



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 08:53 AM UTC

PDB ID : 9OIK / pdb_00009oik
Title : Structure of S. Typhimurium 14028 Gifsy-1 prophage HepS bound to bacteriophage lambda J Tail Tip
Authors : Sargen, M.R.; Antine, S.P.; Grabe, G.J.; Antonellis, G.; Ragucci, A.E.; Kranzusch, P.J.; Helaine, S.A.
Deposited on : 2025-05-06
Resolution : 1.86 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

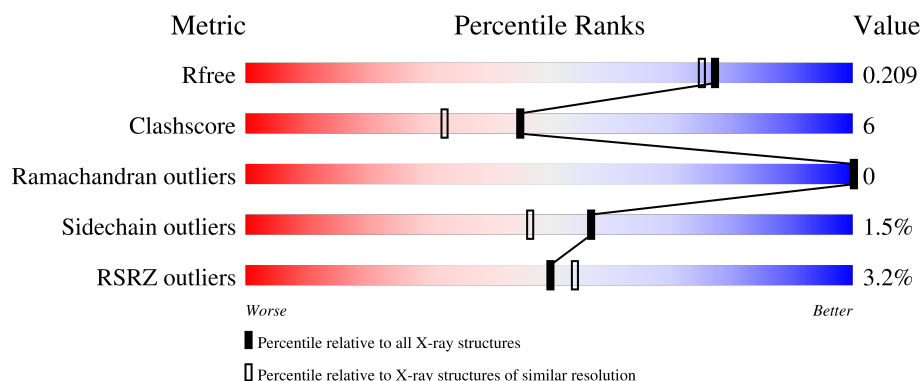
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.86 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}	180053	3428 (1.86-1.86)
Clashscore	190562	3579 (1.86-1.86)
Ramachandran outliers	187476	3553 (1.86-1.86)
Sidechain outliers	187428	3553 (1.86-1.86)
RSRZ outliers	180081	3429 (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 $\leq 5\%$. 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	228	4% 84% 11% • •
1	B	228	2% 79% 8% 14%
2	C	156	% 13% • 84%
2	D	156	18% • 79%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4239 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 HepS-H105A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	0	0	0
			1705	1087	285	322	11			
1	B	197	Total	C	N	O	S	0	0	0
			1551	993	258	290	10			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	expression tag	UNP A0A0F6B549
A	-1	GLY	-	expression tag	UNP A0A0F6B549
A	0	SER	-	expression tag	UNP A0A0F6B549
A	105	ALA	HIS	engineered mutation	UNP A0A0F6B549
A	211	ALA	-	expression tag	UNP A0A0F6B549
A	212	ALA	-	expression tag	UNP A0A0F6B549
A	213	ALA	-	expression tag	UNP A0A0F6B549
A	214	SER	-	expression tag	UNP A0A0F6B549
A	215	ALA	-	expression tag	UNP A0A0F6B549
A	216	ALA	-	expression tag	UNP A0A0F6B549
A	217	GLY	-	expression tag	UNP A0A0F6B549
A	218	ASP	-	expression tag	UNP A0A0F6B549
A	219	TYR	-	expression tag	UNP A0A0F6B549
A	220	LYS	-	expression tag	UNP A0A0F6B549
A	221	ASP	-	expression tag	UNP A0A0F6B549
A	222	ASP	-	expression tag	UNP A0A0F6B549
A	223	ASP	-	expression tag	UNP A0A0F6B549
A	224	ASP	-	expression tag	UNP A0A0F6B549
A	225	LYS	-	expression tag	UNP A0A0F6B549
B	-2	MET	-	expression tag	UNP A0A0F6B549
B	-1	GLY	-	expression tag	UNP A0A0F6B549
B	0	SER	-	expression tag	UNP A0A0F6B549
B	105	ALA	HIS	engineered mutation	UNP A0A0F6B549
B	211	ALA	-	expression tag	UNP A0A0F6B549
B	212	ALA	-	expression tag	UNP A0A0F6B549

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	213	ALA	-	expression tag	UNP A0A0F6B549
B	214	SER	-	expression tag	UNP A0A0F6B549
B	215	ALA	-	expression tag	UNP A0A0F6B549
B	216	ALA	-	expression tag	UNP A0A0F6B549
B	217	GLY	-	expression tag	UNP A0A0F6B549
B	218	ASP	-	expression tag	UNP A0A0F6B549
B	219	TYR	-	expression tag	UNP A0A0F6B549
B	220	LYS	-	expression tag	UNP A0A0F6B549
B	221	ASP	-	expression tag	UNP A0A0F6B549
B	222	ASP	-	expression tag	UNP A0A0F6B549
B	223	ASP	-	expression tag	UNP A0A0F6B549
B	224	ASP	-	expression tag	UNP A0A0F6B549
B	225	LYS	-	expression tag	UNP A0A0F6B549

- Molecule 2 is a protein called J tail tip.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	25	Total	C	N	O	S	0	0	0
			205	136	33	34	2			
2	D	32	Total	C	N	O	S	0	0	0
			254	164	42	46	2			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	778	MET	-	expression tag	UNP Q12306
C	779	GLY	-	expression tag	UNP Q12306
C	780	HIS	-	expression tag	UNP Q12306
C	781	HIS	-	expression tag	UNP Q12306
C	782	HIS	-	expression tag	UNP Q12306
C	783	HIS	-	expression tag	UNP Q12306
C	784	HIS	-	expression tag	UNP Q12306
C	785	HIS	-	expression tag	UNP Q12306
C	786	HIS	-	expression tag	UNP Q12306
C	787	HIS	-	expression tag	UNP Q12306
C	788	HIS	-	expression tag	UNP Q12306
C	789	HIS	-	expression tag	UNP Q12306
C	790	SER	-	expression tag	UNP Q12306
C	791	SER	-	expression tag	UNP Q12306
C	792	GLY	-	expression tag	UNP Q12306
C	793	HIS	-	expression tag	UNP Q12306
C	794	ILE	-	expression tag	UNP Q12306

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	795	GLU	-	expression tag	UNP Q12306
C	796	GLY	-	expression tag	UNP Q12306
C	797	ARG	-	expression tag	UNP Q12306
C	798	HIS	-	expression tag	UNP Q12306
C	799	MET	-	expression tag	UNP Q12306
C	800	ALA	-	expression tag	UNP Q12306
C	801	SER	-	expression tag	UNP Q12306
C	900	SER	-	linker	UNP Q12306
C	901	MET	-	linker	UNP Q12306
D	778	MET	-	expression tag	UNP Q12306
D	779	GLY	-	expression tag	UNP Q12306
D	780	HIS	-	expression tag	UNP Q12306
D	781	HIS	-	expression tag	UNP Q12306
D	782	HIS	-	expression tag	UNP Q12306
D	783	HIS	-	expression tag	UNP Q12306
D	784	HIS	-	expression tag	UNP Q12306
D	785	HIS	-	expression tag	UNP Q12306
D	786	HIS	-	expression tag	UNP Q12306
D	787	HIS	-	expression tag	UNP Q12306
D	788	HIS	-	expression tag	UNP Q12306
D	789	HIS	-	expression tag	UNP Q12306
D	790	SER	-	expression tag	UNP Q12306
D	791	SER	-	expression tag	UNP Q12306
D	792	GLY	-	expression tag	UNP Q12306
D	793	HIS	-	expression tag	UNP Q12306
D	794	ILE	-	expression tag	UNP Q12306
D	795	GLU	-	expression tag	UNP Q12306
D	796	GLY	-	expression tag	UNP Q12306
D	797	ARG	-	expression tag	UNP Q12306
D	798	HIS	-	expression tag	UNP Q12306
D	799	MET	-	expression tag	UNP Q12306
D	800	ALA	-	expression tag	UNP Q12306
D	801	SER	-	expression tag	UNP Q12306
D	900	SER	-	linker	UNP Q12306
D	901	MET	-	linker	UNP Q12306

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	207	Total O 207 207	0	0
3	B	229	Total O 229 229	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	41	Total	O	0	0
			41	41		
3	D	47	Total	O	0	0
			47	47		

PHE
LYS
ILE
LYS
LYS
THR
THR
PRO
LEU
ARG
ARG
LEU
MET
GLU
ALA
PHE
ALA
LYS
ARG
GLN
GLY
LYS
GLU
MET
ASP
SER
LEU
ARG
PHE
LEU
TYR
ASP
GLY
ILE
ARG
ILE
GLN
ALA
ASP
GLN
THR
PRO
GLU
ASP
LEU
ASP
MET
GLU
ASN
ASP
ASN
ILE
ILE
GLU
ALA
HIS
ARG
GLU
GLN
ILE

GLY
GLY
SER
MET
1902
P907
N911
N925
1930
1933

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	78.19Å 93.88Å 92.46Å 90.00° 110.81° 90.00°	Depositor
Resolution (Å)	43.21 – 1.86 43.21 – 1.86	Depositor EDS
% Data completeness (in resolution range)	99.9 (43.21-1.86) 99.9 (43.21-1.86)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 1.86Å)	Xtriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, R_{free}	0.183 , 0.209 0.183 , 0.209	Depositor DCC
R_{free} test set	2531 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	26.3	Xtriage
Anisotropy	0.355	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4239	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.28% 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	0.29	0/1734	0.45	0/2342
1	B	0.33	0/1577	0.48	0/2131
2	C	0.32	0/208	0.50	0/277
2	D	0.34	0/259	0.47	0/348
All	All	0.31	0/3778	0.47	0/5098

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	1705	0	1732	23	0
1	B	1551	0	1581	16	0
2	C	205	0	205	4	0
2	D	254	0	248	7	0
3	A	207	0	0	11	4
3	B	229	0	0	8	3
3	C	41	0	0	1	0
3	D	47	0	0	2	0
All	All	4239	0	3766	44	5

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.

The worst 5 of 44 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:135:SER:OG	3:A:301:HOH:O	1.91	0.89
3:B:523:HOH:O	2:D:925:MET:SD	2.31	0.87
1:B:205:MET:SD	3:D:1030:HOH:O	2.34	0.85
1:B:129:ASP:OD1	3:B:301:HOH:O	2.04	0.74
1:A:117:GLU:OE1	3:A:302:HOH:O	2.04	0.74

All (5) 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)
3:A:410:HOH:O	3:A:471:HOH:O[2_555]	1.56	0.64
3:A:443:HOH:O	3:B:461:HOH:O[4_555]	1.60	0.60
3:B:491:HOH:O	3:B:498:HOH:O[2_656]	1.78	0.42
3:A:507:HOH:O	3:B:521:HOH:O[3_555]	1.89	0.31
3:A:411:HOH:O	3:A:482:HOH:O[2_656]	2.07	0.13

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	214/228 (94%)	210 (98%)	4 (2%)	0	100	100
1	B	193/228 (85%)	191 (99%)	2 (1%)	0	100	100
2	C	21/156 (14%)	19 (90%)	2 (10%)	0	100	100
2	D	30/156 (19%)	28 (93%)	2 (7%)	0	100	100
All	All	458/768 (60%)	448 (98%)	10 (2%)	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	189/198 (96%)	185 (98%)	4 (2%)	47	33
1	B	174/198 (88%)	172 (99%)	2 (1%)	65	57
2	C	22/137 (16%)	22 (100%)	0	100	100
2	D	27/137 (20%)	27 (100%)	0	100	100
All	All	412/670 (62%)	406 (98%)	6 (2%)	57	47

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	100	ARG
1	B	45	LEU
1	B	100	ARG
1	A	45	LEU
1	A	1	MET

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

Mol	Chain	Res	Type
1	B	44	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 oligosaccharides 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	218/228 (95%)	0.04	8 (3%) 45 49	19, 31, 51, 90	0
1	B	197/228 (86%)	-0.15	5 (2%) 58 62	18, 28, 47, 77	0
2	C	25/156 (16%)	0.58	2 (8%) 18 19	21, 30, 47, 62	0
2	D	32/156 (20%)	0.07	0 100 100	19, 27, 51, 55	0
All	All	472/768 (61%)	-0.01	15 (3%) 50 54	18, 29, 51, 90	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	160	PRO	3.7
1	A	0	SER	3.3
2	C	905	ILE	2.9
1	A	216	ALA	2.8
1	B	158	ASP	2.8

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 oligosaccharides 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.