Full wwPDB X-ray Structure Validation Report (i)

Feb 19, 2016 – 07:07 PM GMT

PDB ID	:	3X44
Title	:	Crystal structure of O-ureido-L-serine-bound K43A mutant of O-ureido-L-serine synthase
Authors	:	Matoba, Y.; Uda, N.; Oda, K.; Sugiyama, M.
Deposited on	:	2015-03-13
Resolution	:	1.90 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

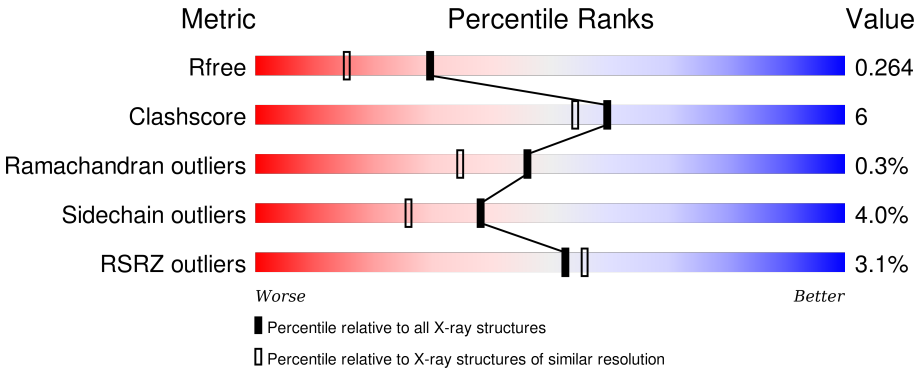
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 1.90 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}	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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 >=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 <=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	332	4% 82% 15% •
1	B	332	2% 80% 14% • 5%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5307 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.

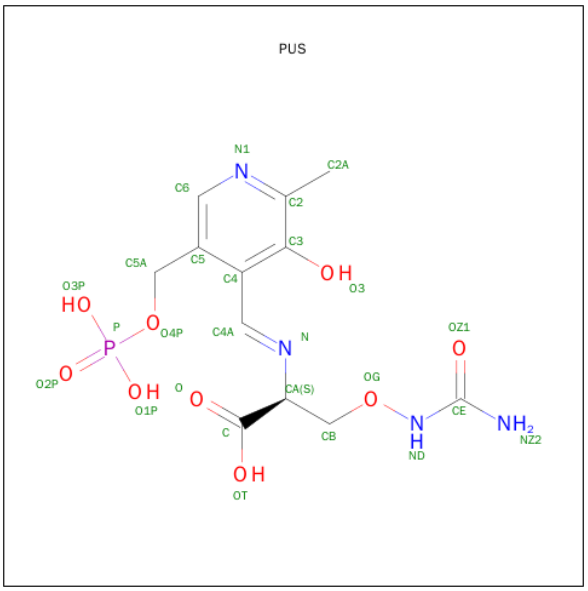
- Molecule 1 is a protein called O-ureido-L-serine synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S	0	0	0
			2410	1518	423	461	8			
1	B	316	Total	C	N	O	S	0	0	0
			2377	1499	418	452	8			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	43	ALA	LYS	ENGINEERED MUTATION	UNP D2Z027
A	325	LEU	-	EXPRESSION TAG	UNP D2Z027
A	326	GLU	-	EXPRESSION TAG	UNP D2Z027
A	327	HIS	-	EXPRESSION TAG	UNP D2Z027
A	328	HIS	-	EXPRESSION TAG	UNP D2Z027
A	329	HIS	-	EXPRESSION TAG	UNP D2Z027
A	330	HIS	-	EXPRESSION TAG	UNP D2Z027
A	331	HIS	-	EXPRESSION TAG	UNP D2Z027
A	332	HIS	-	EXPRESSION TAG	UNP D2Z027
B	43	ALA	LYS	ENGINEERED MUTATION	UNP D2Z027
B	325	LEU	-	EXPRESSION TAG	UNP D2Z027
B	326	GLU	-	EXPRESSION TAG	UNP D2Z027
B	327	HIS	-	EXPRESSION TAG	UNP D2Z027
B	328	HIS	-	EXPRESSION TAG	UNP D2Z027
B	329	HIS	-	EXPRESSION TAG	UNP D2Z027
B	330	HIS	-	EXPRESSION TAG	UNP D2Z027
B	331	HIS	-	EXPRESSION TAG	UNP D2Z027
B	332	HIS	-	EXPRESSION TAG	UNP D2Z027

- Molecule 2 is (E)-O-(CARBAMOYLAMINO)-N-({3-HYDROXY-2-METHYL-5-[(PHOSPHONOOXY)METHYL]PYRIDIN-4-YL}METHYLIDENE)-L-SERINE (three-letter code: PUS) (formula: $C_{12}H_{17}N_4O_9P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			26	12	4	9	1		
2	B	1	Total	C	N	O	P	0	0
			26	12	4	9	1		

- Molecule 3 is water.

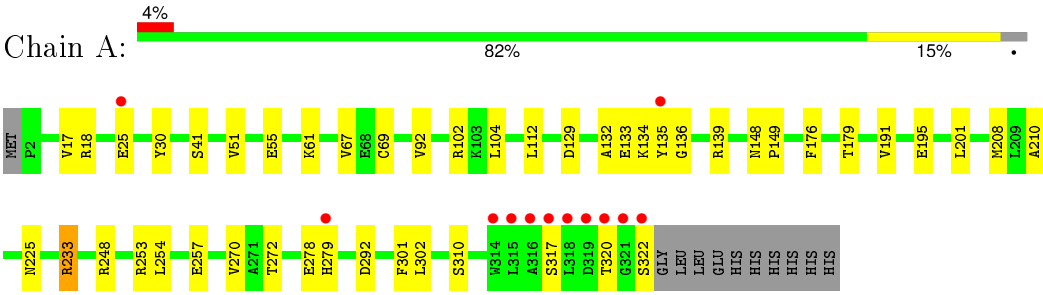
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	222	Total	O	0	0
			222	222		
3	B	246	Total	O	0	0
			246	246		

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 errors 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.

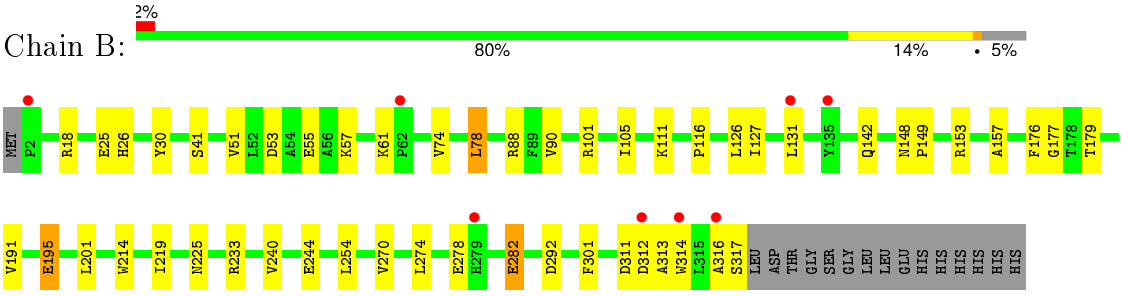
- Molecule 1: O-ureido-L-serine synthase

Chain A:



- Molecule 1: O-ureido-L-serine synthase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	113.35Å 80.26Å 74.67Å 90.00° 117.96° 90.00°	Depositor
Resolution (Å)	29.87 – 1.90 34.28 – 1.90	Depositor EDS
% Data completeness (in resolution range)	92.6 (29.87-1.90) 94.5 (34.28-1.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.15 (at 1.89Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.212 , 0.260 0.215 , 0.264	Depositor DCC
R_{free} test set	2240 reflections (5.07%)	DCC
Wilson B-factor (Å^2)	15.2	Xtriage
Anisotropy	1.482	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 44221 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5307	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.29% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: PUS

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	0.33	0/2456	0.61	0/3336
1	B	0.33	0/2423	0.61	1/3291 (0.0%)
All	All	0.33	0/4879	0.61	1/6627 (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	B	177	GLY	N-CA-C	-5.01	100.58	113.10

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	2410	0	2398	31	0
1	B	2377	0	2368	25	0
2	A	26	0	13	0	0
2	B	26	0	13	0	0
3	A	222	0	0	6	1
3	B	246	0	0	3	1
All	All	5307	0	4792	55	2

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 6.

All (55) 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:191:VAL:HG21	1:B:191:VAL:HG21	1.51	0.91
1:A:208:MET:HE3	1:A:212:GLY:HA2	1.68	0.76
1:B:233:ARG:HD2	3:B:622:HOH:O	1.95	0.65
1:A:148:ASN:HB3	1:A:149:PRO:CD	2.30	0.62
1:A:254:LEU:HD11	1:A:270:VAL:HG21	1.81	0.61
1:A:51:VAL:O	1:A:55:GLU:HG3	2.00	0.61
1:B:26:HIS:HB2	1:B:282:GLU:HG2	1.84	0.60
1:A:211:ARG:HD3	3:A:697:HOH:O	2.01	0.59
1:A:129:ASP:O	1:A:133:GLU:HG3	2.01	0.59
1:A:208:MET:HE2	3:A:569:HOH:O	2.02	0.58
1:B:18:ARG:HB2	1:B:30:TYR:CE2	2.41	0.56
1:A:214:TRP:CD2	1:A:225:ASN:HB3	2.43	0.54
1:B:148:ASN:HB3	1:B:149:PRO:HD3	1.90	0.53
1:B:313:ALA:O	1:B:316:ALA:HB3	2.10	0.52
1:A:41:SER:HB2	1:A:292:ASP:HB3	1.93	0.50
1:B:254:LEU:HD11	1:B:270:VAL:HG21	1.94	0.49
1:A:67:VAL:HG13	1:A:139:ARG:HD2	1.95	0.49
1:B:101:ARG:O	1:B:105:ILE:HG13	2.13	0.49
1:B:148:ASN:HB3	1:B:149:PRO:CD	2.43	0.48
1:B:219:ILE:HD11	1:B:244:GLU:HA	1.95	0.48
1:A:18:ARG:HB2	1:A:30:TYR:CE2	2.49	0.48
1:A:148:ASN:CB	1:A:149:PRO:CD	2.92	0.48
1:A:176:PHE:CD1	1:A:179:THR:HG22	2.49	0.47
1:A:233:ARG:HB3	3:A:562:HOH:O	2.14	0.47
1:A:102:ARG:HG2	1:A:112:LEU:HD21	1.97	0.47
1:B:312:ASP:HB3	3:B:745:HOH:O	2.15	0.47
1:B:311:ASP:HB3	1:B:314:TRP:HB3	1.97	0.46
1:B:53:ASP:OD2	1:B:57:LYS:HE2	2.16	0.46
1:B:214:TRP:CD2	1:B:225:ASN:HB3	2.50	0.46
1:A:210:ALA:HB1	3:A:685:HOH:O	2.16	0.46
1:B:51:VAL:O	1:B:55:GLU:HG3	2.16	0.45
1:B:127:ILE:O	1:B:131:LEU:HD13	2.17	0.45
1:A:69:CYS:HA	1:A:92:VAL:HB	1.99	0.45
1:A:233:ARG:HG3	3:A:560:HOH:O	2.17	0.45
1:B:90:VAL:HG22	1:B:111:LYS:HB2	1.99	0.44
1:A:208:MET:CE	1:A:212:GLY:HA2	2.45	0.44
1:B:195:GLU:HB2	3:B:611:HOH:O	2.17	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:208:MET:O	1:A:212:GLY:HA2	2.17	0.44
1:A:148:ASN:HB3	1:A:149:PRO:HD3	1.98	0.44
1:A:195:GLU:HB2	3:A:604:HOH:O	2.18	0.44
1:B:78:LEU:HA	1:B:78:LEU:HD12	1.90	0.43
1:A:104:LEU:HA	1:A:104:LEU:HD23	1.81	0.43
1:A:248:ARG:HD3	1:A:302:LEU:HA	2.01	0.42
1:A:253:ARG:O	1:A:257:GLU:HB2	2.19	0.42
1:B:153:ARG:O	1:B:157:ALA:HB3	2.20	0.42
1:B:270:VAL:O	1:B:274:LEU:HG	2.20	0.42
1:A:134:LYS:HD3	1:A:135:TYR:CZ	2.56	0.41
1:B:201:LEU:HA	1:B:240:VAL:O	2.20	0.41
1:A:201:LEU:HD22	1:A:272:THR:OG1	2.19	0.41
1:B:74:VAL:HG21	1:B:142:GLN:OE1	2.21	0.41
1:B:41:SER:HB2	1:B:292:ASP:HB3	2.03	0.41
1:A:69:CYS:HA	1:A:92:VAL:O	2.21	0.41
1:B:176:PHE:CD1	1:B:179:THR:HG22	2.56	0.41
1:A:132:ALA:O	1:A:136:GLY:N	2.50	0.41
1:A:134:LYS:HD3	1:A:135:TYR:CE1	2.57	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:614:HOH:O	3:A:614:HOH:O[2_555]	1.86	0.34
3:B:732:HOH:O	3:B:732:HOH:O[2_556]	1.92	0.28

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	319/332 (96%)	304 (95%)	14 (4%)	1 (0%)	46 35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	314/332 (95%)	300 (96%)	13 (4%)	1 (0%)	46 35
All	All	633/664 (95%)	604 (95%)	27 (4%)	2 (0%)	46 35

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	320	THR
1	B	116	PRO

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	250/260 (96%)	240 (96%)	10 (4%)	38 26
1	B	246/260 (95%)	236 (96%)	10 (4%)	37 25
All	All	496/520 (95%)	476 (96%)	20 (4%)	38 26

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	VAL
1	A	25	GLU
1	A	61	LYS
1	A	233	ARG
1	A	278	GLU
1	A	279	HIS
1	A	301	PHE
1	A	310	SER
1	A	317	SER
1	A	322	SER
1	B	25	GLU
1	B	61	LYS
1	B	78	LEU
1	B	88	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	126	LEU
1	B	195	GLU
1	B	278	GLU
1	B	282	GLU
1	B	301	PHE
1	B	317	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	171	HIS
1	A	279	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

2 ligands 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	$\#\|Z\| > 2$
2	PUS	A	400	-	21,26,26	3.65	12 (57%)	23,36,36	1.27	1 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	PUS	B	401	-	21,26,26	3.35	13 (61%)	23,36,36	1.33	2 (8%)

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	PUS	A	400	-	-	0/15/21/21	0/1/1/1
2	PUS	B	401	-	-	0/15/21/21	0/1/1/1

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	PUS	P-O4P	-3.89	1.48	1.59
2	A	400	PUS	P-O4P	-3.61	1.49	1.59
2	A	400	PUS	OZ1-CE	-3.24	1.19	1.24
2	B	401	PUS	OZ1-CE	-2.69	1.20	1.24
2	A	400	PUS	P-O1P	-2.68	1.45	1.54
2	B	401	PUS	P-O1P	-2.52	1.46	1.54
2	B	401	PUS	P-O3P	-2.51	1.46	1.54
2	A	400	PUS	P-O3P	-2.30	1.46	1.54
2	B	401	PUS	C4-C4A	2.09	1.50	1.46
2	B	401	PUS	C6-N1	2.12	1.38	1.34
2	A	400	PUS	C2-N1	2.40	1.38	1.33
2	B	401	PUS	C2-N1	2.58	1.38	1.33
2	A	400	PUS	C4-C4A	2.99	1.51	1.46
2	B	401	PUS	CA-N	3.49	1.49	1.46
2	B	401	PUS	C4A-N	3.88	1.34	1.27
2	A	400	PUS	C4-C5	3.96	1.47	1.42
2	B	401	PUS	C4-C5	4.09	1.47	1.42
2	A	400	PUS	C4A-N	4.10	1.34	1.27
2	B	401	PUS	C6-C5	4.16	1.46	1.37
2	A	400	PUS	C6-C5	4.61	1.47	1.37
2	A	400	PUS	C4-C3	4.88	1.47	1.40
2	B	401	PUS	C4-C3	5.50	1.48	1.40
2	A	400	PUS	CA-N	6.39	1.52	1.46
2	B	401	PUS	C3-C2	9.25	1.47	1.40
2	A	400	PUS	C3-C2	10.03	1.47	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	PUS	P-O4P-C5A	-3.53	112.62	120.79
2	A	400	PUS	P-O4P-C5A	-3.35	113.04	120.79
2	B	401	PUS	O3-C3-C2	2.06	120.55	117.53

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

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

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	321/332 (96%)	0.10	12 (3%) 45 49	13, 21, 37, 79	0
1	B	316/332 (95%)	0.06	8 (2%) 61 64	15, 21, 34, 54	0
All	All	637/664 (95%)	0.08	20 (3%) 52 56	13, 21, 37, 79	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	316	ALA	5.2
1	A	318	LEU	4.4
1	A	320	THR	4.4
1	A	314	TRP	3.3
1	B	135	TYR	3.2
1	B	279	HIS	3.0
1	B	312	ASP	3.0
1	A	315	LEU	3.0
1	A	322	SER	2.9
1	B	314	TRP	2.8
1	A	321	GLY	2.7
1	B	316	ALA	2.6
1	A	25	GLU	2.5
1	B	2	PRO	2.4
1	B	62	PRO	2.4
1	A	135	TYR	2.2
1	B	131	LEU	2.1
1	A	279	HIS	2.1
1	A	319	ASP	2.1
1	A	317	SER	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^{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	LLDF	B-factors(Å^2)	Q<0.9
2	PUS	A	400	26/26	0.94	0.12	1.28	15,18,24,25	0
2	PUS	B	401	26/26	0.94	0.11	0.55	16,19,25,26	0

6.5 Other polymers (i)

There are no such residues in this entry.