



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 20, 2023 – 07:01 PM JST

PDB ID : 7D5Z
Title : Crystal structure of EBV gH/gL bound with neutralizing antibody 1D8
Authors : Zhu, Q.; Shan, S.; Yu, J.; Wang, X.; Zhang, L.; Zeng, M.
Deposited on : 2020-09-28
Resolution : 4.20 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 ⓘ 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
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

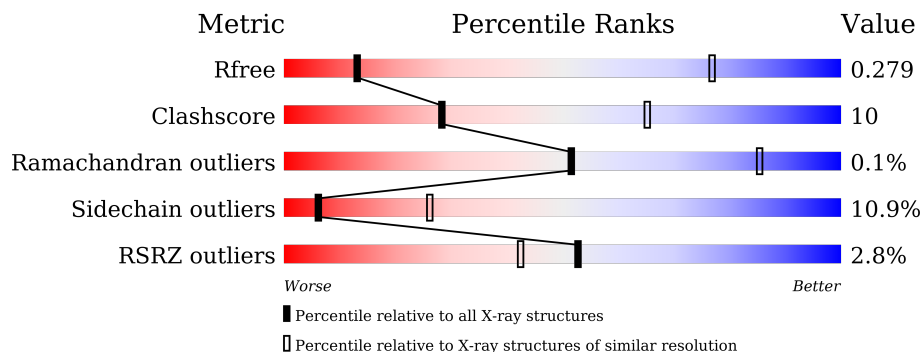
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.20 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}	130704	1005 (4.62-3.78)
Clashscore	141614	1044 (4.60-3.80)
Ramachandran outliers	138981	1000 (4.60-3.80)
Sidechain outliers	138945	1007 (4.62-3.78)
RSRZ outliers	127900	1063 (4.70-3.70)

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 $\leq 5\%$. 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	665	
1	E	665	
1	I	665	
1	M	665	
2	B	112	
2	F	112	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	112	<p>3% 57% 26% 16%</p>
2	N	112	<p>10% 56% 26% 16%</p>
3	C	233	<p>61% 26% 9%</p>
3	G	233	<p>64% 24% 9%</p>
3	K	233	<p>60% 27% 9%</p>
3	O	233	<p>66% 21% 9%</p>
4	D	470	<p>32% 12% 54%</p>
4	H	470	<p>30% 14% 54%</p>
4	L	470	<p>30% 14% 54%</p>
4	P	470	<p>30% 13% 54%</p>

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 36304 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 Envelope glycoprotein H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	M	655	5099	3270	842	955	32	0	0	0
1	A	655	5099	3270	842	955	32	0	0	0
1	E	655	5099	3270	842	955	32	0	0	0
1	I	655	5099	3270	842	955	32	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	680	HIS	-	expression tag	UNP Q3KSQ3
M	681	HIS	-	expression tag	UNP Q3KSQ3
M	682	HIS	-	expression tag	UNP Q3KSQ3
M	683	HIS	-	expression tag	UNP Q3KSQ3
M	684	HIS	-	expression tag	UNP Q3KSQ3
A	680	HIS	-	expression tag	UNP Q3KSQ3
A	681	HIS	-	expression tag	UNP Q3KSQ3
A	682	HIS	-	expression tag	UNP Q3KSQ3
A	683	HIS	-	expression tag	UNP Q3KSQ3
A	684	HIS	-	expression tag	UNP Q3KSQ3
E	680	HIS	-	expression tag	UNP Q3KSQ3
E	681	HIS	-	expression tag	UNP Q3KSQ3
E	682	HIS	-	expression tag	UNP Q3KSQ3
E	683	HIS	-	expression tag	UNP Q3KSQ3
E	684	HIS	-	expression tag	UNP Q3KSQ3
I	680	HIS	-	expression tag	UNP Q3KSQ3
I	681	HIS	-	expression tag	UNP Q3KSQ3
I	682	HIS	-	expression tag	UNP Q3KSQ3
I	683	HIS	-	expression tag	UNP Q3KSQ3
I	684	HIS	-	expression tag	UNP Q3KSQ3

- Molecule 2 is a protein called Envelope glycoprotein L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	94	Total	C	N	O	S	0	0	0
			719	454	120	141	4			
2	B	94	Total	C	N	O	S	0	0	0
			719	454	120	141	4			
2	F	94	Total	C	N	O	S	0	0	0
			719	454	120	141	4			
2	J	94	Total	C	N	O	S	0	0	0
			719	454	120	141	4			

- Molecule 3 is a protein called light chain of 1D8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	O	213	Total	C	N	O	S	0	0	0
			1647	1035	278	329	5			
3	C	213	Total	C	N	O	S	0	0	0
			1647	1035	278	329	5			
3	G	213	Total	C	N	O	S	0	0	0
			1647	1035	278	329	5			
3	K	213	Total	C	N	O	S	0	0	0
			1647	1035	278	329	5			

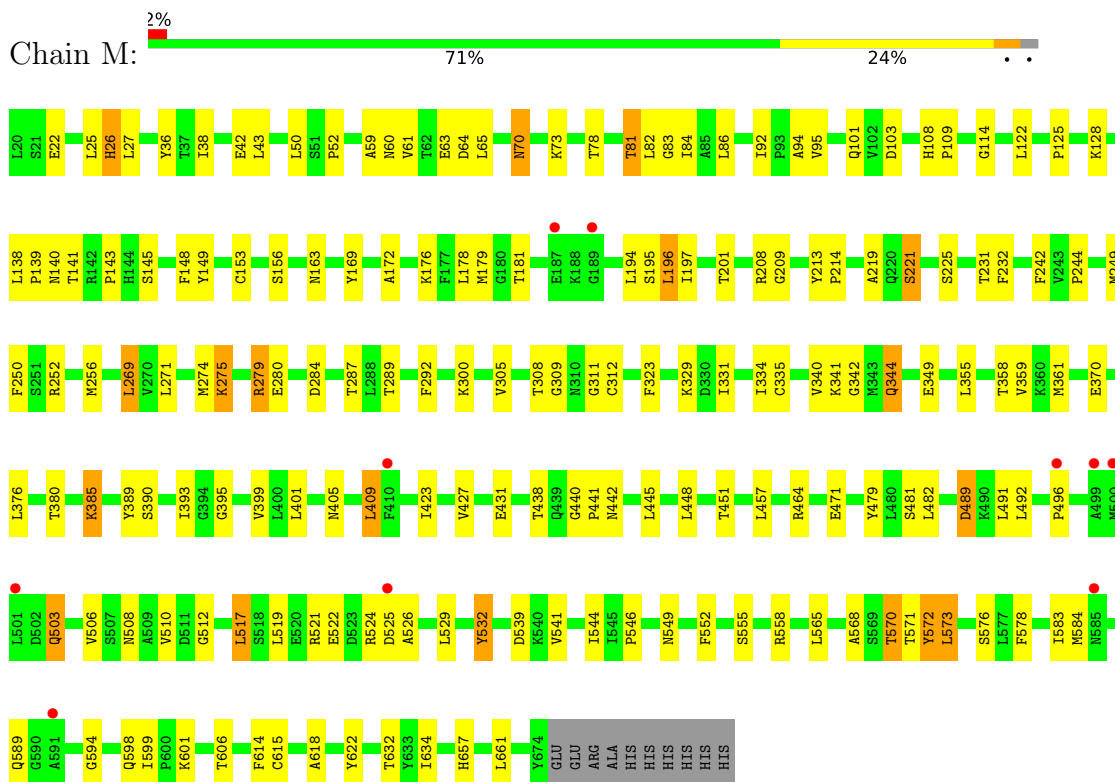
- Molecule 4 is a protein called heavy chain of 1D8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	P	214	Total	C	N	O	S	0	0	0
			1611	1021	273	307	10			
4	D	214	Total	C	N	O	S	0	0	0
			1611	1021	273	307	10			
4	H	214	Total	C	N	O	S	0	0	0
			1611	1021	273	307	10			
4	L	214	Total	C	N	O	S	0	0	0
			1611	1021	273	307	10			

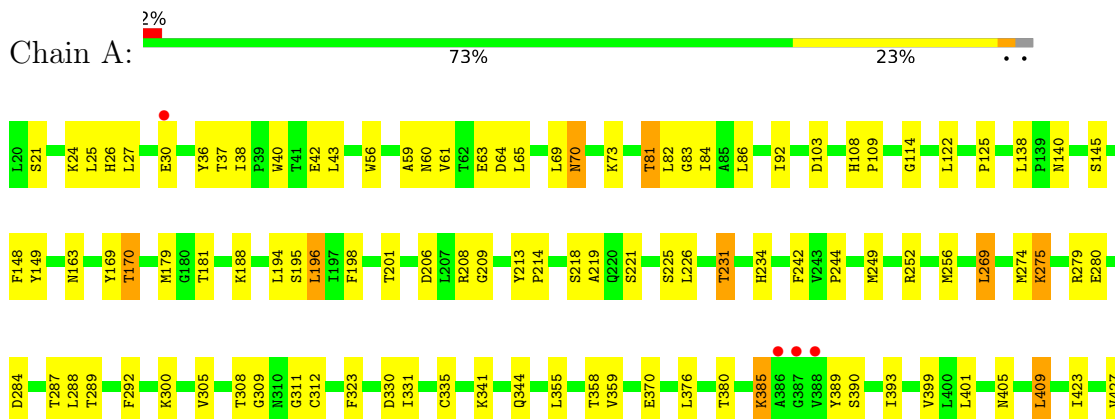
3 Residue-property plots [i](#)

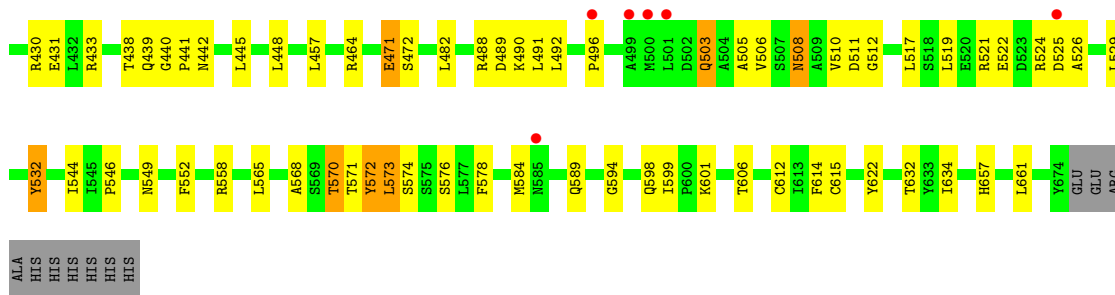
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

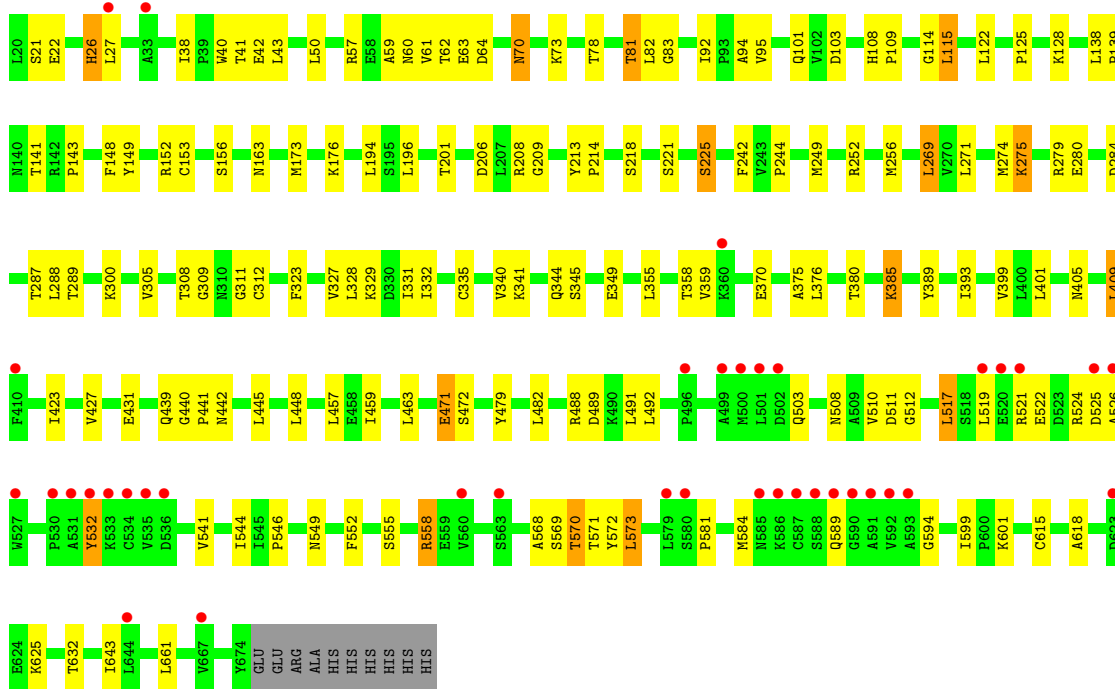
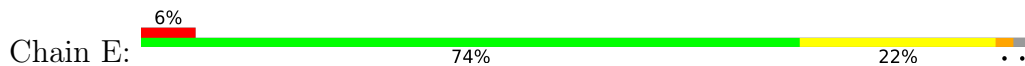


- Molecule 1: Envelope glycoprotein H

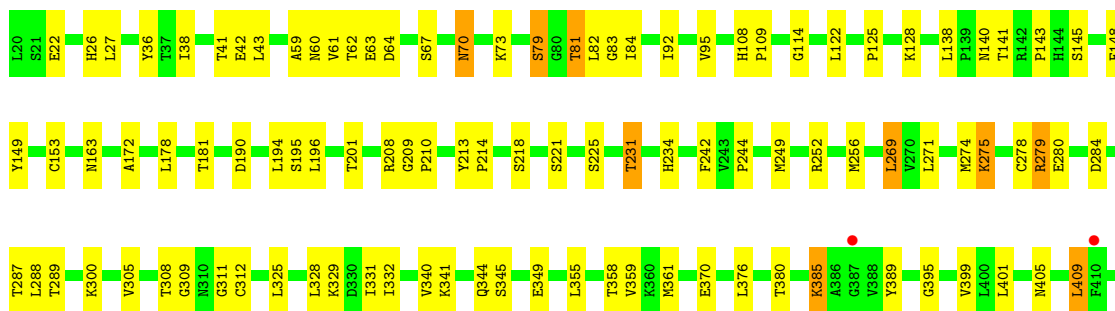


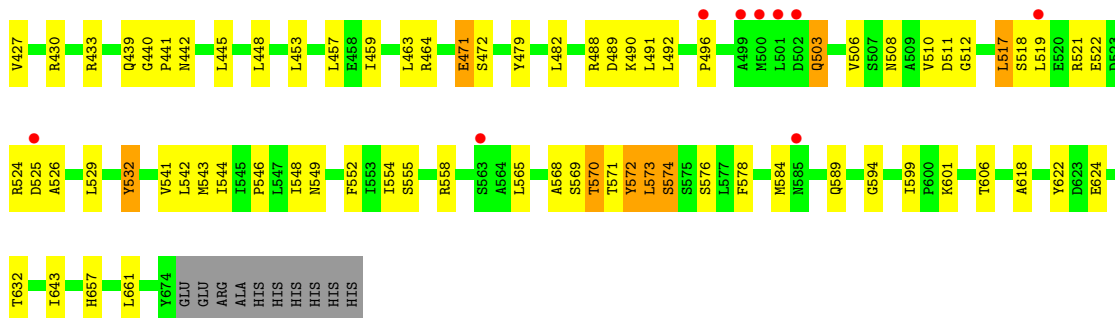


• Molecule 1: Envelope glycoprotein H

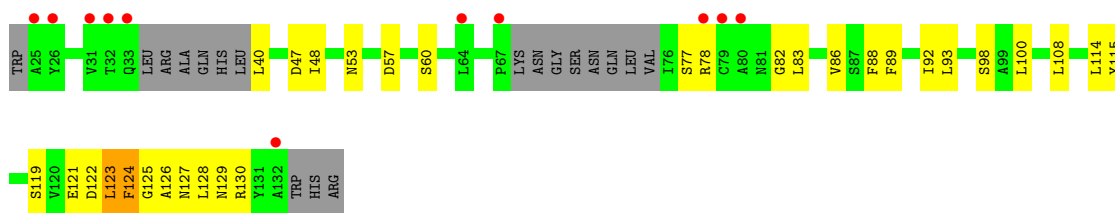


• Molecule 1: Envelope glycoprotein H

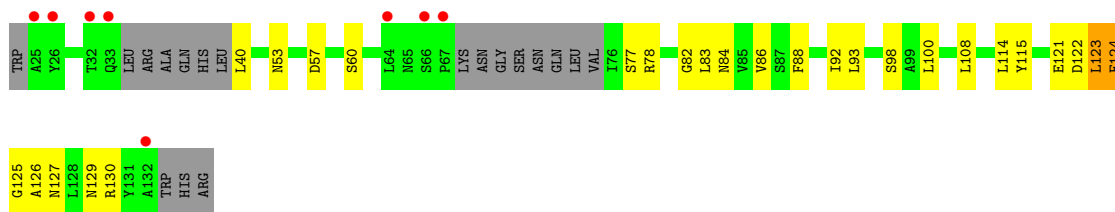




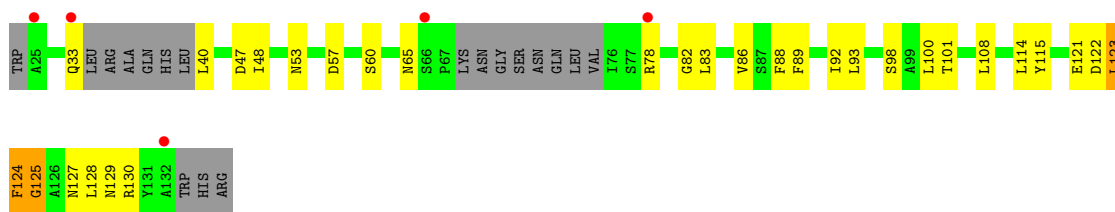
● Molecule 2: Envelope glycoprotein L



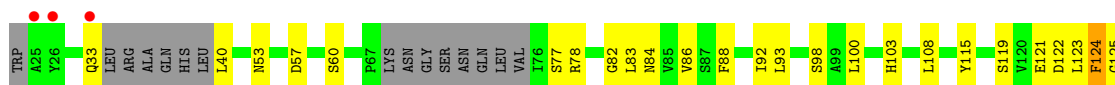
● Molecule 2: Envelope glycoprotein L



● Molecule 2: Envelope glycoprotein L



● Molecule 2: Envelope glycoprotein L

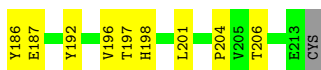




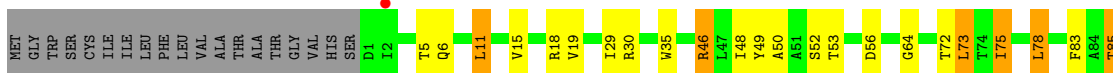
- Molecule 3: light chain of 1D8



- Molecule 3: light chain of 1D8

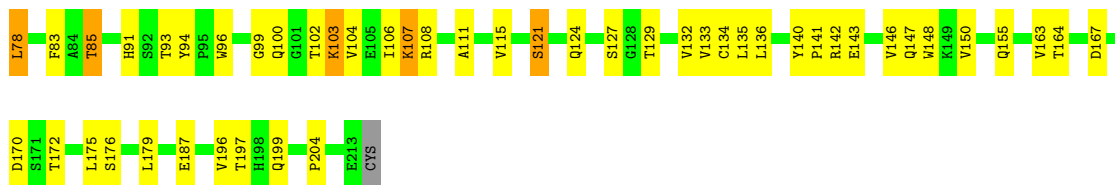


- Molecule 3: light chain of 1D8

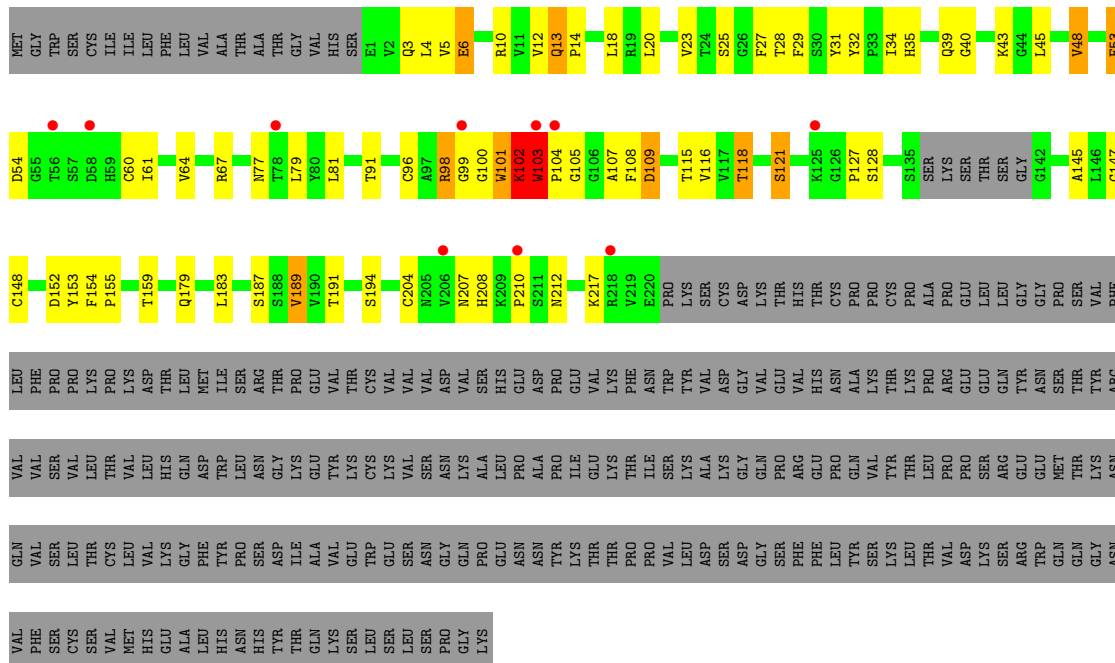
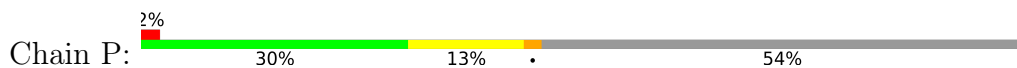


- Molecule 3: light chain of 1D8

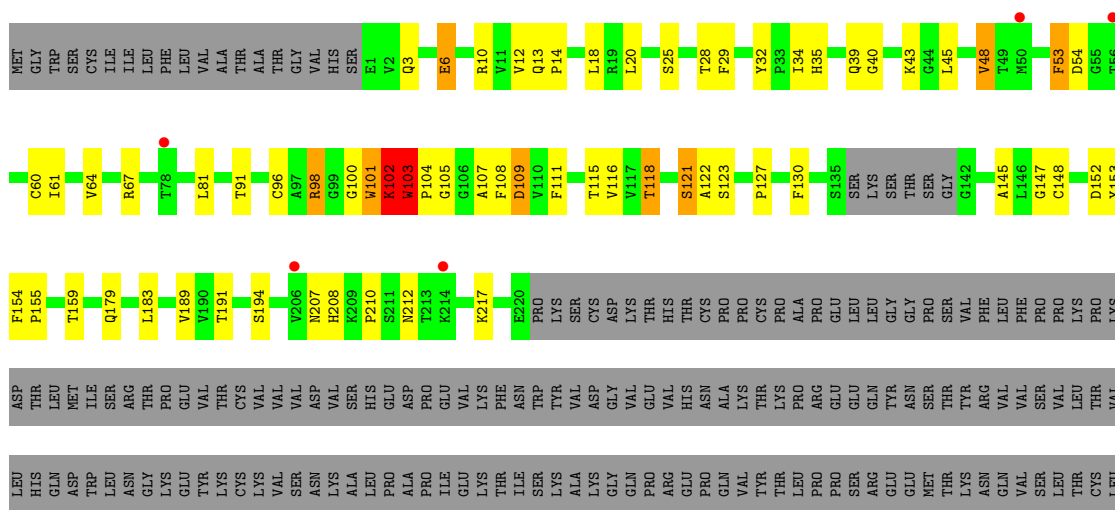
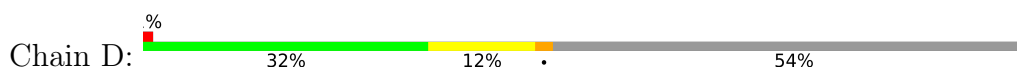




• Molecule 4: heavy chain of 1D8



• Molecule 4: heavy chain of 1D8



PHE
SER
CYS
SER
VAL
MET
HIS
GLU
ALA
LEU
HIS
ASN
HIS
TYR
THR
GLN
LYS
SER
LEU
SER
LEU
SER
PRO
GLY
LYS

4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	212.87Å 212.87Å 598.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.03 – 4.20 50.03 – 4.20	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.03-4.20) 99.9 (50.03-4.20)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.91 (at 4.14Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.246 , 0.277 0.250 , 0.279	Depositor DCC
R_{free} test set	5096 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	127.2	Xtrriage
Anisotropy	0.104	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 163.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	36304	wwPDB-VP
Average B, all atoms (Å ²)	153.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 31.10 % 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 1.1713e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.28	0/5208	0.50	0/7067
1	E	0.28	0/5208	0.51	0/7067
1	I	0.28	0/5208	0.50	0/7067
1	M	0.27	0/5208	0.50	0/7067
2	B	0.28	0/729	0.49	0/987
2	F	0.27	0/729	0.47	0/987
2	J	0.27	0/729	0.49	0/987
2	N	0.27	0/729	0.47	0/987
3	C	0.30	0/1686	0.52	0/2294
3	G	0.29	0/1686	0.53	0/2294
3	K	0.29	0/1686	0.53	0/2294
3	O	0.29	0/1686	0.53	0/2294
4	D	0.32	0/1653	0.55	2/2249 (0.1%)
4	H	0.34	1/1653 (0.1%)	0.56	2/2249 (0.1%)
4	L	0.32	0/1653	0.56	2/2249 (0.1%)
4	P	0.31	0/1653	0.55	2/2249 (0.1%)
All	All	0.29	1/37104 (0.0%)	0.51	8/50388 (0.0%)

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
3	C	0	2
3	G	0	1
3	K	0	2
3	O	0	1
4	D	0	2
4	H	0	2
4	L	0	2
4	P	0	2
All	All	0	14

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	103	TRP	CB-CG	5.80	1.60	1.50

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	103	TRP	N-CA-C	6.17	127.65	111.00
4	L	103	TRP	N-CA-C	6.09	127.44	111.00
4	L	102	LYS	C-N-CA	6.08	136.91	121.70
4	H	102	LYS	C-N-CA	6.02	136.74	121.70
4	D	103	TRP	N-CA-C	5.90	126.93	111.00

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	106	ILE	Peptide
3	C	29	ILE	Peptide
3	O	29	ILE	Peptide
4	P	102	LYS	Peptide
4	P	121	SER	Peptide

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	5099	0	5119	91	0
1	E	5099	0	5119	87	0
1	I	5099	0	5119	99	0
1	M	5099	0	5119	102	0
2	B	719	0	703	15	0
2	F	719	0	703	19	0
2	J	719	0	703	19	0
2	N	719	0	703	18	0
3	C	1647	0	1598	37	0
3	G	1647	0	1598	36	0
3	K	1647	0	1598	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	O	1647	0	1598	29	0
4	D	1611	0	1568	43	0
4	H	1611	0	1568	45	0
4	L	1611	0	1568	47	0
4	P	1611	0	1568	54	0
All	All	36304	0	35952	708	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 10.

The worst 5 of 708 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:M:201:THR:HG22	1:M:225:SER:HB2	1.53	0.88
3:K:75:ILE:HG12	3:K:78:LEU:HD23	1.59	0.84
1:I:201:THR:HG22	1:I:225:SER:HB2	1.59	0.84
1:I:526:ALA:HB1	1:I:532:TYR:HE2	1.44	0.82
3:G:75:ILE:HG12	3:G:78:LEU:HD23	1.62	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	653/665 (98%)	609 (93%)	44 (7%)	0	100	100
1	E	653/665 (98%)	613 (94%)	40 (6%)	0	100	100
1	I	653/665 (98%)	606 (93%)	47 (7%)	0	100	100
1	M	653/665 (98%)	606 (93%)	47 (7%)	0	100	100
2	B	88/112 (79%)	79 (90%)	8 (9%)	1 (1%)	14	52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	88/112 (79%)	80 (91%)	6 (7%)	2 (2%)	6	37
2	J	88/112 (79%)	80 (91%)	7 (8%)	1 (1%)	14	52
2	N	88/112 (79%)	80 (91%)	7 (8%)	1 (1%)	14	52
3	C	211/233 (91%)	200 (95%)	11 (5%)	0	100	100
3	G	211/233 (91%)	198 (94%)	13 (6%)	0	100	100
3	K	211/233 (91%)	201 (95%)	10 (5%)	0	100	100
3	O	211/233 (91%)	199 (94%)	12 (6%)	0	100	100
4	D	210/470 (45%)	197 (94%)	13 (6%)	0	100	100
4	H	210/470 (45%)	197 (94%)	13 (6%)	0	100	100
4	L	210/470 (45%)	196 (93%)	14 (7%)	0	100	100
4	P	210/470 (45%)	198 (94%)	12 (6%)	0	100	100
All	All	4648/5920 (78%)	4339 (93%)	304 (6%)	5 (0%)	51	85

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	N	124	PHE
2	B	124	PHE
2	F	124	PHE
2	J	124	PHE
2	F	125	GLY

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	561/570 (98%)	505 (90%)	56 (10%)	7	28
1	E	561/570 (98%)	507 (90%)	54 (10%)	8	30
1	I	561/570 (98%)	504 (90%)	57 (10%)	7	28
1	M	561/570 (98%)	507 (90%)	54 (10%)	8	30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	83/99 (84%)	74 (89%)	9 (11%)	6	26
2	F	83/99 (84%)	74 (89%)	9 (11%)	6	26
2	J	83/99 (84%)	75 (90%)	8 (10%)	8	30
2	N	83/99 (84%)	74 (89%)	9 (11%)	6	26
3	C	186/202 (92%)	161 (87%)	25 (13%)	4	20
3	G	186/202 (92%)	164 (88%)	22 (12%)	5	23
3	K	186/202 (92%)	162 (87%)	24 (13%)	4	21
3	O	186/202 (92%)	162 (87%)	24 (13%)	4	21
4	D	180/416 (43%)	159 (88%)	21 (12%)	5	23
4	H	180/416 (43%)	158 (88%)	22 (12%)	5	23
4	L	180/416 (43%)	158 (88%)	22 (12%)	5	23
4	P	180/416 (43%)	155 (86%)	25 (14%)	3	20
All	All	4040/5148 (78%)	3599 (89%)	441 (11%)	6	26

5 of 441 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	70	ASN
3	G	5	THR
4	L	148	CYS
3	K	30	ARG
1	E	196	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
4	P	35	HIS

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](#)

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](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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(Å ²)	Q<0.9
1	A	655/665 (98%)	-0.17	10 (1%) 73 64	59, 142, 223, 269	0
1	E	655/665 (98%)	0.18	38 (5%) 23 19	64, 184, 281, 343	0
1	I	655/665 (98%)	-0.16	11 (1%) 70 61	58, 140, 220, 266	0
1	M	655/665 (98%)	-0.11	10 (1%) 73 64	67, 155, 224, 267	0
2	B	94/112 (83%)	0.30	8 (8%) 10 10	96, 169, 218, 269	0
2	F	94/112 (83%)	0.23	5 (5%) 26 23	105, 177, 229, 279	0
2	J	94/112 (83%)	0.12	3 (3%) 47 37	94, 170, 215, 266	0
2	N	94/112 (83%)	0.48	11 (11%) 4 5	104, 186, 232, 277	0
3	C	213/233 (91%)	-0.14	0 100 100	72, 118, 162, 211	0
3	G	213/233 (91%)	-0.03	1 (0%) 91 86	86, 122, 167, 219	0
3	K	213/233 (91%)	-0.16	1 (0%) 91 86	72, 115, 156, 202	0
3	O	213/233 (91%)	-0.15	3 (1%) 75 65	81, 126, 165, 227	0
4	D	214/470 (45%)	0.25	5 (2%) 60 51	82, 148, 198, 280	0
4	H	214/470 (45%)	0.43	11 (5%) 28 24	81, 146, 196, 265	0
4	L	214/470 (45%)	0.18	3 (1%) 75 65	74, 135, 181, 258	0
4	P	214/470 (45%)	0.51	10 (4%) 31 26	79, 159, 203, 291	0
All	All	4704/5920 (79%)	0.02	130 (2%) 53 42	58, 145, 234, 343	0

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	532	TYR	6.3
1	E	501	LEU	5.5
1	E	525	ASP	5.0
1	E	531	ALA	4.9
1	E	535	VAL	4.7

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.