



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 11:52 AM GMT

PDB ID : 4CBY  
Title : Design, synthesis, and biological evaluation of potent and selective Class IIa HDAC inhibitors as a potential therapy for Huntingtons disease  
Authors : Burli, R.W.; Luckhurst, C.A.; Aziz, O.; Matthews, K.L.; Yates, D.; Lyons, K.A.; Beconi, M.; McAllister, G.; Breccia, P.; Stott, A.J.; Penrose, S.D.; Wall, M.; Lamers, M.; Leonard, P.; Mueller, I.; Richardson, C.M.; Jarvis, R.; Stones, L.; Hughes, S.; Wishart, G.; Haughan, A.F.; OConnell, C.; Mead, T.; McNeil, H.; Vann, J.; Mangette, J.; Maillard, M.; Beaumont, V.; Munoz-Sanjuan, I.; Dominguez, C.  
Deposited on : 2013-10-17  
Resolution : 2.72 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

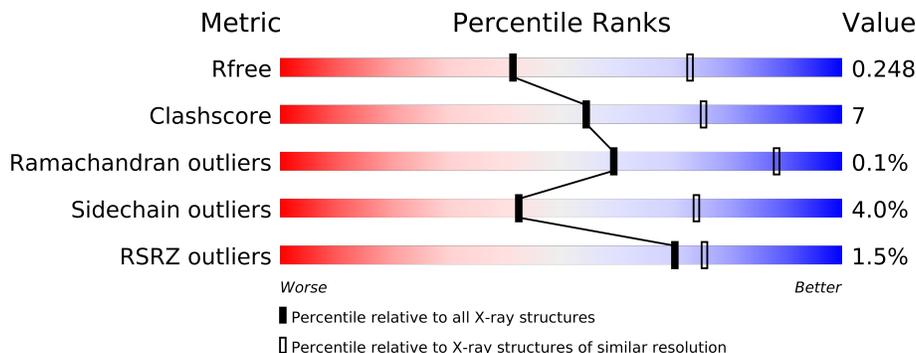
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.72 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))
$R_{free}$	66092	1770 (2.74-2.70)
Clashscore	79885	2183 (2.74-2.70)
Ramachandran outliers	78287	2147 (2.74-2.70)
Sidechain outliers	78261	2148 (2.74-2.70)
RSRZ outliers	66119	1772 (2.74-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	395	
1	B	395	
1	C	395	
1	D	395	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	NA	B	2035	-	X

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 11435 atoms, of which 0 are hydrogen and 0 are deuterium.

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 HISTONE DEACETYLASE 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	380	2823	1782	494	527	20	0	0	0
1	B	380	2795	1767	483	525	20	0	0	0
1	C	380	2802	1764	488	530	20	0	0	0
1	D	383	2812	1772	492	528	20	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

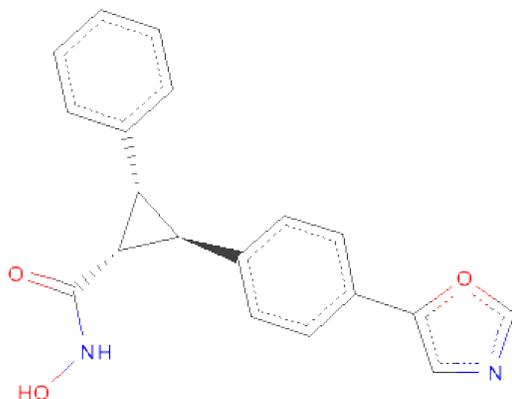
Chain	Residue	Modelled	Actual	Comment	Reference
A	645	MET	-	EXPRESSION TAG	UNP P56524
A	646	GLY	-	EXPRESSION TAG	UNP P56524
A	647	SER	-	EXPRESSION TAG	UNP P56524
A	728	ALA	LEU	ENGINEERED MUTATION	UNP P56524
A	1034	HIS	-	EXPRESSION TAG	UNP P56524
A	1035	HIS	-	EXPRESSION TAG	UNP P56524
A	1036	HIS	-	EXPRESSION TAG	UNP P56524
A	1037	HIS	-	EXPRESSION TAG	UNP P56524
A	1038	HIS	-	EXPRESSION TAG	UNP P56524
A	1039	HIS	-	EXPRESSION TAG	UNP P56524
B	645	MET	-	EXPRESSION TAG	UNP P56524
B	646	GLY	-	EXPRESSION TAG	UNP P56524
B	647	SER	-	EXPRESSION TAG	UNP P56524
B	728	ALA	LEU	ENGINEERED MUTATION	UNP P56524
B	1034	HIS	-	EXPRESSION TAG	UNP P56524
B	1035	HIS	-	EXPRESSION TAG	UNP P56524
B	1036	HIS	-	EXPRESSION TAG	UNP P56524
B	1037	HIS	-	EXPRESSION TAG	UNP P56524
B	1038	HIS	-	EXPRESSION TAG	UNP P56524
B	1039	HIS	-	EXPRESSION TAG	UNP P56524
C	645	MET	-	EXPRESSION TAG	UNP P56524

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	646	GLY	-	EXPRESSION TAG	UNP P56524
C	647	SER	-	EXPRESSION TAG	UNP P56524
C	728	ALA	LEU	ENGINEERED MUTATION	UNP P56524
C	1034	HIS	-	EXPRESSION TAG	UNP P56524
C	1035	HIS	-	EXPRESSION TAG	UNP P56524
C	1036	HIS	-	EXPRESSION TAG	UNP P56524
C	1037	HIS	-	EXPRESSION TAG	UNP P56524
C	1038	HIS	-	EXPRESSION TAG	UNP P56524
C	1039	HIS	-	EXPRESSION TAG	UNP P56524
D	645	MET	-	EXPRESSION TAG	UNP P56524
D	646	GLY	-	EXPRESSION TAG	UNP P56524
D	647	SER	-	EXPRESSION TAG	UNP P56524
D	728	ALA	LEU	ENGINEERED MUTATION	UNP P56524
D	1034	HIS	-	EXPRESSION TAG	UNP P56524
D	1035	HIS	-	EXPRESSION TAG	UNP P56524
D	1036	HIS	-	EXPRESSION TAG	UNP P56524
D	1037	HIS	-	EXPRESSION TAG	UNP P56524
D	1038	HIS	-	EXPRESSION TAG	UNP P56524
D	1039	HIS	-	EXPRESSION TAG	UNP P56524

- Molecule 2 is (1R,2R,3R)-2-[4-(1,3-OXAZOL-5-YL)PHENYL]-N-OXIDANYL-3-PHENYL-CYCLOPROPANE-1-CARBOXAMIDE (three-letter code: KEE) (formula: C<sub>19</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
2	A	1	24	19	2	3	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			24	19	2	3		
2	C	1	Total	C	N	O	0	0
			24	19	2	3		
2	D	1	Total	C	N	O	0	0
			24	19	2	3		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Zn	0	0
			2	2		
3	A	2	Total	Zn	0	0
			2	2		
3	D	2	Total	Zn	0	0
			2	2		
3	C	2	Total	Zn	0	0
			2	2		

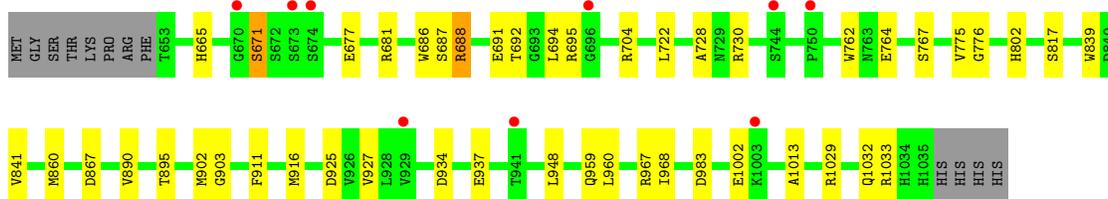
- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Na	0	0
			2	2		
4	A	2	Total	Na	0	0
			2	2		
4	D	2	Total	Na	0	0
			2	2		
4	C	2	Total	Na	0	0
			2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	22	Total	O	0	0
			22	22		
5	B	23	Total	O	0	0
			23	23		
5	C	24	Total	O	0	0
			24	24		
5	D	22	Total	O	0	0
			22	22		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.15Å 161.82Å 83.73Å 90.00° 97.03° 90.00°	Depositor
Resolution (Å)	83.10 – 2.72 29.37 – 2.72	Depositor EDS
% Data completeness (in resolution range)	98.3 (83.10-2.72) 98.5 (29.37-2.72)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 2.72Å)	Xtrriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.209 , 0.246 0.211 , 0.248	Depositor DCC
$R_{free}$ test set	1886 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.7	Xtrriage
Anisotropy	0.093	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 7.7	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Outliers	2 of 37669 reflections (0.005%)	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11435	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 37.36 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 4.3423e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

## 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: NA, ZN, KEE

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	0.58	1/2887 (0.0%)	0.68	2/3924 (0.1%)
1	B	0.60	3/2860 (0.1%)	0.65	0/3894
1	C	0.58	1/2865 (0.0%)	0.68	0/3899
1	D	0.57	0/2876	0.70	3/3912 (0.1%)
All	All	0.58	5/11488 (0.0%)	0.68	5/15629 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	712	GLN	CD-NE2	-5.42	1.19	1.32
1	C	816	ASN	CG-OD1	5.21	1.35	1.24
1	B	686	TRP	CD2-CE2	5.15	1.47	1.41
1	B	762	TRP	CD2-CE2	5.09	1.47	1.41
1	A	762	TRP	CD2-CE2	5.03	1.47	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	688	ARG	NE-CZ-NH1	10.20	125.40	120.30
1	D	688	ARG	NE-CZ-NH2	-7.80	116.40	120.30
1	D	681	ARG	NE-CZ-NH1	6.22	123.41	120.30
1	A	682	ILE	CB-CA-C	-5.32	100.96	111.60
1	A	798	ARG	NE-CZ-NH2	-5.25	117.67	120.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2823	0	2753	44	0
1	B	2795	0	2683	42	0
1	C	2802	0	2693	38	0
1	D	2812	0	2701	35	0
2	A	24	0	13	0	0
2	B	24	0	12	0	0
2	C	24	0	13	0	0
2	D	24	0	13	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
5	A	22	0	0	1	0
5	B	23	0	0	2	0
5	C	24	0	0	1	0
5	D	22	0	0	4	0
All	All	11435	0	10881	152	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

The worst 5 of 152 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:688:ARG:NH2	1:D:983:ASP:OD1	1.96	0.98
1:D:937:GLU:CB	5:D:2017:HOH:O	2.16	0.94
1:D:902:MET:HE3	1:D:1013:ALA:HB2	1.46	0.94
1:A:682:ILE:HG23	1:A:796:VAL:HG12	1.52	0.92
1:B:968:ILE:HB	5:B:2019:HOH:O	1.76	0.86

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 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	378/395 (96%)	363 (96%)	15 (4%)	0	100	100
1	B	378/395 (96%)	361 (96%)	17 (4%)	0	100	100
1	C	378/395 (96%)	364 (96%)	13 (3%)	1 (0%)	50	81
1	D	381/395 (96%)	365 (96%)	15 (4%)	1 (0%)	50	81
All	All	1515/1580 (96%)	1453 (96%)	60 (4%)	2 (0%)	59	88

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	671	SER
1	D	671	SER

### 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	297/321 (92%)	284 (96%)	13 (4%)	39	70
1	B	290/321 (90%)	281 (97%)	9 (3%)	52	83
1	C	292/321 (91%)	276 (94%)	16 (6%)	30	59
1	D	291/321 (91%)	282 (97%)	9 (3%)	52	83
All	All	1170/1284 (91%)	1123 (96%)	47 (4%)	42	74

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1004	VAL
1	C	745	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	895	THR
1	C	702	ARG
1	C	802	HIS

Some 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	729	ASN
1	C	678	HIS
1	D	731	GLN
1	B	843	HIS
1	B	995	ASN

### 5.3.3 RNA [i](#)

There are no RNA chains 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 20 ligands modelled in this entry, 16 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	KEE	A	2033	3	27,27,27	0.91	1 (3%)	35,38,38	1.66	8 (22%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	KEE	B	2032	3	27,27,27	1.11	1 (3%)	35,38,38	1.28	5 (14%)
2	KEE	C	2033	3	27,27,27	0.98	2 (7%)	35,38,38	1.29	4 (11%)
2	KEE	D	2036	3	27,27,27	0.80	0	35,38,38	1.38	5 (14%)

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	KEE	A	2033	3	-	0/16/27/27	0/3/4/4
2	KEE	B	2032	3	-	0/16/27/27	0/3/4/4
2	KEE	C	2033	3	-	0/16/27/27	0/3/4/4
2	KEE	D	2036	3	-	0/16/27/27	0/3/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2032	KEE	C8-C17	3.56	1.50	1.46
2	A	2033	KEE	C8-C17	2.65	1.49	1.46
2	C	2033	KEE	C2-C3	2.41	1.55	1.52
2	C	2033	KEE	C8-C17	2.14	1.48	1.46

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2033	KEE	C5-C3-C2	-4.63	112.38	121.92
2	D	2036	KEE	C5-C3-C2	-4.29	113.08	121.92
2	B	2032	KEE	O22-C1-N23	-3.75	118.71	123.25
2	A	2033	KEE	C7-C8-C17	-3.56	116.61	120.39
2	D	2036	KEE	C11-C4-C3	-3.29	113.45	121.63

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	380/395 (96%)	-0.07	7 (1%) 65 72	15, 29, 45, 57	0
1	B	380/395 (96%)	-0.11	1 (0%) 91 95	16, 28, 45, 57	0
1	C	380/395 (96%)	-0.08	6 (1%) 68 74	18, 29, 54, 79	0
1	D	383/395 (96%)	0.01	9 (2%) 57 64	20, 31, 55, 78	0
All	All	1523/1580 (96%)	-0.06	23 (1%) 70 75	15, 29, 50, 79	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	673	SER	5.2
1	D	673	SER	3.1
1	A	697	LYS	2.9
1	D	1003	LYS	2.8
1	A	972	LEU	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 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NA	B	2035	1/1	0.35	6.58	28,28,28,28	0
4	NA	A	2036	1/1	0.22	1.50	28,28,28,28	0
4	NA	D	2039	1/1	0.21	1.39	23,23,23,23	0
2	KEE	C	2033	24/24	0.21	0.60	26,29,37,38	0
4	NA	C	2036	1/1	0.17	0.51	25,25,25,25	0
4	NA	B	2036	1/1	0.15	0.34	23,23,23,23	0
4	NA	D	2040	1/1	0.13	0.09	24,24,24,24	0
2	KEE	A	2033	24/24	0.18	0.08	24,28,37,38	0
4	NA	C	2037	1/1	0.13	0.05	22,22,22,22	0
2	KEE	B	2032	24/24	0.16	-0.15	23,26,35,36	0
2	KEE	D	2036	24/24	0.17	-0.31	28,33,41,43	0
4	NA	A	2037	1/1	0.12	-0.48	20,20,20,20	0
3	ZN	A	2034	1/1	0.09	-1.92	19,19,19,19	0
3	ZN	A	2035	1/1	0.04	-1.97	25,25,25,25	0
3	ZN	C	2035	1/1	0.04	-2.07	46,46,46,46	0
3	ZN	D	2038	1/1	0.05	-2.14	45,45,45,45	0
3	ZN	C	2034	1/1	0.07	-3.01	25,25,25,25	0
3	ZN	B	2034	1/1	0.03	-4.04	28,28,28,28	0
3	ZN	D	2037	1/1	0.07	-4.83	25,25,25,25	0
3	ZN	B	2033	1/1	0.06	-7.58	23,23,23,23	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.