



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 10, 2020 – 11:14 AM BST

PDB ID : 4UZL  
Title : STRUCTURE OF THE WNT DEACYLASE NOTUM - CRYSTAL FORM I  
MYRISTOLEATE COMPLEX - 2.1A  
Authors : Zebisch, M.; Jones, E.Y.  
Deposited on : 2014-09-05  
Resolution : 2.10 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

<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

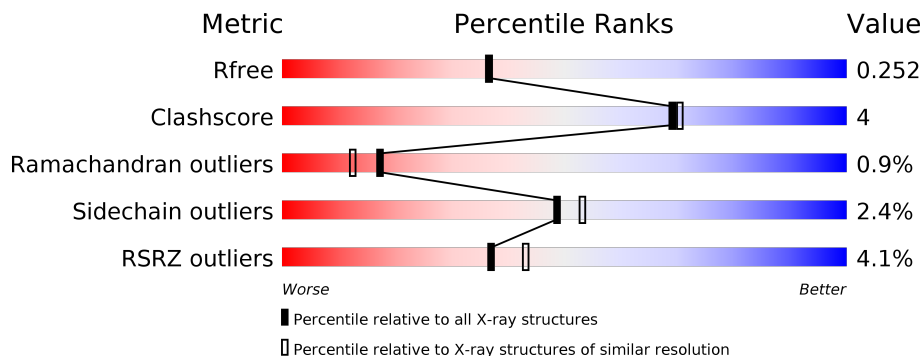
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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	386	
1	B	386	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5626 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 PROTEIN NOTUM HOMOLOG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	339	2754	1748	499	489	18	0	5	0
1	B	346	2781	1761	499	503	18	0	1	0

There are 34 discrepancies between the modelled and reference sequences:

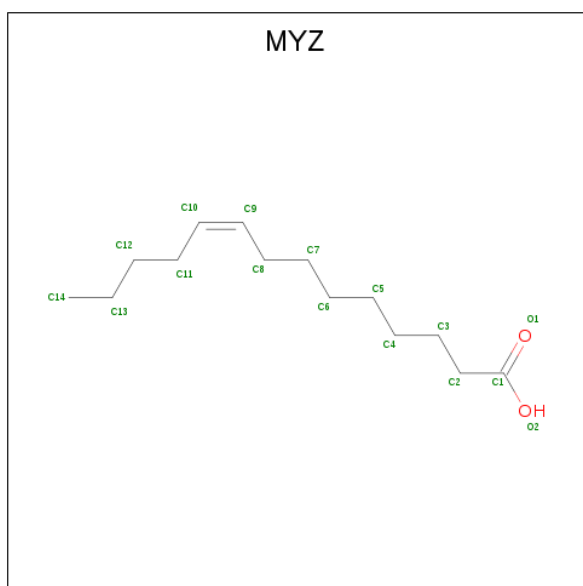
Chain	Residue	Modelled	Actual	Comment	Reference
A	78	GLU	-	expression tag	UNP Q6P988
A	79	THR	-	expression tag	UNP Q6P988
A	80	GLY	-	expression tag	UNP Q6P988
A	452	GLY	-	expression tag	UNP Q6P988
A	453	THR	-	expression tag	UNP Q6P988
A	454	HIS	-	expression tag	UNP Q6P988
A	455	HIS	-	expression tag	UNP Q6P988
A	456	HIS	-	expression tag	UNP Q6P988
A	457	HIS	-	expression tag	UNP Q6P988
A	458	HIS	-	expression tag	UNP Q6P988
A	459	HIS	-	expression tag	UNP Q6P988
A	460	HIS	-	expression tag	UNP Q6P988
A	461	HIS	-	expression tag	UNP Q6P988
A	462	HIS	-	expression tag	UNP Q6P988
A	463	HIS	-	expression tag	UNP Q6P988
A	232	ALA	SER	engineered mutation	UNP Q6P988
A	330	SER	CYS	engineered mutation	UNP Q6P988
B	78	GLU	-	expression tag	UNP Q6P988
B	79	THR	-	expression tag	UNP Q6P988
B	80	GLY	-	expression tag	UNP Q6P988
B	452	GLY	-	expression tag	UNP Q6P988
B	453	THR	-	expression tag	UNP Q6P988
B	454	HIS	-	expression tag	UNP Q6P988
B	455	HIS	-	expression tag	UNP Q6P988
B	456	HIS	-	expression tag	UNP Q6P988

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	457	HIS	-	expression tag	UNP Q6P988
B	458	HIS	-	expression tag	UNP Q6P988
B	459	HIS	-	expression tag	UNP Q6P988
B	460	HIS	-	expression tag	UNP Q6P988
B	461	HIS	-	expression tag	UNP Q6P988
B	462	HIS	-	expression tag	UNP Q6P988
B	463	HIS	-	expression tag	UNP Q6P988
B	232	ALA	SER	engineered mutation	UNP Q6P988
B	330	SER	CYS	engineered mutation	UNP Q6P988

- Molecule 2 is Myristoleic acid (three-letter code: MYZ) (formula: C<sub>14</sub>H<sub>26</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			16	14	2		
2	B	1	Total	C	O	0	0
			16	14	2		

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	B	1	14	8	1	5	0	0

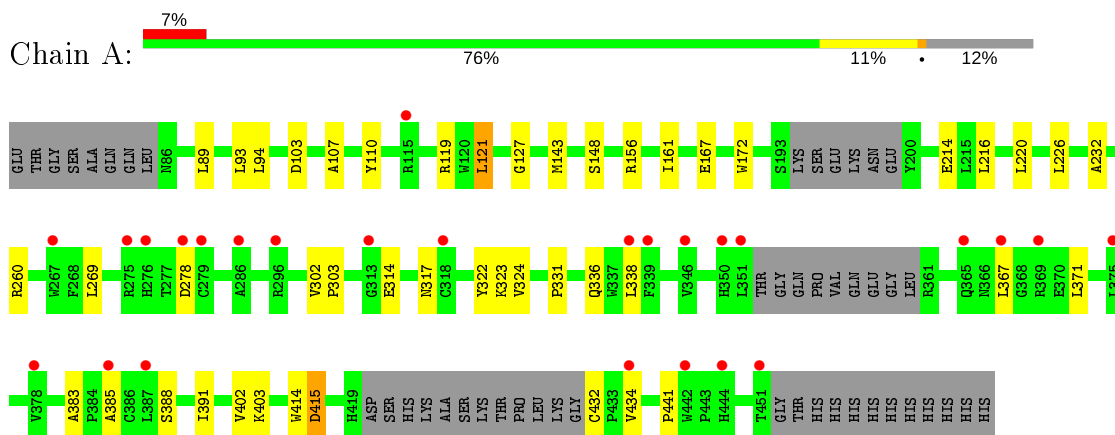
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	15	Total	O	0	0
			15	15		
4	B	30	Total	O	0	0
			30	30		

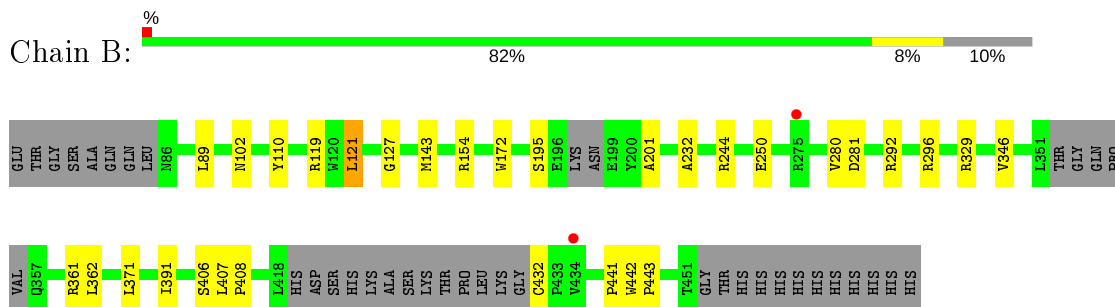
### 3 Residue-property plots [i](#)

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: PROTEIN NOTUM HOMOLOG



- Molecule 1: PROTEIN NOTUM HOMOLOG



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.12Å 170.72Å 61.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	85.36 – 2.10 67.99 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (85.36-2.10) 98.9 (67.99-2.10)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.214 , 0.253 0.220 , 0.252	Depositor DCC
$R_{free}$ test set	1115 reflections (2.43%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.9	Xtrriage
Anisotropy	0.492	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 42.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5626	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MYZ, 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	0.53	0/2845	0.66	0/3868
1	B	0.57	0/2859	0.69	0/3888
All	All	0.55	0/5704	0.68	0/7756

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	2754	0	2664	25	2
1	B	2781	0	2668	17	2
2	A	16	0	25	1	0
2	B	16	0	25	3	0
3	B	14	0	13	0	0
4	A	15	0	0	1	0
4	B	30	0	0	2	0
All	All	5626	0	5395	41	2

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 4.



All (41) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:119[A]:ARG:HG2	1:A:172[A]:TRP:CZ3	2.06	0.88
1:A:93:LEU:HD11	1:B:154:ARG:HD2	1.58	0.84
1:B:232:ALA:HB1	2:B:1453:MYZ:C1	2.24	0.68
1:A:119[A]:ARG:CG	1:A:172[A]:TRP:CZ3	2.80	0.64
1:A:232:ALA:HB1	2:A:1452:MYZ:C1	2.29	0.63
1:A:119[A]:ARG:HG2	1:A:172[A]:TRP:HZ3	1.65	0.60
1:A:119[B]:ARG:NH2	1:A:415:ASP:OD2	2.36	0.57
1:A:103:ASP:OD2	1:A:148:SER:HB2	2.05	0.56
1:B:250:GLU:OE2	1:B:329:ARG:NH2	2.39	0.56
1:A:161:ILE:HD11	1:A:172[B]:TRP:CE3	2.43	0.53
1:B:361:ARG:HG3	1:B:442:TRP:CH2	2.45	0.51
1:B:143:MET:HE3	4:B:2008:HOH:O	2.12	0.50
1:B:89:LEU:HB3	1:B:110:TYR:HB3	1.95	0.48
1:B:121:LEU:HG	1:B:172[B]:TRP:CZ3	2.48	0.47
1:B:244:ARG:HD2	4:B:2015:HOH:O	2.14	0.47
1:A:338:LEU:HD12	1:A:385:ALA:HB2	1.97	0.47
1:A:317:ASN:HB3	1:A:323:LYS:HD3	1.97	0.47
1:B:119:ARG:HG2	1:B:172[B]:TRP:CZ3	2.49	0.46
1:A:269:LEU:HD12	1:A:367:LEU:HD12	1.97	0.45
1:A:260:ARG:HG2	1:A:331:PRO:HG2	1.98	0.45
1:A:302:VAL:HB	1:A:303:PRO:CD	2.46	0.45
1:B:232:ALA:HB1	2:B:1453:MYZ:O2	2.16	0.45
1:A:414:TRP:HB2	1:A:434:VAL:HG23	1.98	0.45
1:A:143:MET:HE3	4:A:2003:HOH:O	2.16	0.44
1:B:119:ARG:NH1	1:B:172[B]:TRP:CZ2	2.84	0.44
1:A:314:GLU:O	1:A:317:ASN:ND2	2.51	0.43
1:A:103:ASP:OD2	1:A:148:SER:CB	2.66	0.43
1:A:226:LEU:C	1:A:226:LEU:HD23	2.38	0.43
1:A:336:GLN:O	1:A:383:ALA:HA	2.19	0.43
1:B:102:ASN:HA	1:B:201:ALA:O	2.18	0.43
1:B:407:LEU:N	1:B:408:PRO:CD	2.82	0.43
1:A:94:LEU:HA	1:A:214:GLU:OE2	2.20	0.42
1:A:402:VAL:HG12	1:A:403:LYS:HG3	2.01	0.42
1:B:442:TRP:CD2	1:B:443:PRO:HD2	2.55	0.42
1:B:280:VAL:HG23	1:B:281:ASP:N	2.35	0.42
1:A:89:LEU:HB3	1:A:110:TYR:HB3	2.02	0.41
1:B:407:LEU:HB3	1:B:408:PRO:HD3	2.03	0.41
1:B:346:VAL:HG11	2:B:1453:MYZ:H131	2.03	0.41
1:A:94:LEU:HD12	1:A:107:ALA:HB3	2.03	0.40
1:A:121:LEU:HD11	1:A:161:ILE:HG12	2.04	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:216:LEU:HA	1:A:220:LEU:HB3	2.03	0.40

All (2) 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:167:GLU:OE2	1:B:406:SER:OG[1_554]	2.10	0.10
1:A:278:ASP:OD1	1:B:296:ARG:NH1[1_655]	2.15	0.05

### 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	336/386 (87%)	315 (94%)	18 (5%)	3 (1%)	17	12
1	B	339/386 (88%)	326 (96%)	10 (3%)	3 (1%)	17	12
All	All	675/772 (87%)	641 (95%)	28 (4%)	6 (1%)	17	12

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	391	ILE
1	A	127	GLY
1	A	391	ILE
1	A	441	PRO
1	B	127	GLY
1	B	441	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	298/335 (89%)	289 (97%)	9 (3%)	41	44
1	B	300/335 (90%)	294 (98%)	6 (2%)	55	60
All	All	598/670 (89%)	583 (98%)	15 (2%)	49	52

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	121	LEU
1	A	156[A]	ARG
1	A	156[B]	ARG
1	A	322	TYR
1	A	324	VAL
1	A	371	LEU
1	A	388	SER
1	A	415	ASP
1	A	432	CYS
1	B	121	LEU
1	B	195	SER
1	B	292	ARG
1	B	362	LEU
1	B	371	LEU
1	B	432	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	276	HIS
1	A	309	GLN
1	A	435	HIS
1	B	135	ASN
1	B	248	GLN
1	B	258	GLN
1	B	366	ASN

### 5.3.3 RNA

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](#)

3 ligands 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	$\# Z  > 2$
2	MYZ	A	1452	-	12,15,15	0.24	0	11,15,15	0.70	0
2	MYZ	B	1453	-	12,15,15	0.23	0	11,15,15	0.57	0
3	NAG	B	1452	1	14,14,15	0.65	0	17,19,21	1.45	4 (23%)

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
2	MYZ	A	1452	-	-	6/11/13/13	-
2	MYZ	B	1453	-	-	3/11/13/13	-
3	NAG	B	1452	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1452	NAG	C4-C3-C2	3.78	116.55	111.02
3	B	1452	NAG	O7-C7-C8	-2.21	117.96	122.06

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1452	NAG	C3-C4-C5	2.06	113.91	110.24
3	B	1452	NAG	O5-C5-C6	2.00	110.34	107.20

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1452	MYZ	C10-C11-C12-C13
2	B	1453	MYZ	C3-C4-C5-C6
2	B	1453	MYZ	C5-C6-C7-C8
2	B	1453	MYZ	C10-C11-C12-C13
2	A	1452	MYZ	C6-C7-C8-C9
2	A	1452	MYZ	C2-C3-C4-C5
2	A	1452	MYZ	C7-C8-C9-C10
2	A	1452	MYZ	C4-C5-C6-C7
2	A	1452	MYZ	C1-C2-C3-C4

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1452	MYZ	1	0
2	B	1453	MYZ	3	0

## 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	A	339/386 (87%)	0.46	26 (7%) <span style="border: 1px solid red; padding: 2px;">13</span> <span style="border: 1px solid red; padding: 2px;">17</span>	34, 70, 134, 161	0
1	B	346/386 (89%)	0.03	2 (0%) <span style="border: 1px solid blue; padding: 2px;">89</span> <span style="border: 1px solid blue; padding: 2px;">91</span>	32, 53, 82, 101	0
All	All	685/772 (88%)	0.24	28 (4%) <span style="border: 1px solid red; padding: 2px;">37</span> <span style="border: 1px solid red; padding: 2px;">43</span>	32, 59, 115, 161	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	367	LEU	7.6
1	A	339	PHE	7.5
1	A	313	GLY	4.4
1	A	286	ALA	3.7
1	A	275	ARG	3.5
1	A	434	VAL	3.2
1	A	387	LEU	3.1
1	A	444	HIS	3.1
1	A	375	LEU	3.1
1	A	338	LEU	3.0
1	A	369	ARG	2.9
1	A	351	LEU	2.8
1	B	275	ARG	2.8
1	A	279	CYS	2.6
1	A	267	TRP	2.5
1	A	350	HIS	2.5
1	A	442	TRP	2.4
1	A	346	VAL	2.4
1	A	276	HIS	2.3
1	A	278	ASP	2.3
1	A	451	THR	2.3
1	B	434	VAL	2.3
1	A	385	ALA	2.2
1	A	115	ARG	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	296	ARG	2.1
1	A	318	CYS	2.1
1	A	378	VAL	2.1
1	A	365	GLN	2.0

## 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](#)

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, 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	B	1452	14/15	0.85	0.15	85,92,96,96	0
2	MYZ	A	1452	16/16	0.91	0.20	47,68,78,82	0
2	MYZ	B	1453	16/16	0.93	0.14	35,54,62,62	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.