wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 01:43 pm BST

PDB ID	:	3VXU
Title	:	The complex between T36-5 TCR and HLA-A24 bound to HIV-1 Nef134-10(2F) peptide
Authors	:	Shimizu, A.; Fukai, S.; Yamagata, A.; Iwamoto, A.
Deposited on	:	2012-09-20
Resolution	:	2.70 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 following versions of software and data (see references (1)) were used in the production of this report:

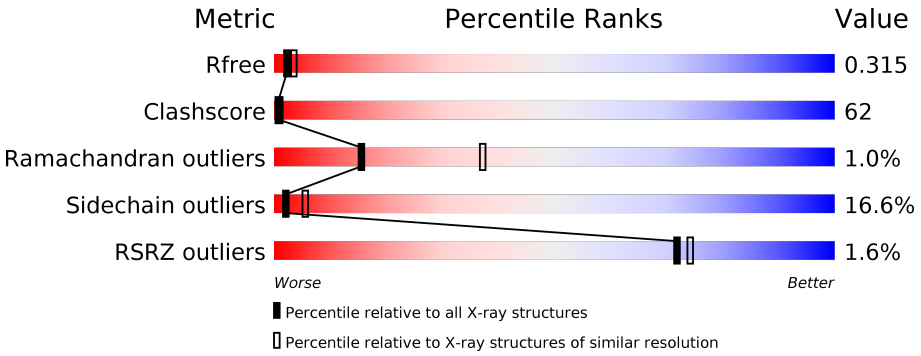
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.70 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.



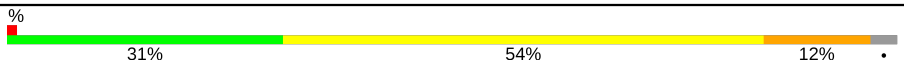
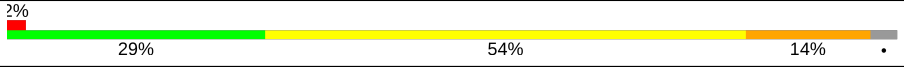
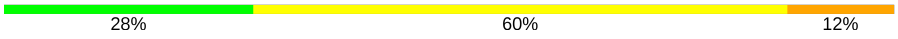
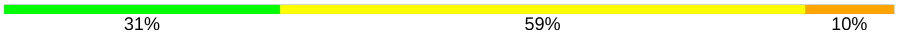
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 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% 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	275	3% 28% 60% 11%
1	F	275	% 35% 55% 11%
2	B	100	2% 38% 48% 13%
2	G	100	5% 45% 42% 12%
3	C	10	10% 90%
3	H	10	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	205	
4	I	205	
5	E	242	
5	J	242	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13256 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 HLA class I histocompatibility antigen, A-24 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2222	1382	403	427	10			
1	F	274	Total	C	N	O	S	0	0	0
			2222	1382	403	427	10			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP P05534
F	0	MET	-	EXPRESSION TAG	UNP P05534

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			
2	G	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	EXPRESSION TAG	UNP P61769
G	0	MET	-	EXPRESSION TAG	UNP P61769

- Molecule 3 is a protein called 10-mer peptide from Protein Nef.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	S	0	0	0
			91	64	14	12	1			
3	H	10	Total	C	N	O	S	0	0	0
			91	64	14	12	1			

- Molecule 4 is a protein called T36-5 TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	199	Total	C	N	O	S	0	0	0
			1553	968	257	321	7			
4	I	199	Total	C	N	O	S	0	0	0
			1553	968	257	321	7			

- Molecule 5 is a protein called T36-5 TCR beta chain.

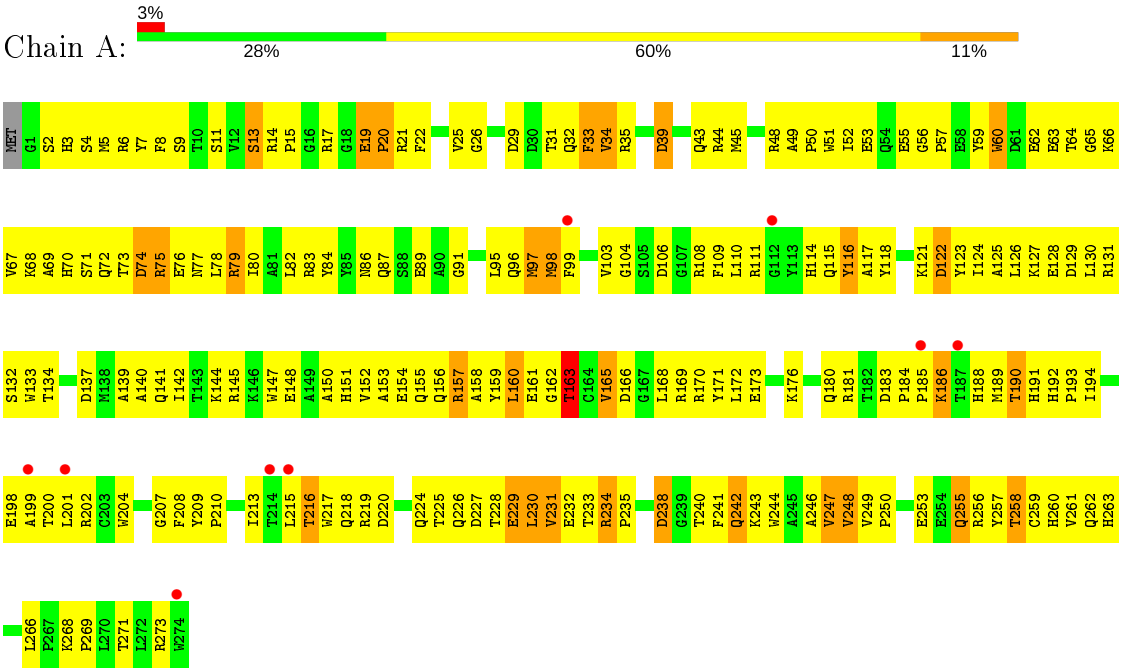
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	241	Total	C	N	O	S	0	0	0
			1933	1217	336	372	8			
5	J	241	Total	C	N	O	S	0	0	0
			1933	1217	336	372	8			

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 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.

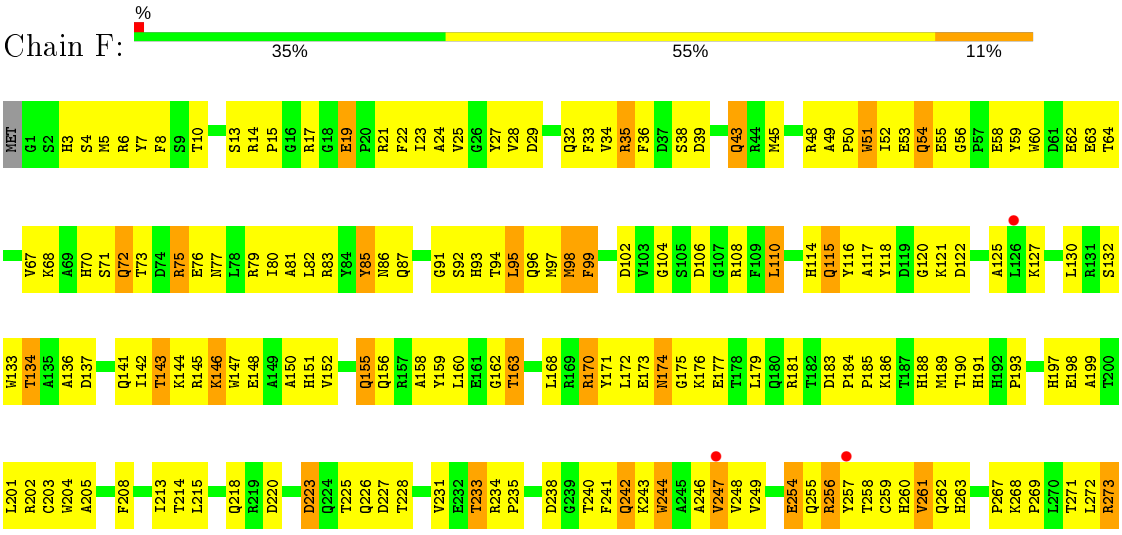
- Molecule 1: HLA class I histocompatibility antigen, A-24 alpha chain

Chain A:



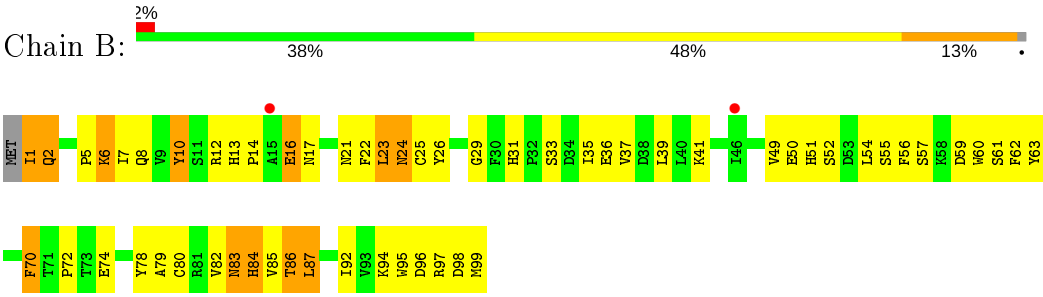
- Molecule 1: HLA class I histocompatibility antigen, A-24 alpha chain

Chain F:



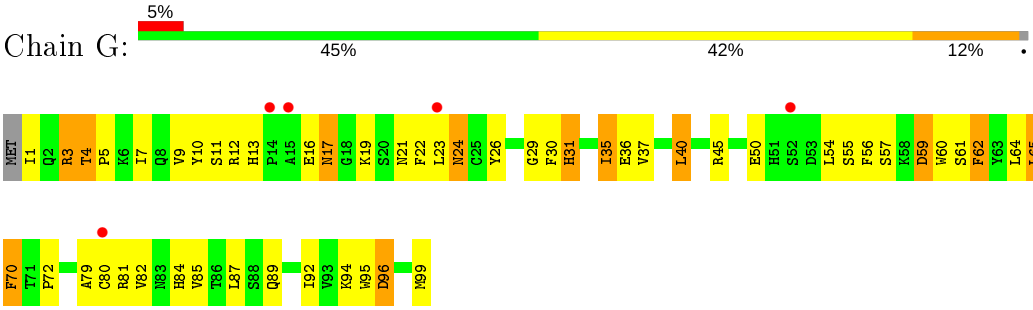
• Molecule 2: Beta-2-microglobulin

Chain B:



• Molecule 2: Beta-2-microglobulin

Chain G:

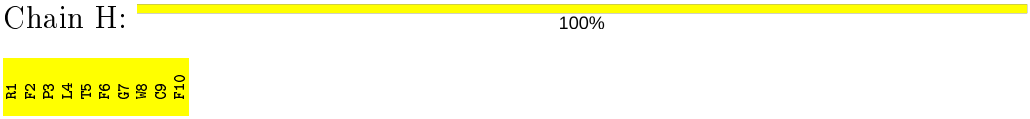


• Molecule 3: 10-mer peptide from Protein Nef

Chain C:

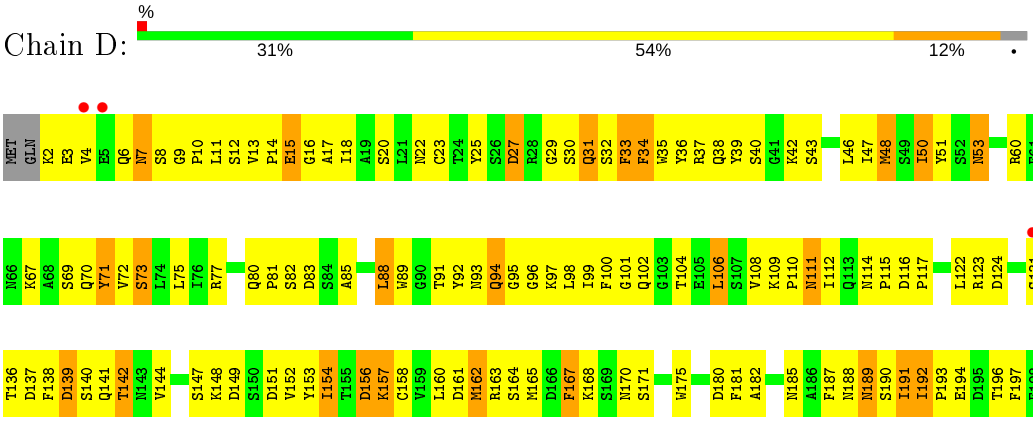
• Molecule 3: 10-mer peptide from Protein Nef

Chain H:



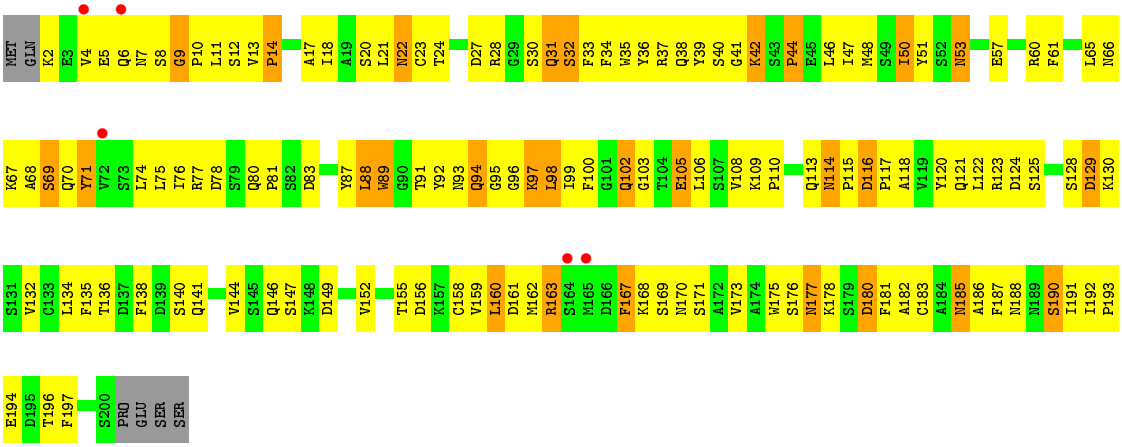
• Molecule 4: T36-5 TCR alpha chain

Chain D:



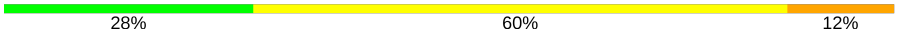
• Molecule 4: T36-5 TCR alpha chain

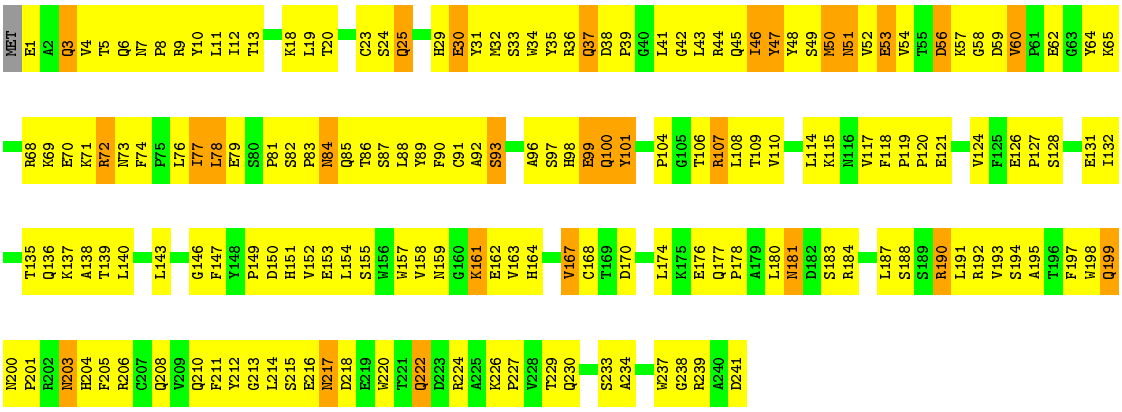
Chain I:



• Molecule 5: T36-5 TCR beta chain

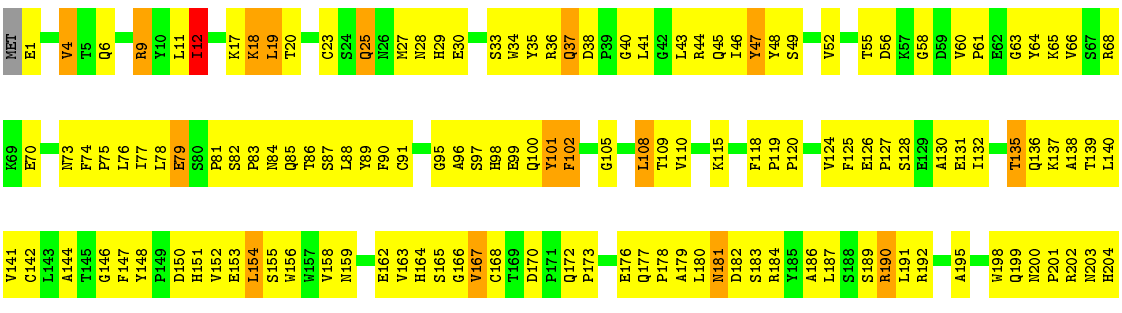
Chain E:





• Molecule 5: T36-5 TCR beta chain

Chain J:




F205
R206
C207
Q208
V209
Q210
F211
Y212
G213
L214
S215
E216
N217
D218
E219
W220
T221
Q222
D223
R224
A225
K226
P227
Q230
I231
V232
S233
A234
E235
A236
W237
G238
R239
A240
D241

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	73.16Å 73.16Å 415.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.35 – 2.70 32.35 – 2.60	Depositor EDS
% Data completeness (in resolution range)	92.2 (32.35-2.70) 95.9 (32.35-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.36 (at 2.61Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.272 , 0.323 0.267 , 0.315	Depositor DCC
R_{free} test set	3517 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	38.6	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.21 , -40.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.069 for -h,-k,l 0.499 for h,-h-k,-l 0.069 for -k,-h,-l	Xtriage
Reported twinning fraction	0.500 for h,k,l 0.500 for h,-h-k,-l	Depositor
Outliers	0 of 73574 reflections	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	13256	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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

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	0.53	0/2282	0.70	0/3092
1	F	0.55	0/2282	0.74	1/3092 (0.0%)
2	B	0.54	0/852	0.71	0/1152
2	G	0.60	0/852	0.71	0/1152
3	C	0.65	0/96	0.83	0/128
3	H	0.65	0/96	0.71	0/128
4	D	0.58	0/1587	0.80	0/2149
4	I	0.59	0/1587	0.80	0/2149
5	E	0.54	0/1986	0.71	0/2705
5	J	0.52	0/1986	0.74	1/2705 (0.0%)
All	All	0.55	0/13606	0.74	2/18452 (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
5	J	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	J	12	ILE	CB-CA-C	-5.57	100.45	111.60
1	F	95	LEU	CA-CB-CG	-5.03	103.74	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	J	101	TYR	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
5	J	47	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	2222	0	2082	315	0
1	F	2222	0	2082	278	0
2	B	829	0	794	92	0
2	G	829	0	794	92	0
3	C	91	0	85	24	0
3	H	91	0	85	31	0
4	D	1553	0	1461	212	0
4	I	1553	0	1461	233	0
5	E	1933	0	1845	267	0
5	J	1933	0	1845	282	0
All	All	13256	0	12534	1612	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 62.

The worst 5 of 1612 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:69:ALA:HA	5:E:51:ASN:HB2	1.20	1.19
5:J:177:GLN:HB3	5:J:180:LEU:HD13	1.19	1.19
5:J:35:TYR:HB3	5:J:43:LEU:HD11	1.24	1.16
1:A:4:SER:HB2	1:A:6:ARG:HH12	1.08	1.14
4:D:114:ASN:HA	1:F:108:ARG:HH22	0.98	1.11

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	272/275 (99%)	259 (95%)	12 (4%)	1 (0%)	34	60
1	F	272/275 (99%)	257 (94%)	14 (5%)	1 (0%)	34	60
2	B	97/100 (97%)	93 (96%)	4 (4%)	0	100	100
2	G	97/100 (97%)	93 (96%)	4 (4%)	0	100	100
3	C	8/10 (80%)	8 (100%)	0	0	100	100
3	H	8/10 (80%)	8 (100%)	0	0	100	100
4	D	197/205 (96%)	167 (85%)	24 (12%)	6 (3%)	4	10
4	I	197/205 (96%)	169 (86%)	22 (11%)	6 (3%)	4	10
5	E	239/242 (99%)	223 (93%)	15 (6%)	1 (0%)	34	60
5	J	239/242 (99%)	220 (92%)	18 (8%)	1 (0%)	34	60
All	All	1626/1664 (98%)	1497 (92%)	113 (7%)	16 (1%)	15	37

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	94	GLN
4	D	8	SER
4	D	40	SER
4	I	7	ASN
4	I	40	SER

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	230/231 (100%)	194 (84%)	36 (16%)	2 6
1	F	230/231 (100%)	190 (83%)	40 (17%)	2 5
2	B	94/95 (99%)	77 (82%)	17 (18%)	1 4
2	G	94/95 (99%)	78 (83%)	16 (17%)	2 5
3	C	9/9 (100%)	9 (100%)	0	100 100
3	H	9/9 (100%)	9 (100%)	0	100 100
4	D	177/183 (97%)	145 (82%)	32 (18%)	1 4
4	I	177/183 (97%)	144 (81%)	33 (19%)	1 4
5	E	214/215 (100%)	176 (82%)	38 (18%)	2 4
5	J	214/215 (100%)	185 (86%)	29 (14%)	3 8
All	All	1448/1466 (99%)	1207 (83%)	241 (17%)	2 5

5 of 241 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	107	ARG
1	F	86	ASN
5	J	56	ASP
5	E	149	PRO
5	E	241	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 70 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	32	GLN
1	F	188	HIS
5	J	181	ASN
1	F	43	GLN
1	F	141	GLN

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 carbohydrates 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

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	274/275 (99%)	0.02	9 (3%) 46 46	36, 69, 105, 120	0
1	F	274/275 (99%)	-0.16	3 (1%) 80 82	33, 67, 86, 91	0
2	B	99/100 (99%)	0.02	2 (2%) 65 67	53, 68, 93, 97	0
2	G	99/100 (99%)	0.25	5 (5%) 28 26	49, 79, 96, 101	0
3	C	10/10 (100%)	-0.46	0 100 100	44, 48, 56, 57	0
3	H	10/10 (100%)	-0.16	0 100 100	34, 40, 59, 64	0
4	D	199/205 (97%)	0.06	3 (1%) 73 76	32, 61, 97, 115	0
4	I	199/205 (97%)	0.09	5 (2%) 57 59	20, 64, 102, 119	0
5	E	241/242 (99%)	-0.27	0 100 100	35, 64, 81, 92	0
5	J	241/242 (99%)	-0.25	0 100 100	29, 62, 84, 94	0
All	All	1646/1664 (98%)	-0.07	27 (1%) 72 74	20, 66, 96, 120	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	I	164	SER	5.2
4	D	131	SER	4.8
4	I	165	MET	4.5
1	A	199	ALA	4.4
1	A	215	LEU	3.9

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 (i)

There are no carbohydrates 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.