



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 02:33 PM UTC

PDB ID : 5WTB / pdb_00005wtb
Title : Complex Structure of Staphylococcus aureus SdrE with human complement factor H
Authors : Wu, M.; Zhang, Y.; Hang, T.; Wang, C.; Yang, Y.; Zang, J.; Zhang, M.; Zhang, X.
Deposited on : 2016-12-10
Resolution : 3.30 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

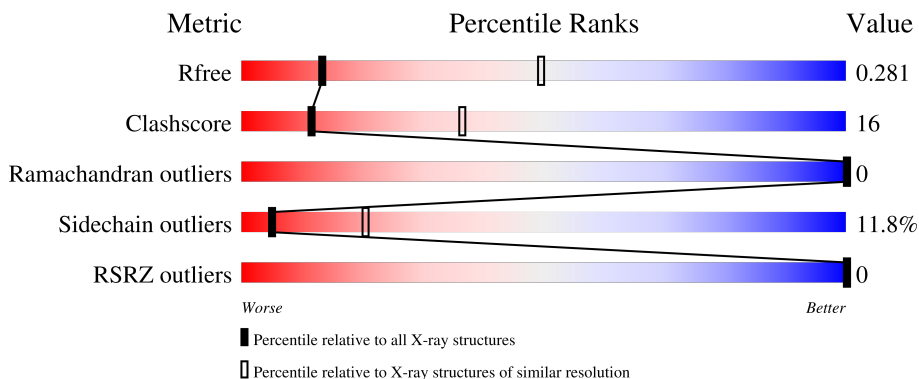
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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.




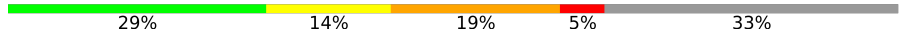

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1169 (3.32-3.28)
Clashscore	190562	1209 (3.32-3.28)
Ramachandran outliers	187476	1188 (3.32-3.28)
Sidechain outliers	187428	1187 (3.32-3.28)
RSRZ outliers	180081	1169 (3.32-3.28)

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	338	
1	B	338	
1	C	338	
1	D	338	
2	E	21	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	21	
2	G	21	
2	H	21	

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 10427 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 Serine-aspartate repeat-containing protein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	313	2447	1524	397	522	4	0	0	0
1	A	313	2447	1524	397	522	4	0	0	0
1	B	319	2491	1549	405	533	4	0	0	0
1	C	323	2528	1575	410	539	4	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	489	MET	ILE	engineered mutation	UNP Q932F7
D	600	LEU	-	expression tag	UNP Q932F7
D	601	GLU	-	expression tag	UNP Q932F7
D	602	HIS	-	expression tag	UNP Q932F7
D	603	HIS	-	expression tag	UNP Q932F7
D	604	HIS	-	expression tag	UNP Q932F7
D	605	HIS	-	expression tag	UNP Q932F7
D	606	HIS	-	expression tag	UNP Q932F7
D	607	HIS	-	expression tag	UNP Q932F7
A	489	MET	ILE	engineered mutation	UNP Q932F7
A	600	LEU	-	expression tag	UNP Q932F7
A	601	GLU	-	expression tag	UNP Q932F7
A	602	HIS	-	expression tag	UNP Q932F7
A	603	HIS	-	expression tag	UNP Q932F7
A	604	HIS	-	expression tag	UNP Q932F7
A	605	HIS	-	expression tag	UNP Q932F7
A	606	HIS	-	expression tag	UNP Q932F7
A	607	HIS	-	expression tag	UNP Q932F7
B	489	MET	ILE	engineered mutation	UNP Q932F7
B	600	LEU	-	expression tag	UNP Q932F7
B	601	GLU	-	expression tag	UNP Q932F7

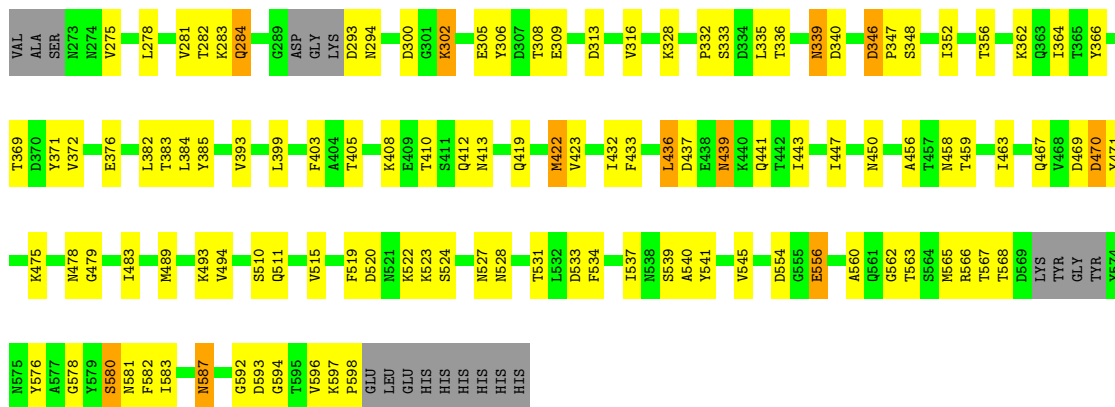
Continued on next page...

Continued from previous page...

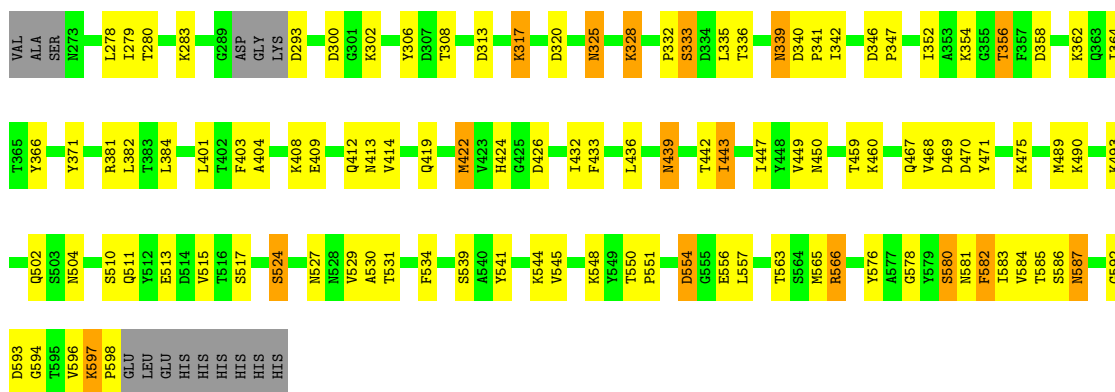
Chain	Residue	Modelled	Actual	Comment	Reference
B	602	HIS	-	expression tag	UNP Q932F7
B	603	HIS	-	expression tag	UNP Q932F7
B	604	HIS	-	expression tag	UNP Q932F7
B	605	HIS	-	expression tag	UNP Q932F7
B	606	HIS	-	expression tag	UNP Q932F7
B	607	HIS	-	expression tag	UNP Q932F7
C	489	MET	ILE	engineered mutation	UNP Q932F7
C	600	LEU	-	expression tag	UNP Q932F7
C	601	GLU	-	expression tag	UNP Q932F7
C	602	HIS	-	expression tag	UNP Q932F7
C	603	HIS	-	expression tag	UNP Q932F7
C	604	HIS	-	expression tag	UNP Q932F7
C	605	HIS	-	expression tag	UNP Q932F7
C	606	HIS	-	expression tag	UNP Q932F7
C	607	HIS	-	expression tag	UNP Q932F7

- Molecule 2 is a protein called Peptide from Complement factor H.

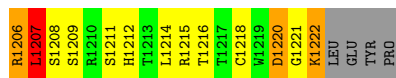
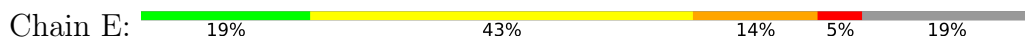
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	17	Total	C	N	O	S	0	0	0
			139	83	30	25	1			
2	F	14	Total	C	N	O	S	0	0	0
			118	71	26	20	1			
2	G	14	Total	C	N	O	S	0	0	0
			118	71	26	20	1			
2	H	17	Total	C	N	O	S	0	0	0
			139	83	30	25	1			



• Molecule 1: Serine-aspartate repeat-containing protein E



• Molecule 2: Peptide from Complement factor H



• Molecule 2: Peptide from Complement factor H

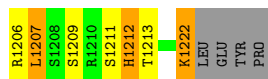


• Molecule 2: Peptide from Complement factor H



• Molecule 2: Peptide from Complement factor H

Chain H:  48% 19% 14% 19%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	117.50Å 117.50Å 154.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.79 – 3.30 49.79 – 3.30	Depositor EDS
% Data completeness (in resolution range)	95.2 (49.79-3.30) 95.2 (49.79-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 3.33Å)	Xtrriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.222 , 0.281 0.223 , 0.281	Depositor DCC
R_{free} test set	1580 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	94.2	Xtrriage
Anisotropy	0.052	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 102.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.418 for k,h,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10427	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% 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	0.70	0/2485	0.91	3/3372 (0.1%)
1	B	0.68	0/2528	0.91	2/3431 (0.1%)
1	C	0.68	0/2568	0.90	2/3486 (0.1%)
1	D	0.72	0/2485	0.93	1/3372 (0.0%)
2	E	0.74	0/141	1.25	2/188 (1.1%)
2	F	0.88	0/120	1.09	0/161
2	G	0.88	0/120	1.21	1/161 (0.6%)
2	H	0.80	0/141	1.00	0/188
All	All	0.70	0/10588	0.92	11/14359 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	440	LYS	N-CA-C	6.97	121.61	113.18
2	G	1217	THR	N-CA-C	6.69	118.56	108.46
1	B	439	ASN	N-CA-C	6.00	118.35	111.02
2	E	1207	LEU	N-CA-C	5.95	118.65	109.07
1	A	331	ILE	CA-C-N	-5.54	113.31	119.19

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	2447	0	2356	75	0
1	B	2491	0	2400	97	0
1	C	2528	0	2435	89	0
1	D	2447	0	2356	64	0
2	E	139	0	138	24	0
2	F	118	0	118	16	0
2	G	118	0	118	13	0
2	H	139	0	138	13	0
All	All	10427	0	10059	328	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 16.

The worst 5 of 328 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:493:LYS:HG2	1:C:541:TYR:CE1	1.68	1.28
1:B:493:LYS:HG2	1:B:541:TYR:CE1	1.68	1.26
1:D:493:LYS:HG2	1:D:541:TYR:CE2	1.82	1.13
1:C:493:LYS:HG2	1:C:541:TYR:HE1	0.97	1.10
1:B:493:LYS:HG2	1:B:541:TYR:HE1	1.18	0.98

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	307/338 (91%)	278 (91%)	29 (9%)	0	100	100
1	B	313/338 (93%)	285 (91%)	28 (9%)	0	100	100
1	C	319/338 (94%)	287 (90%)	32 (10%)	0	100	100
1	D	307/338 (91%)	275 (90%)	32 (10%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	15/21 (71%)	14 (93%)	1 (7%)	0	100	100
2	F	12/21 (57%)	10 (83%)	2 (17%)	0	100	100
2	G	12/21 (57%)	11 (92%)	1 (8%)	0	100	100
2	H	15/21 (71%)	15 (100%)	0	0	100	100
All	All	1300/1436 (90%)	1175 (90%)	125 (10%)	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	279/301 (93%)	244 (88%)	35 (12%)	4	19
1	B	285/301 (95%)	261 (92%)	24 (8%)	10	34
1	C	288/301 (96%)	256 (89%)	32 (11%)	6	22
1	D	279/301 (93%)	248 (89%)	31 (11%)	6	22
2	E	16/20 (80%)	10 (62%)	6 (38%)	0	0
2	F	14/20 (70%)	9 (64%)	5 (36%)	0	1
2	G	14/20 (70%)	9 (64%)	5 (36%)	0	1
2	H	16/20 (80%)	13 (81%)	3 (19%)	1	7
All	All	1191/1284 (93%)	1050 (88%)	141 (12%)	5	20

5 of 141 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	557	LEU
1	C	583	ILE
2	F	1208	SER
1	A	439	ASN
1	A	436	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	274	ASN
1	C	424	HIS
1	C	419	GLN
1	C	430	GLN
1	A	339	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](#)

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	313/338 (92%)	-1.35	0 100 100	51, 85, 109, 128	18 (5%)
1	B	319/338 (94%)	-1.28	0 100 100	55, 91, 128, 154	19 (5%)
1	C	323/338 (95%)	-1.28	0 100 100	51, 92, 121, 132	25 (7%)
1	D	313/338 (92%)	-1.29	0 100 100	44, 85, 105, 126	20 (6%)
2	E	17/21 (80%)	-1.30	0 100 100	66, 82, 98, 98	1 (5%)
2	F	14/21 (66%)	-0.88	0 100 100	71, 81, 102, 104	2 (14%)
2	G	14/21 (66%)	-1.13	0 100 100	51, 81, 92, 93	3 (21%)
2	H	17/21 (80%)	-1.21	0 100 100	67, 80, 99, 104	1 (5%)
All	All	1330/1436 (92%)	-1.29	0 100 100	44, 87, 119, 154	89 (6%)

There are no RSRZ outliers to report.

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.