wwPDB X-ray Structure Validation Summary Report (i)

Jan 4, 2024 – 03:00 pm GMT

PDB ID	:	5A4T
Title	:	DYRK1A IN COMPLEX WITH NITRILE BENZOTHIAZOLE FRAGMENT
Authors	:	Rothweiler, U.
Deposited on	:	2015-06-15
Resolution	:	2.15 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 (i)) were used in the production of this report:

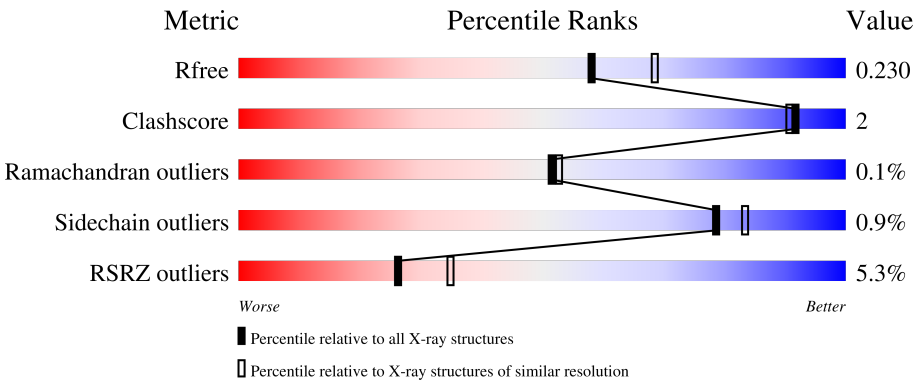
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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 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	368	2%, 90%, • 7%
1	B	368	6%, 88%, • 8%
1	C	368	6%, 86%, • 10%
1	D	368	6%, 85%, 5% 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11182 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.

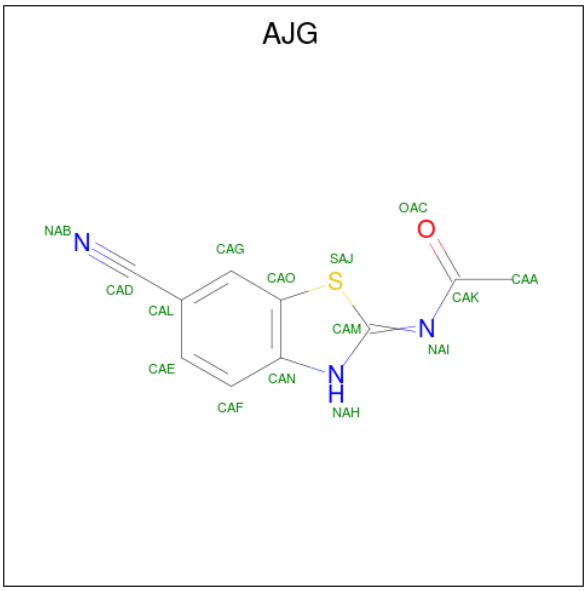
- Molecule 1 is a protein called DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION-REGULATED KINASE 1A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	342	2794	1798	477	501	1	17	0	0	0
1	B	340	2761	1777	471	495	1	17	0	0	0
1	C	333	2678	1724	452	484	1	17	0	0	0
1	D	333	2711	1752	455	486	1	17	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	GLY	-	expression tag	UNP Q13627
A	124	ALA	-	expression tag	UNP Q13627
A	125	SER	-	expression tag	UNP Q13627
B	123	GLY	-	expression tag	UNP Q13627
B	124	ALA	-	expression tag	UNP Q13627
B	125	SER	-	expression tag	UNP Q13627
C	123	GLY	-	expression tag	UNP Q13627
C	124	ALA	-	expression tag	UNP Q13627
C	125	SER	-	expression tag	UNP Q13627
D	123	GLY	-	expression tag	UNP Q13627
D	124	ALA	-	expression tag	UNP Q13627
D	125	SER	-	expression tag	UNP Q13627

- Molecule 2 is N-(6-CYANO-3H-1,3-BENZOTHIAZOL-2-YLIDENE)ETHANAMIDE (three-letter code: AJG) (formula: $C_{10}H_7N_3OS$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			15	10	3	1	1		
2	B	1	Total	C	N	O	S	0	0
			15	10	3	1	1		
2	C	1	Total	C	N	O	S	0	0
			15	10	3	1	1		
2	D	1	Total	C	N	O	S	0	0
			15	10	3	1	1		

- Molecule 3 is water.

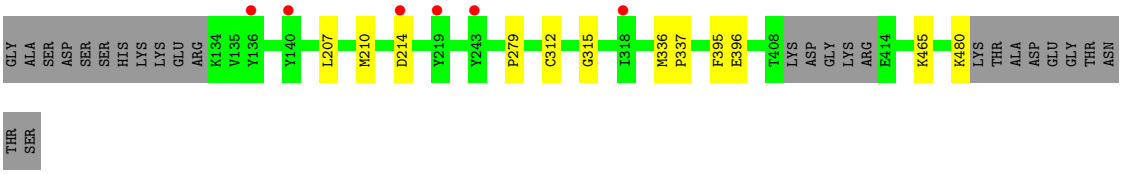
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	70	Total	O	0	0
			70	70		
3	B	35	Total	O	0	0
			35	35		
3	C	36	Total	O	0	0
			36	36		
3	D	37	Total	O	0	0
			37	37		

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.

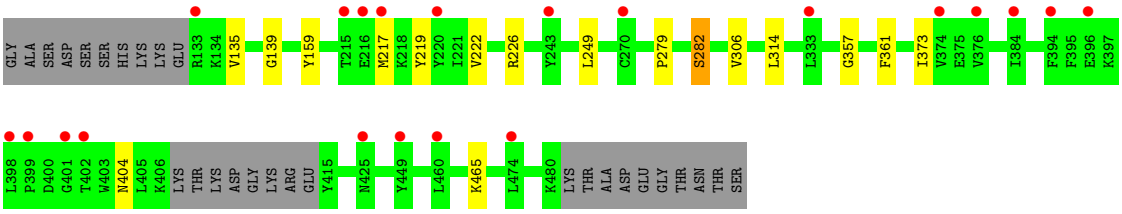
• Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A

Chain A:



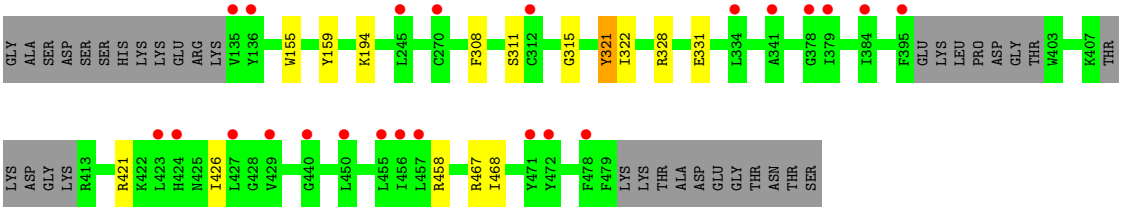
• Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A

Chain B:



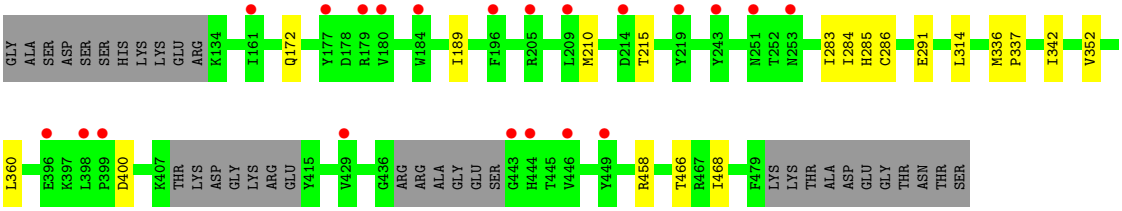
• Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A

Chain C:



• Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A

Chain D:


GLY
ALA
SER
ASP
SER
SER
HIS
LYS
LYS
GLU
ARG
K134
I161
Q172
Y177
D178
R179
V180
W184
I189
F196
R205
L209
M210
D214
T215
Y219
Y243
N251
T252
N253
I283
I284
H285
C286
E291
L314
M336
P337
I342
V352
L360
E396
K397
L398
P399
D400
K407
THR
LYS
ASP
GLY
LYS
ARG
GLU
Y415
V429
G436
ARG
ARG
ALA
GLY
GLU
SER
G443
H444
T445
V446
Y449
R458
T466
R467
I468
F479
LYS
LYS
THR
ALA
ASP
GLU
GLY
THR
ASN
THR
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	87.75Å 87.70Å 229.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.19 – 2.15 48.19 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.19-2.15) 98.7 (48.19-2.05)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.27 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.204 , 0.234 0.200 , 0.230	Depositor DCC
R_{free} test set	5538 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	39.9	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 19.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.37$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	0.426 for k,h,-l	Xtriage
Reported twinning fraction	0.571 for H, K, L 0.429 for K, H, -L	Depositor
Outliers	0 of 110551 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11182	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: AJG, PTR

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	0.35	0/2842	0.53	0/3833
1	B	0.31	0/2808	0.51	0/3789
1	C	0.33	0/2724	0.52	0/3686
1	D	0.31	0/2758	0.50	0/3723
All	All	0.32	0/11132	0.51	0/15031

There are no bond length outliers.

There are no bond angle outliers.

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	2794	0	2787	8	0
1	B	2761	0	2745	8	0
1	C	2678	0	2599	8	0
1	D	2711	0	2693	9	0
2	A	15	0	0	1	0
2	B	15	0	0	1	0
2	C	15	0	0	1	0
2	D	15	0	0	1	0
3	A	70	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	35	0	0	0	0
3	C	36	0	0	0	0
3	D	37	0	0	0	0
All	All	11182	0	10824	34	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 2.

The worst 5 of 34 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:458:ARG:NH2	1:D:466:THR:O	2.16	0.76
1:A:395:PHE:HB3	3:A:2053:HOH:O	1.88	0.72
2:B:1482:AJG:SAJ	2:B:1482:AJG:CAA	2.85	0.64
2:A:1482:AJG:SAJ	2:A:1482:AJG:CAA	2.88	0.62
2:C:1481:AJG:SAJ	2:C:1481:AJG:CAA	2.91	0.59

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	337/368 (92%)	317 (94%)	20 (6%)	0	100 100
1	B	335/368 (91%)	312 (93%)	22 (7%)	1 (0%)	41 37
1	C	326/368 (89%)	308 (94%)	18 (6%)	0	100 100
1	D	326/368 (89%)	305 (94%)	21 (6%)	0	100 100
All	All	1324/1472 (90%)	1242 (94%)	81 (6%)	1 (0%)	51 53

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	217	MET

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	301/324 (93%)	299 (99%)	2 (1%)	84	89
1	B	296/324 (91%)	292 (99%)	4 (1%)	67	72
1	C	282/324 (87%)	281 (100%)	1 (0%)	91	93
1	D	292/324 (90%)	288 (99%)	4 (1%)	67	72
All	All	1171/1296 (90%)	1160 (99%)	11 (1%)	78	83

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	215	THR
1	D	291	GLU
1	D	400	ASP
1	D	336	MET
1	B	404	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	469	GLN
1	B	260	ASN
1	D	383	HIS
1	A	251	ASN
1	A	211	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 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	PTR	B	321	1	15,16,17	0.81	0	19,22,24	0.88	0
1	PTR	A	321	1	15,16,17	0.77	0	19,22,24	0.87	0
1	PTR	D	321	1	15,16,17	0.69	0	19,22,24	0.76	0
1	PTR	C	321	1	15,16,17	0.95	1 (6%)	19,22,24	0.76	1 (5%)

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	PTR	B	321	1	-	0/10/11/13	0/1/1/1
1	PTR	A	321	1	-	0/10/11/13	0/1/1/1
1	PTR	D	321	1	-	2/10/11/13	0/1/1/1
1	PTR	C	321	1	-	0/10/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	321	PTR	P-OH	2.54	1.63	1.59

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	321	PTR	O3P-P-O2P	2.37	116.71	107.64

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	D	321	PTR	O-C-CA-CB
1	D	321	PTR	CZ-OH-P-O2P

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	321	PTR	1	0

5.5 Carbohydrates (i)

There are no monosaccharides 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	$\#\|Z\| > 2$
2	AJG	A	1482	-	16,16,16	3.36	5 (31%)	21,22,22	3.02	5 (23%)
2	AJG	C	1481	-	16,16,16	2.98	6 (37%)	21,22,22	2.65	5 (23%)
2	AJG	B	1482	-	16,16,16	4.09	6 (37%)	21,22,22	3.28	5 (23%)
2	AJG	D	1481	-	16,16,16	3.63	6 (37%)	21,22,22	2.82	5 (23%)

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	AJG	A	1482	-	-	2/6/6/6	0/2/2/2
2	AJG	C	1481	-	-	2/6/6/6	0/2/2/2
2	AJG	B	1482	-	-	2/6/6/6	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	AJG	D	1481	-	-	4/6/6/6	0/2/2/2

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1482	AJG	CAM-SAJ	-12.02	1.60	1.75
2	D	1481	AJG	CAM-SAJ	-11.40	1.61	1.75
2	B	1482	AJG	CAO-SAJ	-8.65	1.56	1.74
2	C	1481	AJG	CAM-SAJ	-8.48	1.65	1.75
2	A	1482	AJG	CAM-SAJ	-8.42	1.65	1.75

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1482	AJG	CAO-SAJ-CAM	12.59	98.95	88.31
2	A	1482	AJG	CAO-SAJ-CAM	11.35	97.90	88.31
2	D	1481	AJG	CAO-SAJ-CAM	10.53	97.21	88.31
2	C	1481	AJG	CAO-SAJ-CAM	9.13	96.02	88.31
2	C	1481	AJG	CAA-CAK-NAI	3.92	120.92	114.85

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1482	AJG	OAC-CAK-NAI-CAM
2	A	1482	AJG	CAA-CAK-NAI-CAM
2	B	1482	AJG	OAC-CAK-NAI-CAM
2	B	1482	AJG	CAA-CAK-NAI-CAM
2	C	1481	AJG	OAC-CAK-NAI-CAM

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1482	AJG	1	0
2	C	1481	AJG	1	0
2	B	1482	AJG	1	0
2	D	1481	AJG	1	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	341/368 (92%)	0.46	6 (1%)	68	75	23, 36, 56, 68	0
1	B	339/368 (92%)	0.77	21 (6%)	20	27	27, 47, 76, 91	0
1	C	332/368 (90%)	0.75	23 (6%)	16	23	25, 47, 68, 80	0
1	D	332/368 (90%)	0.69	21 (6%)	20	27	25, 47, 72, 89	0
All	All	1344/1472 (91%)	0.67	71 (5%)	26	35	23, 44, 69, 91	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	423	LEU	4.7
1	C	472	TYR	4.7
1	D	446	VAL	4.4
1	C	384	ILE	4.3
1	B	243	TYR	4.0

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. 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	B-factors(Å^2)	Q<0.9
1	PTR	C	321	16/17	0.87	0.16	37,41,43,44	0
1	PTR	D	321	16/17	0.91	0.15	32,35,43,45	0
1	PTR	B	321	16/17	0.93	0.13	42,46,49,49	0
1	PTR	A	321	16/17	0.94	0.15	27,30,35,36	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, 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	B-factors(Å^2)	Q<0.9
2	AJG	D	1481	15/15	0.84	0.20	52,57,64,64	0
2	AJG	B	1482	15/15	0.92	0.18	46,49,59,63	0
2	AJG	C	1481	15/15	0.93	0.13	37,40,42,45	0
2	AJG	A	1482	15/15	0.94	0.12	39,40,45,49	0

6.5 Other polymers (i)

There are no such residues in this entry.