Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 09:26 AM GMT

PDB ID	:	3IG2
Title	:	The Crystal Structure of a Putative Phenylalanyl-tRNA synthetase (PheRS) beta chain domain from Bacteroides fragilis to 2.1A
Authors	:	Stein, A.J.; Sather, A.; Hendricks, R.; Keigher, L.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2009-07-27
Resolution	:	2.09 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:

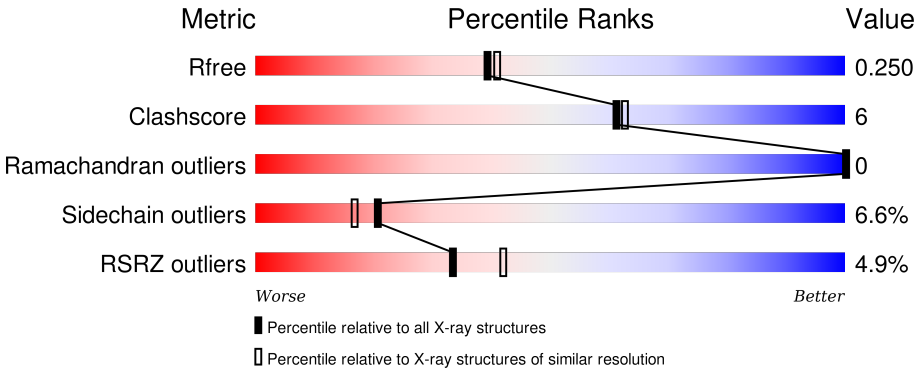
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.09 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	213	% 73% 15% • 11%
1	B	213	3% 81% 10% • 8%
1	C	213	4% 77% 11% 12%
1	D	213	9% 77% 13% • 9%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6285 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 Phenylalanyl-tRNA synthetase beta chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	190	Total	C	N	O	S	Se	0	2	0
			1502	965	245	283	3	6			
1	B	197	Total	C	N	O	S	Se	0	2	0
			1547	989	254	296	3	5			
1	C	187	Total	C	N	O	S	Se	0	3	0
			1475	944	240	284	3	4			
1	D	194	Total	C	N	O	S	Se	0	0	0
			1502	958	246	291	3	4			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	501	SER	-	expression tag	UNP Q64T65
A	502	ASN	-	expression tag	UNP Q64T65
A	503	ALA	-	expression tag	UNP Q64T65
B	501	SER	-	expression tag	UNP Q64T65
B	502	ASN	-	expression tag	UNP Q64T65
B	503	ALA	-	expression tag	UNP Q64T65
C	501	SER	-	expression tag	UNP Q64T65
C	502	ASN	-	expression tag	UNP Q64T65
C	503	ALA	-	expression tag	UNP Q64T65
D	501	SER	-	expression tag	UNP Q64T65
D	502	ASN	-	expression tag	UNP Q64T65
D	503	ALA	-	expression tag	UNP Q64T65

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

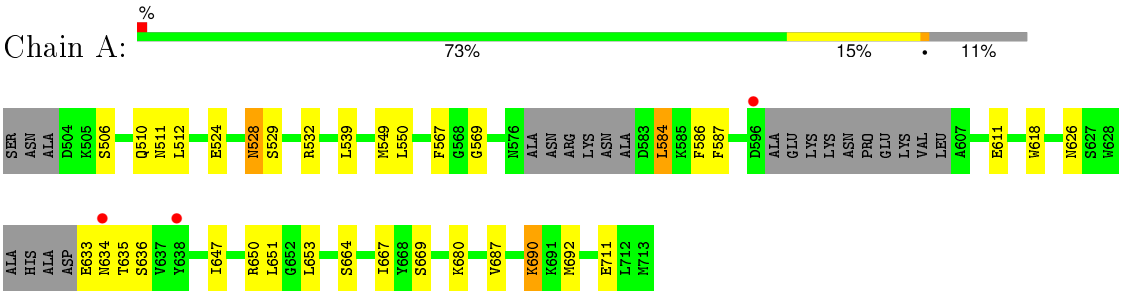
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	73	Total O 73 73	0	0
3	B	89	Total O 89 89	0	0
3	C	64	Total O 64 64	0	0
3	D	32	Total O 32 32	0	0

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.

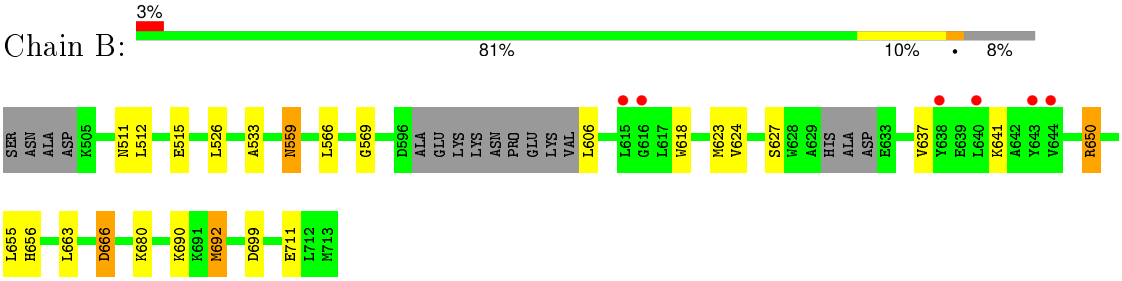
- Molecule 1: Phenylalanyl-tRNA synthetase beta chain

Chain A:



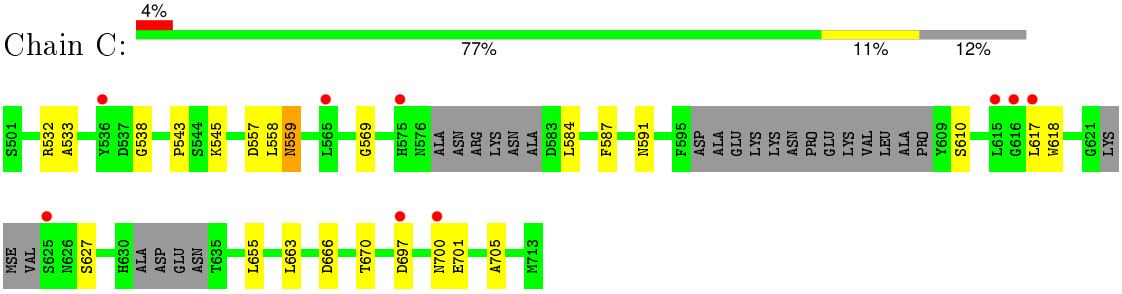
- Molecule 1: Phenylalanyl-tRNA synthetase beta chain

Chain B:



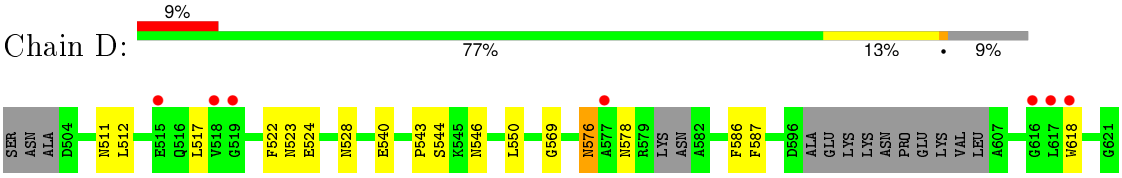
- Molecule 1: Phenylalanyl-tRNA synthetase beta chain

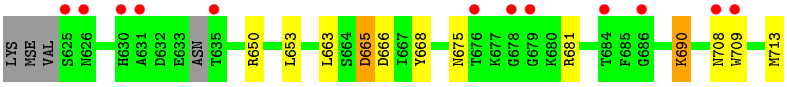
Chain C:



- Molecule 1: Phenylalanyl-tRNA synthetase beta chain

Chain D:




LYS
MSE
VAL
S625
N626
H630
A631
D632
E633
ASN
T635
R650
L653
L663
S664
D665
D666
I667
Y668
N675
T676
K677
G678
G679
K680
R681
T684
F685
G686
K690
N708
W709
M713

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	48.50Å 60.34Å 78.81Å 99.04° 100.37° 105.08°	Depositor
Resolution (Å)	28.45 – 2.09 28.45 – 2.09	Depositor EDS
% Data completeness (in resolution range)	97.5 (28.45-2.09) 90.6 (28.45-2.09)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.51 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.195 , 0.236 0.216 , 0.250	Depositor DCC
R_{free} test set	2416 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	33.9	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 40.7	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 47599 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6285	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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: MG

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	0.57	0/1533	0.66	0/2067
1	B	0.67	0/1576	0.78	2/2125 (0.1%)
1	C	0.60	0/1507	0.68	0/2034
1	D	0.48	0/1526	0.62	0/2064
All	All	0.58	0/6142	0.69	2/8290 (0.0%)

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	B	692	MSE	CG-SE-CE	-5.70	86.36	98.90
1	B	666	ASP	CB-CG-OD1	5.10	122.89	118.30

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.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1502	0	1449	21	0
1	B	1547	0	1503	15	0
1	C	1475	0	1415	10	0
1	D	1502	0	1411	23	0
2	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	73	0	0	2	0
3	B	89	0	0	1	0
3	C	64	0	0	0	0
3	D	32	0	0	0	0
All	All	6285	0	5778	65	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 6.

All (65) 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:680:LYS:NZ	1:B:711:GLU:OE1	1.82	1.13
1:D:690:LYS:HE3	1:D:690:LYS:H	1.35	0.91
1:B:680:LYS:NZ	1:B:711:GLU:CD	2.31	0.83
1:B:533:ALA:H	1:B:559:ASN:HD21	1.27	0.80
1:C:533:ALA:H	1:C:559:ASN:HD21	1.27	0.79
1:D:576:ASN:HD22	1:D:578:ASN:H	1.31	0.78
1:D:543:PRO:HG2	1:D:546:ASN:OD1	1.86	0.75
1:D:709:TRP:NE1	1:D:713:MSE:CE	2.52	0.72
1:D:690:LYS:CE	1:D:690:LYS:H	2.02	0.70
1:D:576:ASN:ND2	1:D:578:ASN:H	1.89	0.70
1:A:511:ASN:HD21	1:D:511:ASN:ND2	1.91	0.68
1:D:665:ASP:HB3	1:D:668:TYR:H	1.59	0.68
1:A:584:LEU:HG	1:A:586:PHE:CZ	2.30	0.66
1:D:690:LYS:HE3	1:D:690:LYS:N	2.09	0.66
1:C:543:PRO:HB3	1:C:545:LYS:HE3	1.79	0.65
1:A:690:LYS:H	1:A:690:LYS:CD	2.10	0.64
1:B:680:LYS:HZ2	1:B:711:GLU:CD	2.01	0.63
1:B:680:LYS:HZ1	1:B:711:GLU:CD	1.89	0.62
1:C:532:ARG:CZ	1:C:557:ASP:OD1	2.48	0.61
1:C:587:PHE:HB2	1:C:617[A]:LEU:HD23	1.86	0.58
1:A:690:LYS:HD2	1:A:690:LYS:N	2.19	0.57
1:A:539:LEU:HD11	1:A:567:PHE:HZ	1.69	0.57
1:A:528:ASN:HD22	1:A:529:SER:N	2.04	0.54
1:D:576:ASN:HD22	1:D:578:ASN:N	2.04	0.54
1:B:559:ASN:H	1:B:559:ASN:HD22	1.56	0.54
3:A:42:HOH:O	1:D:528:ASN:HB3	2.08	0.53
1:A:647:ILE:O	1:A:651:LEU:HG	2.10	0.52
1:C:617[B]:LEU:HB2	1:C:705:ALA:HB3	1.92	0.51
1:C:587:PHE:CB	1:C:617[A]:LEU:HD23	2.41	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:569:GLY:HA3	1:B:618:TRP:CE2	2.47	0.50
1:C:569:GLY:HA3	1:C:618:TRP:CE2	2.47	0.50
1:C:532:ARG:NH1	1:C:557:ASP:OD1	2.46	0.49
1:A:506:SER:O	1:A:510:GLN:HG3	2.13	0.49
1:A:569:GLY:HA3	1:A:618:TRP:CE2	2.49	0.48
1:D:675:ASN:HD21	1:D:681:ARG:HE	1.62	0.48
1:B:511:ASN:O	1:B:515:GLU:HG3	2.13	0.48
1:D:709:TRP:CE2	1:D:713:MSE:CE	2.97	0.47
1:D:524:GLU:HB2	1:D:587:PHE:CZ	2.50	0.46
1:D:709:TRP:CE2	1:D:713:MSE:HE1	2.51	0.46
1:B:559:ASN:HD22	1:B:559:ASN:N	2.13	0.46
1:D:523:ASN:O	1:D:586:PHE:HA	2.15	0.46
1:A:690:LYS:CD	1:A:690:LYS:N	2.74	0.45
1:A:511:ASN:ND2	1:D:511:ASN:ND2	2.63	0.45
1:D:709:TRP:NE1	1:D:713:MSE:HE1	2.30	0.44
1:D:569:GLY:HA3	1:D:618:TRP:CE2	2.52	0.44
1:B:656:HIS:CE1	1:C:538:GLY:O	2.70	0.44
1:A:664:SER:HB2	1:A:669:SER:O	2.18	0.44
1:A:667:ILE:HD13	1:A:692:MSE:HE3	1.99	0.44
1:D:517:LEU:O	1:D:522:PHE:HB2	2.18	0.43
1:A:528:ASN:HD22	1:A:529:SER:H	1.66	0.43
1:B:650:ARG:HA	1:B:650:ARG:HD3	1.60	0.43
1:A:611:GLU:OE2	1:D:528:ASN:HB3	2.19	0.43
1:B:637:VAL:O	1:B:641:LYS:HB2	2.19	0.43
1:B:566:LEU:HD21	1:B:692:MSE:HE1	2.00	0.43
1:C:700:ASN:HB3	1:C:701:GLU:H	1.66	0.42
1:A:690:LYS:CE	1:A:690:LYS:H	2.32	0.41
1:A:633:GLU:HB3	1:A:687:VAL:HG21	2.01	0.41
1:A:532:ARG:NH1	3:A:124:HOH:O	2.53	0.41
1:D:650:ARG:HA	1:D:650:ARG:HD2	1.92	0.41
1:A:549:MSE:HE3	1:A:549:MSE:HB3	1.94	0.41
1:B:623:MSE:HG3	3:B:38:HOH:O	2.20	0.41
1:B:624:VAL:HG13	1:B:690:LYS:HE2	2.03	0.40
1:A:524:GLU:HB2	1:A:587:PHE:CZ	2.57	0.40
1:D:653:LEU:HA	1:D:653:LEU:HD12	1.89	0.40
1:A:680:LYS:HD2	1:A:711:GLU:OE1	2.22	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	184/213 (86%)	182 (99%)	2 (1%)	0	100	100
1	B	193/213 (91%)	191 (99%)	2 (1%)	0	100	100
1	C	180/213 (84%)	178 (99%)	2 (1%)	0	100	100
1	D	184/213 (86%)	183 (100%)	1 (0%)	0	100	100
All	All	741/852 (87%)	734 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	162/178 (91%)	151 (93%)	11 (7%)	20	16
1	B	166/178 (93%)	155 (93%)	11 (7%)	21	17
1	C	160/178 (90%)	148 (92%)	12 (8%)	17	13
1	D	157/178 (88%)	147 (94%)	10 (6%)	22	18
All	All	645/712 (91%)	601 (93%)	44 (7%)	21	16

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

Mol	Chain	Res	Type
1	A	512	LEU
1	A	528	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	550	LEU
1	A	584	LEU
1	A	626	ASN
1	A	634	ASN
1	A	635	THR
1	A	636	SER
1	A	650	ARG
1	A	653	LEU
1	A	690	LYS
1	B	512	LEU
1	B	526	LEU
1	B	559	ASN
1	B	606	LEU
1	B	627	SER
1	B	650	ARG
1	B	655	LEU
1	B	663	LEU
1	B	666	ASP
1	B	699[A]	ASP
1	B	699[B]	ASP
1	C	558	LEU
1	C	559	ASN
1	C	584	LEU
1	C	591	ASN
1	C	610	SER
1	C	627	SER
1	C	655	LEU
1	C	663	LEU
1	C	666	ASP
1	C	670[A]	THR
1	C	670[B]	THR
1	C	697	ASP
1	D	512	LEU
1	D	540	GLU
1	D	544	SER
1	D	550	LEU
1	D	576	ASN
1	D	663	LEU
1	D	665	ASP
1	D	666	ASP
1	D	690	LYS
1	D	708	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	528	ASN
1	A	552	ASN
1	A	626	ASN
1	A	634	ASN
1	A	700	ASN
1	B	559	ASN
1	B	581	ASN
1	C	559	ASN
1	C	591	ASN
1	D	511	ASN
1	D	523	ASN
1	D	576	ASN
1	D	591	ASN
1	D	675	ASN

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

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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, 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	185/213 (86%)	0.13	3 (1%) 74 79	7, 15, 19, 20	0
1	B	192/213 (90%)	0.05	6 (3%) 52 61	2, 14, 17, 20	0
1	C	183/213 (85%)	0.22	9 (4%) 33 42	6, 13, 17, 18	0
1	D	190/213 (89%)	0.48	19 (10%) 9 13	5, 16, 20, 23	0
All	All	750/852 (88%)	0.22	37 (4%) 33 42	2, 14, 19, 23	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	640	LEU	5.5
1	D	709	TRP	5.2
1	D	617	LEU	4.1
1	C	625	SER	4.0
1	B	643	TYR	3.8
1	B	638	TYR	3.4
1	C	700	ASN	3.2
1	C	565	LEU	3.1
1	D	577	ALA	3.1
1	D	678	GLY	3.0
1	B	644	VAL	2.8
1	D	515	GLU	2.8
1	A	596	ASP	2.7
1	D	684	THR	2.6
1	C	575	HIS	2.6
1	D	625	SER	2.6
1	D	679	GLY	2.6
1	D	618	TRP	2.5
1	A	638	TYR	2.5
1	C	616	GLY	2.5
1	D	519	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	631	ALA	2.4
1	A	634	ASN	2.4
1	C	697	ASP	2.4
1	D	518	VAL	2.3
1	C	617[A]	LEU	2.3
1	D	630	HIS	2.3
1	D	635	THR	2.2
1	D	676	THR	2.2
1	D	708	ASN	2.2
1	D	616	GLY	2.1
1	D	626	ASN	2.1
1	B	616	GLY	2.1
1	C	536	TYR	2.1
1	C	615	LEU	2.1
1	D	686	GLY	2.1
1	B	615	LEU	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)

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, 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
2	MG	C	1	1/1	0.99	0.09	0.29	28,28,28,28	0

6.5 Other polymers (i)

There are no such residues in this entry.