wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:58 PM BST

PDB ID	:	4C2I
EMDB ID:	:	EMD-2442
Title	:	Cryo-EM structure of Dengue virus serotype 1 complexed with Fab fragments of human antibody 1F4
Authors	:	Fibriansah, G.; Tan, J.L.; de Alwis, R.; Smith, S.A.; Ng, T.-S.; Kostyuchenko, V.A.; Ibarra, K.D.; Harris, E.; de Silva, A.; Crowe Junior, J.E.; Lok, S.-M.
Deposited on	:	2013-08-18
Resolution	:	6.00 Å(reported)
Based on PDB ID	:	4AZX

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

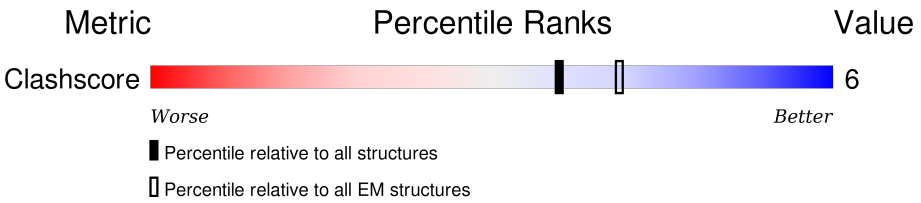
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 6.00 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)	EM structures (#Entries)
Clashscore	114402	924

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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%

Mol	Chain	Length	Quality of chain
1	A	495	100%
1	C	495	100%
1	E	495	100%
2	B	75	96% .
2	D	75	96% .
2	F	75	96% .
3	H	232	97% . .
3	M	232	98% .
4	L	239	88% 12%
4	N	239	88% 12%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 2761 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 ENVELOPE PROTEIN.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
1	A	495	495	495	0	495
1	C	495	495	495	0	495
1	E	495	495	495	0	495

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	SER	GLY	CONFLICT	UNP Q7TGE4
A	8	ASN	SER	CONFLICT	UNP Q7TGE4
A	17	ALA	GLY	CONFLICT	UNP Q7TGE4
A	18	THR	ALA	CONFLICT	UNP Q7TGE4
A	19	GLY	THR	CONFLICT	UNP Q7TGE4
C	7	SER	GLY	CONFLICT	UNP Q7TGE4
C	8	ASN	SER	CONFLICT	UNP Q7TGE4
C	17	ALA	GLY	CONFLICT	UNP Q7TGE4
C	18	THR	ALA	CONFLICT	UNP Q7TGE4
C	19	GLY	THR	CONFLICT	UNP Q7TGE4
E	7	SER	GLY	CONFLICT	UNP Q7TGE4
E	8	ASN	SER	CONFLICT	UNP Q7TGE4
E	17	ALA	GLY	CONFLICT	UNP Q7TGE4
E	18	THR	ALA	CONFLICT	UNP Q7TGE4
E	19	GLY	THR	CONFLICT	UNP Q7TGE4

- Molecule 2 is a protein called POLYPROTEIN.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
2	B	72	72	72	0	72
2	D	72	72	72	0	72

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
2	F	72	72	72	0	72

- Molecule 3 is a protein called HEAVY CHAIN FAB FRAGMENT OF ANTIBODY 1F4.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
3	H	228	228	228	0	228
			Total	C		
3	M	228	228	228	0	228

- Molecule 4 is a protein called LIGHT CHAIN FAB FRAGMENT OF ANTIBODY 1F4.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
4	L	211	211	211	0	211
			Total	C		
4	N	211	211	211	0	211

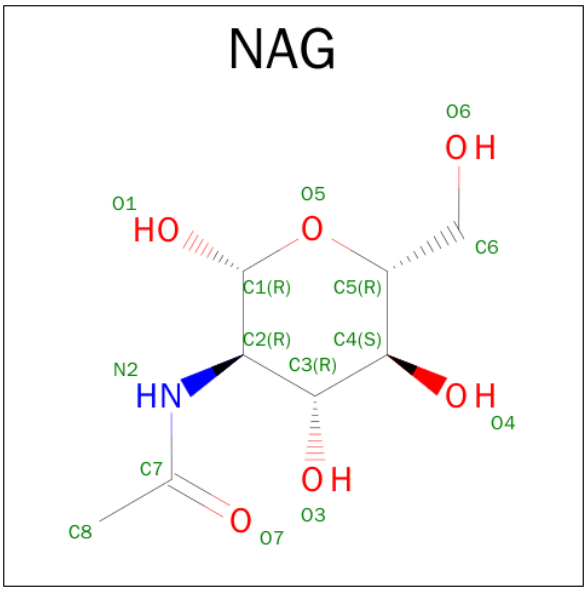
- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
5	A	3	42	24	3	15	0
			Total	C	N	O	
5	E	3	42	24	3	15	0

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
6	A	2	28	16	2	10	0
			Total	C	N	O	
6	C	2	28	16	2	10	0
			Total	C	N	O	
6	E	2	28	16	2	10	0

- Molecule 7 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
7	C	1	14	8	1	5	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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: ENVELOPE PROTEIN

Chain A: 100%

There are no outlier residues recorded for this chain.

- Molecule 1: ENVELOPE PROTEIN

Chain C: 100%

There are no outlier residues recorded for this chain.

- Molecule 1: ENVELOPE PROTEIN

Chain E: 100%

There are no outlier residues recorded for this chain.

- Molecule 2: POLYPROTEIN

Chain B: 96%

S1 P72 SER MET ALA

- Molecule 2: POLYPROTEIN

Chain D: 96%

S1 P72 SER MET ALA

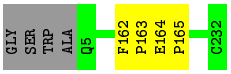
- Molecule 2: POLYPROTEIN

Chain F: 96%

S1 P72 SER MET ALA

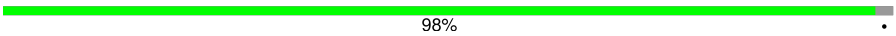
- Molecule 3: HEAVY CHAIN FAB FRAGMENT OF ANTIBODY 1F4

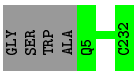
Chain H: 97%



- Molecule 3: HEAVY CHAIN FAB FRAGMENT OF ANTIBODY 1F4

Chain M:





- Molecule 4: LIGHT CHAIN FAB FRAGMENT OF ANTIBODY 1F4

Chain L:

- Molecule 4: LIGHT CHAIN FAB FRAGMENT OF ANTIBODY 1F4

Chain N:

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PARTICLE	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	17.5	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	47000	Depositor
Image detector	DIRECT ELECTRON DETECTOR (FALCON, FEI)	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

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

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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

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	495	0	0	0	0
1	C	495	0	0	0	0
1	E	495	0	0	0	0
2	B	72	0	0	0	0
2	D	72	0	0	0	0
2	F	72	0	0	0	0
3	H	228	0	0	2	0
3	M	228	0	0	0	0
4	L	211	0	0	0	0
4	N	211	0	0	0	0
5	A	42	0	37	4	0
5	E	42	0	37	5	0
6	A	28	0	25	4	0
6	C	28	0	24	2	0
6	E	28	0	25	0	0
7	C	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2761	0	161	17	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 17 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)
5:A:1501:NAG:H62	5:A:1502:NAG:O7	1.41	1.18
5:A:1501:NAG:C6	5:A:1502:NAG:O7	1.98	1.11
5:A:1501:NAG:H61	5:A:1502:NAG:O7	1.82	0.78
5:E:1496:NAG:O6	5:E:1497:NAG:C7	2.33	0.77
5:E:1496:NAG:H5	5:E:1497:NAG:O7	1.87	0.74

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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)

12 carbohydrates are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	NAG	A	1501	5	14,14,15	1.79	4 (28%)	15,19,21	1.95	4 (26%)
5	NAG	A	1502	5	14,14,15	1.55	3 (21%)	15,19,21	1.28	2 (13%)
5	NAG	A	1503	5	14,14,15	1.68	4 (28%)	15,19,21	1.70	3 (20%)
6	NAG	A	1601	6	14,14,15	1.29	2 (14%)	15,19,21	2.27	6 (40%)
6	NAG	A	1602	6	14,14,15	1.53	3 (21%)	15,19,21	1.51	3 (20%)
6	NAG	C	1601	6	14,14,15	1.39	3 (21%)	15,19,21	2.26	6 (40%)
6	NAG	C	1602	6	14,14,15	1.99	3 (21%)	15,19,21	1.90	4 (26%)
5	NAG	E	1496	5	14,14,15	1.15	1 (7%)	15,19,21	1.74	5 (33%)
5	NAG	E	1497	5	14,14,15	1.07	1 (7%)	15,19,21	1.49	1 (6%)
5	NAG	E	1498	5	14,14,15	1.63	4 (28%)	15,19,21	1.16	1 (6%)
6	NAG	E	1601	6	14,14,15	1.19	2 (14%)	15,19,21	2.12	5 (33%)
6	NAG	E	1602	6	14,14,15	2.00	2 (14%)	15,19,21	1.51	4 (26%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	A	1501	5	-	0/6/23/26	0/1/1/1
5	NAG	A	1502	5	-	0/6/23/26	0/1/1/1
5	NAG	A	1503	5	-	0/6/23/26	0/1/1/1
6	NAG	A	1601	6	-	0/6/23/26	0/1/1/1
6	NAG	A	1602	6	-	0/6/23/26	0/1/1/1
6	NAG	C	1601	6	-	0/6/23/26	0/1/1/1
6	NAG	C	1602	6	-	0/6/23/26	0/1/1/1
5	NAG	E	1496	5	-	0/6/23/26	0/1/1/1
5	NAG	E	1497	5	-	0/6/23/26	0/1/1/1
5	NAG	E	1498	5	-	0/6/23/26	0/1/1/1
6	NAG	E	1601	6	-	0/6/23/26	0/1/1/1
6	NAG	E	1602	6	-	0/6/23/26	0/1/1/1

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	1602	NAG	C1-C2	-5.78	1.44	1.52
6	A	1602	NAG	C1-C2	-4.25	1.46	1.52
6	E	1602	NAG	C2-N2	-3.95	1.39	1.46
5	E	1498	NAG	C4-C5	-3.52	1.45	1.53
5	A	1503	NAG	C1-C2	-3.05	1.48	1.52

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1601	NAG	O5-C5-C4	-4.78	102.22	110.13
5	A	1501	NAG	O5-C5-C4	-4.37	102.90	110.13
6	E	1601	NAG	C3-C4-C5	-3.92	103.23	110.23
6	C	1601	NAG	O5-C5-C4	-3.84	103.78	110.13
6	C	1601	NAG	O4-C4-C3	-3.34	102.84	110.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1501	NAG	4	0
5	A	1502	NAG	4	0
6	A	1601	NAG	4	0
6	A	1602	NAG	4	0
6	C	1601	NAG	2	0
6	C	1602	NAG	2	0
5	E	1496	NAG	5	0
5	E	1497	NAG	5	0

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is

the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
7	NAG	C	1501	-	14,14,15	1.74	3 (21%)	15,19,21	1.63	2 (13%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	NAG	C	1501	-	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	C	1501	NAG	C1-C2	-3.85	1.47	1.52
7	C	1501	NAG	O3-C3	2.55	1.49	1.43
7	C	1501	NAG	O5-C1	3.58	1.49	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	C	1501	NAG	O5-C5-C4	-2.04	106.75	110.13
7	C	1501	NAG	C2-N2-C7	4.57	129.04	123.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.