Full wwPDB X-ray Structure Validation Report (i)

Jun 16, 2025 – 09:47 PM JST

PDB ID	:	9KP0 / pdb_00009kp0
Title	:	Crystal structure of Oryza sativa HPPD
Authors	:	Nishio, T.; Nishijima, N.; Kubota, T.; Furuhata, Y.; Kato, Y.
Deposited on	:	2024-11-22
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 (1)) were used in the production of this report:

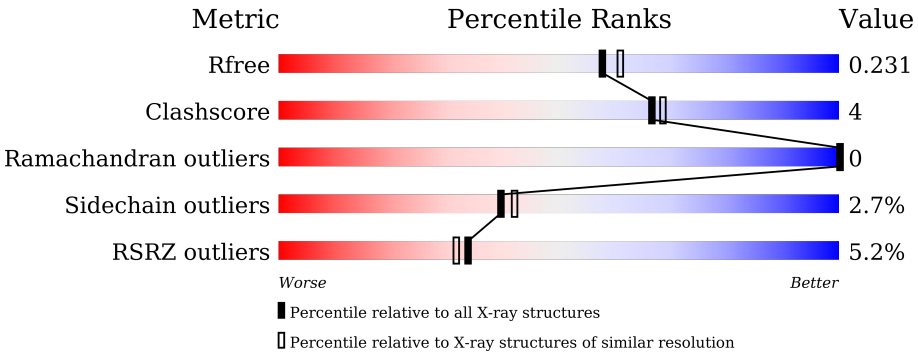
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

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}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (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	446	4% 80% 9% 11%
1	B	446	5% 81% 8% 11%

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
3	ACY	A	502	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6510 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.

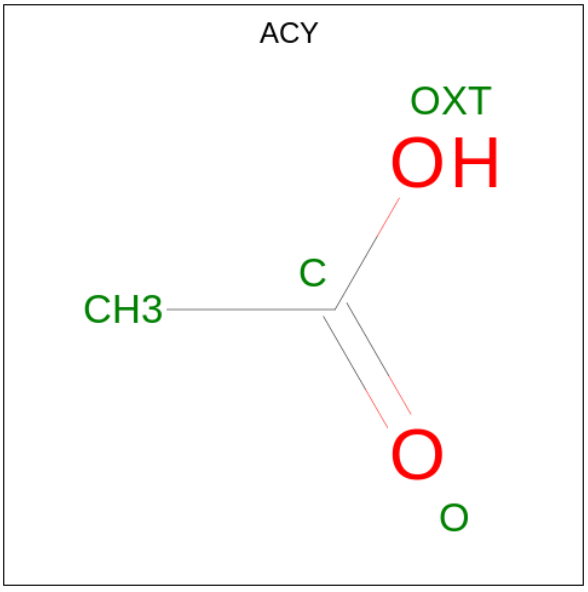
- Molecule 1 is a protein called 4-hydroxyphenylpyruvate dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	399	Total	C	N	O	S	0	0	0
			3009	1900	538	563	8			
1	B	396	Total	C	N	O	S	0	0	0
			2986	1887	534	557	8			

- Molecule 2 is COBALT (II) ION (CCD ID: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Co	0	0
			1	1		
2	B	1	Total	Co	0	0
			1	1		

- Molecule 3 is ACETIC ACID (CCD ID: ACY) (formula: $C_2H_4O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0

- Molecule 4 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	7	Total Zn 7 7	0	0
4	B	5	Total Zn 5 5	0	0

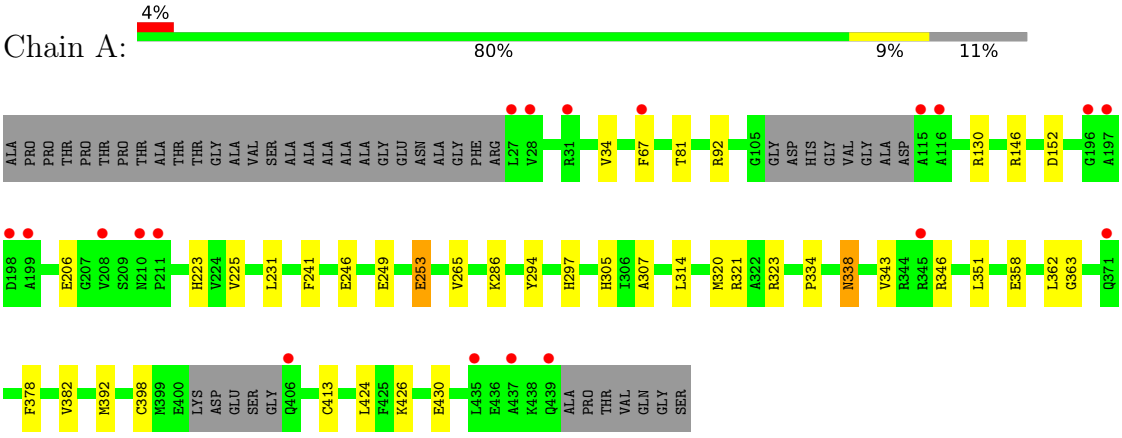
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	256	Total O 256 256	0	0
5	B	241	Total O 241 241	0	0

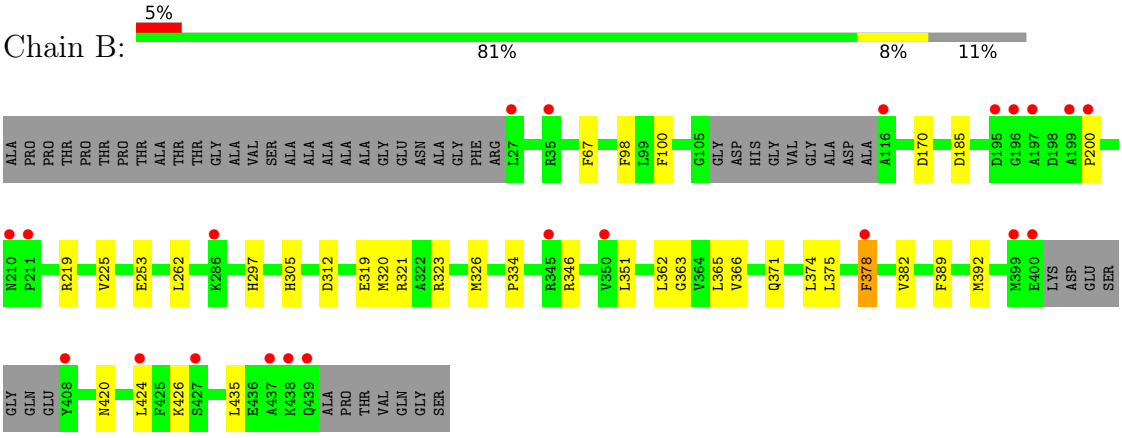
3 Residue-property plots (i)

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: 4-hydroxyphenylpyruvate dioxygenase



• Molecule 1: 4-hydroxyphenylpyruvate dioxygenase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	59.11Å 64.59Å 73.16Å 101.39° 107.96° 97.00°	Depositor
Resolution (Å)	45.94 – 2.00 45.94 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.3 (45.94-2.00) 98.3 (45.94-2.00)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0430	Depositor
R, R_{free}	0.194 , 0.227 0.201 , 0.231	Depositor DCC
R_{free} test set	3288 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	25.9	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 43.2	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	0.94	EDS
Total number of atoms	6510	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.96% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CO, ZN, ACY

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	0.54	0/3080	0.98	2/4176 (0.0%)
1	B	0.53	0/3057	0.97	6/4145 (0.1%)
All	All	0.53	0/6137	0.97	8/8321 (0.1%)

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	4
1	B	0	2
All	All	0	6

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	378	PHE	CA-CB-CG	-7.22	106.58	113.80
1	A	241	PHE	CA-CB-CG	7.07	120.87	113.80
1	B	170	ASP	CA-CB-CG	5.82	118.42	112.60
1	B	67	PHE	CA-CB-CG	5.75	119.55	113.80
1	A	152	ASP	CA-CB-CG	5.54	118.14	112.60
1	B	371	GLN	CB-CA-C	5.42	118.50	109.56
1	B	185	ASP	CA-CB-CG	5.28	117.88	112.60
1	B	312	ASP	CA-CB-CG	5.25	117.85	112.60

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	130	ARG	Sidechain
1	A	146	ARG	Sidechain
1	A	346	ARG	Sidechain
1	A	92	ARG	Sidechain
1	B	219	ARG	Sidechain
1	B	346	ARG	Sidechain

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	3009	0	2898	20	0
1	B	2986	0	2879	21	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	4	0	3	3	0
4	A	7	0	0	0	0
4	B	5	0	0	0	0
5	A	256	0	0	3	0
5	B	241	0	0	2	0
All	All	6510	0	5780	40	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 4.

All (40) 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:378:PHE:CE1	1:B:424:LEU:HD22	2.07	0.90
1:B:420:ASN:O	1:B:424:LEU:HG	1.77	0.85
1:A:430:GLU:HG2	5:A:834:HOH:O	1.79	0.82
1:B:378:PHE:CZ	1:B:424:LEU:HD13	2.25	0.71
1:A:297:HIS:ND1	3:A:502:ACY:OXT	2.27	0.67
1:A:297:HIS:HD1	3:A:502:ACY:C	2.10	0.64
1:B:378:PHE:CD1	1:B:424:LEU:HD22	2.36	0.60
1:B:225:VAL:HG21	1:B:305:HIS:CE1	2.38	0.58
1:B:320:MET:HE1	1:B:392:MET:HE3	1.87	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:225:VAL:HG21	1:A:305:HIS:CE1	2.41	0.56
1:A:67:PHE:HB3	1:B:326:MET:O	2.06	0.56
1:B:378:PHE:CE1	1:B:424:LEU:CD2	2.85	0.55
1:A:320:MET:HE1	1:A:392:MET:HE3	1.92	0.52
1:B:365:LEU:CD1	1:B:424:LEU:HD23	2.42	0.50
1:A:378:PHE:CE2	1:A:424:LEU:HB3	2.46	0.50
1:A:297:HIS:HD1	3:A:502:ACY:CH3	2.26	0.48
1:B:334:PRO:HB3	1:B:363:GLY:CA	2.44	0.47
1:B:378:PHE:CD2	1:B:389:PHE:HE2	2.32	0.47
1:A:253:GLU:CD	5:A:658:HOH:O	2.57	0.47
1:B:297:HIS:ND1	5:B:602:HOH:O	2.35	0.47
1:B:378:PHE:HD2	1:B:389:PHE:CE2	2.32	0.46
1:A:321:ARG:HD2	1:A:362:LEU:HD22	1.96	0.46
1:B:378:PHE:HD2	1:B:389:PHE:HE2	1.64	0.46
1:B:98:PHE:HB3	1:B:100:PHE:CE1	2.51	0.46
1:B:366:VAL:HG22	1:B:375:LEU:CD2	2.46	0.46
1:A:426:LYS:O	1:A:430:GLU:HG3	2.16	0.45
1:B:321:ARG:NH2	1:B:362:LEU:O	2.50	0.44
1:A:338:ASN:HB2	5:A:826:HOH:O	2.17	0.44
1:A:231:LEU:HD11	1:A:265:VAL:HG12	1.99	0.44
1:B:378:PHE:CD1	1:B:378:PHE:N	2.86	0.43
1:B:334:PRO:HB3	1:B:363:GLY:HA2	2.01	0.43
1:A:398:CYS:CB	1:A:413:CYS:HG	2.28	0.43
1:A:294:TYR:CD1	1:A:294:TYR:C	2.98	0.42
1:A:223:HIS:CE1	1:A:307:ALA:CB	3.03	0.41
1:A:314:LEU:HD13	1:A:358:GLU:HB3	2.02	0.41
1:A:81:THR:HG22	1:A:206:GLU:CD	2.45	0.41
1:A:246:GLU:OE2	1:A:249:GLU:OE1	2.38	0.41
1:A:334:PRO:HB3	1:A:363:GLY:HA2	2.03	0.41
1:B:200:PRO:HA	5:B:833:HOH:O	2.20	0.41
1:B:262:LEU:C	1:B:262:LEU:HD12	2.47	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	393/446 (88%)	386 (98%)	7 (2%)	0	100	100
1	B	390/446 (87%)	384 (98%)	6 (2%)	0	100	100
All	All	783/892 (88%)	770 (98%)	13 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	300/328 (92%)	292 (97%)	8 (3%)	40	42
1	B	298/328 (91%)	290 (97%)	8 (3%)	40	42
All	All	598/656 (91%)	582 (97%)	16 (3%)	40	42

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

Mol	Chain	Res	Type
1	A	34	VAL
1	A	253	GLU
1	A	286	LYS
1	A	323	ARG
1	A	338	ASN
1	A	343	VAL
1	A	351	LEU
1	A	382	VAL
1	B	253	GLU
1	B	319	GLU
1	B	323	ARG
1	B	351	LEU
1	B	374	LEU
1	B	382	VAL
1	B	426	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	435	LEU

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	292	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 15 ligands modelled in this entry, 14 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	$\#\|Z\| > 2$
3	ACY	A	502	-	3,3,3	1.17	0	3,3,3	0.69	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.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	ACY	3	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

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^{th} 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	399/446 (89%)	0.08	19 (4%)	36	35	18, 27, 50, 96	0
1	B	396/446 (88%)	0.14	22 (5%)	31	29	18, 27, 53, 94	0
All	All	795/892 (89%)	0.11	41 (5%)	34	32	18, 27, 51, 96	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	116	ALA	4.6
1	A	197	ALA	4.6
1	B	197	ALA	4.2
1	B	211	PRO	4.1
1	A	211	PRO	3.9
1	A	406	GLN	3.5
1	B	424	LEU	3.5
1	A	199	ALA	3.4
1	A	196	GLY	3.2
1	B	437	ALA	3.2
1	B	427	SER	3.1
1	A	115	ALA	3.1
1	A	31	ARG	3.0
1	B	439	GLN	3.0
1	B	350	VAL	3.0
1	A	439	GLN	2.9
1	A	210	ASN	2.7
1	B	196	GLY	2.7
1	A	198	ASP	2.7
1	B	200	PRO	2.6
1	B	199	ALA	2.6
1	A	27	LEU	2.5
1	B	210	ASN	2.5
1	A	67	PHE	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	28	VAL	2.4
1	A	371	GLN	2.3
1	B	438	LYS	2.3
1	B	27	LEU	2.3
1	B	399	MET	2.2
1	A	435	LEU	2.2
1	B	408	TYR	2.2
1	A	116	ALA	2.2
1	B	400	GLU	2.2
1	B	286	LYS	2.2
1	A	345	ARG	2.2
1	A	437	ALA	2.1
1	B	195	ASP	2.1
1	B	378	PHE	2.1
1	B	35	ARG	2.0
1	B	345	ARG	2.0
1	A	208	VAL	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 oligosaccharides 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^{th} 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	ACY	A	502	4/4	0.90	0.11	27,30,31,33	0
4	ZN	A	509	1/1	0.92	0.09	75,75,75,75	0
4	ZN	B	506	1/1	0.95	0.07	76,76,76,76	0
4	ZN	B	503	1/1	0.96	0.06	44,44,44,44	0
4	ZN	B	505	1/1	0.97	0.06	56,56,56,56	0
4	ZN	A	505	1/1	0.97	0.09	50,50,50,50	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	ZN	A	504	1/1	0.98	0.04	37,37,37,37	0
4	ZN	B	502	1/1	0.99	0.03	54,54,54,54	0
4	ZN	A	507	1/1	0.99	0.02	29,29,29,29	0
4	ZN	A	508	1/1	0.99	0.02	39,39,39,39	0
4	ZN	A	503	1/1	0.99	0.03	46,46,46,46	0
2	CO	B	501	1/1	1.00	0.03	19,19,19,19	0
4	ZN	B	504	1/1	1.00	0.01	24,24,24,24	0
4	ZN	A	506	1/1	1.00	0.02	27,27,27,27	0
2	CO	A	501	1/1	1.00	0.01	20,20,20,20	0

6.5 Other polymers (i)

There are no such residues in this entry.