Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 11:46 AM GMT

PDB ID	:	3PYX
Title	:	Crystals Structure of Aspartate beta-Semialdehyde Dehydrogenase complex with NADP and 2-aminoterephthalate
Authors	:	Pavlovsky, A.G.; Viola, R.E.
Deposited on	:	2010-12-13
Resolution	:	1.60 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:

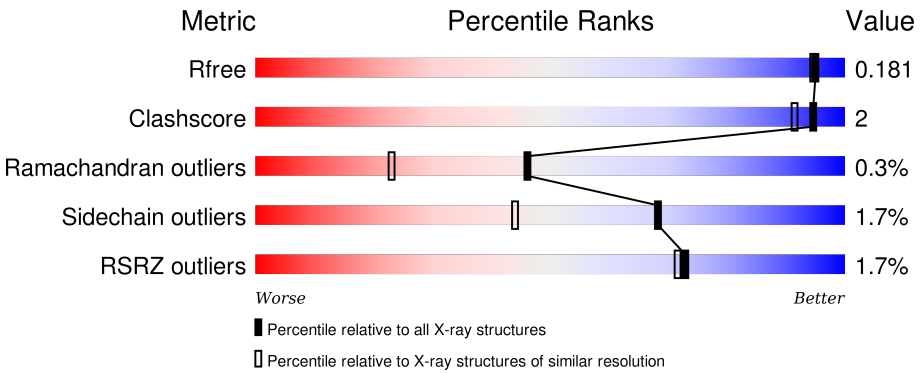
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.60 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	2475 (1.60-1.60)
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (1.60-1.60)
RSRZ outliers	91569	2479 (1.60-1.60)

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	366	2% 92% 5%
1	B	366	% 95%

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	NA	A	368	-	-	-	X
4	EDO	A	369	-	-	-	X
4	EDO	B	368	-	-	-	X
5	12T	B	371	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6334 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.

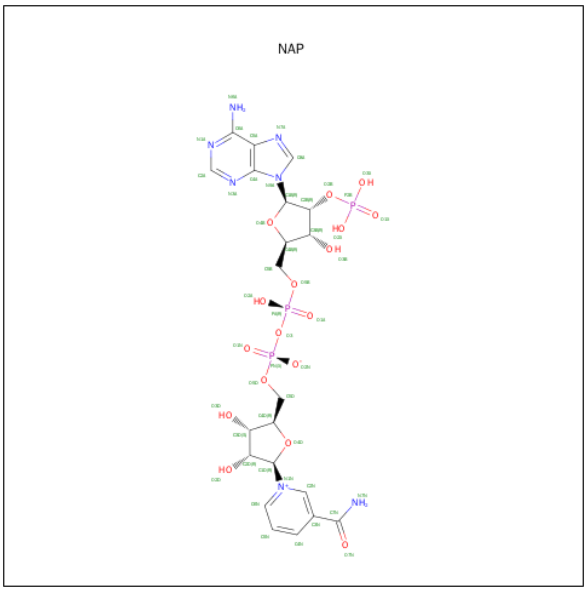
- Molecule 1 is a protein called Aspartate-semialdehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	357	2730	1726	459	531	14	0	2	0
1	B	359	2757	1745	461	534	17	0	4	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	LEU	-	EXPRESSION TAG	UNP A5MTN0
A	360	GLU	-	EXPRESSION TAG	UNP A5MTN0
A	361	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	362	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	363	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	364	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	365	HIS	-	EXPRESSION TAG	UNP A5MTN0
A	366	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	359	LEU	-	EXPRESSION TAG	UNP A5MTN0
B	360	GLU	-	EXPRESSION TAG	UNP A5MTN0
B	361	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	362	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	363	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	364	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	365	HIS	-	EXPRESSION TAG	UNP A5MTN0
B	366	HIS	-	EXPRESSION TAG	UNP A5MTN0

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$).

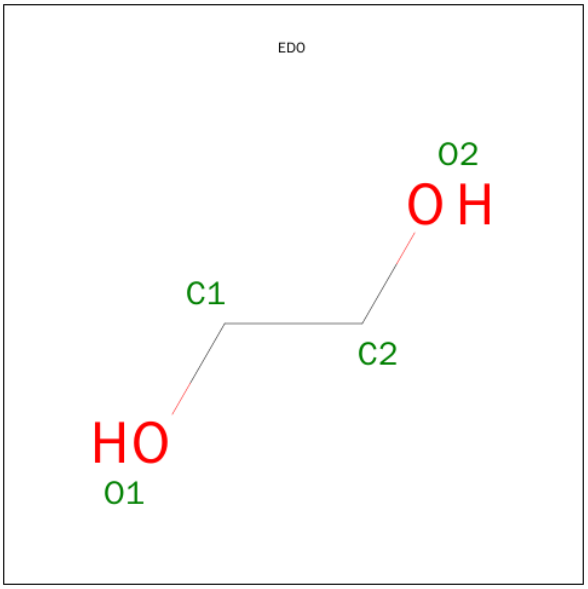


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	48	21	7	17	3	0	0
			Total	C	N	O	P		
2	B	1	48	21	7	17	3	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

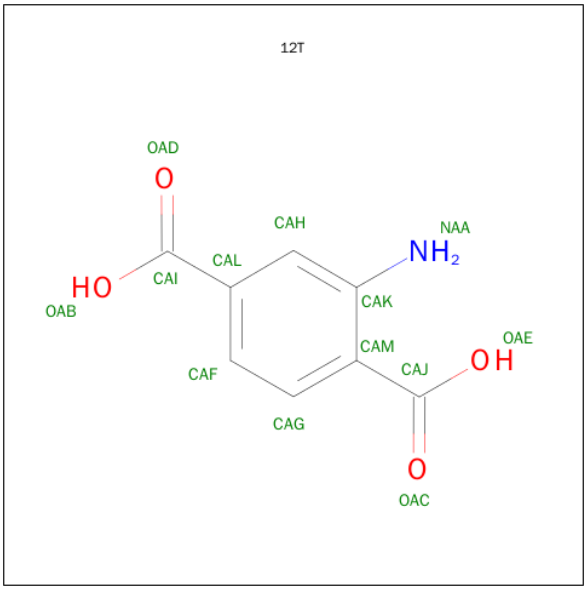
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
3	A	1	1	1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 5 is 2-AMINOBENZENE-1,4-DICARBOXYLIC ACID (three-letter code: 12T) (formula: $C_8H_7NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	B	1	13	8	1	4	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	323	323	323	0	0

Continued on next page...

Continued from previous page...

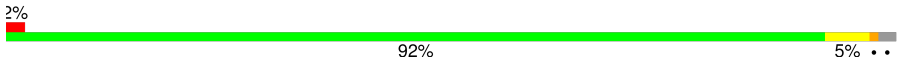
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	B	386	386	386	0	0

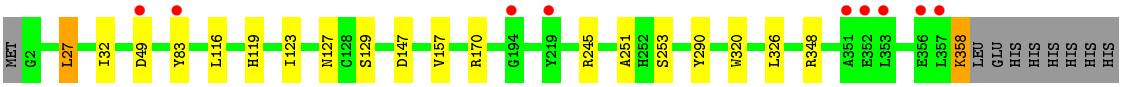
3 Residue-property plots (i)

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: Aspartate-semialdehyde dehydrogenase

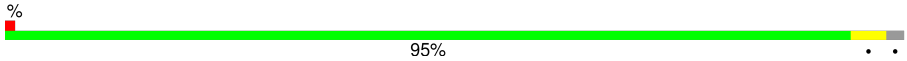
Chain A:

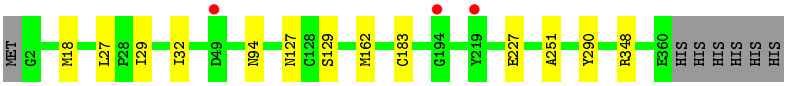




- Molecule 1: Aspartate-semialdehyde dehydrogenase

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.10Å 99.28Å 64.54Å 90.00° 101.02° 90.00°	Depositor
Resolution (Å)	29.33 – 1.60 29.33 – 1.60	Depositor EDS
% Data completeness (in resolution range)	92.2 (29.33-1.60) 92.3 (29.33-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.76 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.160 , 0.182 0.158 , 0.181	Depositor DCC
R_{free} test set	4491 reflections (5.22%)	DCC
Wilson B-factor (Å^2)	17.3	Xtriage
Anisotropy	0.672	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 44.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 90545 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6334	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.81% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, NAP, 12T, EDO

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	0.55	0/2785	0.66	2/3784 (0.1%)
1	B	0.61	2/2818 (0.1%)	0.67	0/3827
All	All	0.58	2/5603 (0.0%)	0.67	2/7611 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	183[A]	CYS	CB-SG	-5.38	1.73	1.81
1	B	183[B]	CYS	CB-SG	-5.38	1.73	1.81

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	170	ARG	NE-CZ-NH2	-5.34	117.63	120.30
1	A	27	LEU	CA-CB-CG	5.19	127.24	115.30

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	2730	0	2740	8	0
1	B	2757	0	2775	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	48	0	25	0	0
2	B	48	0	25	0	0
3	A	1	0	0	0	0
4	A	16	0	24	1	0
4	B	12	0	18	0	0
5	B	13	0	5	1	0
6	A	323	0	0	0	0
6	B	386	0	0	2	0
All	All	6334	0	5612	19	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 2.

All (19) 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)
5:B:371:12T:OAC	6:B:630:HOH:O	1.62	1.16
1:B:18[B]:MET:CE	1:B:94:ASN:ND2	2.33	0.92
1:B:18[B]:MET:CE	1:B:94:ASN:HD21	1.86	0.87
1:B:18[B]:MET:HE3	1:B:94:ASN:HD21	1.39	0.86
1:B:18[B]:MET:HE1	1:B:94:ASN:ND2	1.93	0.84
1:B:18[B]:MET:HE1	1:B:94:ASN:HD22	1.55	0.71
1:B:18[B]:MET:HE3	1:B:94:ASN:ND2	2.00	0.69
1:B:127:ASN:OD1	1:B:129:SER:OG	2.20	0.53
1:A:116:LEU:HD23	1:A:123:ILE:HD12	1.92	0.52
1:A:119:HIS:CB	1:A:123:ILE:HD11	2.45	0.47
1:A:119:HIS:HB3	1:A:123:ILE:HD11	1.97	0.46
1:B:162[B]:MET:HE3	6:B:618:HOH:O	2.17	0.45
1:B:127:ASN:HB3	1:B:227:GLU:OE1	2.17	0.45
1:A:83:TYR:OH	1:A:358:LYS:HE2	2.17	0.44
1:A:253:SER:HB3	1:A:320:TRP:CZ2	2.52	0.44
1:A:326:LEU:HD12	4:A:370:EDO:H22	1.99	0.44
1:A:127:ASN:OD1	1:A:129[A]:SER:OG	2.25	0.42
1:A:157:VAL:HA	1:A:245:ARG:HD2	2.02	0.41
1:B:29:ILE:HG21	1:B:32:ILE:HG12	2.02	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	357/366 (98%)	341 (96%)	15 (4%)	1 (0%)	46	23
1	B	361/366 (99%)	347 (96%)	13 (4%)	1 (0%)	46	23
All	All	718/732 (98%)	688 (96%)	28 (4%)	2 (0%)	46	23

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	251	ALA
1	B	251	ALA

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	295/302 (98%)	288 (98%)	7 (2%)	57	27
1	B	299/302 (99%)	296 (99%)	3 (1%)	82	67
All	All	594/604 (98%)	584 (98%)	10 (2%)	68	44

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	LEU
1	A	32	ILE
1	A	49	ASP
1	A	147	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	290	TYR
1	A	348	ARG
1	A	358	LYS
1	B	27	LEU
1	B	290	TYR
1	B	348	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	ASN
1	A	114	HIS
1	A	132	GLN
1	B	94	ASN
1	B	114	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 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	NAP	A	367	-	42,52,52	1.64	3 (7%)	54,80,80	1.97	5 (9%)
4	EDO	A	369	-	3,3,3	0.55	0	2,2,2	0.12	0
4	EDO	A	370	-	3,3,3	0.51	0	2,2,2	0.35	0
4	EDO	A	371	-	3,3,3	0.60	0	2,2,2	0.25	0
4	EDO	A	372	-	3,3,3	0.53	0	2,2,2	0.33	0
2	NAP	B	367	-	42,52,52	1.69	4 (9%)	54,80,80	1.86	5 (9%)
4	EDO	B	368	-	3,3,3	0.58	0	2,2,2	0.28	0
4	EDO	B	369	-	3,3,3	0.54	0	2,2,2	0.37	0
4	EDO	B	370	-	3,3,3	0.45	0	2,2,2	0.42	0
5	12T	B	371	-	7,13,13	0.68	0	9,18,18	1.28	2 (22%)

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	NAP	A	367	-	-	0/27/67/67	0/5/5/5
4	EDO	A	369	-	-	0/1/1/1	0/0/0/0
4	EDO	A	370	-	-	0/1/1/1	0/0/0/0
4	EDO	A	371	-	-	0/1/1/1	0/0/0/0
4	EDO	A	372	-	-	0/1/1/1	0/0/0/0
2	NAP	B	367	-	-	0/27/67/67	0/5/5/5
4	EDO	B	368	-	-	0/1/1/1	0/0/0/0
4	EDO	B	369	-	-	0/1/1/1	0/0/0/0
4	EDO	B	370	-	-	0/1/1/1	0/0/0/0
5	12T	B	371	-	-	0/0/8/8	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	367	NAP	PA-O2A	-2.05	1.46	1.54
2	B	367	NAP	C2A-N1A	2.46	1.38	1.33
2	A	367	NAP	C2A-N1A	2.64	1.38	1.33
2	B	367	NAP	C2A-N3A	3.50	1.38	1.32
2	A	367	NAP	C2A-N3A	3.54	1.38	1.32
2	A	367	NAP	O7N-C7N	8.65	1.42	1.24
2	B	367	NAP	O7N-C7N	8.72	1.42	1.24

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	367	NAP	N3A-C2A-N1A	-12.05	119.67	128.89
2	B	367	NAP	N3A-C2A-N1A	-10.32	121.00	128.89
2	B	367	NAP	O7N-C7N-C3N	-3.19	116.10	119.59
2	A	367	NAP	O7N-C7N-C3N	-3.11	116.19	119.59
2	B	367	NAP	C4A-C5A-N7A	-2.72	106.97	109.48
2	B	367	NAP	C1B-N9A-C4A	-2.69	122.89	126.94
2	A	367	NAP	C1B-N9A-C4A	-2.23	123.58	126.94
5	B	371	12T	CAL-CAH-CAK	-2.20	119.18	121.84
2	A	367	NAP	O4B-C1B-N9A	2.41	113.15	108.10
5	B	371	12T	CAF-CAL-CAH	2.59	121.60	118.17
2	B	367	NAP	O4D-C1D-N1N	2.71	111.11	108.13
2	A	367	NAP	O4D-C1D-N1N	2.93	111.35	108.13

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	370	EDO	1	0
5	B	371	12T	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	357/366 (97%)	-0.12	9 (2%)	61	58	12, 20, 36, 47	0
1	B	359/366 (98%)	-0.28	3 (0%)	87	87	12, 17, 28, 35	0
All	All	716/732 (97%)	-0.20	12 (1%)	73	71	12, 18, 33, 47	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	49	ASP	3.7
1	B	49	ASP	3.6
1	B	194	GLY	3.6
1	A	194	GLY	3.4
1	A	353	LEU	3.2
1	A	351	ALA	3.2
1	A	356	GLU	3.1
1	A	83	TYR	2.5
1	B	219	TYR	2.2
1	A	357	LEU	2.2
1	A	219	TYR	2.1
1	A	352	GLU	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, 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	LLDF	B-factors(Å^2)	Q<0.9
5	12T	B	371	13/13	0.75	0.37	17.63	24,27,28,29	13
4	EDO	A	369	4/4	0.97	0.13	7.90	18,18,19,20	0
3	NA	A	368	1/1	0.99	0.15	5.35	19,19,19,19	0
4	EDO	B	368	4/4	0.98	0.15	5.34	17,18,18,20	0
4	EDO	A	372	4/4	0.82	0.12	1.73	50,50,50,52	0
4	EDO	B	370	4/4	0.84	0.12	1.03	47,47,49,49	0
4	EDO	A	370	4/4	0.97	0.08	0.88	30,32,33,33	0
4	EDO	A	371	4/4	0.86	0.11	0.81	34,36,37,37	0
2	NAP	B	367	48/48	0.96	0.09	0.42	18,25,36,37	0
4	EDO	B	369	4/4	0.95	0.08	0.33	24,26,27,28	0
2	NAP	A	367	48/48	0.95	0.09	-0.05	27,35,39,40	0

6.5 Other polymers (i)

There are no such residues in this entry.