wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 04:24 PM GMT

PDB ID	:	4EOK
Title	:	Thr 160 phosphorylated CDK2 H84S, Q85M, K89D - human cyclin A3 complex with the inhibitor NU6102
Authors	:	Echalier, A.; Cot, E.; Camasses, A.; Hodimont, E.; Hoh, F.; Sheinerman, F.; Krasinska, L.; Fisher, D.
Deposited on	:	2012-04-14
Resolution	:	2.57 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
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:

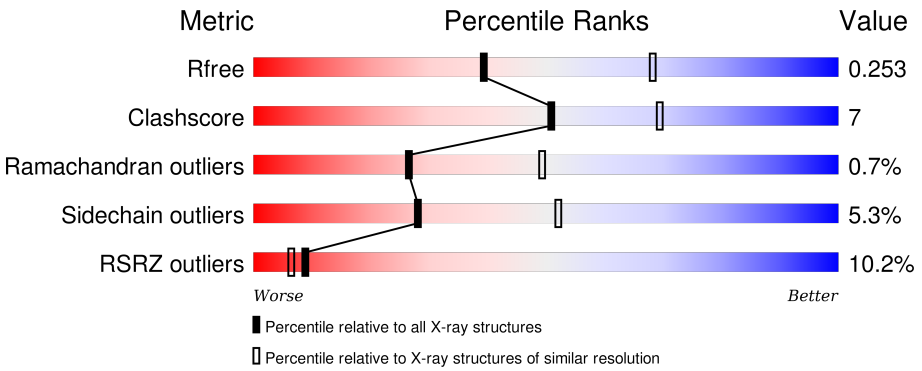
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.57 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	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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	300	3%, 83%, 14%
1	C	300	15%, 83%, 14%
2	B	258	6%, 85%, 13%
2	D	258	16%, 86%, 12%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8994 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 Cyclin-dependent kinase 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	294	2362	1534	400	418	1	9	0	0	0
1	C	300	2405	1559	406	430	1	9	0	0	0

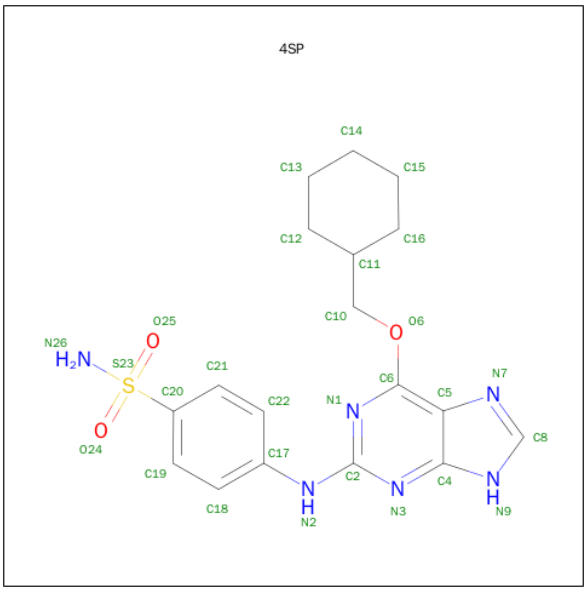
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	LEU	-	EXPRESSION TAG	UNP P24941
A	-1	GLY	-	EXPRESSION TAG	UNP P24941
A	0	SER	-	EXPRESSION TAG	UNP P24941
A	84	SER	HIS	ENGINEERED MUTATION	UNP P24941
A	85	MET	GLN	ENGINEERED MUTATION	UNP P24941
A	89	ASP	LYS	ENGINEERED MUTATION	UNP P24941
C	-2	LEU	-	EXPRESSION TAG	UNP P24941
C	-1	GLY	-	EXPRESSION TAG	UNP P24941
C	0	SER	-	EXPRESSION TAG	UNP P24941
C	84	SER	HIS	ENGINEERED MUTATION	UNP P24941
C	85	MET	GLN	ENGINEERED MUTATION	UNP P24941
C	89	ASP	LYS	ENGINEERED MUTATION	UNP P24941

- Molecule 2 is a protein called Cyclin-A2.

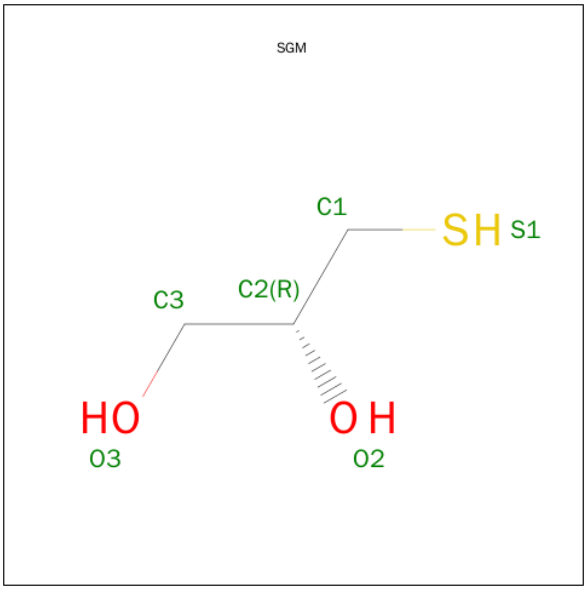
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	258	2083	1350	339	383	11	0	0	0
2	D	256	2068	1339	337	381	11	0	0	0

- Molecule 3 is O6-CYCLOHEXYLMETHOXY-2-(4'-SULPHAMOYLANILINO) PURINE (three-letter code: 4SP) (formula: $C_{18}H_{22}N_6O_3S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	28	18	6	3	1	0	0
			Total	C	N	O	S		
3	C	1	28	18	6	3	1	0	0

- Molecule 4 is MONOTHIOGLYCEROL (three-letter code: SGM) (formula: $C_3H_8O_2S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
4	B	1	6	3	2	1	0	0
			Total	C	O	S		
4	D	1	6	3	2	1	0	0

- Molecule 5 is water.

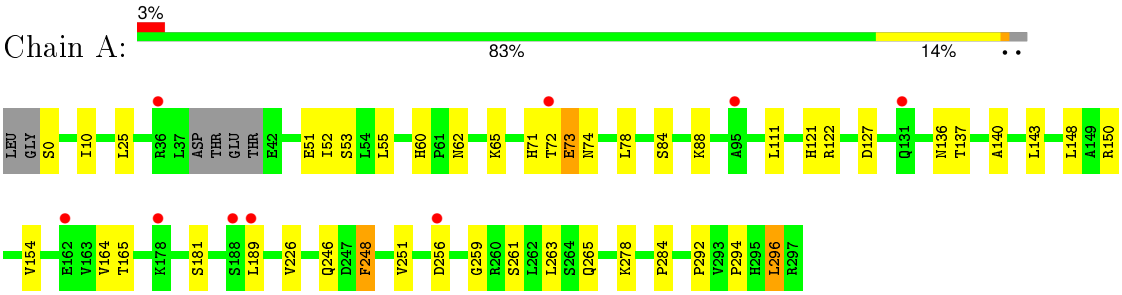
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	7	Total O 7 7	0	0
5	B	1	Total O 1 1	0	0

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.

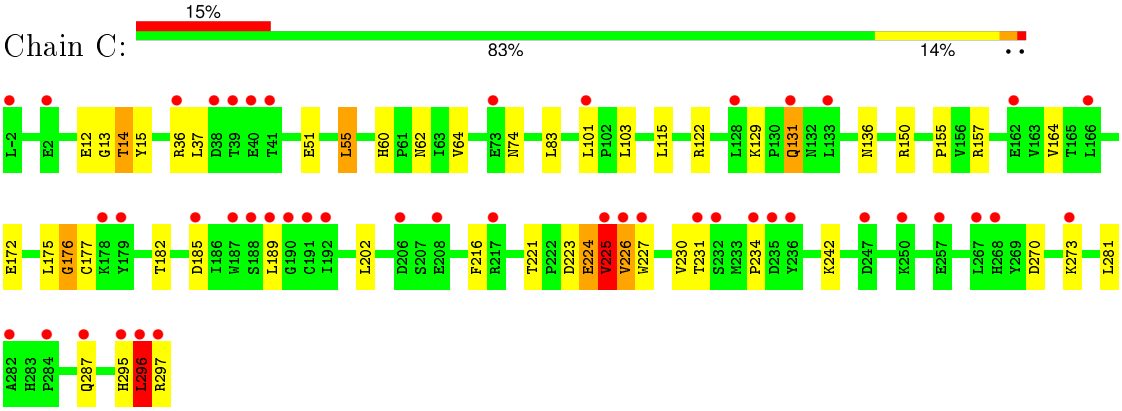
- Molecule 1: Cyclin-dependent kinase 2

Chain A:



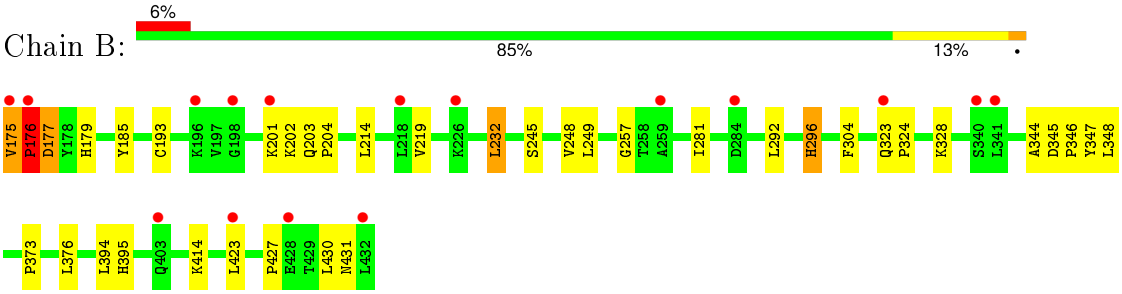
- Molecule 1: Cyclin-dependent kinase 2

Chain C:



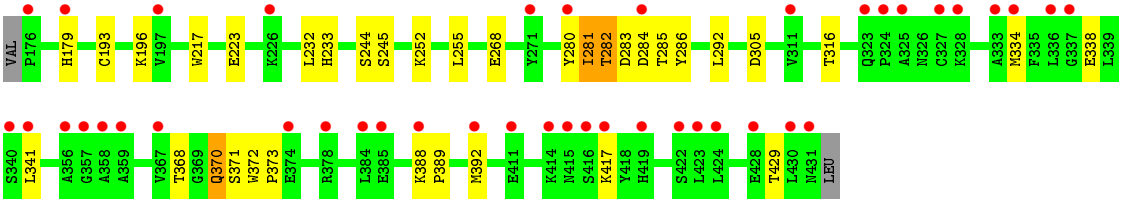
- Molecule 2: Cyclin-A2

Chain B:



- Molecule 2: Cyclin-A2

Chain D:


VAL
P176
H179
C193
K196
V197
W217
E223
K226
L232
H233
S244
S245
K252
L255
E268
Y271
Y280
I281
T282
D283
D284
T285
Y286
L292
D305
V311
T316
Q323
P324
A325
N326
G327
K328
A333
N334
F335
L336
G337
E338
L339

S340
L341
A356
G357
A358
A359
V367
T368
G369
Q370
S371
W372
P373
E374
R378
L384
E385
K388
P389
M392
E411
K414
N415
S416
K417
Y418
H419
S422
L423
L424
E428
T429
L430
N431
LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	74.23Å 136.01Å 171.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.57 29.23 – 2.57	Depositor EDS
% Data completeness (in resolution range)	99.2 (30.00-2.57) 99.2 (29.23-2.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.57Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.223 , 0.259 0.220 , 0.253	Depositor DCC
R_{free} test set	2845 reflections (5.39%)	DCC
Wilson B-factor (Å^2)	52.3	Xtriage
Anisotropy	0.180	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 55634 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8994	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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: TPO, SGM, 4SP

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	0.43	0/2410	0.58	0/3267
1	C	0.47	0/2454	0.60	3/3329 (0.1%)
2	B	0.48	0/2133	0.56	0/2897
2	D	0.38	0/2118	0.52	0/2875
All	All	0.44	0/9115	0.56	3/12368 (0.0%)

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	296	LEU	CB-CA-C	-9.99	91.22	110.20
1	C	297	ARG	N-CA-CB	8.68	126.23	110.60
1	C	224	GLU	N-CA-C	-5.17	97.05	111.00

There are no chirality outliers.

There are no planarity outliers.

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	2362	0	2408	27	0
1	C	2405	0	2447	59	0
2	B	2083	0	2106	26	0
2	D	2068	0	2088	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	28	0	22	0	0
3	C	28	0	22	2	0
4	B	6	0	7	1	0
4	D	6	0	8	2	0
5	A	7	0	0	0	0
5	B	1	0	0	0	0
All	All	8994	0	9108	125	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.

The worst 5 of 125 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)
2:B:193:CYS:SG	4:B:501:SGM:S1	2.32	1.27
1:C:224:GLU:OE2	1:C:231:THR:OG1	1.64	1.14
1:C:225:VAL:O	1:C:226:VAL:HG12	1.47	1.12
1:C:224:GLU:O	1:C:225:VAL:HG12	1.56	1.05
1:C:227:TRP:O	1:C:230:VAL:HG23	1.57	1.04

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	289/300 (96%)	272 (94%)	15 (5%)	2 (1%)	26 49
1	C	297/300 (99%)	280 (94%)	13 (4%)	4 (1%)	15 29
2	B	256/258 (99%)	250 (98%)	5 (2%)	1 (0%)	39 63
2	D	254/258 (98%)	240 (94%)	13 (5%)	1 (0%)	39 63
All	All	1096/1116 (98%)	1042 (95%)	46 (4%)	8 (1%)	26 49

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	176	PRO
1	C	176	GLY
1	C	164	VAL
1	C	225	VAL
1	A	164	VAL

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	258/263 (98%)	244 (95%)	14 (5%)	27	51
1	C	263/263 (100%)	245 (93%)	18 (7%)	20	38
2	B	232/232 (100%)	221 (95%)	11 (5%)	32	58
2	D	230/232 (99%)	221 (96%)	9 (4%)	39	66
All	All	983/990 (99%)	931 (95%)	52 (5%)	28	52

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	296	HIS
1	C	55	LEU
2	D	292	LEU
2	B	328	LYS
1	C	14	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	317	GLN
2	B	395	HIS
2	D	254	GLN
2	B	233	HIS
2	D	233	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 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	TPO	A	160	1	8,10,11	0.64	0	7,14,16	1.46	0
1	TPO	C	160	1	8,10,11	0.65	0	7,14,16	1.42	1 (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
1	TPO	A	160	1	-	0/8/11/13	0/0/0/0
1	TPO	C	160	1	-	0/8/11/13	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	160	TPO	O-C-CA	-2.02	120.11	125.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

4 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	4SP	A	301	-	28,31,31	2.38	7 (25%)	35,44,44	2.27	9 (25%)
4	SGM	B	501	-	5,5,5	0.40	0	5,5,5	0.66	0
3	4SP	C	301	-	28,31,31	2.51	8 (28%)	35,44,44	2.16	8 (22%)
4	SGM	D	501	-	5,5,5	0.42	0	5,5,5	0.62	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
3	4SP	A	301	-	-	0/15/23/23	0/4/4/4
4	SGM	B	501	-	-	0/4/4/4	0/0/0/0
3	4SP	C	301	-	-	0/15/23/23	0/4/4/4
4	SGM	D	501	-	-	0/4/4/4	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	301	4SP	C20-S23	-10.42	1.61	1.77
3	A	301	4SP	C20-S23	-9.93	1.62	1.77
3	A	301	4SP	S23-N26	-3.75	1.52	1.60
3	C	301	4SP	S23-N26	-3.46	1.52	1.60
3	C	301	4SP	C17-N2	-2.03	1.36	1.40

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	301	4SP	O25-S23-O24	-7.16	108.74	118.80
3	C	301	4SP	O25-S23-O24	-5.98	110.40	118.80
3	A	301	4SP	N3-C2-N1	-3.89	120.26	126.22
3	A	301	4SP	C4-C5-N7	-3.74	106.04	109.48
3	C	301	4SP	C4-C5-N7	-3.49	106.27	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	501	SGM	1	0
3	C	301	4SP	2	0
4	D	501	SGM	2	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	293/300 (97%)	0.06	9 (3%) 52 48	29, 43, 71, 84	2 (0%)
1	C	299/300 (99%)	0.71	46 (15%) 3 2	54, 75, 109, 130	1 (0%)
2	B	258/258 (100%)	0.23	16 (6%) 24 20	33, 54, 75, 86	0
2	D	256/258 (99%)	0.82	42 (16%) 2 1	46, 75, 119, 136	0
All	All	1106/1116 (99%)	0.45	113 (10%) 9 6	29, 61, 107, 136	3 (0%)

The worst 5 of 113 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	175	VAL	9.5
1	C	40	GLU	7.8
2	D	428	GLU	6.9
2	D	423	LEU	5.0
1	C	178	LYS	4.8

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	TPO	A	160	11/12	0.97	0.11	-	39,42,43,43	0
1	TPO	C	160	11/12	0.95	0.12	-	65,71,74,75	0

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
4	SGM	B	501	6/6	0.86	0.27	1.87	68,68,69,70	0
3	4SP	C	301	28/28	0.87	0.19	0.18	68,69,75,75	0
4	SGM	D	501	6/6	0.86	0.18	0.18	79,79,80,81	0
3	4SP	A	301	28/28	0.95	0.13	-0.62	43,47,52,53	0

6.5 Other polymers (i)

There are no such residues in this entry.