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 (i)

Dec 23, 2024 – 04:23 PM JST

PDB ID	:	8XO7
Title	:	Crystal structure of measles virus fusion inhibitor MEK35GE complexed with F protein HR1 (HR1-42) (P2 space group)
Authors	:	Oishi, S.; Takahara, A.; Nakatsu, T.
Deposited on	:	2023-12-31
Resolution	:	2.17 Å(reported)

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

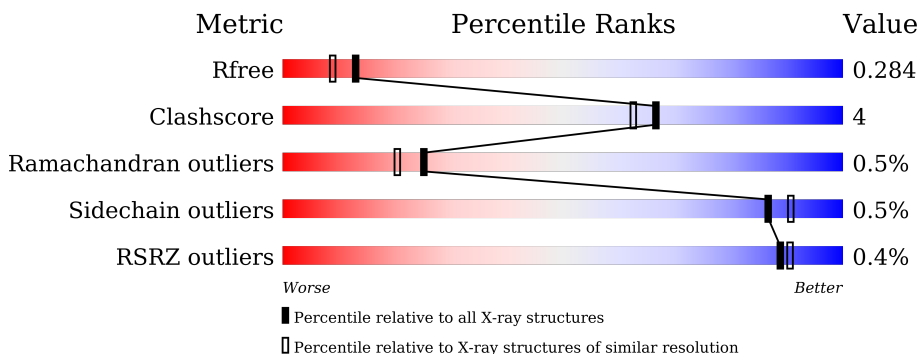|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.17 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	Percentile Ranks	Value
Rfree		0.284
Clashscore		4
Ramachandran outliers		0.5%
Sidechain outliers		0.5%
RSRZ outliers		0.4%

Worse ▮ *Better*

▮ Percentile relative to all X-ray structures

❙ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	164625	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (2.16-2.16)

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	44	89% ... 11%
1	C	44	93% ... 7%
1	E	44	95% ... 5%
1	G	44	91% ... 9%
1	I	44	93% ... 5% •
1	K	44	93% ... 7%

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	37	76% 14% 11%
2	D	37	92% 8%
2	F	37	84% 5% 11%
2	H	37	5% 86% 5% 8%
2	J	37	70% 14% 5% 11%
2	L	37	84% 5% 11%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3835 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 Fusion glycoprotein F1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	44	Total	C	N	O	S	0	1	1
			330	199	60	69	2			
1	C	44	Total	C	N	O	S	0	1	1
			330	199	60	69	2			
1	E	44	Total	C	N	O	S	0	2	1
			338	204	63	69	2			
1	G	44	Total	C	N	O	S	0	2	1
			336	203	61	70	2			
1	I	44	Total	C	N	O	S	0	2	1
			336	203	60	71	2			
1	K	44	Total	C	N	O	S	0	3	1
			341	206	62	71	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	142	ACE	-	acetylation	UNP P69353
A	185	NH2	-	amidation	UNP P69353
C	142	ACE	-	acetylation	UNP P69353
C	185	NH2	-	amidation	UNP P69353
E	142	ACE	-	acetylation	UNP P69353
E	185	NH2	-	amidation	UNP P69353
G	142	ACE	-	acetylation	UNP P69353
G	185	NH2	-	amidation	UNP P69353
I	142	ACE	-	acetylation	UNP P69353
I	185	NH2	-	amidation	UNP P69353
K	142	ACE	-	acetylation	UNP P69353
K	185	NH2	-	amidation	UNP P69353

- Molecule 2 is a protein called Measles virus fusion inhibitor MEK35GE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	33	Total 282	C 181	N 49	O 52	0	2	1
2	D	37	Total 295	C 187	N 49	O 59	0	1	1
2	F	33	Total 276	C 175	N 47	O 54	0	2	1
2	H	34	Total 278	C 176	N 47	O 55	0	1	1
2	J	33	Total 270	C 170	N 46	O 54	0	1	1
2	L	33	Total 270	C 171	N 47	O 52	0	1	1

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	16	Total 16	O 16	0	1
3	B	14	Total 14	O 14	0	0
3	C	22	Total 22	O 22	0	0
3	D	17	Total 17	O 17	0	0
3	E	19	Total 19	O 19	0	0
3	F	9	Total 9	O 9	0	0
3	G	8	Total 8	O 8	0	0
3	H	6	Total 6	O 6	0	0
3	I	10	Total 10	O 10	0	0
3	J	8	Total 8	O 8	0	0
3	K	15	Total 15	O 15	0	0
3	L	9	Total 9	O 9	0	0

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.
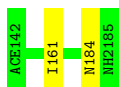
• Molecule 1: Fusion glycoprotein F1

Chain A:

• Molecule 1: Fusion glycoprotein F1

Chain C:

• Molecule 1: Fusion glycoprotein F1

Chain E:

• Molecule 1: Fusion glycoprotein F1

Chain G:

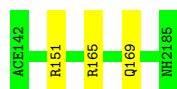• Molecule 1: Fusion glycoprotein F1

Chain I:

• Molecule 1: Fusion glycoprotein F1

Chain K:

93%   7%

ACE142 R151 R165 Q169 NH2185

• Molecule 2: Measles virus fusion inhibitor MEK35GE

Chain B:

76%   14%   11%
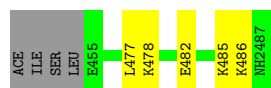
ACE ILE SER LEU E455 L477 K478 E482 K485 K486 NH2487

• Molecule 2: Measles virus fusion inhibitor MEK35GE

Chain D:

92%   8%

ACE451 K465 K469 L477 NH2487

• Molecule 2: Measles virus fusion inhibitor MEK35GE

Chain F:

84%   5%   11%
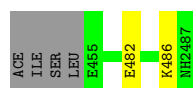
ACE ILE SER LEU E455 E482 K486 NH2487

• Molecule 2: Measles virus fusion inhibitor MEK35GE

5%

Chain H:

86%   5%   8%

ACE ILE SER L454 E475 L484 K485 K486 NH2487

• Molecule 2: Measles virus fusion inhibitor MEK35GE

Chain J:

70%   14%   5%   11%

ACE ILE SER LEU E455 V459 K472 E475 K479 E482 I483 L484 NH2487

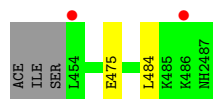• Molecule 2: Measles virus fusion inhibitor MEK35GE

Chain L:

84%   5%   11%

ACE ILE SER LEU E455 K471 L484 NH2487
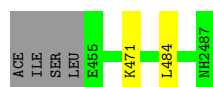
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	52.40Å 35.57Å 123.26Å 90.00° 91.26° 90.00°	Depositor
Resolution (Å)	48.60 – 2.17 48.60 – 2.17	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.60-2.17) 99.0 (48.60-2.17)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.68 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.203 , 0.281 0.209 , 0.284	Depositor DCC
R_{free} test set	1257 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	40.1	Xtriage
Anisotropy	0.537	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 34.0	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.020 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3835	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.85 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.4332e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, ACE

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	0.76	0/330	0.85	1/446 (0.2%)
1	C	0.76	0/330	0.80	0/446
1	E	0.75	0/341	0.84	0/460
1	G	0.71	0/339	0.79	0/458
1	I	0.72	0/339	0.86	0/458
1	K	0.75	0/347	0.87	0/469
2	B	0.79	0/287	0.88	0/369
2	D	0.76	0/294	0.83	0/386
2	F	0.77	0/280	0.87	0/365
2	H	0.73	0/279	0.79	0/365
2	J	0.81	0/271	0.91	0/354
2	L	0.73	0/271	0.81	0/353
All	All	0.75	0/3708	0.84	1/4929 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	151	ARG	NE-CZ-NH2	-5.06	117.77	120.30

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	330	0	326	3	0
1	C	330	0	326	2	0
1	E	338	0	339	2	0
1	G	336	0	334	5	0
1	I	336	0	332	5	0
1	K	341	0	340	2	0
2	B	282	0	330	4	0
2	D	295	0	328	2	0
2	F	276	0	310	1	0
2	H	278	0	308	2	0
2	J	270	0	297	5	0
2	L	270	0	304	5	0
3	A	16	0	0	1	0
3	B	14	0	0	1	0
3	C	22	0	0	1	0
3	D	17	0	0	0	0
3	E	19	0	0	0	0
3	F	9	0	0	0	0
3	G	8	0	0	0	0
3	H	6	0	0	0	0
3	I	10	0	0	0	0
3	J	8	0	0	0	0
3	K	15	0	0	0	0
3	L	9	0	0	0	0
All	All	3835	0	3874	27	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 4.

The worst 5 of 27 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:G:151:ARG:NH1	2:J:484:LEU:O	2.31	0.59
1:I:151:ARG:NH2	2:L:484:LEU:O	2.34	0.59
1:I:155[A]:GLU:HG2	2:L:484:LEU:HD23	1.86	0.56
1:I:155[B]:GLU:HG3	2:L:484:LEU:HD23	1.88	0.55
2:B:477:LEU:HD11	1:E:161:ILE:HG22	1.91	0.53

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	43/44 (98%)	42 (98%)	1 (2%)	0	100	100
1	C	43/44 (98%)	43 (100%)	0	0	100	100
1	E	44/44 (100%)	44 (100%)	0	0	100	100
1	G	44/44 (100%)	44 (100%)	0	0	100	100
1	I	44/44 (100%)	44 (100%)	0	0	100	100
1	K	45/44 (102%)	44 (98%)	1 (2%)	0	100	100
2	B	34/37 (92%)	31 (91%)	3 (9%)	0	100	100
2	D	36/37 (97%)	35 (97%)	1 (3%)	0	100	100
2	F	33/37 (89%)	31 (94%)	1 (3%)	1 (3%)	3	0
2	H	33/37 (89%)	32 (97%)	1 (3%)	0	100	100
2	J	32/37 (86%)	29 (91%)	2 (6%)	1 (3%)	3	0
2	L	32/37 (86%)	30 (94%)	2 (6%)	0	100	100
All	All	463/486 (95%)	449 (97%)	12 (3%)	2 (0%)	25	27

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	486	LYS
2	J	479	LYS

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	35/34 (103%)	35 (100%)	0	100	100
1	C	35/34 (103%)	35 (100%)	0	100	100
1	E	36/34 (106%)	36 (100%)	0	100	100
1	G	36/34 (106%)	36 (100%)	0	100	100
1	I	36/34 (106%)	34 (94%)	2 (6%)	17	14
1	K	37/34 (109%)	37 (100%)	0	100	100
2	B	32/32 (100%)	32 (100%)	0	100	100
2	D	33/32 (103%)	33 (100%)	0	100	100
2	F	31/32 (97%)	31 (100%)	0	100	100
2	H	31/32 (97%)	31 (100%)	0	100	100
2	J	30/32 (94%)	29 (97%)	1 (3%)	33	33
2	L	30/32 (94%)	30 (100%)	0	100	100
All	All	402/396 (102%)	399 (99%)	3 (1%)	86	86

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	155[A]	GLU
1	I	155[B]	GLU
2	J	482	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	E	169	GLN
1	E	184	ASN
1	G	149	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(Å²)	Q<0.9
1	A	42/44 (95%)	-0.23	0	100	100	27, 43, 60, 86	1 (2%)
1	C	42/44 (95%)	-0.41	0	100	100	26, 39, 52, 59	1 (2%)
1	E	42/44 (95%)	-0.34	0	100	100	24, 43, 61, 67	2 (4%)
1	G	42/44 (95%)	0.06	0	100	100	28, 52, 71, 80	2 (4%)
1	I	42/44 (95%)	0.11	0	100	100	27, 52, 75, 76	2 (4%)
1	K	42/44 (95%)	-0.18	0	100	100	21, 41, 72, 96	3 (7%)
2	B	32/37 (86%)	0.01	0	100	100	22, 53, 68, 87	2 (6%)
2	D	35/37 (94%)	-0.10	0	100	100	33, 52, 65, 80	1 (2%)
2	F	32/37 (86%)	-0.18	0	100	100	27, 50, 69, 82	2 (6%)
2	H	33/37 (89%)	0.31	2 (6%)	28	34	36, 55, 90, 106	1 (3%)
2	J	32/37 (86%)	0.56	0	100	100	36, 68, 94, 107	1 (3%)
2	L	32/37 (86%)	-0.03	0	100	100	35, 58, 81, 91	1 (3%)
All	All	448/486 (92%)	-0.05	2 (0%)	89	90	21, 50, 80, 107	19 (4%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	454	LEU	4.0
2	H	486	LYS	2.5

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 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.