W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 2, 2022 – 06:05 PM EDT

|		
PDB ID	:	7SUE
Title	:	Crystal Structure of Human Fab S24-188 in the complex with the N-teminal Domain of Nucleocapsid protein from SARS CoV-2
Authors	:	Kim, Y.; Maltseva, N.; Tesar, C.; Jedrzejczak, R.; Dugan, H.; Stamper, C.; Wilson, P.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2021-11-17
Resolution	:	2.90 Å(reported)

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 2.90 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.
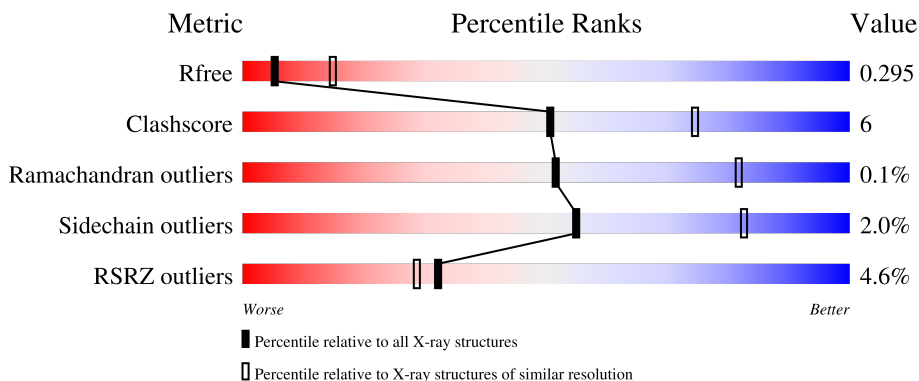
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	216	84% ... 14% ·
1	E	216	3% / 38% ... 13% ... 50%
1	G	216	6% / 44% ... 6% ... 50%
1	L	216	2% / 84% ... 14% ·
2	B	231	86% ... 12% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	231	3% / 46% / 10% / 44%
2	H	231	83% / 14% / ··
2	I	231	6% / 46% / 9% / 45%
3	C	130	5% / 76% / 12% / · / 12%
3	D	130	2% / 72% / 16% / 12%
3	J	130	11% / 75% / 13% / 12%
3	K	130	9% / 72% / 16% / · / 12%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 13707 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 S24-188 Fab Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	212	Total 1577	C 984	N 263	O 325	S 5	0	0	0
1	A	212	Total 1577	C 984	N 263	O 325	S 5	0	0	0
1	E	109	Total 799	C 495	N 133	O 168	S 3	0	0	0
1	G	109	Total 799	C 495	N 133	O 168	S 3	0	0	0

- Molecule 2 is a protein called S24-188 Fab Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	227	Total 1699	C 1076	N 283	O 331	S 9	0	0	0
2	B	227	Total 1699	C 1076	N 283	O 331	S 9	0	0	0
2	F	129	Total 995	C 630	N 167	O 192	S 6	0	0	0
2	I	128	Total 983	C 623	N 166	O 188	S 6	0	0	0

- Molecule 3 is a protein called Nucleoprotein.

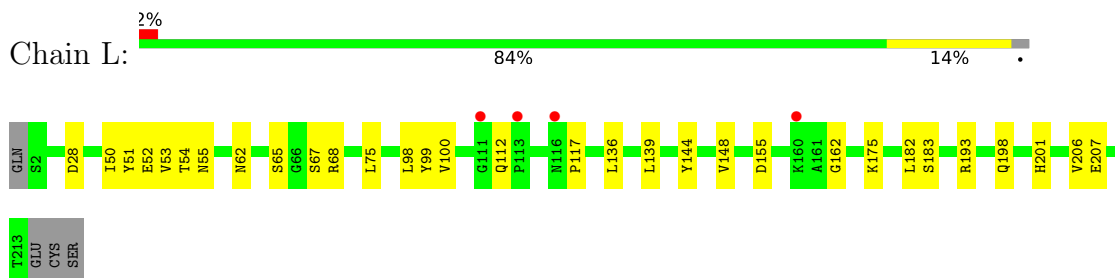| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	115	Total 896	C 571	N 156	O 169	0	0	0
3	D	115	Total 896	C 571	N 156	O 169	0	0	0
3	J	114	Total 891	C 567	N 158	O 166	0	0	0
3	K	115	Total 896	C 571	N 156	O 169	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

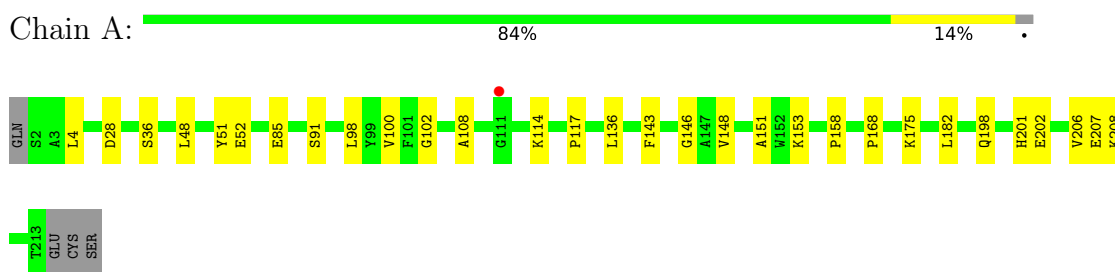| Chain	Residue	Modelled	Actual	Comment	Reference
C	44	SER	-	expression tag	UNP P0DTC9
C	45	ASN	-	expression tag	UNP P0DTC9
C	46	MET	-	expression tag	UNP P0DTC9
D	44	SER	-	expression tag	UNP P0DTC9
D	45	ASN	-	expression tag	UNP P0DTC9
D	46	MET	-	expression tag	UNP P0DTC9
J	44	SER	-	expression tag	UNP P0DTC9
J	45	ASN	-	expression tag	UNP P0DTC9
J	46	MET	-	expression tag	UNP P0DTC9
K	44	SER	-	expression tag	UNP P0DTC9
K	45	ASN	-	expression tag	UNP P0DTC9
K	46	MET	-	expression tag	UNP P0DTC9

3 Residue-property plots ⓘ

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.
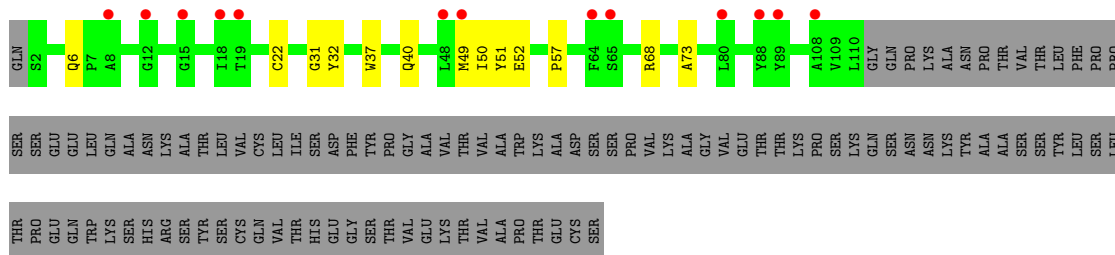
- Molecule 1: S24-188 Fab Light chain

Chain L:

- Molecule 1: S24-188 Fab Light chain

Chain A:

- Molecule 1: S24-188 Fab Light chain
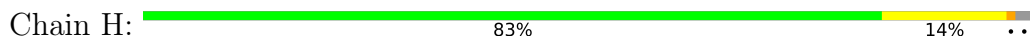
Chain E:

- Molecule 1: S24-188 Fab Light chain

Chain G:

• Molecule 2: S24-188 Fab Heavy chain

Chain H:

• Molecule 2: S24-188 Fab Heavy chain

Chain B:

• Molecule 2: S24-188 Fab Heavy chain

Chain F:

• Molecule 2: S24-188 Fab Heavy chain

Chain I:

GLY PRO SER VAL PHE PRO LEU ALA PRO CYS SER ARG SER THR SER GLY GLY THR ALA ALA LEU GLY CYS LEU VAL LYS ASP TYR PHE PRO GLU PRO VAL THR VAL SER TRP ASN SER GLY ALA LEU THR SER GLY VAL HIS THR PHE PRO ALA VAL LEU GLN SER SER GLY LEU TYR SER

LEU SER SER VAL VAL THR VAL PRO SER SER SER LEU GLY THR GLN THR TYR ILE CYS ASN VAL ASN HIS LYS PRO SER ASN THR LYS VAL ASP LYS LYS VAL GLU PRO LYS

- Molecule 3: Nucleoprotein

Chain C:

5%

76% 12% · 12%

SER ASN MET N47 A55 L56 T57 L64 R68 S78 D82 Y87 T91 ARG ARG ILE ARG GLY GLY ASP GLY LYS MET LYS D103 R107 W108 Y109 E118 L121 I130 I131 L139 N140 K143 L159 Q160 L161 T165 G170 F171 Y172 ALA

- Molecule 3: Nucleoprotein

Chain D:

2%

72% 16% 12%

SER ASN MET N47 F53 T54 A55 H59 L64 R68 P73 I74 N75 D82 Y87 T91 ARG ARG ILE ARG GLY GLY ASP GLY LYS MET LYS D103 R107 W108 Y109 F110 G114 Y123 W132 P142 K143 R149 N154 L159 Q163 G170

F171 Y172 ALA

- Molecule 3: Nucleoprotein

Chain J:

11%

75% 13% 12%

SER ASN MET N47 A55 Q58 H59 L64 K65 F66 P67 R68 G69 Q70 G71 V72 N77 D82 Y86 Y87 T91 R92 ARG ARG ILE ARG GLY GLY ASP GLY LYS MET LYS ASP LEU S105 P106 Y112 E118 L121 P122 Y123 I131 W132 V133 A134 T135 E136 G137

N140 P162 Q163 K169 Y172 ALA

- Molecule 3: Nucleoprotein

Chain K:

9%

72% 16% · 12%

SER ASN MET N47 A55 L56 T57 Q58 H59 G60 K61 F66 P67 R68 V72 D82 Q83 Y87 R88 R89 A90 T91 ARG ARG ILE ARG GLY GLY ASP GLY LYS MET LYS D103 L104 S105 P106 R107 W108 Y109 Y112 E118 L121 N126 G129 T135 L139

L161 T166 G170 F171 Y172 ALA

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	71.31Å 78.50Å 127.15Å 97.60° 89.96° 90.02°	Depositor
Resolution (Å)	39.20 – 2.90 39.20 – 2.90	Depositor EDS
% Data completeness (in resolution range)	91.5 (39.20-2.90) 91.3 (39.20-2.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 2.90Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.262 , 0.292 0.265 , 0.295	Depositor DCC
R_{free} test set	2724 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	89.1	Xtriage
Anisotropy	0.506	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 26.9	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.440 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	13707	wwPDB-VP
Average B, all atoms (Å^2)	119.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.67% 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	0.26	0/1615	0.48	0/2205
1	E	0.25	0/816	0.49	0/1111
1	G	0.27	0/816	0.50	0/1111
1	L	0.26	0/1615	0.48	0/2205
2	B	0.25	0/1740	0.49	0/2366
2	F	0.25	0/1018	0.54	0/1379
2	H	0.27	0/1740	0.51	0/2366
2	I	0.28	0/1006	0.51	0/1361
3	C	0.25	0/923	0.52	0/1261
3	D	0.24	0/923	0.49	0/1261
3	J	0.25	0/918	0.51	0/1253
3	K	0.27	0/923	0.57	0/1261
All	All	0.26	0/14053	0.50	0/19140

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	1577	0	1519	17	0
1	E	799	0	755	18	0
1	G	799	0	755	12	0
1	L	1577	0	1519	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1699	0	1663	12	0
2	F	995	0	961	16	0
2	H	1699	0	1663	23	0
2	I	983	0	951	14	0
3	C	896	0	855	11	0
3	D	896	0	855	13	0
3	J	891	0	853	11	0
3	K	896	0	855	13	0
All	All	13707	0	13204	154	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 6.

The worst 5 of 154 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)
3:K:72:VAL:HG11	3:K:83:GLN:HB3	1.81	0.61
2:H:6:GLN:H	2:H:122:GLN:HE22	1.49	0.59
2:I:41:PRO:HD3	2:I:92:ALA:HA	1.84	0.59
1:E:51:TYR:CZ	1:E:55:ASN:HB3	2.38	0.59
3:J:118:GLU:HB3	3:J:121:LEU:HD22	1.85	0.58

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	210/216 (97%)	206 (98%)	4 (2%)	0	100	100
1	E	107/216 (50%)	103 (96%)	4 (4%)	0	100	100
1	G	107/216 (50%)	103 (96%)	4 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	210/216 (97%)	206 (98%)	4 (2%)	0	100	100
2	B	221/231 (96%)	216 (98%)	5 (2%)	0	100	100
2	F	125/231 (54%)	118 (94%)	7 (6%)	0	100	100
2	H	221/231 (96%)	217 (98%)	4 (2%)	0	100	100
2	I	124/231 (54%)	118 (95%)	6 (5%)	0	100	100
3	C	111/130 (85%)	109 (98%)	2 (2%)	0	100	100
3	D	111/130 (85%)	109 (98%)	2 (2%)	0	100	100
3	J	110/130 (85%)	109 (99%)	1 (1%)	0	100	100
3	K	111/130 (85%)	107 (96%)	3 (3%)	1 (1%)	17	48
All	All	1768/2308 (77%)	1721 (97%)	46 (3%)	1 (0%)	51	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	K	89	ARG

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	178/183 (97%)	174 (98%)	4 (2%)	52	81
1	E	89/183 (49%)	88 (99%)	1 (1%)	73	92
1	G	89/183 (49%)	89 (100%)	0	100	100
1	L	178/183 (97%)	175 (98%)	3 (2%)	60	86
2	B	188/192 (98%)	182 (97%)	6 (3%)	39	73
2	F	105/192 (55%)	103 (98%)	2 (2%)	57	84
2	H	188/192 (98%)	181 (96%)	7 (4%)	34	68
2	I	103/192 (54%)	101 (98%)	2 (2%)	57	84
3	C	92/103 (89%)	91 (99%)	1 (1%)	73	92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	D	92/103 (89%)	92 (100%)	0	100	100
3	J	91/103 (88%)	90 (99%)	1 (1%)	73	92
3	K	92/103 (89%)	90 (98%)	2 (2%)	52	81
All	All	1485/1912 (78%)	1456 (98%)	29 (2%)	55	82

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	29	PHE
3	K	61	LYS
2	B	102	PHE
2	I	87	ARG
2	B	54	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	6	GLN
3	J	126	ASN
3	K	126	ASN
1	L	40	GLN
1	L	39	GLN

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 monosaccharides 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(Å²)	Q<0.9
1	A	212/216 (98%)	0.08	1 (0%)	91	91	88, 108, 126, 134	0
1	E	109/216 (50%)	0.40	6 (5%)	25	21	116, 135, 146, 152	0
1	G	109/216 (50%)	0.60	13 (11%)	4	3	125, 142, 156, 170	0
1	L	212/216 (98%)	0.14	4 (1%)	66	65	83, 107, 128, 147	0
2	B	227/231 (98%)	0.08	1 (0%)	92	93	87, 99, 122, 136	0
2	F	129/231 (55%)	0.47	8 (6%)	20	16	118, 139, 158, 167	0
2	H	227/231 (98%)	0.09	1 (0%)	92	93	83, 100, 117, 130	0
2	I	128/231 (55%)	0.38	13 (10%)	6	5	106, 141, 165, 178	0
3	C	115/130 (88%)	0.39	7 (6%)	21	17	94, 112, 147, 157	0
3	D	115/130 (88%)	0.08	3 (2%)	56	52	87, 96, 123, 150	0
3	J	114/130 (87%)	0.62	14 (12%)	4	3	115, 144, 166, 171	0
3	K	115/130 (88%)	0.60	12 (10%)	6	5	113, 140, 176, 197	0
All	All	1812/2308 (78%)	0.27	83 (4%)	32	29	83, 116, 154, 197	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	12	GLY	7.7
2	I	129	SER	6.6
3	K	104	LEU	6.1
3	J	71	GLY	5.8
3	J	123	TYR	5.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

WORLDWIDE PDB PROTEIN DATA BANK

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.