wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 04:44 AM GMT

PDB ID	:	2NZW
Title	:	Crystal Structure of alpha1,3-Fucosyltransferase
Authors	:	Sun, H.Y.; Ko, T.P.
Deposited on	:	2006-11-27
Resolution	:	1.90 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:

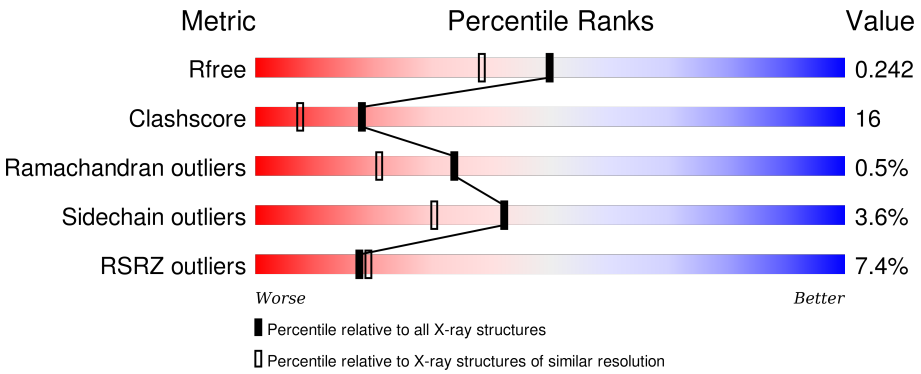
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.90 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	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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	371	8% 69% 23% • 6%
1	B	371	8% 68% 21% • • 9%
1	C	371	5% 76% 16% • 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10265 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 Alpha1,3-fucosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	349	2860	1852	467	534	7	0	0	0
1	B	338	2776	1797	452	520	7	0	0	0
1	C	351	2879	1867	469	536	7	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

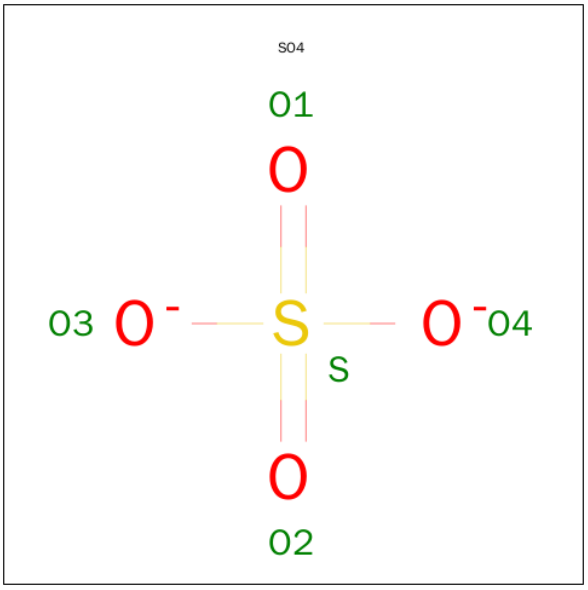
Chain	Residue	Modelled	Actual	Comment	Reference
A	364	LEU	-	EXPRESSION TAG	UNP O30511
A	365	GLU	-	EXPRESSION TAG	UNP O30511
A	366	HIS	-	EXPRESSION TAG	UNP O30511
A	367	HIS	-	EXPRESSION TAG	UNP O30511
A	368	HIS	-	EXPRESSION TAG	UNP O30511
A	369	HIS	-	EXPRESSION TAG	UNP O30511
A	370	HIS	-	EXPRESSION TAG	UNP O30511
A	371	HIS	-	EXPRESSION TAG	UNP O30511
B	364	LEU	-	EXPRESSION TAG	UNP O30511
B	365	GLU	-	EXPRESSION TAG	UNP O30511
B	366	HIS	-	EXPRESSION TAG	UNP O30511
B	367	HIS	-	EXPRESSION TAG	UNP O30511
B	368	HIS	-	EXPRESSION TAG	UNP O30511
B	369	HIS	-	EXPRESSION TAG	UNP O30511
B	370	HIS	-	EXPRESSION TAG	UNP O30511
B	371	HIS	-	EXPRESSION TAG	UNP O30511
C	364	LEU	-	EXPRESSION TAG	UNP O30511
C	365	GLU	-	EXPRESSION TAG	UNP O30511
C	366	HIS	-	EXPRESSION TAG	UNP O30511
C	367	HIS	-	EXPRESSION TAG	UNP O30511
C	368	HIS	-	EXPRESSION TAG	UNP O30511
C	369	HIS	-	EXPRESSION TAG	UNP O30511
C	370	HIS	-	EXPRESSION TAG	UNP O30511

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	371	HIS	-	EXPRESSION TAG	UNP O30511

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	B	1	5	4	1	0	0
			Total	O	S		
2	C	1	5	4	1	0	0
			Total	O	S		
2	C	1	5	4	1	0	0

- Molecule 3 is water.

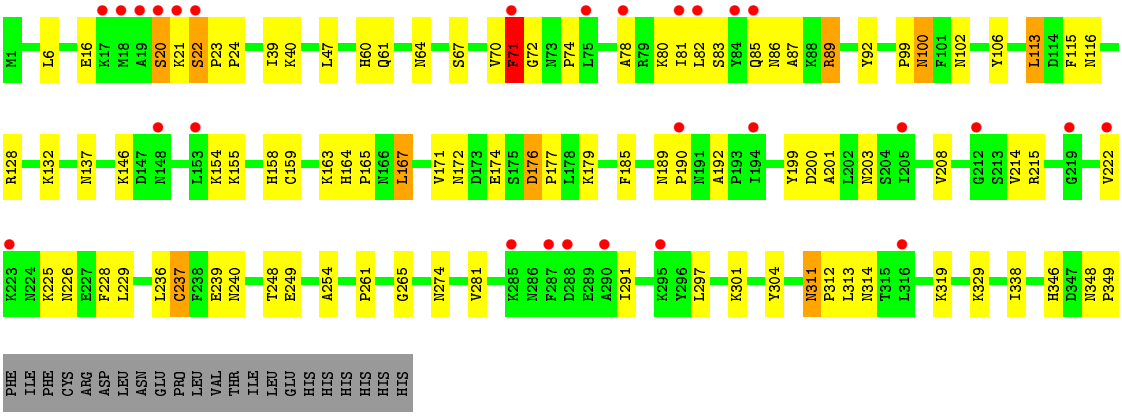
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	559	559	559	0	0
			Total	O		
3	B	522	522	522	0	0
			Total	O		
3	C	654	654	654	0	0

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.

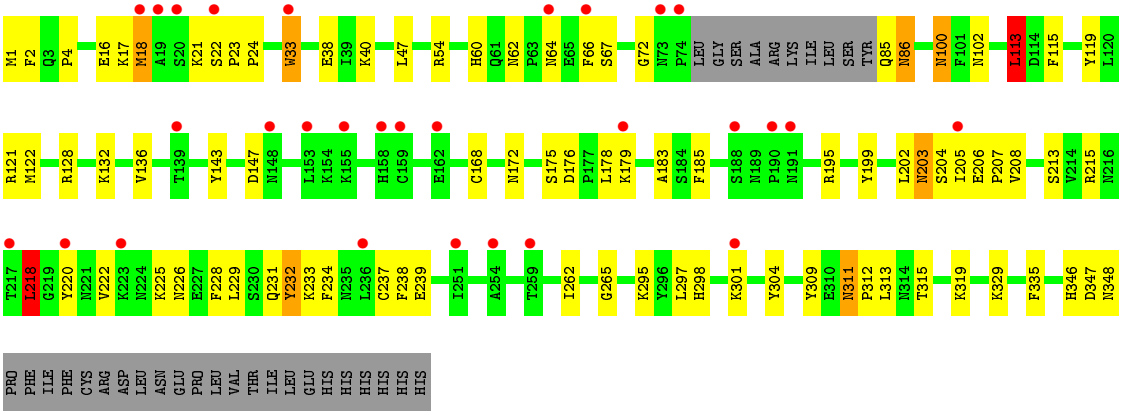
- Molecule 1: Alpha1,3-fucosyltransferase

Chain A:



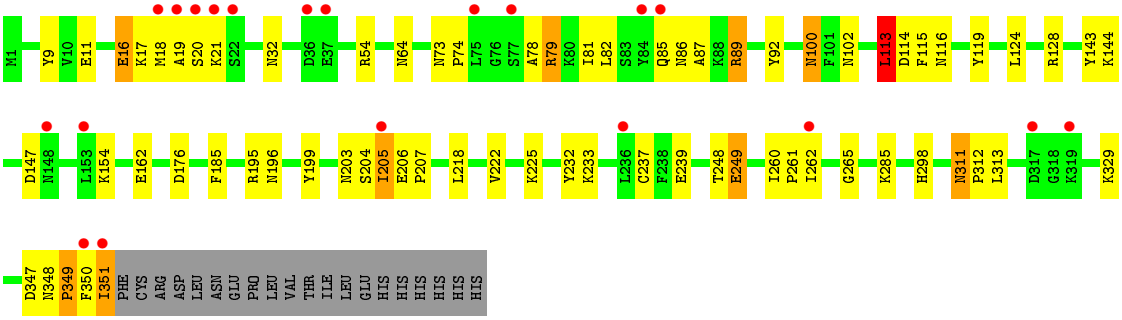
- Molecule 1: Alpha1,3-fucosyltransferase

Chain B:



- Molecule 1: Alpha1,3-fucosyltransferase

Chain C:


M1
Y9
V10
E11
E16
K17
M18
A19
S20
K21
S22
N32
D36
E37
R54
N64
N73
P74
L75
G76
S77
A78
R79
K80
I81
L82
S83
Y84
Q85
N86
A87
K88
R89
Y92
N100
F101
N102
L113
D114
F115
N116
Y119
L124
R128
Y143
K144
D147
N148
L153
K154
E162
D176
F185
R195
N196
Y199
N203
S204
I205
E206
P207
L218
V222
K225
Y232
K233
L236
C237
F238
E239
T248
E249
I260
P261
I262
G265
K285
H298
N311
P312
L313
D317
G318
K319
K329
D347
N348
P349
F350
I351
PHE
CYS
ARG
ASP
LEU
ASN
GLU
PRO
LEU
VAL
THR
ILE
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	104.55Å 136.20Å 96.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.90 29.54 – 1.90	Depositor EDS
% Data completeness (in resolution range)	91.9 (30.00-1.90) 92.1 (29.54-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.88 (at 1.91Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.191 , 0.242 0.192 , 0.242	Depositor DCC
R_{free} test set	5075 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	27.4	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 60.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 100376 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10265	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% 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: SO4

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	0.93	2/2946 (0.1%)	0.87	4/3999 (0.1%)
1	B	0.88	1/2859 (0.0%)	0.85	3/3880 (0.1%)
1	C	1.01	2/2966 (0.1%)	0.93	7/4026 (0.2%)
All	All	0.94	5/8771 (0.1%)	0.89	14/11905 (0.1%)

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
1	B	0	1
1	C	0	2
All	All	0	4

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	237	CYS	CB-SG	-8.98	1.67	1.82
1	C	249	GLU	CG-CD	6.81	1.62	1.51
1	A	71	PHE	CB-CG	-5.49	1.42	1.51
1	C	237	CYS	CB-SG	-5.35	1.73	1.81
1	B	237	CYS	CB-SG	-5.11	1.73	1.81

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	89	ARG	NE-CZ-NH2	-7.80	116.40	120.30
1	A	113	LEU	CA-CB-CG	7.57	132.71	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	113	LEU	CA-CB-CG	7.33	132.15	115.30
1	C	113	LEU	CA-CB-CG	7.15	131.75	115.30
1	B	121	ARG	NE-CZ-NH2	-6.33	117.13	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	92	TYR	Sidechain
1	B	232	TYR	Sidechain
1	C	232	TYR	Sidechain
1	C	9	TYR	Sidechain

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	2860	0	2747	95	0
1	B	2776	0	2653	92	0
1	C	2879	0	2767	79	0
2	B	5	0	0	0	0
2	C	10	0	0	0	0
3	A	559	0	0	27	0
3	B	522	0	0	13	1
3	C	654	0	0	23	0
All	All	10265	0	8167	262	1

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 262 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:122:MET:CE	1:B:122:MET:SD	2.02	1.48
1:C:87:ALA:HB1	3:C:2615:HOH:O	1.49	1.10

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:87:ALA:HB1	3:A:530:HOH:O	1.55	1.07
1:B:132:LYS:NZ	1:B:226:ASN:HD21	1.62	0.97
1:B:60:HIS:HD2	1:B:62:ASN:H	1.10	0.96

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:2249:HOH:O	3:B:2249:HOH:O[2_665]	2.12	0.08

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	347/371 (94%)	326 (94%)	20 (6%)	1 (0%)	46	35
1	B	334/371 (90%)	317 (95%)	14 (4%)	3 (1%)	21	9
1	C	349/371 (94%)	333 (95%)	15 (4%)	1 (0%)	46	35
All	All	1030/1113 (92%)	976 (95%)	49 (5%)	5 (0%)	34	21

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	19	ALA
1	A	301	LYS
1	B	18	MET
1	B	213	SER
1	B	218	LEU

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	312/334 (93%)	300 (96%)	12 (4%)	40	28
1	B	303/334 (91%)	292 (96%)	11 (4%)	42	30
1	C	314/334 (94%)	304 (97%)	10 (3%)	46	35
All	All	929/1002 (93%)	896 (96%)	33 (4%)	42	30

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	100	ASN
1	B	147	ASP
1	C	311	ASN
1	B	113	LEU
1	B	115	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	86	ASN
1	B	226	ASN
1	C	298	HIS
1	B	100	ASN
1	B	172	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)

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)

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	$\#\|Z\| > 2$
2	SO4	B	2001	-	4,4,4	0.24	0	6,6,6	0.11	0
2	SO4	C	2002	-	4,4,4	0.54	0	6,6,6	0.99	1 (16%)
2	SO4	C	2003	-	4,4,4	0.17	0	6,6,6	0.17	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	SO4	B	2001	-	-	0/0/0/0	0/0/0/0
2	SO4	C	2002	-	-	0/0/0/0	0/0/0/0
2	SO4	C	2003	-	-	0/0/0/0	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})
2	C	2002	SO4	O2-S-O1	-2.07	102.94	109.50

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 (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	349/371 (94%)	0.41	28 (8%)	15	17	18, 37, 67, 81	0
1	B	338/371 (91%)	0.42	29 (8%)	13	14	18, 37, 69, 80	0
1	C	351/371 (94%)	0.18	20 (5%)	27	30	16, 28, 54, 78	0
All	All	1038/1113 (93%)	0.33	77 (7%)	17	19	16, 33, 66, 81	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	20	SER	8.0
1	C	20	SER	7.6
1	C	84	TYR	6.3
1	C	351	ILE	5.6
1	C	19	ALA	5.6

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	SO4	C	2002	5/5	0.97	0.09	-0.63	45,48,49,50	0
2	SO4	B	2001	5/5	0.95	0.12	-0.64	76,76,77,77	0
2	SO4	C	2003	5/5	0.98	0.10	-	61,62,63,63	0

6.5 Other polymers (i)

There are no such residues in this entry.