Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 07:57 pm BST

PDB ID	:	1LLA
Title	:	CRYSTAL STRUCTURE OF DEOXYGENATED LIMULUS POLYPHEMUS SUBUNIT II HEMOCYANIN AT 2.18 ANGSTROMS RESOLUTION: CLUES FOR A MECHANISM FOR ALLOSTERIC REGULATION
Authors	:	Hazes, B.; Hol, W.G.J.
Deposited on	:	1992-09-07
Resolution	:	2.18 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 ⓘ symbol.

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

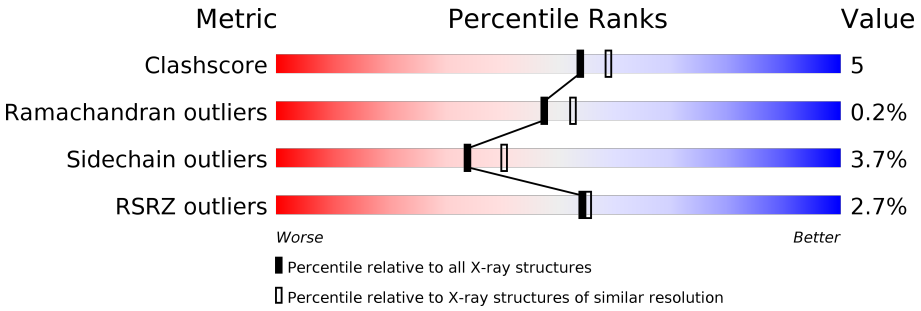
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.18 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))
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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 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	628	3% 77% 16%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5196 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 HEMOCYANIN (SUBUNIT TYPE II).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	600	Total	C	N	O	S	0	0	0
			4860	3101	849	888	22			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	ILE	VAL	CONFLICT	UNP P04253
A	408	THR	PHE	CONFLICT	UNP P04253
A	409	PHE	THR	CONFLICT	UNP P04253

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

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

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

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

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

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

- Molecule 5 is water.

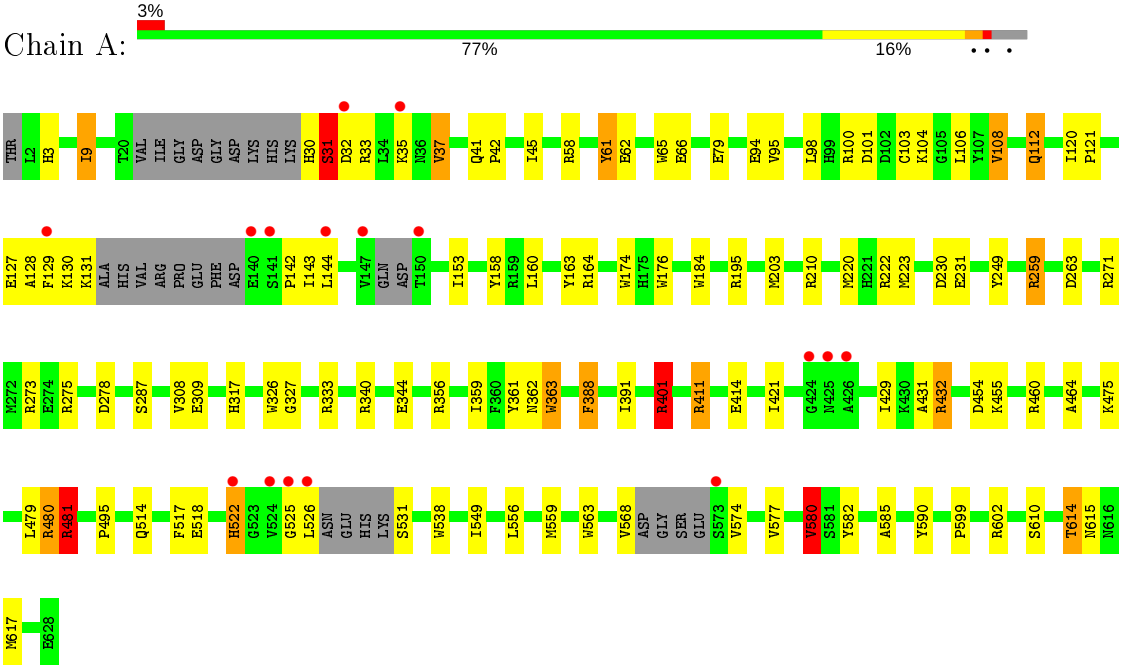
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	332	Total O 332 332	0	0

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 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: HEMOCYANIN (SUBUNIT TYPE II)

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	R 3 2	Depositor
Cell constants a, b, c, α, β, γ	117.00Å 117.00Å 117.00Å 60.02° 60.02° 60.02°	Depositor
Resolution (Å)	10.00 – 2.18 37.96 – 2.19	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.18) 95.6 (37.96-2.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.174 , (Not available) 0.171 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 70.3	EDS
L-test for twinning[1]	$< \|L\| >$ = (Not available), $< L^2 >$ = (Not available)	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5196	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *(Not available)*

[1] 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: NA, CU, CL

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	0.89	2/5001 (0.0%)	1.45	70/6776 (1.0%)

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	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	610	SER	CA-CB	6.68	1.62	1.52
1	A	563	TRP	CG-CD2	-5.62	1.34	1.43

All (70) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	481	ARG	NE-CZ-NH1	12.90	126.75	120.30
1	A	563	TRP	CD1-CG-CD2	9.01	113.51	106.30
1	A	411	ARG	NE-CZ-NH1	-8.35	116.13	120.30
1	A	176	TRP	CD1-CG-CD2	8.08	112.76	106.30
1	A	273	ARG	NE-CZ-NH1	8.00	124.30	120.30
1	A	65	TRP	CD1-CG-CD2	7.96	112.67	106.30
1	A	580	VAL	CB-CA-C	-7.84	96.50	111.40
1	A	363	TRP	CD1-CG-CD2	7.42	112.23	106.30
1	A	326	TRP	CD1-CG-CD2	7.40	112.22	106.30
1	A	100	ARG	NE-CZ-NH2	7.37	123.98	120.30
1	A	480	ARG	NE-CZ-NH1	7.30	123.95	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	356	ARG	NE-CZ-NH1	7.27	123.93	120.30
1	A	326	TRP	CE2-CD2-CG	-7.25	101.50	107.30
1	A	432	ARG	NE-CZ-NH2	-7.16	116.72	120.30
1	A	460	ARG	NE-CZ-NH1	7.15	123.88	120.30
1	A	522	HIS	CA-C-N	7.13	130.46	116.20
1	A	65	TRP	CE2-CD2-CG	-6.96	101.73	107.30
1	A	58	ARG	NE-CZ-NH1	6.93	123.77	120.30
1	A	176	TRP	CE2-CD2-CG	-6.87	101.80	107.30
1	A	174	TRP	CE2-CD2-CG	-6.87	101.80	107.30
1	A	614	THR	N-CA-CB	-6.76	97.45	110.30
1	A	163	TYR	CB-CG-CD1	-6.72	116.97	121.00
1	A	563	TRP	CE2-CD2-CG	-6.56	102.05	107.30
1	A	103	CYS	CA-CB-SG	-6.54	102.23	114.00
1	A	174	TRP	CG-CD2-CE3	6.41	139.67	133.90
1	A	184	TRP	CE2-CD2-CG	-6.24	102.31	107.30
1	A	363	TRP	CE2-CD2-CG	-6.22	102.33	107.30
1	A	538	TRP	CE2-CD2-CG	-6.21	102.33	107.30
1	A	108	VAL	CG1-CB-CG2	-6.21	100.97	110.90
1	A	340	ARG	NE-CZ-NH2	-6.14	117.23	120.30
1	A	184	TRP	CD1-CG-CD2	6.13	111.21	106.30
1	A	33	ARG	NE-CZ-NH1	6.13	123.36	120.30
1	A	522	HIS	O-C-N	-6.11	112.82	123.20
1	A	195	ARG	NE-CZ-NH2	-6.10	117.25	120.30
1	A	582	TYR	CB-CG-CD2	-6.09	117.35	121.00
1	A	58	ARG	NE-CZ-NH2	-6.05	117.28	120.30
1	A	538	TRP	CD1-CG-CD2	6.03	111.12	106.30
1	A	33	ARG	NE-CZ-NH2	-5.99	117.31	120.30
1	A	563	TRP	CG-CD1-NE1	-5.83	104.27	110.10
1	A	164	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	A	309	GLU	CA-C-N	5.77	129.90	117.20
1	A	9	ILE	CA-CB-CG1	-5.76	100.05	111.00
1	A	176	TRP	CB-CG-CD1	-5.75	119.52	127.00
1	A	590	TYR	CB-CG-CD2	-5.74	117.56	121.00
1	A	176	TRP	CG-CD2-CE3	5.73	139.05	133.90
1	A	174	TRP	CB-CG-CD1	-5.63	119.68	127.00
1	A	568	VAL	CG1-CB-CG2	-5.60	101.94	110.90
1	A	356	ARG	NE-CZ-NH2	-5.59	117.51	120.30
1	A	230	ASP	CB-CA-C	-5.52	99.36	110.40
1	A	176	TRP	CG-CD1-NE1	-5.45	104.65	110.10
1	A	432	ARG	NE-CZ-NH1	5.45	123.02	120.30
1	A	95	VAL	CG1-CB-CG2	-5.40	102.26	110.90
1	A	556	LEU	CA-CB-CG	5.39	127.69	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	79	GLU	CA-CB-CG	-5.38	101.57	113.40
1	A	37	VAL	N-CA-CB	-5.33	99.76	111.50
1	A	61	TYR	CA-CB-CG	5.33	123.53	113.40
1	A	363	TRP	CG-CD1-NE1	-5.31	104.79	110.10
1	A	308	VAL	CG1-CB-CG2	-5.30	102.43	110.90
1	A	31	SER	N-CA-C	-5.28	96.74	111.00
1	A	388	PHE	CB-CG-CD2	-5.23	117.14	120.80
1	A	210	ARG	NE-CZ-NH1	5.23	122.91	120.30
1	A	333	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	A	142	PRO	N-CA-CB	5.22	109.56	103.30
1	A	164	ARG	NE-CZ-NH2	-5.15	117.73	120.30
1	A	275	ARG	NE-CZ-NH1	5.13	122.86	120.30
1	A	563	TRP	CE2-CD2-CE3	5.10	124.82	118.70
1	A	401	ARG	NE-CZ-NH2	5.08	122.84	120.30
1	A	259	ARG	NE-CZ-NH1	5.05	122.83	120.30
1	A	203	MET	CG-SD-CE	5.02	108.23	100.20
1	A	249	TYR	CB-CG-CD2	-5.00	118.00	121.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	222	ARG	Sidechain
1	A	271	ARG	Sidechain
1	A	401	ARG	Sidechain
1	A	411	ARG	Sidechain
1	A	525	GLY	Peptide

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	4860	0	4587	50	0
2	A	2	0	0	0	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	332	0	0	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5196	0	4587	50	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 5.

All (50) 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:9:ILE:HD12	1:A:106:LEU:HD13	1.63	0.81
1:A:112:GLN:H	1:A:112:GLN:HE21	1.30	0.80
1:A:144:LEU:HD21	1:A:432:ARG:NH2	1.97	0.78
1:A:104:LYS:NZ	1:A:104:LYS:HB3	2.06	0.71
1:A:481:ARG:NH1	1:A:481:ARG:HG3	2.14	0.62
1:A:62:GLU:O	1:A:66:GLU:HG2	2.01	0.60
1:A:464:ALA:HB1	1:A:479:LEU:HD22	1.86	0.57
1:A:128:ALA:HB1	1:A:429:ILE:HG21	1.86	0.56
1:A:112:GLN:HE21	1:A:112:GLN:N	2.03	0.54
1:A:98:LEU:HD23	5:A:914:HOH:O	2.10	0.52
1:A:104:LYS:HZ3	1:A:104:LYS:HB3	1.75	0.51
1:A:41:GLN:HA	1:A:41:GLN:OE1	2.10	0.51
1:A:549:ILE:HG23	5:A:930:HOH:O	2.11	0.51
1:A:129:PHE:CZ	1:A:421:ILE:HB	2.46	0.51
1:A:481:ARG:HD2	5:A:767:HOH:O	2.11	0.50
1:A:32:ASP:O	1:A:35:LYS:HG2	2.12	0.50
1:A:112:GLN:H	1:A:112:GLN:NE2	2.03	0.50
1:A:3:HIS:NE2	1:A:522:HIS:O	2.45	0.50
1:A:259:ARG:HG3	5:A:863:HOH:O	2.10	0.50
1:A:317:HIS:CD2	1:A:317:HIS:C	2.86	0.50
1:A:45:ILE:HD12	1:A:344:GLU:HB3	1.94	0.49
1:A:327:GLY:HA3	1:A:363:TRP:CZ2	2.48	0.49
1:A:414:GLU:HA	1:A:431:ALA:O	2.14	0.48
1:A:104:LYS:NZ	1:A:104:LYS:CB	2.76	0.48
1:A:615:ASN:HB2	5:A:895:HOH:O	2.14	0.48
1:A:94:GLU:HG3	1:A:108:VAL:HG21	1.96	0.47
1:A:317:HIS:CD2	1:A:317:HIS:O	2.67	0.47
1:A:98:LEU:CD2	5:A:914:HOH:O	2.63	0.47
1:A:143:ILE:HD12	1:A:429:ILE:HG13	1.96	0.46
1:A:614:THR:OG1	1:A:617:MET:CE	2.64	0.46
1:A:577:VAL:O	1:A:580:VAL:HG22	2.16	0.45
1:A:518:GLU:HA	1:A:518:GLU:OE1	2.16	0.44
1:A:481:ARG:HH11	1:A:481:ARG:HG3	1.82	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:127:GLU:O	1:A:131:LYS:HG2	2.17	0.44
1:A:391:ILE:HD11	1:A:455:LYS:HG3	2.00	0.43
1:A:454:ASP:OD1	1:A:495:PRO:HD3	2.19	0.43
1:A:574:VAL:HG11	1:A:585:ALA:HB1	2.00	0.43
1:A:153:ILE:HG12	1:A:158:TYR:CE1	2.54	0.43
1:A:359:ILE:HA	1:A:362:ASN:HD22	1.83	0.42
1:A:480:ARG:HD3	1:A:602:ARG:O	2.19	0.42
1:A:42:PRO:HD3	5:A:802:HOH:O	2.18	0.42
1:A:559:MET:HG3	1:A:617:MET:HG2	2.00	0.42
1:A:223:MET:SD	1:A:361:TYR:HB3	2.60	0.42
1:A:475:LYS:HD3	1:A:475:LYS:HA	1.78	0.41
1:A:104:LYS:HZ2	1:A:104:LYS:HB3	1.81	0.41
1:A:30:HIS:O	1:A:31:SER:HB2	2.20	0.41
1:A:160:LEU:HD21	1:A:220:MET:SD	2.61	0.41
1:A:153:ILE:HG12	1:A:158:TYR:CD1	2.56	0.40
1:A:120:ILE:HA	1:A:121:PRO:HD3	1.96	0.40
1:A:522:HIS:N	1:A:522:HIS:CD2	2.89	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	588/628 (94%)	573 (97%)	14 (2%)	1 (0%)	47 52

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	SER

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	518/558 (93%)	499 (96%)	19 (4%)	34 40

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	SER
1	A	37	VAL
1	A	61	TYR
1	A	101	ASP
1	A	112	GLN
1	A	130	LYS
1	A	231	GLU
1	A	263	ASP
1	A	278	ASP
1	A	287	SER
1	A	388	PHE
1	A	401	ARG
1	A	481	ARG
1	A	514	GLN
1	A	517	PHE
1	A	526	LEU
1	A	531	SER
1	A	580	VAL
1	A	599	PRO

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	112	GLN
1	A	317	HIS
1	A	362	ASN
1	A	413	GLN
1	A	522	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

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

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.

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	600/628 (95%)	-0.47	16 (2%) 54 55	4, 19, 38, 48	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	129	PHE	3.9
1	A	425	ASN	3.8
1	A	141	SER	3.7
1	A	426	ALA	3.6
1	A	524	VAL	3.1
1	A	35	LYS	2.9
1	A	140	GLU	2.8
1	A	526	LEU	2.7
1	A	424	GLY	2.5
1	A	150	THR	2.5
1	A	32	ASP	2.5
1	A	522	HIS	2.3
1	A	525	GLY	2.2
1	A	144	LEU	2.1
1	A	573	SER	2.1
1	A	147	VAL	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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	NA	A	631	1/1	0.98	0.13	17,17,17,17	0
2	CU	A	630	1/1	0.99	0.04	9,9,9,9	0
2	CU	A	629	1/1	0.99	0.06	10,10,10,10	0
4	CL	A	632	1/1	0.99	0.06	18,18,18,18	0

6.5 Other polymers

There are no such residues in this entry.