wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 07:06 pm BST

PDB ID	:	1G2C
Title	:	HUMAN RESPIRATORY SYNCYTIAL VIRUS FUSION PROTEIN CORE
Authors	:	Zhao, X.; Singh, M.; Malashkevich, V.N.; Kim, P.S.
Deposited on	:	2000-10-18
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*
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 (1)) were used in the production of this report:

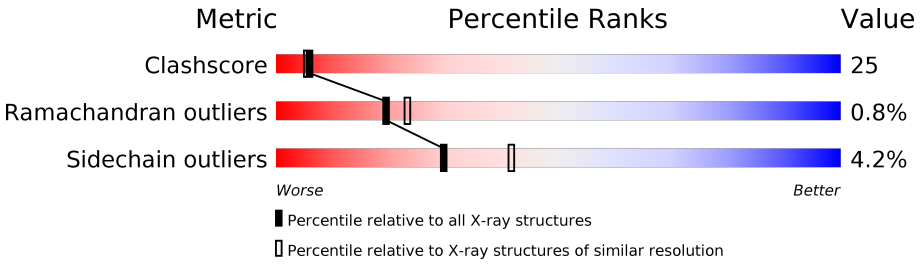
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 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)

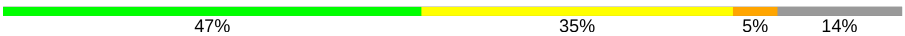
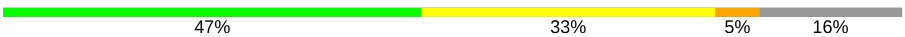
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	52	56% 37%
1	C	52	48% 42% 8%
1	E	52	62% 29% 8%
1	G	52	67% 25% 6%
1	I	52	56% 40%
1	K	52	71% 19% 8%
1	M	52	56% 33% 10%
1	O	52	69% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	Q	52	
1	S	52	
1	U	52	
1	W	52	
2	B	43	
2	D	43	
2	F	43	
2	H	43	
2	J	43	
2	L	43	
2	N	43	
2	P	43	
2	R	43	
2	T	43	
2	V	43	
2	X	43	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8942 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 PROTEIN (F).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	50	Total	C	N	O	0	0	0
			379	242	63	74			
1	C	48	Total	C	N	O	0	0	0
			361	232	59	70			
1	E	48	Total	C	N	O	0	0	0
			361	232	59	70			
1	G	49	Total	C	N	O	0	0	0
			369	236	61	72			
1	I	50	Total	C	N	O	0	0	0
			379	242	64	73			
1	K	48	Total	C	N	O	0	0	0
			362	231	61	70			
1	M	47	Total	C	N	O	0	0	0
			354	227	58	69			
1	O	50	Total	C	N	O	0	0	0
			379	244	63	72			
1	Q	49	Total	C	N	O	0	0	0
			369	236	61	72			
1	S	50	Total	C	N	O	0	0	0
			379	242	64	73			
1	U	50	Total	C	N	O	0	0	0
			379	242	63	74			
1	W	50	Total	C	N	O	0	0	0
			379	242	63	74			

- Molecule 2 is a protein called FUSION PROTEIN (F).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	40	Total	C	N	O	0	0	0
			314	197	53	64			
2	D	36	Total	C	N	O	0	0	0
			289	183	48	58			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	F	36	Total	C	N	O	0	0	0
			289	183	48	58			
2	H	37	Total	C	N	O	0	0	0
			296	188	49	59			
2	J	36	Total	C	N	O	0	0	0
			289	183	48	58			
2	L	40	Total	C	N	O	0	0	0
			327	210	52	65			
2	N	37	Total	C	N	O	0	0	0
			296	188	49	59			
2	P	37	Total	C	N	O	0	0	0
			296	188	49	59			
2	R	37	Total	C	N	O	0	0	0
			296	188	49	59			
2	T	38	Total	C	N	O	0	0	0
			304	192	51	61			
2	V	38	Total	C	N	O	0	0	0
			304	192	51	61			
2	X	38	Total	C	N	O	0	0	0
			304	192	51	61			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total	O	0	0
			28	28		
3	B	41	Total	O	0	0
			41	41		
3	C	26	Total	O	0	0
			26	26		
3	D	38	Total	O	0	0
			38	38		
3	E	28	Total	O	0	0
			28	28		
3	F	41	Total	O	0	0
			41	41		
3	G	38	Total	O	0	0
			38	38		
3	H	53	Total	O	0	0
			53	53		
3	I	70	Total	O	0	0
			70	70		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	J	45	Total O 45 45	0	0
3	K	30	Total O 30 30	0	0
3	L	40	Total O 40 40	0	0
3	M	36	Total O 36 36	0	0
3	N	42	Total O 42 42	0	0
3	O	35	Total O 35 35	0	0
3	P	32	Total O 32 32	0	0
3	Q	32	Total O 32 32	0	0
3	R	42	Total O 42 42	0	0
3	S	26	Total O 26 26	0	0
3	T	39	Total O 39 39	0	0
3	U	22	Total O 22 22	0	0
3	V	42	Total O 42 42	0	0
3	W	17	Total O 17 17	0	0
3	X	45	Total O 45 45	0	0

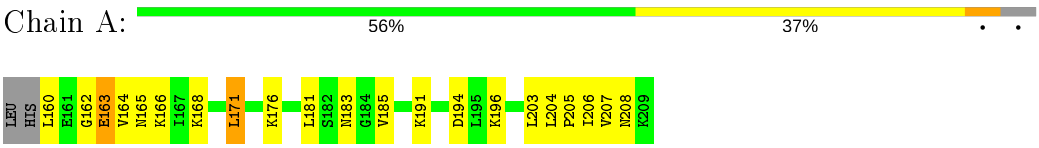
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

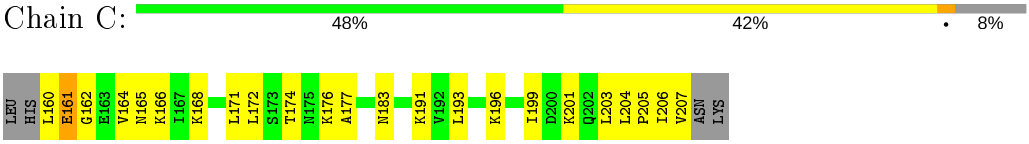
- Molecule 1: FUSION PROTEIN (F)

Chain A:



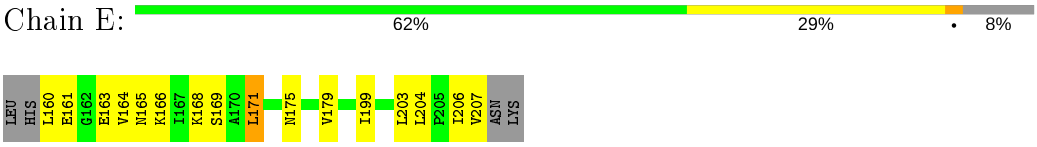
- Molecule 1: FUSION PROTEIN (F)

Chain C:



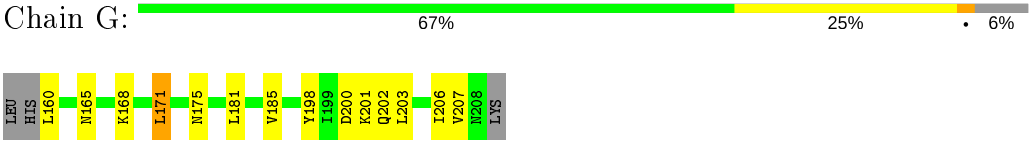
- Molecule 1: FUSION PROTEIN (F)

Chain E:



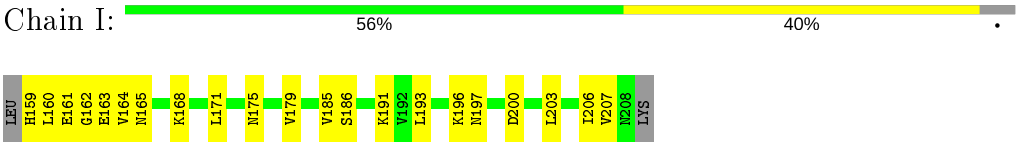
- Molecule 1: FUSION PROTEIN (F)

Chain G:



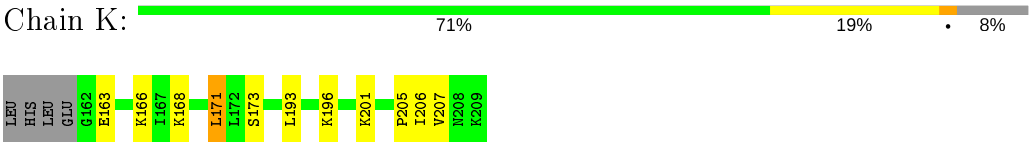
- Molecule 1: FUSION PROTEIN (F)

Chain I:



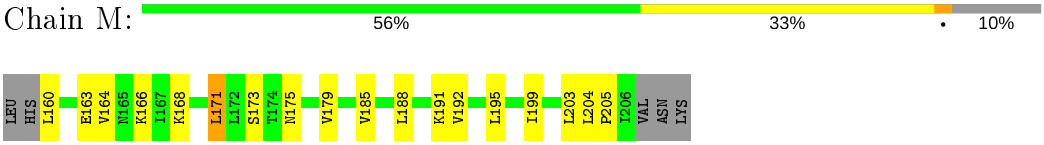
- Molecule 1: FUSION PROTEIN (F)

Chain K:



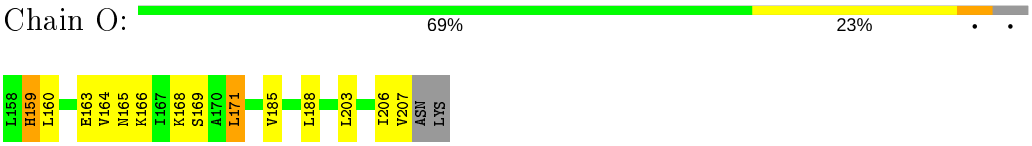
• Molecule 1: FUSION PROTEIN (F)

Chain M:



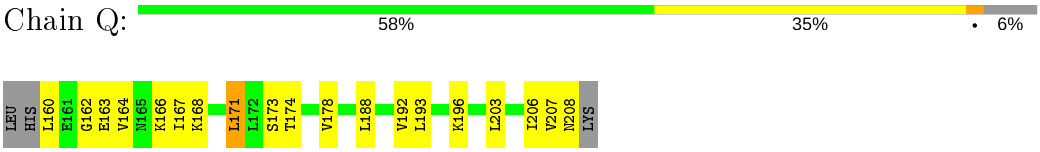
• Molecule 1: FUSION PROTEIN (F)

Chain O:



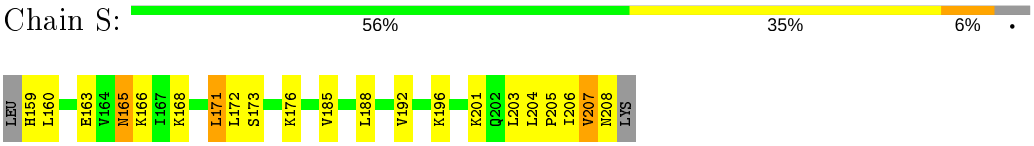
• Molecule 1: FUSION PROTEIN (F)

Chain Q:



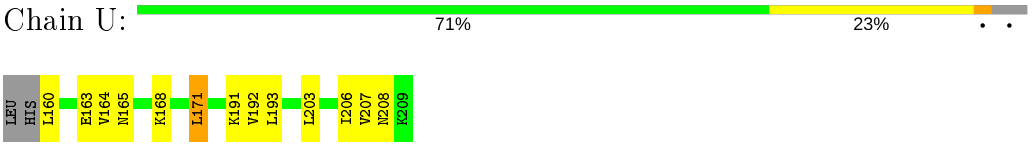
• Molecule 1: FUSION PROTEIN (F)

Chain S:



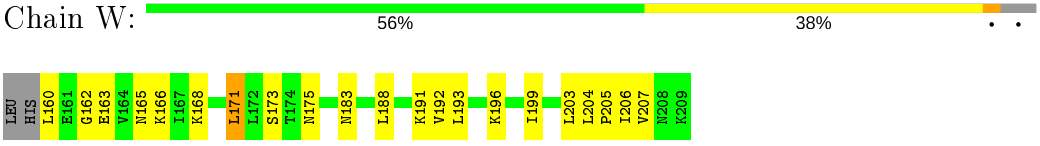
• Molecule 1: FUSION PROTEIN (F)

Chain U:



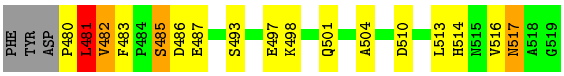
• Molecule 1: FUSION PROTEIN (F)

Chain W:



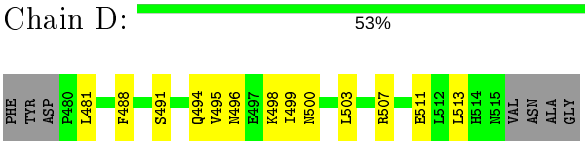
• Molecule 2: FUSION PROTEIN (F)

Chain B:



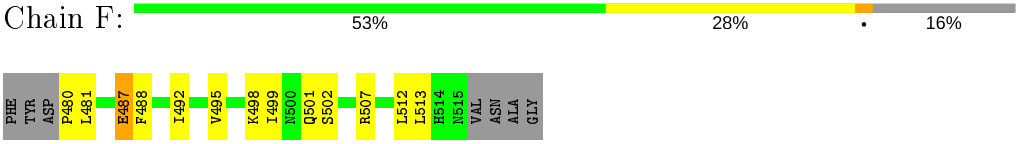
• Molecule 2: FUSION PROTEIN (F)

Chain D:



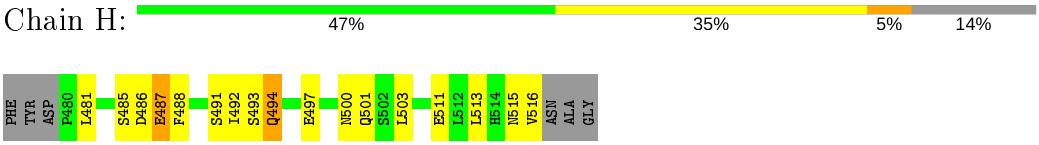
• Molecule 2: FUSION PROTEIN (F)

Chain F:



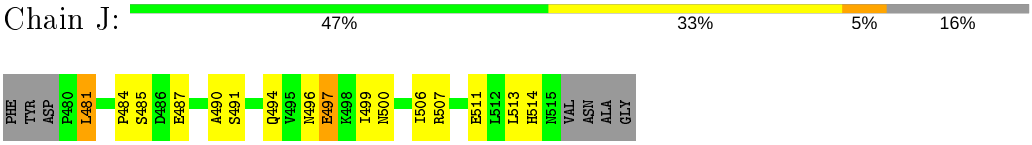
• Molecule 2: FUSION PROTEIN (F)

Chain H:



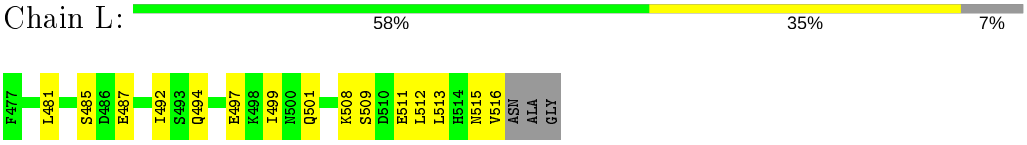
• Molecule 2: FUSION PROTEIN (F)

Chain J:



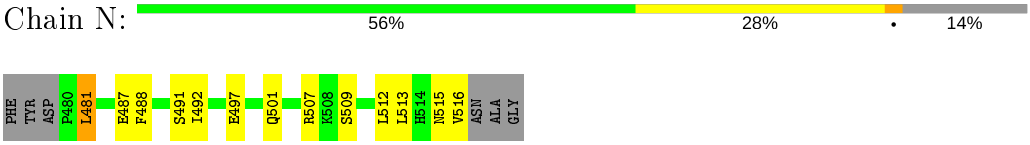
• Molecule 2: FUSION PROTEIN (F)

Chain L:



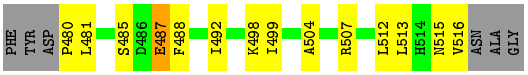
• Molecule 2: FUSION PROTEIN (F)

Chain N:



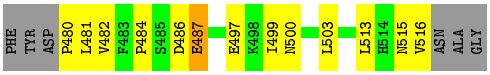
• Molecule 2: FUSION PROTEIN (F)

Chain P:



• Molecule 2: FUSION PROTEIN (F)

Chain R:



• Molecule 2: FUSION PROTEIN (F)

Chain T:

• Molecule 2: FUSION PROTEIN (F)

Chain V:

• Molecule 2: FUSION PROTEIN (F)

Chain X:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	67.90Å 71.54Å 76.45Å 81.34° 73.80° 60.72°	Depositor
Resolution (Å)	10.00 – 2.30	Depositor
% Data completeness (in resolution range)	94.7 (10.00-2.30)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.5	Depositor
R, R_{free}	0.233 , 0.286	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8942	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.35	0/380	0.57	0/511
1	C	0.42	0/362	0.60	0/489
1	E	0.41	0/362	0.61	0/489
1	G	0.35	0/370	0.57	0/500
1	I	0.32	0/381	0.58	0/515
1	K	0.37	0/363	0.61	0/488
1	M	0.46	0/355	0.62	0/479
1	O	0.43	0/381	0.61	0/515
1	Q	0.36	0/370	0.60	0/500
1	S	0.33	0/381	0.56	0/515
1	U	0.35	0/380	0.56	0/511
1	W	0.33	0/380	0.55	0/511
2	B	0.42	0/319	0.62	0/429
2	D	0.40	0/294	0.60	0/396
2	F	0.40	0/294	0.60	0/396
2	H	0.44	0/301	0.61	0/406
2	J	0.44	0/294	0.60	0/396
2	L	0.46	0/334	0.66	0/452
2	N	0.41	0/301	0.66	1/406 (0.2%)
2	P	0.47	0/301	0.74	1/406 (0.2%)
2	R	0.41	0/301	0.58	0/406
2	T	0.44	0/309	0.61	0/417
2	V	0.46	0/309	0.63	0/417
2	X	0.38	0/309	0.58	0/417
All	All	0.40	0/8131	0.60	2/10967 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	N	481	LEU	N-CA-C	-5.41	96.40	111.00
2	P	480	PRO	N-CA-CB	5.14	109.47	103.30

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	379	0	419	34	0
1	C	361	0	400	23	0
1	E	361	0	400	23	0
1	G	369	0	406	21	0
1	I	379	0	413	25	0
1	K	362	0	402	17	0
1	M	354	0	391	32	0
1	O	379	0	418	23	0
1	Q	369	0	406	36	0
1	S	379	0	413	33	0
1	U	379	0	419	27	0
1	W	379	0	419	26	0
2	B	314	0	304	21	0
2	D	289	0	281	19	0
2	F	289	0	281	15	0
2	H	296	0	290	25	0
2	J	289	0	281	20	0
2	L	327	0	311	32	0
2	N	296	0	290	22	0
2	P	296	0	290	21	0
2	R	296	0	290	19	0
2	T	304	0	296	16	0
2	V	304	0	296	18	0
2	X	304	0	296	21	0
3	A	28	0	0	9	0
3	B	41	0	0	3	0
3	C	26	0	0	1	0
3	D	38	0	0	1	0
3	E	28	0	0	1	0
3	F	41	0	0	2	0
3	G	38	0	0	0	0
3	H	53	0	0	2	0
3	I	70	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	45	0	0	3	0
3	K	30	0	0	2	0
3	L	40	0	0	2	0
3	M	36	0	0	5	0
3	N	42	0	0	5	0
3	O	35	0	0	1	0
3	P	32	0	0	0	0
3	Q	32	0	0	3	0
3	R	42	0	0	7	0
3	S	26	0	0	1	0
3	T	39	0	0	1	0
3	U	22	0	0	1	0
3	V	42	0	0	3	0
3	W	17	0	0	1	0
3	X	45	0	0	3	0
All	All	8942	0	8412	411	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 25.

The worst 5 of 411 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:A:162:GLY:HA2	1:A:165:ASN:ND2	1.75	1.01
1:I:206:ILE:HD12	2:J:481:LEU:HD11	1.40	0.99
1:A:160:LEU:HD21	1:E:160:LEU:HG	1.44	0.96
1:O:171:LEU:HD13	2:R:513:LEU:HD11	1.51	0.91
3:M:751:HOH:O	1:Q:164:VAL:HG21	1.71	0.91

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	48/52 (92%)	47 (98%)	1 (2%)	0	100 100
1	C	46/52 (88%)	45 (98%)	0	1 (2%)	6 5
1	E	46/52 (88%)	44 (96%)	2 (4%)	0	100 100
1	G	47/52 (90%)	46 (98%)	1 (2%)	0	100 100
1	I	48/52 (92%)	48 (100%)	0	0	100 100
1	K	46/52 (88%)	44 (96%)	2 (4%)	0	100 100
1	M	45/52 (86%)	45 (100%)	0	0	100 100
1	O	48/52 (92%)	48 (100%)	0	0	100 100
1	Q	47/52 (90%)	46 (98%)	1 (2%)	0	100 100
1	S	48/52 (92%)	47 (98%)	0	1 (2%)	7 5
1	U	48/52 (92%)	45 (94%)	3 (6%)	0	100 100
1	W	48/52 (92%)	47 (98%)	1 (2%)	0	100 100
2	B	38/43 (88%)	33 (87%)	2 (5%)	3 (8%)	1 0
2	D	34/43 (79%)	32 (94%)	2 (6%)	0	100 100
2	F	34/43 (79%)	34 (100%)	0	0	100 100
2	H	35/43 (81%)	34 (97%)	1 (3%)	0	100 100
2	J	34/43 (79%)	31 (91%)	2 (6%)	1 (3%)	4 3
2	L	38/43 (88%)	35 (92%)	3 (8%)	0	100 100
2	N	35/43 (81%)	33 (94%)	2 (6%)	0	100 100
2	P	35/43 (81%)	35 (100%)	0	0	100 100
2	R	35/43 (81%)	35 (100%)	0	0	100 100
2	T	36/43 (84%)	34 (94%)	1 (3%)	1 (3%)	5 3
2	V	36/43 (84%)	35 (97%)	1 (3%)	0	100 100
2	X	36/43 (84%)	35 (97%)	0	1 (3%)	5 3
All	All	991/1140 (87%)	958 (97%)	25 (2%)	8 (1%)	19 23

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	481	LEU
2	T	516	VAL
2	X	516	VAL
2	J	485	SER
2	B	517	ASN

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	46/48 (96%)	44 (96%)	2 (4%)	29	40
1	C	44/48 (92%)	41 (93%)	3 (7%)	16	21
1	E	44/48 (92%)	43 (98%)	1 (2%)	50	67
1	G	45/48 (94%)	43 (96%)	2 (4%)	28	39
1	I	46/48 (96%)	43 (94%)	3 (6%)	17	23
1	K	44/48 (92%)	42 (96%)	2 (4%)	27	39
1	M	43/48 (90%)	42 (98%)	1 (2%)	50	67
1	O	46/48 (96%)	44 (96%)	2 (4%)	29	40
1	Q	45/48 (94%)	43 (96%)	2 (4%)	28	39
1	S	46/48 (96%)	42 (91%)	4 (9%)	10	12
1	U	46/48 (96%)	44 (96%)	2 (4%)	29	40
1	W	46/48 (96%)	43 (94%)	3 (6%)	17	23
2	B	36/39 (92%)	34 (94%)	2 (6%)	21	29
2	D	34/39 (87%)	34 (100%)	0	100	100
2	F	34/39 (87%)	33 (97%)	1 (3%)	42	58
2	H	35/39 (90%)	33 (94%)	2 (6%)	20	28
2	J	34/39 (87%)	32 (94%)	2 (6%)	19	27
2	L	38/39 (97%)	38 (100%)	0	100	100
2	N	35/39 (90%)	34 (97%)	1 (3%)	42	58
2	P	35/39 (90%)	33 (94%)	2 (6%)	20	28
2	R	35/39 (90%)	34 (97%)	1 (3%)	42	58
2	T	36/39 (92%)	35 (97%)	1 (3%)	43	60
2	V	36/39 (92%)	35 (97%)	1 (3%)	43	60
2	X	36/39 (92%)	35 (97%)	1 (3%)	43	60
All	All	965/1044 (92%)	924 (96%)	41 (4%)	30	42

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	171	LEU
1	O	159	HIS
1	W	165	ASN
1	K	173	SER
1	M	171	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	208	ASN
2	L	501	GLN
2	T	517	ASN
2	J	500	ASN
2	N	500	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 carbohydrates 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

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.