WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 15, 2024 – 05:27 PM EDT

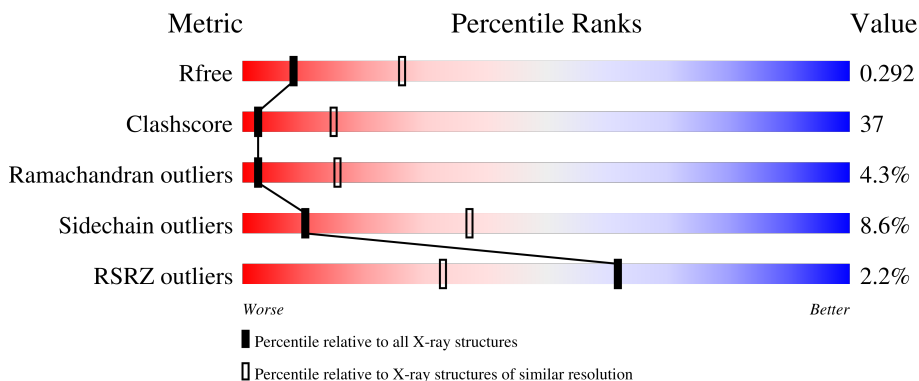|		
PDB ID	:	1YVL
Title	:	Structure of Unphosphorylated STAT1
Authors	:	Mao, X.; Ren, Z.; Parker, G.N.; Sondermann, H.; Pastorello, M.A.; Wang, W.; McMurray, J.S.; Demeler, B.; Darnell Jr., J.E.; Chen, X.
Deposited on	:	2005-02-16
Resolution	:	3.00 Å(reported)

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.00 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	Percentile Ranks	Value
Rfree		0.292
Clashscore		37
Ramachandran outliers		4.3%
Sidechain outliers		8.6%
RSRZ outliers		2.2%

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	683	2% / 39% 49% 7% 5%
1	B	683	3% / 38% 49% 8% · ·
2	C	5	60% 20% 20%
2	D	5	60% 40%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 10863 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 Signal transducer and activator of transcription 1-alpha/beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	652	Total	C	N	O	S	0	0	0
			5372	3435	922	991	24			
1	B	653	Total	C	N	O	S	0	0	0
			5387	3442	924	997	24			

- Molecule 2 is a protein called 5-residue peptide.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	P	0	0	0
			51	30	8	12	1			
2	D	5	Total	C	N	O	P	0	0	0
			51	30	8	12	1			

- Molecule 3 is GOLD ION (three-letter code: AU) (formula: Au).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Au	0	0
			1	1		
3	B	1	Total	Au	0	0
			1	1		

3 Residue-property plots ⓘ

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: Signal transducer and activator of transcription 1-alpha/beta
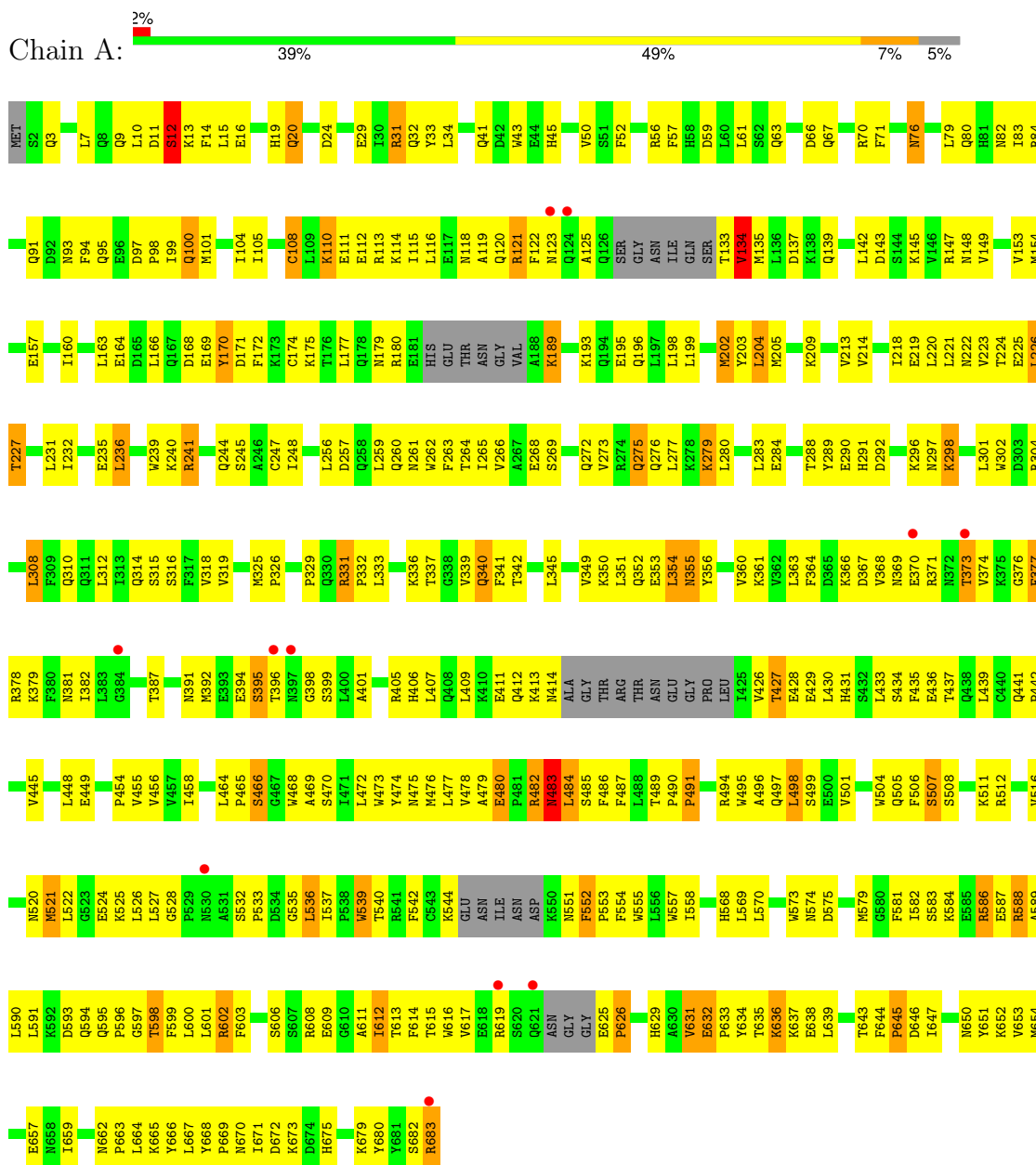
• Molecule 1: Signal transducer and activator of transcription 1-alpha/beta

Chain B:

• Molecule 2: 5-residue peptide

Chain C:

• Molecule 2: 5-residue peptide

Chain D:
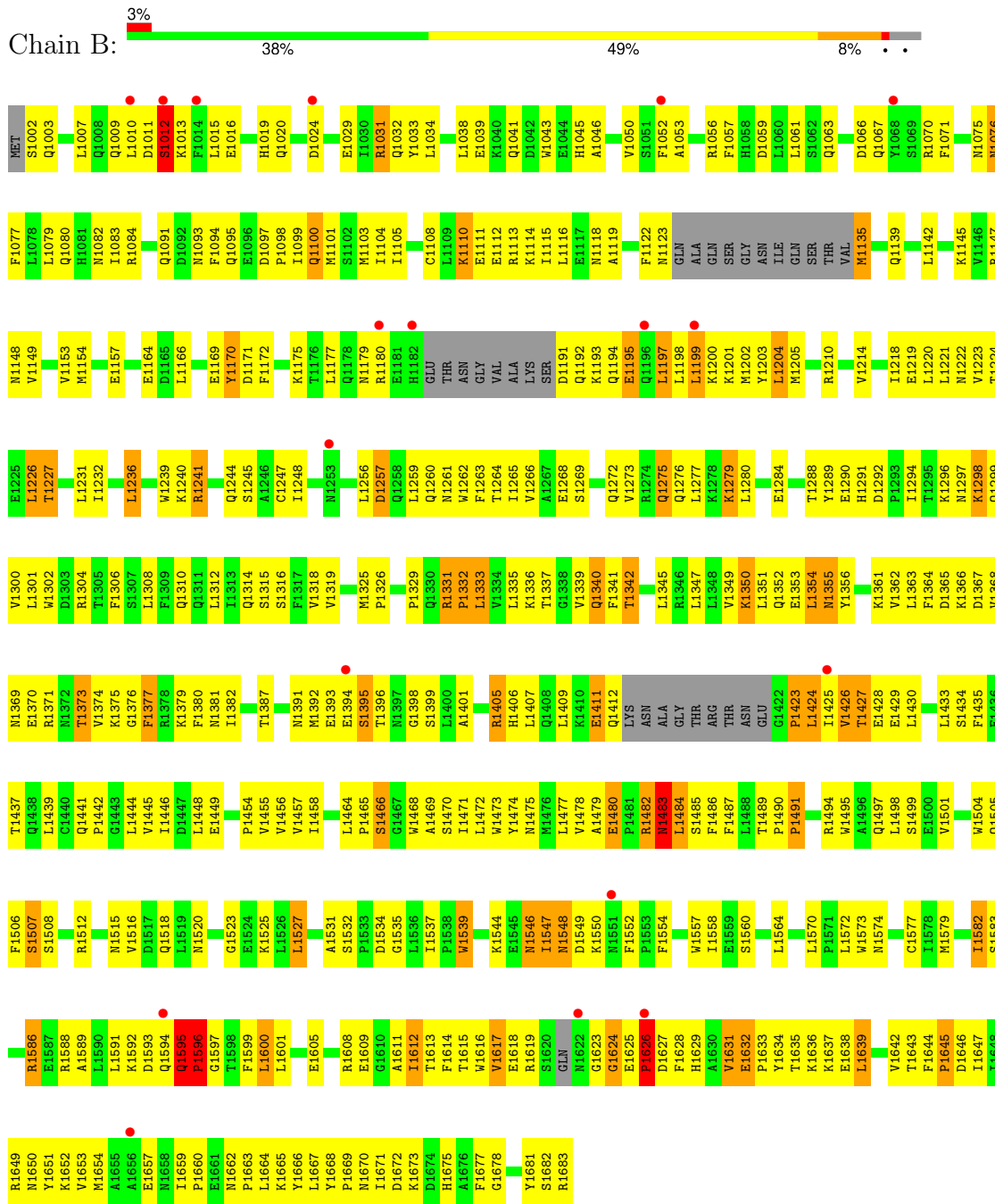
Y1440 D1441 H1444

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	102.55Å 102.55Å 646.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 3.00 29.79 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-3.00) 95.0 (29.79-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.30 (at 2.80Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.240 , 0.281 0.248 , 0.292	Depositor DCC
R_{free} test set	2037 reflections (2.20%)	wwPDB-VP
Wilson B-factor (Å^2)	67.6	Xtriage
Anisotropy	0.483	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 71.6	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	10863	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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: PTR, AU

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	0.42	0/5486	0.67	2/7413 (0.0%)
1	B	0.45	1/5503 (0.0%)	0.69	4/7437 (0.1%)
2	C	0.42	0/36	1.10	1/46 (2.2%)
2	D	0.45	0/36	1.23	1/46 (2.2%)
All	All	0.44	1/11061 (0.0%)	0.68	8/14942 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1577	CYS	CB-SG	5.97	1.92	1.82

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	1441	ASP	N-CA-CB	7.17	123.50	110.60
1	B	1405	ARG	N-CA-C	-7.12	91.77	111.00
1	A	405	ARG	N-CA-C	-6.20	94.25	111.00
2	C	441	ASP	N-CA-CB	6.14	121.66	110.60
1	B	1595	GLN	N-CA-C	-5.85	95.22	111.00

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5372	0	5335	398	0
1	B	5387	0	5349	407	0
2	C	51	0	39	2	0
2	D	51	0	39	0	0
3	A	1	0	0	1	0
3	B	1	0	0	1	0
All	All	10863	0	10762	798	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 37.

The worst 5 of 798 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:B:1194:GLN:O	1:B:1195:GLU:HG3	1.47	1.12
1:B:1612:ILE:HD13	1:B:1612:ILE:H	1.08	1.11
1:A:551:ASN:HB3	1:A:608:ARG:HD3	1.23	1.11
1:B:1441:GLN:HB3	1:B:1442:PRO:HD2	1.34	1.07
1:A:441:GLN:HB3	1:A:442:PRO:HD2	1.34	1.03

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	640/683 (94%)	525 (82%)	92 (14%)	23 (4%)	3	19
1	B	643/683 (94%)	520 (81%)	91 (14%)	32 (5%)	2	12
2	C	3/5 (60%)	3 (100%)	0	0	100	100
2	D	3/5 (60%)	3 (100%)	0	0	100	100
All	All	1289/1376 (94%)	1051 (82%)	183 (14%)	55 (4%)	2	15

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	427	THR
1	B	1426	VAL
1	B	1547	ILE
1	B	1548	ASN
1	B	1595	GLN

5.3.2   Protein sidechains ⓘ

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	599/634 (94%)	549 (92%)	50 (8%)	11	39
1	B	603/634 (95%)	549 (91%)	54 (9%)	9	35
2	C	4/4 (100%)	4 (100%)	0	100	100
2	D	4/4 (100%)	4 (100%)	0	100	100
All	All	1210/1276 (95%)	1106 (91%)	104 (9%)	10	37

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1135	MET
1	B	1331	ARG
1	B	1626	PRO
1	B	1170	TYR
1	B	1236	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	594	GLN
1	B	1520	ASN
1	B	1093	ASN
1	B	1505	GLN
1	B	1670	ASN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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	$\#|Z| > 2$
2	PTR	C	440	2	15,16,17	1.16	0	17,22,24	1.28	1 (5%)
2	PTR	D	1440	2	15,16,17	1.24	1 (6%)	17,22,24	1.31	2 (11%)

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	PTR	C	440	2	-	1/10/11/13	0/1/1/1
2	PTR	D	1440	2	-	1/10/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1440	PTR	CE1-CZ	2.19	1.42	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	1440	PTR	CD2-CE2-CZ	-2.70	116.65	119.73
2	C	440	PTR	CD2-CE2-CZ	-2.66	116.69	119.73
2	D	1440	PTR	O2P-P-OH	2.08	111.46	105.32

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	440	PTR	N-CA-CB-CG
2	D	1440	PTR	N-CA-CB-CG

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	440	PTR	2	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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 ⓘ

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, 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	652/683 (95%)	-0.26	11 (1%)	70	41	41, 70, 112, 144	0
1	B	653/683 (95%)	-0.24	18 (2%)	53	25	29, 68, 131, 150	0
2	C	4/5 (80%)	-0.61	0	100	100	53, 58, 59, 60	0
2	D	4/5 (80%)	-0.28	0	100	100	87, 89, 89, 90	0
All	All	1313/1376 (95%)	-0.25	29 (2%)	62	33	29, 69, 122, 150	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1594	GLN	4.7
1	B	1182	HIS	4.4
1	A	396	THR	3.0
1	A	621	GLN	3.0
1	B	1394	GLU	2.9

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, 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	PTR	C	440	16/17	0.93	0.20	36,47,59,61	0
2	PTR	D	1440	16/17	0.94	0.13	48,52,55,62	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
3	AU	B	2002	1/1	0.95	0.27	6,6,6,6	1
3	AU	A	2001	1/1	0.98	0.25	24,24,24,24	1

6.5 Other polymers ⓘ

There are no such residues in this entry.