



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 12, 2016 – 04:51 PM EST

PDB ID : 4Z6A  
Title : Crystal Structure of a FVIIa-Trypsin Chimera (YT) in Complex with Soluble Tissue Factor  
Authors : Sorensen, A.B.; Svensson, L.A.; Gandhi, P.S.  
Deposited on : 2015-04-04  
Resolution : 2.25 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028442  
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) : rb-20028442

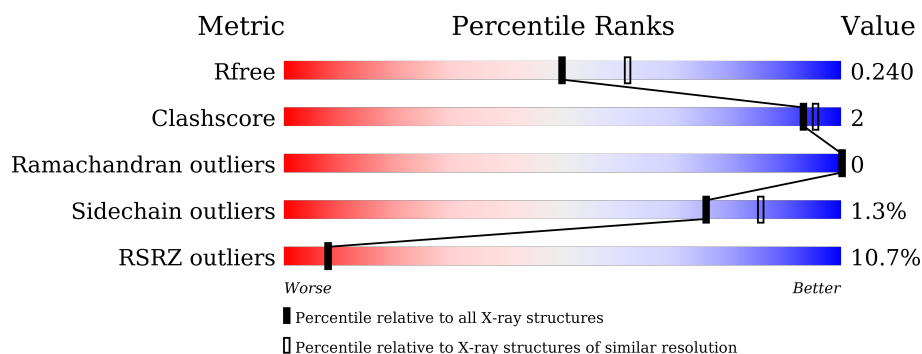
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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	1640 (2.28-2.24)
Clashscore	102246	1095 (2.26-2.26)
Ramachandran outliers	100387	1063 (2.26-2.26)
Sidechain outliers	100360	1063 (2.26-2.26)
RSRZ outliers	91569	1647 (2.28-2.24)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	L	96	<div> <div>25%</div> <div> <div></div> <div>89%</div> <div>8%</div> <div>..</div> </div> </div>
2	H	249	<div> <div>9%</div> <div> <div></div> <div>97%</div> <div>.</div> </div> </div>
3	T	207	<div> <div>5%</div> <div> <div></div> <div>92%</div> <div>5%</div> <div>.</div> </div> </div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 8483 atoms, of which 3992 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 Coagulation factor VII.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	L	94	Total	C	H	N	O	S	0	0	0
			1238	406	563	117	139	13			

- Molecule 2 is a protein called Coagulation factor VII.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	H	249	Total	C	H	N	O	S	0	4	0
			3797	1228	1870	337	349	13			

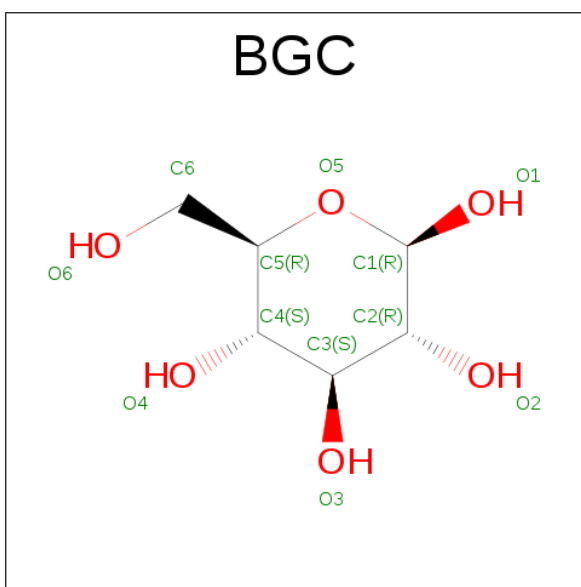
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	?	-	LEU	deletion	UNP P08709
H	?	-	GLN	deletion	UNP P08709
H	?	-	GLN	deletion	UNP P08709
H	?	-	SER	deletion	UNP P08709
H	?	-	ARG	deletion	UNP P08709
H	169	GLU	LYS	engineered mutation	UNP P08709
H	170	ALA	VAL	engineered mutation	UNP P08709
H	171	SER	GLY	engineered mutation	UNP P08709
H	172	TYR	ASP	engineered mutation	UNP P08709
H	173	PRO	SER	engineered mutation	UNP P08709
H	174	GLY	PRO	engineered mutation	UNP P08709
H	175	LYS	ASN	engineered mutation	UNP P08709

- Molecule 3 is a protein called Tissue factor.

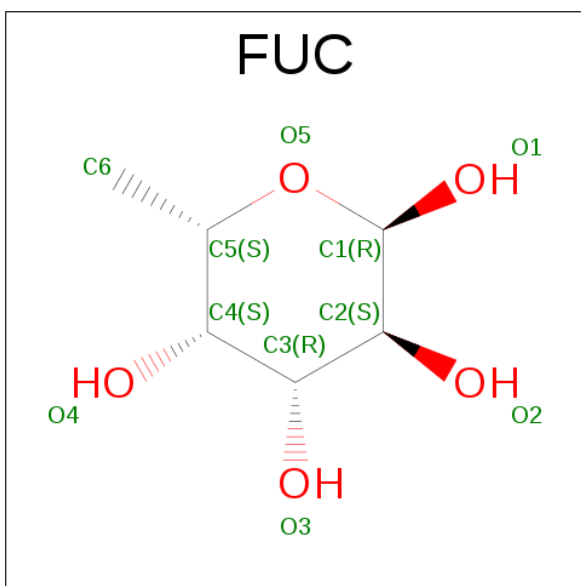
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	T	201	Total	C	H	N	O	S	0	0	0
			3087	1011	1497	255	319	5			

- Molecule 4 is BETA-D-GLUCOSE (three-letter code: BGC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



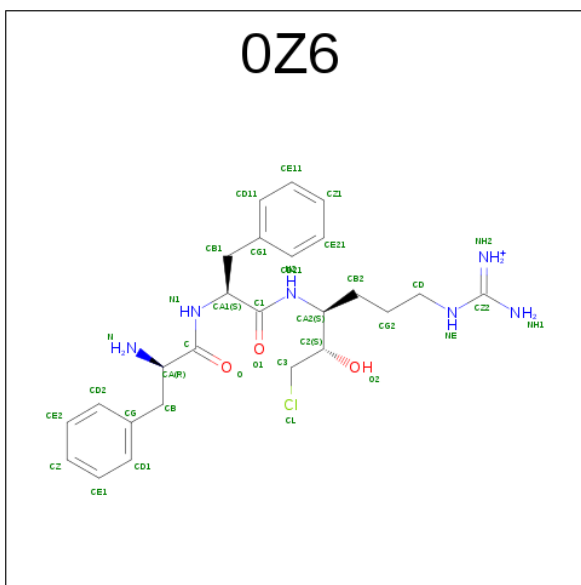
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	L	1	Total	C	H	O	0	0
			22	6	11	5		

- Molecule 5 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula:  $C_6H_{12}O_5$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	L	1	Total	C	H	O	0	0
			20	6	10	4		

- Molecule 6 is D-phenylalanyl-N-[(2S,3S)-6-{[amino(iminio)methyl]amino}-1-chloro-2-hydroxyhexan-3-yl]-L-phenylalaninamide (three-letter code: 0Z6) (formula:  $C_{25}H_{36}ClN_6O_3$ ).

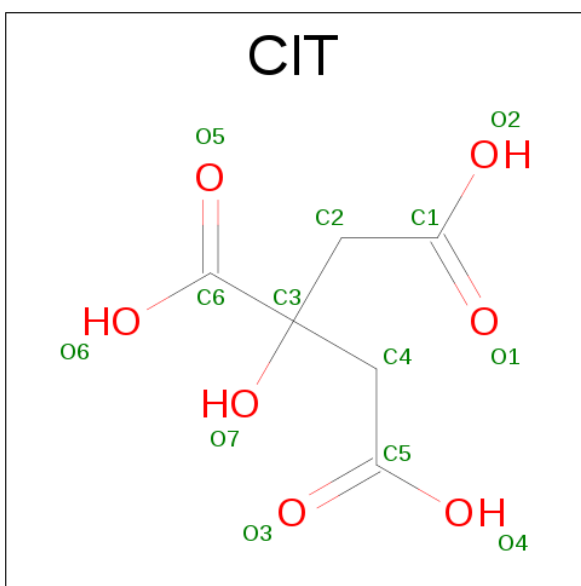


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	H	1	Total	C	H	N	O	0	0
			70	25	36	6	3		

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	H	1	Total Ca 1 1	0	0

- Molecule 8 is CITRIC ACID (three-letter code: CIT) (formula:  $\text{C}_6\text{H}_8\text{O}_7$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	T	1	Total	C	H	O	0	0
			18	6	5	7		


- Molecule 9 is water.

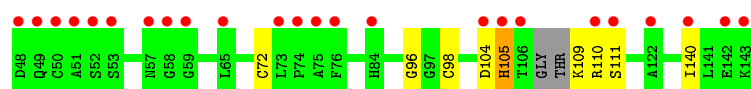
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	L	26	Total	O	0	0
			26	26		
9	H	141	Total	O	0	0
			141	141		
9	T	63	Total	O	0	0
			63	63		

### 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 ( $\text{RSRZ} > 2$ ). 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.

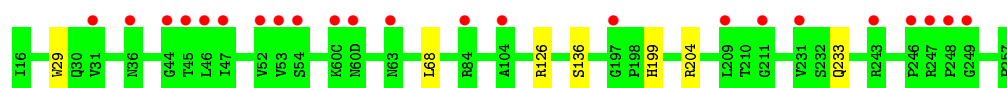
- Molecule 1: Coagulation factor VII

Chain L: 



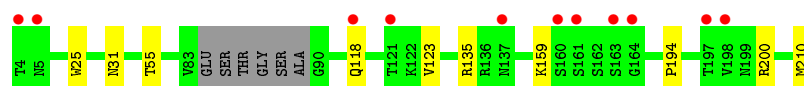
- Molecule 2: Coagulation factor VII

Chain H: 



- Molecule 3: Tissue factor

Chain T: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.03 Å 81.70 Å 124.56 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.51 – 2.25 29.51 – 2.25	Depositor EDS
% Data completeness (in resolution range)	96.6 (29.51-2.25) 96.6 (29.51-2.25)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 2.24 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.188 , 0.240 0.177 , 0.240	Depositor DCC
$R_{free}$ test set	1605 reflections (5.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.9	Xtriage
Anisotropy	0.269	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 48.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8483	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, BGC, CIT, FUC, OZ6

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	L	0.26	0/687	0.48	0/931
2	H	0.28	0/1987	0.47	0/2712
3	T	0.28	0/1625	0.45	0/2216
All	All	0.28	0/4299	0.46	0/5859

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

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	L	675	563	571	9	0
2	H	1927	1870	1876	4	0
3	T	1590	1497	1503	5	0
4	L	11	11	10	0	0
5	L	10	10	10	0	0
6	H	34	36	31	0	0
7	H	1	0	0	0	0
8	T	13	5	5	0	0
9	H	141	0	0	1	0
9	L	26	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	T	63	0	0	0	0
All	All	4491	3992	4006	16	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 2.

All (16) 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:L:96:GLY:O	1:L:109:LYS:N	2.06	0.88
3:T:118:GLN:NE2	3:T:210:MET:SD	2.60	0.74
1:L:104:ASP:CB	1:L:105:HIS:HA	2.21	0.71
1:L:104:ASP:HB3	1:L:105:HIS:HA	1.74	0.69
2:H:204:ARG:NH1	9:H:601:HOH:O	2.27	0.68
3:T:118:GLN:HG2	3:T:123:VAL:HG22	1.80	0.62
2:H:126[A]:ARG:NH2	2:H:233:GLN:OE1	2.34	0.61
1:L:109:LYS:HA	1:L:110:ARG:CB	2.36	0.56
3:T:194:PRO:O	3:T:200:ARG:NH1	2.41	0.53
1:L:109:LYS:CA	1:L:110:ARG:CB	2.95	0.45
3:T:25:TRP:O	3:T:55:THR:HB	2.17	0.44
2:H:136:SER:HB2	2:H:199:HIS:CE1	2.53	0.43
1:L:98:CYS:O	2:H:204:ARG:NH1	2.52	0.43
1:L:72:CYS:O	3:T:135:ARG:NH2	2.47	0.42
1:L:109:LYS:CB	1:L:110:ARG:C	2.89	0.42
1:L:111:SER:O	1:L:111:SER:OG	2.40	0.40

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	L	90/96 (94%)	83 (92%)	7 (8%)	0	100	100
2	H	250/249 (100%)	245 (98%)	5 (2%)	0	100	100
3	T	197/207 (95%)	194 (98%)	3 (2%)	0	100	100
All	All	537/552 (97%)	522 (97%)	15 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	L	73/84 (87%)	71 (97%)	2 (3%)	52	63
2	H	207/210 (99%)	205 (99%)	2 (1%)	82	89
3	T	176/191 (92%)	174 (99%)	2 (1%)	80	88
All	All	456/485 (94%)	450 (99%)	6 (1%)	76	85

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	105	HIS
1	L	140	ILE
2	H	29	TRP
2	H	68	LEU
3	T	31	ASN
3	T	159	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

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$
6	0Z6	H	501	2	34,35,36	2.09	4 (11%)	42,45,46	1.01	2 (4%)
4	BGC	L	201	-	11,11,12	0.78	0	15,15,17	1.20	2 (13%)
5	FUC	L	202	1	10,10,11	0.78	0	13,14,16	0.76	0
8	CIT	T	301	-	3,12,12	1.23	0	3,17,17	2.35	2 (66%)

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
6	0Z6	H	501	2	-	0/35/35/37	0/2/2/2
4	BGC	L	201	-	-	0/2/19/22	0/1/1/1
5	FUC	L	202	1	-	0/0/17/20	0/1/1/1
8	CIT	T	301	-	-	0/6/16/16	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	501	0Z6	C1-N2	5.28	1.46	1.34
6	H	501	0Z6	C-N1	5.29	1.46	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	501	0Z6	CZ2-NH1	5.67	1.43	1.32
6	H	501	0Z6	CZ2-NE	7.04	1.45	1.33

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	T	301	CIT	C3-C2-C1	-3.31	109.77	114.95
8	T	301	CIT	C3-C4-C5	-2.35	111.27	114.95
6	H	501	0Z6	CG2-CB2-CA2	-2.06	109.70	114.03
4	L	201	BGC	C1-C2-C3	-2.05	107.07	109.55
6	H	501	0Z6	CA-C-N1	2.10	122.09	116.12
4	L	201	BGC	C1-O5-C5	2.47	115.77	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	L	94/96 (97%)	0.73	24 (25%) 1 1	30, 55, 98, 106	0
2	H	249/249 (100%)	0.28	23 (9%) 11 12	23, 33, 61, 95	1 (0%)
3	T	201/207 (97%)	0.25	11 (5%) 29 32	29, 46, 87, 109	0
All	All	544/552 (98%)	0.35	58 (10%) 8 8	23, 42, 82, 109	1 (0%)

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	51	ALA	4.8
3	T	163	SER	4.2
3	T	4	THR	4.1
1	L	105	HIS	4.0
2	H	248	PRO	3.9
1	L	106	THR	3.7
3	T	160	SER	3.7
1	L	50	CYS	3.6
2	H	53	VAL	3.6
1	L	122	ALA	3.4
2	H	31	VAL	3.3
1	L	75	ALA	3.2
2	H	46	LEU	3.2
1	L	48	ASP	3.2
2	H	45	THR	3.1
3	T	198	VAL	3.1
1	L	53	SER	3.1
1	L	143	LYS	3.1
1	L	58	GLY	3.1
1	L	111	SER	3.1
1	L	73	LEU	3.0
2	H	209	LEU	3.0
1	L	84	HIS	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	L	110	ARG	2.9
2	H	60(C)	LYS	2.9
2	H	52	VAL	2.8
3	T	161	SER	2.8
3	T	5	ASN	2.8
3	T	164	GLY	2.8
1	L	57	ASN	2.8
2	H	36	ASN	2.8
1	L	49	GLN	2.7
1	L	65	LEU	2.6
2	H	246	PRO	2.5
2	H	211	GLY	2.5
1	L	140	ILE	2.5
3	T	137	ASN	2.4
1	L	76	PHE	2.4
2	H	231	VAL	2.4
1	L	142	GLU	2.4
2	H	104	ALA	2.4
1	L	74	PRO	2.4
2	H	44	GLY	2.4
2	H	60(D)	ASN	2.4
1	L	52	SER	2.3
2	H	54	SER	2.3
3	T	118	GLN	2.3
2	H	84	ARG	2.3
2	H	249	GLY	2.2
2	H	247	ARG	2.2
3	T	197	THR	2.2
2	H	47	ILE	2.2
2	H	63	ASN	2.1
2	H	197	GLY	2.1
1	L	104	ASP	2.1
3	T	121	THR	2.1
1	L	59	GLY	2.0
2	H	243	ARG	2.0

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	FUC	L	202	10/11	0.89	0.29	1.01	61,71,86,86	0
8	CIT	T	301	13/13	0.71	0.21	0.67	87,91,109,110	0
4	BGC	L	201	11/12	0.67	0.28	0.46	93,100,118,120	0
6	OZ6	H	501	34/35	0.90	0.14	0.06	23,34,46,48	0
7	CA	H	502	1/1	0.90	0.10	-1.30	64,64,64,64	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.