Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 09:09 PM GMT

PDB ID	:	4V3E
Title	:	The CIDRa domain from IT4var07 PfEMP1 bound to endothelial protein C receptor
Authors	:	Lau, C.K.Y.; Turner, L.; Jespersen, J.S.; Lowe, E.D.; Petersen, B.; Wang, C.W.; Petersen, J.E.V.; Lusingu, J.; Theander, T.G.; Lavstsen, T.; Higgins, M.K.
Deposited on	:	2014-10-17
Resolution	:	2.90 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 (i) symbol.

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

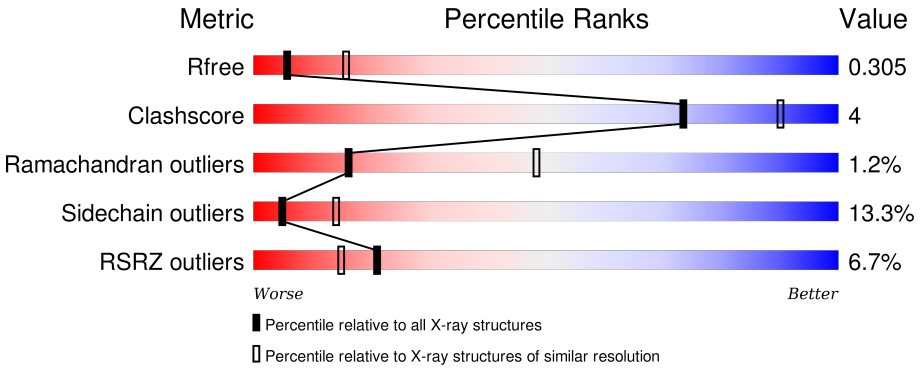
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	251	6% 60% 13% • 24%
2	B	171	6% 79% 17% • •

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3126 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.

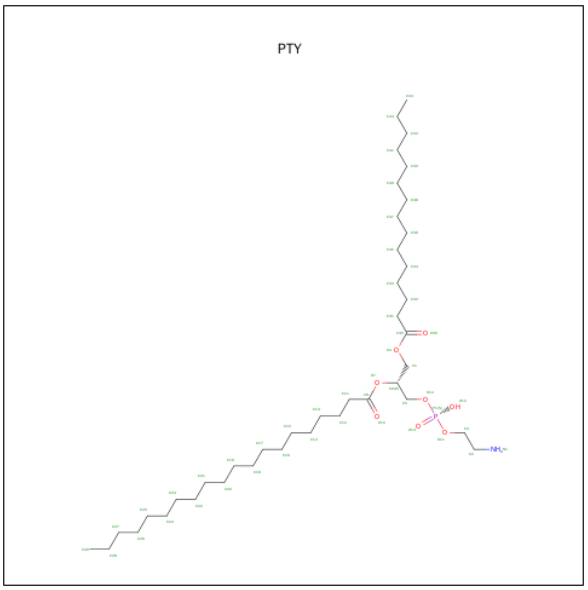
- Molecule 1 is a protein called IT4VAR07 CIDRA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total	C	N	O	S	0	0	0
			1614	1036	263	304	11			

- Molecule 2 is a protein called ENDOTHELIAL PROTEIN C RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	167	Total	C	N	O	S	0	0	0
			1367	873	243	247	4			

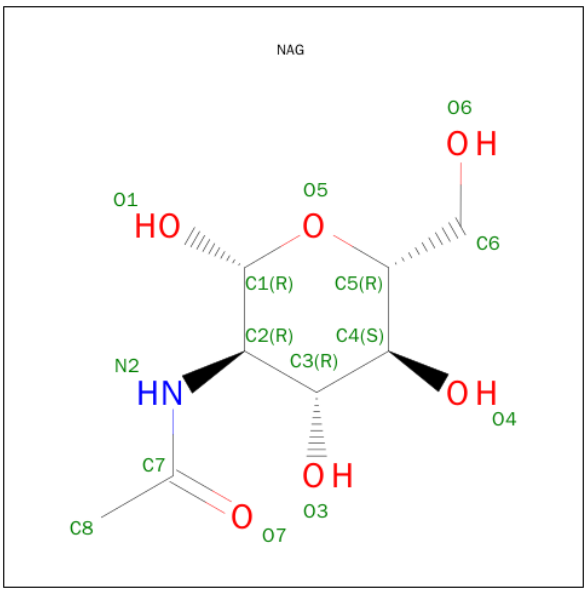
- Molecule 3 is PHOSPHATIDYLETHANOLAMINE (three-letter code: PTY) (formula: $C_{40}H_{80}NO_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	P	50	0
			50	40	1	8	1		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:

$C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

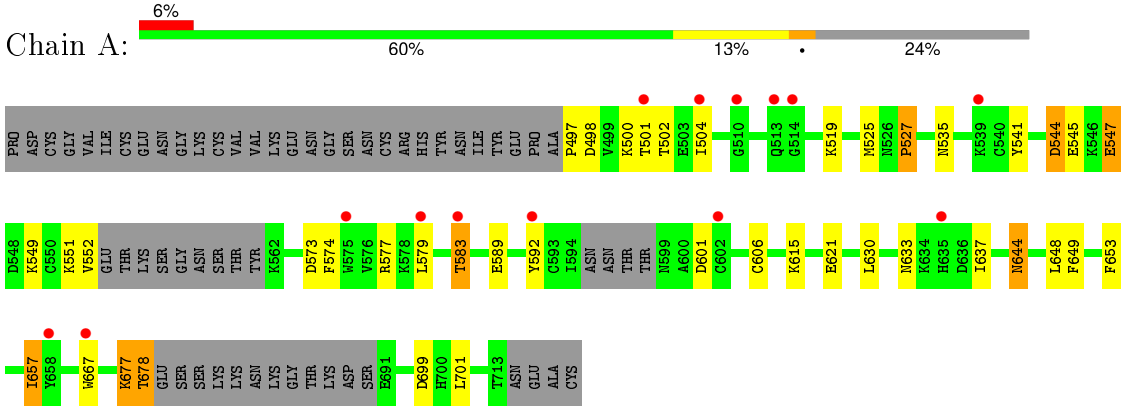
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	2	Total	C	N	O	0	0
			28	16	2	10		

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.

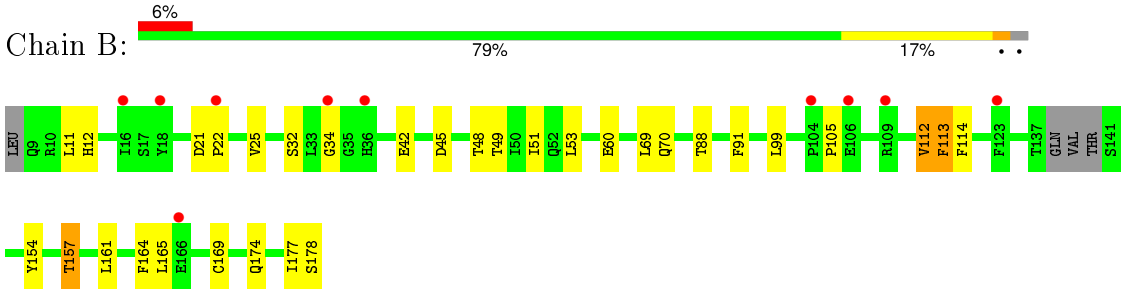
- Molecule 1: IT4VAR07 CIDRA

Chain A:



- Molecule 2: ENDOTHELIAL PROTEIN C RECEPTOR

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	56.02Å 56.02Å 250.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	83.50 – 2.90 48.51 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (83.50-2.90) 99.8 (48.51-2.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.74 (at 2.91Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.247 , 0.276 0.259 , 0.305	Depositor DCC
R_{free} test set	520 reflections (5.04%)	DCC
Wilson B-factor (Å^2)	103.0	Xtriage
Anisotropy	0.413	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 97.7	EDS
Estimated twinning fraction	0.047 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 10832 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3126	wwPDB-VP
Average B, all atoms (Å^2)	122.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, PTY, MAN

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	0.55	0/1646	0.80	2/2208 (0.1%)
2	B	0.54	0/1404	0.74	0/1909
All	All	0.55	0/3050	0.77	2/4117 (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	A	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	678	THR	N-CA-C	-7.24	91.45	111.00
1	A	677	LYS	N-CA-C	7.08	130.12	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	497	PRO	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	1614	0	1578	13	0
2	B	1367	0	1315	12	0
3	B	50	0	79	0	0
4	B	28	0	26	0	0
5	B	39	0	34	0	0
6	B	28	0	25	0	0
All	All	3126	0	3057	25	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 4.

All (25) 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:579:LEU:O	1:A:583:THR:HG23	1.63	0.95
1:A:644:ASN:H	1:A:644:ASN:HD22	1.34	0.76
1:A:527:PRO:HB3	1:A:615:LYS:HD3	1.69	0.72
1:A:653:PHE:O	1:A:657:ILE:HB	1.94	0.68
1:A:677:LYS:HE3	1:A:699:ASP:HB3	1.76	0.67
2:B:112:VAL:HG21	2:B:165:LEU:HB3	1.79	0.63
2:B:22:PRO:HG3	2:B:91:PHE:CE2	2.35	0.62
2:B:112:VAL:CG2	2:B:165:LEU:HB3	2.33	0.59
2:B:161:LEU:O	2:B:165:LEU:HD23	2.05	0.57
1:A:649:PHE:HA	1:A:667:TRP:CZ2	2.42	0.55
2:B:154:TYR:HB2	2:B:157:THR:HG22	1.91	0.53
1:A:644:ASN:H	1:A:644:ASN:ND2	2.05	0.51
1:A:573:ASP:OD1	1:A:577:ARG:NH1	2.44	0.51
2:B:174:GLN:O	2:B:178:SER:HB3	2.13	0.49
2:B:25:VAL:CG1	2:B:45:ASP:HB3	2.42	0.49
2:B:25:VAL:HG12	2:B:45:ASP:HB3	1.96	0.46
1:A:544:ASP:HB2	1:A:547:GLU:H	1.81	0.46
1:A:501:THR:HG22	1:A:541:TYR:HD1	1.81	0.45
1:A:501:THR:HG22	1:A:541:TYR:CD1	2.53	0.44
1:A:583:THR:HB	1:A:701:LEU:HD21	1.99	0.43
2:B:21:ASP:HB2	2:B:22:PRO:CD	2.49	0.42
1:A:649:PHE:HA	1:A:667:TRP:HZ2	1.81	0.41
2:B:164:PHE:HD2	2:B:165:LEU:HD22	1.85	0.41
2:B:165:LEU:O	2:B:169:CYS:HB3	2.20	0.40
2:B:113:PHE:CD1	2:B:113:PHE:N	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	184/251 (73%)	172 (94%)	11 (6%)	1 (0%)	34 71
2	B	163/171 (95%)	151 (93%)	9 (6%)	3 (2%)	11 37
All	All	347/422 (82%)	323 (93%)	20 (6%)	4 (1%)	16 48

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	527	PRO
2	B	34	GLY
2	B	105	PRO
2	B	88	THR

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	183/235 (78%)	156 (85%)	27 (15%)	4 11
2	B	149/153 (97%)	132 (89%)	17 (11%)	7 21
All	All	332/388 (86%)	288 (87%)	44 (13%)	5 14

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	498	ASP
1	A	500	LYS
1	A	502	THR
1	A	504	ILE
1	A	519	LYS
1	A	525	MET
1	A	535	ASN
1	A	544	ASP
1	A	545	GLU
1	A	547	GLU
1	A	549	LYS
1	A	551	LYS
1	A	552	VAL
1	A	574	PHE
1	A	583	THR
1	A	589	GLU
1	A	592	TYR
1	A	601	ASP
1	A	606	CYS
1	A	621	GLU
1	A	630	LEU
1	A	633	ASN
1	A	637	ILE
1	A	644	ASN
1	A	648	LEU
1	A	657	ILE
1	A	678	THR
2	B	11	LEU
2	B	12	HIS
2	B	32	SER
2	B	42	GLU
2	B	48	THR
2	B	49	THR
2	B	51	ILE
2	B	53	LEU
2	B	60	GLU
2	B	69	LEU
2	B	70	GLN
2	B	99	LEU
2	B	112	VAL
2	B	113	PHE
2	B	114	PHE
2	B	157	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	177	ILE

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	505	ASN
1	A	534	ASN
1	A	607	ASN
1	A	644	ASN
2	B	171	GLN

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

5 carbohydrates 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	$\#\|Z\| > 2$
5	NAG	B	202	2,5	14,14,15	0.35	0	15,19,21	0.72	0
5	NAG	B	203	5	14,14,15	0.36	0	15,19,21	1.88	2 (13%)
6	NAG	B	204	2,6	14,14,15	0.45	0	15,19,21	1.67	1 (6%)
6	NAG	B	205	6	14,14,15	0.37	0	15,19,21	1.09	1 (6%)
5	MAN	B	207	5	11,11,12	0.51	0	14,15,17	1.15	2 (14%)

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
5	NAG	B	202	2,5	-	0/6/23/26	0/1/1/1
5	NAG	B	203	5	-	0/6/23/26	0/1/1/1
6	NAG	B	204	2,6	-	0/6/23/26	0/1/1/1
6	NAG	B	205	6	-	0/6/23/26	0/1/1/1
5	MAN	B	207	5	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	207	MAN	C1-O5-C5	2.27	115.13	112.25
5	B	203	NAG	O4-C4-C3	2.54	116.05	110.34
5	B	207	MAN	C1-C2-C3	3.41	113.57	109.54
6	B	205	NAG	C1-O5-C5	3.73	116.99	112.25
6	B	204	NAG	C1-O5-C5	5.70	119.49	112.25
5	B	203	NAG	C1-O5-C5	6.56	120.57	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry (i)

3 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	$\#\|Z\| > 2$
3	PTY	B	200	-	48,49,49	0.32	0	49,54,54	0.61	2 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	B	201	2	14,14,15	0.38	0	15,19,21	1.15	1 (6%)
4	NAG	B	206	2	14,14,15	0.34	0	15,19,21	0.95	1 (6%)

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	PTY	B	200	-	-	0/53/53/53	0/0/0/0
4	NAG	B	201	2	-	0/6/23/26	0/1/1/1
4	NAG	B	206	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	200	PTY	O14-P1-O13	-2.48	99.99	109.62
3	B	200	PTY	O11-P1-O13	-2.29	100.71	109.62
4	B	206	NAG	C1-O5-C5	3.42	116.59	112.25
4	B	201	NAG	C1-O5-C5	3.99	117.31	112.25

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	192/251 (76%)	0.52	14 (7%)	18	12	74, 126, 184, 207	0
2	B	167/171 (97%)	0.53	10 (5%)	25	18	69, 119, 155, 180	0
All	All	359/422 (85%)	0.52	24 (6%)	21	15	69, 123, 173, 207	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	513	GLN	8.6
2	B	106	GLU	6.3
1	A	514	GLY	4.6
1	A	602	CYS	4.6
1	A	592	TYR	4.0
2	B	18	TYR	4.0
1	A	635	HIS	3.7
2	B	36	HIS	3.4
1	A	510	GLY	3.1
1	A	667	TRP	2.8
1	A	504	ILE	2.7
1	A	658	TYR	2.7
1	A	583	THR	2.7
1	A	539	LYS	2.5
2	B	34	GLY	2.5
2	B	16	ILE	2.5
1	A	575	TRP	2.5
2	B	22	PRO	2.4
1	A	501	THR	2.3
2	B	109	ARG	2.2
2	B	123	PHE	2.2
2	B	166	GLU	2.1
1	A	579	LEU	2.0
2	B	104	PRO	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)

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, 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	LLDF	B-factors(Å^2)	Q<0.9
5	NAG	B	202	14/15	0.93	0.14	-1.22	123,139,151,161	0
6	NAG	B	205	14/15	0.62	0.31	-	193,200,205,207	0
5	MAN	B	207	11/12	0.83	0.19	-	199,202,205,205	0
5	NAG	B	203	14/15	0.92	0.14	-	171,180,189,196	0
6	NAG	B	204	14/15	0.68	0.29	-	187,195,201,204	0

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, 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	LLDF	B-factors(Å^2)	Q<0.9
4	NAG	B	206	14/15	0.43	0.28	1.28	169,175,190,193	0
4	NAG	B	201	14/15	0.81	0.19	-1.06	125,128,129,129	0
3	PTY	B	200	50/50	-	-	-	18,20,31,36	50

6.5 Other polymers (i)

There are no such residues in this entry.