



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:03 PM GMT

PDB ID : 4H2I
Title : Human ecto-5'-nucleotidase (CD73): crystal form III (closed) in complex with AMPCP
Authors : Straeter, N.; Knapp, K.M.; Zebisch, M.; Pippel, J.
Deposited on : 2012-09-12
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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

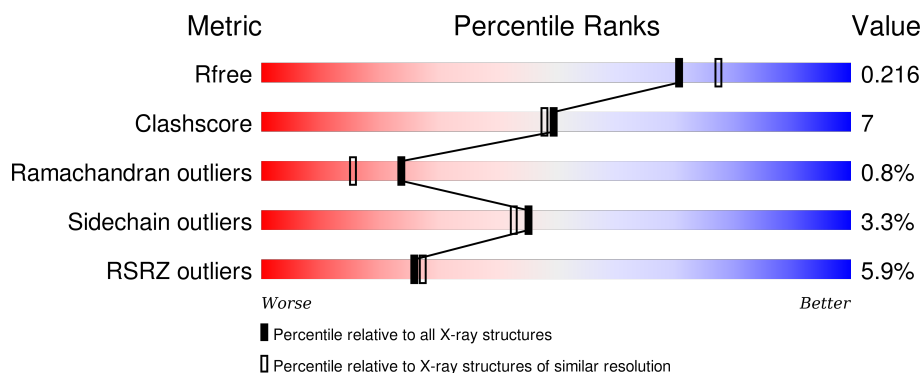
1 Overall quality at a glance

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}	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 $\leq 5\%$. 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	532	<div> <div>6%</div> <div>82%</div> <div>14%</div> <div>• •</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4518 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 5'-nucleotidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	524	Total	C	N	O	S	0	1	0
			4095	2604	697	774	20			

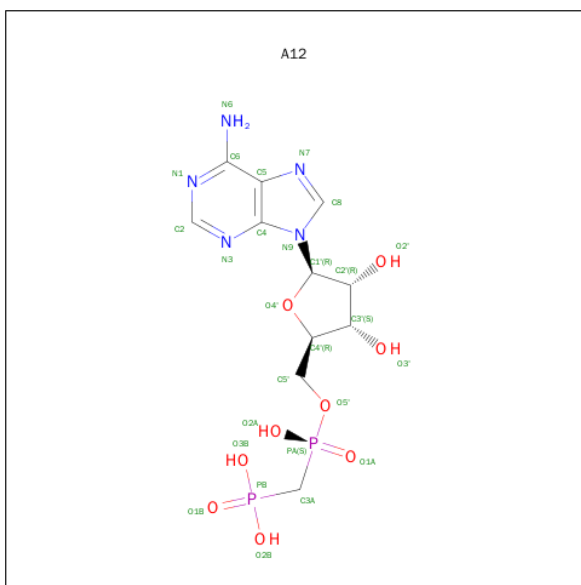
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	INITIATING METHIONINE	UNP P21589
A	53	ASP	ASN	ENGINEERED MUTATION	UNP P21589
A	311	ASP	ASN	ENGINEERED MUTATION	UNP P21589
A	333	ASP	ASN	ENGINEERED MUTATION	UNP P21589
A	376	ALA	THR	VARIANT	UNP P21589
A	403	ASP	ASN	ENGINEERED MUTATION	UNP P21589
A	550	LEU	-	EXPRESSION TAG	UNP P21589
A	551	GLU	-	EXPRESSION TAG	UNP P21589
A	552	HIS	-	EXPRESSION TAG	UNP P21589
A	553	HIS	-	EXPRESSION TAG	UNP P21589
A	554	HIS	-	EXPRESSION TAG	UNP P21589
A	555	HIS	-	EXPRESSION TAG	UNP P21589
A	556	HIS	-	EXPRESSION TAG	UNP P21589
A	557	HIS	-	EXPRESSION TAG	UNP P21589

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

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

- Molecule 3 is PHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: A12) (formula: C₁₁H₁₇N₅O₉P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			27	11	5	9	2		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Ca 1 1	0	0

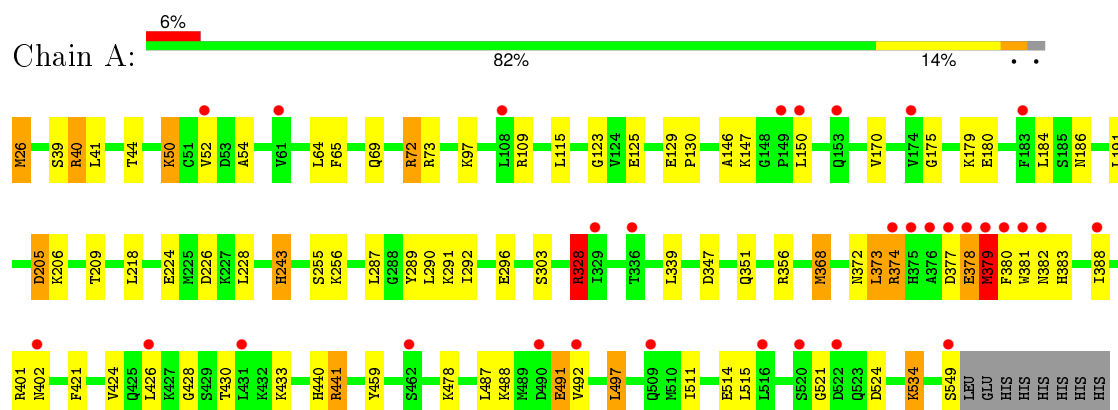
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	392	Total O 392 392	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: 5'-nucleotidase



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	53.16 Å 94.86 Å 234.28 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.22 – 2.00 24.20 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (24.22-2.00) 99.9 (24.20-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.34 (at 1.99 Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.170 , 0.215 0.170 , 0.216	Depositor DCC
R_{free} test set	1224 reflections (3.11%)	DCC
Wilson B-factor (Å ²)	29.6	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 53.5	EDS
Estimated twinning fraction	0.025 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.037 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 40528 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4518	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: ZN, CL, CA, A12

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	0.98	0/4184	1.06	16/5665 (0.3%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	72	ARG	NE-CZ-NH1	-9.80	115.40	120.30
1	A	40	ARG	NE-CZ-NH1	-9.32	115.64	120.30
1	A	328	ARG	NE-CZ-NH1	8.15	124.37	120.30
1	A	226	ASP	CB-CG-OD1	8.13	125.61	118.30
1	A	328	ARG	NE-CZ-NH2	-7.09	116.76	120.30
1	A	401	ARG	NE-CZ-NH2	-6.88	116.86	120.30
1	A	226	ASP	CB-CG-OD2	-6.72	112.25	118.30
1	A	524	ASP	CB-CG-OD1	6.30	123.97	118.30
1	A	205	ASP	CB-CG-OD1	6.20	123.88	118.30
1	A	401	ARG	NE-CZ-NH1	5.76	123.18	120.30
1	A	347	ASP	CB-CG-OD2	-5.75	113.13	118.30
1	A	224	GLU	OE1-CD-OE2	-5.73	116.43	123.30
1	A	441	ARG	NE-CZ-NH1	5.52	123.06	120.30
1	A	50	LYS	CD-CE-NZ	-5.36	99.37	111.70
1	A	368	MET	CG-SD-CE	-5.22	91.85	100.20
1	A	41	LEU	CA-CB-CG	-5.14	103.47	115.30

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	4095	0	4074	57	0
2	A	2	0	0	0	0
3	A	27	0	14	2	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
6	A	392	0	0	10	0
All	All	4518	0	4088	58	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 7.

All (58) 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:356:ARG:HD2	6:A:835:HOH:O	1.50	1.10
1:A:380:PHE:HB3	1:A:478:LYS:HA	1.56	0.88
1:A:491:GLU:HG3	1:A:492:VAL:H	1.50	0.75
1:A:290:LEU:HD21	1:A:292:ILE:HD11	1.73	0.70
1:A:54:ALA:HB2	6:A:1019:HOH:O	1.93	0.68
1:A:378:GLU:HB3	1:A:382:ASN:HD22	1.59	0.68
1:A:351:GLN:NE2	6:A:782:HOH:O	2.27	0.68
1:A:491:GLU:HG3	1:A:492:VAL:N	2.09	0.67
1:A:228:LEU:HD11	6:A:1047:HOH:O	1.99	0.63
1:A:147:LYS:HA	6:A:1045:HOH:O	1.98	0.62
1:A:184:LEU:HD22	6:A:1050:HOH:O	2.00	0.61
1:A:290:LEU:CD2	1:A:292:ILE:HD11	2.30	0.61
1:A:40:ARG:HG2	1:A:44:THR:HG21	1.84	0.59
1:A:289:TYR:OH	1:A:291:LYS:HE3	2.03	0.59
1:A:378:GLU:O	1:A:379:MET:CG	2.54	0.56
1:A:328:ARG:HD3	1:A:328:ARG:O	2.04	0.56
1:A:374:ARG:HE	1:A:534:LYS:HZ3	1.55	0.55
1:A:69[B]:GLN:OE1	1:A:72:ARG:HD2	2.07	0.54
1:A:52:VAL:HG11	1:A:402:ASN:HB2	1.89	0.53
1:A:123:GLY:HA2	1:A:186:ASN:O	2.08	0.52
1:A:351:GLN:HG3	6:A:924:HOH:O	2.08	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:69[B]:GLN:HE22	1:A:109:ARG:NH2	2.07	0.51
1:A:26:MET:HE2	6:A:838:HOH:O	2.11	0.51
1:A:146:ALA:HA	1:A:191:LEU:HD23	1.91	0.51
1:A:373:LEU:HD21	1:A:383:HIS:HA	1.93	0.51
1:A:125:GLU:OE1	1:A:125:GLU:HA	2.11	0.50
1:A:380:PHE:CB	1:A:478:LYS:HA	2.35	0.50
1:A:255:SER:OG	1:A:256:LYS:N	2.45	0.50
1:A:372:ASN:C	1:A:374:ARG:H	2.15	0.49
1:A:368:MET:HG2	1:A:421:PHE:CD2	2.48	0.49
1:A:292:ILE:HD13	1:A:303:SER:HB2	1.94	0.48
1:A:379:MET:HB3	1:A:380:PHE:H	1.28	0.48
1:A:426:LEU:HD22	1:A:430:THR:HG21	1.95	0.48
1:A:205:ASP:O	1:A:209:THR:HG23	2.15	0.47
1:A:459:TYR:O	1:A:549:SER:N	2.44	0.47
1:A:69[B]:GLN:HE22	1:A:109:ARG:HH21	1.61	0.47
1:A:511:ILE:O	1:A:515:LEU:HD23	2.14	0.47
1:A:65:PHE:CE2	1:A:69[A]:GLN:HG3	2.51	0.46
1:A:440:HIS:CE1	1:A:441:ARG:HG2	2.50	0.46
1:A:175:GLY:HA2	1:A:218:LEU:O	2.16	0.46
1:A:373:LEU:CD2	1:A:383:HIS:HA	2.46	0.45
1:A:129:GLU:HB2	1:A:130:PRO:HD3	1.99	0.45
1:A:150:LEU:HD22	1:A:150:LEU:O	2.17	0.44
1:A:428:GLY:HA3	1:A:488:LYS:O	2.17	0.44
1:A:421:PHE:O	1:A:521:GLY:HA3	2.17	0.43
1:A:339:LEU:HD23	1:A:339:LEU:HA	1.73	0.42
1:A:243:HIS:O	3:A:603:A12:H3A1	2.19	0.42
1:A:381:TRP:O	1:A:382:ASN:HB2	2.19	0.42
3:A:603:A12:H2	6:A:1015:HOH:O	2.20	0.42
1:A:206:LYS:HE2	1:A:206:LYS:HB3	1.55	0.42
1:A:433:LYS:NZ	1:A:514:GLU:OE2	2.52	0.42
1:A:64:LEU:HD11	1:A:287:LEU:HD13	2.02	0.42
1:A:426:LEU:HD23	1:A:515:LEU:HA	2.03	0.41
1:A:50:LYS:HE2	6:A:1029:HOH:O	2.20	0.41
1:A:115:LEU:HA	1:A:115:LEU:HD23	1.83	0.41
1:A:487:LEU:HD23	1:A:487:LEU:C	2.42	0.41
1:A:388:ILE:HD11	1:A:497:LEU:HD23	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	523/532 (98%)	495 (95%)	24 (5%)	4 (1%)	24	15

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	379	MET
1	A	243	HIS
1	A	373	LEU
1	A	377	ASP

5.3.2 Protein sidechains

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	452/459 (98%)	437 (97%)	15 (3%)	45	43

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	MET
1	A	39	SER
1	A	97	LYS
1	A	170	VAL
1	A	179	LYS
1	A	180	GLU
1	A	296	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	328	ARG
1	A	374	ARG
1	A	378	GLU
1	A	379	MET
1	A	424	VAL
1	A	491	GLU
1	A	497	LEU
1	A	534	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	299	ASN
1	A	371	ASN
1	A	382	ASN
1	A	503	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](#)

Of 5 ligands modelled in this entry, 4 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	A12	A	603	2	22,29,29	1.50	3 (13%)	27,45,45	2.57	11 (40%)

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	A12	A	603	2	-	0/12/32/32	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	603	A12	PB-O1B	-2.55	1.44	1.50
3	A	603	A12	C5-C4	3.18	1.47	1.40
3	A	603	A12	PB-O2B	3.23	1.62	1.54

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	603	A12	N3-C2-N1	-6.36	124.03	128.89
3	A	603	A12	C1'-N9-C4	-4.53	120.11	126.94
3	A	603	A12	C2'-C1'-N9	-4.30	107.72	114.29
3	A	603	A12	O2B-PB-O1B	-3.73	102.85	112.40
3	A	603	A12	O3B-PB-O2B	-3.40	98.17	108.13
3	A	603	A12	O3'-C3'-C4'	-3.09	101.80	111.05
3	A	603	A12	O4'-C4'-C3'	-2.41	100.30	105.15
3	A	603	A12	N6-C6-N1	2.10	123.72	119.20
3	A	603	A12	C2-N1-C6	2.16	122.62	118.77
3	A	603	A12	O3B-PB-O1B	2.44	118.65	112.40
3	A	603	A12	O1B-PB-C3A	4.05	120.36	111.13

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	603	A12	2	0

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, 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	524/532 (98%)	-0.06	31 (5%)	26 27	17, 36, 70, 128	7 (1%)

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	375	HIS	7.6
1	A	378	GLU	6.5
1	A	431	LEU	4.9
1	A	61	VAL	4.8
1	A	382	ASN	4.4
1	A	380	PHE	4.0
1	A	402	ASN	3.9
1	A	381	TRP	3.7
1	A	336	THR	3.5
1	A	516	LEU	3.3
1	A	108	LEU	3.2
1	A	374	ARG	3.1
1	A	376	ALA	3.0
1	A	52	VAL	3.0
1	A	174	VAL	2.8
1	A	377	ASP	2.7
1	A	379	MET	2.6
1	A	183	PHE	2.5
1	A	490	ASP	2.4
1	A	509	GLN	2.4
1	A	150	LEU	2.4
1	A	153	GLN	2.3
1	A	426	LEU	2.3
1	A	549	SER	2.2
1	A	329	ILE	2.2
1	A	149	PRO	2.2
1	A	388	ILE	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	520	SER	2.1
1	A	522	ASP	2.0
1	A	462	SER	2.0
1	A	492	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 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(\AA^2)	Q<0.9
5	CA	A	605	1/1	0.98	0.07	-0.37	19,19,19,19	1
3	A12	A	603	27/27	0.99	0.07	-1.20	22,25,41,46	0
2	ZN	A	602	1/1	1.00	0.06	-2.53	22,22,22,22	0
2	ZN	A	601	1/1	1.00	0.05	-5.14	21,21,21,21	0
4	CL	A	604	1/1	0.95	0.28	-	59,59,59,59	0

6.5 Other polymers [i](#)

There are no such residues in this entry.