wwPDB X-ray Structure Validation Summary Report (i)

Feb 11, 2024 – 12:28 AM EST

PDB ID	:	2QQP
Title	:	Crystal Structure of Authentic Providence Virus
Authors	:	Speir, J.A.; Taylor, D.J.; Johnson, J.E.
Deposited on	:	2007-07-26
Resolution	:	3.80 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

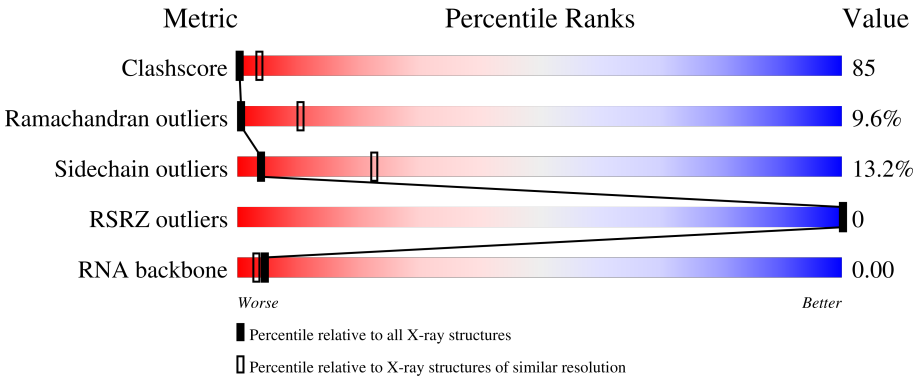
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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.



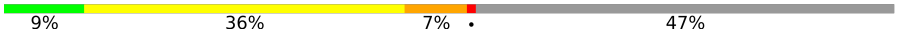
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)
RNA backbone	3102	1036 (4.60-3.00)

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% 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	556	17% 55% 14% • 13%
1	C	556	15% 56% 15% • 13%
1	E	556	18% 63% 12% • 7%
1	G	556	18% 58% 15% • 7%
2	B	75	20% 24% 8% 48%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	75	
2	F	75	
2	H	75	
3	R	4	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 16634 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 p81.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	484	Total	C	N	O	S	0	0	0
			3714	2359	604	740	11			
1	C	484	Total	C	N	O	S	0	0	0
			3714	2359	604	740	11			
1	E	517	Total	C	N	O	S	0	0	0
			3948	2505	645	785	13			
1	G	518	Total	C	N	O	S	0	0	0
			3957	2510	646	788	13			

- Molecule 2 is a protein called p81.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	39	Total	C	N	O	0	0	0
			255	162	44	49			
2	D	40	Total	C	N	O	0	0	0
			263	166	46	51			
2	F	67	Total	C	N	O	0	0	0
			448	282	84	82			
2	H	38	Total	C	N	O	0	0	0
			250	159	43	48			

- Molecule 3 is a RNA chain called RNA (5'-R(*UP*UP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	4	Total	C	N	O	P	0	0	0
			77	36	8	30	3			

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	1	Total Ca 1 1	0	0

- Molecule 5 is water.

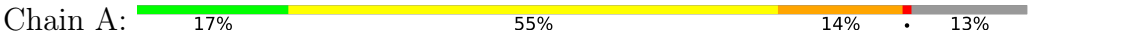
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O 1 1	0	0
5	C	2	Total O 2 2	0	0
5	E	2	Total O 2 2	0	0
5	G	1	Total O 1 1	0	0

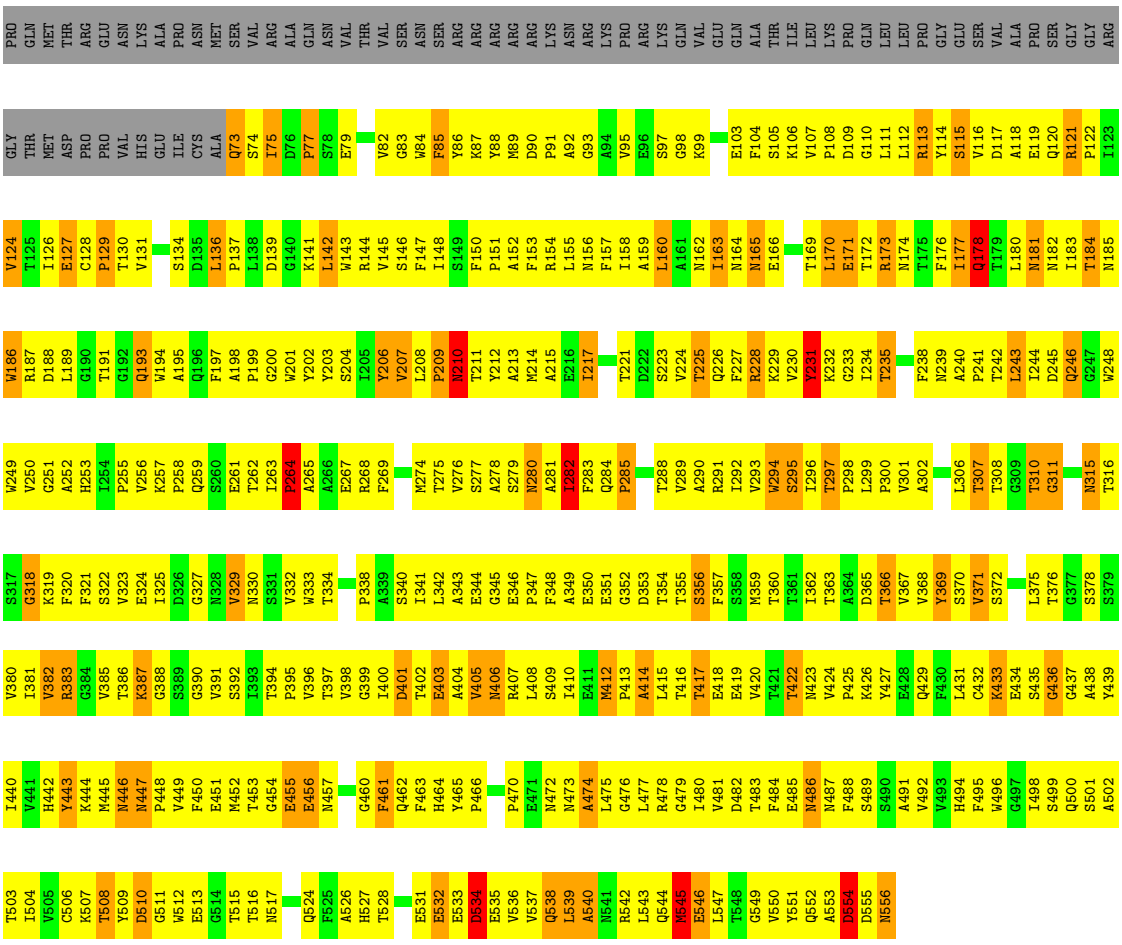
3 Residue-property plots (i)

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 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: p81

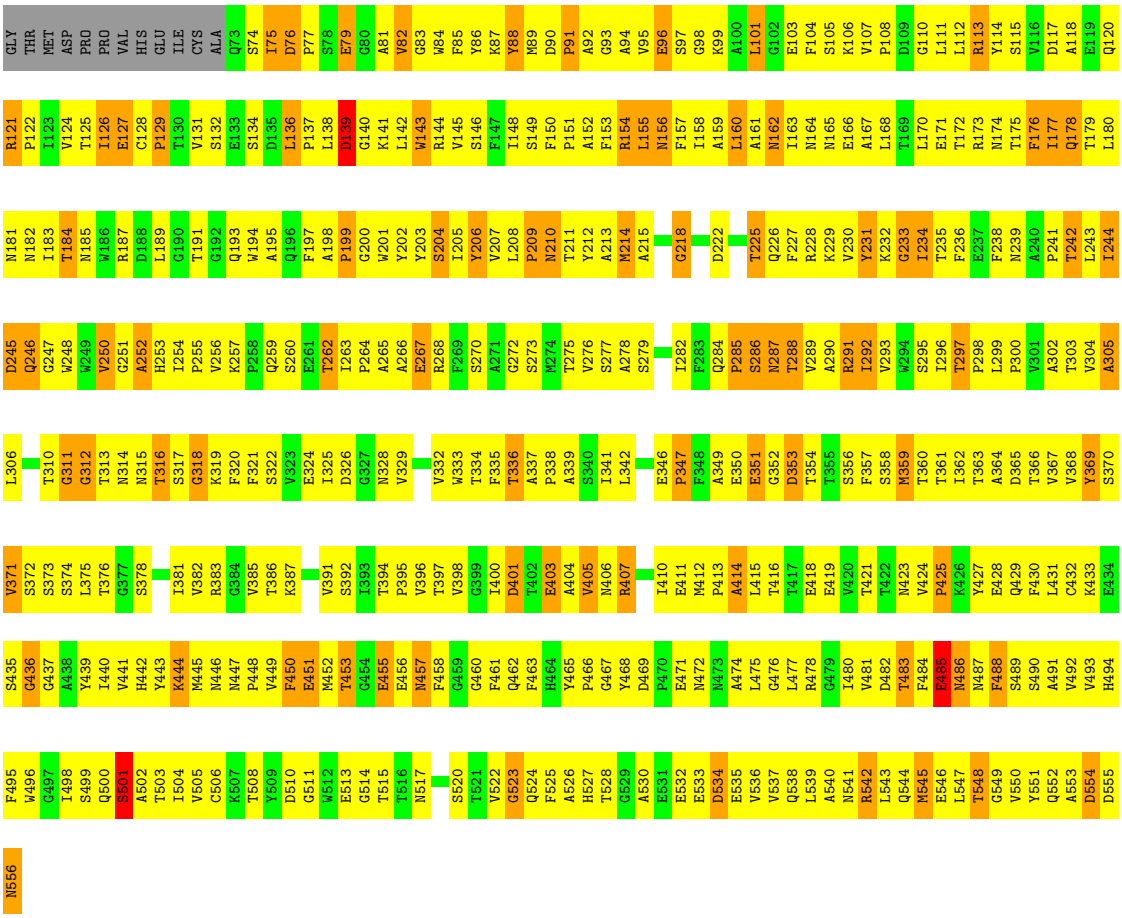
Chain A:





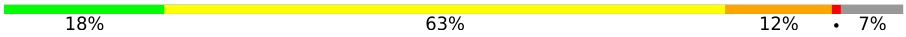
- Molecule 1: p81

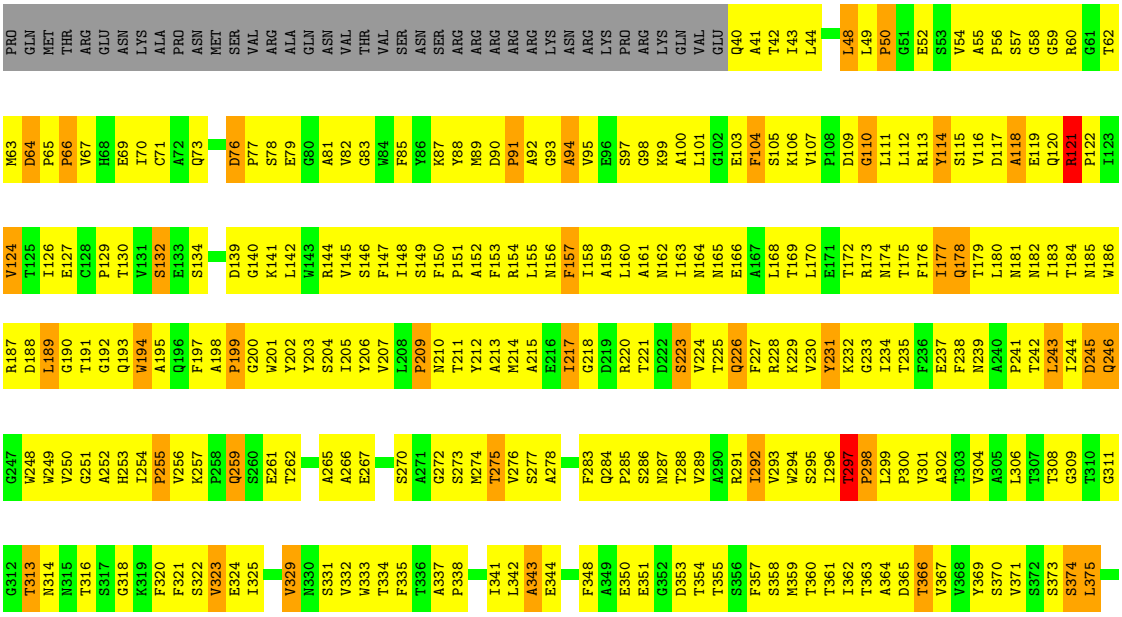
Chain C:

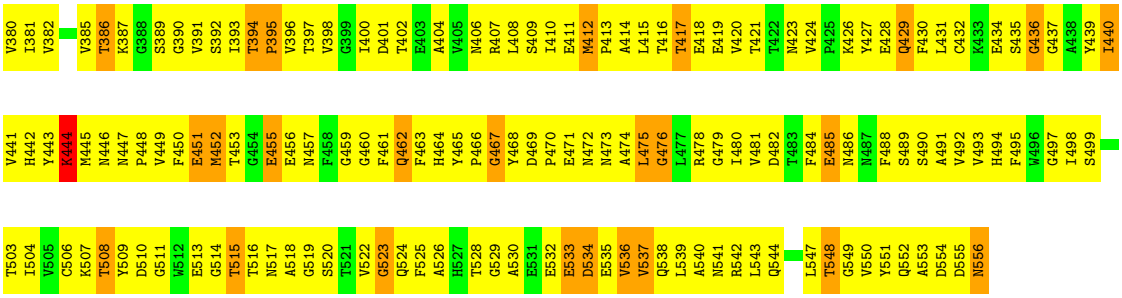


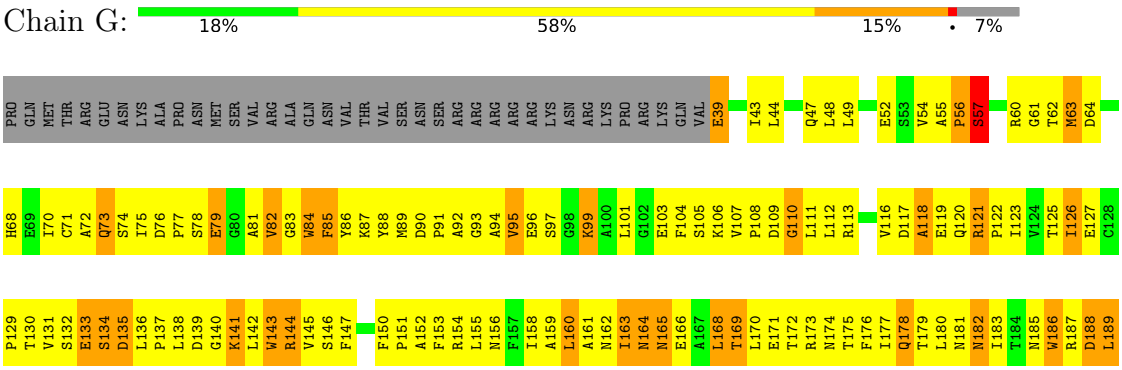
• Molecule 1: p81

Chain E:







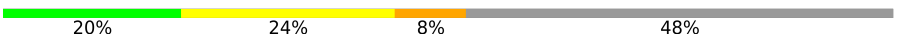


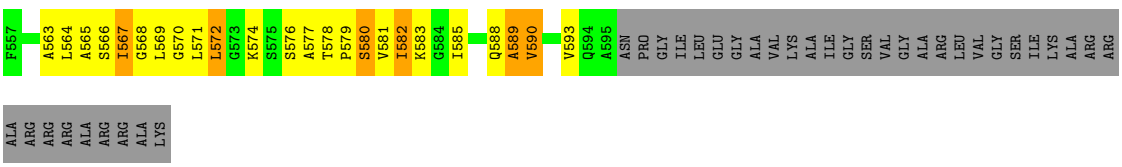
• Molecule 1: p81

Chain G:

• Molecule 2: p81

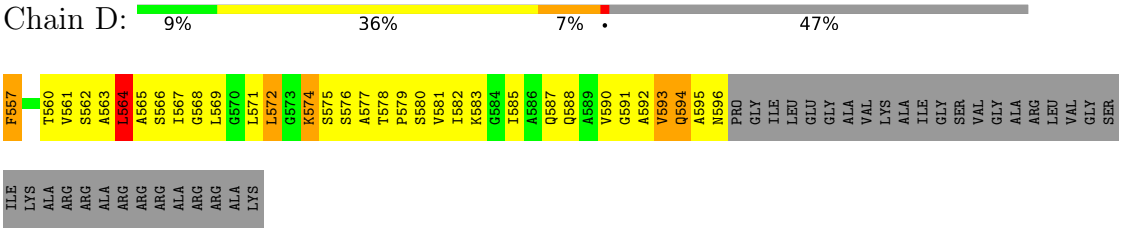
Chain B:





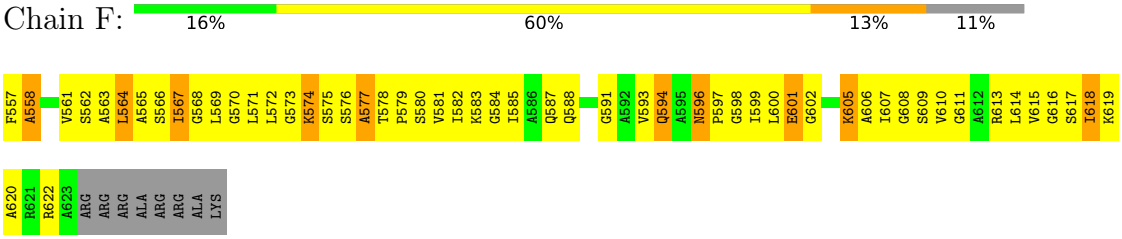
• Molecule 2: p81

Chain D:



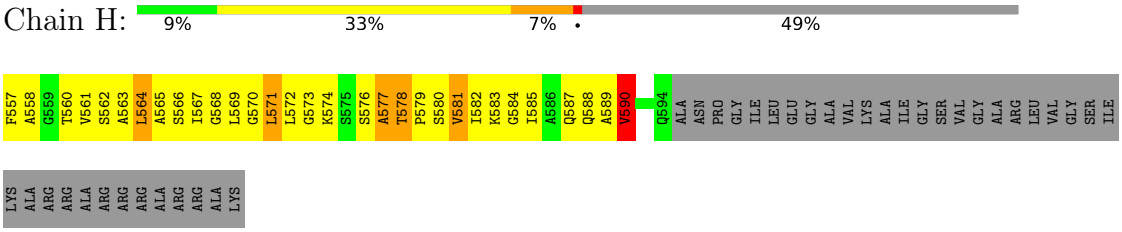
• Molecule 2: p81

Chain F:



• Molecule 2: p81

Chain H:



• Molecule 3: RNA (5'-R(*UP*UP*UP*U)-3')

Chain R:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	659.79Å 434.07Å 415.85Å 90.00° 126.13° 90.00°	Depositor
Resolution (Å)	49.42 – 3.80 50.01 – 3.80	Depositor EDS
% Data completeness (in resolution range)	28.6 (49.42-3.80) 28.6 (50.01-3.80)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 3.77Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.285 , (Not available) 0.279 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	35.1	Xtriage
Anisotropy	0.086	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , -9.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.059 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.75	EDS
Total number of atoms	16634	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.81% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.54	0/3807	0.78	0/5203
1	C	0.53	0/3807	0.78	0/5203
1	E	0.55	0/4047	0.78	0/5532
1	G	0.52	0/4056	0.79	0/5544
2	B	0.60	0/256	0.72	0/346
2	D	0.55	0/264	0.83	0/357
2	F	0.58	0/450	0.85	0/606
2	H	0.51	0/251	0.74	0/339
3	R	1.57	0/84	0.91	0/128
All	All	0.55	0/17022	0.78	0/23258

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	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	206	TYR	Sidechain
1	C	206	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	3714	0	3548	661	0
1	C	3714	0	3548	689	0
1	E	3948	0	3786	744	0
1	G	3957	0	3792	703	0
2	B	255	0	275	40	0
2	D	263	0	281	66	0
2	F	448	0	491	90	0
2	H	250	0	270	47	0
3	R	77	0	42	4	0
4	A	1	0	0	0	0
4	G	1	0	0	0	0
5	A	1	0	0	0	0
5	C	2	0	0	0	0
5	E	2	0	0	0	0
5	G	1	0	0	0	0
All	All	16634	0	16033	2780	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 85.

The worst 5 of 2780 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:C:475:LEU:HD11	1:E:267:GLU:CG	1.36	1.52
1:C:297:THR:HG23	1:C:476:GLY:CA	1.55	1.35
1:C:120:GLN:NE2	1:C:208:LEU:HB3	1.45	1.29
1:E:297:THR:HG22	1:E:298:PRO:CD	1.64	1.27
1:C:475:LEU:HD11	1:E:267:GLU:CB	1.68	1.21

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	482/556 (87%)	357 (74%)	80 (17%)	45 (9%)	0 11
1	C	482/556 (87%)	358 (74%)	74 (15%)	50 (10%)	0 9
1	E	515/556 (93%)	379 (74%)	95 (18%)	41 (8%)	1 14
1	G	516/556 (93%)	373 (72%)	87 (17%)	56 (11%)	0 8
2	B	37/75 (49%)	21 (57%)	12 (32%)	4 (11%)	0 8
2	D	38/75 (51%)	30 (79%)	6 (16%)	2 (5%)	2 23
2	F	65/75 (87%)	47 (72%)	12 (18%)	6 (9%)	1 12
2	H	36/75 (48%)	16 (44%)	16 (44%)	4 (11%)	0 7
All	All	2171/2524 (86%)	1581 (73%)	382 (18%)	208 (10%)	0 10

5 of 208 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	209	PRO
1	A	297	THR
1	A	318	GLY
1	A	405	VAL
1	A	455	GLU

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	403/466 (86%)	339 (84%)	64 (16%)	2 16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	403/466 (86%)	348 (86%)	55 (14%)	3	22
1	E	429/466 (92%)	385 (90%)	44 (10%)	7	31
1	G	430/466 (92%)	376 (87%)	54 (13%)	4	24
2	B	25/49 (51%)	22 (88%)	3 (12%)	5	25
2	D	26/49 (53%)	21 (81%)	5 (19%)	1	10
2	F	43/49 (88%)	36 (84%)	7 (16%)	2	15
2	H	25/49 (51%)	21 (84%)	4 (16%)	2	16
All	All	1784/2060 (87%)	1548 (87%)	236 (13%)	4	22

5 of 236 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	501	SER
1	G	464	HIS
1	E	255	PRO
1	G	458	PHE
1	G	274	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	120	GLN
1	G	47	GLN
1	C	462	GLN
1	C	181	ASN
1	C	464	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	3/4 (75%)	3 (100%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	2	U
3	R	3	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	R	4	U

There are no RNA pucker outliers to report.

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 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	484/556 (87%)	-0.50	0	100	100	7, 20, 30, 48	0
1	C	484/556 (87%)	-0.49	0	100	100	8, 19, 30, 43	0
1	E	517/556 (92%)	-0.49	0	100	100	7, 21, 33, 52	0
1	G	518/556 (93%)	-0.51	0	100	100	7, 20, 31, 53	0
2	B	39/75 (52%)	-0.53	0	100	100	10, 20, 36, 40	0
2	D	40/75 (53%)	-0.51	0	100	100	10, 21, 35, 40	0
2	F	67/75 (89%)	-0.36	0	100	100	13, 20, 35, 39	0
2	H	38/75 (50%)	-0.54	0	100	100	11, 18, 29, 38	0
3	R	4/4 (100%)	0.18	0	100	100	54, 56, 57, 61	0
All	All	2191/2528 (86%)	-0.50	0	100	100	7, 20, 32, 61	0

There are no RSRZ outliers to report.

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CA	A	557	1/1	0.98	0.10	44,44,44,44	0
4	CA	G	557	1/1	0.99	0.04	73,73,73,73	0

6.5 Other polymers (i)

There are no such residues in this entry.