



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2020 – 01:25 PM BST

PDB ID : 4O3T
Title : Zymogen HGF-beta/MET with Zymogen Activator Peptide ZAP.14
Authors : Eigenbrot, C.; Landgraf, K.E.; Steffek, M.
Deposited on : 2013-12-18
Resolution : 2.99 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.13.1
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) : 2.13.1

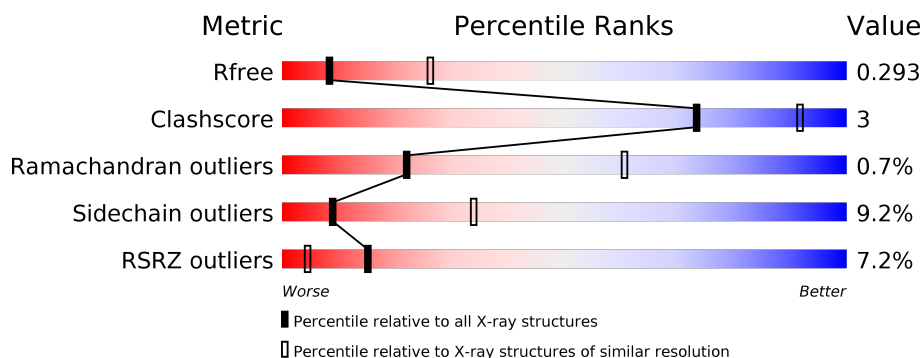
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.99 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 $\leq 5\%$. 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	240	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>10%</div> <div>9%</div> </div> </div>
2	B	551	<div> <div>9%</div> <div> <div></div> <div>75%</div> <div>15%</div> <div>9%</div> </div> </div>
3	P	11	<div> <div></div> <div> <div></div> <div>73%</div> <div>9%</div> <div>18%</div> </div> </div>
4	C	3	<div> <div></div> <div> <div></div> <div>100%</div> </div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5804 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 Hepatocyte growth factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	219	Total	C	N	O	S	0	0	0
			1707	1086	302	305	14			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	495	GLY	VAL	engineered mutation	UNP P14210
A	604	SER	CYS	engineered mutation	UNP P14210
A	729	HIS	-	expression tag	UNP P14210
A	730	HIS	-	expression tag	UNP P14210
A	731	HIS	-	expression tag	UNP P14210
A	732	HIS	-	expression tag	UNP P14210
A	733	HIS	-	expression tag	UNP P14210
A	734	HIS	-	expression tag	UNP P14210

- Molecule 2 is a protein called Hepatocyte growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	499	Total	C	N	O	S	0	0	0
			3965	2520	674	742	29			

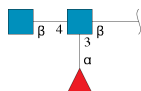
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	568	HIS	-	expression tag	UNP P08581
B	569	HIS	-	expression tag	UNP P08581
B	570	HIS	-	expression tag	UNP P08581
B	571	HIS	-	expression tag	UNP P08581
B	572	HIS	-	expression tag	UNP P08581
B	573	HIS	-	expression tag	UNP P08581
B	574	HIS	-	expression tag	UNP P08581
B	575	HIS	-	expression tag	UNP P08581

- Molecule 3 is a protein called ZAP.14.

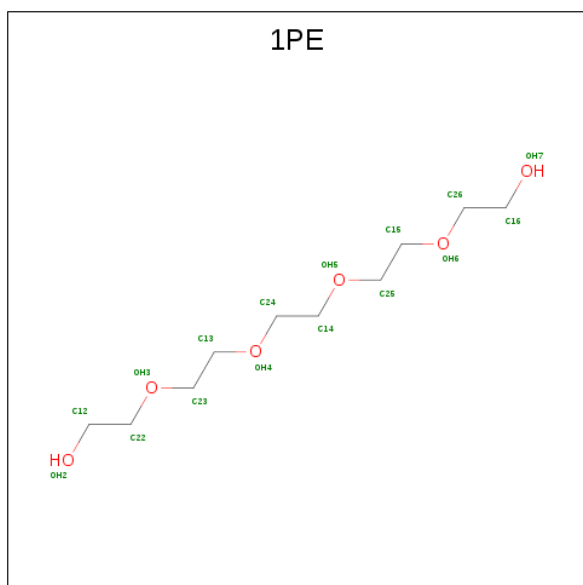
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	9	Total	C	N	O	S	0	0	0
			78	56	11	10	1			

- Molecule 4 is an oligosaccharide called alpha-L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	3	Total	C	N	O		0	0	0
			38	22	2	14				

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C₁₀H₂₂O₆).

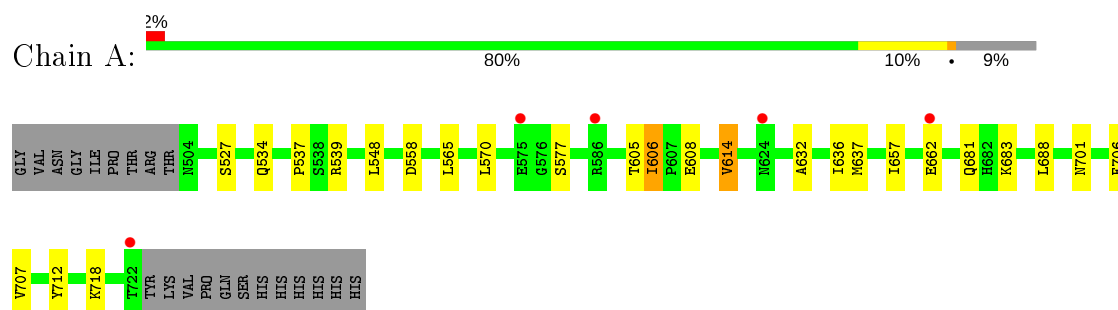


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			16	10	6		

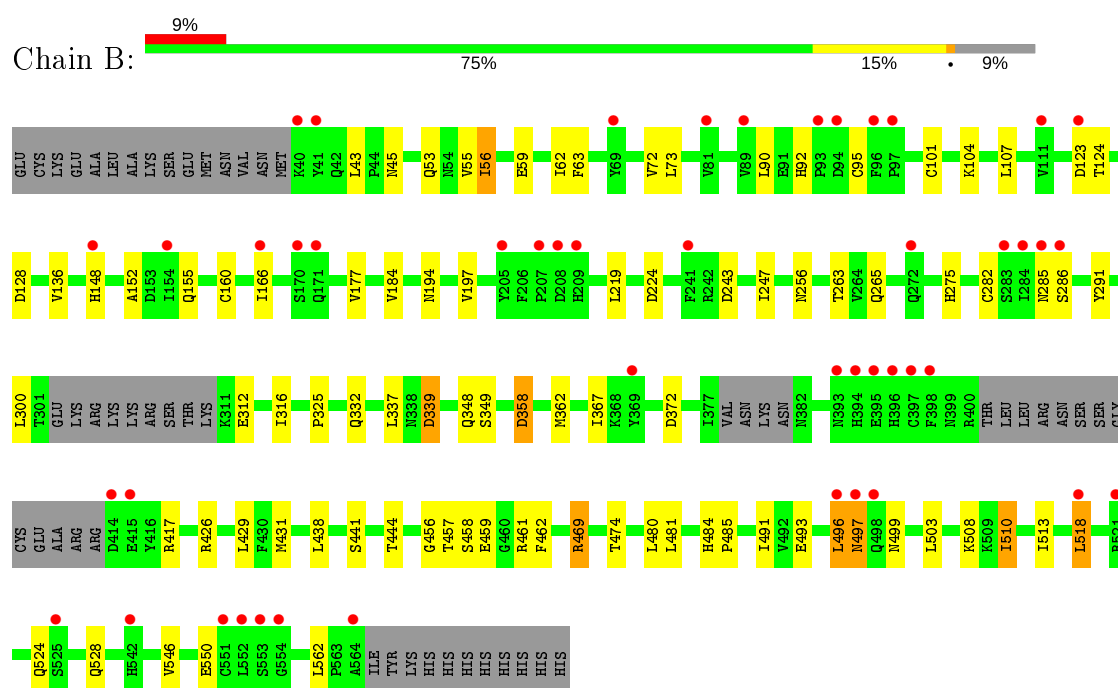
3 Residue-property plots [i](#)

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.

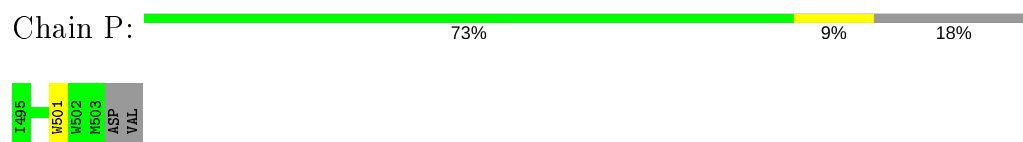
- Molecule 1: Hepatocyte growth factor



- Molecule 2: Hepatocyte growth factor receptor



- Molecule 3: ZAP.14



- Molecule 4: α -L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

100%

MAG1
FUC2
MAG3

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.79Å 121.47Å 137.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.54 – 2.99 45.54 – 2.99	Depositor EDS
% Data completeness (in resolution range)	99.5 (45.54-2.99) 99.5 (45.54-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.18 (at 3.01Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.232 , 0.276 0.243 , 0.293	Depositor DCC
R_{free} test set	877 reflections (4.21%)	wwPDB-VP
Wilson B-factor (Å ²)	63.9	Xtriage
Anisotropy	0.693	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	5804	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 1PE, FUC, NAG

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	0.37	0/1748	0.64	0/2366
2	B	0.38	0/4065	0.67	0/5519
3	P	0.40	0/83	0.61	0/114
All	All	0.38	0/5896	0.66	0/7999

There are no bond length outliers.

There are no bond angle outliers.

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	A	1707	0	1699	10	0
2	B	3965	0	3814	26	0
3	P	78	0	70	0	0
4	C	38	0	34	3	0
5	A	16	0	22	0	0
All	All	5804	0	5639	38	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 3.

The worst 5 of 38 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:B:263:THR:HG21	2:B:265:GLN:HE21	1.44	0.83
2:B:43:LEU:HD21	2:B:518:LEU:HD21	1.73	0.69
2:B:358:ASP:HA	2:B:438:LEU:HB2	1.77	0.67
1:A:606:ILE:HD11	1:A:657:ILE:HD11	1.78	0.66
1:A:577:SER:HB3	1:A:712:TYR:CZ	2.35	0.61

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	217/240 (90%)	202 (93%)	15 (7%)	0	100	100
2	B	491/551 (89%)	454 (92%)	33 (7%)	4 (1%)	19	57
3	P	7/11 (64%)	5 (71%)	1 (14%)	1 (14%)	0	1
All	All	715/802 (89%)	661 (92%)	49 (7%)	5 (1%)	22	60

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	148	HIS
2	B	496	LEU
2	B	128	ASP
3	P	501	TRP
2	B	497	ASN

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	184/203 (91%)	170 (92%)	14 (8%)	13	43
2	B	451/499 (90%)	406 (90%)	45 (10%)	7	29
3	P	7/9 (78%)	7 (100%)	0	100	100
All	All	642/711 (90%)	583 (91%)	59 (9%)	9	34

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	177	VAL
2	B	337	LEU
2	B	510	ILE
2	B	219	LEU
2	B	243	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	194	ASN
2	B	256	ASN
2	B	394	HIS
2	B	171	GLN
2	B	318	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

3 monosaccharides 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
4	NAG	C	1	2,4	14,14,15	0.52	0	17,19,21	1.45	3 (17%)
4	FUC	C	2	4	10,10,11	0.48	0	14,14,16	1.24	2 (14%)
4	NAG	C	3	4	14,14,15	0.29	0	17,19,21	0.77	1 (5%)

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
4	NAG	C	1	2,4	-	0/6/23/26	0/1/1/1
4	FUC	C	2	4	-	-	0/1/1/1
4	NAG	C	3	4	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	2	FUC	C1-O5-C5	3.59	120.91	112.78
4	C	1	NAG	C1-O5-C5	3.57	117.03	112.19
4	C	1	NAG	O5-C1-C2	-3.57	105.65	111.29
4	C	3	NAG	C1-O5-C5	2.98	116.23	112.19
4	C	1	NAG	O3-C3-C4	2.66	116.50	110.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	2	FUC	1	0
4	C	1	NAG	3	0
4	C	3	NAG	2	0

5.6 Ligand geometry

1 ligand is 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	$\# Z > 2$
5	1PE	A	801	-	15,15,15	0.57	0	14,14,14	0.15	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
5	1PE	A	801	-	-	9/13/13/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	801	1PE	OH7-C16-C26-OH6
5	A	801	1PE	OH2-C12-C22-OH3
5	A	801	1PE	OH6-C15-C25-OH5
5	A	801	1PE	C16-C26-OH6-C15
5	A	801	1PE	C12-C22-OH3-C23

There are no ring outliers.

No monomer is involved in short contacts.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	219/240 (91%)	0.24	5 (2%) 60 31	39, 58, 85, 100	0
2	B	499/551 (90%)	0.58	47 (9%) 8 3	42, 65, 107, 135	0
3	P	9/11 (81%)	0.29	0 100 100	50, 53, 57, 64	0
All	All	727/802 (90%)	0.48	52 (7%) 15 4	39, 62, 102, 135	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	497	ASN	7.3
2	B	208	ASP	7.3
2	B	209	HIS	7.0
2	B	564	ALA	6.8
2	B	207	PRO	6.5

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	NAG	C	1	14/15	0.71	0.28	78,84,85,86	0
4	NAG	C	3	14/15	0.81	0.28	85,87,89,89	0
4	FUC	C	2	10/11	0.83	0.28	85,86,87,87	0

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	1PE	A	801	16/16	0.81	0.21	54,61,66,66	0

6.5 Other polymers

There are no such residues in this entry.