Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:07 AM GMT

PDB ID	:	3HB5
Title	:	Binary and ternary crystal structures of a novel inhibitor of 17 beta-HSD type 1: a lead compound for breast cancer therapy
Authors	:	Mazumdar, M.; Lin, S.-X.
Deposited on	:	2009-05-04
Resolution	:	2.00 Å(reported)

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

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	Percentile Ranks	Value
Rfree		0.250
Clashscore		15
Ramachandran outliers		1.1%
Sidechain outliers		11.9%
RSRZ outliers		7.0%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
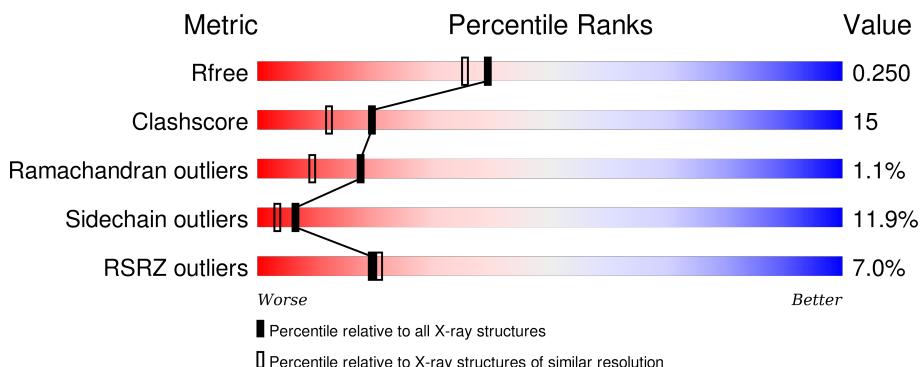
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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 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	X	327	6% — 67% / 15% / 5% / 13%
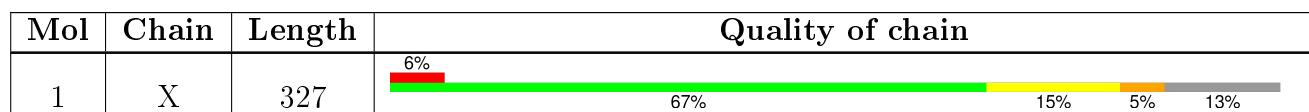
The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAP	X	360	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2349 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 Estradiol 17-beta-dehydrogenase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	X	284	2177	1382	385	398	12	0	0	0

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$).
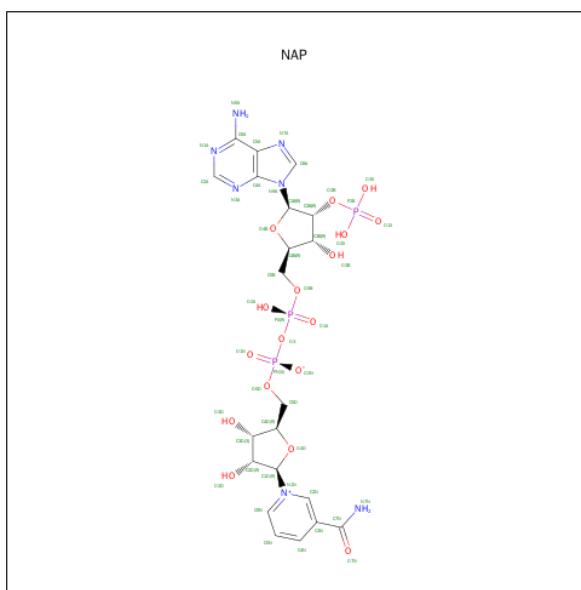
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	X	1	48	21	7	17	3	0	0

- Molecule 3 is 3-{[(9BETA,14BETA,16ALPHA,17ALPHA)-3,17-DIHYDROXYESTRA-1,3,5(10)-TRIEN-16-YL]METHYL}BENZAMIDE (three-letter code: E2B) (formula: $C_{26}H_{31}NO_3$).
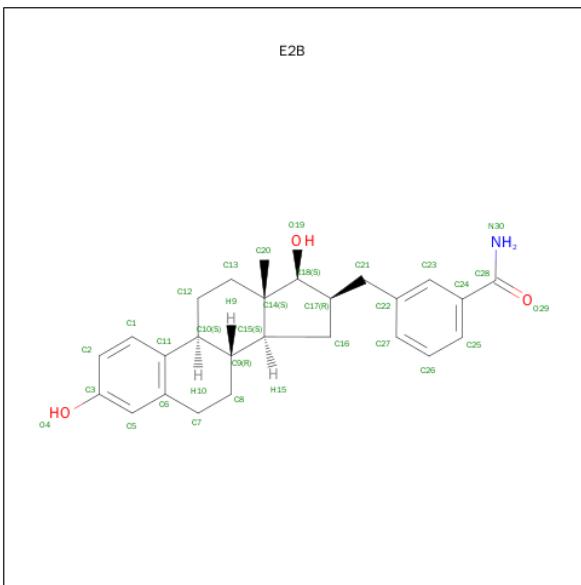
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	X	1	Total 30	C 26	N 1	O 3	30	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	X	94	Total 94	O 94	0	0

3 Residue-property plots ⓘ

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: Estradiol 17-beta-dehydrogenase 1

Chain X:
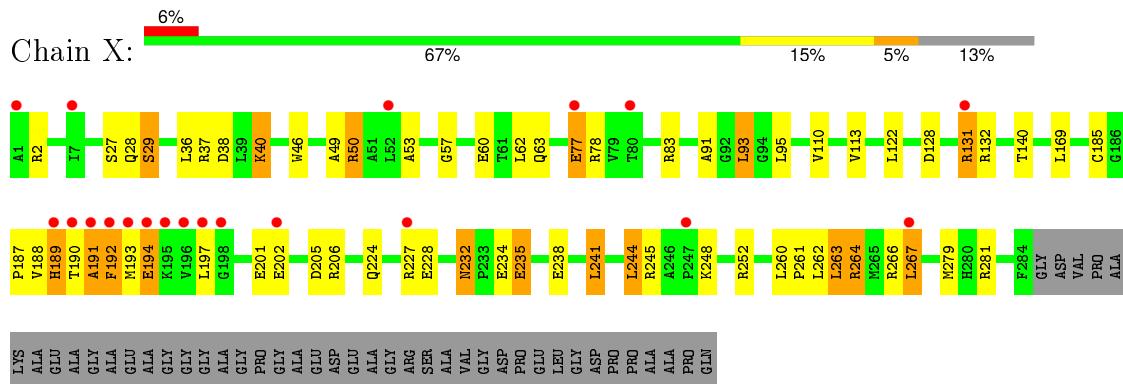
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	123.20Å 44.00Å 60.80Å 90.00° 100.30° 90.00°	Depositor
Resolution (Å)	19.40 – 2.00 19.94 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.40-2.00) 98.8 (19.94-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.83 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.202 , 0.248 0.211 , 0.250	Depositor DCC
R_{free} test set	1083 reflections (5.26%)	DCC
Wilson B-factor (Å^2)	25.7	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 60.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 21662 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2349	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 6.80% 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: NAP, E2B

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	X	1.04	2/2218 (0.1%)	0.96	5/3009 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	X	235	GLU	CB-CG	-5.85	1.41	1.52
1	X	60	GLU	CG-CD	5.11	1.59	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	X	244	LEU	CA-CB-CG	7.22	131.90	115.30
1	X	252	ARG	CG-CD-NE	-5.61	100.02	111.80
1	X	245	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	X	241	LEU	CA-CB-CG	5.14	127.12	115.30
1	X	245	ARG	NE-CZ-NH2	-5.01	117.80	120.30

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	X	2177	0	2218	60	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	X	48	0	25	20	0
3	X	30	0	29	0	0
4	X	94	0	0	26	0
All	All	2349	0	2272	67	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 15.

All (67) 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:X:50:ARG:HA	4:X:380:HOH:O	1.26	1.28
1:X:131:ARG:HB3	4:X:375:HOH:O	1.56	1.02
2:X:360:NAP:O1N	4:X:419:HOH:O	1.80	1.00
1:X:263:LEU:HG	4:X:354:HOH:O	1.63	0.97
1:X:264:ARG:HB3	1:X:264:ARG:HH11	1.27	0.97
2:X:360:NAP:O7N	4:X:404:HOH:O	1.86	0.92
1:X:46:TRP:HA	4:X:398:HOH:O	1.70	0.89
1:X:263:LEU:HB2	4:X:354:HOH:O	1.74	0.86
2:X:360:NAP:C7N	4:X:404:HOH:O	2.26	0.80
1:X:140:THR:HG21	2:X:360:NAP:O4D	1.82	0.79
2:X:360:NAP:N7N	4:X:404:HOH:O	2.19	0.75
1:X:40:LYS:HG2	4:X:371:HOH:O	1.88	0.73
1:X:264:ARG:HB3	1:X:264:ARG:NH1	2.01	0.72
1:X:185:CYS:HB2	2:X:360:NAP:C6N	2.20	0.71
1:X:46:TRP:CA	4:X:398:HOH:O	2.35	0.68
2:X:360:NAP:H51A	4:X:419:HOH:O	1.94	0.68
1:X:140:THR:HG22	2:X:360:NAP:H1D	1.77	0.66
1:X:264:ARG:CB	1:X:264:ARG:HH11	2.07	0.66
1:X:46:TRP:O	4:X:398:HOH:O	2.14	0.65
1:X:232:ASN:ND2	1:X:235:GLU:H	1.93	0.65
1:X:93:LEU:HD22	1:X:113:VAL:HG21	1.78	0.64
1:X:57:GLY:O	4:X:356:HOH:O	2.15	0.63
1:X:185:CYS:HB2	2:X:360:NAP:C5N	2.28	0.63
1:X:232:ASN:HB3	4:X:389:HOH:O	2.02	0.60
1:X:261:PRO:HD2	4:X:376:HOH:O	2.02	0.58
1:X:28:GLN:NE2	4:X:403:HOH:O	2.38	0.56
1:X:140:THR:CG2	2:X:360:NAP:H1D	2.35	0.56
1:X:191:ALA:HA	4:X:378:HOH:O	2.06	0.55
1:X:49:ALA:HB3	4:X:398:HOH:O	2.07	0.54
1:X:185:CYS:HB2	2:X:360:NAP:H6N	1.90	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:131:ARG:CB	4:X:375:HOH:O	2.32	0.52
1:X:95:LEU:HB2	1:X:110:VAL:HG21	1.92	0.51
1:X:189:HIS:CD2	4:X:374:HOH:O	2.64	0.50
1:X:37:ARG:H	2:X:360:NAP:P2B	2.35	0.49
1:X:232:ASN:HD22	1:X:235:GLU:H	1.59	0.49
1:X:93:LEU:HD22	1:X:113:VAL:CG2	2.41	0.49
1:X:260:LEU:HD23	1:X:263:LEU:HD12	1.95	0.49
1:X:224:GLN:O	1:X:228:GLU:HG3	2.13	0.48
1:X:77:GLU:HG2	1:X:78:ARG:N	2.27	0.48
1:X:263:LEU:HD22	1:X:267:LEU:HD22	1.96	0.47
1:X:266:ARG:HB3	1:X:267:LEU:HD13	1.97	0.47
1:X:279:MET:HE2	1:X:279:MET:HA	1.96	0.46
1:X:131:ARG:HG3	1:X:132:ARG:N	2.31	0.46
1:X:91:ALA:HB1	2:X:360:NAP:N3A	2.30	0.45
1:X:131:ARG:NH1	4:X:365:HOH:O	2.49	0.45
1:X:122:LEU:HD13	1:X:169:LEU:CD1	2.47	0.45
1:X:128:ASP:HA	1:X:131:ARG:HG2	1.98	0.45
1:X:185:CYS:CB	2:X:360:NAP:C5N	2.94	0.45
1:X:232:ASN:ND2	1:X:234:GLU:HB2	2.31	0.45
1:X:27:SER:OG	1:X:29:SER:OG	2.35	0.45
1:X:36:LEU:HA	2:X:360:NAP:O2X	2.17	0.44
1:X:224:GLN:OE1	1:X:227:ARG:NH2	2.37	0.44
1:X:263:LEU:HD22	1:X:267:LEU:CD2	2.48	0.44
1:X:53:ALA:HA	4:X:380:HOH:O	2.18	0.43
1:X:140:THR:CG2	2:X:360:NAP:C1D	2.97	0.43
2:X:360:NAP:C5D	2:X:360:NAP:O5B	2.68	0.42
1:X:122:LEU:HD13	1:X:169:LEU:HD11	2.02	0.42
1:X:263:LEU:CB	4:X:354:HOH:O	2.40	0.41
1:X:140:THR:HG21	2:X:360:NAP:C1D	2.50	0.41
1:X:188:VAL:HG12	1:X:190:THR:HG23	2.01	0.41
1:X:50:ARG:CZ	4:X:380:HOH:O	2.69	0.41
1:X:2:ARG:O	1:X:83:ARG:NH2	2.54	0.41
1:X:140:THR:HG21	2:X:360:NAP:C4D	2.51	0.41
2:X:360:NAP:H52N	2:X:360:NAP:O5B	2.20	0.40
1:X:36:LEU:HD12	1:X:36:LEU:N	2.36	0.40
1:X:206:ARG:NE	4:X:379:HOH:O	2.54	0.40
1:X:232:ASN:C	1:X:232:ASN:HD22	2.25	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	X	282/327 (86%)	270 (96%)	9 (3%)	3 (1%)	17	9

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	192	PHE
1	X	191	ALA
1	X	194	GLU

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	X	235/257 (91%)	207 (88%)	28 (12%)	6	3

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	29	SER
1	X	38	ASP
1	X	40	LYS
1	X	50	ARG
1	X	62	LEU
1	X	63	GLN
1	X	77	GLU
1	X	93	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	X	131	ARG
1	X	187	PRO
1	X	189	HIS
1	X	192	PHE
1	X	193	MET
1	X	194	GLU
1	X	197	LEU
1	X	201	GLU
1	X	202	GLU
1	X	205	ASP
1	X	232	ASN
1	X	238	GLU
1	X	241	LEU
1	X	244	LEU
1	X	248	LYS
1	X	262	LEU
1	X	263	LEU
1	X	264	ARG
1	X	267	LEU
1	X	281	ARG

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

Mol	Chain	Res	Type
1	X	231	GLN
1	X	232	ASN

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)

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	NAP	X	360	-	42,52,52	1.70	5 (11%)	54,80,80	2.59	15 (27%)
3	E2B	X	700	-	34,34,34	1.52	4 (11%)	49,52,52	1.24	5 (10%)

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	X	360	-	-	0/27/67/67	0/5/5/5
3	E2B	X	700	-	-	0/8/51/51	0/5/5/5

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	360	NAP	P2B-O2X	-6.34	1.31	1.54
2	X	360	NAP	P2B-O2B	-5.03	1.44	1.60
3	X	700	E2B	C10-C9	-4.64	1.49	1.54
2	X	360	NAP	P2B-O1X	-3.90	1.38	1.51
3	X	700	E2B	O19-C18	-3.63	1.35	1.42
2	X	360	NAP	P2B-O3X	-3.16	1.43	1.54
3	X	700	E2B	C28-N30	-2.25	1.28	1.33
2	X	360	NAP	PA-O2A	-2.07	1.46	1.54
3	X	700	E2B	C16-C15	-2.00	1.49	1.54

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	X	360	NAP	N3A-C2A-N1A	-9.49	121.63	128.89
2	X	360	NAP	O5B-C5B-C4B	-3.69	95.52	109.12

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	X	360	NAP	C4B-O4B-C1B	-3.36	106.03	109.72
3	X	700	E2B	C16-C15-C14	-3.14	99.40	103.97
2	X	360	NAP	C3N-C7N-N7N	-2.96	114.58	117.82
3	X	700	E2B	C24-C28-N30	-2.84	114.71	117.82
2	X	360	NAP	C4A-C5A-N7A	-2.84	106.87	109.48
2	X	360	NAP	C1B-N9A-C4A	-2.02	123.89	126.94
2	X	360	NAP	O2X-P2B-O1X	-2.00	104.14	110.58
2	X	360	NAP	O4B-C4B-C5B	2.13	116.94	109.32
2	X	360	NAP	C2B-C3B-C4B	2.17	106.98	101.85
2	X	360	NAP	O3D-C3D-C2D	2.25	119.14	111.83
2	X	360	NAP	O5D-C5D-C4D	2.51	118.38	109.12
2	X	360	NAP	O2B-C2B-C1B	2.61	120.20	110.02
3	X	700	E2B	O29-C28-C24	2.66	122.49	119.59
3	X	700	E2B	C22-C21-C17	2.81	118.55	113.75
3	X	700	E2B	C8-C9-C15	3.17	117.26	112.02
2	X	360	NAP	P2B-O2B-C2B	3.31	129.51	121.56
2	X	360	NAP	O2B-P2B-O1X	4.94	119.45	107.11
2	X	360	NAP	O4D-C1D-N1N	10.59	119.76	108.13

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	X	360	NAP	20	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(Å²)	Q<0.9
1	X	284/327 (86%)	0.27	20 (7%)	19	21	11, 29, 50, 78	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	192	PHE	7.6
1	X	193	MET	7.6
1	X	1	ALA	6.9
1	X	194	GLU	6.4
1	X	195	LYS	5.3
1	X	191	ALA	5.2
1	X	197	LEU	3.9
1	X	196	VAL	3.7
1	X	190	THR	3.3
1	X	52	LEU	3.3
1	X	267	LEU	3.0
1	X	247	PRO	2.8
1	X	202	GLU	2.6
1	X	198	GLY	2.5
1	X	80	THR	2.5
1	X	131	ARG	2.4
1	X	227	ARG	2.4
1	X	77	GLU	2.3
1	X	7	ILE	2.3
1	X	189	HIS	2.1

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
2	NAP	X	360	48/48	0.61	0.44	4.01	17,30,33,34	48
3	E2B	X	700	30/30	-	-	-	0,0,0,0	30

6.5 Other polymers (i)

There are no such residues in this entry.