



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 8, 2024 – 05:33 AM EDT

PDB ID : 3U1N  
Title : Structure of the catalytic core of human SAMHD1  
Authors : Goldstone, D.C.; Ennis-Adeniran, V.; Walker, P.A.; Haire, L.F.; Webb, M.; Taylor, I.A.  
Deposited on : 2011-09-30  
Resolution : 3.10 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

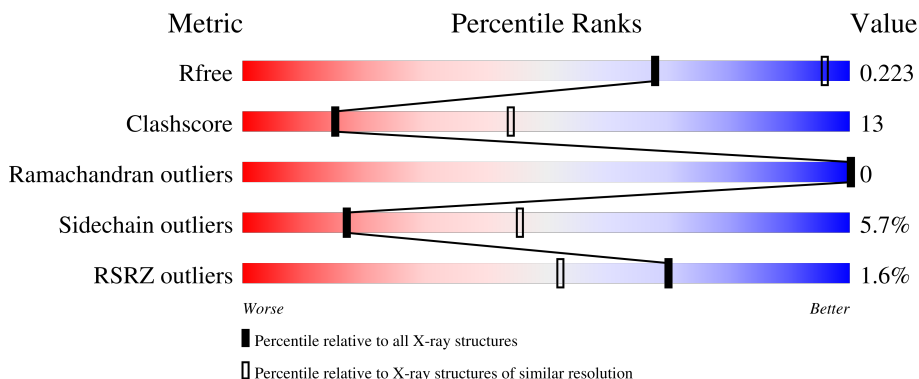
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	164625	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)

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 of 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	528	
1	B	528	
1	C	528	
1	D	528	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13595 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 SAM domain and HD domain-containing protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	435	Total	C	N	O	S	Se	0	0	0
			3423	2200	590	616	9	8			
1	B	434	Total	C	N	O	S	Se	0	0	0
			3423	2199	597	609	10	8			
1	C	427	Total	C	N	O	S	Se	0	0	0
			3353	2153	576	606	10	8			
1	D	424	Total	C	N	O	S	Se	0	0	0
			3344	2152	578	597	9	8			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	99	ALA	-	expression tag	UNP Q9Y3Z3
A	100	SER	-	expression tag	UNP Q9Y3Z3
A	101	TRP	-	expression tag	UNP Q9Y3Z3
A	102	SER	-	expression tag	UNP Q9Y3Z3
A	103	HIS	-	expression tag	UNP Q9Y3Z3
A	104	PRO	-	expression tag	UNP Q9Y3Z3
A	105	GLN	-	expression tag	UNP Q9Y3Z3
A	106	PHE	-	expression tag	UNP Q9Y3Z3
A	107	GLU	-	expression tag	UNP Q9Y3Z3
A	108	LYS	-	expression tag	UNP Q9Y3Z3
A	109	GLY	-	expression tag	UNP Q9Y3Z3
A	110	ALA	-	expression tag	UNP Q9Y3Z3
A	111	LEU	-	expression tag	UNP Q9Y3Z3
A	112	GLU	-	expression tag	UNP Q9Y3Z3
A	113	VAL	-	expression tag	UNP Q9Y3Z3
A	114	LEU	-	expression tag	UNP Q9Y3Z3
A	115	PHE	-	expression tag	UNP Q9Y3Z3
A	116	GLN	-	expression tag	UNP Q9Y3Z3
A	117	GLY	-	expression tag	UNP Q9Y3Z3
A	118	PRO	-	expression tag	UNP Q9Y3Z3
A	119	GLY	-	expression tag	UNP Q9Y3Z3

*Continued on next page...*

*Continued from previous page...*

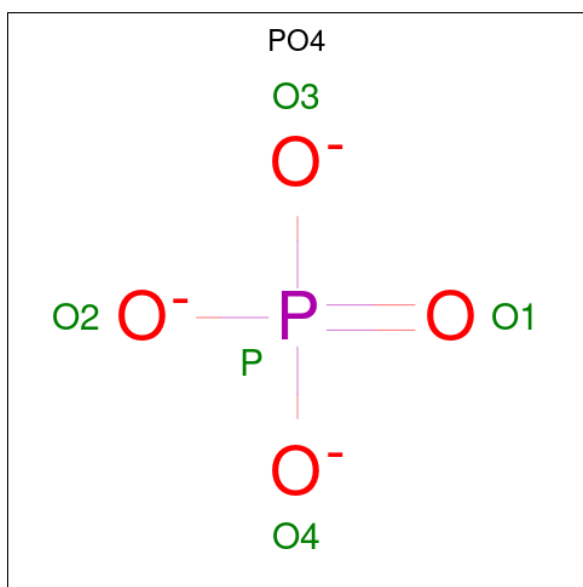
Chain	Residue	Modelled	Actual	Comment	Reference
B	99	ALA	-	expression tag	UNP Q9Y3Z3
B	100	SER	-	expression tag	UNP Q9Y3Z3
B	101	TRP	-	expression tag	UNP Q9Y3Z3
B	102	SER	-	expression tag	UNP Q9Y3Z3
B	103	HIS	-	expression tag	UNP Q9Y3Z3
B	104	PRO	-	expression tag	UNP Q9Y3Z3
B	105	GLN	-	expression tag	UNP Q9Y3Z3
B	106	PHE	-	expression tag	UNP Q9Y3Z3
B	107	GLU	-	expression tag	UNP Q9Y3Z3
B	108	LYS	-	expression tag	UNP Q9Y3Z3
B	109	GLY	-	expression tag	UNP Q9Y3Z3
B	110	ALA	-	expression tag	UNP Q9Y3Z3
B	111	LEU	-	expression tag	UNP Q9Y3Z3
B	112	GLU	-	expression tag	UNP Q9Y3Z3
B	113	VAL	-	expression tag	UNP Q9Y3Z3
B	114	LEU	-	expression tag	UNP Q9Y3Z3
B	115	PHE	-	expression tag	UNP Q9Y3Z3
B	116	GLN	-	expression tag	UNP Q9Y3Z3
B	117	GLY	-	expression tag	UNP Q9Y3Z3
B	118	PRO	-	expression tag	UNP Q9Y3Z3
B	119	GLY	-	expression tag	UNP Q9Y3Z3
C	99	ALA	-	expression tag	UNP Q9Y3Z3
C	100	SER	-	expression tag	UNP Q9Y3Z3
C	101	TRP	-	expression tag	UNP Q9Y3Z3
C	102	SER	-	expression tag	UNP Q9Y3Z3
C	103	HIS	-	expression tag	UNP Q9Y3Z3
C	104	PRO	-	expression tag	UNP Q9Y3Z3
C	105	GLN	-	expression tag	UNP Q9Y3Z3
C	106	PHE	-	expression tag	UNP Q9Y3Z3
C	107	GLU	-	expression tag	UNP Q9Y3Z3
C	108	LYS	-	expression tag	UNP Q9Y3Z3
C	109	GLY	-	expression tag	UNP Q9Y3Z3
C	110	ALA	-	expression tag	UNP Q9Y3Z3
C	111	LEU	-	expression tag	UNP Q9Y3Z3
C	112	GLU	-	expression tag	UNP Q9Y3Z3
C	113	VAL	-	expression tag	UNP Q9Y3Z3
C	114	LEU	-	expression tag	UNP Q9Y3Z3
C	115	PHE	-	expression tag	UNP Q9Y3Z3
C	116	GLN	-	expression tag	UNP Q9Y3Z3
C	117	GLY	-	expression tag	UNP Q9Y3Z3
C	118	PRO	-	expression tag	UNP Q9Y3Z3
C	119	GLY	-	expression tag	UNP Q9Y3Z3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	99	ALA	-	expression tag	UNP Q9Y3Z3
D	100	SER	-	expression tag	UNP Q9Y3Z3
D	101	TRP	-	expression tag	UNP Q9Y3Z3
D	102	SER	-	expression tag	UNP Q9Y3Z3
D	103	HIS	-	expression tag	UNP Q9Y3Z3
D	104	PRO	-	expression tag	UNP Q9Y3Z3
D	105	GLN	-	expression tag	UNP Q9Y3Z3
D	106	PHE	-	expression tag	UNP Q9Y3Z3
D	107	GLU	-	expression tag	UNP Q9Y3Z3
D	108	LYS	-	expression tag	UNP Q9Y3Z3
D	109	GLY	-	expression tag	UNP Q9Y3Z3
D	110	ALA	-	expression tag	UNP Q9Y3Z3
D	111	LEU	-	expression tag	UNP Q9Y3Z3
D	112	GLU	-	expression tag	UNP Q9Y3Z3
D	113	VAL	-	expression tag	UNP Q9Y3Z3
D	114	LEU	-	expression tag	UNP Q9Y3Z3
D	115	PHE	-	expression tag	UNP Q9Y3Z3
D	116	GLN	-	expression tag	UNP Q9Y3Z3
D	117	GLY	-	expression tag	UNP Q9Y3Z3
D	118	PRO	-	expression tag	UNP Q9Y3Z3
D	119	GLY	-	expression tag	UNP Q9Y3Z3

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	P	0	0
			5	4	1		
2	C	1	Total	O	P	0	0
			5	4	1		
2	D	1	Total	O	P	0	0
			5	4	1		

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

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

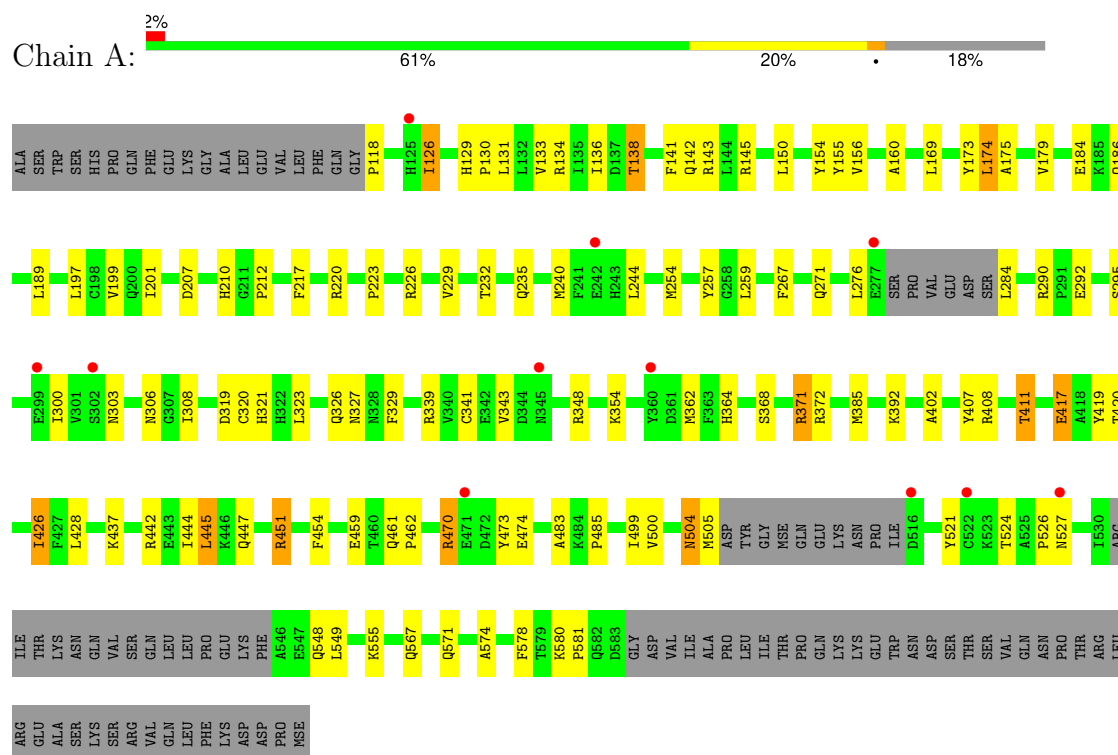
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	7	Total	O	0	0
			7	7		
4	B	8	Total	O	0	0
			8	8		
4	C	7	Total	O	0	0
			7	7		
4	D	6	Total	O	0	0
			6	6		

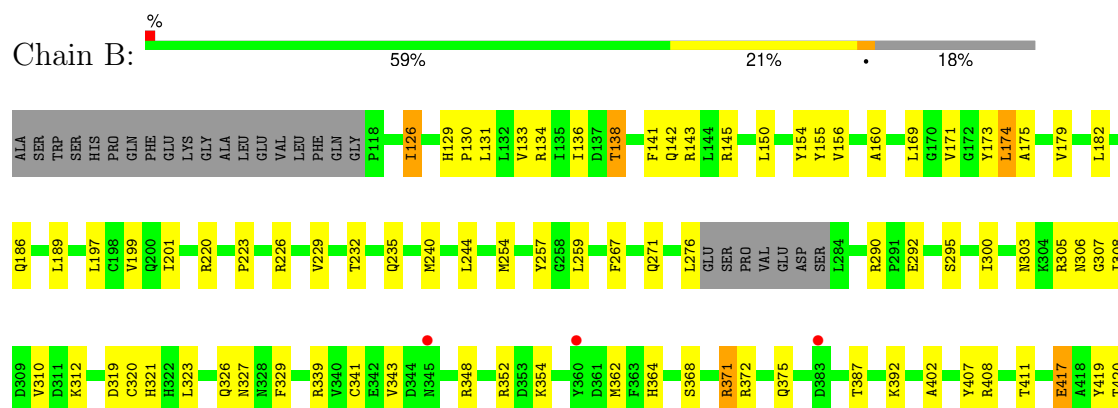
### 3 Residue-property plots [i](#)

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: SAM domain and HD domain-containing protein 1



- Molecule 1: SAM domain and HD domain-containing protein 1









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.52Å 95.82Å 96.66Å 91.18° 109.24° 115.20°	Depositor
Resolution (Å)	34.56 – 3.10 34.56 – 3.10	Depositor EDS
% Data completeness (in resolution range)	94.5 (34.56-3.10) 94.4 (34.56-3.10)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 3.06Å)	Xtriage
Refinement program	PHENIX 1.7_650	Depositor
R, $R_{free}$	0.196 , 0.228 0.192 , 0.223	Depositor DCC
$R_{free}$ test set	1971 reflections (4.46%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	62.8	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 63.0	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	0.93	EDS
Total number of atoms	13595	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.75% 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: PO4, ZN

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	0.56	1/3499 (0.0%)	0.68	4/4732 (0.1%)
1	B	0.56	0/3498	0.69	4/4725 (0.1%)
1	C	0.50	0/3426	0.78	6/4630 (0.1%)
1	D	0.51	0/3418	0.65	5/4615 (0.1%)
All	All	0.53	1/13841 (0.0%)	0.70	19/18702 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	184	GLU	CD-OE2	7.92	1.34	1.25

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	290	ARG	NE-CZ-NH1	-21.25	109.68	120.30
1	C	290	ARG	NE-CZ-NH2	19.10	129.85	120.30
1	B	451	ARG	NE-CZ-NH2	14.49	127.55	120.30
1	A	451	ARG	NE-CZ-NH2	14.07	127.33	120.30
1	B	451	ARG	NE-CZ-NH1	-13.61	113.50	120.30

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	3423	0	3252	90	0
1	B	3423	0	3292	101	0
1	C	3353	0	3188	92	0
1	D	3344	0	3197	90	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	1	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	7	0	0	1	0
4	B	8	0	0	2	0
4	C	7	0	0	2	0
4	D	6	0	0	0	0
All	All	13595	0	12929	355	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 13.

The worst 5 of 355 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:C:408:ARG:H	1:C:411:THR:HG22	1.28	0.98
1:A:408:ARG:H	1:A:411:THR:HG22	1.28	0.97
1:D:408:ARG:H	1:D:411:THR:HG22	1.31	0.96
1:B:408:ARG:H	1:B:411:THR:HG22	1.31	0.94
1:C:485:PRO:HG2	1:C:489:LEU:HD11	1.56	0.87

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	427/528 (81%)	404 (95%)	23 (5%)	0	100	100
1	B	426/528 (81%)	403 (95%)	23 (5%)	0	100	100
1	C	419/528 (79%)	400 (96%)	19 (4%)	0	100	100
1	D	414/528 (78%)	395 (95%)	19 (5%)	0	100	100
All	All	1686/2112 (80%)	1602 (95%)	84 (5%)	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	A	346/458 (76%)	326 (94%)	20 (6%)	17	45
1	B	350/458 (76%)	331 (95%)	19 (5%)	18	47
1	C	342/458 (75%)	319 (93%)	23 (7%)	13	40
1	D	341/458 (74%)	325 (95%)	16 (5%)	22	52
All	All	1379/1832 (75%)	1301 (94%)	78 (6%)	17	46

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	500	VAL
1	D	445	LEU
1	C	505	MSE
1	D	276	LEU
1	D	500	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	504	ASN
1	C	425	ASN
1	D	447	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	326	GLN
1	C	447	GLN

### 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 oligosaccharides 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	PO4	C	3	3	4,4,4	0.91	0	6,6,6	0.83	0
2	PO4	B	2	3	4,4,4	1.05	0	6,6,6	0.56	0
2	PO4	A	1	3	4,4,4	1.11	0	6,6,6	0.25	0
2	PO4	D	4	3	4,4,4	0.95	0	6,6,6	0.93	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	PO4	1	0
2	D	4	PO4	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	427/528 (80%)	-0.26	11 (2%) 57 38	24, 55, 100, 135	0
1	B	426/528 (80%)	-0.26	6 (1%) 73 56	24, 56, 99, 129	0
1	C	419/528 (79%)	-0.15	5 (1%) 76 60	33, 65, 106, 143	0
1	D	416/528 (78%)	-0.15	5 (1%) 76 60	29, 64, 105, 172	0
All	All	1688/2112 (79%)	-0.21	27 (1%) 70 52	24, 60, 103, 172	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	516	ASP	4.0
1	B	471	GLU	3.1
1	A	345	ASN	3.0
1	A	471	GLU	2.9
1	C	403	GLY	2.8

### 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( $\text{\AA}^2$ )	Q<0.9
2	PO4	B	2	5/5	0.95	0.07	71,71,71,71	0
2	PO4	A	1	5/5	0.96	0.06	70,70,70,70	0
2	PO4	C	3	5/5	0.96	0.06	63,63,63,63	0
2	PO4	D	4	5/5	0.98	0.05	64,64,64,64	0
3	ZN	A	627	1/1	0.99	0.02	53,53,53,53	0
3	ZN	B	4	1/1	0.99	0.03	50,50,50,50	0
3	ZN	C	627	1/1	0.99	0.02	53,53,53,53	0
3	ZN	D	2	1/1	0.99	0.02	51,51,51,51	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.