wwPDB X-ray Structure Validation Summary Report (i)

Mar 23, 2024 – 07:55 PM EDT

PDB ID	:	1AFV
Title	:	HIV-1 CAPSID PROTEIN (P24) COMPLEX WITH FAB25.3
Authors	:	Momany, C.; Kovari, L.C.; Prongay, A.J.; Keller, W.; Gitti, R.K.; Lee, B.M.; Gorbalenya, A.E.; Tong, L.; Mcclure, J.; Ehrlich, L.S.; Summers, M.F.; Carter, C.; Rossmann, M.G.
Deposited on	:	1997-03-14
Resolution	:	3.70 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 (1)) were used in the production of this report:

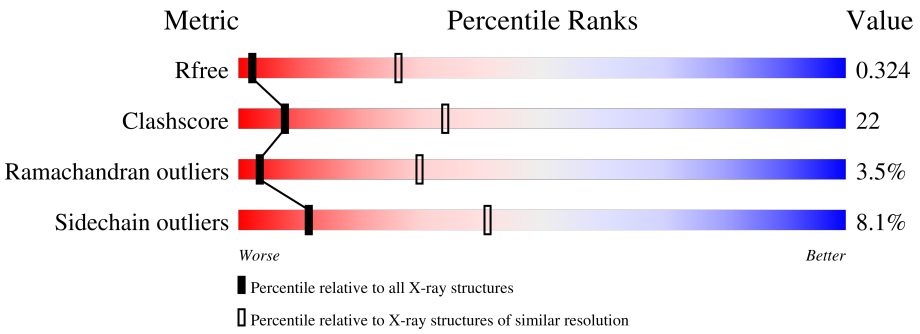
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.1

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 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)

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%

Mol	Chain	Length	Quality of chain
1	A	151	42% 50% 8%
1	B	151	42% 50% 7% •
2	L	217	50% 45% 5%
2	M	217	53% 42% 5%
3	H	220	59% 37% •
3	K	220	60% 35% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9024 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 HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	151	1172	742	205	217	8	0	0	0
1	B	151	1172	742	205	217	8	0	0	0

- Molecule 2 is a protein called ANTIBODY FAB25.3 FRAGMENT (LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	217	1667	1038	281	340	8	0	0	0
2	M	217	1667	1038	281	340	8	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	4	LEU	MET	conflict	UNP 600718
L	18	ARG	SER	conflict	UNP 600718
L	19	ALA	VAL	conflict	UNP 600718
L	30	ASP	GLU	conflict	UNP 600718
L	31	ASN	TYR	conflict	UNP 600718
L	34	ILE	THR	conflict	UNP 600718
L	36	PHE	LEU	conflict	UNP 600718
L	38	ASN	GLN	conflict	UNP 600718
L	40	PHE	TYR	conflict	UNP 600718
L	54	ALA	GLY	conflict	UNP 600718
L	58	LEU	VAL	conflict	UNP 600718
L	59	GLY	GLU	conflict	UNP 600718
L	82	MET	VAL	conflict	UNP 600718
L	85	GLU	ASP	conflict	UNP 600718
L	87	THR	ILE	conflict	UNP 600718
L	89	MET	ILE	conflict	UNP 600718

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	96	LYS	ARG	conflict	UNP 600718
L	97	GLU	LYS	conflict	UNP 600718
L	100	LEU	ALA	conflict	UNP 600718
L	104	ALA	SER	conflict	UNP 600718
L	108	VAL	LEU	conflict	UNP 600718
L	110	LEU	ILE	conflict	UNP 600718
M	4	LEU	MET	conflict	UNP 600718
M	18	ARG	SER	conflict	UNP 600718
M	19	ALA	VAL	conflict	UNP 600718
M	30	ASP	GLU	conflict	UNP 600718
M	31	ASN	TYR	conflict	UNP 600718
M	34	ILE	THR	conflict	UNP 600718
M	36	PHE	LEU	conflict	UNP 600718
M	38	ASN	GLN	conflict	UNP 600718
M	40	PHE	TYR	conflict	UNP 600718
M	54	ALA	GLY	conflict	UNP 600718
M	58	LEU	VAL	conflict	UNP 600718
M	59	GLY	GLU	conflict	UNP 600718
M	82	MET	VAL	conflict	UNP 600718
M	85	GLU	ASP	conflict	UNP 600718
M	87	THR	ILE	conflict	UNP 600718
M	89	MET	ILE	conflict	UNP 600718
M	96	LYS	ARG	conflict	UNP 600718
M	97	GLU	LYS	conflict	UNP 600718
M	100	LEU	ALA	conflict	UNP 600718
M	104	ALA	SER	conflict	UNP 600718
M	108	VAL	LEU	conflict	UNP 600718
M	110	LEU	ILE	conflict	UNP 600718

- Molecule 3 is a protein called ANTIBODY FAB25.3 FRAGMENT (HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	H	220	1672	1062	275	330	5	0	0	0
			Total	C	N	O	S			
3	K	220	1672	1062	275	330	5	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	1	GLN	GLU	conflict	UNP Q99LC4
H	3	GLN	LYS	conflict	UNP Q99LC4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	5	GLN	HIS	conflict	UNP Q99LC4
H	6	GLN	GLU	conflict	UNP Q99LC4
H	7	PRO	SER	conflict	UNP Q99LC4
H	9	SER	PRO	conflict	UNP Q99LC4
H	10	VAL	GLY	conflict	UNP Q99LC4
H	16	ALA	THR	conflict	UNP Q99LC4
H	20	LEU	ILE	conflict	UNP Q99LC4
H	31	SER	ASN	conflict	UNP Q99LC4
H	32	SER	TYR	conflict	UNP Q99LC4
H	34	ILE	LEU	conflict	UNP Q99LC4
H	35	HIS	GLY	conflict	UNP Q99LC4
H	37	ALA	VAL	conflict	UNP Q99LC4
H	43	GLN	HIS	conflict	UNP Q99LC4
H	50	GLU	ASP	conflict	UNP Q99LC4
H	52	HIS	TYR	conflict	UNP Q99LC4
H	54	ASN	GLY	conflict	UNP Q99LC4
H	55	SER	GLY	conflict	UNP Q99LC4
H	57	ASN	TYR	conflict	UNP Q99LC4
H	72	VAL	ALA	conflict	UNP Q99LC4
H	81	VAL	MET	conflict	UNP Q99LC4
H	82	ASP	GLN	conflict	UNP Q99LC4
H	95	TYR	PHE	conflict	UNP Q99LC4
H	99	TRP	PHE	conflict	UNP Q99LC4
H	100	ARG	TYR	conflict	UNP Q99LC4
H	104	PRO	SER	conflict	UNP Q99LC4
H	?	-	TRP	deletion	UNP Q99LC4
H	109	TYR	VAL	conflict	UNP Q99LC4
H	116	LEU	VAL	conflict	UNP Q99LC4
K	1	GLN	GLU	conflict	UNP Q99LC4
K	3	GLN	LYS	conflict	UNP Q99LC4
K	5	GLN	HIS	conflict	UNP Q99LC4
K	6	GLN	GLU	conflict	UNP Q99LC4
K	7	PRO	SER	conflict	UNP Q99LC4
K	9	SER	PRO	conflict	UNP Q99LC4
K	10	VAL	GLY	conflict	UNP Q99LC4
K	16	ALA	THR	conflict	UNP Q99LC4
K	20	LEU	ILE	conflict	UNP Q99LC4
K	31	SER	ASN	conflict	UNP Q99LC4
K	32	SER	TYR	conflict	UNP Q99LC4
K	34	ILE	LEU	conflict	UNP Q99LC4
K	35	HIS	GLY	conflict	UNP Q99LC4
K	37	ALA	VAL	conflict	UNP Q99LC4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	43	GLN	HIS	conflict	UNP Q99LC4
K	50	GLU	ASP	conflict	UNP Q99LC4
K	52	HIS	TYR	conflict	UNP Q99LC4
K	54	ASN	GLY	conflict	UNP Q99LC4
K	55	SER	GLY	conflict	UNP Q99LC4
K	57	ASN	TYR	conflict	UNP Q99LC4
K	72	VAL	ALA	conflict	UNP Q99LC4
K	81	VAL	MET	conflict	UNP Q99LC4
K	82	ASP	GLN	conflict	UNP Q99LC4
K	95	TYR	PHE	conflict	UNP Q99LC4
K	99	TRP	PHE	conflict	UNP Q99LC4
K	100	ARG	TYR	conflict	UNP Q99LC4
K	104	PRO	SER	conflict	UNP Q99LC4
K	?	-	TRP	deletion	UNP Q99LC4
K	109	TYR	VAL	conflict	UNP Q99LC4
K	116	LEU	VAL	conflict	UNP Q99LC4

- Molecule 4 is LEAD (II) ION (three-letter code: PB) (formula: Pb).

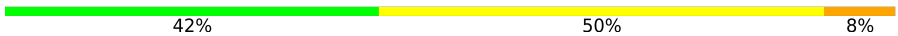
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	L	1	Total Pb 1 1	0	0
4	M	1	Total Pb 1 1	0	0

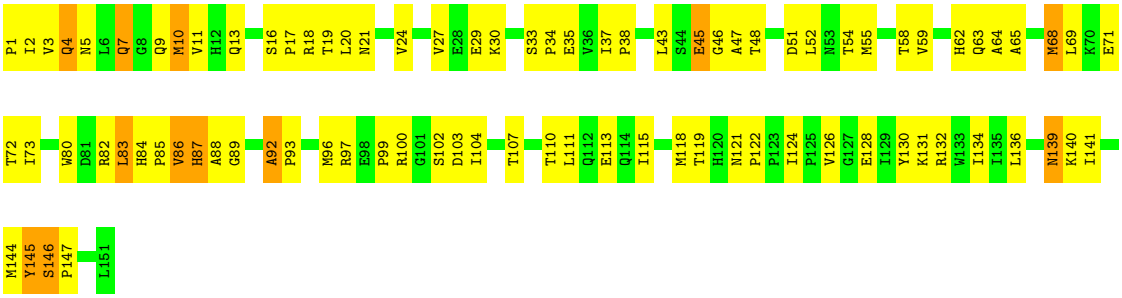
3 Residue-property plots (i)

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. 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.

• Molecule 1: HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID PROTEIN

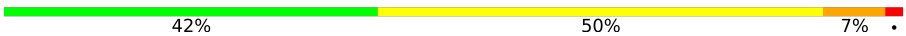
Chain A:

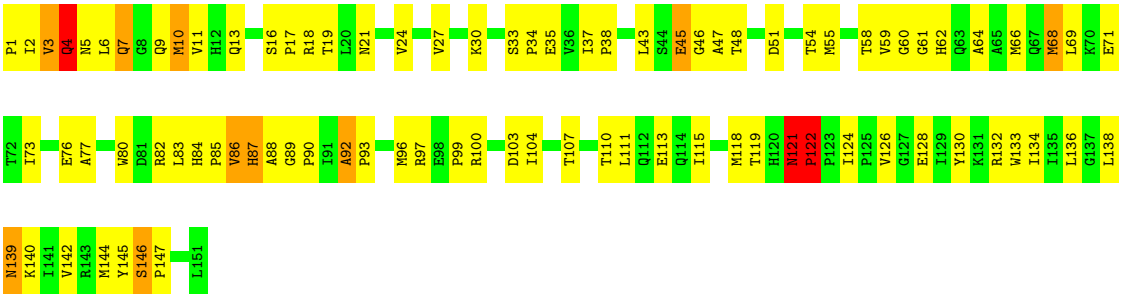




• Molecule 1: HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID PROTEIN

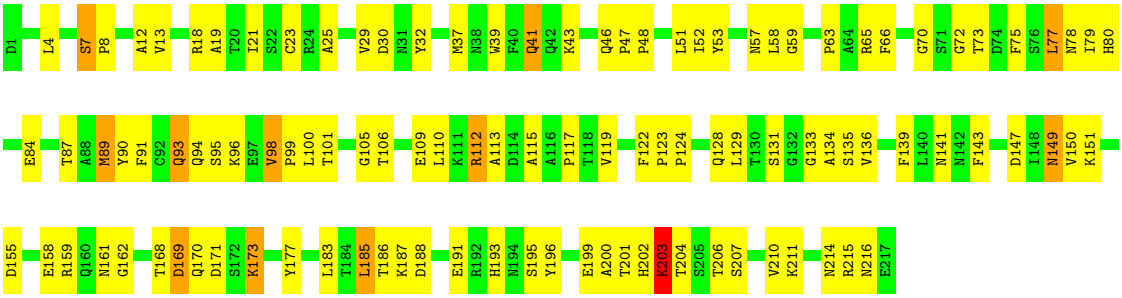
Chain B:





• Molecule 2: ANTIBODY FAB25.3 FRAGMENT (LIGHT CHAIN)

Chain L:



- Molecule 2: ANTIBODY FAB25.3 FRAGMENT (LIGHT CHAIN)

Chain M: 53% 42% 5%

D1 S7 P8 A12 V13 R18 A19 T20 I21 S22 C23 R24 A25 S28 V29 M37 M38 W39 F40 Q41 Q42 K43 Q46 P47 P48 L51 I52 Y53 N57 L58 G59 V62 P63 A64 R65 F66 G70 S71 G72 T73 D74 F75 S76 L77 N78 I79 H80 E84

T87 A88 M89 Y90 F91 C92 Q93 Q94 S95 V98 P99 L100 T101 G105 T106 E109 L110 K111 R112 A113 D114 A115 V119 S125 Q128 L129 T130 S131 G132 G133 A134 S135 F139 D147 I148 N149 V150 K151 W152 K153 I154 D155 E158 R159 Q160 N161 G162 N165 D169

Q170 D171 S172 K173 D174 S175 S181 T182 L183 T184 L185 T186 K187 D188 E191 R192 H193 N194 S195 Y196 T197 C198 E199 A200 T201 H202 K203 T204 V210 K211 S212 F213 N214 R215 N216 E217

- Molecule 3: ANTIBODY FAB25.3 FRAGMENT (HEAVY CHAIN)

Chain H: 59% 37% •

Q1 V2 V12 R13 P14 V18 K19 C22 K23 A24 S25 G26 Y27 T28 F29 T30 S31 S32 W33 I34 H35 Q39 R40 P41 G42 Q43 G44 L45 E46 W47 I48 G49 E50 P53 N54 T58 N59 Y60 N61 E62 K63 F64 V72 S76 S77 T78 A79 Y80 V81 D82 L83

L86 R98 W99 R100 Y101 G102 S103 P104 Y109 Q112 A121 K122 T123 T124 P125 P126 L131 A132 P133 G134 S135 A136 A137 N140 V143 G146 V149 Y152 F153 P154 E155 P156 V157 W161 V170 H171 T172 F173 P174 A175 V176 L177 Q178 L181 Y182 T183 L184

S185 S186 S187 V190 P191 S192 S193 T194 W195 P196 T201 C202 N203 S209 K216 I217 V218 P219 K220

- Molecule 3: ANTIBODY FAB25.3 FRAGMENT (HEAVY CHAIN)

Chain K: 60% 35% •

Q1 V2 V12 R13 P14 V18 K19 K23 A24 S25 G26 F29 T30 S31 S32 H35 Q39 R40 P41 G42 Q43 G44 L45 E46 W47 E50 P53 N54 N59 Y60 N61 E62 K63 F64 K65 L70 T71 V72 S75 S76 S77 T78 A79 Y80 V81 D82 L83 L86

T87 S88 E89 R98 W99 R100 Y101 G102 S103 P104 Y105 Y106 F107 D108 Y109 Q112 A121 P125 P130 L131 S135 A136 A137 N140 S141 M142 V143 G146 V149 F153 P156 V157 W161 V170 V176 L177 Q178 L181 Y182 T183 L184 S185 S186 S187

V190 P191 S192 S193 T194 W195 P196 T201 C202 N203 S209 K216 I217 V218 P219 K220

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, α, β, γ	119.90Å 92.30Å 149.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 3.70 19.98 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (12.00-3.70) 83.9 (19.98-3.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.29 (at 2.98Å)	Xtriage
Refinement program	X-PLOR 3.8	Depositor
R, R_{free}	0.217 , 0.324 0.263 , 0.324	Depositor DCC
R_{free} test set	1433 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	47.3	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 41.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	9024	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.98 % 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 7.9592e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

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

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	0/1201	0.70	2/1635 (0.1%)
1	B	0.36	0/1201	0.65	2/1635 (0.1%)
2	L	0.42	0/1704	0.69	0/2311
2	M	0.41	0/1704	0.69	0/2311
3	H	0.43	0/1722	0.71	0/2360
3	K	0.43	0/1722	0.70	0/2360
All	All	0.41	0/9254	0.69	4/12612 (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	H	0	1
3	K	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	121	ASN	C-N-CD	-13.21	91.53	120.60
1	B	121	ASN	C-N-CD	-9.51	99.68	120.60
1	A	121	ASN	C-N-CA	7.08	151.75	122.00
1	B	121	ASN	N-CA-C	-5.38	96.48	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	H	101	TYR	Sidechain
3	K	101	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	1172	0	1176	62	0
1	B	1172	0	1176	67	0
2	L	1667	0	1599	78	0
2	M	1667	0	1599	74	0
3	H	1672	0	1621	69	0
3	K	1672	0	1621	66	0
4	L	1	0	0	0	0
4	M	1	0	0	0	0
All	All	9024	0	8792	391	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 391 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:B:85:PRO:HB3	2:M:98:VAL:HB	1.51	0.93
2:L:203:LYS:H	2:L:203:LYS:HD3	1.37	0.89
1:A:99:PRO:HG3	1:A:124:ILE:HG21	1.58	0.86
1:B:99:PRO:HG3	1:B:124:ILE:HG21	1.59	0.85
2:L:84:GLU:HA	2:L:110:LEU:HD13	1.59	0.85

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	149/151 (99%)	114 (76%)	27 (18%)	8 (5%)	2	20
1	B	149/151 (99%)	113 (76%)	27 (18%)	9 (6%)	1	18
2	L	215/217 (99%)	190 (88%)	19 (9%)	6 (3%)	5	33
2	M	215/217 (99%)	187 (87%)	21 (10%)	7 (3%)	4	31
3	H	218/220 (99%)	185 (85%)	27 (12%)	6 (3%)	5	33
3	K	218/220 (99%)	187 (86%)	26 (12%)	5 (2%)	6	36
All	All	1164/1176 (99%)	976 (84%)	147 (13%)	41 (4%)	3	30

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	GLU
1	A	145	TYR
1	B	45	GLU
1	B	122	PRO
1	B	145	TYR

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	128/128 (100%)	119 (93%)	9 (7%)	15	45
1	B	128/128 (100%)	116 (91%)	12 (9%)	8	35
2	L	189/189 (100%)	173 (92%)	16 (8%)	10	40
2	M	189/189 (100%)	175 (93%)	14 (7%)	13	44
3	H	190/190 (100%)	175 (92%)	15 (8%)	12	42
3	K	190/190 (100%)	174 (92%)	16 (8%)	11	40
All	All	1014/1014 (100%)	932 (92%)	82 (8%)	11	41

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	98	VAL
3	K	43	GLN
2	M	112	ARG
2	M	203	LYS
3	K	112	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	149	ASN
3	H	112	GLN
2	M	216	ASN
3	H	35	HIS
2	M	38	ASN

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 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 (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.