



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 07:10 pm GMT

PDB ID : 3VNJ  
Title : Crystal structures of D-Psicose 3-epimerase with D-psicose from *Clostridium cellulolyticum* H10  
Authors : Chan, H.C.; Zhu, Y.; Hu, Y.; Ko, T.P.; Huang, C.H.; Ren, F.; Chen, C.C.; Guo, R.T.; Sun, Y.  
Deposited on : 2012-01-16  
Resolution : 2.08 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](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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

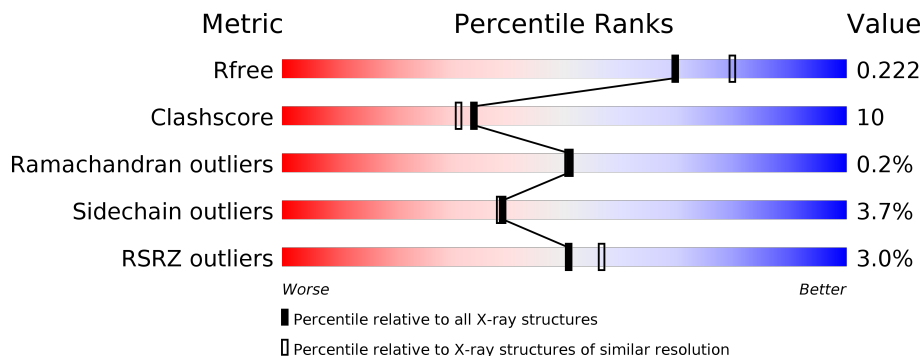
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.08 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))
$R_{free}$	100719	4955 (2.10-2.06)
Clashscore	112137	5537 (2.10-2.06)
Ramachandran outliers	110173	5483 (2.10-2.06)
Sidechain outliers	110143	5484 (2.10-2.06)
RSRZ outliers	101464	4991 (2.10-2.06)

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	A	294	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 15%, green 82%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>82%</span> <span>15%</span> </div> </div>
1	B	294	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 2%, yellow 17%, green 80%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>80%</span> <span>17%</span> </div> </div>
1	C	294	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, orange 3%, yellow 20%, green 77%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>77%</span> <span>20%</span> </div> </div>
1	D	294	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 5%, orange 5%, yellow 22%, green 73%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>73%</span> <span>22%</span> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10049 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.

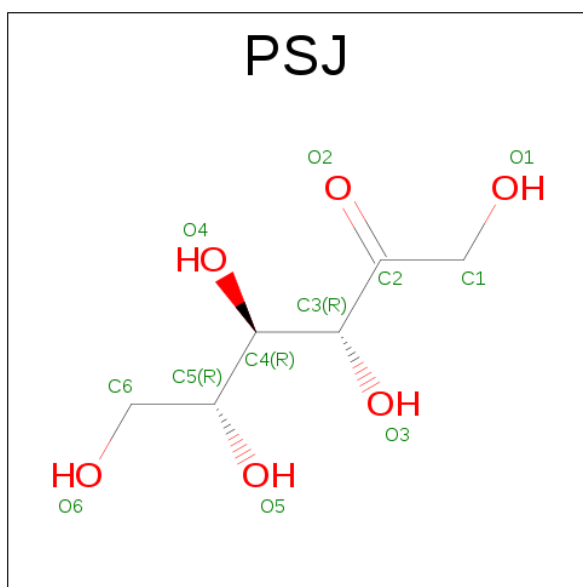
- Molecule 1 is a protein called Xylose isomerase domain protein TIM barrel.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	290	Total	C	N	O	S	0	0	0
			2298	1462	389	436	11			
1	B	288	Total	C	N	O	S	0	0	0
			2287	1456	387	434	10			
1	C	288	Total	C	N	O	S	0	0	0
			2287	1456	387	434	10			
1	D	289	Total	C	N	O	S	0	0	0
			2293	1459	388	435	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ALA	-	EXPRESSION TAG	UNP B8I944
B	0	ALA	-	EXPRESSION TAG	UNP B8I944
C	0	ALA	-	EXPRESSION TAG	UNP B8I944
D	0	ALA	-	EXPRESSION TAG	UNP B8I944

- Molecule 2 is SUGAR (D-PSICOSE) (three-letter code: PSJ) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		
2	C	1	Total	C	O	0	0
			12	6	6		
2	D	1	Total	C	O	0	0
			12	6	6		

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mn	0	0
			1	1		
3	A	1	Total	Mn	0	0
			1	1		
3	D	1	Total	Mn	0	0
			1	1		
3	C	1	Total	Mn	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	247	Total	O	0	0
			247	247		

*Continued on next page...*

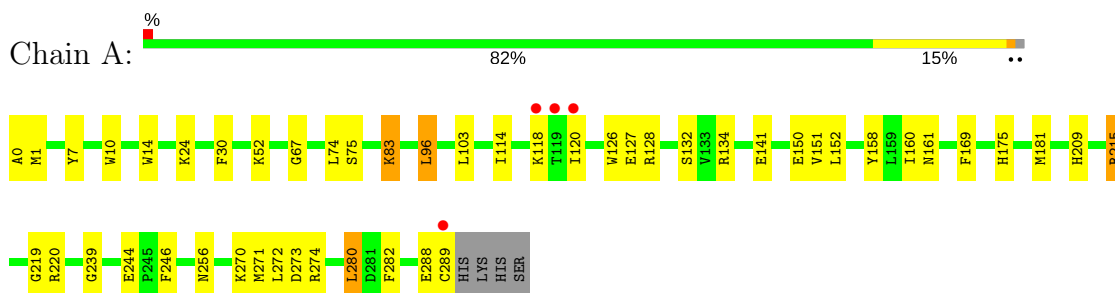
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	221	Total 221	O 221	0	0
4	C	191	Total 191	O 191	0	0
4	D	173	Total 173	O 173	0	0

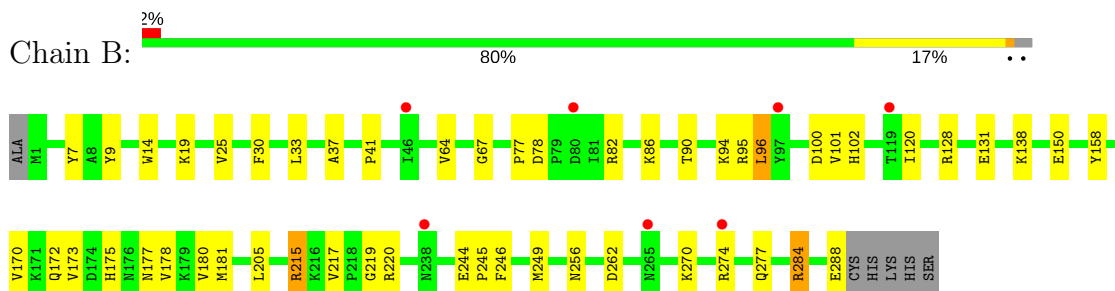
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

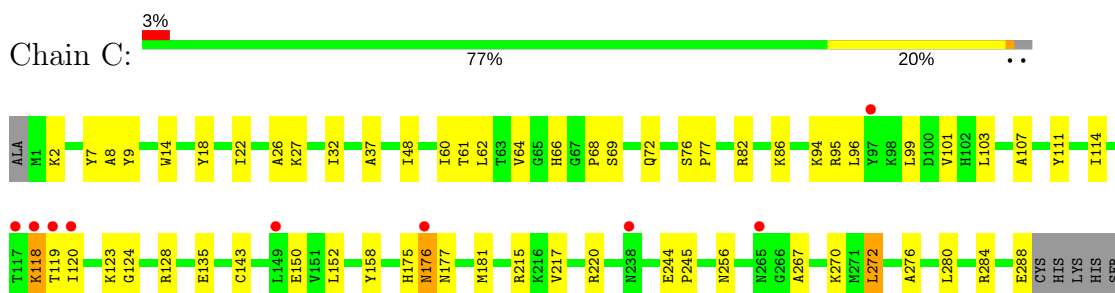
- Molecule 1: Xylose isomerase domain protein TIM barrel



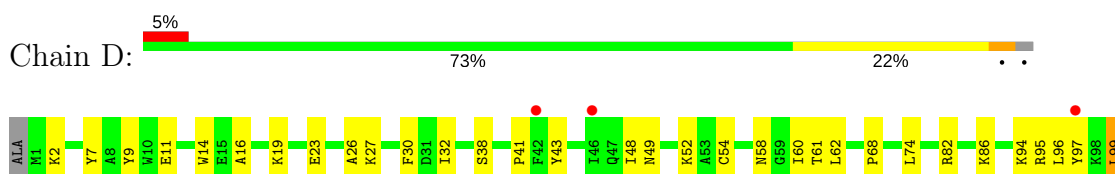
- Molecule 1: Xylose isomerase domain protein TIM barrel

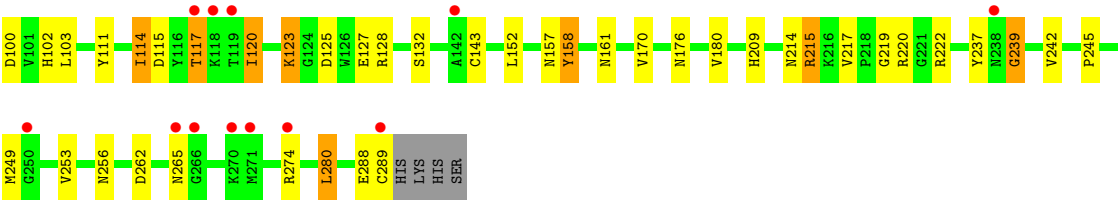


- Molecule 1: Xylose isomerase domain protein TIM barrel



- Molecule 1: Xylose isomerase domain protein TIM barrel





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.88Å 115.50Å 91.58Å 90.00° 105.16° 90.00°	Depositor
Resolution (Å)	25.00 – 2.08 24.78 – 2.08	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.08) 92.6 (24.78-2.08)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.41 (at 2.08Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.182 , 0.220 0.183 , 0.222	Depositor DCC
$R_{free}$ test set	4494 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.2	Xtriage
Anisotropy	0.328	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 55.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10049	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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: PSJ, MN

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	0.77	0/2351	0.84	3/3184 (0.1%)
1	B	0.75	0/2340	0.82	4/3169 (0.1%)
1	C	0.68	0/2340	0.84	1/3169 (0.0%)
1	D	0.67	0/2346	0.80	4/3177 (0.1%)
All	All	0.72	0/9377	0.82	12/12699 (0.1%)

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
1	D	0	2
All	All	0	3

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	215	ARG	NE-CZ-NH2	-7.83	116.38	120.30
1	D	215	ARG	NE-CZ-NH2	-6.54	117.03	120.30
1	D	100	ASP	CB-CG-OD1	6.13	123.81	118.30
1	B	284	ARG	NE-CZ-NH1	-5.84	117.38	120.30
1	A	215	ARG	NE-CZ-NH2	-5.73	117.44	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	7	TYR	Sidechain
1	D	158	TYR	Sidechain
1	D	7	TYR	Sidechain

## 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	2298	0	2233	35	0
1	B	2287	0	2223	40	0
1	C	2287	0	2223	49	0
1	D	2293	0	2228	54	0
2	A	12	0	12	2	0
2	B	12	0	11	1	0
2	C	12	0	11	1	0
2	D	12	0	11	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	247	0	0	6	0
4	B	221	0	0	5	0
4	C	191	0	0	3	0
4	D	173	0	0	6	0
All	All	10049	0	8952	174	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 10.

The worst 5 of 174 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:D:114:ILE:HD11	1:D:157:ASN:HA	1.12	1.11
1:C:220:ARG:NH2	1:D:288:GLU:O	1.92	1.03
1:D:115:ASP:OD2	1:D:117:THR:HB	1.67	0.93
1:B:86:LYS:O	1:B:90:THR:HG23	1.74	0.87
1:D:23:GLU:O	1:D:27:LYS:HG2	1.77	0.83

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	288/294 (98%)	280 (97%)	7 (2%)	1 (0%)	44	43
1	B	286/294 (97%)	278 (97%)	8 (3%)	0	100	100
1	C	286/294 (97%)	275 (96%)	11 (4%)	0	100	100
1	D	287/294 (98%)	273 (95%)	13 (4%)	1 (0%)	44	43
All	All	1147/1176 (98%)	1106 (96%)	39 (3%)	2 (0%)	51	50

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	239	GLY
1	D	239	GLY

### 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	241/245 (98%)	232 (96%)	9 (4%)	39	38
1	B	240/245 (98%)	233 (97%)	7 (3%)	48	50
1	C	240/245 (98%)	233 (97%)	7 (3%)	48	50
1	D	241/245 (98%)	228 (95%)	13 (5%)	26	22
All	All	962/980 (98%)	926 (96%)	36 (4%)	39	38

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	27	LYS
1	C	176	ASN
1	D	265	ASN
1	C	118	LYS
1	C	270	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	84	ASN
1	C	175	HIS
1	D	12	GLN
1	C	66	HIS
1	D	84	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are 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
2	PSJ	A	301	3	11,11,11	0.54	0	10,14,14	1.35	2 (20%)
2	PSJ	B	301	3	11,11,11	0.55	0	10,14,14	1.28	2 (20%)
2	PSJ	C	301	3	11,11,11	0.60	0	10,14,14	0.99	1 (10%)
2	PSJ	D	301	3	11,11,11	0.49	0	10,14,14	1.45	3 (30%)

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	PSJ	A	301	3	-	0/16/16/16	0/0/0/0
2	PSJ	B	301	3	-	0/16/16/16	0/0/0/0
2	PSJ	C	301	3	-	0/16/16/16	0/0/0/0
2	PSJ	D	301	3	-	0/16/16/16	0/0/0/0

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	301	PSJ	C6-C5-C4	-2.78	106.22	112.41
2	A	301	PSJ	C6-C5-C4	-2.78	106.22	112.41
2	B	301	PSJ	C6-C5-C4	-2.75	106.28	112.41
2	D	301	PSJ	O3-C3-C2	-2.37	106.81	110.95
2	C	301	PSJ	C6-C5-C4	-2.32	107.24	112.41

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	PSJ	2	0
2	B	301	PSJ	1	0
2	C	301	PSJ	1	0

## 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, 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	290/294 (98%)	-0.29	4 (1%) 75 78	24, 33, 50, 69	0
1	B	288/294 (97%)	-0.13	7 (2%) 59 63	23, 38, 55, 71	0
1	C	288/294 (97%)	0.01	9 (3%) 49 55	28, 42, 57, 74	0
1	D	289/294 (98%)	0.06	15 (5%) 28 33	27, 42, 63, 70	0
All	All	1155/1176 (98%)	-0.09	35 (3%) 51 56	23, 39, 58, 74	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	119	THR	6.7
1	C	238	ASN	4.9
1	B	119	THR	4.8
1	A	119	THR	4.6
1	D	238	ASN	4.4

### 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](#)

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
2	PSJ	A	301	12/12	0.95	0.14	0.67	36,41,45,46	0
2	PSJ	D	301	12/12	0.96	0.15	0.41	34,41,44,44	0
2	PSJ	C	301	12/12	0.95	0.13	0.04	33,42,51,51	0
2	PSJ	B	301	12/12	0.95	0.13	-0.06	35,42,47,48	0
3	MN	D	302	1/1	0.99	0.12	-0.64	34,34,34,34	0
3	MN	A	302	1/1	1.00	0.11	-0.67	27,27,27,27	0
3	MN	B	302	1/1	1.00	0.07	-2.19	27,27,27,27	0
3	MN	C	302	1/1	1.00	0.08	-2.43	31,31,31,31	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.