wwPDB X-ray Structure Validation Summary Report (i)

Oct 23, 2021 – 09:48 AM EDT

PDB ID	:	2VIS
Title	:	INFLUENZA VIRUS HEMAGGLUTININ, (ESCAPE) MUTANT WITH THR 131 REPLACED BY ILE, COMPLEXED WITH A NEUTRALIZING ANTIBODY
Authors	:	Bizebard, T.; Fleury, D.; Gigant, B.; Wharton, S.A.; Skehel, J.J.; Knossow, M.
Deposited on	:	1997-12-22
Resolution	:	3.25 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 (i) symbol.

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

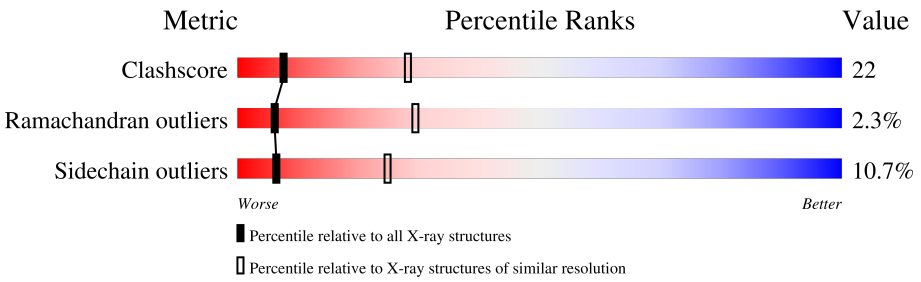
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.25 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))
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	210	61% 33% 6%
2	B	221	55% 40% 5%
3	C	282	46% 43% 5% 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5354 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 IMMUNOGLOBULIN (IGG1, LAMBDA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	210	1589	994	266	323	6	0	0	0

- Molecule 2 is a protein called IMMUNOGLOBULIN (IGG1, LAMBDA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	221	1672	1065	270	328	9	0	0	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	3	GLN	LYS	conflict	GB 4096752
B	5	LYS	GLN	conflict	GB 4096752
B	28	LEU	SER	conflict	GB 4096752
B	30	ILE	THR	conflict	GB 4096752
B	32	ASN	TYR	conflict	GB 4096752
B	63	LEU	HIS	conflict	GB 4096752
B	69	ILE	PHE	conflict	GB 4096752
B	83	LYS	ASN	conflict	GB 4096752
B	92	MET	LEU	conflict	GB 4096752
B	98	ASP	-	insertion	GB 4096752
B	99	PHE	-	insertion	GB 4096752
B	100	TYR	-	insertion	GB 4096752
B	102	TYR	HIS	conflict	GB 4096752
B	103	ASP	GLY	conflict	GB 4096752
B	105	PHE	-	insertion	GB 4096752
B	106	TYR	-	insertion	GB 4096752
B	107	TYR	-	insertion	GB 4096752
B	108	ALA	-	insertion	GB 4096752
B	109	MET	-	insertion	GB 4096752
B	110	ASP	-	insertion	GB 4096752

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	117	SER	LEU	conflict	GB 4096752
B	122	SER	ALA	conflict	GB 4096752
B	135	PRO	SER	conflict	GB 4096752

- Molecule 3 is a protein called HEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	267	2076	1306	363	396	11	0	0	0

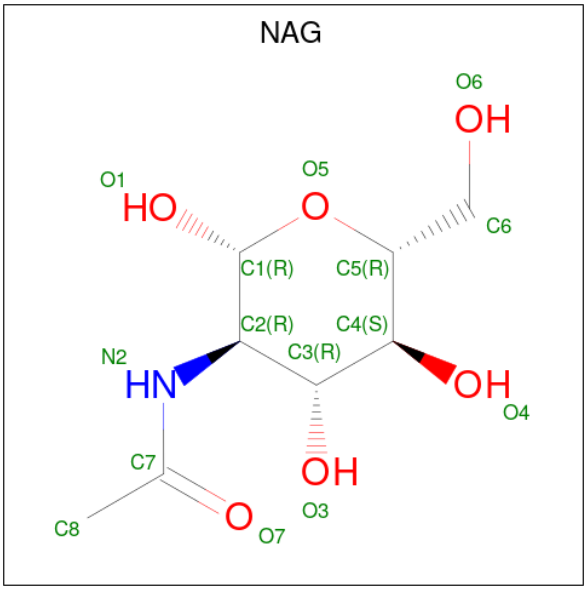
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	131	ILE	THR	engineered mutation	UNP P03437

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	A	2	2	2	0	0
			Total	Zn		
4	C	1	1	1	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

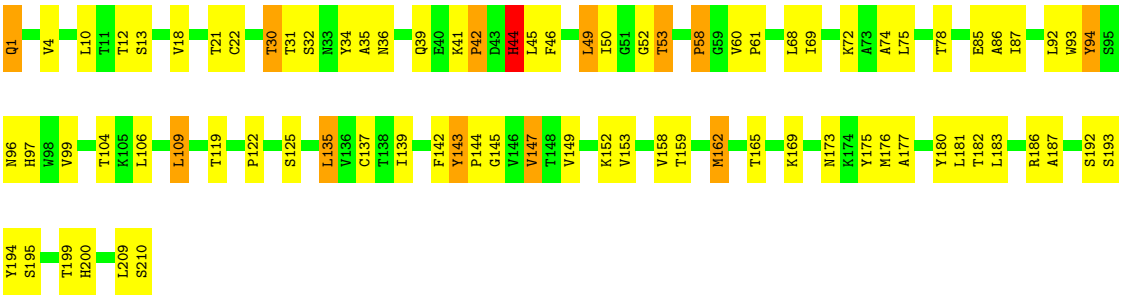
3 Residue-property plots (i)

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. 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.

Note EDS was not executed.

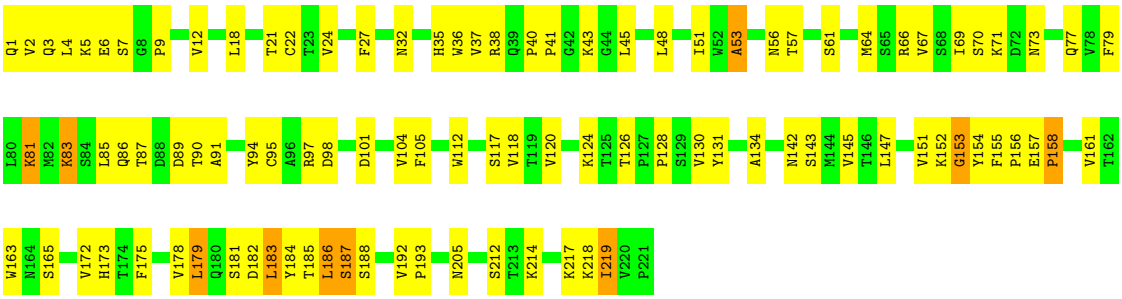
• Molecule 1: IMMUNOGLOBULIN (IGG1, LAMBDA)

Chain A: 61% 33% 6%



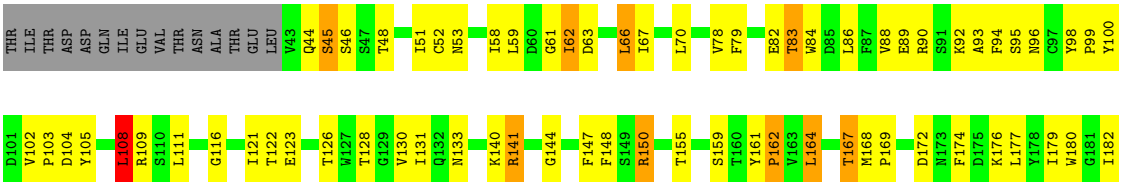
• Molecule 2: IMMUNOGLOBULIN (IGG1, LAMBDA)

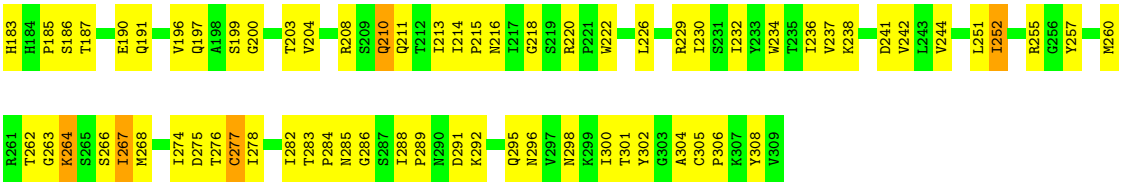
Chain B: 55% 40% 5%



• Molecule 3: HEMAGGLUTININ

Chain C: 46% 43% 5% 5%




H183
H184
P185
S186
T187
E190
Q191
V196
Q197
A198
S199
G200
T203
V204
R208
S209
Q210
Q211
T212
I213
I214
P215
N216
I217
G218
S219
R220
P221
W222
L226
R229
I230
S231
I232
Y233
W234
T235
I236
V237
K238
D241
V242
L243
V244
L251
I252
R255
G256
Y257
M260
R261
T262
G263
K264
S265
S266
I267
M268
I274
D275
T276
C277
I278
I282
T283
P284
N285
G286
S287
I288
P289
N290
D291
K292
Q295
N296
V297
N298
K299
I300
T301
Y302
G303
A304
C305
P306
K307
Y308
V309

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	85.50Å 85.50Å 515.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	7.00 – 3.25	Depositor
% Data completeness (in resolution range)	(Not available) (7.00-3.25)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.84	Depositor
R, R_{free}	0.188 , 0.267	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5354	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

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, ZN

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	0.58	0/1627	0.85	3/2224 (0.1%)
2	B	0.61	0/1718	0.86	1/2351 (0.0%)
3	C	0.60	0/2128	0.87	2/2897 (0.1%)
All	All	0.60	0/5473	0.86	6/7472 (0.1%)

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	A	0	2
2	B	0	1
3	C	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	45	LEU	CA-CB-CG	6.55	130.37	115.30
2	B	153	GLY	N-CA-C	6.29	128.82	113.10
1	A	135	LEU	CA-CB-CG	5.89	128.85	115.30
3	C	108	LEU	CA-CB-CG	5.38	127.67	115.30
1	A	45	LEU	N-CA-C	-5.21	96.94	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	143	TYR	Sidechain
1	A	94	TYR	Sidechain
2	B	53	ALA	Mainchain
3	C	302	TYR	Sidechain

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	1589	0	1532	67	0
2	B	1672	0	1635	75	0
3	C	2076	0	2020	97	0
4	A	2	0	0	0	0
4	C	1	0	0	0	0
5	C	14	0	13	0	0
All	All	5354	0	5200	232	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 22.

The worst 5 of 232 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)
2:B:134:ALA:HB2	2:B:219:ILE:HD11	1.50	0.89
3:C:283:THR:HG23	3:C:285:ASN:H	1.38	0.89
2:B:1:GLN:HE21	2:B:3:GLN:HB2	1.42	0.84
1:A:44:HIS:HB2	2:B:94:TYR:OH	1.80	0.82
3:C:45:SER:HB3	3:C:296:ASN:OD1	1.80	0.81

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	208/210 (99%)	180 (86%)	23 (11%)	5 (2%)	6 28
2	B	219/221 (99%)	186 (85%)	30 (14%)	3 (1%)	11 40
3	C	265/282 (94%)	232 (88%)	25 (9%)	8 (3%)	4 24
All	All	692/713 (97%)	598 (86%)	78 (11%)	16 (2%)	6 29

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	173	HIS
2	B	212	SER
3	C	210	GLN
3	C	263	GLY
1	A	44	HIS

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	178/178 (100%)	160 (90%)	18 (10%)	7 27
2	B	192/192 (100%)	175 (91%)	17 (9%)	9 32
3	C	236/250 (94%)	206 (87%)	30 (13%)	4 18
All	All	606/620 (98%)	541 (89%)	65 (11%)	6 25

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	210	GLN
3	C	255	ARG
2	B	147	LEU
2	B	142	ASN
3	C	264	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	210	GLN
3	C	216	ASN
3	C	298	ASN
1	A	96	ASN
2	B	3	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

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	C	4	3	14,14,15	1.63	4 (28%)	17,19,21	2.72	6 (35%)

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	C	4	3	-	1/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	4	NAG	O5-C1	-2.96	1.39	1.43
5	C	4	NAG	C4-C3	2.91	1.59	1.52
5	C	4	NAG	C6-C5	2.35	1.59	1.51
5	C	4	NAG	C4-C5	2.11	1.57	1.53

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	4	NAG	C1-O5-C5	5.99	120.31	112.19
5	C	4	NAG	C6-C5-C4	5.50	125.88	113.00
5	C	4	NAG	O5-C5-C6	-5.43	98.69	107.20
5	C	4	NAG	C1-C2-N2	3.02	115.65	110.49
5	C	4	NAG	O6-C6-C5	2.45	119.70	111.29

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	4	NAG	O5-C5-C6-O6

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.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.