



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 02:53 am GMT

PDB ID : 5G0J  
Title : Crystal structure of Danio rerio HDAC6 CD1 and CD2 (linker intact) in complex with Nexturastat A  
Authors : Miyake, Y.; Keusch, J.J.; Wang, L.; Saito, M.; Hess, D.; Wang, X.; Melancon, B.J.; Helquist, P.; Gut, H.; Matthias, P.  
Deposited on : 2016-03-18  
Resolution : 2.88 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

<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

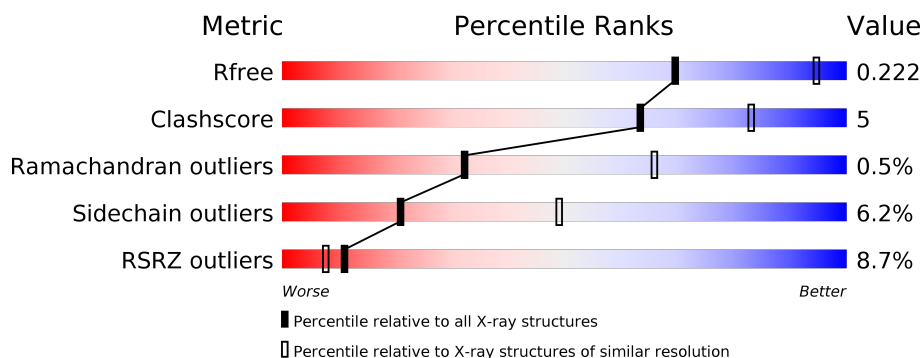
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.88 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	2135 (2.90-2.86)
Clashscore	112137	2400 (2.90-2.86)
Ramachandran outliers	110173	2346 (2.90-2.86)
Sidechain outliers	110143	2349 (2.90-2.86)
RSRZ outliers	101464	2149 (2.90-2.86)

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	794	

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	N4R	A	1817	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6061 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 HDAC6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	756	Total	C	N	O	S	0	1	0
			5866	3706	1021	1100	39			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	GLY	-	EXPRESSION TAG	UNP F8W4B7
A	39	PRO	-	EXPRESSION TAG	UNP F8W4B7

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Cl	0	0
			2	2		

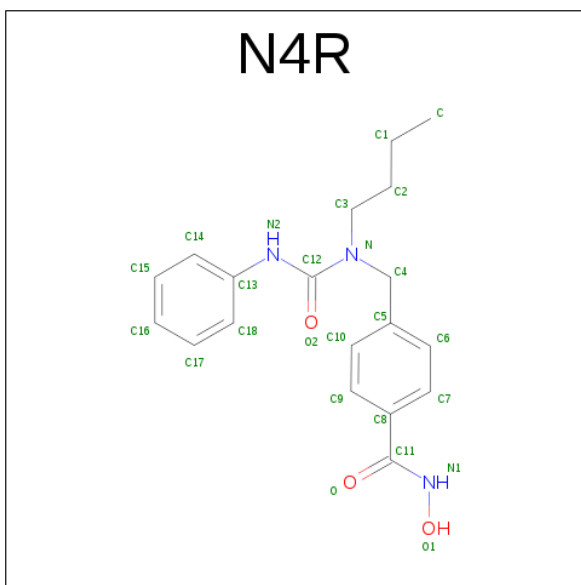
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	5	Total	K	0	0
			5	5		

- Molecule 5 is NEXTURASTAT A (three-letter code: N4R) (formula: C<sub>19</sub>H<sub>23</sub>N<sub>3</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			25	19	3	3		

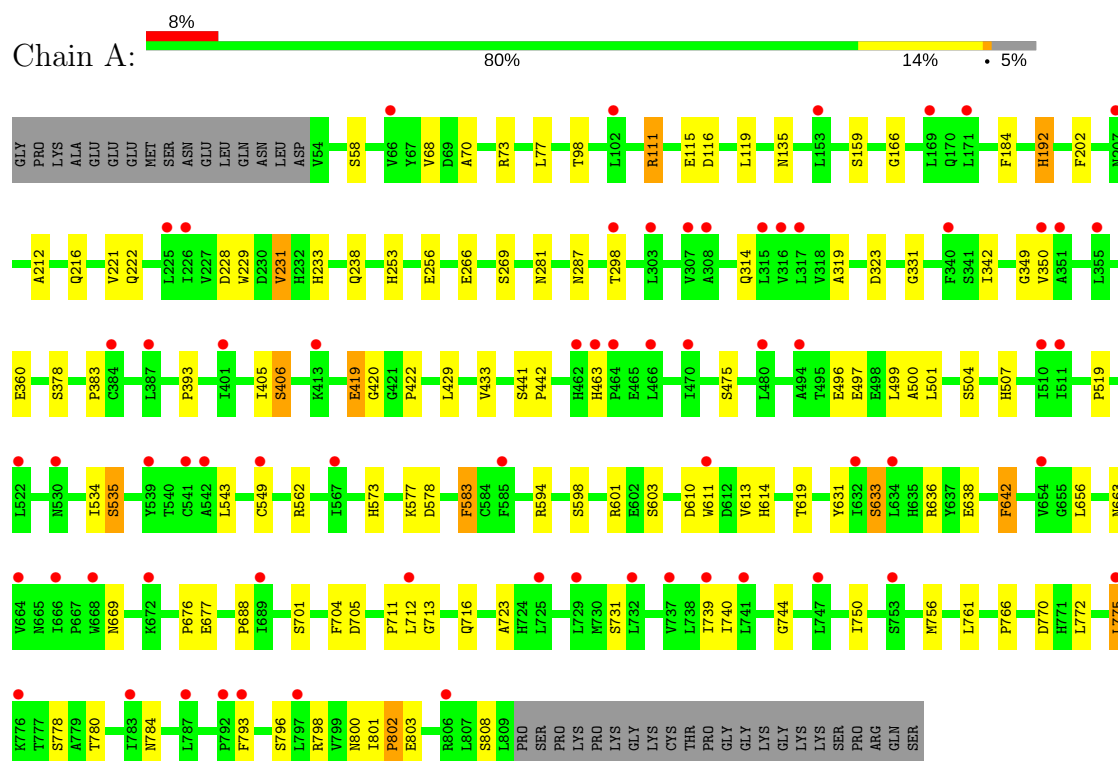
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	162	Total	O	0	0
			162	162		

### 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 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: HDAC6



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	187.22Å 187.22Å 102.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.88 162.14 – 2.88	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-2.88) 99.9 (162.14-2.88)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.77 (at 2.86Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.176 , 0.214 0.187 , 0.222	Depositor DCC
$R_{free}$ test set	2383 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	74.9	Xtriage
Anisotropy	0.727	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 76.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6061	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.77% 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: ZN, N4R, K, CL

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	0/6011	0.74	0/8160

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

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	5866	0	5729	57	0
2	A	2	0	0	0	0
3	A	1	0	0	0	0
4	A	5	0	0	0	0
5	A	25	0	0	0	0
6	A	162	0	0	5	0
All	All	6061	0	5729	57	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 5.

All (57) 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:420:GLY:HA2	1:A:801:ILE:HG23	1.82	0.60
1:A:111:ARG:O	1:A:166:GLY:HA3	2.03	0.59
1:A:504:SER:HB3	1:A:577:LYS:HE3	1.84	0.58
1:A:610:ASP:HA	1:A:701:SER:HB2	1.87	0.57
1:A:212:ALA:O	1:A:216:GLN:HG3	2.05	0.57
1:A:669:ASN:O	1:A:778:SER:HB2	2.06	0.56
1:A:463:HIS:HE1	1:A:583:PHE:CE1	2.24	0.55
1:A:238:GLN:HE22	1:A:266:GLU:HB2	1.72	0.54
1:A:463:HIS:HB3	6:A:2119:HOH:O	2.07	0.54
1:A:803:GLU:HA	1:A:803:GLU:OE1	2.08	0.53
1:A:519:PRO:HB3	6:A:2117:HOH:O	2.08	0.53
1:A:636:ARG:HG2	1:A:638:GLU:HG2	1.90	0.53
1:A:497:GLU:O	1:A:500:ALA:HB3	2.10	0.52
1:A:405:ILE:HG22	1:A:801:ILE:HD12	1.92	0.52
1:A:614:HIS:HB2	6:A:2141:HOH:O	2.10	0.52
1:A:429:LEU:CD1	1:A:793:PHE:HB3	2.39	0.52
1:A:298:THR:HG21	1:A:393:PRO:HG3	1.92	0.51
1:A:222:GLN:HB2	1:A:314:GLN:HE22	1.75	0.51
1:A:507:HIS:HB2	1:A:578:ASP:HA	1.94	0.50
1:A:711:PRO:HG2	1:A:712:LEU:HD12	1.92	0.50
1:A:631:TYR:CE2	1:A:633:SER:HB2	2.47	0.49
1:A:704:PHE:HB2	1:A:750:ILE:HG22	1.94	0.49
1:A:319:ALA:HB1	1:A:360:GLU:HG3	1.94	0.49
1:A:419:GLU:HG2	1:A:784:ASN:CG	2.34	0.47
1:A:115:GLU:O	1:A:119:LEU:HG	2.14	0.47
1:A:501:LEU:HD22	1:A:594:ARG:HB3	1.97	0.47
1:A:233:HIS:CD2	1:A:253:HIS:CE1	3.03	0.47
1:A:429:LEU:HD12	1:A:793:PHE:HB3	1.98	0.46
1:A:583:PHE:HB3	6:A:2115:HOH:O	2.13	0.46
1:A:549:CYS:SG	1:A:740:ILE:HG12	2.56	0.46
1:A:676:PRO:HG3	1:A:772:LEU:HA	1.98	0.46
1:A:221:VAL:HG13	1:A:314:GLN:HB2	1.97	0.46
1:A:406:SER:HB2	1:A:801:ILE:HG13	1.97	0.46
1:A:256:GLU:HA	1:A:287:ASN:HA	1.98	0.45
1:A:613:VAL:HG11	1:A:713:GLY:HA2	1.97	0.45
1:A:705:ASP:HB3	1:A:744:GLY:H	1.82	0.44
1:A:636:ARG:HD3	1:A:642:PHE:CE1	2.52	0.44
1:A:419:GLU:HG2	1:A:784:ASN:OD1	2.17	0.44
1:A:723:ALA:HB2	1:A:756:MET:HB3	1.99	0.44
1:A:562:ARG:HD3	6:A:2106:HOH:O	2.19	0.43
1:A:688:PRO:HB3	1:A:796:SER:HB2	2.01	0.43
1:A:705:ASP:O	1:A:713:GLY:HA3	2.17	0.43

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:677:GLU:HG3	1:A:775:LEU:HB2	2.00	0.43
1:A:231:VAL:HG11	1:A:331:GLY:HA2	2.01	0.43
1:A:228:ASP:HA	1:A:319:ALA:HB3	2.00	0.43
1:A:349:GLY:HA2	1:A:766:PRO:HG2	2.00	0.42
1:A:70:ALA:HA	1:A:73:ARG:HG3	2.01	0.42
1:A:383:PRO:HD3	1:A:442:PRO:HB3	2.02	0.42
1:A:534:ILE:HG22	1:A:535:SER:N	2.34	0.42
1:A:613:VAL:HG12	1:A:705:ASP:OD2	2.19	0.42
1:A:496:GLU:HA	1:A:499:LEU:HD12	2.02	0.41
1:A:342:ILE:HG23	1:A:731:SER:HB2	2.03	0.41
1:A:739:ILE:HD12	1:A:761:LEU:HD11	2.03	0.41
1:A:422:PRO:HA	1:A:800:ASN:OD1	2.21	0.41
1:A:534:ILE:HG22	1:A:535:SER:H	1.86	0.41
1:A:463:HIS:CE1	1:A:583:PHE:CE1	3.09	0.40
1:A:802:PRO:HB2	1:A:803:GLU:H	1.73	0.40

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	755/794 (95%)	691 (92%)	60 (8%)	4 (0%)	32 66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	798	ARG
1	A	802	PRO
1	A	192	HIS
1	A	808	SER

### 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	641/672 (95%)	601 (94%)	40 (6%)	21	50

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	SER
1	A	68	VAL
1	A	77	LEU
1	A	98	THR
1	A	111	ARG
1	A	116	ASP
1	A	135	ASN
1	A	159	SER
1	A	184	PHE
1	A	192	HIS
1	A	202	PHE
1	A	229	TRP
1	A	231	VAL
1	A	269	SER
1	A	281	ASN
1	A	323	ASP
1	A	350	VAL
1	A	378	SER
1	A	406	SER
1	A	419	GLU
1	A	433	VAL
1	A	441	SER
1	A	475	SER
1	A	535	SER
1	A	543	LEU
1	A	573	HIS
1	A	583	PHE
1	A	598	SER
1	A	601	ARG
1	A	603	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	611	TRP
1	A	619	THR
1	A	633	SER
1	A	642	PHE
1	A	656	LEU
1	A	663	ASN
1	A	716	GLN
1	A	770	ASP
1	A	775	LEU
1	A	780	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	55	GLN
1	A	233	HIS
1	A	386	HIS
1	A	463	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 8 are monoatomic - leaving 1 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	$\# Z  > 2$
5	N4R	A	1817	3	26,26,26	0.35	0	32,33,33	0.59	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
5	N4R	A	1817	3	-	0/22/22/22	0/2/2/2

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.

No monomer is involved in short contacts.

## 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	756/794 (95%)	0.94	66 (8%) 11 8	55, 88, 118, 152	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	672	LYS	3.9
1	A	585	PHE	3.6
1	A	666	ILE	3.2
1	A	712	LEU	3.2
1	A	806	ARG	3.2
1	A	664	VAL	3.0
1	A	511	ILE	3.0
1	A	632	ILE	3.0
1	A	567	ILE	3.0
1	A	689	ILE	2.9
1	A	739	ILE	2.9
1	A	462	HIS	2.9
1	A	725	LEU	2.8
1	A	470	ILE	2.8
1	A	308	ALA	2.8
1	A	737	VAL	2.8
1	A	783	ILE	2.8
1	A	775	LEU	2.8
1	A	792	PRO	2.7
1	A	169	LEU	2.7
1	A	317	LEU	2.7
1	A	66	VAL	2.7
1	A	355	LEU	2.6
1	A	729	LEU	2.6
1	A	316	VAL	2.5
1	A	226	ILE	2.5
1	A	384	CYS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	350	VAL	2.4
1	A	315	LEU	2.4
1	A	539	TYR	2.4
1	A	464	PRO	2.4
1	A	351	ALA	2.4
1	A	510	ILE	2.4
1	A	634	LEU	2.4
1	A	530	ASN	2.3
1	A	797	LEU	2.3
1	A	340	PHE	2.3
1	A	542	ALA	2.3
1	A	303	LEU	2.3
1	A	668	TRP	2.2
1	A	747	LEU	2.2
1	A	793	PHE	2.2
1	A	387	LEU	2.2
1	A	307	VAL	2.2
1	A	153	LEU	2.2
1	A	298	THR	2.2
1	A	753	SER	2.2
1	A	611	TRP	2.2
1	A	102	LEU	2.2
1	A	413	LYS	2.2
1	A	466	LEU	2.2
1	A	541	CYS	2.2
1	A	463	HIS	2.2
1	A	522	LEU	2.1
1	A	401	ILE	2.1
1	A	654	VAL	2.1
1	A	225	LEU	2.1
1	A	207	ASN	2.1
1	A	787	LEU	2.1
1	A	549	CYS	2.1
1	A	171	LEU	2.0
1	A	732	LEU	2.0
1	A	494	ALA	2.0
1	A	480	LEU	2.0
1	A	741	LEU	2.0
1	A	776	LYS	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 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
5	N4R	A	1817	25/25	0.89	0.44	1.87	112,130,135,135	0
3	ZN	A	1811	1/1	0.99	0.24	0.37	123,123,123,123	0
2	CL	A	1809	1/1	0.83	0.23	0.27	110,110,110,110	0
2	CL	A	1810	1/1	0.96	0.12	-1.80	113,113,113,113	0
4	K	A	1816	1/1	0.94	0.10	-2.12	97,97,97,97	0
4	K	A	1813	1/1	0.90	0.09	-2.64	132,132,132,132	0
4	K	A	1812	1/1	0.97	0.15	-3.06	88,88,88,88	0
4	K	A	1815	1/1	0.92	0.07	-3.37	137,137,137,137	0
4	K	A	1814	1/1	0.86	0.08	-3.96	113,113,113,113	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.