



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 14, 2023 – 11:51 PM EDT

PDB ID : 1RU7  
Title : 1934 Human H1 Hemagglutinin  
Authors : Skehel, J.J.; Gamblin, S.J.; Haire, L.F.; Russell, R.J.; Stevens, D.J.; Xiao, B.;  
Ha, Y.; Vasisht, N.; Steinhauer, D.A.; Daniels, R.S.  
Deposited on : 2003-12-11  
Resolution : 2.30 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](mailto: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 : 4.02b-467  
Xtrriage (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) : 2.35

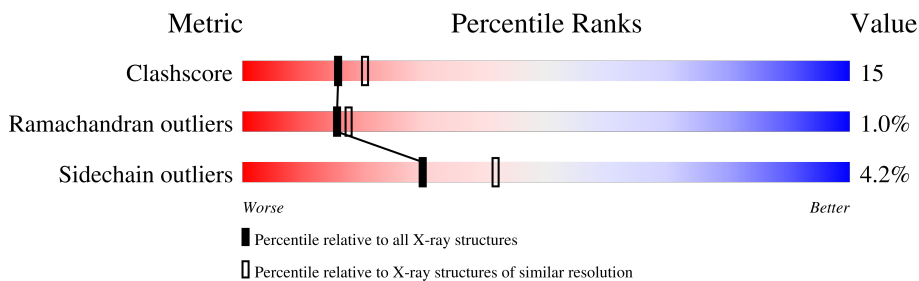
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)





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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	327	
1	C	327	
1	E	327	
1	G	327	
1	I	327	
1	K	327	
2	B	160	
2	D	160	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	160	 78% 18% ...
2	H	160	 66% 29% . .
2	J	160	 69% 28% .
2	L	160	 75% 22% ..

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 24314 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 hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	323	2557	1612	450	482	13	0	0	0
1	C	323	2557	1612	450	482	13	0	0	0
1	E	323	2557	1612	450	482	13	0	0	0
1	G	323	2557	1612	450	482	13	0	0	0
1	I	323	2557	1612	450	482	13	0	0	0
1	K	323	2557	1612	450	482	13	0	0	0

- Molecule 2 is a protein called hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	160	1282	805	218	252	7	0	0	0
2	D	160	1282	805	218	252	7	0	0	0
2	F	159	1275	800	217	251	7	0	0	0
2	H	160	1282	805	218	252	7	0	0	0
2	J	160	1282	805	218	252	7	0	0	0
2	L	159	1275	800	217	251	7	0	0	0

- Molecule 3 is water.

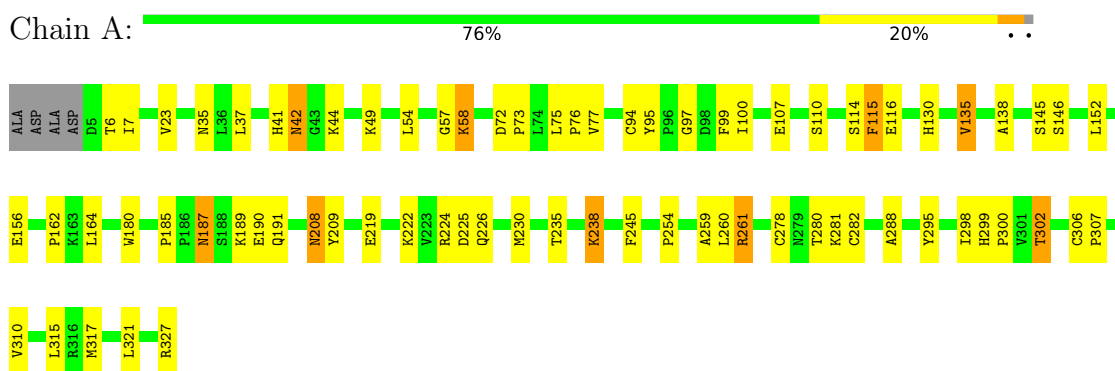
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	165	Total 165	O 165	0	0
3	B	39	Total 39	O 39	0	0
3	C	152	Total 152	O 152	0	0
3	D	56	Total 56	O 56	0	0
3	E	151	Total 151	O 151	0	0
3	F	57	Total 57	O 57	0	0
3	G	178	Total 178	O 178	0	0
3	H	46	Total 46	O 46	0	0
3	I	181	Total 181	O 181	0	0
3	J	51	Total 51	O 51	0	0
3	K	157	Total 157	O 157	0	0
3	L	61	Total 61	O 61	0	0

### 3 Residue-property plots [i](#)

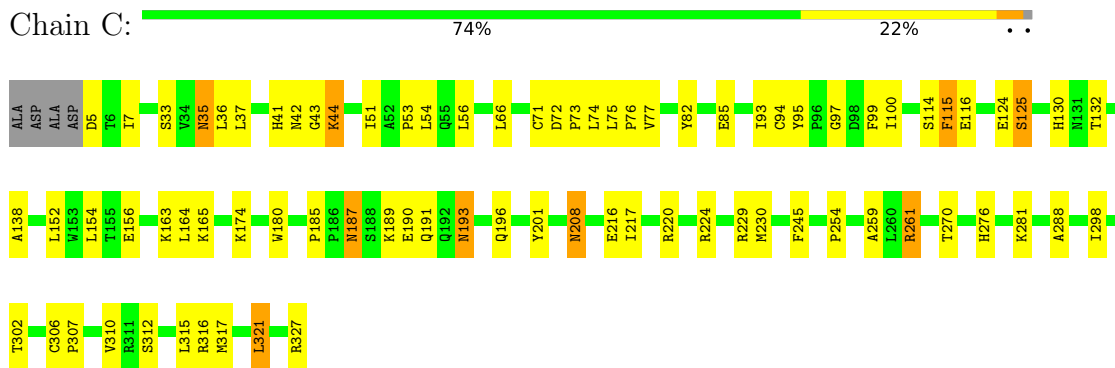
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

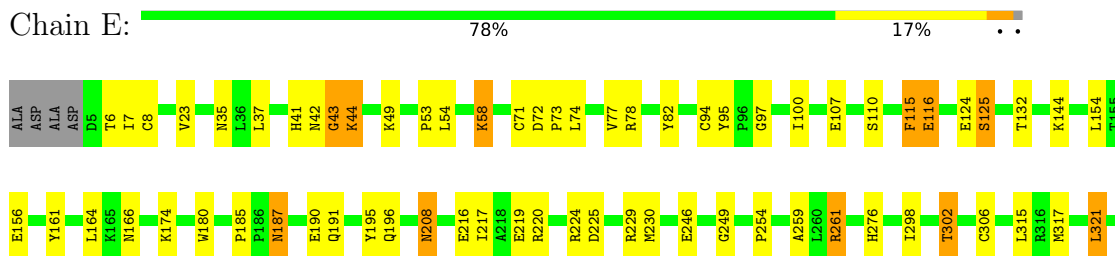
- Molecule 1: hemagglutinin



- Molecule 1: hemagglutinin

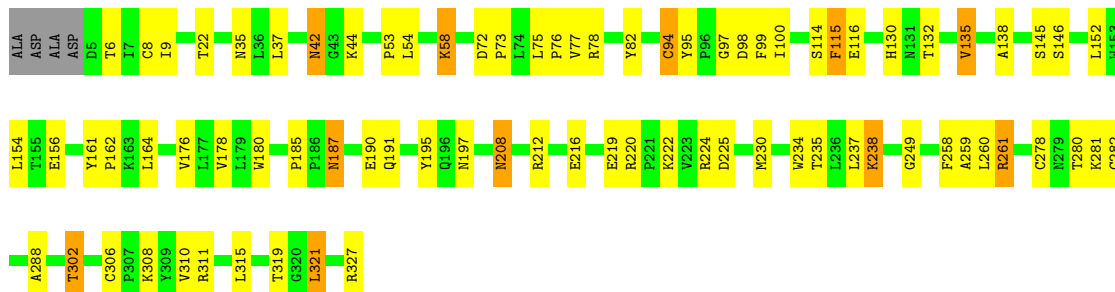


- Molecule 1: hemagglutinin

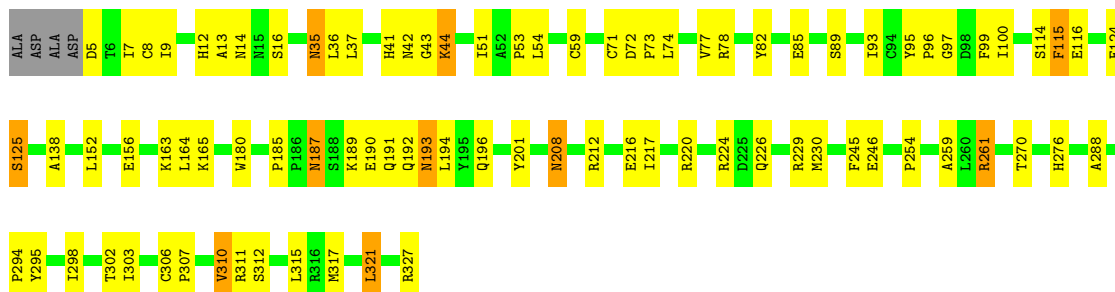


R327

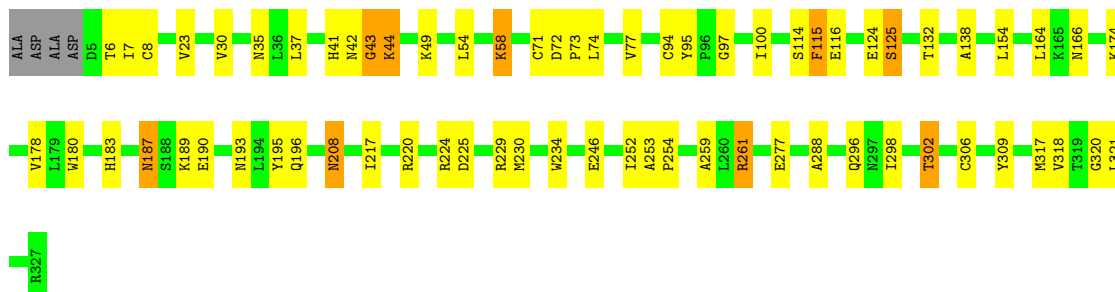
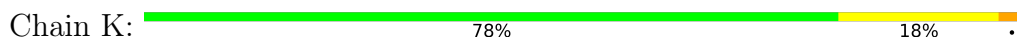
• Molecule 1: hemagglutinin



• Molecule 1: hemagglutinin



• Molecule 1: hemagglutinin



• Molecule 2: hemagglutinin

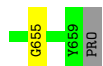
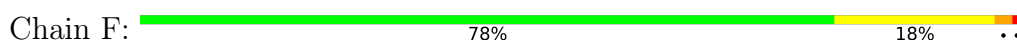




• Molecule 2: hemagglutinin



• Molecule 2: hemagglutinin



• Molecule 2: hemagglutinin



• Molecule 2: hemagglutinin



• Molecule 2: hemagglutinin





M649	E650	N654	D658	Y659	PRU
------	------	------	------	------	-----

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	227.51Å 131.08Å 174.38Å 90.00° 109.81° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.30)	Depositor
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.235 , 0.266	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	24314	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

## 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	0.36	0/2621	0.64	0/3563
1	C	0.35	0/2620	0.63	0/3560
1	E	0.35	0/2621	0.64	1/3563 (0.0%)
1	G	0.35	0/2621	0.64	0/3563
1	I	0.35	0/2620	0.62	0/3560
1	K	0.35	0/2621	0.65	1/3563 (0.0%)
2	B	0.33	0/1308	0.51	0/1761
2	D	0.34	0/1308	0.54	0/1761
2	F	0.35	0/1300	0.53	0/1749
2	H	0.33	0/1308	0.52	0/1761
2	J	0.34	0/1308	0.53	0/1761
2	L	0.35	0/1300	0.52	0/1749
All	All	0.35	0/23556	0.60	2/31914 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	43	GLY	N-CA-C	-8.50	91.86	113.10
1	E	43	GLY	N-CA-C	-6.70	96.36	113.10

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	2557	0	2495	76	0
1	C	2557	0	2494	80	0
1	E	2557	0	2495	61	0
1	G	2557	0	2495	97	0
1	I	2557	0	2494	82	0
1	K	2557	0	2495	64	0
2	B	1282	0	1207	47	0
2	D	1282	0	1207	53	0
2	F	1275	0	1200	27	0
2	H	1282	0	1207	54	0
2	J	1282	0	1207	45	0
2	L	1275	0	1200	32	0
3	A	165	0	0	3	0
3	B	39	0	0	1	0
3	C	152	0	0	2	0
3	D	56	0	0	1	0
3	E	151	0	0	5	0
3	F	57	0	0	1	0
3	G	178	0	0	7	0
3	H	46	0	0	1	0
3	I	181	0	0	5	0
3	J	51	0	0	1	0
3	K	157	0	0	2	0
3	L	61	0	0	1	0
All	All	24314	0	22196	664	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 15.

The worst 5 of 664 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:633:ILE:HD11	2:B:639:GLU:HB2	1.29	1.15
1:G:280:THR:HG22	1:G:282:CYS:H	1.14	1.12
1:A:238:LYS:H	1:A:238:LYS:CE	1.66	1.07
1:A:238:LYS:H	1:A:238:LYS:HE3	0.93	1.05
2:H:633:ILE:HD11	2:H:639:GLU:HB2	1.36	1.05

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	321/327 (98%)	308 (96%)	13 (4%)	0	100	100
1	C	319/327 (98%)	309 (97%)	7 (2%)	3 (1%)	17	20
1	E	321/327 (98%)	308 (96%)	11 (3%)	2 (1%)	25	31
1	G	321/327 (98%)	311 (97%)	10 (3%)	0	100	100
1	I	319/327 (98%)	306 (96%)	11 (3%)	2 (1%)	25	31
1	K	321/327 (98%)	308 (96%)	11 (3%)	2 (1%)	25	31
2	B	158/160 (99%)	137 (87%)	14 (9%)	7 (4%)	2	1
2	D	158/160 (99%)	141 (89%)	15 (10%)	2 (1%)	12	12
2	F	157/160 (98%)	146 (93%)	10 (6%)	1 (1%)	25	31
2	H	158/160 (99%)	139 (88%)	12 (8%)	7 (4%)	2	1
2	J	158/160 (99%)	143 (90%)	13 (8%)	2 (1%)	12	12
2	L	157/160 (98%)	145 (92%)	12 (8%)	0	100	100
All	All	2868/2922 (98%)	2701 (94%)	139 (5%)	28 (1%)	15	17

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	659	TYR
2	H	633	ILE
2	H	659	TYR
2	B	526	HIS
2	B	633	ILE

### 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	285/287 (99%)	271 (95%)	14 (5%)	25	35
1	C	285/287 (99%)	276 (97%)	9 (3%)	39	54
1	E	285/287 (99%)	275 (96%)	10 (4%)	36	50
1	G	285/287 (99%)	271 (95%)	14 (5%)	25	35
1	I	285/287 (99%)	275 (96%)	10 (4%)	36	50
1	K	285/287 (99%)	278 (98%)	7 (2%)	47	65
2	B	136/136 (100%)	128 (94%)	8 (6%)	19	27
2	D	136/136 (100%)	129 (95%)	7 (5%)	24	33
2	F	135/136 (99%)	129 (96%)	6 (4%)	28	39
2	H	136/136 (100%)	129 (95%)	7 (5%)	24	33
2	J	136/136 (100%)	129 (95%)	7 (5%)	24	33
2	L	135/136 (99%)	128 (95%)	7 (5%)	23	32
All	All	2524/2538 (99%)	2418 (96%)	106 (4%)	30	42

5 of 106 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	99	PHE
2	H	598	LEU
2	L	502	LEU
1	G	135	VAL
1	G	278	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	528	ASN
1	I	193	ASN
2	H	560	ASN
1	I	42	ASN
2	J	530	GLN

### 5.3.3 RNA

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](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	I	1
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	I	42:ASN	C	43:GLY	N	2.63
1	C	42:ASN	C	43:GLY	N	2.60

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.