



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 04:31 PM UTC

PDB ID : 4ES9 / pdb_00004es9
Title : Crystal Structure of the adhesin domain of Epf from Streptococcus pyogenes in P21
Authors : Linke, C.; Siemens, N.; Kreikemeyer, B.; Baker, E.N.
Deposited on : 2012-04-23
Resolution : 2.00 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 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

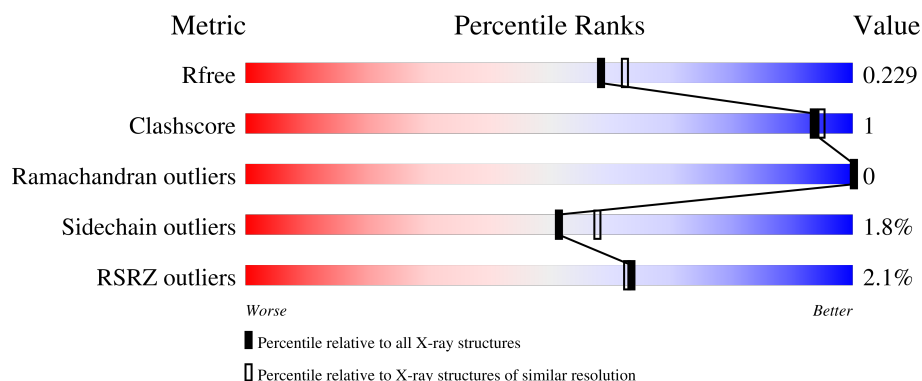
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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	320	0% 88% 8%
1	B	320	7% 85% 8% 6%
1	C	320	89% 5% 6%
1	D	320	87% 7% 7%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10563 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 Epf.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	1	0
			2344	1478	396	468	2			
1	B	300	Total	C	N	O	S	0	4	0
			2395	1511	403	479	2			
1	C	301	Total	C	N	O	S	0	2	0
			2389	1509	402	476	2			
1	D	299	Total	C	N	O	S	0	7	0
			2400	1517	402	479	2			

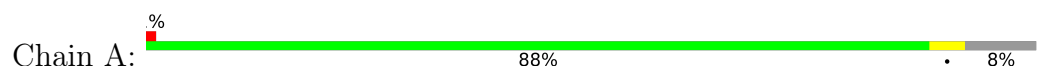
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	259	Total	O	0	0
			259	259		
2	B	248	Total	O	0	0
			248	248		
2	C	238	Total	O	0	0
			238	238		
2	D	290	Total	O	0	0
			290	290		

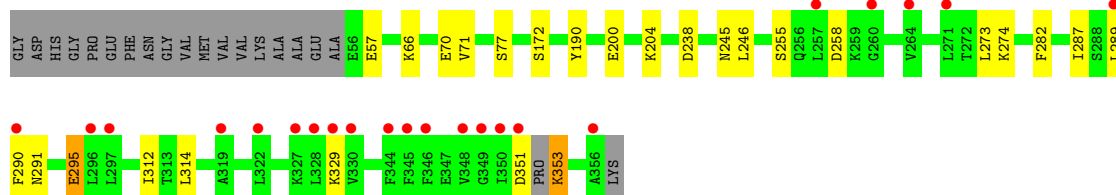
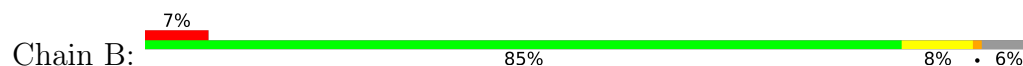
3 Residue-property plots [i](#)

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: Epf



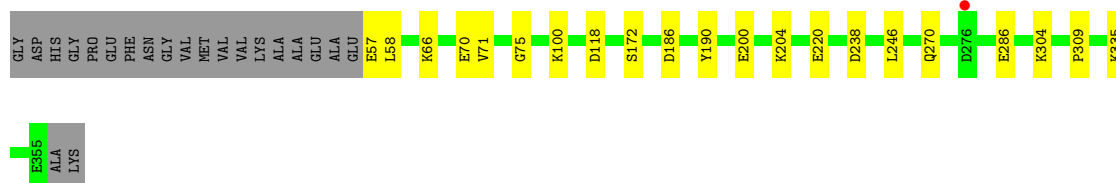
• Molecule 1: Epf



• Molecule 1: Epf



• Molecule 1: Epf



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.08Å 117.13Å 73.04Å 90.00° 106.46° 90.00°	Depositor
Resolution (Å)	19.73 – 2.00 19.73 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (19.73-2.00) 98.3 (19.73-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.06 (at 2.01Å)	Xtriage
Refinement program	BUSTER-TNT, BUSTER 2.8.0	Depositor
R, R_{free}	0.172 , 0.215 (Not available) , 0.229	Depositor DCC
R_{free} test set	3809 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	24.5	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 45.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.038 for l,-k,h	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10563	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.95% 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

5.1 Standard geometry

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	0.84	1/2386 (0.0%)	1.10	4/3205 (0.1%)
1	B	0.82	0/2444	1.15	10/3279 (0.3%)
1	C	0.80	0/2434	1.09	3/3267 (0.1%)
1	D	0.82	0/2460	1.10	3/3303 (0.1%)
All	All	0.82	1/9724 (0.0%)	1.11	20/13054 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	94	ILE	CA-C	5.68	1.57	1.52

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	186	ASP	CA-CB-CG	6.97	119.57	112.60
1	C	74	ASP	CA-CB-CG	6.52	119.12	112.60
1	B	70	GLU	CA-C-N	6.50	128.98	120.60
1	B	70	GLU	C-N-CA	6.50	128.98	120.60
1	A	70	GLU	CA-C-N	6.36	128.80	120.60
1	A	70	GLU	C-N-CA	6.36	128.80	120.60
1	A	245	ASN	CA-CB-CG	6.25	118.85	112.60
1	B	295	GLU	CA-C-N	5.95	129.82	121.02
1	B	295	GLU	C-N-CA	5.95	129.82	121.02
1	D	238	ASP	CA-CB-CG	5.57	118.17	112.60
1	D	186	ASP	CA-CB-CG	5.48	118.08	112.60
1	B	245	ASN	CA-CB-CG	5.45	118.05	112.60
1	C	238	ASP	CA-CB-CG	5.43	118.03	112.60
1	D	118	ASP	CA-CB-CG	5.40	118.00	112.60
1	A	196	ILE	N-CA-CB	5.32	117.07	111.00
1	B	77	SER	CA-C-N	5.28	125.94	120.60
1	B	77	SER	C-N-CA	5.28	125.94	120.60
1	B	238	ASP	CA-CB-CG	5.14	117.74	112.60
1	B	57	GLU	CA-C-N	5.02	130.04	121.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	57	GLU	C-N-CA	5.02	130.04	121.35

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	2344	0	2346	3	0
1	B	2395	0	2409	10	0
1	C	2389	0	2409	6	0
1	D	2400	0	2427	7	0
2	A	259	0	0	0	0
2	B	248	0	0	0	0
2	C	238	0	0	2	0
2	D	290	0	0	1	0
All	All	10563	0	9591	26	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 (26) 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:223:MET:HE3	2:C:629:HOH:O	1.81	0.80
1:B:200:GLU:HB2	1:B:204:LYS:HE3	1.82	0.61
1:D:200:GLU:HB2	1:D:204:LYS:HE2	1.83	0.60
1:B:258:ASP:HB2	1:B:274:LYS:HB2	1.87	0.56
1:B:290:PHE:HB2	1:B:329[B]:LYS:HB2	1.91	0.53
1:D:270[A]:GLN:NE2	1:D:309:PRO:HG2	2.24	0.52
1:C:225:LEU:HD23	1:C:246[B]:LEU:HD23	1.92	0.51
1:B:273:LEU:HD11	1:B:287:ILE:HD11	1.93	0.50
1:D:70:GLU:HB2	1:D:100:LYS:HD3	1.94	0.49
1:B:290:PHE:HB2	1:B:329[A]:LYS:HB3	1.97	0.47
1:A:200:GLU:HB2	1:A:204:LYS:HE3	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:66:LYS:HG3	1:D:75:GLY:HA3	1.99	0.45
1:D:286:GLU:HG3	1:D:335[A]:LYS:HE3	1.98	0.45
1:A:316:GLU:OE1	1:A:320:LYS:HE2	2.17	0.45
1:D:172:SER:HB3	1:D:190:TYR:CZ	2.52	0.44
1:B:66:LYS:HD2	2:C:671:HOH:O	2.16	0.43
1:B:172:SER:HB3	1:B:190:TYR:CZ	2.54	0.43
1:C:94:ILE:HD11	1:C:99:VAL:HA	2.00	0.43
1:B:351:ASP:C	1:B:353:LYS:HB3	2.43	0.43
1:C:148:ASP:HB3	1:C:281:ASN:HB3	2.00	0.43
1:D:304:LYS:NZ	2:D:644:HOH:O	2.49	0.42
1:B:291:ASN:HD21	1:B:295:GLU:HB2	1.85	0.42
1:A:148:ASP:HB3	1:A:281:ASN:HB3	2.02	0.41
1:B:255:SER:HB3	1:B:282:PHE:CD1	2.56	0.41
1:C:140:VAL:HG11	1:C:246[A]:LEU:HD13	2.03	0.41
1:C:172:SER:HB3	1:C:190:TYR:CZ	2.55	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	Percentiles	
1	A	295/320 (92%)	290 (98%)	5 (2%)	0	100	100
1	B	300/320 (94%)	294 (98%)	6 (2%)	0	100	100
1	C	301/320 (94%)	294 (98%)	7 (2%)	0	100	100
1	D	304/320 (95%)	298 (98%)	6 (2%)	0	100	100
All	All	1200/1280 (94%)	1176 (98%)	24 (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	264/282 (94%)	260 (98%)	4 (2%)	57	64
1	B	271/282 (96%)	265 (98%)	6 (2%)	45	50
1	C	270/282 (96%)	265 (98%)	5 (2%)	50	56
1	D	274/282 (97%)	269 (98%)	5 (2%)	51	58
All	All	1079/1128 (96%)	1059 (98%)	20 (2%)	51	56

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	VAL
1	A	220	GLU
1	A	246	LEU
1	A	335	LYS
1	B	71	VAL
1	B	246	LEU
1	B	289	LEU
1	B	312	ILE
1	B	314	LEU
1	B	353	LYS
1	C	71	VAL
1	C	246[A]	LEU
1	C	246[B]	LEU
1	C	286	GLU
1	C	355	GLU
1	D	57	GLU
1	D	58	LEU
1	D	71	VAL
1	D	220	GLU
1	D	246	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	108	ASN
1	A	205	GLN
1	A	245	ASN
1	A	268	ASN
1	A	324	ASN
1	B	108	ASN
1	B	270	GLN
1	B	324	ASN
1	C	120	GLN
1	C	159	GLN
1	C	236	GLN
1	C	324	ASN
1	D	64	ASN
1	D	159	GLN
1	D	268	ASN
1	D	324	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 ⓘ

There are no ligands in this entry.

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	296/320 (92%)	-0.35	2 (0%) 84 84	15, 26, 52, 67	1 (0%)
1	B	300/320 (93%)	0.06	22 (7%) 21 19	15, 28, 63, 76	4 (1%)
1	C	301/320 (94%)	-0.23	0 100 100	14, 29, 49, 61	2 (0%)
1	D	299/320 (93%)	-0.47	1 (0%) 90 89	13, 25, 42, 56	7 (2%)
All	All	1196/1280 (93%)	-0.25	25 (2%) 63 63	13, 27, 55, 76	14 (1%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	322	LEU	4.5
1	B	289	LEU	3.9
1	B	296	LEU	3.8
1	B	297	LEU	3.3
1	B	328	LEU	3.2
1	B	349	GLY	3.2
1	B	330	VAL	3.2
1	B	346	PHE	3.0
1	B	348	VAL	2.9
1	B	264	VAL	2.9
1	B	350	ILE	2.8
1	B	290	PHE	2.8
1	B	260	GLY	2.8
1	A	305	ASP	2.7
1	B	319	ALA	2.6
1	B	356	ALA	2.6
1	B	327	LYS	2.4
1	B	344	PHE	2.4
1	D	276	ASP	2.4
1	B	271	LEU	2.3
1	B	345	PHE	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	350	ILE	2.1
1	B	329[A]	LYS	2.1
1	B	351	ASP	2.1
1	B	257	LEU	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 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.