



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Feb 26, 2018 – 11:57 AM EST

PDB ID : 5U1F
EMDB ID: : EMD-8427
Title : Initial contact of HIV-1 Env with CD4: Cryo-EM structure of BG505 DS-SOSIP trimer in complex with CD4 and antibody PGT145
Authors : Acharya, P.; Kwong, P.D.; Potter, C.S.; Carragher, B.
Deposited on : 2016-11-28
Resolution : 6.80 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030736

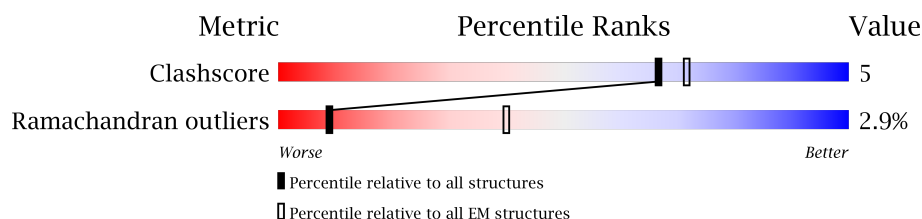
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY










The reported resolution of this entry is 6.80 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	501	 83% 6% • 11%
1	C	501	 84% 5% • 11%
1	D	501	 85% • • 11%
2	B	178	 65% 9% • 26%
2	E	178	 63% 10% • 26%
2	F	178	 66% 7% • 26%
3	H	267	 86% • 12%
4	L	239	 90% 10%
5	M	363	 84% 12% •

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12578 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 BG505 DS-SOSIP gp120.

Mol	Chain	Residues	Atoms				AltConf	Trace
1	C	447	Total	C	N	O	0	0
			2210	1316	447	447		
1	D	447	Total	C	N	O	0	0
			2210	1316	447	447		
1	A	447	Total	C	N	O	0	0
			2210	1316	447	447		

There are 93 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	SER	-	expression tag	UNP Q2N0S6
C	4	ARG	-	expression tag	UNP Q2N0S6
C	5	ALA	-	expression tag	UNP Q2N0S6
C	6	THR	-	expression tag	UNP Q2N0S6
C	7	MET	-	expression tag	UNP Q2N0S6
C	8	PRO	-	expression tag	UNP Q2N0S6
C	9	MET	-	expression tag	UNP Q2N0S6
C	10	GLY	-	expression tag	UNP Q2N0S6
C	11	SER	-	expression tag	UNP Q2N0S6
C	12	LEU	-	expression tag	UNP Q2N0S6
C	13	GLN	-	expression tag	UNP Q2N0S6
C	14	PRO	-	expression tag	UNP Q2N0S6
C	15	LEU	-	expression tag	UNP Q2N0S6
C	16	ALA	-	expression tag	UNP Q2N0S6
C	17	THR	-	expression tag	UNP Q2N0S6
C	18	LEU	-	expression tag	UNP Q2N0S6
C	19	TYR	-	expression tag	UNP Q2N0S6
C	20	LEU	-	expression tag	UNP Q2N0S6
C	21	LEU	-	expression tag	UNP Q2N0S6
C	22	GLY	-	expression tag	UNP Q2N0S6
C	23	MET	-	expression tag	UNP Q2N0S6
C	24	LEU	-	expression tag	UNP Q2N0S6
C	25	VAL	-	expression tag	UNP Q2N0S6
C	26	ALA	-	expression tag	UNP Q2N0S6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	27	SER	-	expression tag	UNP Q2N0S6
C	28	VAL	-	expression tag	UNP Q2N0S6
C	29	LEU	-	expression tag	UNP Q2N0S6
C	201	CYS	ILE	conflict	UNP Q2N0S6
C	332	ASN	THR	conflict	UNP Q2N0S6
C	433	CYS	ALA	conflict	UNP Q2N0S6
C	501	CYS	ALA	conflict	UNP Q2N0S6
D	3	SER	-	expression tag	UNP Q2N0S6
D	4	ARG	-	expression tag	UNP Q2N0S6
D	5	ALA	-	expression tag	UNP Q2N0S6
D	6	THR	-	expression tag	UNP Q2N0S6
D	7	MET	-	expression tag	UNP Q2N0S6
D	8	PRO	-	expression tag	UNP Q2N0S6
D	9	MET	-	expression tag	UNP Q2N0S6
D	10	GLY	-	expression tag	UNP Q2N0S6
D	11	SER	-	expression tag	UNP Q2N0S6
D	12	LEU	-	expression tag	UNP Q2N0S6
D	13	GLN	-	expression tag	UNP Q2N0S6
D	14	PRO	-	expression tag	UNP Q2N0S6
D	15	LEU	-	expression tag	UNP Q2N0S6
D	16	ALA	-	expression tag	UNP Q2N0S6
D	17	THR	-	expression tag	UNP Q2N0S6
D	18	LEU	-	expression tag	UNP Q2N0S6
D	19	TYR	-	expression tag	UNP Q2N0S6
D	20	LEU	-	expression tag	UNP Q2N0S6
D	21	LEU	-	expression tag	UNP Q2N0S6
D	22	GLY	-	expression tag	UNP Q2N0S6
D	23	MET	-	expression tag	UNP Q2N0S6
D	24	LEU	-	expression tag	UNP Q2N0S6
D	25	VAL	-	expression tag	UNP Q2N0S6
D	26	ALA	-	expression tag	UNP Q2N0S6
D	27	SER	-	expression tag	UNP Q2N0S6
D	28	VAL	-	expression tag	UNP Q2N0S6
D	29	LEU	-	expression tag	UNP Q2N0S6
D	201	CYS	ILE	conflict	UNP Q2N0S6
D	332	ASN	THR	conflict	UNP Q2N0S6
D	433	CYS	ALA	conflict	UNP Q2N0S6
D	501	CYS	ALA	conflict	UNP Q2N0S6
A	3	SER	-	expression tag	UNP Q2N0S6
A	4	ARG	-	expression tag	UNP Q2N0S6
A	5	ALA	-	expression tag	UNP Q2N0S6
A	6	THR	-	expression tag	UNP Q2N0S6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	MET	-	expression tag	UNP Q2N0S6
A	8	PRO	-	expression tag	UNP Q2N0S6
A	9	MET	-	expression tag	UNP Q2N0S6
A	10	GLY	-	expression tag	UNP Q2N0S6
A	11	SER	-	expression tag	UNP Q2N0S6
A	12	LEU	-	expression tag	UNP Q2N0S6
A	13	GLN	-	expression tag	UNP Q2N0S6
A	14	PRO	-	expression tag	UNP Q2N0S6
A	15	LEU	-	expression tag	UNP Q2N0S6
A	16	ALA	-	expression tag	UNP Q2N0S6
A	17	THR	-	expression tag	UNP Q2N0S6
A	18	LEU	-	expression tag	UNP Q2N0S6
A	19	TYR	-	expression tag	UNP Q2N0S6
A	20	LEU	-	expression tag	UNP Q2N0S6
A	21	LEU	-	expression tag	UNP Q2N0S6
A	22	GLY	-	expression tag	UNP Q2N0S6
A	23	MET	-	expression tag	UNP Q2N0S6
A	24	LEU	-	expression tag	UNP Q2N0S6
A	25	VAL	-	expression tag	UNP Q2N0S6
A	26	ALA	-	expression tag	UNP Q2N0S6
A	27	SER	-	expression tag	UNP Q2N0S6
A	28	VAL	-	expression tag	UNP Q2N0S6
A	29	LEU	-	expression tag	UNP Q2N0S6
A	201	CYS	ILE	conflict	UNP Q2N0S6
A	332	ASN	THR	conflict	UNP Q2N0S6
A	433	CYS	ALA	conflict	UNP Q2N0S6
A	501	CYS	ALA	conflict	UNP Q2N0S6

- Molecule 2 is a protein called BG505 SOSIP gp41.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	B	132	Total	C	N	O	0	0
			650	386	132	132		
2	E	132	Total	C	N	O	0	0
			650	386	132	132		
2	F	132	Total	C	N	O	0	0
			650	386	132	132		

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	504	VAL	-	expression tag	UNP Q2N0S6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	505	GLY	-	expression tag	UNP Q2N0S6
B	506	ARG	-	expression tag	UNP Q2N0S6
B	507	ARG	-	expression tag	UNP Q2N0S6
B	508	ARG	-	expression tag	UNP Q2N0S6
B	509	ARG	-	expression tag	UNP Q2N0S6
B	510	ARG	-	expression tag	UNP Q2N0S6
B	559	PRO	ILE	conflict	UNP Q2N0S6
B	605	CYS	THR	conflict	UNP Q2N0S6
B	665	GLY	-	expression tag	UNP Q2N0S6
B	666	SER	-	expression tag	UNP Q2N0S6
B	667	ALA	-	expression tag	UNP Q2N0S6
B	668	PRO	-	expression tag	UNP Q2N0S6
B	669	THR	-	expression tag	UNP Q2N0S6
B	670	LYS	-	expression tag	UNP Q2N0S6
B	671	ALA	-	expression tag	UNP Q2N0S6
B	672	LYS	-	expression tag	UNP Q2N0S6
B	673	ARG	-	expression tag	UNP Q2N0S6
B	674	ARG	-	expression tag	UNP Q2N0S6
B	675	VAL	-	expression tag	UNP Q2N0S6
B	676	VAL	-	expression tag	UNP Q2N0S6
B	677	GLN	-	expression tag	UNP Q2N0S6
B	678	ARG	-	expression tag	UNP Q2N0S6
B	679	GLU	-	expression tag	UNP Q2N0S6
B	680	LYS	-	expression tag	UNP Q2N0S6
B	681	ARG	-	expression tag	UNP Q2N0S6
E	504	VAL	-	expression tag	UNP Q2N0S6
E	505	GLY	-	expression tag	UNP Q2N0S6
E	506	ARG	-	expression tag	UNP Q2N0S6
E	507	ARG	-	expression tag	UNP Q2N0S6
E	508	ARG	-	expression tag	UNP Q2N0S6
E	509	ARG	-	expression tag	UNP Q2N0S6
E	510	ARG	-	expression tag	UNP Q2N0S6
E	559	PRO	ILE	conflict	UNP Q2N0S6
E	605	CYS	THR	conflict	UNP Q2N0S6
E	665	GLY	-	expression tag	UNP Q2N0S6
E	666	SER	-	expression tag	UNP Q2N0S6
E	667	ALA	-	expression tag	UNP Q2N0S6
E	668	PRO	-	expression tag	UNP Q2N0S6
E	669	THR	-	expression tag	UNP Q2N0S6
E	670	LYS	-	expression tag	UNP Q2N0S6
E	671	ALA	-	expression tag	UNP Q2N0S6
E	672	LYS	-	expression tag	UNP Q2N0S6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	673	ARG	-	expression tag	UNP Q2N0S6
E	674	ARG	-	expression tag	UNP Q2N0S6
E	675	VAL	-	expression tag	UNP Q2N0S6
E	676	VAL	-	expression tag	UNP Q2N0S6
E	677	GLN	-	expression tag	UNP Q2N0S6
E	678	ARG	-	expression tag	UNP Q2N0S6
E	679	GLU	-	expression tag	UNP Q2N0S6
E	680	LYS	-	expression tag	UNP Q2N0S6
E	681	ARG	-	expression tag	UNP Q2N0S6
F	504	VAL	-	expression tag	UNP Q2N0S6
F	505	GLY	-	expression tag	UNP Q2N0S6
F	506	ARG	-	expression tag	UNP Q2N0S6
F	507	ARG	-	expression tag	UNP Q2N0S6
F	508	ARG	-	expression tag	UNP Q2N0S6
F	509	ARG	-	expression tag	UNP Q2N0S6
F	510	ARG	-	expression tag	UNP Q2N0S6
F	559	PRO	ILE	conflict	UNP Q2N0S6
F	605	CYS	THR	conflict	UNP Q2N0S6
F	665	GLY	-	expression tag	UNP Q2N0S6
F	666	SER	-	expression tag	UNP Q2N0S6
F	667	ALA	-	expression tag	UNP Q2N0S6
F	668	PRO	-	expression tag	UNP Q2N0S6
F	669	THR	-	expression tag	UNP Q2N0S6
F	670	LYS	-	expression tag	UNP Q2N0S6
F	671	ALA	-	expression tag	UNP Q2N0S6
F	672	LYS	-	expression tag	UNP Q2N0S6
F	673	ARG	-	expression tag	UNP Q2N0S6
F	674	ARG	-	expression tag	UNP Q2N0S6
F	675	VAL	-	expression tag	UNP Q2N0S6
F	676	VAL	-	expression tag	UNP Q2N0S6
F	677	GLN	-	expression tag	UNP Q2N0S6
F	678	ARG	-	expression tag	UNP Q2N0S6
F	679	GLU	-	expression tag	UNP Q2N0S6
F	680	LYS	-	expression tag	UNP Q2N0S6
F	681	ARG	-	expression tag	UNP Q2N0S6

- Molecule 3 is a protein called PGT145 heavy chain.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	H	235	Total	C	N	O	0	0
			1151	681	235	235		

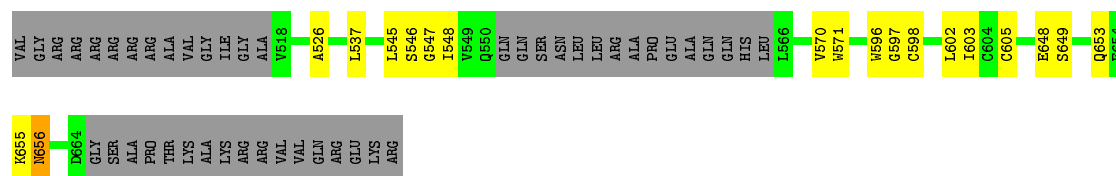
- Molecule 4 is a protein called PGT145 light chain.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	L	214	Total	C	N	O	0	0
			1055	627	214	214		

- Molecule 5 is a protein called T-cell surface glycoprotein CD4.

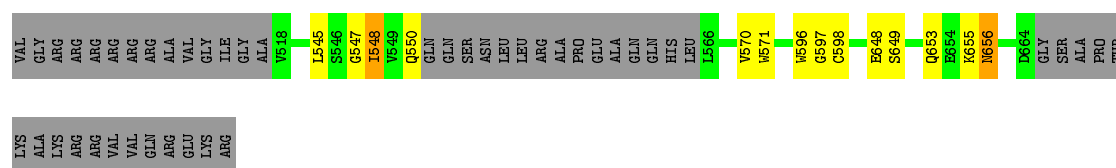
Mol	Chain	Residues	Atoms				AltConf	Trace
5	M	363	Total	C	N	O	0	0
			1792	1066	363	363		

Chain E:  63% 10% 26%




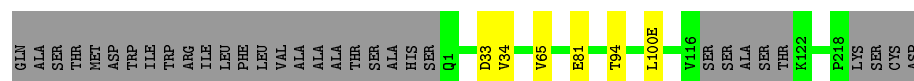
- Molecule 2: BG505 SOSIP gp41

Chain F:  66% 7% 26%




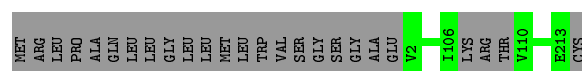
- Molecule 3: PGT145 heavy chain

Chain H:  86% 12%



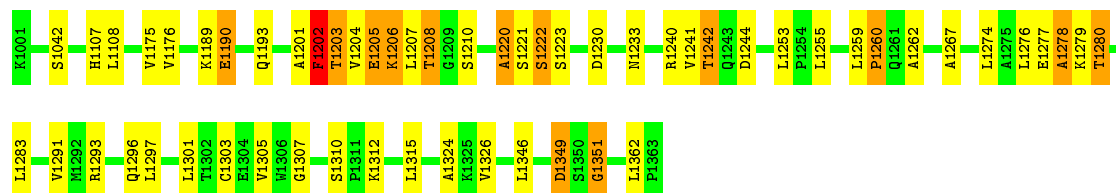
- Molecule 4: PGT145 light chain

Chain L:  90% 10%



- Molecule 5: T-cell surface glycoprotein CD4

Chain M:  84% 12%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	55000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	8.6	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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 >2	RMSZ	# Z >2
1	A	0.38	1/2206 (0.0%)	0.84	16/3067 (0.5%)
1	C	0.38	2/2206 (0.1%)	0.83	15/3067 (0.5%)
1	D	0.38	1/2206 (0.0%)	0.84	15/3067 (0.5%)
2	B	0.27	0/648	0.84	3/898 (0.3%)
2	E	0.27	0/648	0.84	3/898 (0.3%)
2	F	0.27	0/648	0.84	3/898 (0.3%)
3	H	0.39	0/1137	0.60	1/1571 (0.1%)
4	L	0.39	0/1053	0.58	0/1462
5	M	0.52	1/1791 (0.1%)	0.96	6/2492 (0.2%)
All	All	0.39	5/12543 (0.0%)	0.82	62/17420 (0.4%)

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	2
1	C	0	1
1	D	0	1
All	All	0	4

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	84	ILE	C-N	6.01	1.47	1.34
1	C	84	ILE	C-N	5.96	1.47	1.34
1	D	84	ILE	C-N	5.95	1.47	1.34
5	M	1202	PHE	C-O	-5.23	1.13	1.23
1	C	393	SER	C-N	5.02	1.45	1.34

The worst 5 of 62 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	B	648	GLU	N-CA-C	14.21	149.38	111.00
2	E	648	GLU	N-CA-C	14.16	149.24	111.00
2	F	648	GLU	N-CA-C	14.12	149.13	111.00
1	C	84	ILE	O-C-N	12.81	143.20	122.70
1	D	86	LEU	O-C-N	-12.72	102.35	122.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	65	LYS	Mainchain
1	A	86	LEU	Mainchain
1	C	86	LEU	Mainchain
1	D	86	LEU	Mainchain

5.2 Too-close contacts ⓘ

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	2210	0	953	18	0
1	C	2210	0	953	18	0
1	D	2210	0	953	6	0
2	B	650	0	295	14	0
2	E	650	0	295	16	0
2	F	650	0	295	5	0
3	H	1151	0	525	3	0
4	L	1055	0	468	0	0
5	M	1792	0	781	28	0
All	All	12578	0	5518	84	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 5.

The worst 5 of 84 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:35:TRP:HA	2:B:610:TRP:CB	1.87	1.04
1:C:43:PRO:CB	2:E:526:ALA:HA	1.89	1.02
1:A:43:PRO:CB	2:B:526:ALA:CB	2.42	0.98
1:A:43:PRO:CB	2:B:526:ALA:HA	1.93	0.97
1:A:503:ARG:CB	2:B:605:CYS:CB	2.45	0.94

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 PDB entries followed by that with respect to all EM entries.

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	439/501 (88%)	391 (89%)	33 (8%)	15 (3%)	4	35
1	C	439/501 (88%)	398 (91%)	31 (7%)	10 (2%)	7	43
1	D	439/501 (88%)	388 (88%)	40 (9%)	11 (2%)	6	41
2	B	128/178 (72%)	103 (80%)	21 (16%)	4 (3%)	5	37
2	E	128/178 (72%)	103 (80%)	21 (16%)	4 (3%)	5	37
2	F	128/178 (72%)	103 (80%)	21 (16%)	4 (3%)	5	37
3	H	229/267 (86%)	219 (96%)	10 (4%)	0	100	100
4	L	210/239 (88%)	207 (99%)	3 (1%)	0	100	100
5	M	361/363 (99%)	298 (82%)	38 (10%)	25 (7%)	1	20
All	All	2501/2906 (86%)	2210 (88%)	218 (9%)	73 (3%)	9	38

5 of 73 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	396	ILE
1	D	396	ILE
1	A	55	ALA
1	A	396	ILE
5	M	1190	GLU

5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	TYS	H	100(F)	3	4,4,17	0.77	0	1,4,24	0.08	0
3	TYS	H	100(I)	3	4,4,17	0.77	0	1,4,24	0.16	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	TYS	H	100(F)	3	-	0/0/2/13	0/0/0/1
3	TYS	H	100(I)	3	-	0/0/2/13	0/0/0/1

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.

5.5 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.