Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 08:49 PM UTC

PDB ID	:	9KMO / pdb_00009kmo
Title	:	HR1 domain from HCoV-OC43 in complex with a pan-CoV inhibitor EK1
Authors	:	Yan, L.
Deposited on	:	2024-11-16
Resolution	:	2.70 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:

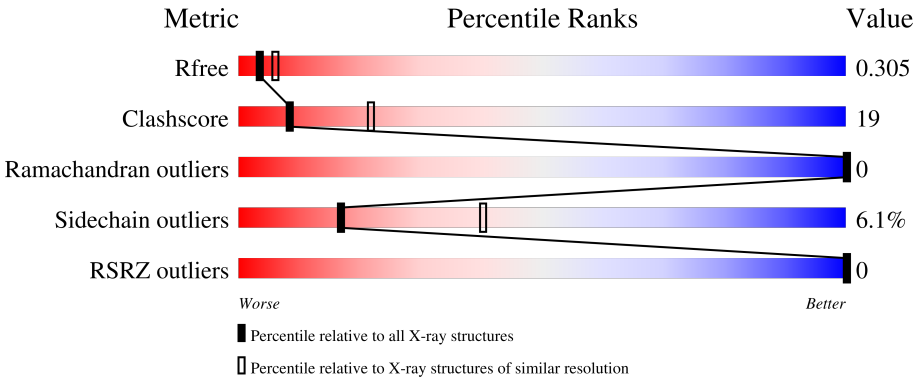
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.70 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	3538 (2.70-2.70)
Clashscore	190562	3843 (2.70-2.70)
Ramachandran outliers	187476	3778 (2.70-2.70)
Sidechain outliers	187428	3778 (2.70-2.70)
RSRZ outliers	180081	3538 (2.70-2.70)

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	70	64% 30%
1	B	70	67% 20% 9%
1	C	70	67% 23%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1550 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 Spike protein S2'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	68	Total	C	N	O	S	0	0	0
			530	336	85	108	1			
1	B	64	Total	C	N	O	S	0	0	0
			499	315	80	103	1			
1	C	67	Total	C	N	O	S	0	0	0
			521	331	83	106	1			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	ACE	-	acetylation	UNP P36334
A	25	ASP	TYR	conflict	UNP P36334
A	100	GLU	GLN	conflict	UNP P36334
A	101	TYR	VAL	conflict	UNP P36334
A	104	LYS	ASN	conflict	UNP P36334
A	105	LYS	ARG	conflict	UNP P36334
A	107	GLU	GLN	conflict	UNP P36334
A	112	LYS	VAL	conflict	UNP P36334
A	114	GLU	ASN	conflict	UNP P36334
A	115	GLU	GLN	conflict	UNP P36334
B	16	ACE	-	acetylation	UNP P36334
B	25	ASP	TYR	conflict	UNP P36334
B	100	GLU	GLN	conflict	UNP P36334
B	101	TYR	VAL	conflict	UNP P36334
B	104	LYS	ASN	conflict	UNP P36334
B	105	LYS	ARG	conflict	UNP P36334
B	107	GLU	GLN	conflict	UNP P36334
B	112	LYS	VAL	conflict	UNP P36334
B	114	GLU	ASN	conflict	UNP P36334
B	115	GLU	GLN	conflict	UNP P36334
C	16	ACE	-	acetylation	UNP P36334
C	25	ASP	TYR	conflict	UNP P36334
C	100	GLU	GLN	conflict	UNP P36334

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	101	TYR	VAL	conflict	UNP P36334
C	104	LYS	ASN	conflict	UNP P36334
C	105	LYS	ARG	conflict	UNP P36334
C	107	GLU	GLN	conflict	UNP P36334
C	112	LYS	VAL	conflict	UNP P36334
C	114	GLU	ASN	conflict	UNP P36334
C	115	GLU	GLN	conflict	UNP P36334

3 Residue-property plots (i)

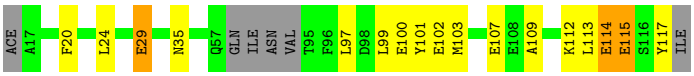
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Spike protein S2'

Chain A:

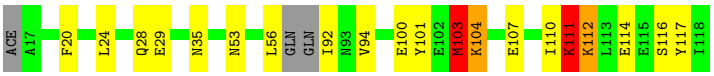
• Molecule 1: Spike protein S2'

Chain B:



• Molecule 1: Spike protein S2'

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	53.01Å 53.01Å 62.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.74 – 2.70 25.74 – 2.71	Depositor EDS
% Data completeness (in resolution range)	97.0 (25.74-2.70) 99.5 (25.74-2.71)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.73 (at 2.72Å)	Xtriage
Refinement program	PHENIX 1.16_3549, PHENIX 1.16_3549	Depositor
R, R_{free}	0.282 , 0.311 0.297 , 0.305	Depositor DCC
R_{free} test set	287 reflections (5.35%)	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.037	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 68.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	0.396 for -h,-k,l 0.340 for h,-h-k,-l 0.367 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	1550	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.19% 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

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	0.80	1/533 (0.2%)	0.97	1/719 (0.1%)
1	B	0.75	1/502 (0.2%)	1.00	4/676 (0.6%)
1	C	2.49	3/524 (0.6%)	1.36	11/707 (1.6%)
All	All	1.58	5/1559 (0.3%)	1.12	16/2102 (0.8%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	103	MET	C-N	-55.75	0.57	1.34
1	C	111	LYS	C-N	-8.73	1.21	1.33
1	C	104	LYS	C-N	7.19	1.43	1.34
1	A	109	ALA	C-O	-7.06	1.14	1.24
1	B	112	LYS	C-O	-5.26	1.17	1.24

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	104	LYS	O-C-N	-13.82	105.27	122.27
1	C	112	LYS	O-C-N	12.90	139.77	122.49
1	C	103	MET	CA-C-N	11.98	138.52	120.31
1	C	103	MET	C-N-CA	11.98	138.52	120.31
1	C	104	LYS	CA-C-N	9.59	134.88	120.31
1	C	104	LYS	C-N-CA	9.59	134.88	120.31
1	C	112	LYS	CA-C-N	-9.27	104.92	121.14
1	C	112	LYS	C-N-CA	-9.27	104.92	121.14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	111	LYS	O-C-N	-8.92	109.19	122.39
1	A	108	GLU	CB-CG-CD	6.80	124.17	112.60
1	B	114	GLU	CB-CG-CD	-6.39	101.74	112.60
1	B	115	GLU	N-CA-C	-5.93	106.29	112.93
1	B	102	GLU	CB-CG-CD	5.67	122.24	112.60
1	C	111	LYS	CA-C-N	5.44	131.34	121.92
1	C	111	LYS	C-N-CA	5.44	131.34	121.92
1	B	100	GLU	CB-CG-CD	5.18	121.41	112.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	103	MET	Mainchain
1	C	111	LYS	Mainchain

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	530	0	526	12	15
1	B	499	0	488	29	1
1	C	521	0	518	40	16
All	All	1550	0	1532	59	16

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 19.

All (59) 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:103:MET:C	1:C:104:LYS:CA	1.92	1.41
1:C:103:MET:CA	1:C:104:LYS:N	1.88	1.34
1:C:103:MET:O	1:C:104:LYS:N	1.57	1.29
1:B:101:TYR:CD1	1:C:111:LYS:NZ	2.09	1.20

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:101:TYR:CD1	1:C:114:GLU:OE2	2.04	1.11
1:B:101:TYR:CE1	1:C:114:GLU:OE2	2.07	1.07
1:B:101:TYR:HB3	1:C:111:LYS:NZ	1.69	1.05
1:B:101:TYR:CG	1:C:111:LYS:NZ	2.24	1.00
1:B:101:TYR:CB	1:C:111:LYS:HZ2	1.75	0.99
1:B:101:TYR:CD2	1:C:114:GLU:CD	2.41	0.98
1:B:101:TYR:CB	1:C:111:LYS:NZ	2.29	0.94
1:B:101:TYR:CD2	1:C:114:GLU:OE1	2.21	0.94
1:B:101:TYR:CE2	1:C:114:GLU:CD	2.51	0.88
1:B:101:TYR:CG	1:C:114:GLU:OE2	2.26	0.88
1:C:112:LYS:O	1:C:116:SER:OG	1.94	0.86
1:B:101:TYR:CZ	1:C:114:GLU:OE2	2.29	0.85
1:B:101:TYR:HB3	1:C:111:LYS:HZ1	1.49	0.76
1:A:56:LEU:HD13	1:A:92:ILE:HA	1.69	0.75
1:B:35:ASN:HD21	1:B:107:GLU:HG3	1.49	0.75
1:B:101:TYR:CG	1:C:114:GLU:CD	2.66	0.74
1:C:56:LEU:HD13	1:C:92:ILE:HA	1.69	0.74
1:A:114:GLU:HG3	1:A:117:TYR:HE2	1.54	0.72
1:B:101:TYR:HD1	1:C:111:LYS:NZ	1.80	0.72
1:C:103:MET:O	1:C:104:LYS:CA	2.24	0.67
1:C:103:MET:C	1:C:104:LYS:N	0.57	0.66
1:A:35:ASN:ND2	1:A:107:GLU:OE1	2.29	0.66
1:B:101:TYR:CE2	1:C:114:GLU:OE2	2.48	0.65
1:B:35:ASN:ND2	1:B:107:GLU:HG3	2.11	0.64
1:B:101:TYR:CG	1:C:114:GLU:OE1	2.50	0.63
1:B:115:GLU:O	1:B:115:GLU:HG3	1.98	0.63
1:B:101:TYR:CD2	1:C:114:GLU:OE2	2.48	0.62
1:C:35:ASN:HD21	1:C:107:GLU:HG3	1.65	0.61
1:C:53:ASN:HB2	1:C:94:VAL:HG21	1.82	0.61
1:C:28:GLN:HG2	1:C:110:ILE:HG23	1.83	0.61
1:A:53:ASN:HB2	1:A:94:VAL:HG21	1.82	0.60
1:B:101:TYR:HB3	1:C:111:LYS:HZ2	1.42	0.60
1:B:101:TYR:HB3	1:C:111:LYS:CE	2.38	0.52
1:A:100:GLU:O	1:A:104:LYS:HG3	2.10	0.51
1:C:103:MET:O	1:C:104:LYS:HA	2.12	0.49
1:C:111:LYS:NZ	1:C:114:GLU:OE2	2.47	0.48
1:C:56:LEU:HB3	1:C:92:ILE:HG12	1.97	0.47
1:A:56:LEU:HB3	1:A:92:ILE:HG12	1.97	0.47
1:C:100:GLU:O	1:C:104:LYS:CG	2.64	0.46
1:C:103:MET:N	1:C:104:LYS:N	2.59	0.46
1:B:101:TYR:CB	1:C:111:LYS:HZ1	2.12	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:101:TYR:N	1:A:101:TYR:CD1	2.81	0.45
1:A:28:GLN:HG2	1:A:110:ILE:HG23	1.98	0.45
1:B:101:TYR:HD1	1:C:111:LYS:HZ1	1.56	0.45
1:B:97:LEU:HD21	1:B:99:LEU:HD11	1.99	0.44
1:B:109:ALA:O	1:B:113:LEU:HG	2.19	0.43
1:A:114:GLU:HG3	1:A:117:TYR:CE2	2.44	0.43
1:C:103:MET:C	1:C:104:LYS:C	2.79	0.43
1:C:24:LEU:HD12	1:C:117:TYR:CD1	2.55	0.42
1:A:24:LEU:HD12	1:A:117:TYR:CD1	2.55	0.42
1:B:20:PHE:O	1:B:24:LEU:HG	2.21	0.41
1:C:20:PHE:O	1:C:24:LEU:HG	2.21	0.41
1:B:24:LEU:HD12	1:B:117:TYR:CD1	2.55	0.41
1:A:20:PHE:O	1:A:24:LEU:HG	2.21	0.41
1:A:100:GLU:HA	1:A:103:MET:HE3	2.03	0.40

All (16) 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)
1:A:108:GLU:OE2	1:C:101:TYR:CD2[3_545]	0.68	1.52
1:A:108:GLU:OE1	1:C:101:TYR:CG[3_545]	0.80	1.40
1:A:108:GLU:CD	1:C:101:TYR:CD2[3_545]	1.15	1.05
1:A:108:GLU:OE2	1:C:101:TYR:CE2[3_545]	1.27	0.93
1:A:108:GLU:CD	1:C:101:TYR:CG[3_545]	1.34	0.86
1:A:108:GLU:OE1	1:C:101:TYR:CD1[3_545]	1.52	0.68
1:A:108:GLU:OE1	1:C:101:TYR:CB[3_545]	1.56	0.64
1:A:108:GLU:CD	1:C:101:TYR:CE2[3_545]	1.67	0.53
1:A:108:GLU:OE1	1:C:101:TYR:CD2[3_545]	1.81	0.39
1:A:108:GLU:CG	1:C:101:TYR:CZ[3_545]	1.81	0.39
1:A:108:GLU:CD	1:C:101:TYR:CD1[3_545]	1.90	0.30
1:A:108:GLU:OE2	1:C:101:TYR:CG[3_545]	1.95	0.25
1:A:108:GLU:CG	1:C:101:TYR:CE1[3_545]	2.02	0.18
1:A:108:GLU:CG	1:C:101:TYR:CE2[3_545]	2.09	0.11
1:B:29:GLU:OE1	1:C:104:LYS:NZ[2_555]	2.09	0.11
1:A:108:GLU:CD	1:C:101:TYR:CZ[3_545]	2.16	0.04

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	64/70 (91%)	63 (98%)	1 (2%)	0	100	100
1	B	60/70 (86%)	60 (100%)	0	0	100	100
1	C	63/70 (90%)	62 (98%)	1 (2%)	0	100	100
All	All	187/210 (89%)	185 (99%)	2 (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	56/57 (98%)	50 (89%)	6 (11%)	6	16
1	B	52/57 (91%)	49 (94%)	3 (6%)	18	42
1	C	55/57 (96%)	54 (98%)	1 (2%)	51	78
All	All	163/171 (95%)	153 (94%)	10 (6%)	17	40

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	GLU
1	A	102	GLU
1	A	105	LYS
1	A	111	LYS
1	A	114	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	115	GLU
1	B	29	GLU
1	B	103	MET
1	B	114	GLU
1	C	29	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	35	ASN
1	C	28	GLN
1	C	35	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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	103:MET	C	104:LYS	N	0.57

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, 95^{th} 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	68/70 (97%)	-1.60	0	100	100	9, 44, 74, 88	0
1	B	64/70 (91%)	-1.44	0	100	100	26, 59, 95, 122	0
1	C	67/70 (95%)	-1.42	0	100	100	36, 72, 117, 135	0
All	All	199/210 (94%)	-1.49	0	100	100	9, 54, 104, 135	0

There are no RSRZ outliers to report.

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

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

6.3 Carbohydrates

There are no oligosaccharides 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.