



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2016 – 10:47 AM EDT

PDB ID : 5EDU  
Title : Crystal structure of human histone deacetylase 6 catalytic domain 2 in complex with trichostatin A  
Authors : Hai, Y.; Christianson, D.  
Deposited on : 2015-10-22  
Resolution : 2.79 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
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-20027939

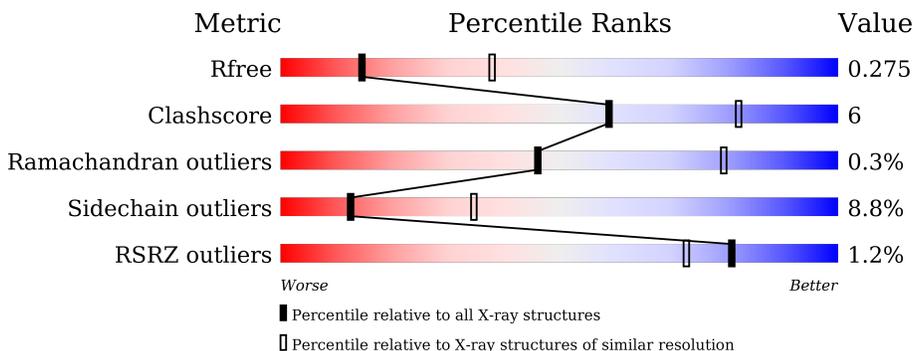
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 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	745	
1	B	745	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
2	TSN	A	904	-	-	-	X
2	TSN	B	2501	-	-	-	X
4	K	A	902	-	-	-	X
4	K	B	2503	-	-	-	X

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 11179 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 Maltose-binding periplasmic protein, Histone deacetylase 6 chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	723	5542	3521	952	1043	26	0	1	0
1	A	720	5483	3486	946	1025	26	0	1	0

There are 50 discrepancies between the modelled and reference sequences:

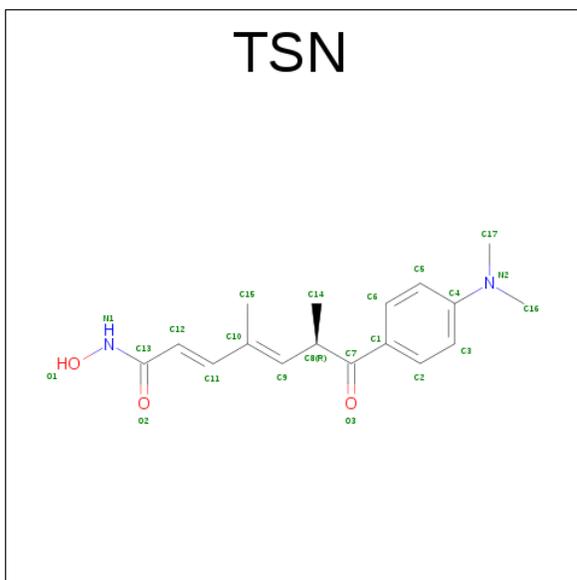
Chain	Residue	Modelled	Actual	Comment	Reference
B	91	MET	-	expression tag	UNP P0AEY0
B	92	GLY	-	expression tag	UNP P0AEY0
B	93	SER	-	expression tag	UNP P0AEY0
B	94	SER	-	expression tag	UNP P0AEY0
B	95	HIS	-	expression tag	UNP P0AEY0
B	96	HIS	-	expression tag	UNP P0AEY0
B	97	HIS	-	expression tag	UNP P0AEY0
B	98	HIS	-	expression tag	UNP P0AEY0
B	99	HIS	-	expression tag	UNP P0AEY0
B	100	HIS	-	expression tag	UNP P0AEY0
B	101	GLU	-	expression tag	UNP P0AEY0
B	102	ASN	-	expression tag	UNP P0AEY0
B	103	LEU	-	expression tag	UNP P0AEY0
B	104	TYR	-	expression tag	UNP P0AEY0
B	105	PHE	-	expression tag	UNP P0AEY0
B	106	GLN	-	expression tag	UNP P0AEY0
B	107	GLY	-	expression tag	UNP P0AEY0
B	108	SER	-	expression tag	UNP P0AEY0
B	467	ALA	GLU	conflict	UNP P0AEY0
B	470	ALA	LYS	conflict	UNP P0AEY0
B	471	ALA	ASP	conflict	UNP P0AEY0
B	475	ASN	-	linker	UNP P0AEY0
B	476	ALA	-	linker	UNP P0AEY0
B	477	ALA	-	linker	UNP P0AEY0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	478	ALA	-	linker	UNP P0AEY0
A	91	MET	-	expression tag	UNP P0AEY0
A	92	GLY	-	expression tag	UNP P0AEY0
A	93	SER	-	expression tag	UNP P0AEY0
A	94	SER	-	expression tag	UNP P0AEY0
A	95	HIS	-	expression tag	UNP P0AEY0
A	96	HIS	-	expression tag	UNP P0AEY0
A	97	HIS	-	expression tag	UNP P0AEY0
A	98	HIS	-	expression tag	UNP P0AEY0
A	99	HIS	-	expression tag	UNP P0AEY0
A	100	HIS	-	expression tag	UNP P0AEY0
A	101	GLU	-	expression tag	UNP P0AEY0
A	102	ASN	-	expression tag	UNP P0AEY0
A	103	LEU	-	expression tag	UNP P0AEY0
A	104	TYR	-	expression tag	UNP P0AEY0
A	105	PHE	-	expression tag	UNP P0AEY0
A	106	GLN	-	expression tag	UNP P0AEY0
A	107	GLY	-	expression tag	UNP P0AEY0
A	108	SER	-	expression tag	UNP P0AEY0
A	467	ALA	GLU	conflict	UNP P0AEY0
A	470	ALA	LYS	conflict	UNP P0AEY0
A	471	ALA	ASP	conflict	UNP P0AEY0
A	475	ASN	-	linker	UNP P0AEY0
A	476	ALA	-	linker	UNP P0AEY0
A	477	ALA	-	linker	UNP P0AEY0
A	478	ALA	-	linker	UNP P0AEY0

- Molecule 2 is TRICHOSTATIN A (three-letter code: TSN) (formula: C<sub>17</sub>H<sub>22</sub>N<sub>2</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
2	B	1	Total	C	N	O	0	0
			22	17	2	3		
2	A	1	Total	C	N	O	0	0
			22	17	2	3		

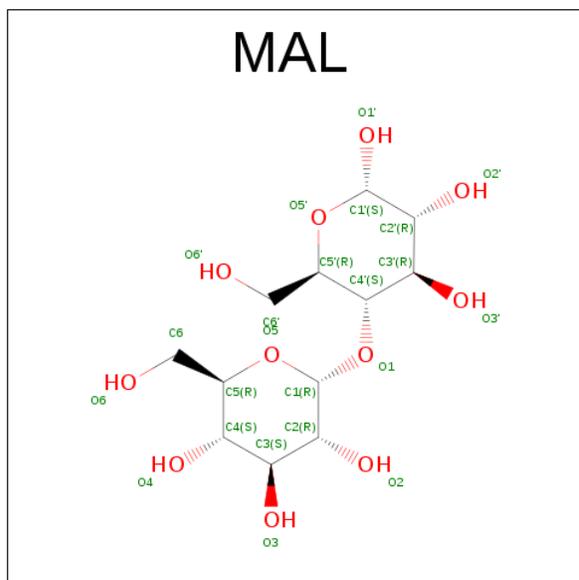
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	K	0	0
			2	2		
4	A	2	Total	K	0	0
			2	2		

- Molecule 5 is MALTOSE (three-letter code: MAL) (formula: C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	C O	0	0
			23	12 11		
5	A	1	Total	C O	0	0
			23	12 11		

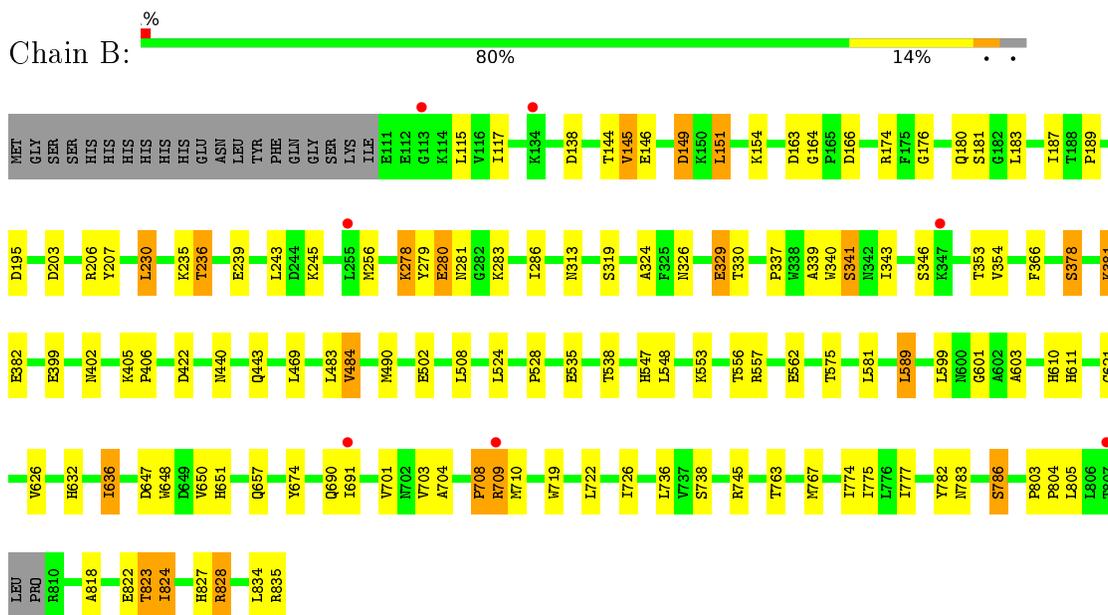
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	29	Total	O	0	0
			29	29		
6	A	29	Total	O	0	0
			29	29		

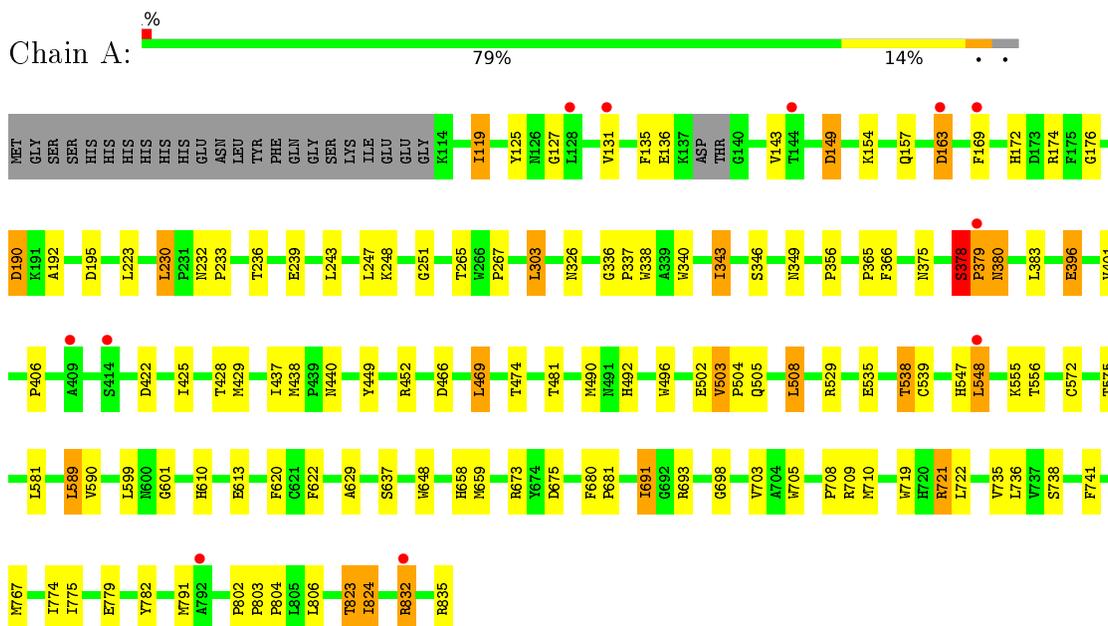
### 3 Residue-property plots [i](#)

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 ( $RSRZ > 2$ ). 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: Maltose-binding periplasmic protein, Histone deacetylase 6 chimera



- Molecule 1: Maltose-binding periplasmic protein, Histone deacetylase 6 chimera



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.26Å 149.03Å 216.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.54 – 2.79 122.73 – 2.79	Depositor EDS
% Data completeness (in resolution range)	100.0 (87.54-2.79) 100.0 (122.73-2.79)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 2.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.3_1479)	Depositor
R, $R_{free}$	0.213 , 0.275 0.215 , 0.275	Depositor DCC
$R_{free}$ test set	2055 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	55.6	Xtriage
Anisotropy	0.366	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	11179	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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: TSN, ZN, K, MAL

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	0/5624	0.43	3/7665 (0.0%)
1	B	0.27	0/5683	0.41	1/7734 (0.0%)
All	All	0.26	0/11307	0.42	4/15399 (0.0%)

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

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	378	SER	C-N-CD	5.94	140.87	128.40
1	A	378	SER	C-N-CD	5.62	140.20	128.40
1	A	452	ARG	NE-CZ-NH1	-5.21	117.69	120.30
1	A	379	PRO	CA-N-CD	-5.19	104.24	111.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	378	SER	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	5483	0	5314	65	0
1	B	5542	0	5398	62	0
2	A	22	0	21	3	0
2	B	22	0	21	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	A	23	0	22	2	0
5	B	23	0	22	1	0
6	A	29	0	0	0	0
6	B	29	0	0	0	0
All	All	11179	0	10798	127	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 6.

The worst 5 of 127 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:379:PRO:O	1:A:380:ASN:ND2	1.97	0.98
1:A:378:SER:OG	1:A:379:PRO:HA	1.78	0.84
1:B:378:SER:O	1:B:381:LYS:HE2	1.78	0.83
1:B:187:ILE:HG22	1:B:189:PRO:HD3	1.65	0.78
1:B:230:LEU:HD11	1:B:243:LEU:HD21	1.72	0.72

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	717/745 (96%)	680 (95%)	33 (5%)	4 (1%)	30	65
1	B	720/745 (97%)	686 (95%)	33 (5%)	1 (0%)	56	87
All	All	1437/1490 (96%)	1366 (95%)	66 (5%)	5 (0%)	46	79

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	136	GLU
1	A	378	SER
1	A	375	ASN
1	A	708	PRO
1	B	708	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	559/595 (94%)	509 (91%)	50 (9%)	12	34
1	B	571/595 (96%)	521 (91%)	50 (9%)	12	35
All	All	1130/1190 (95%)	1030 (91%)	100 (9%)	12	35

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	786	SER
1	A	190	ASP
1	A	710	MET
1	B	805	LEU
1	A	119	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	194	GLN
1	A	463	GLN
1	A	380	ASN
1	A	180	GLN
1	A	457	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 10 ligands modelled in this entry, 6 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	TSN	A	904	3	22,22,22	0.71	1 (4%)	24,29,29	1.37	1 (4%)
5	MAL	A	905	-	24,24,24	0.48	0	35,35,35	0.72	0
2	TSN	B	2501	3	22,22,22	0.66	1 (4%)	24,29,29	1.04	1 (4%)
5	MAL	B	2505	-	24,24,24	0.51	0	35,35,35	0.51	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
2	TSN	A	904	3	-	0/23/23/23	0/1/1/1
5	MAL	A	905	-	-	0/8/48/48	0/2/2/2
2	TSN	B	2501	3	-	0/23/23/23	0/1/1/1
5	MAL	B	2505	-	-	0/8/48/48	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	904	TSN	C8-C7	-2.63	1.50	1.53
2	B	2501	TSN	C8-C7	-2.39	1.50	1.53

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	904	TSN	C12-C11-C10	-4.63	119.21	126.21
2	B	2501	TSN	C12-C11-C10	-2.54	122.36	126.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	904	TSN	3	0
5	A	905	MAL	2	0
2	B	2501	TSN	3	0
5	B	2505	MAL	1	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 [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	720/745 (96%)	0.40	11 (1%) 76 68	26, 50, 87, 125	0
1	B	723/745 (97%)	0.35	7 (0%) 84 77	30, 52, 83, 126	0
All	All	1443/1490 (96%)	0.38	18 (1%) 81 73	26, 51, 85, 126	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	807	THR	3.1
1	A	144	THR	3.0
1	A	131	VAL	3.0
1	B	347	LYS	2.6
1	A	409	ALA	2.5

### 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 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	TSN	A	904	22/22	0.89	0.30	7.25	52,56,64,65	0
4	K	B	2503	1/1	0.79	0.34	6.53	105,105,105,105	0
4	K	A	902	1/1	0.97	0.27	4.43	68,68,68,68	0
2	TSN	B	2501	22/22	0.91	0.27	2.11	58,61,73,75	0
4	K	A	903	1/1	0.91	0.23	1.51	82,82,82,82	0
5	MAL	A	905	23/23	0.84	0.26	1.42	62,66,69,73	0
5	MAL	B	2505	23/23	0.91	0.20	-0.25	51,54,56,56	0
3	ZN	B	2502	1/1	0.99	0.13	-2.41	62,62,62,62	0
4	K	B	2504	1/1	0.63	0.16	-3.09	116,116,116,116	0
3	ZN	A	901	1/1	0.99	0.14	-4.08	41,41,41,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.