wwPDB X-ray Structure Validation Summary Report (i)

Jun 23, 2024 – 04:38 AM EDT

PDB ID	:	5NLY
Title	:	Brag2 Sec7-PH (390-763), P212121
Authors	:	Nawrotek, A.; Cherfils, J.
Deposited on	:	2017-04-05
Resolution	:	2.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

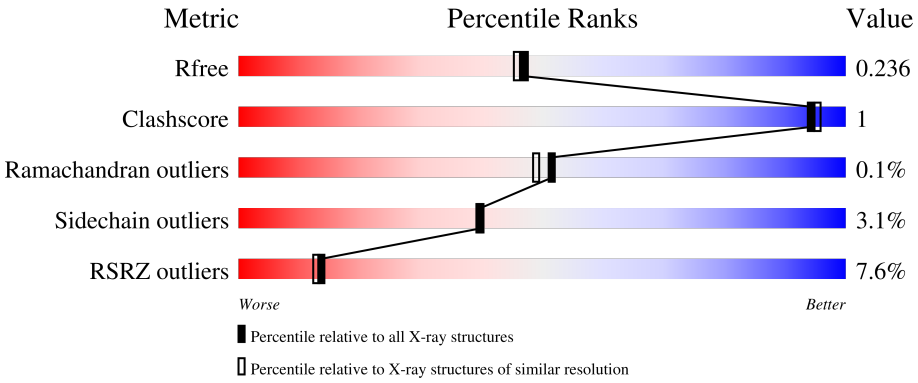
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 of 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	405	7% 77% 9% 13%
2	B	405	6% 82% · 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12215 atoms, of which 5950 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 IQ motif and SEC7 domain-containing protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	353	5903	1860	2976	521	527	19	0	1	0

There are 31 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	MET	-	initiating methionine	UNP Q6DN90
A	360	SER	-	expression tag	UNP Q6DN90
A	361	TYR	-	expression tag	UNP Q6DN90
A	362	TYR	-	expression tag	UNP Q6DN90
A	363	HIS	-	expression tag	UNP Q6DN90
A	364	HIS	-	expression tag	UNP Q6DN90
A	365	HIS	-	expression tag	UNP Q6DN90
A	366	HIS	-	expression tag	UNP Q6DN90
A	367	HIS	-	expression tag	UNP Q6DN90
A	368	HIS	-	expression tag	UNP Q6DN90
A	369	ASP	-	expression tag	UNP Q6DN90
A	370	TYR	-	expression tag	UNP Q6DN90
A	371	ASP	-	expression tag	UNP Q6DN90
A	372	ILE	-	expression tag	UNP Q6DN90
A	373	PRO	-	expression tag	UNP Q6DN90
A	374	THR	-	expression tag	UNP Q6DN90
A	375	THR	-	expression tag	UNP Q6DN90
A	376	GLU	-	expression tag	UNP Q6DN90
A	377	ASN	-	expression tag	UNP Q6DN90
A	378	LEU	-	expression tag	UNP Q6DN90
A	379	TYR	-	expression tag	UNP Q6DN90
A	380	PHE	-	expression tag	UNP Q6DN90
A	381	GLN	-	expression tag	UNP Q6DN90
A	382	GLY	-	expression tag	UNP Q6DN90
A	383	ALA	-	expression tag	UNP Q6DN90
A	384	MET	-	expression tag	UNP Q6DN90
A	385	GLY	-	expression tag	UNP Q6DN90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	386	SER	-	expression tag	UNP Q6DN90
A	387	GLU	-	expression tag	UNP Q6DN90
A	388	PHE	-	expression tag	UNP Q6DN90
A	389	MET	-	expression tag	UNP Q6DN90

- Molecule 2 is a protein called IQ motif and SEC7 domain-containing protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	351	Total	C	H	N	O	S	0	1	0
			5886	1852	2974	521	519	20			

There are 31 discrepancies between the modelled and reference sequences:

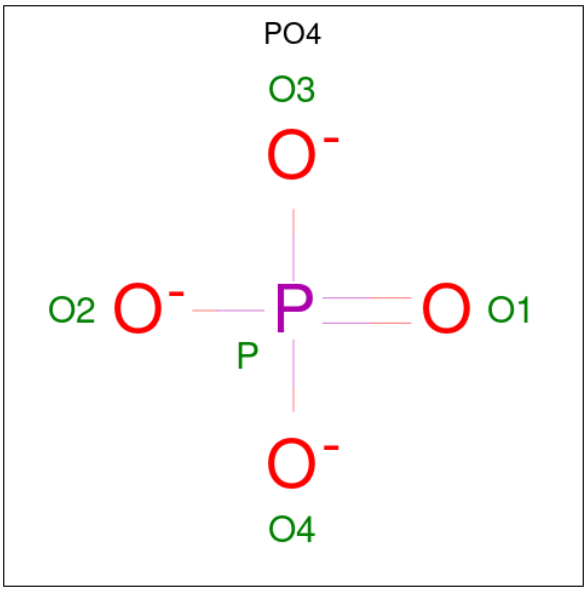
Chain	Residue	Modelled	Actual	Comment	Reference
B	359	MET	-	initiating methionine	UNP Q6DN90
B	360	SER	-	expression tag	UNP Q6DN90
B	361	TYR	-	expression tag	UNP Q6DN90
B	362	TYR	-	expression tag	UNP Q6DN90
B	363	HIS	-	expression tag	UNP Q6DN90
B	364	HIS	-	expression tag	UNP Q6DN90
B	365	HIS	-	expression tag	UNP Q6DN90
B	366	HIS	-	expression tag	UNP Q6DN90
B	367	HIS	-	expression tag	UNP Q6DN90
B	368	HIS	-	expression tag	UNP Q6DN90
B	369	ASP	-	expression tag	UNP Q6DN90
B	370	TYR	-	expression tag	UNP Q6DN90
B	371	ASP	-	expression tag	UNP Q6DN90
B	372	ILE	-	expression tag	UNP Q6DN90
B	373	PRO	-	expression tag	UNP Q6DN90
B	374	THR	-	expression tag	UNP Q6DN90
B	375	THR	-	expression tag	UNP Q6DN90
B	376	GLU	-	expression tag	UNP Q6DN90
B	377	ASN	-	expression tag	UNP Q6DN90
B	378	LEU	-	expression tag	UNP Q6DN90
B	379	TYR	-	expression tag	UNP Q6DN90
B	380	PHE	-	expression tag	UNP Q6DN90
B	381	GLN	-	expression tag	UNP Q6DN90
B	382	GLY	-	expression tag	UNP Q6DN90
B	383	ALA	-	expression tag	UNP Q6DN90
B	384	MET	-	expression tag	UNP Q6DN90
B	385	GLY	-	expression tag	UNP Q6DN90
B	386	SER	-	expression tag	UNP Q6DN90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	387	GLU	-	expression tag	UNP Q6DN90
B	388	PHE	-	expression tag	UNP Q6DN90
B	389	MET	-	expression tag	UNP Q6DN90

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O P 5 4 1	0	0

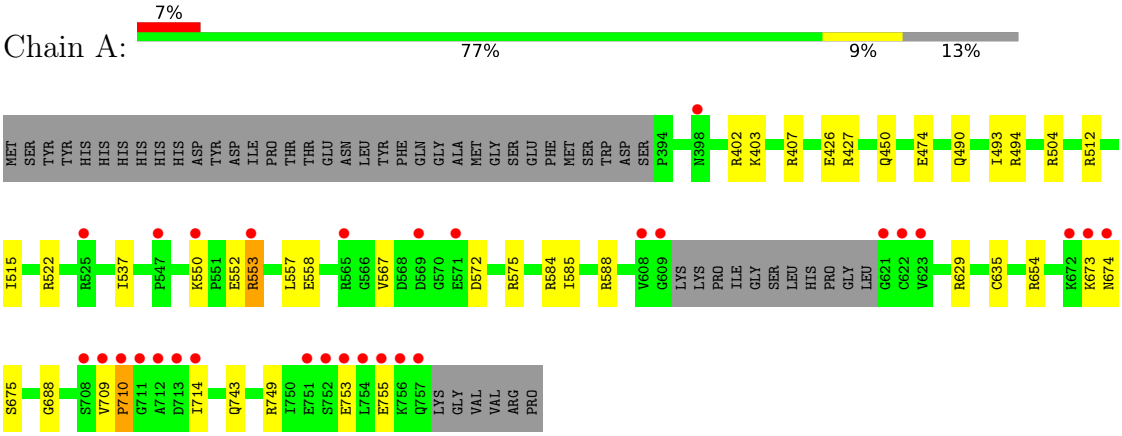
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	237	Total O 237 237	0	0
4	B	184	Total O 184 184	0	0

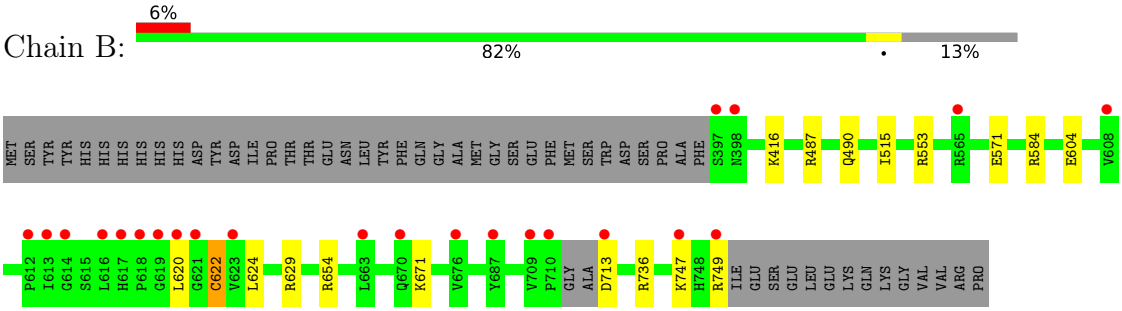
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: IQ motif and SEC7 domain-containing protein 1



• Molecule 2: IQ motif and SEC7 domain-containing protein 1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.89Å 66.04Å 206.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.99 – 2.00 28.99 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (28.99-2.00) 99.9 (28.99-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 2.00Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.208 , 0.233 0.215 , 0.236	Depositor DCC
R_{free} test set	2881 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	26.7	Xtriage
Anisotropy	0.298	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 48.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.027 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12215	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.12% 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, 0.333 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: PO4, CME

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	0.57	0/2940	0.75	9/3943 (0.2%)
2	B	0.58	0/2915	0.75	8/3909 (0.2%)
All	All	0.58	0/5855	0.75	17/7852 (0.2%)

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	584	ARG	NE-CZ-NH2	-11.10	114.75	120.30
1	A	629	ARG	NE-CZ-NH2	-9.02	115.79	120.30
2	B	629	ARG	NE-CZ-NH1	-8.36	116.12	120.30
2	B	584	ARG	NE-CZ-NH1	7.88	124.24	120.30
1	A	553	ARG	NE-CZ-NH2	-7.29	116.65	120.30

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	2927	2976	2976	13	0
2	B	2912	2974	2973	3	0
3	B	5	0	0	0	0
4	A	237	0	0	5	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	184	0	0	0	0
All	All	6265	5950	5949	15	1

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 1.

The worst 5 of 15 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:688:GLY:HA3	1:A:709:VAL:HG12	1.73	0.71
1:A:426:GLU:OE1	4:A:801:HOH:O	2.17	0.58
1:A:407:ARG:NH1	4:A:802:HOH:O	2.24	0.54
1:A:427:ARG:NH2	4:A:801:HOH:O	2.39	0.53
1:A:537:ILE:HG23	1:A:585:ILE:HD11	1.91	0.53

All (1) 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)
4:A:802:HOH:O	4:A:975:HOH:O[3_544]	1.97	0.23

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	346/405 (85%)	333 (96%)	12 (4%)	1 (0%)	41 37
2	B	343/405 (85%)	335 (98%)	8 (2%)	0	100 100
All	All	689/810 (85%)	668 (97%)	20 (3%)	1 (0%)	51 49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	710	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	321/366 (88%)	308 (96%)	13 (4%)	31	29
2	B	319/365 (87%)	312 (98%)	7 (2%)	52	55
All	All	640/731 (88%)	620 (97%)	20 (3%)	40	40

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	490	GLN
2	B	713	ASP
2	B	749	ARG
2	B	747	LYS
1	A	674	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	460	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

9 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	CME	A	516	1	8,9,10	0.66	0	5,9,11	0.88	0
2	CME	B	635	2	8,9,10	0.65	0	5,9,11	0.92	0
2	CME	B	514	2	8,9,10	0.53	0	5,9,11	0.73	0
1	CME	A	514	1	8,9,10	0.42	0	5,9,11	0.69	0
1	CME	A	635	1	8,9,10	1.22	1 (12%)	5,9,11	1.07	0
1	CME	A	470	1	8,9,10	0.66	0	5,9,11	0.52	0
2	CME	B	622	2	8,9,10	0.95	1 (12%)	5,9,11	1.18	1 (20%)
2	CME	B	633	2	8,9,10	0.83	0	5,9,11	0.73	0
2	CME	B	516	2	8,9,10	0.49	0	5,9,11	0.58	0

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
1	CME	A	516	1	-	0/5/8/10	-
2	CME	B	635	2	-	2/5/8/10	-
2	CME	B	514	2	-	0/5/8/10	-
1	CME	A	514	1	-	0/5/8/10	-
1	CME	A	635	1	-	0/5/8/10	-
1	CME	A	470	1	-	2/5/8/10	-
2	CME	B	622	2	-	4/5/8/10	-
2	CME	B	633	2	-	2/5/8/10	-
2	CME	B	516	2	-	0/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	635	CME	CB-CA	2.38	1.59	1.53
2	B	622	CME	CA-N	-2.02	1.42	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	622	CME	CE-SD-SG	2.22	113.66	103.45

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	470	CME	CE-SD-SG-CB
2	B	622	CME	CZ-CE-SD-SG
2	B	635	CME	SD-CE-CZ-OH
1	A	470	CME	SD-CE-CZ-OH
2	B	622	CME	SD-CE-CZ-OH

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	622	CME	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\| > 2$
3	PO4	B	801	-	4,4,4	2.49	1 (25%)	6,6,6	0.75	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	801	PO4	P-O1	4.16	1.60	1.50

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 (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

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	349/405 (86%)	0.44	30 (8%)	10	9	19, 32, 74, 86	0
2	B	346/405 (85%)	0.31	23 (6%)	18	17	21, 35, 66, 89	0
All	All	695/810 (85%)	0.38	53 (7%)	13	13	19, 33, 70, 89	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	712	ALA	12.1
1	A	711	GLY	11.1
1	A	622	CYS	8.9
1	A	710	PRO	8.5
2	B	398	ASN	8.2

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

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
2	CME	B	622	10/11	0.41	0.35	66,74,96,102	0
1	CME	A	470	10/11	0.89	0.17	24,35,84,89	0
1	CME	A	516	10/11	0.93	0.16	23,31,82,82	0
2	CME	B	514	10/11	0.94	0.14	23,32,72,74	0
2	CME	B	516	10/11	0.94	0.13	22,31,82,85	0
1	CME	A	635	10/11	0.94	0.18	21,40,80,88	0
2	CME	B	633	10/11	0.94	0.17	34,57,84,88	0
1	CME	A	514	10/11	0.95	0.12	22,35,83,89	0
2	CME	B	635	10/11	0.95	0.17	29,52,88,91	0

6.3 Carbohydrates (i)

There are no monosaccharides 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. 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	PO4	B	801	5/5	0.72	0.19	70,71,71,72	0

6.5 Other polymers (i)

There are no such residues in this entry.