Full wwPDB X-ray Structure Validation Report (i)

Oct 17, 2023 – 03:45 AM EDT

PDB ID	:	2EG7
Title	:	The crystal structure of E. coli dihydroorotase complexed with HDDP
Authors	:	Lee, M.; Maher, M.J.; Guss, J.M.
Deposited on	:	2007-02-28
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

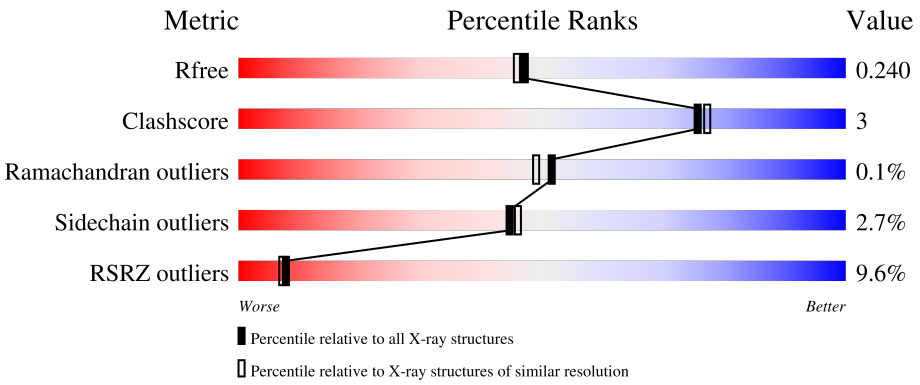
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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 <=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	347	10% 91% 8% •
1	B	347	9% 86% 13% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5772 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 Dihydroorotase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	343	2702	1707	478	501	16	7	2	0
1	B	343	2694	1702	475	501	16	12	1	0

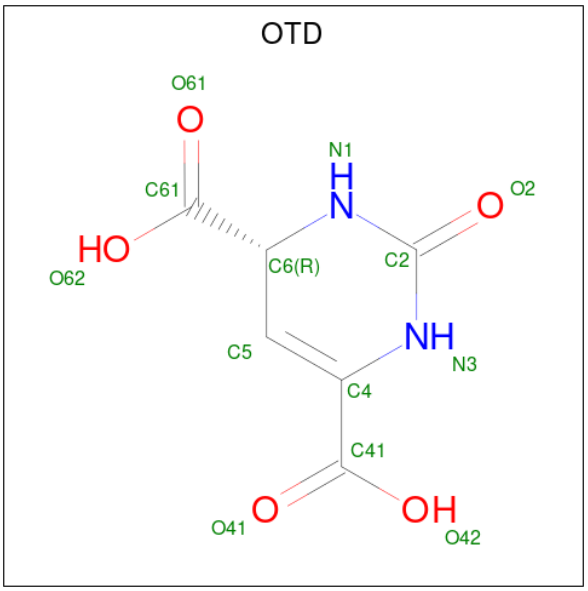
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	KCX	LYS	modified residue	UNP P05020
A	119	VAL	ILE	conflict	UNP P05020
B	102	KCX	LYS	modified residue	UNP P05020
B	119	VAL	ILE	conflict	UNP P05020

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is 2-OXO-1,2,3,6-TETRAHYDROPYRIMIDINE-4,6-DICARBOXYLIC ACID (three-letter code: OTD) (formula: $C_6H_6N_2O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	13	6	2	5	0	0
			Total	C	N	O		
3	B	1	13	6	2	5	0	0

- Molecule 4 is water.

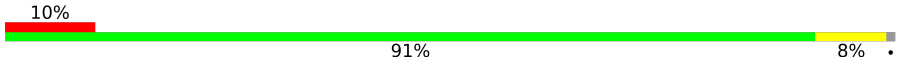
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	206	206	206	0	0
			Total	O		
4	B	140	140	140	0	0

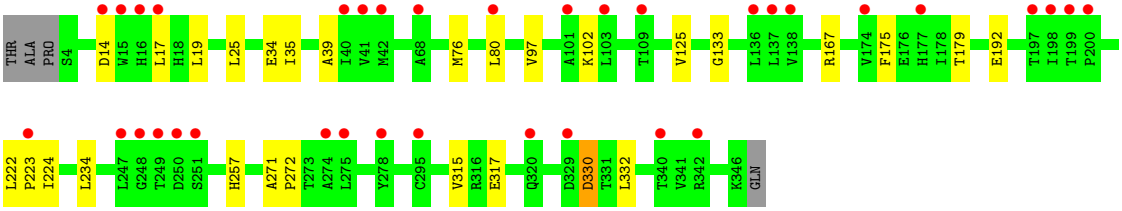
3 Residue-property plots (i)

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: Dihydroorotase

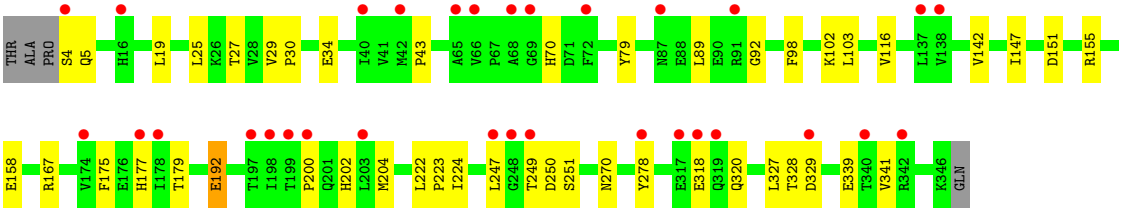
Chain A:





- Molecule 1: Dihydroorotase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.59Å 79.63Å 180.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 29.87 – 2.00	Depositor EDS
% Data completeness (in resolution range)	90.3 (30.00-2.00) 90.3 (29.87-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.68 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.180 , 0.235 0.189 , 0.240	Depositor DCC
R_{free} test set	2384 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å^2)	31.0	Xtriage
Anisotropy	0.679	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 41.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5772	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: OTD, KCX, 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	0.68	0/2754	0.71	1/3747 (0.0%)
1	B	0.59	0/2743	0.70	0/3732
All	All	0.64	0/5497	0.71	1/7479 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	330	ASP	CB-CG-OD1	5.23	123.01	118.30

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	2702	0	2664	14	0
1	B	2694	0	2651	19	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	13	0	4	0	0
3	B	13	0	4	1	0
4	A	206	0	0	1	0
4	B	140	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5772	0	5323	33	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 3.

All (33) 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:B:192:GLU:HG2	4:B:434:HOH:O	1.85	0.75
1:A:17:LEU:HD21	1:A:19:LEU:HD21	1.86	0.57
1:A:192:GLU:CD	1:A:192:GLU:H	2.10	0.55
1:A:80:LEU:HD22	1:A:125:VAL:HG11	1.89	0.54
1:B:200:PRO:O	1:B:204:MET:HG3	2.07	0.54
1:B:34:GLU:HB3	1:B:70:HIS:CE1	2.43	0.54
1:A:315:VAL:CG1	1:A:317:GLU:HG3	2.39	0.53
1:A:315:VAL:HG12	1:A:317:GLU:HG3	1.92	0.50
1:B:92:GLY:HA3	1:B:98:PHE:CD1	2.47	0.50
1:A:257:HIS:ND1	1:A:330:ASP:OD2	2.30	0.49
1:A:14:ASP:O	1:A:39:ALA:HA	2.13	0.48
1:B:151:ASP:O	1:B:155:ARG:HG2	2.14	0.48
1:B:25:LEU:HD12	1:B:29:VAL:HG13	1.94	0.48
1:A:25:LEU:C	1:A:25:LEU:HD23	2.34	0.48
1:B:19:LEU:HD22	1:B:29:VAL:HG12	1.97	0.46
1:B:27:THR:HG21	1:B:327:LEU:HD23	1.97	0.45
1:B:247:LEU:HG	1:B:278:TYR:CE1	2.51	0.45
1:A:76:MET:HB2	1:A:97:VAL:O	2.17	0.44
1:B:43:PRO:HB2	1:B:79:TYR:HB2	2.00	0.44
1:B:29:VAL:N	1:B:30:PRO:CD	2.82	0.43
1:B:142:VAL:HG21	1:B:147:ILE:HD13	2.01	0.42
1:A:222:LEU:HA	1:A:223:PRO:C	2.39	0.42
1:B:222:LEU:HA	1:B:223:PRO:C	2.39	0.42
1:A:133:GLY:HA2	4:A:612:HOH:O	2.19	0.42
1:A:179:THR:HA	1:A:234:LEU:HD11	2.01	0.42
1:B:318:GLU:HA	1:B:341:VAL:O	2.20	0.42
1:A:271:ALA:N	1:A:272:PRO:CD	2.83	0.42
1:B:103:LEU:CD1	1:B:116:VAL:HB	2.50	0.41
1:A:34:GLU:HG3	1:A:35:ILE:HG13	2.02	0.41
1:B:179:THR:HG22	1:B:202:HIS:CD2	2.56	0.41
1:B:249:THR:HB	1:B:270:ASN:OD1	2.21	0.41
1:B:250:ASP:OD2	3:B:410:OTD:H6	2.20	0.41
1:B:250:ASP:O	1:B:251:SER:C	2.60	0.40

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	341/347 (98%)	327 (96%)	14 (4%)	0	100	100
1	B	340/347 (98%)	329 (97%)	10 (3%)	1 (0%)	41	37
All	All	681/694 (98%)	656 (96%)	24 (4%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	177	HIS

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	292/294 (99%)	288 (99%)	4 (1%)	67	72
1	B	291/294 (99%)	279 (96%)	12 (4%)	30	28
All	All	583/588 (99%)	567 (97%)	16 (3%)	44	46

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	167	ARG
1	A	175	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	224	ILE
1	A	332	LEU
1	B	4	SER
1	B	5	GLN
1	B	89	LEU
1	B	158	GLU
1	B	167	ARG
1	B	175	PHE
1	B	192	GLU
1	B	224	ILE
1	B	320	GLN
1	B	328	THR
1	B	329	ASP
1	B	339	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	232	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)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\#\|Z\| > 2$
1	KCX	B	102	1,2	9,11,12	0.94	0	5,12,14	2.10	1 (20%)
1	KCX	A	102	1,2	9,11,12	1.24	0	5,12,14	1.92	1 (20%)

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
1	KCX	B	102	1,2	-	0/9/10/12	-
1	KCX	A	102	1,2	-	0/9/10/12	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	102	KCX	OQ1-CX-NZ	-4.58	117.86	124.96
1	A	102	KCX	OQ1-CX-NZ	-4.05	118.68	124.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
3	OTD	B	410	2	10,13,13	1.30	1 (10%)	14,18,18	1.51	1 (7%)
3	OTD	A	410	2	10,13,13	1.55	2 (20%)	14,18,18	2.08	4 (28%)

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
3	OTD	B	410	2	-	0/7/20/20	0/1/1/1
3	OTD	A	410	2	-	0/7/20/20	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	410	OTD	O62-C61	-2.60	1.22	1.30
3	B	410	OTD	C2-N1	2.05	1.38	1.34
3	A	410	OTD	C2-N3	2.03	1.41	1.37

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	410	OTD	N3-C2-N1	4.04	120.38	116.12
3	A	410	OTD	O41-C41-C4	-3.74	113.38	121.09
3	B	410	OTD	N3-C2-N1	3.19	119.48	116.12
3	A	410	OTD	O2-C2-N1	-3.12	116.71	122.92
3	A	410	OTD	O42-C41-C4	2.42	119.42	114.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	410	OTD	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, 95th 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^2)	Q<0.9
1	A	342/347 (98%)	0.40	35 (10%)	6	6	32, 38, 45, 54	2 (0%)
1	B	342/347 (98%)	0.44	31 (9%)	9	8	29, 39, 47, 56	3 (0%)
All	All	684/694 (98%)	0.42	66 (9%)	8	7	29, 39, 46, 56	5 (0%)

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	249	THR	4.9
1	A	137	LEU	4.6
1	B	137	LEU	4.3
1	A	40	ILE	4.1
1	A	247	LEU	3.5
1	A	17	LEU	3.5
1	A	199	THR	3.4
1	B	68	ALA	3.4
1	A	248	GLY	3.4
1	A	320	GLN	3.3
1	B	198	ILE	3.3
1	A	174	VAL	3.3
1	B	199	THR	3.3
1	B	247	LEU	3.2
1	B	4	SER	3.1
1	B	138	VAL	3.1
1	B	319	GLN	3.1
1	A	109	THR	3.0
1	A	197	THR	3.0
1	B	178	ILE	3.0
1	A	200	PRO	2.9
1	B	249	THR	2.9
1	B	16	HIS	2.9
1	A	329	ASP	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	138	VAL	2.8
1	B	174	VAL	2.8
1	B	318	GLU	2.8
1	B	248	GLY	2.8
1	A	251	SER	2.8
1	A	198	ILE	2.7
1	A	274	ALA	2.7
1	A	16	HIS	2.7
1	B	197	THR	2.7
1	A	101	ALA	2.6
1	B	317	GLU	2.6
1	B	203	LEU	2.6
1	B	69	GLY	2.6
1	B	40	ILE	2.6
1	A	177	HIS	2.4
1	B	91	ARG	2.4
1	B	177	HIS	2.4
1	A	42	MET	2.4
1	B	340	THR	2.4
1	A	278	TYR	2.3
1	A	68	ALA	2.2
1	A	136	LEU	2.2
1	A	340	THR	2.2
1	B	42	MET	2.2
1	A	41	VAL	2.2
1	A	15	TRP	2.2
1	B	65	ALA	2.2
1	A	250	ASP	2.2
1	A	14	ASP	2.2
1	A	223	PRO	2.2
1	B	342	ARG	2.2
1	A	80	LEU	2.1
1	A	275	LEU	2.1
1	B	87	ASN	2.1
1	B	200	PRO	2.1
1	B	66	VAL	2.1
1	A	295	CYS	2.1
1	B	72	PHE	2.0
1	B	329	ASP	2.0
1	A	103	LEU	2.0
1	B	278	TYR	2.0
1	A	342	ARG	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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, 95th 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(Å^2)	Q<0.9
1	KCX	A	102	12/13	0.95	0.23	34,37,38,39	0
1	KCX	B	102	12/13	0.95	0.26	33,36,37,37	0

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, 95th 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(Å^2)	Q<0.9
3	OTD	B	410	13/13	0.95	0.09	39,41,43,46	0
3	OTD	A	410	13/13	0.96	0.10	29,33,35,38	0
2	ZN	B	401	1/1	0.99	0.02	38,38,38,38	0
2	ZN	A	400	1/1	1.00	0.02	30,30,30,30	0
2	ZN	A	401	1/1	1.00	0.01	28,28,28,28	0
2	ZN	B	400	1/1	1.00	0.02	35,35,35,35	0

6.5 Other polymers (i)

There are no such residues in this entry.