Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 09:01 am BST

PDB ID	:	6HYF
Title	:	Crystal structure of the third FNIII domain from rat beta4 integrin, a binding site for periaxin
Authors	:	Raasakka, A.; Kursula, P.
Deposited on	:	2018-10-20
Resolution	:	1.60 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

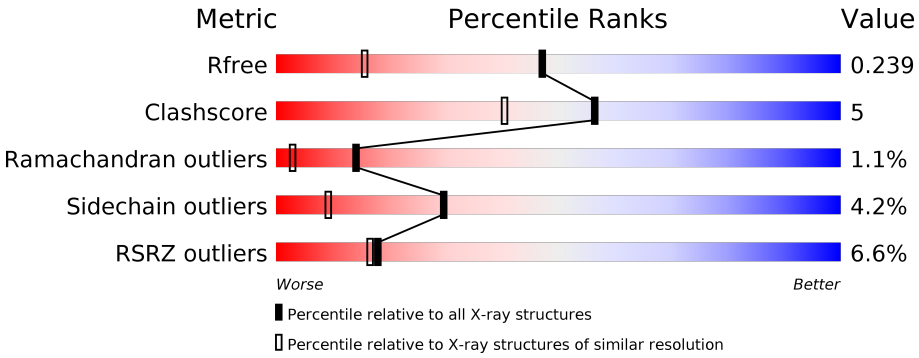
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 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	113	82% 12% 6%
1	B	113	10% 72% 14% 12%
1	C	113	3% 71% 8% 19%
1	D	113	9% 64% 8% 27%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6201 atoms, of which 2981 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 Integrin beta-4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	106	Total	C	H	N	O	S	0	7	0
			1657	527	812	156	160	2			
1	B	100	Total	C	H	N	O	S	0	7	0
			1596	508	788	142	156	2			
1	C	91	Total	C	H	N	O	S	0	10	0
			1467	462	727	133	143	2			
1	D	82	Total	C	H	N	O	S	0	6	0
			1319	419	654	118	127	1			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1436	MET	-	initiating methionine	UNP Q64632
A	1437	HIS	-	expression tag	UNP Q64632
A	1438	HIS	-	expression tag	UNP Q64632
A	1439	HIS	-	expression tag	UNP Q64632
A	1440	HIS	-	expression tag	UNP Q64632
A	1441	HIS	-	expression tag	UNP Q64632
A	1442	HIS	-	expression tag	UNP Q64632
A	1443	SER	-	expression tag	UNP Q64632
A	1444	SER	-	expression tag	UNP Q64632
A	1445	GLY	-	expression tag	UNP Q64632
A	1446	VAL	-	expression tag	UNP Q64632
A	1447	ASP	-	expression tag	UNP Q64632
A	1448	LEU	-	expression tag	UNP Q64632
A	1449	GLY	-	expression tag	UNP Q64632
A	1450	THR	-	expression tag	UNP Q64632
A	1451	GLU	-	expression tag	UNP Q64632
A	1452	ASN	-	expression tag	UNP Q64632
A	1453	LEU	-	expression tag	UNP Q64632
A	1454	TYR	-	expression tag	UNP Q64632
A	1455	PHE	-	expression tag	UNP Q64632
A	1456	GLN	-	expression tag	UNP Q64632

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1457	SER	-	expression tag	UNP Q64632
B	1436	MET	-	initiating methionine	UNP Q64632
B	1437	HIS	-	expression tag	UNP Q64632
B	1438	HIS	-	expression tag	UNP Q64632
B	1439	HIS	-	expression tag	UNP Q64632
B	1440	HIS	-	expression tag	UNP Q64632
B	1441	HIS	-	expression tag	UNP Q64632
B	1442	HIS	-	expression tag	UNP Q64632
B	1443	SER	-	expression tag	UNP Q64632
B	1444	SER	-	expression tag	UNP Q64632
B	1445	GLY	-	expression tag	UNP Q64632
B	1446	VAL	-	expression tag	UNP Q64632
B	1447	ASP	-	expression tag	UNP Q64632
B	1448	LEU	-	expression tag	UNP Q64632
B	1449	GLY	-	expression tag	UNP Q64632
B	1450	THR	-	expression tag	UNP Q64632
B	1451	GLU	-	expression tag	UNP Q64632
B	1452	ASN	-	expression tag	UNP Q64632
B	1453	LEU	-	expression tag	UNP Q64632
B	1454	TYR	-	expression tag	UNP Q64632
B	1455	PHE	-	expression tag	UNP Q64632
B	1456	GLN	-	expression tag	UNP Q64632
B	1457	SER	-	expression tag	UNP Q64632
C	1436	MET	-	initiating methionine	UNP Q64632
C	1437	HIS	-	expression tag	UNP Q64632
C	1438	HIS	-	expression tag	UNP Q64632
C	1439	HIS	-	expression tag	UNP Q64632
C	1440	HIS	-	expression tag	UNP Q64632
C	1441	HIS	-	expression tag	UNP Q64632
C	1442	HIS	-	expression tag	UNP Q64632
C	1443	SER	-	expression tag	UNP Q64632
C	1444	SER	-	expression tag	UNP Q64632
C	1445	GLY	-	expression tag	UNP Q64632
C	1446	VAL	-	expression tag	UNP Q64632
C	1447	ASP	-	expression tag	UNP Q64632
C	1448	LEU	-	expression tag	UNP Q64632
C	1449	GLY	-	expression tag	UNP Q64632
C	1450	THR	-	expression tag	UNP Q64632
C	1451	GLU	-	expression tag	UNP Q64632
C	1452	ASN	-	expression tag	UNP Q64632
C	1453	LEU	-	expression tag	UNP Q64632
C	1454	TYR	-	expression tag	UNP Q64632

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1455	PHE	-	expression tag	UNP Q64632
C	1456	GLN	-	expression tag	UNP Q64632
C	1457	SER	-	expression tag	UNP Q64632
D	1436	MET	-	initiating methionine	UNP Q64632
D	1437	HIS	-	expression tag	UNP Q64632
D	1438	HIS	-	expression tag	UNP Q64632
D	1439	HIS	-	expression tag	UNP Q64632
D	1440	HIS	-	expression tag	UNP Q64632
D	1441	HIS	-	expression tag	UNP Q64632
D	1442	HIS	-	expression tag	UNP Q64632
D	1443	SER	-	expression tag	UNP Q64632
D	1444	SER	-	expression tag	UNP Q64632
D	1445	GLY	-	expression tag	UNP Q64632
D	1446	VAL	-	expression tag	UNP Q64632
D	1447	ASP	-	expression tag	UNP Q64632
D	1448	LEU	-	expression tag	UNP Q64632
D	1449	GLY	-	expression tag	UNP Q64632
D	1450	THR	-	expression tag	UNP Q64632
D	1451	GLU	-	expression tag	UNP Q64632
D	1452	ASN	-	expression tag	UNP Q64632
D	1453	LEU	-	expression tag	UNP Q64632
D	1454	TYR	-	expression tag	UNP Q64632
D	1455	PHE	-	expression tag	UNP Q64632
D	1456	GLN	-	expression tag	UNP Q64632
D	1457	SER	-	expression tag	UNP Q64632

- Molecule 2 is water.

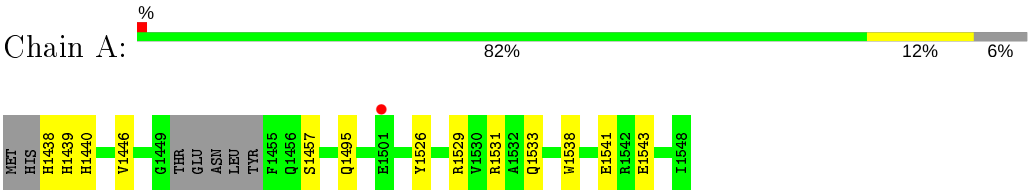
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	64	Total O 64 64	0	0
2	B	37	Total O 37 37	0	0
2	C	39	Total O 39 39	0	0
2	D	22	Total O 22 22	0	0

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 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.

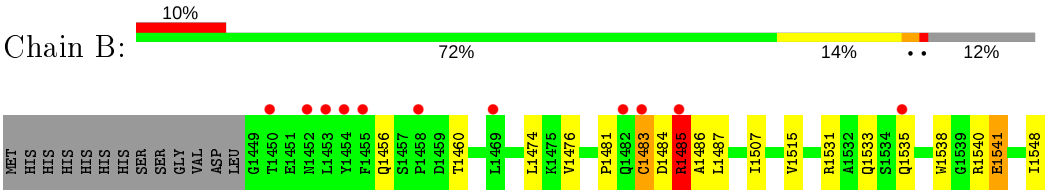
- Molecule 1: Integrin beta-4

Chain A:



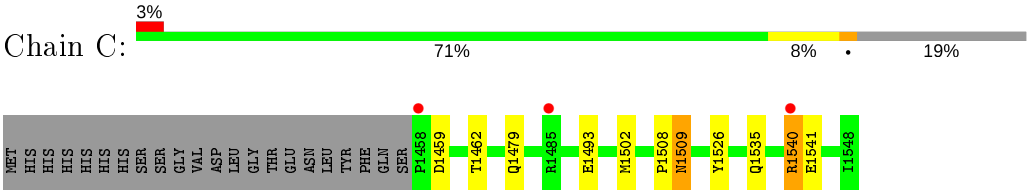
- Molecule 1: Integrin beta-4

Chain B:



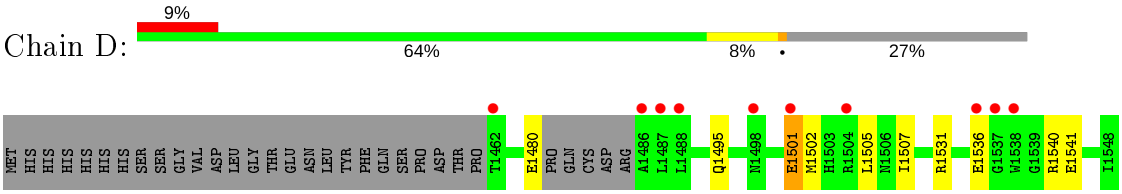
- Molecule 1: Integrin beta-4

Chain C:



- Molecule 1: Integrin beta-4

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.22Å 52.84Å 144.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.79 – 1.60 41.79 – 1.60	Depositor EDS
% Data completeness (in resolution range)	92.2 (41.79-1.60) 85.0 (41.79-1.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.30 (at 1.60Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.190 , 0.239 0.190 , 0.239	Depositor DCC
R_{free} test set	1991 reflections (3.88%)	wwPDB-VP
Wilson B-factor (Å^2)	25.7	Xtriage
Anisotropy	0.232	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 51.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.058 for k,h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6201	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.54% 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

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	0.72	1/909 (0.1%)	0.67	0/1233
1	B	0.71	0/862	0.72	1/1172 (0.1%)
1	C	0.61	1/816 (0.1%)	0.68	0/1110
1	D	0.55	0/709	0.62	0/958
All	All	0.66	2/3296 (0.1%)	0.68	1/4473 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1526	TYR	CD2-CE2	-5.25	1.31	1.39
1	C	1526	TYR	CE1-CZ	-5.06	1.31	1.38

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1485	ARG	CB-CA-C	5.07	120.54	110.40

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	845	812	758	5	0
1	B	808	788	744	17	0
1	C	740	727	660	6	3

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	665	654	614	3	3
2	A	64	0	0	1	0
2	B	37	0	0	0	0
2	C	39	0	0	1	0
2	D	22	0	0	0	0
All	All	3220	2981	2776	30	3

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 5.

All (30) 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:A:1543:GLU:OE1	2:A:1601:HOH:O	2.01	0.79
1:B:1485:ARG:HB3	1:B:1486:ALA:HA	1.68	0.74
1:B:1483:CYS:SG	1:B:1484:ASP:N	2.62	0.73
1:C:1479:GLN:NE2	2:C:1601:HOH:O	2.22	0.71
1:A:1495:GLN:OE1	1:A:1529:ARG:NH2	2.32	0.60
1:B:1474:LEU:C	1:B:1474:LEU:HD12	2.24	0.56
1:B:1485:ARG:CB	1:B:1486:ALA:HA	2.34	0.56
1:A:1440:HIS:HB2	1:B:1535:GLN:HG3	1.87	0.55
1:B:1507:ILE:HD11	1:B:1515[B]:VAL:CG1	2.37	0.53
1:D:1495:GLN:HA	1:D:1501[B]:GLU:O	2.08	0.53
1:B:1474:LEU:HD23	1:B:1548:ILE:CD1	2.42	0.50
1:B:1485:ARG:HD2	1:B:1485:ARG:H	1.76	0.50
1:B:1460[A]:THR:HG22	1:B:1540:ARG:CB	2.42	0.49
1:D:1495:GLN:HA	1:D:1501[A]:GLU:O	2.12	0.49
1:C:1493:GLU:HG2	1:C:1502:MET:SD	2.52	0.49
1:B:1531:ARG:CG	1:B:1541:GLU:HG3	2.44	0.48
1:A:1531:ARG:HD3	1:A:1541:GLU:HG3	1.96	0.47
1:C:1508:PRO:O	1:C:1509:ASN:HB2	2.15	0.45
1:D:1531:ARG:HG3	1:D:1541:GLU:OE1	2.17	0.45
1:A:1533:GLN:HB2	1:A:1538:TRP:CZ3	2.52	0.44
1:B:1485:ARG:N	1:B:1485:ARG:HD2	2.33	0.44
1:B:1507:ILE:HD11	1:B:1515[B]:VAL:HG11	1.98	0.43
1:B:1460[A]:THR:HG22	1:B:1540:ARG:HB2	1.99	0.43
1:C:1462:THR:HG21	1:C:1479:GLN:HE21	1.84	0.42
1:B:1474:LEU:CD1	1:B:1476:VAL:CG2	2.98	0.42
1:B:1533:GLN:HB2	1:B:1538:TRP:CZ3	2.55	0.42
1:C:1462:THR:OG1	1:C:1479:GLN:HB2	2.21	0.41
1:B:1474:LEU:HD11	1:B:1476:VAL:CG2	2.51	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1481:PRO:HD2	1:B:1487:LEU:HD11	2.03	0.41

All (3) 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)
1:C:1540:ARG:HE	1:D:1541:GLU:OE2[3_545]	1.51	0.09
1:C:1540:ARG:NE	1:D:1541:GLU:OE2[3_545]	2.15	0.05
1:C:1541:GLU:OE2	1:D:1540:ARG:NH1[3_545]	2.17	0.03

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	109/113 (96%)	108 (99%)	1 (1%)	0	100 100
1	B	105/113 (93%)	99 (94%)	5 (5%)	1 (1%)	15 3
1	C	99/113 (88%)	93 (94%)	4 (4%)	2 (2%)	7 1
1	D	84/113 (74%)	81 (96%)	1 (1%)	2 (2%)	6 1
All	All	397/452 (88%)	381 (96%)	11 (3%)	5 (1%)	14 2

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1509	ASN
1	B	1483	CYS
1	D	1501[A]	GLU
1	D	1501[B]	GLU
1	C	1459	ASP

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	98/100 (98%)	94 (96%)	4 (4%)	30 9
1	B	93/100 (93%)	90 (97%)	3 (3%)	39 15
1	C	88/100 (88%)	86 (98%)	2 (2%)	50 25
1	D	76/100 (76%)	71 (93%)	5 (7%)	16 3
All	All	355/400 (89%)	341 (96%)	14 (4%)	30 10

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1438	HIS
1	A	1439	HIS
1	A	1446	VAL
1	A	1457	SER
1	B	1456	GLN
1	B	1485	ARG
1	B	1541	GLU
1	C	1535	GLN
1	C	1540	ARG
1	D	1480	GLU
1	D	1502	MET
1	D	1505	LEU
1	D	1507	ILE
1	D	1536	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1535	GLN

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

There are no ligands in this entry.

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	106/113 (93%)	0.18	1 (0%) 84 84	26, 38, 62, 85	0
1	B	100/113 (88%)	0.39	11 (11%) 5 4	25, 44, 100, 126	0
1	C	91/113 (80%)	0.16	3 (3%) 46 43	28, 45, 79, 116	0
1	D	82/113 (72%)	0.54	10 (12%) 4 3	30, 52, 99, 127	0
All	All	379/452 (83%)	0.31	25 (6%) 18 17	25, 44, 91, 127	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1483	CYS	6.7
1	D	1538	TRP	5.3
1	B	1455	PHE	4.7
1	B	1482	GLN	4.4
1	D	1536	GLU	4.4
1	B	1485	ARG	4.2
1	C	1458	PRO	4.0
1	D	1462	THR	3.9
1	D	1501[A]	GLU	3.7
1	D	1486	ALA	3.5
1	D	1498	ASN	3.2
1	B	1450	THR	3.2
1	B	1458	PRO	3.1
1	B	1469	LEU	3.0
1	B	1452	ASN	2.9
1	C	1540	ARG	2.9
1	B	1454	TYR	2.7
1	D	1488	LEU	2.7
1	C	1485	ARG	2.4
1	A	1501	GLU	2.3
1	D	1487	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	1537	GLY	2.3
1	D	1504	ARG	2.3
1	B	1453	LEU	2.2
1	B	1535	GLN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.