wwPDB X-ray Structure Validation Summary Report (i)

Sep 26, 2023 – 10:09 PM EDT

PDB ID	:	6CHB
Title	:	Crystal structure of a natively-glycosylated BG505 SOSIP.664 HIV-1 Envelope Trimer in complex with the broadly-neutralizing antibodies BG18 and IOMA
Authors	:	Barnes, C.O.; Bjorkman, P.J.
Deposited on	:	2018-02-22
Resolution	:	6.80 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 (1)) were used in the production of this report:

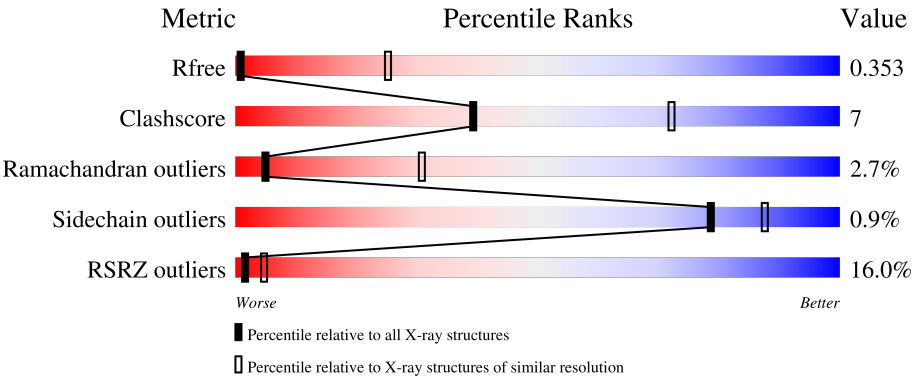
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1003 (9.70-3.90)
Clashscore	141614	1067 (9.70-3.90)
Ramachandran outliers	138981	1001 (9.70-3.90)
Sidechain outliers	138945	1001 (9.70-3.86)
RSRZ outliers	127900	1004 (9.50-3.80)

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	153	4% 63% 17% • 18%
1	B	153	14% 59% 22% • 18%
1	C	153	14% 56% 25% • 18%
2	F	479	7% 75% 19% 6%
2	G	479	11% 70% 23% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	479	10% 76% 16% 6%
3	I	241	19% 69% 25% 6%
3	J	241	19% 74% 19% 6%
3	Q	241	28% 74% 20% 6%
4	K	215	15% 80% 16% 5%
4	L	215	20% 74% 20% 5%
4	R	215	27% 80% 15% 5%
5	D	232	9% 42% 10% 46%
5	M	232	14% 41% 11% 46%
5	O	232	6% 43% 10% 46%
6	E	214	7% 44% 5% 50%
6	N	214	12% 44% 6% 49%
6	P	214	5% 37% 12% 50%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 28753 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 Envelope glycoprotein gp41.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	126	Total	C	N	O	S	0	0	0
			1001	633	172	190	6			
1	A	126	Total	C	N	O	S	0	0	0
			1001	633	172	190	6			
1	C	126	Total	C	N	O	S	0	0	0
			1001	633	172	190	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	605	CYS	THR	engineered mutation	UNP Q2N0S7
A	605	CYS	THR	engineered mutation	UNP Q2N0S7
C	605	CYS	THR	engineered mutation	UNP Q2N0S7

- Molecule 2 is a protein called Envelope glycoprotein gp120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	450	Total	C	N	O	S	0	0	0
			3538	2221	624	666	27			
2	F	450	Total	C	N	O	S	0	0	0
			3538	2221	624	666	27			
2	H	450	Total	C	N	O	S	0	0	0
			3538	2221	624	666	27			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	332	ASN	THR	conflict	UNP Q2N0S6
G	501	CYS	ALA	engineered mutation	UNP Q2N0S6
F	332	ASN	THR	conflict	UNP Q2N0S6
F	501	CYS	ALA	engineered mutation	UNP Q2N0S6
H	332	ASN	THR	conflict	UNP Q2N0S6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	501	CYS	ALA	engineered mutation	UNP Q2N0S6

- Molecule 3 is a protein called BG18 Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	227	Total	C	N	O	S	0	0	0
			1723	1086	297	332	8			
3	I	227	Total	C	N	O	S	0	0	0
			1723	1086	297	332	8			
3	Q	227	Total	C	N	O	S	0	0	0
			1723	1086	297	332	8			

- Molecule 4 is a protein called BG18 Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	K	205	Total	C	N	O	S	0	0	0
			1540	963	257	314	6			
4	L	205	Total	C	N	O	S	0	0	0
			1540	963	257	314	6			
4	R	205	Total	C	N	O	S	0	0	0
			1540	963	257	314	6			

- Molecule 5 is a protein called IOMA Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	125	Total	C	N	O	S	0	0	0
			992	629	173	181	9			
5	M	125	Total	C	N	O	S	0	0	0
			992	629	173	181	9			
5	O	125	Total	C	N	O	S	0	0	0
			992	629	173	181	9			

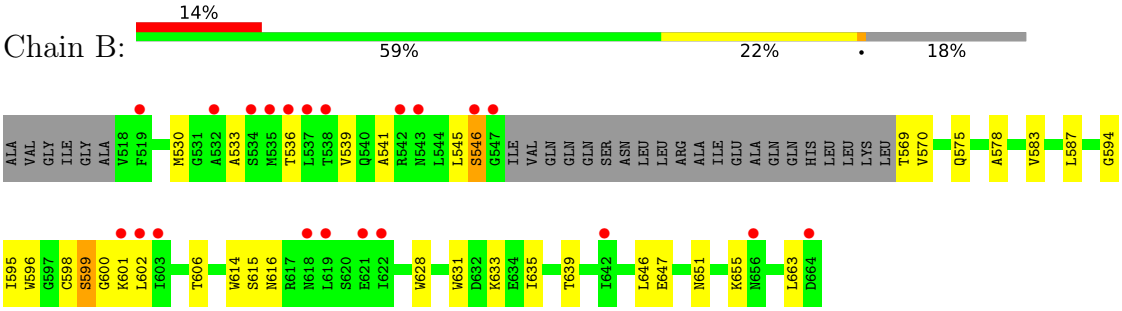
- Molecule 6 is a protein called IOMA Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	E	107	Total	C	N	O	S	0	0	0
			786	491	133	160	2			
6	N	109	Total	C	N	O	S	0	0	0
			799	498	136	163	2			
6	P	107	Total	C	N	O	S	0	0	0
			786	491	133	160	2			

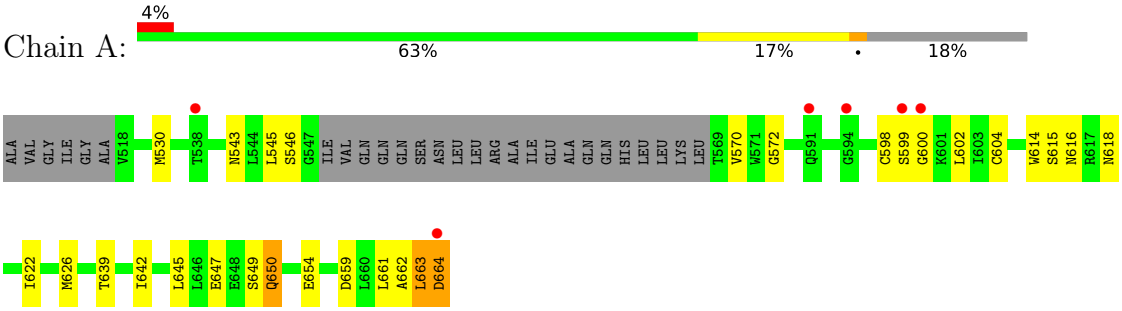
3 Residue-property plots (i)

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.

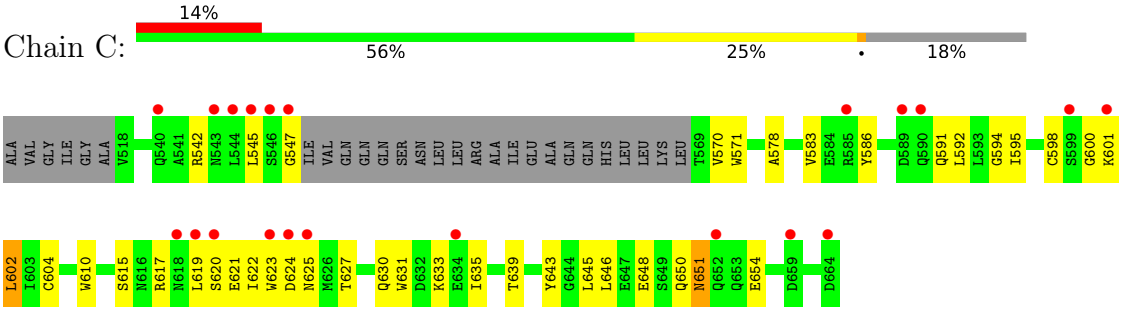
• Molecule 1: Envelope glycoprotein gp41



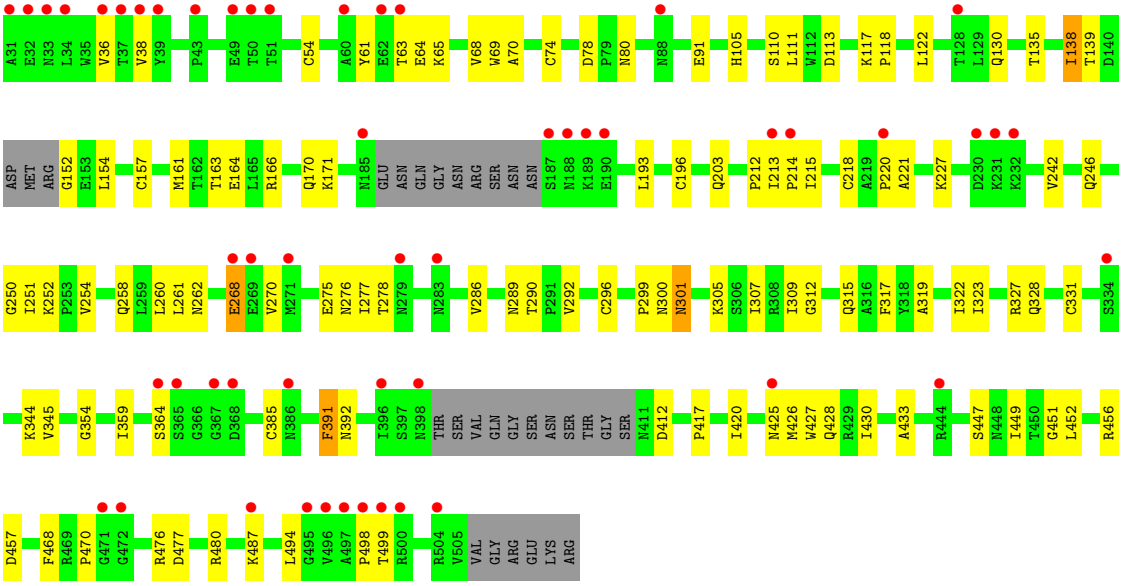
• Molecule 1: Envelope glycoprotein gp41



• Molecule 1: Envelope glycoprotein gp41

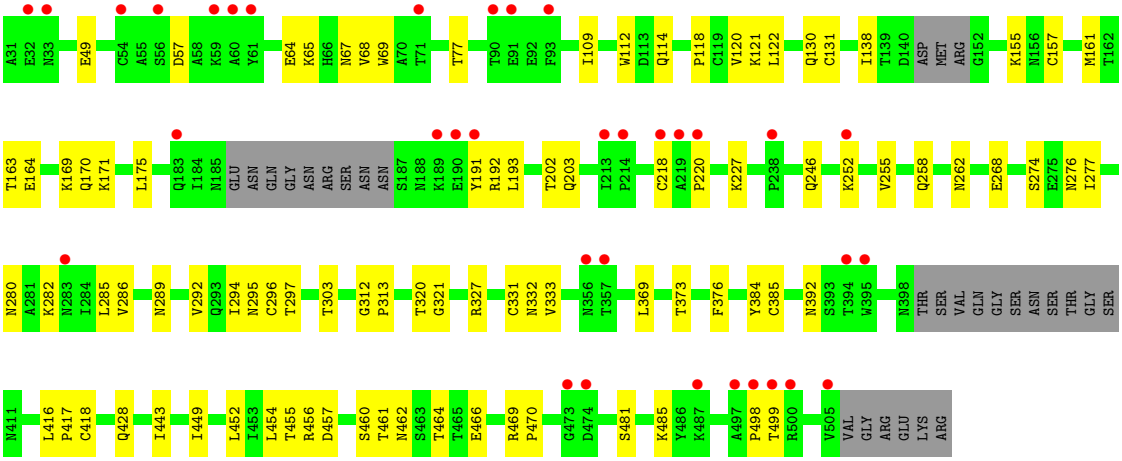


• Molecule 2: Envelope glycoprotein gp120



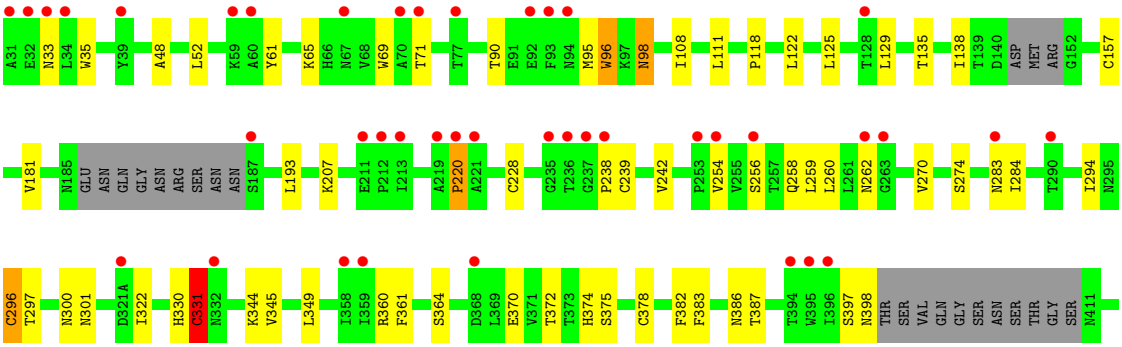
• Molecule 2: Envelope glycoprotein gp120

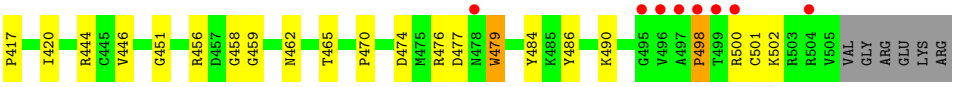
Chain F:



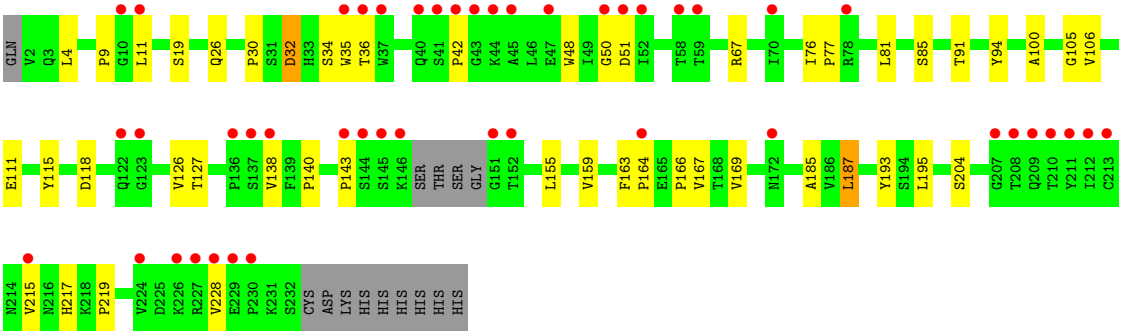
• Molecule 2: Envelope glycoprotein gp120

Chain H:

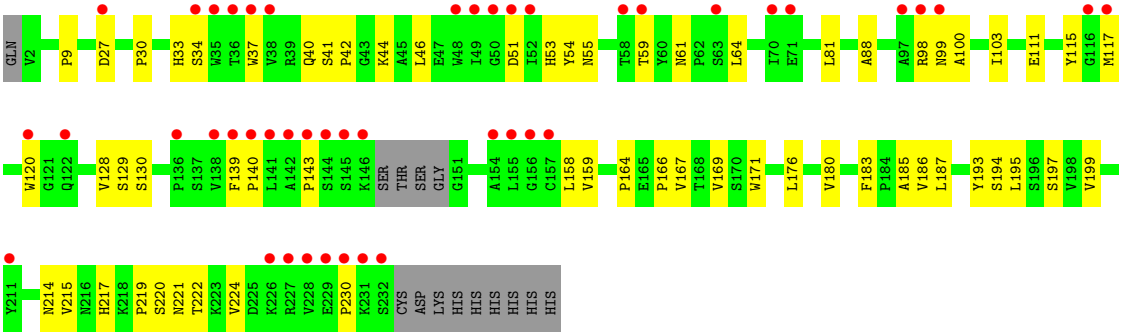




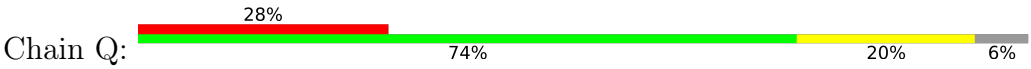
• Molecule 3: BG18 Heavy Chain

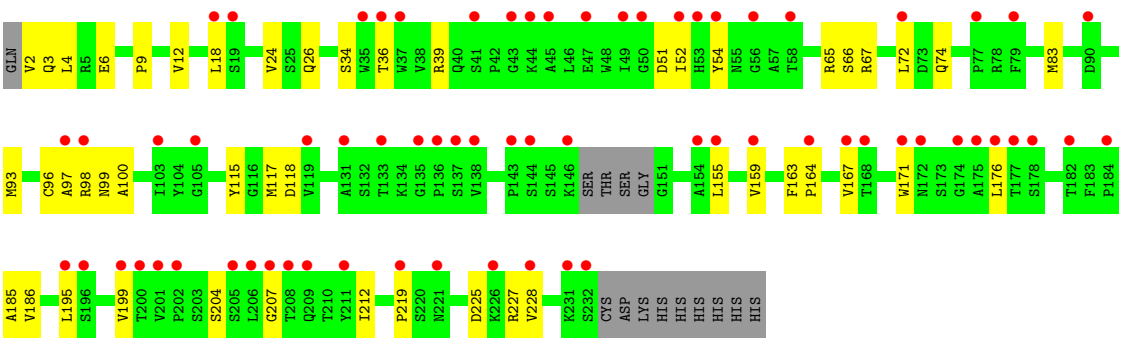


• Molecule 3: BG18 Heavy Chain

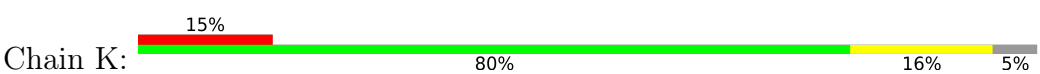


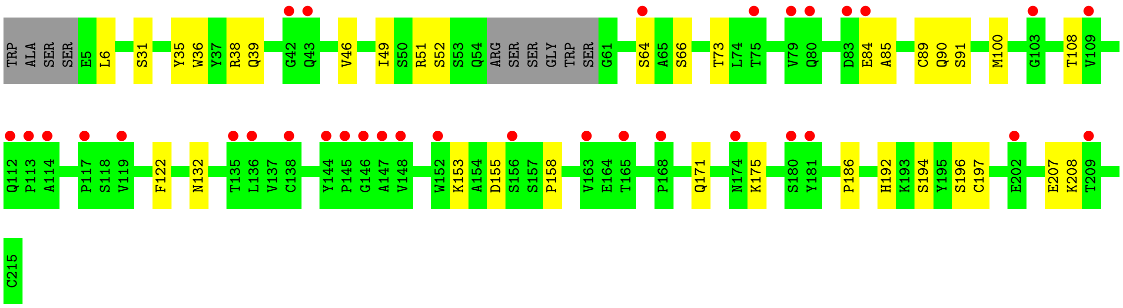
• Molecule 3: BG18 Heavy Chain





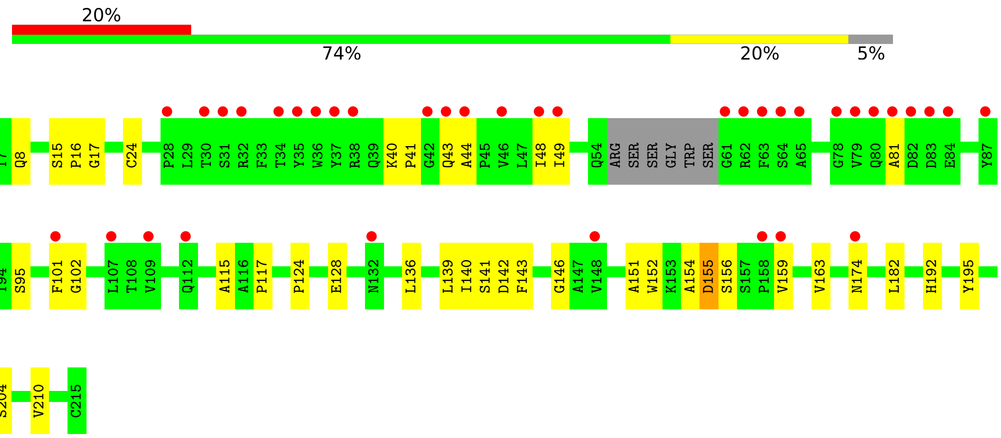
• Molecule 4: BG18 Light Chain





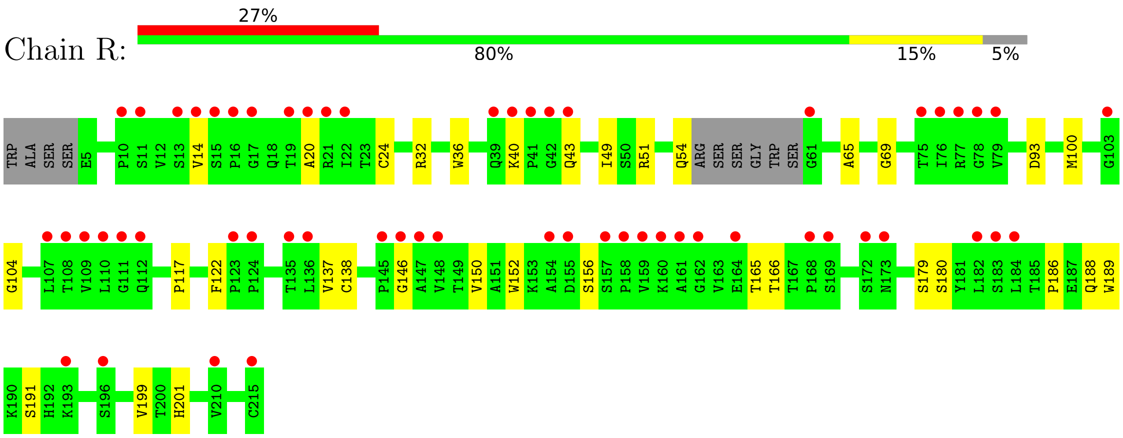
• Molecule 4: BG18 Light Chain

Chain L:



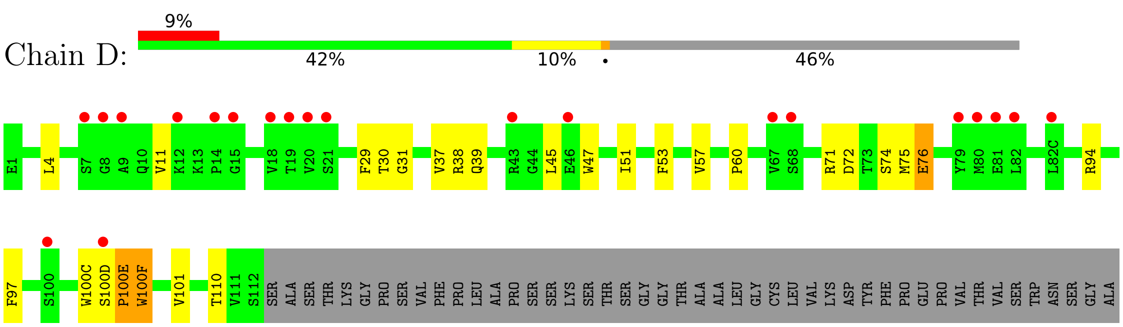
• Molecule 4: BG18 Light Chain

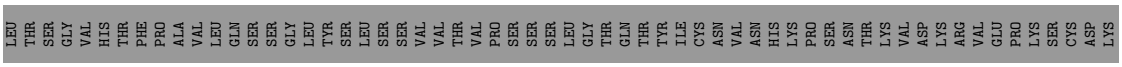
Chain R:



• Molecule 5: IOMA Heavy Chain

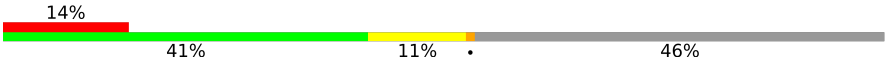
Chain D:

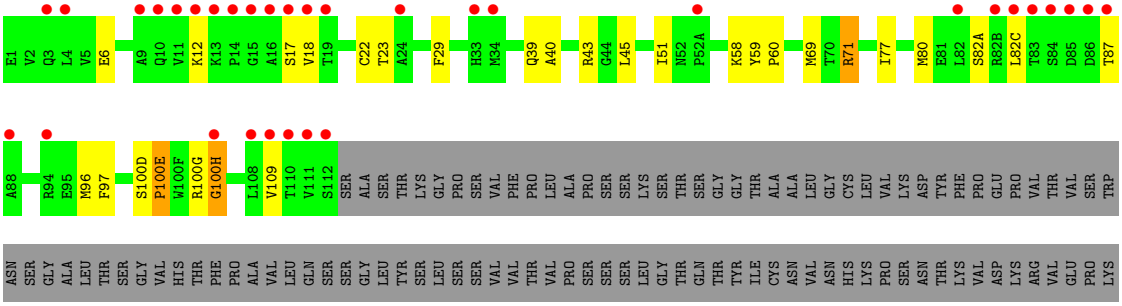




• Molecule 5: IOMA Heavy Chain

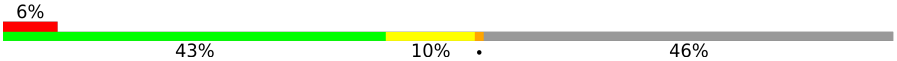
Chain M:

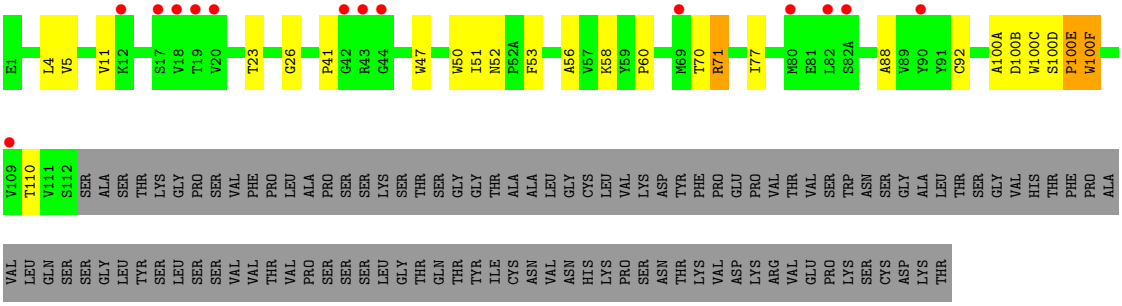




• Molecule 5: IOMA Heavy Chain

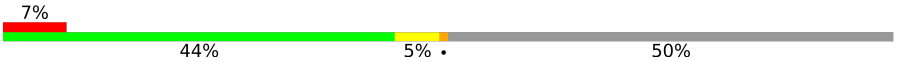
Chain O:

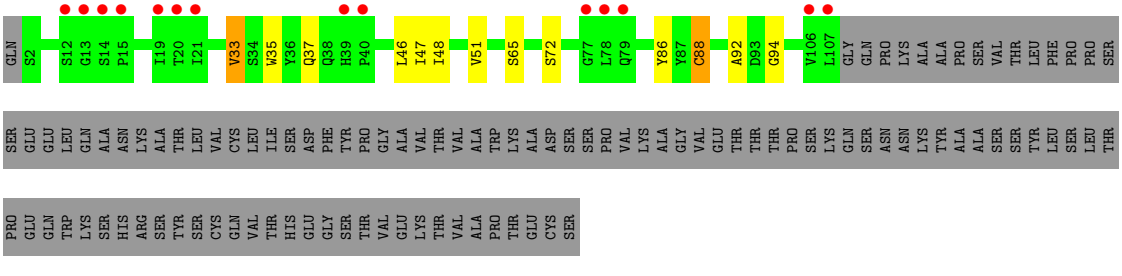




• Molecule 6: IOMA Light Chain

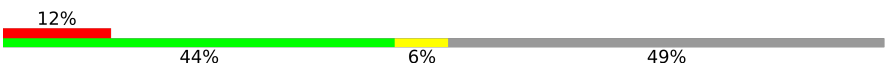
Chain E:

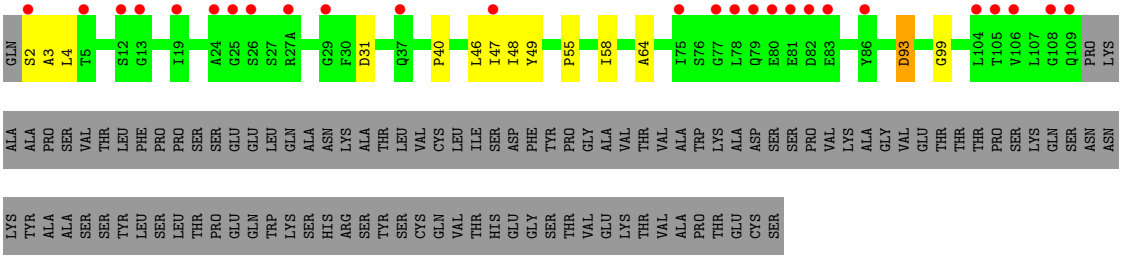




• Molecule 6: IOMA Light Chain

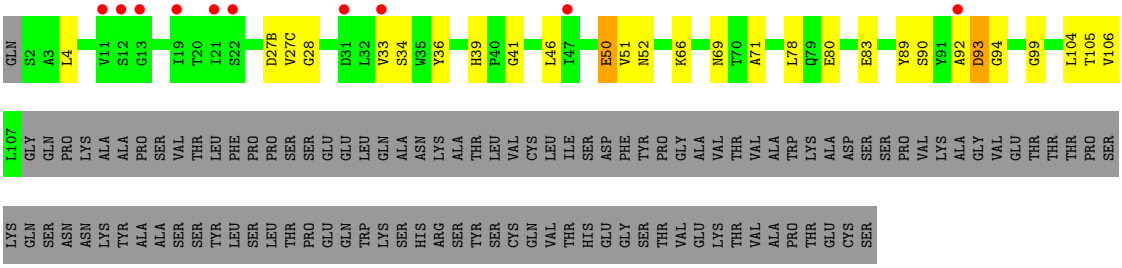
Chain N:





• Molecule 6: IOMA Light Chain

Chain P:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	176.75Å 176.75Å 458.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.63 – 6.80 39.63 – 6.68	Depositor EDS
% Data completeness (in resolution range)	99.7 (39.63-6.80) 99.4 (39.63-6.68)	Depositor EDS
R_{merge}	0.33	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.36 (at 6.65Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.318 , 0.417 0.328 , 0.353	Depositor DCC
R_{free} test set	670 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å^2)	369.1	Xtriage
Anisotropy	0.560	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 533.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.80	EDS
Total number of atoms	28753	wwPDB-VP
Average B, all atoms (Å^2)	411.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.64% 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	0.25	0/1019	0.47	0/1382
1	B	0.25	0/1019	0.46	0/1382
1	C	0.25	0/1019	0.48	0/1382
2	F	0.25	0/3611	0.48	0/4903
2	G	0.26	0/3611	0.49	0/4903
2	H	0.26	0/3611	0.49	0/4903
3	I	0.25	0/1767	0.51	0/2410
3	J	0.25	0/1767	0.51	0/2410
3	Q	0.25	0/1767	0.49	0/2410
4	K	0.25	0/1577	0.48	0/2153
4	L	0.25	0/1577	0.48	0/2153
4	R	0.26	0/1577	0.46	0/2153
5	D	0.28	0/1021	0.51	0/1384
5	M	0.25	0/1021	0.51	0/1384
5	O	0.27	0/1021	0.51	0/1384
6	E	0.25	0/803	0.45	0/1088
6	N	0.26	0/816	0.47	0/1105
6	P	0.25	0/803	0.50	0/1088
All	All	0.26	0/29407	0.49	0/39977

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
5	D	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	659	ASP	Peptide
5	D	100(C)	TRP	Peptide

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	1001	0	978	16	1
1	B	1001	0	979	27	0
1	C	1001	0	979	25	0
2	F	3538	0	3488	49	0
2	G	3538	0	3493	64	0
2	H	3538	0	3491	45	0
3	I	1723	0	1696	36	0
3	J	1723	0	1694	27	0
3	Q	1723	0	1694	29	0
4	K	1540	0	1492	19	0
4	L	1540	0	1492	26	0
4	R	1540	0	1492	16	0
5	D	992	0	949	16	0
5	M	992	0	947	14	0
5	O	992	0	947	16	0
6	E	786	0	758	7	0
6	N	799	0	769	7	0
6	P	786	0	758	16	0
All	All	28753	0	28096	419	1

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 419 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)
3:Q:159:VAL:HG11	3:Q:167:VAL:HG11	1.67	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:P:83:GLU:HG3	6:P:105:THR:HA	1.66	0.77
5:M:69:MET:HE2	5:M:80:MET:HG3	1.68	0.73
3:J:140:PRO:HB3	3:J:228:VAL:HG13	1.73	0.71
2:F:295:ASN:HB2	2:F:332:ASN:HB2	1.72	0.70

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:661:LEU:O	1:A:663:LEU:N[8_554]	2.16	0.04

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	122/153 (80%)	93 (76%)	24 (20%)	5 (4%)	3 22
1	B	122/153 (80%)	104 (85%)	15 (12%)	3 (2%)	5 32
1	C	122/153 (80%)	97 (80%)	20 (16%)	5 (4%)	3 22
2	F	442/479 (92%)	363 (82%)	71 (16%)	8 (2%)	8 40
2	G	442/479 (92%)	377 (85%)	51 (12%)	14 (3%)	4 26
2	H	442/479 (92%)	370 (84%)	57 (13%)	15 (3%)	3 26
3	I	223/241 (92%)	195 (87%)	21 (9%)	7 (3%)	4 27
3	J	223/241 (92%)	188 (84%)	27 (12%)	8 (4%)	3 25
3	Q	223/241 (92%)	195 (87%)	24 (11%)	4 (2%)	8 40
4	K	201/215 (94%)	175 (87%)	24 (12%)	2 (1%)	15 54
4	L	201/215 (94%)	174 (87%)	24 (12%)	3 (2%)	10 46
4	R	201/215 (94%)	175 (87%)	22 (11%)	4 (2%)	7 38
5	D	123/232 (53%)	103 (84%)	16 (13%)	4 (3%)	4 26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	M	123/232 (53%)	106 (86%)	12 (10%)	5 (4%)	3 22
5	O	123/232 (53%)	98 (80%)	22 (18%)	3 (2%)	6 33
6	E	105/214 (49%)	88 (84%)	16 (15%)	1 (1%)	15 54
6	N	107/214 (50%)	91 (85%)	13 (12%)	3 (3%)	5 30
6	P	105/214 (49%)	81 (77%)	20 (19%)	4 (4%)	3 24
All	All	3650/4602 (79%)	3073 (84%)	479 (13%)	98 (3%)	5 31

5 of 98 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	D	74	SER
5	D	100(E)	PRO
5	M	100(E)	PRO
3	I	30	PRO
3	I	166	PRO

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	108/129 (84%)	106 (98%)	2 (2%)	57 75
1	B	108/129 (84%)	107 (99%)	1 (1%)	78 87
1	C	108/129 (84%)	105 (97%)	3 (3%)	43 65
2	F	401/426 (94%)	397 (99%)	4 (1%)	76 86
2	G	401/426 (94%)	400 (100%)	1 (0%)	93 96
2	H	401/426 (94%)	396 (99%)	5 (1%)	71 83
3	I	195/208 (94%)	195 (100%)	0	100 100
3	J	195/208 (94%)	194 (100%)	1 (0%)	88 93
3	Q	195/208 (94%)	194 (100%)	1 (0%)	88 93
4	K	174/182 (96%)	173 (99%)	1 (1%)	86 92
4	L	174/182 (96%)	174 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	R	174/182 (96%)	173 (99%)	1 (1%)	86 92
5	D	104/197 (53%)	100 (96%)	4 (4%)	33 57
5	M	104/197 (53%)	102 (98%)	2 (2%)	57 75
5	O	104/197 (53%)	102 (98%)	2 (2%)	57 75
6	E	85/177 (48%)	84 (99%)	1 (1%)	71 83
6	N	86/177 (49%)	86 (100%)	0	100 100
6	P	85/177 (48%)	85 (100%)	0	100 100
All	All	3202/3957 (81%)	3173 (99%)	29 (1%)	78 87

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	469	ARG
3	Q	225	ASP
1	C	604	CYS
2	H	479	TRP
5	M	71	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	F	170	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 monosaccharides 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	126/153 (82%)	0.37	6 (4%) 30 29	365, 471, 570, 691	0
1	B	126/153 (82%)	0.94	21 (16%) 1 4	309, 438, 494, 548	0
1	C	126/153 (82%)	0.83	21 (16%) 1 4	329, 445, 536, 592	0
2	F	450/479 (93%)	0.51	35 (7%) 13 15	209, 400, 510, 597	0
2	G	450/479 (93%)	0.76	53 (11%) 4 8	212, 372, 474, 595	0
2	H	450/479 (93%)	0.68	49 (10%) 5 9	241, 392, 532, 589	0
3	I	227/241 (94%)	1.11	45 (19%) 1 3	277, 411, 505, 576	0
3	J	227/241 (94%)	0.94	46 (20%) 1 3	270, 380, 474, 594	0
3	Q	227/241 (94%)	1.48	68 (29%) 0 2	341, 442, 578, 700	0
4	K	205/215 (95%)	0.95	33 (16%) 1 4	297, 386, 473, 543	0
4	L	205/215 (95%)	1.16	43 (20%) 1 3	324, 407, 503, 613	0
4	R	205/215 (95%)	1.39	57 (27%) 0 2	354, 521, 620, 698	0
5	D	125/232 (53%)	1.04	21 (16%) 1 4	237, 362, 466, 541	0
5	M	125/232 (53%)	1.20	33 (26%) 0 2	303, 386, 510, 581	0
5	O	125/232 (53%)	0.60	14 (11%) 5 9	223, 325, 420, 472	0
6	E	107/214 (50%)	0.80	14 (13%) 3 7	319, 436, 537, 605	0
6	N	109/214 (50%)	1.24	26 (23%) 0 2	351, 457, 555, 582	0
6	P	107/214 (50%)	0.56	10 (9%) 8 11	290, 406, 493, 559	0
All	All	3722/4602 (80%)	0.89	595 (15%) 1 5	209, 409, 546, 700	0

The worst 5 of 595 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	187	SER	12.6
3	Q	232	SER	9.6
2	G	188	ASN	8.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	G	31	ALA	7.7
2	G	32	GLU	7.5

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 monosaccharides 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.