wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 07:37 PM GMT

PDB ID	:	4P5H
Title	:	Structure of Clostridium perfringens Enterotoxin with a peptide derived from a modified version of ECL-2 of Claudin 2
Authors	:	Naylor, C.E.; Yelland, T.S.; Basak, A.K.
Deposited on	:	2014-03-17
Resolution	:	3.38 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
http://wwpdb.org/validation/2016/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:

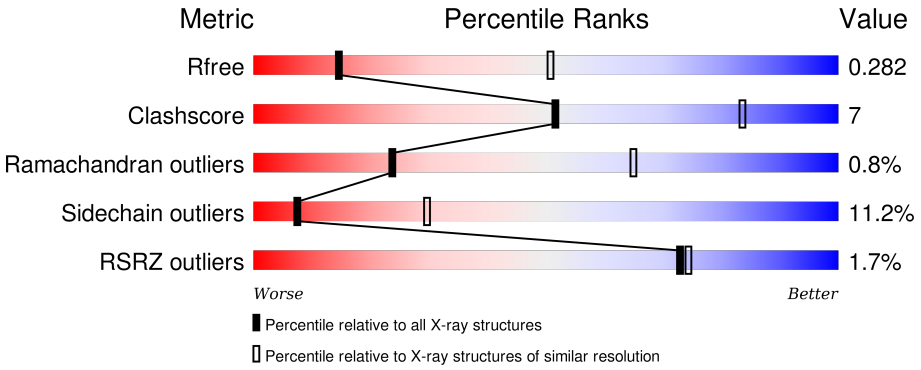
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.38 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}	91344	1084 (3.46-3.30)
Clashscore	102246	1158 (3.46-3.30)
Ramachandran outliers	100387	1139 (3.46-3.30)
Sidechain outliers	100360	1138 (3.46-3.30)
RSRZ outliers	91569	1089 (3.46-3.30)

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. 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	1	20	25% 5% 15% 55%
1	2	20	25% 15% 5% 55%
1	3	20	20% 15% 10% 55%
1	4	20	30% 5% 10% 55%
1	P	20	30% 15% 55%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	Q	20	25% 10% 10% 55%
1	R	20	20% 15% 5% 5% 55%
1	S	20	15% 20% 5% 5% 55%
1	T	20	25% 5% 10% 5% 55%
1	U	20	25% 5% 10% 5% 55%
1	V	20	25% 5% 10% 5% 55%
1	W	20	20% 20% 5% 55%
1	X	20	25% 10% 5% 5% 55%
1	Y	20	20% 15% 5% 5% 55%
1	Z	20	30% 10% 5% 55%
2	A	286	2% 91% 9%
2	B	286	3% 90% 9%
2	C	286	2% 90% 9%
2	D	286	2% 92% 8%
2	E	286	3% 91% 8%
2	F	286	2% 90% 9%
2	G	286	2% 67% 28% 6%
2	H	286	% 71% 26%
2	I	286	% 73% 25%
2	J	286	3% 67% 30%
2	K	286	% 66% 31%
2	L	286	2% 70% 27%
2	M	286	% 63% 33%
2	N	286	67% 28%
2	O	286	% 76% 22%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 34251 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 Claudin-2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	1	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	2	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	3	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	4	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	P	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	Q	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	R	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	S	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	T	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	U	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	V	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	W	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	X	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	Y	9	Total	C	N	O	0	0	0
			60	38	10	12			
1	Z	9	Total	C	N	O	0	0	0
			60	38	10	12			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	149	ASN	SER	engineered mutation	UNP O88552
1	155	ALA	SER	engineered mutation	UNP O88552
2	149	ASN	SER	engineered mutation	UNP O88552
2	155	ALA	SER	engineered mutation	UNP O88552
3	149	ASN	SER	engineered mutation	UNP O88552
3	155	ALA	SER	engineered mutation	UNP O88552
4	149	ASN	SER	engineered mutation	UNP O88552
4	155	ALA	SER	engineered mutation	UNP O88552
P	149	ASN	SER	engineered mutation	UNP O88552
P	155	ALA	SER	engineered mutation	UNP O88552
Q	149	ASN	SER	engineered mutation	UNP O88552
Q	155	ALA	SER	engineered mutation	UNP O88552
R	149	ASN	SER	engineered mutation	UNP O88552
R	155	ALA	SER	engineered mutation	UNP O88552
S	149	ASN	SER	engineered mutation	UNP O88552
S	155	ALA	SER	engineered mutation	UNP O88552
T	149	ASN	SER	engineered mutation	UNP O88552
T	155	ALA	SER	engineered mutation	UNP O88552
U	149	ASN	SER	engineered mutation	UNP O88552
U	155	ALA	SER	engineered mutation	UNP O88552
V	149	ASN	SER	engineered mutation	UNP O88552
V	155	ALA	SER	engineered mutation	UNP O88552
W	149	ASN	SER	engineered mutation	UNP O88552
W	155	ALA	SER	engineered mutation	UNP O88552
X	149	ASN	SER	engineered mutation	UNP O88552
X	155	ALA	SER	engineered mutation	UNP O88552
Y	149	ASN	SER	engineered mutation	UNP O88552
Y	155	ALA	SER	engineered mutation	UNP O88552
Z	149	ASN	SER	engineered mutation	UNP O88552
Z	155	ALA	SER	engineered mutation	UNP O88552

- Molecule 2 is a protein called Heat-labile enterotoxin B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	B	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	C	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	D	286	Total	C	N	O	S	0	4	0
			2229	1413	360	452	4			
2	E	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	G	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	H	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	I	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	J	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	K	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	L	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	M	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	N	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			
2	O	286	Total	C	N	O	S	0	4	0
			2223	1410	357	452	4			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	GLY	-	expression tag	UNP P01558
A	35	ALA	-	expression tag	UNP P01558
A	36	MET	-	expression tag	UNP P01558
A	37	GLY	-	expression tag	UNP P01558
B	34	GLY	-	expression tag	UNP P01558
B	35	ALA	-	expression tag	UNP P01558
B	36	MET	-	expression tag	UNP P01558
B	37	GLY	-	expression tag	UNP P01558
C	34	GLY	-	expression tag	UNP P01558
C	35	ALA	-	expression tag	UNP P01558
C	36	MET	-	expression tag	UNP P01558
C	37	GLY	-	expression tag	UNP P01558
D	34	GLY	-	expression tag	UNP P01558
D	35	ALA	-	expression tag	UNP P01558
D	36	MET	-	expression tag	UNP P01558
D	37	GLY	-	expression tag	UNP P01558
E	34	GLY	-	expression tag	UNP P01558
E	35	ALA	-	expression tag	UNP P01558
E	36	MET	-	expression tag	UNP P01558

Continued on next page...

Continued from previous page...

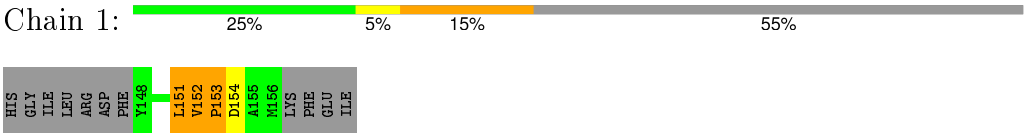
Chain	Residue	Modelled	Actual	Comment	Reference
E	37	GLY	-	expression tag	UNP P01558
F	34	GLY	-	expression tag	UNP P01558
F	35	ALA	-	expression tag	UNP P01558
F	36	MET	-	expression tag	UNP P01558
F	37	GLY	-	expression tag	UNP P01558
G	34	GLY	-	expression tag	UNP P01558
G	35	ALA	-	expression tag	UNP P01558
G	36	MET	-	expression tag	UNP P01558
G	37	GLY	-	expression tag	UNP P01558
H	34	GLY	-	expression tag	UNP P01558
H	35	ALA	-	expression tag	UNP P01558
H	36	MET	-	expression tag	UNP P01558
H	37	GLY	-	expression tag	UNP P01558
I	34	GLY	-	expression tag	UNP P01558
I	35	ALA	-	expression tag	UNP P01558
I	36	MET	-	expression tag	UNP P01558
I	37	GLY	-	expression tag	UNP P01558
J	34	GLY	-	expression tag	UNP P01558
J	35	ALA	-	expression tag	UNP P01558
J	36	MET	-	expression tag	UNP P01558
J	37	GLY	-	expression tag	UNP P01558
K	34	GLY	-	expression tag	UNP P01558
K	35	ALA	-	expression tag	UNP P01558
K	36	MET	-	expression tag	UNP P01558
K	37	GLY	-	expression tag	UNP P01558
L	34	GLY	-	expression tag	UNP P01558
L	35	ALA	-	expression tag	UNP P01558
L	36	MET	-	expression tag	UNP P01558
L	37	GLY	-	expression tag	UNP P01558
M	34	GLY	-	expression tag	UNP P01558
M	35	ALA	-	expression tag	UNP P01558
M	36	MET	-	expression tag	UNP P01558
M	37	GLY	-	expression tag	UNP P01558
N	34	GLY	-	expression tag	UNP P01558
N	35	ALA	-	expression tag	UNP P01558
N	36	MET	-	expression tag	UNP P01558
N	37	GLY	-	expression tag	UNP P01558
O	34	GLY	-	expression tag	UNP P01558
O	35	ALA	-	expression tag	UNP P01558
O	36	MET	-	expression tag	UNP P01558
O	37	GLY	-	expression tag	UNP P01558

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 errors 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.

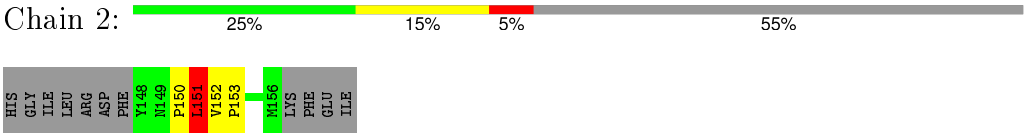
- Molecule 1: Claudin-2

Chain 1:



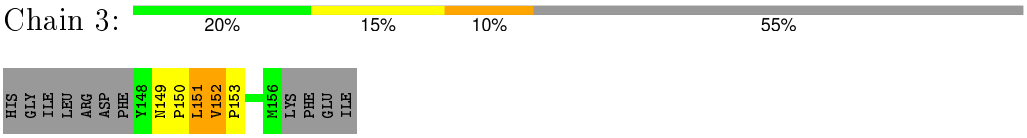
- Molecule 1: Claudin-2

Chain 2:



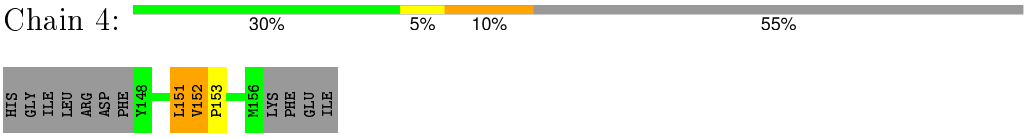
- Molecule 1: Claudin-2

Chain 3:



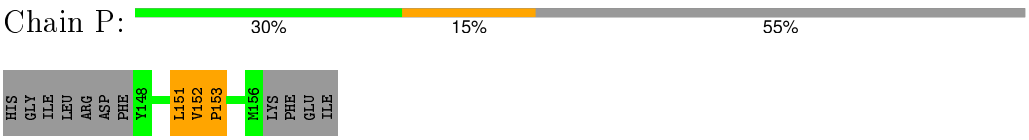
- Molecule 1: Claudin-2

Chain 4:



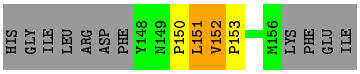
- Molecule 1: Claudin-2

Chain P:



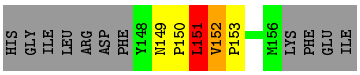
- Molecule 1: Claudin-2

Chain Q:



• Molecule 1: Claudin-2

Chain R:

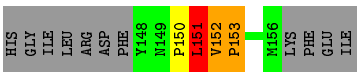


• Molecule 1: Claudin-2

Chain S:

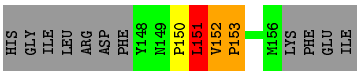
• Molecule 1: Claudin-2

Chain T:



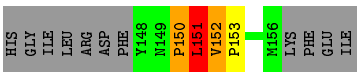
• Molecule 1: Claudin-2

Chain U:



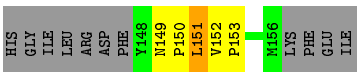
• Molecule 1: Claudin-2

Chain V:



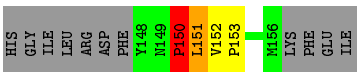
• Molecule 1: Claudin-2

Chain W:



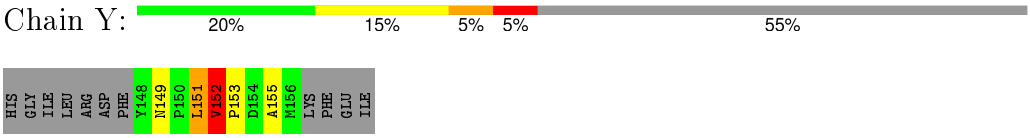
• Molecule 1: Claudin-2

Chain X:



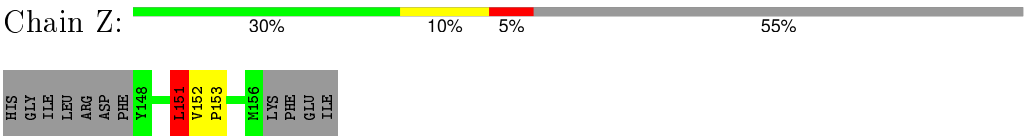
• Molecule 1: Claudin-2

Chain Y:



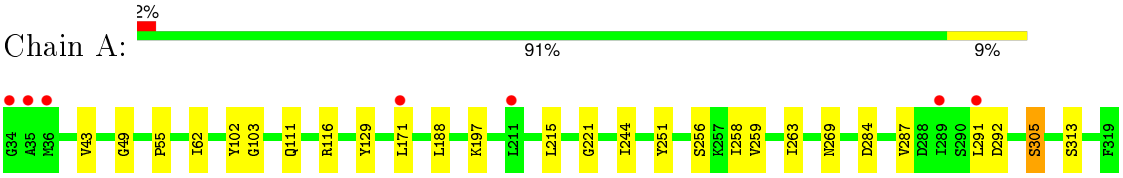
• Molecule 1: Claudin-2

Chain Z:



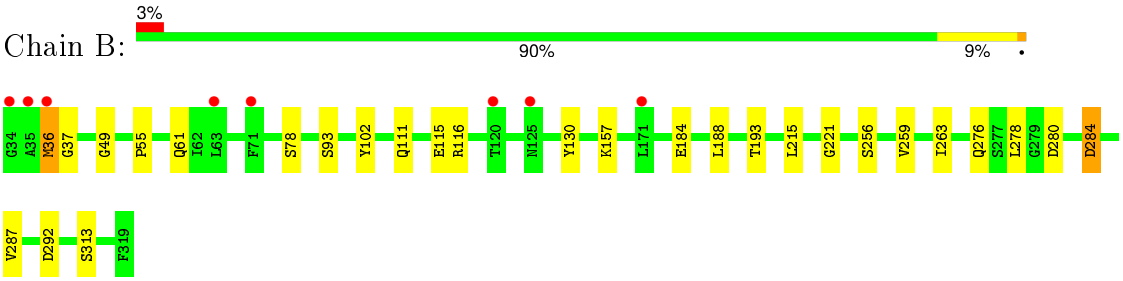
• Molecule 2: Heat-labile enterotoxin B chain

Chain A:



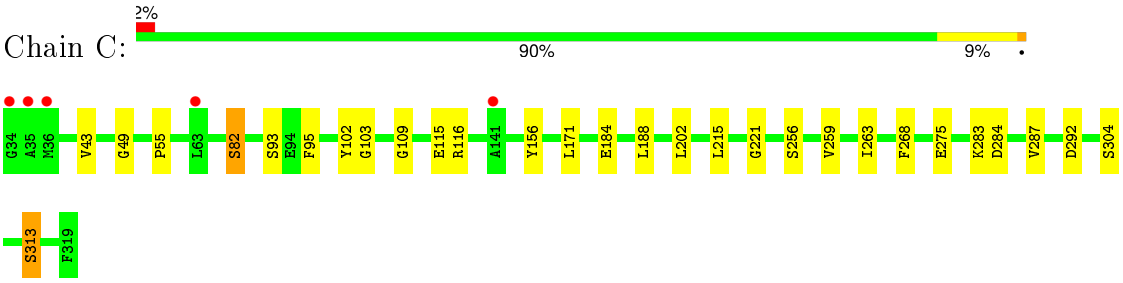
• Molecule 2: Heat-labile enterotoxin B chain

Chain B:



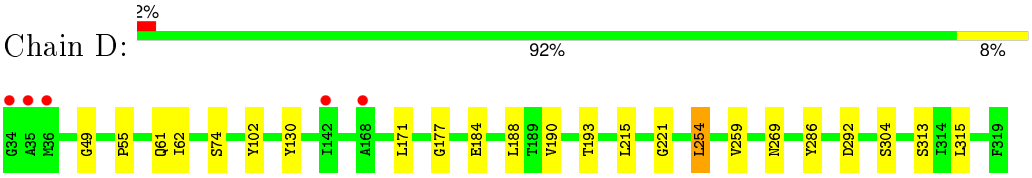
• Molecule 2: Heat-labile enterotoxin B chain

Chain C:



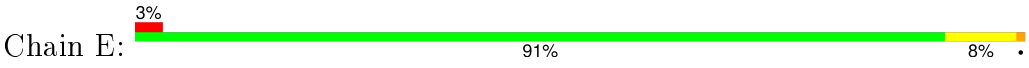
• Molecule 2: Heat-labile enterotoxin B chain

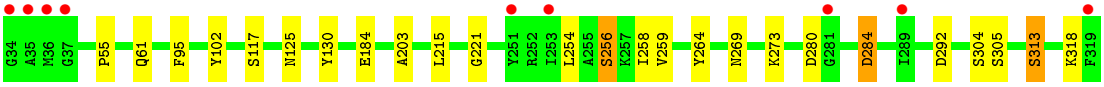
Chain D:



• Molecule 2: Heat-labile enterotoxin B chain

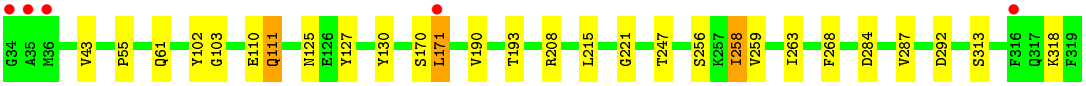
Chain E:





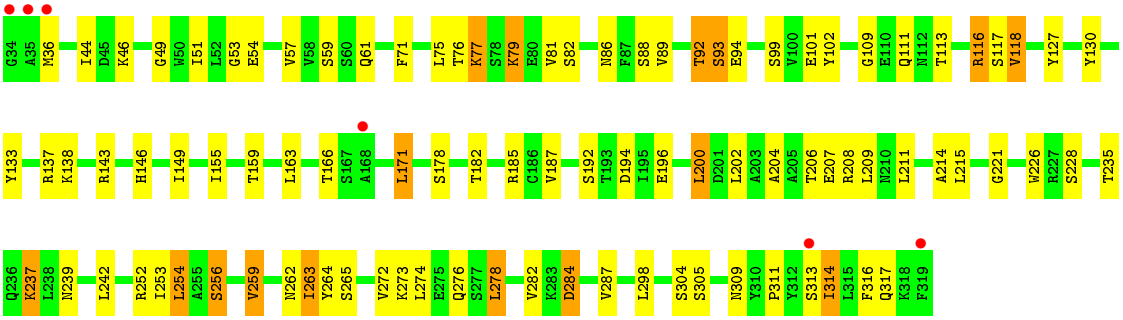
• Molecule 2: Heat-labile enterotoxin B chain

Chain F:



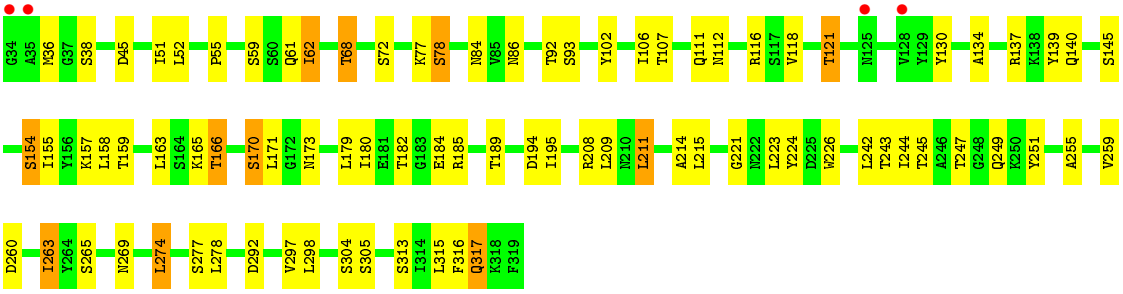
• Molecule 2: Heat-labile enterotoxin B chain

Chain G:



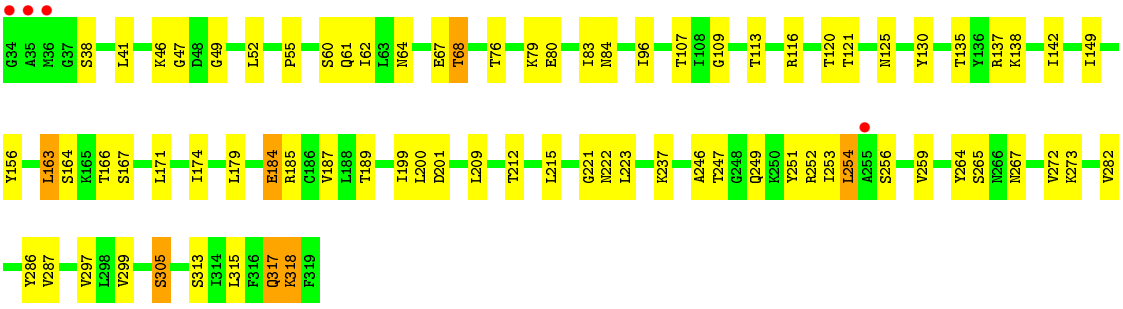
• Molecule 2: Heat-labile enterotoxin B chain

Chain H:



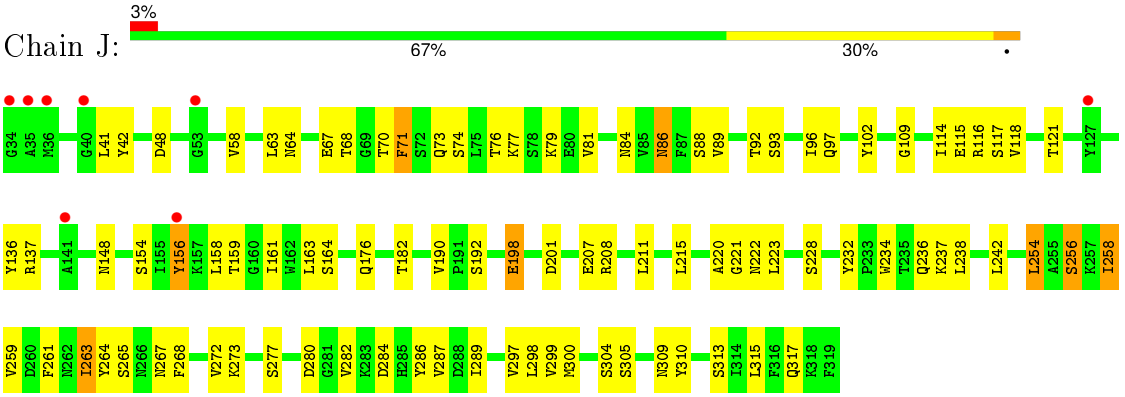
• Molecule 2: Heat-labile enterotoxin B chain

Chain I:



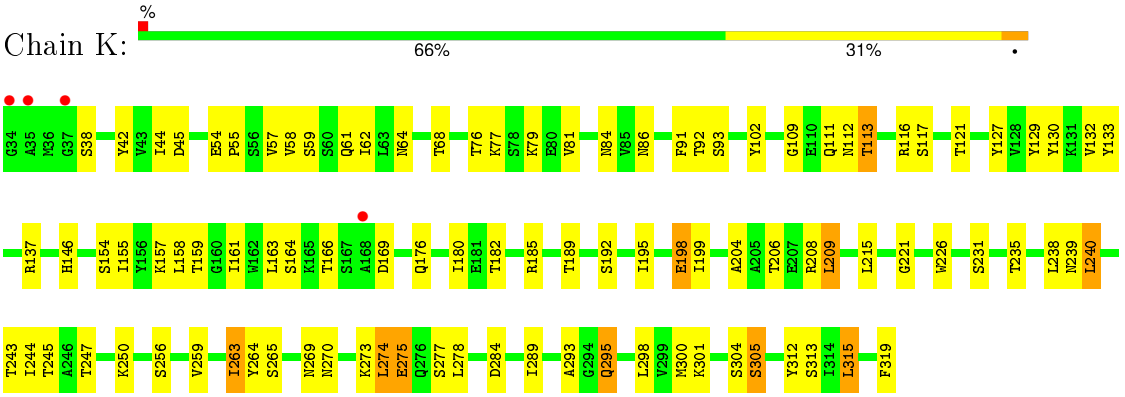
• Molecule 2: Heat-labile enterotoxin B chain

Chain J:



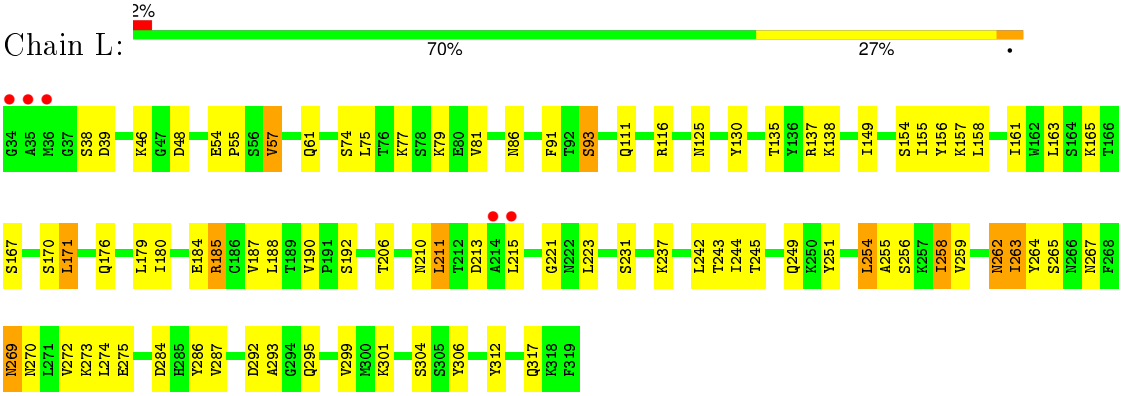
• Molecule 2: Heat-labile enterotoxin B chain

Chain K:



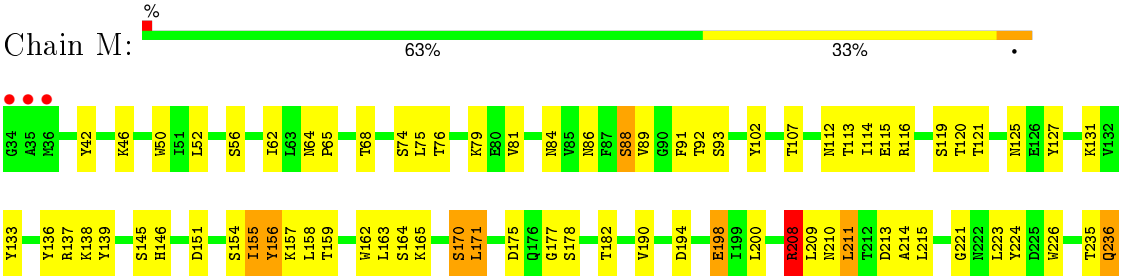
• Molecule 2: Heat-labile enterotoxin B chain

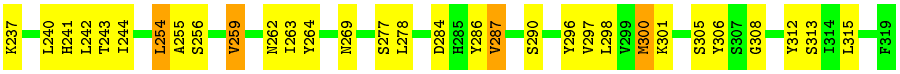
Chain L:



• Molecule 2: Heat-labile enterotoxin B chain

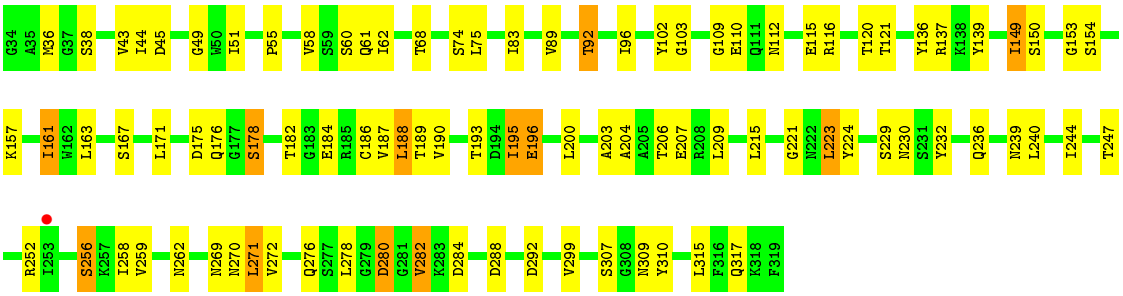
Chain M:





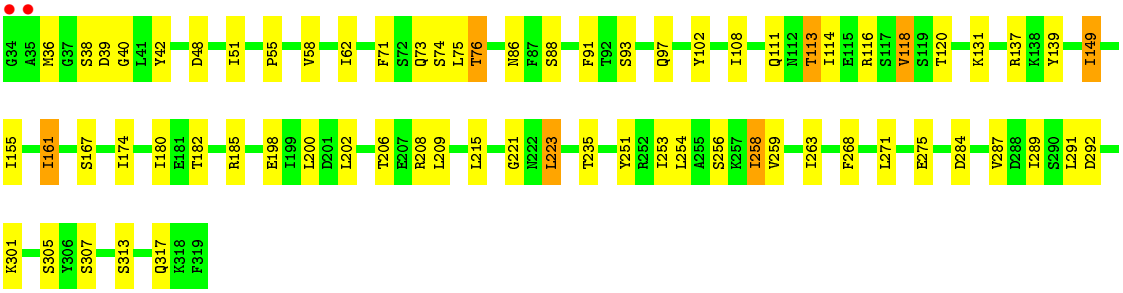
• Molecule 2: Heat-labile enterotoxin B chain

Chain N:



• Molecule 2: Heat-labile enterotoxin B chain

Chain O:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	369.60Å 100.26Å 265.36Å 90.00° 119.74° 90.00°	Depositor
Resolution (Å)	48.98 – 3.38 48.98 – 3.37	Depositor EDS
% Data completeness (in resolution range)	80.8 (48.98-3.38) 81.2 (48.98-3.37)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 3.40Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.204 , 0.240 0.244 , 0.282	Depositor DCC
R_{free} test set	4848 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	70.0	Xtriage
Anisotropy	1.076	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 101.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.21$	Xtriage
Outliers	0 of 96759 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	34251	wwPDB-VP
Average B, all atoms (Å^2)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.55% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	1	0.98	0/61	1.53	2/85 (2.4%)
1	2	1.13	0/61	1.61	2/85 (2.4%)
1	3	1.02	0/61	1.54	2/85 (2.4%)
1	4	1.11	0/61	1.59	2/85 (2.4%)
1	P	0.91	0/61	1.48	2/85 (2.4%)
1	Q	1.21	0/61	1.76	2/85 (2.4%)
1	R	1.14	0/61	1.56	3/85 (3.5%)
1	S	1.29	0/61	1.63	2/85 (2.4%)
1	T	1.05	0/61	1.41	2/85 (2.4%)
1	U	1.08	0/61	1.60	2/85 (2.4%)
1	V	0.97	0/61	1.65	2/85 (2.4%)
1	W	1.17	0/61	1.45	0/85
1	X	1.12	0/61	2.06	4/85 (4.7%)
1	Y	1.16	0/61	1.51	2/85 (2.4%)
1	Z	1.06	0/61	1.50	1/85 (1.2%)
2	A	0.46	0/2276	0.71	0/3094
2	B	0.47	0/2276	0.71	0/3094
2	C	0.48	0/2276	0.75	1/3094 (0.0%)
2	D	0.49	0/2282	0.72	0/3101
2	E	0.44	0/2276	0.71	0/3094
2	F	0.46	0/2276	0.72	0/3094
2	G	0.52	0/2276	0.85	1/3094 (0.0%)
2	H	0.54	0/2276	0.86	0/3094
2	I	0.51	0/2276	0.82	1/3094 (0.0%)
2	J	0.51	0/2276	0.81	0/3094
2	K	0.53	0/2276	0.84	0/3094
2	L	0.52	0/2276	0.84	0/3094
2	M	0.53	0/2276	0.85	1/3094 (0.0%)
2	N	0.54	0/2276	0.84	1/3094 (0.0%)
2	O	0.55	0/2276	0.84	3/3094 (0.1%)
All	All	0.53	0/35061	0.83	38/47692 (0.1%)

There are no bond length outliers.

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	X	150	PRO	C-N-CA	7.56	140.60	121.70
1	X	151	LEU	N-CA-C	7.19	130.40	111.00
2	N	190	VAL	N-CA-CB	6.89	126.66	111.50
1	Q	151	LEU	N-CA-C	6.57	128.74	111.00
1	X	150	PRO	CB-CA-C	6.52	128.31	112.00

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	1	60	0	52	5	0
1	2	60	0	52	6	0
1	3	60	0	52	1	0
1	4	60	0	52	3	0
1	P	60	0	52	4	0
1	Q	60	0	52	3	0
1	R	60	0	52	4	0
1	S	60	0	52	5	0
1	T	60	0	52	7	0
1	U	60	0	52	5	0
1	V	60	0	52	4	0
1	W	60	0	52	5	0
1	X	60	0	52	2	0
1	Y	60	0	52	2	0
1	Z	60	0	52	4	0
2	A	2223	0	2147	21	0
2	B	2223	0	2147	19	0
2	C	2223	0	2147	21	0
2	D	2229	0	2158	15	0
2	E	2223	0	2147	15	0
2	F	2223	0	2147	16	0
2	G	2223	0	2147	46	0
2	H	2223	0	2147	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	I	2223	0	2147	34	0
2	J	2223	0	2147	46	0
2	K	2223	0	2147	53	0
2	L	2223	0	2147	45	0
2	M	2223	0	2147	55	0
2	N	2223	0	2147	50	0
2	O	2223	0	2147	34	0
All	All	34251	0	32996	455	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 7.

The worst 5 of 455 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)
2:G:111:GLN:HE22	2:I:109:GLY:HA2	1.19	1.07
2:A:259:VAL:HG21	1:P:152:VAL:HG21	1.46	0.97
2:M:254:LEU:HG	2:M:286:TYR:HB3	1.50	0.94
2:B:259:VAL:HG21	1:Q:152:VAL:HG21	1.51	0.93
2:E:259:VAL:HG21	1:T:152:VAL:HG21	1.52	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	1	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0 0
1	2	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0 0
1	3	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0 0
1	4	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0 0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	P	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
1	Q	7/20 (35%)	1 (14%)	4 (57%)	2 (29%)	0	0
1	R	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
1	S	7/20 (35%)	2 (29%)	1 (14%)	4 (57%)	0	0
1	T	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
1	U	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
1	V	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
1	W	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
1	X	7/20 (35%)	1 (14%)	3 (43%)	3 (43%)	0	0
1	Y	7/20 (35%)	2 (29%)	1 (14%)	4 (57%)	0	0
1	Z	7/20 (35%)	2 (29%)	3 (43%)	2 (29%)	0	0
2	A	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	B	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	C	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	D	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	E	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	F	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	G	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	H	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	I	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	J	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	K	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	L	287/286 (100%)	282 (98%)	5 (2%)	0	100	100
2	M	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	N	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
2	O	287/286 (100%)	283 (99%)	4 (1%)	0	100	100
All	All	4410/4590 (96%)	4272 (97%)	103 (2%)	35 (1%)	24	65

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	151	LEU
1	1	153	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	2	151	LEU
1	2	153	PRO
1	3	151	LEU

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	1	6/18 (33%)	6 (100%)	0	100 100
1	2	6/18 (33%)	6 (100%)	0	100 100
1	3	6/18 (33%)	4 (67%)	2 (33%)	0 1
1	4	6/18 (33%)	6 (100%)	0	100 100
1	P	6/18 (33%)	6 (100%)	0	100 100
1	Q	6/18 (33%)	6 (100%)	0	100 100
1	R	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	S	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	T	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	U	6/18 (33%)	6 (100%)	0	100 100
1	V	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	W	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	X	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	Y	6/18 (33%)	5 (83%)	1 (17%)	3 13
1	Z	6/18 (33%)	6 (100%)	0	100 100
2	A	245/247 (99%)	239 (98%)	6 (2%)	57 84
2	B	245/247 (99%)	234 (96%)	11 (4%)	34 72
2	C	245/247 (99%)	236 (96%)	9 (4%)	41 76
2	D	246/247 (100%)	236 (96%)	10 (4%)	37 74
2	E	245/247 (99%)	234 (96%)	11 (4%)	34 72
2	F	245/247 (99%)	233 (95%)	12 (5%)	31 69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	G	245/247 (99%)	199 (81%)	46 (19%)	2 8
2	H	245/247 (99%)	201 (82%)	44 (18%)	2 10
2	I	245/247 (99%)	209 (85%)	36 (15%)	4 18
2	J	245/247 (99%)	206 (84%)	39 (16%)	3 15
2	K	245/247 (99%)	207 (84%)	38 (16%)	3 16
2	L	245/247 (99%)	202 (82%)	43 (18%)	2 10
2	M	245/247 (99%)	203 (83%)	42 (17%)	2 12
2	N	245/247 (99%)	201 (82%)	44 (18%)	2 10
2	O	245/247 (99%)	212 (86%)	33 (14%)	5 21
All	All	3766/3975 (95%)	3333 (88%)	433 (12%)	7 28

5 of 433 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	86	ASN
2	K	189	THR
2	O	74	SER
2	J	148	ASN
2	J	284	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
2	K	64	ASN
2	L	111	GLN
2	O	86	ASN
2	K	111	GLN
2	L	173	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

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

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	1	9/20 (45%)	-0.70	0	100	100	133, 141, 145, 158	0
1	2	9/20 (45%)	-0.42	0	100	100	120, 130, 150, 156	0
1	3	9/20 (45%)	-0.59	0	100	100	123, 129, 138, 144	0
1	4	9/20 (45%)	-0.20	0	100	100	108, 117, 143, 145	0
1	P	9/20 (45%)	-0.49	0	100	100	125, 137, 144, 155	0
1	Q	9/20 (45%)	-0.54	0	100	100	103, 113, 129, 169	0
1	R	9/20 (45%)	-0.32	0	100	100	103, 111, 128, 147	0
1	S	9/20 (45%)	-0.33	0	100	100	84, 93, 127, 141	0
1	T	9/20 (45%)	-0.27	0	100	100	140, 148, 156, 157	0
1	U	9/20 (45%)	-0.32	0	100	100	114, 128, 150, 168	0
1	V	9/20 (45%)	-0.42	0	100	100	117, 123, 150, 176	0
1	W	9/20 (45%)	-0.41	0	100	100	116, 127, 138, 142	0
1	X	9/20 (45%)	-0.63	0	100	100	126, 138, 147, 152	0
1	Y	9/20 (45%)	-0.33	0	100	100	115, 123, 137, 137	0
1	Z	9/20 (45%)	-0.79	0	100	100	111, 121, 144, 145	0
2	A	286/286 (100%)	0.11	7 (2%)	62	63	86, 119, 163, 203	1 (0%)
2	B	286/286 (100%)	0.05	8 (2%)	56	58	75, 113, 184, 198	1 (0%)
2	C	286/286 (100%)	0.04	5 (1%)	73	74	71, 104, 161, 207	1 (0%)
2	D	286/286 (100%)	-0.04	5 (1%)	73	74	72, 103, 144, 187	1 (0%)
2	E	286/286 (100%)	0.15	9 (3%)	52	55	81, 123, 167, 195	1 (0%)
2	F	286/286 (100%)	-0.00	5 (1%)	73	74	82, 118, 165, 206	1 (0%)
2	G	286/286 (100%)	0.04	6 (2%)	67	68	75, 113, 165, 210	1 (0%)
2	H	286/286 (100%)	-0.03	4 (1%)	78	80	79, 111, 154, 190	1 (0%)
2	I	286/286 (100%)	0.03	4 (1%)	78	80	84, 117, 160, 184	1 (0%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
2	J	286/286 (100%)	0.06	8 (2%) 56 58	80, 123, 168, 222	1 (0%)
2	K	286/286 (100%)	0.00	4 (1%) 78 80	84, 112, 155, 199	1 (0%)
2	L	286/286 (100%)	0.03	5 (1%) 73 74	79, 120, 159, 180	1 (0%)
2	M	286/286 (100%)	0.00	3 (1%) 84 86	69, 114, 152, 189	1 (0%)
2	N	286/286 (100%)	-0.06	1 (0%) 94 95	71, 110, 154, 186	1 (0%)
2	O	286/286 (100%)	-0.04	2 (0%) 89 90	72, 111, 153, 184	1 (0%)
All	All	4425/4590 (96%)	0.01	76 (1%) 73 74	69, 115, 161, 222	15 (0%)

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	34	GLY	7.6
2	F	34	GLY	7.3
2	C	35	ALA	6.1
2	M	34	GLY	5.2
2	C	34	GLY	5.2

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

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.