W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:44 PM GMT

|		
PDB ID	:	4PR9
Title	:	Human Vinculin (residues 891-1066) in complex with PIP
Authors	:	Chinthalapudi, K.; Rangarajan, E.S.; Patil, D.; Izard, T.
Deposited on	:	2014-03-05
Resolution	:	3.20 Å(reported)

---

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 3.20 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		0.282
Clashscore		16
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		7.6%

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
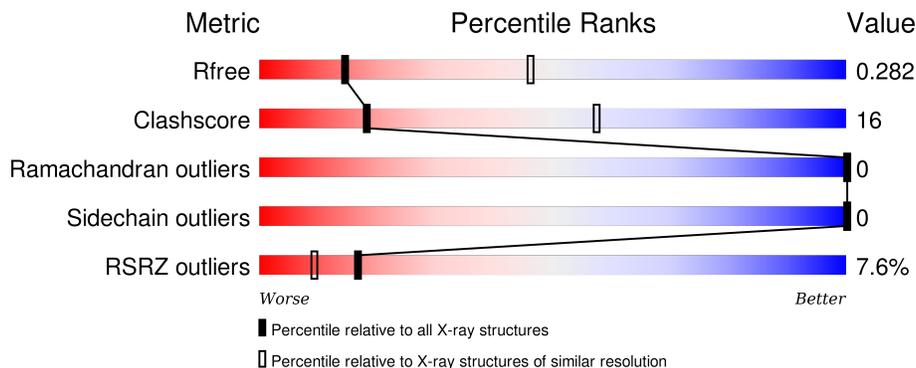
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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. 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	176	6% 81% 19% •
1	B	176	5% 74% 25% •
1	C	176	% 84% 15% •
1	D	176	5% 76% 24%
1	E	176	9% 82% 16% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	176	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8174 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 Vinculin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	175	Total	C	N	O	S	0	0	0
			1358	837	253	256	12			
1	B	176	Total	C	N	O	S	0	0	0
			1361	836	255	258	12			
1	C	174	Total	C	N	O	S	0	0	0
			1346	828	252	254	12			
1	D	176	Total	C	N	O	S	0	0	0
			1367	842	255	258	12			
1	E	172	Total	C	N	O	S	0	0	0
			1327	815	250	250	12			
1	F	160	Total	C	N	O	S	0	0	0
			1235	757	233	233	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1060	ALA	ARG	ENGINEERED MUTATION	UNP P18206
B	1060	ALA	ARG	ENGINEERED MUTATION	UNP P18206
C	1060	ALA	ARG	ENGINEERED MUTATION	UNP P18206
D	1060	ALA	ARG	ENGINEERED MUTATION	UNP P18206
E	1060	ALA	ARG	ENGINEERED MUTATION	UNP P18206
F	1060	ALA	ARG	ENGINEERED MUTATION	UNP P18206

- Molecule 2 is [(2R)-2-OCTANOYLOXY-3-[OXIDANYL-[(1R,2R,3S,4R,5R,6S)-2,3,6-TRIS(OXIDANYL)-4,5-DIPHOSPHONOOXY-CYCLOHEXYL]OXY-PHOSPHORYL]OXY-PR OPYL] OCTANOATE (three-letter code: PIO) (formula: $C_{25}H_{49}O_{19}P_3$).
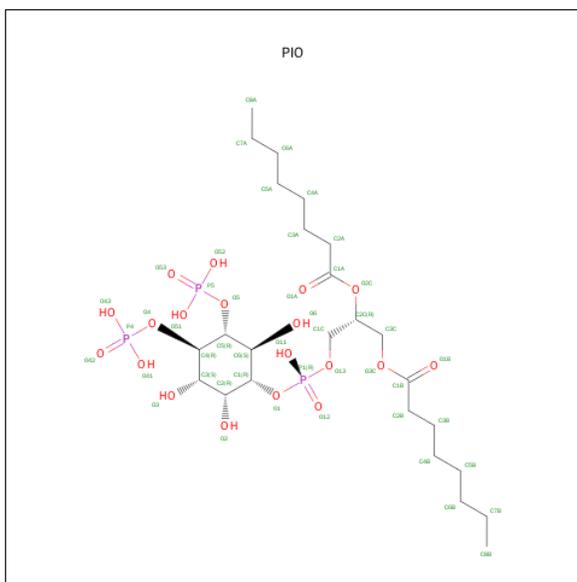
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 38	C 16	O 19	P 3	0	0
2	B	1	Total 37	C 15	O 19	P 3	0	0
2	E	1	Total 47	C 25	O 19	P 3	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
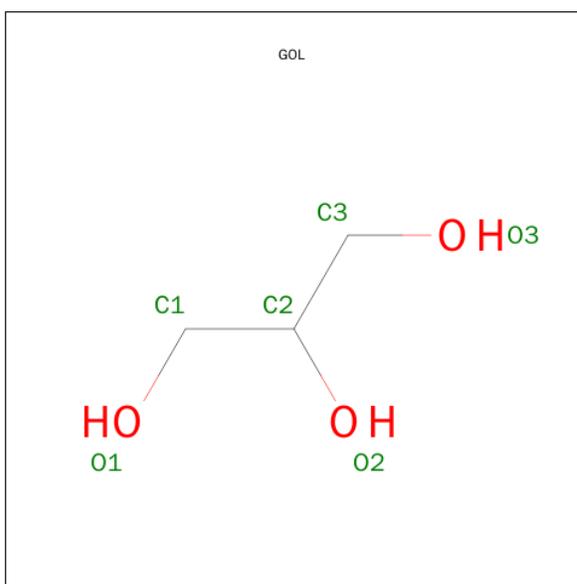
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 6	C 3	O 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total 6	C 3	O 3	0	0
3	F	1	Total 6	C 3	O 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	7	Total 7	O 7	0	0
4	B	18	Total 18	O 18	0	0
4	C	5	Total 5	O 5	0	0
4	D	9	Total 9	O 9	0	0
4	F	1	Total 1	O 1	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: Vinculin
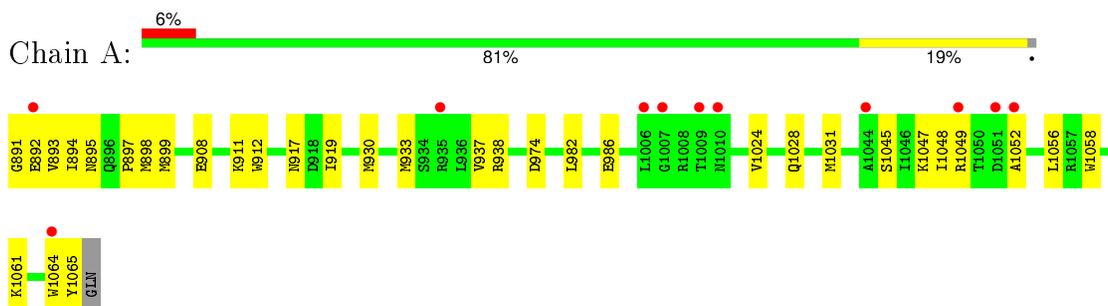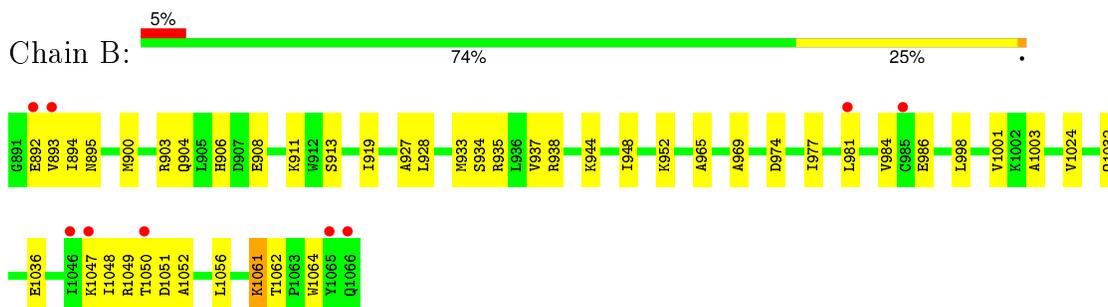
- Molecule 1: Vinculin

- Molecule 1: Vinculin

- Molecule 1: Vinculin

- Molecule 1: Vinculin

Chain E:

- Molecule 1: Vinculin

Chain F:
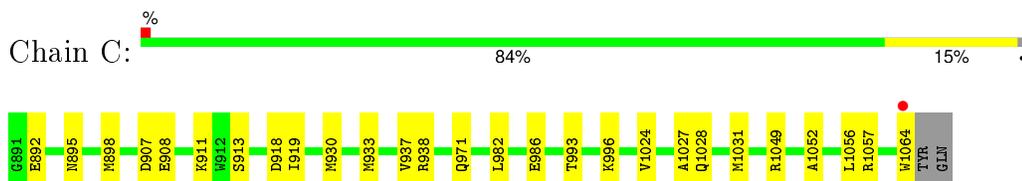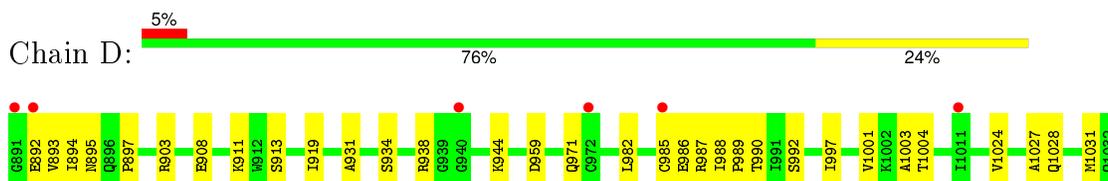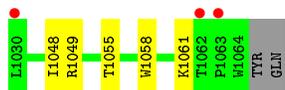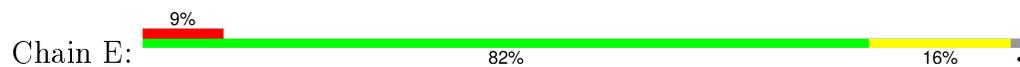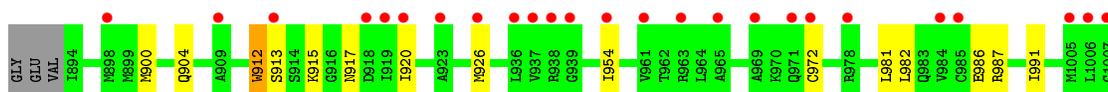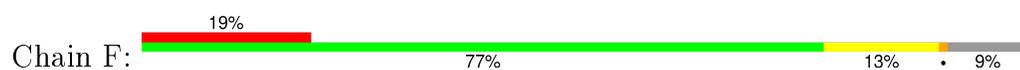
# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α, β, γ	102.58Å 102.58Å 190.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.17 – 3.20 45.17 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.17-3.20) 99.7 (45.17-3.20)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.23 (at 3.19Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, R_{free}	0.204 , 0.250 0.237 , 0.282	Depositor DCC
R_{free} test set	967 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	114.0	Xtriage
Anisotropy	0.192	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 123.6	EDS
Estimated twinning fraction	0.049 for -h,-k,l	Xtriage
L-test for twinning[2]	$<|L|> = 0.48, <L^2> = 0.31$	Xtriage
Outliers	3 of 19179 reflections (0.016%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8174	wwPDB-VP
Average B, all atoms (Å^2)	141.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.69% 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: GOL, PIO

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	0.48	0/1370	0.63	0/1840
1	B	0.45	0/1372	0.57	0/1843
1	C	0.55	0/1357	0.70	2/1822 (0.1%)
1	D	0.46	0/1379	0.64	1/1852 (0.1%)
1	E	0.44	0/1337	0.63	0/1796
1	F	0.45	0/1243	0.60	0/1665
All	All	0.47	0/8058	0.63	3/10818 (0.0%)

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	B	0	1
1	F	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	918	ASP	CB-CG-OD2	5.18	122.97	118.30
1	D	959	ASP	CB-CG-OD2	5.16	122.95	118.30
1	C	907	ASP	CB-CG-OD2	5.13	122.92	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1061	LYS	Peptide
1	F	912	TRP	Peptide

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	1358	0	1416	51	0
1	B	1361	0	1417	66	0
1	C	1346	0	1407	21	0
1	D	1367	0	1424	63	0
1	E	1327	0	1391	34	0
1	F	1235	0	1283	27	0
2	A	38	0	20	4	0
2	B	37	0	18	3	0
2	E	47	0	44	18	0
3	B	6	0	8	0	0
3	C	6	0	8	0	0
3	F	6	0	8	0	0
4	A	7	0	0	0	0
4	B	18	0	0	0	0
4	C	5	0	0	0	0
4	D	9	0	0	0	0
4	F	1	0	0	0	0
All	All	8174	0	8444	259	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 16.

The worst 5 of 259 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:D:985:CYS:HA	1:D:988:ILE:CD1	1.59	1.30
1:B:1049:ARG:O	1:B:1050:THR:HG22	1.21	1.27
1:D:1049:ARG:HD3	1:D:1051:ASP:OD1	1.31	1.24
1:E:932:GLU:OE1	1:E:935:ARG:NH2	1.72	1.22
1:B:969:ALA:CB	1:B:981:LEU:HD22	1.73	1.18

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	173/176 (98%)	167 (96%)	6 (4%)	0	100	100
1	B	174/176 (99%)	169 (97%)	5 (3%)	0	100	100
1	C	172/176 (98%)	168 (98%)	4 (2%)	0	100	100
1	D	174/176 (99%)	168 (97%)	6 (3%)	0	100	100
1	E	170/176 (97%)	163 (96%)	7 (4%)	0	100	100
1	F	154/176 (88%)	150 (97%)	4 (3%)	0	100	100
All	All	1017/1056 (96%)	985 (97%)	32 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	146/147 (99%)	146 (100%)	0	100	100
1	B	146/147 (99%)	146 (100%)	0	100	100
1	C	145/147 (99%)	145 (100%)	0	100	100
1	D	147/147 (100%)	147 (100%)	0	100	100
1	E	143/147 (97%)	143 (100%)	0	100	100
1	F	132/147 (90%)	132 (100%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	859/882 (97%)	859 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1028	GLN
1	D	904	GLN
1	F	1028	GLN
1	C	1032	GLN
1	D	994	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)

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)

6 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	PIO	A	1101	-	38,38,47	0.33	0	48,56,65	1.26	5 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	PIO	B	1101	-	37,37,47	0.33	0	47,55,65	1.27	5 (10%)
3	GOL	B	1102	-	5,5,5	0.12	0	5,5,5	0.22	0
3	GOL	C	1101	-	5,5,5	0.12	0	5,5,5	0.27	0
2	PIO	E	1101	-	47,47,47	0.29	0	57,65,65	1.16	5 (8%)
3	GOL	F	1101	-	5,5,5	0.11	0	5,5,5	0.26	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
2	PIO	A	1101	-	-	1/35/59/68	0/1/1/1
2	PIO	B	1101	-	-	0/34/58/68	0/1/1/1
3	GOL	B	1102	-	-	0/4/4/4	0/0/0/0
3	GOL	C	1101	-	-	0/4/4/4	0/0/0/0
2	PIO	E	1101	-	-	0/44/68/68	0/1/1/1
3	GOL	F	1101	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	1101	PIO	O11-P1-O1	-2.02	98.38	106.49
2	B	1101	PIO	O11-P1-O1	-2.02	98.38	106.49
2	A	1101	PIO	O11-P1-O1	-2.01	98.40	106.49
2	A	1101	PIO	P5-O5-C5	3.19	129.21	121.56
2	E	1101	PIO	P5-O5-C5	3.19	129.22	121.56

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1101	PIO	P5-O5-C5-C6

There are no ring outliers.

3 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1101	PIO	4	0

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1101	PIO	3	0
2	E	1101	PIO	18	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	175/176 (99%)	0.34	11 (6%)	23	13	70, 116, 163, 180	0
1	B	176/176 (100%)	0.30	9 (5%)	32	18	75, 117, 175, 210	0
1	C	174/176 (98%)	0.07	1 (0%)	90	84	68, 99, 151, 176	0
1	D	176/176 (100%)	0.29	8 (4%)	37	23	80, 119, 166, 183	0
1	E	172/176 (97%)	0.59	16 (9%)	11	6	120, 171, 199, 215	0
1	F	160/176 (90%)	1.14	34 (21%)	1	1	164, 207, 226, 241	0
All	All	1033/1056 (97%)	0.45	79 (7%)	17	9	68, 134, 212, 241	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	1009	THR	7.2
1	F	918	ASP	6.4
1	A	1051	ASP	4.9
1	F	1006	LEU	4.9
1	F	909	ALA	4.7

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	PIO	B	1101	37/47	0.60	0.34	1.44	205,218,228,228	37
2	PIO	E	1101	47/47	0.68	0.39	0.42	172,208,232,233	0
2	PIO	A	1101	38/47	0.61	0.27	0.05	201,230,242,242	0
3	GOL	F	1101	6/6	0.50	0.38	-	143,145,146,147	0
3	GOL	B	1102	6/6	0.45	0.41	-	175,176,177,178	0
3	GOL	C	1101	6/6	0.88	0.20	-	143,144,145,146	0

6.5 Other polymers (i)

There are no such residues in this entry.