SER	2.4
1	C	280	ASP	2.4
1	B	280	ASP	2.4
1	B	375	HIS	2.4
1	B	386	LEU	2.4
1	B	374	TYR	2.3
1	C	295	GLU	2.2
1	C	294	ASN	2.2
1	A	289	ILE	2.1
1	B	384	LYS	2.1
1	A	283	GLU	2.0
1	C	282	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 monosaccharides 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.