Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 12:36 pm BST

PDB ID	:	6GC1
Title	:	Crystal structure of Trx-like and NHL repeat containing domains of human NHLRC2
Authors	:	Biterova, E.; Uusimaa, J.; Hinttala, R.; Ruddock, L.W.
Deposited on	:	2018-04-16
Resolution	:	2.70 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:

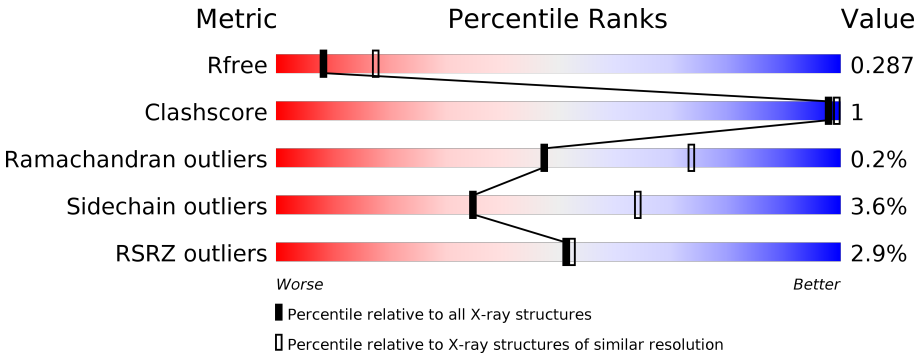
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance

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

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	579	91% · 5%
1	B	579	89% 5% 6%
1	C	579	89% 6% 5%
1	D	579	7% 90% · 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 17055 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 NHL repeat-containing protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	549	Total	C	N	O	S	0	0	0
			4259	2710	717	816	16			
1	B	546	Total	C	N	O	S	0	0	0
			4238	2695	713	814	16			
1	C	551	Total	C	N	O	S	0	0	0
			4269	2714	719	820	16			
1	D	548	Total	C	N	O	S	0	0	0
			4252	2705	716	815	16			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP Q8NBF2
A	-5	HIS	-	expression tag	UNP Q8NBF2
A	-4	HIS	-	expression tag	UNP Q8NBF2
A	-3	HIS	-	expression tag	UNP Q8NBF2
A	-2	HIS	-	expression tag	UNP Q8NBF2
A	-1	HIS	-	expression tag	UNP Q8NBF2
A	0	HIS	-	expression tag	UNP Q8NBF2
B	-6	MET	-	initiating methionine	UNP Q8NBF2
B	-5	HIS	-	expression tag	UNP Q8NBF2
B	-4	HIS	-	expression tag	UNP Q8NBF2
B	-3	HIS	-	expression tag	UNP Q8NBF2
B	-2	HIS	-	expression tag	UNP Q8NBF2
B	-1	HIS	-	expression tag	UNP Q8NBF2
B	0	HIS	-	expression tag	UNP Q8NBF2
C	-6	MET	-	initiating methionine	UNP Q8NBF2
C	-5	HIS	-	expression tag	UNP Q8NBF2
C	-4	HIS	-	expression tag	UNP Q8NBF2
C	-3	HIS	-	expression tag	UNP Q8NBF2
C	-2	HIS	-	expression tag	UNP Q8NBF2
C	-1	HIS	-	expression tag	UNP Q8NBF2
C	0	HIS	-	expression tag	UNP Q8NBF2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-6	MET	-	initiating methionine	UNP Q8NBF2
D	-5	HIS	-	expression tag	UNP Q8NBF2
D	-4	HIS	-	expression tag	UNP Q8NBF2
D	-3	HIS	-	expression tag	UNP Q8NBF2
D	-2	HIS	-	expression tag	UNP Q8NBF2
D	-1	HIS	-	expression tag	UNP Q8NBF2
D	0	HIS	-	expression tag	UNP Q8NBF2

- Molecule 2 is water.

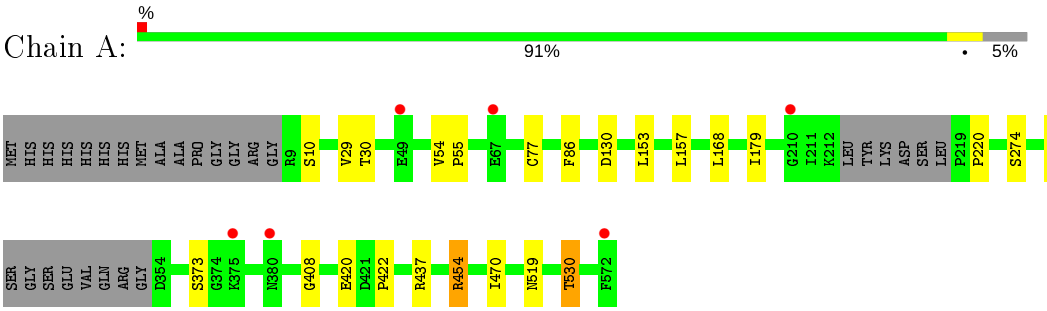
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	13	Total O 13 13	0	0
2	B	10	Total O 10 10	0	0
2	C	6	Total O 6 6	0	0
2	D	8	Total O 8 8	0	0

3 Residue-property plots

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.

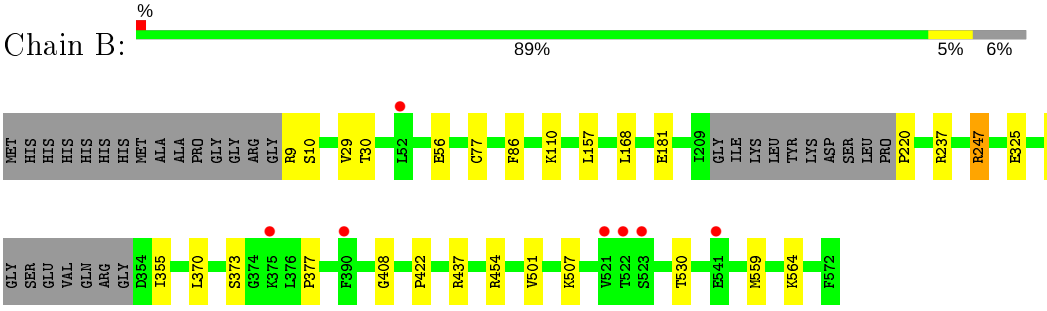
- Molecule 1: NHL repeat-containing protein 2

Chain A:



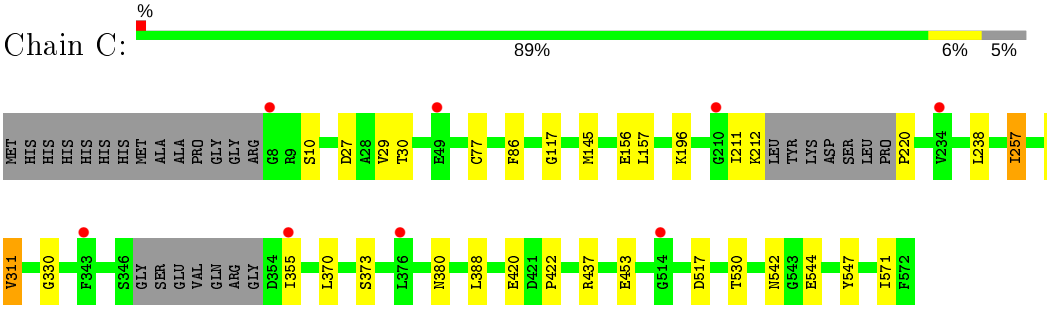
- Molecule 1: NHL repeat-containing protein 2

Chain B:



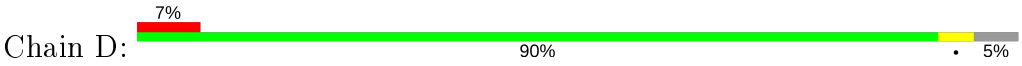
- Molecule 1: NHL repeat-containing protein 2

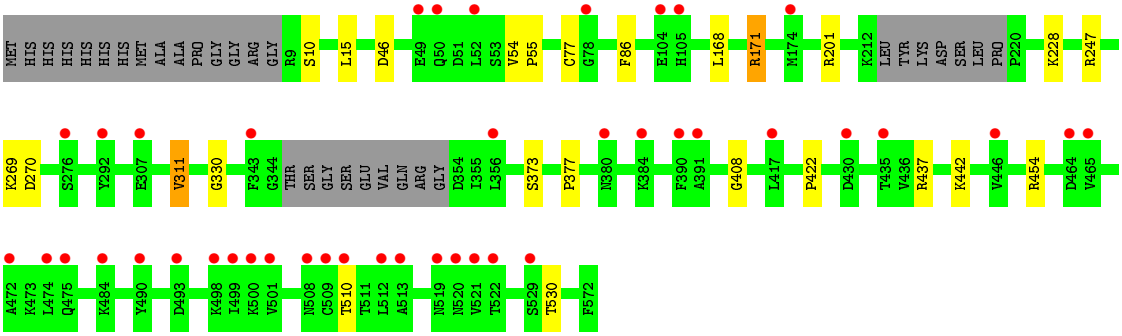
Chain C:



- Molecule 1: NHL repeat-containing protein 2

Chain D:




MET HIS HIS HIS HIS HIS MET ALA ALA PRO GLY GLY ARG GLY R9 S10 L15 D46 E49 Q50 D51 L52 S53 V54 P55 C77 G78 F86 E104 H105 L168 R171 M174 R201 K212 LEU TYR LYS ASP SER LEU PRO P220 K228 R247
K269 D270 S276 Y292 E307 V311 G330 F343 G344 THR SER GLY SER GLU VAL GLN ARG GLY D354 I355 L356 S373 P377 N380 K384 F390 A391 G408 L417 P422 D430 T435 V436 R437 K442 V446 R454 D464 V465
A472 K473 L474 Q475 K484 Y490 D493 K498 I499 K500 V501 N508 C509 T510 T511 L512 A513 N519 N520 V521 T522 S529 T530 F572

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	103.28Å 103.81Å 114.98Å 90.00° 100.77° 90.00°	Depositor
Resolution (Å)	47.77 – 2.70 47.77 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.9 (47.77-2.70) 98.9 (47.77-2.70)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.36 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, R_{free}	0.246 , 0.287 0.249 , 0.287	Depositor DCC
R_{free} test set	3254 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	60.2	Xtriage
Anisotropy	0.262	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 40.6	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	17055	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.47% 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	0.50	0/4352	0.72	0/5910
1	B	0.49	0/4330	0.74	0/5881
1	C	0.49	0/4361	0.72	0/5921
1	D	0.49	0/4344	0.72	0/5898
All	All	0.49	0/17387	0.72	0/23610

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	B	0	3
1	C	0	1
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	237	ARG	Sidechain
1	B	247	ARG	Sidechain
1	B	454	ARG	Sidechain
1	C	437	ARG	Sidechain

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	4259	0	4225	6	0
1	B	4238	0	4198	5	0
1	C	4269	0	4233	9	0
1	D	4252	0	4218	7	0
2	A	13	0	0	0	0
2	B	10	0	0	0	0
2	C	6	0	0	0	0
2	D	8	0	0	1	0
All	All	17055	0	16874	25	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 1.

All (25) 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:C:257:ILE:HD11	1:C:571:ILE:HG23	1.76	0.66
1:D:46:ASP:OD1	1:D:171:ARG:NH2	2.38	0.57
1:C:117:GLY:HA3	1:C:145:MET:HE2	1.87	0.57
1:D:408:GLY:O	1:D:437:ARG:NH2	2.42	0.51
1:A:54:VAL:HG12	1:A:55:PRO:O	2.12	0.50
1:B:408:GLY:O	1:B:437:ARG:NH2	2.42	0.49
1:A:408:GLY:O	1:A:437:ARG:NH2	2.42	0.48
1:B:29:VAL:HG23	1:B:30:THR:HG23	1.97	0.47
1:B:56:GLU:HB2	1:D:15:LEU:HD21	1.97	0.47
1:D:54:VAL:HG12	1:D:55:PRO:O	2.15	0.46
1:A:454:ARG:NH1	1:C:517:ASP:OD1	2.46	0.44
1:C:29:VAL:HG23	1:C:30:THR:HG23	1.98	0.44
1:D:270:ASP:HA	1:D:311:VAL:CG1	2.48	0.44
1:C:238:LEU:HD11	1:C:547:TYR:CE2	2.53	0.43
1:A:29:VAL:HG23	1:A:30:THR:HG23	2.00	0.43
1:C:257:ILE:N	1:C:257:ILE:CD1	2.81	0.43
1:C:270:ASP:HA	1:C:311:VAL:CG1	2.49	0.43
1:C:542:ASN:HB2	1:C:544:GLU:HG2	2.01	0.43
1:A:168:LEU:N	1:A:168:LEU:HD12	2.34	0.42
1:B:355:ILE:HG22	1:B:370:LEU:HA	2.00	0.42
1:C:355:ILE:HG22	1:C:370:LEU:HA	2.00	0.41
1:D:168:LEU:N	1:D:168:LEU:HD12	2.35	0.41
1:D:269:LYS:NZ	2:D:601:HOH:O	2.52	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:519:ASN:O	1:A:530:THR:HG21	2.22	0.40
1:B:168:LEU:N	1:B:168:LEU:HD12	2.36	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	543/579 (94%)	519 (96%)	22 (4%)	2 (0%)	34	60
1	B	540/579 (93%)	517 (96%)	22 (4%)	1 (0%)	47	73
1	C	545/579 (94%)	521 (96%)	23 (4%)	1 (0%)	47	73
1	D	542/579 (94%)	518 (96%)	23 (4%)	1 (0%)	47	73
All	All	2170/2316 (94%)	2075 (96%)	90 (4%)	5 (0%)	47	73

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	330	GLY
1	B	330	GLY
1	C	330	GLY
1	D	330	GLY
1	A	220	PRO

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	472/495 (95%)	458 (97%)	14 (3%)	41	70
1	B	470/495 (95%)	452 (96%)	18 (4%)	33	62
1	C	473/495 (96%)	453 (96%)	20 (4%)	30	58
1	D	471/495 (95%)	456 (97%)	15 (3%)	39	68
All	All	1886/1980 (95%)	1819 (96%)	67 (4%)	35	64

All (67) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	SER
1	A	77	CYS
1	A	86	PHE
1	A	130	ASP
1	A	153	LEU
1	A	157	LEU
1	A	179	ILE
1	A	274	SER
1	A	373	SER
1	A	420	GLU
1	A	422	PRO
1	A	454	ARG
1	A	470	ILE
1	A	530	THR
1	B	9	ARG
1	B	10	SER
1	B	77	CYS
1	B	86	PHE
1	B	110	LYS
1	B	157	LEU
1	B	181	GLU
1	B	220	PRO
1	B	247	ARG
1	B	325	GLU
1	B	373	SER
1	B	377	PRO
1	B	422	PRO
1	B	501	VAL
1	B	507	LYS
1	B	530	THR
1	B	559	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	564	LYS
1	C	10	SER
1	C	27	ASP
1	C	77	CYS
1	C	86	PHE
1	C	156	GLU
1	C	157	LEU
1	C	196	LYS
1	C	211	ILE
1	C	212	LYS
1	C	220	PRO
1	C	257	ILE
1	C	268	ARG
1	C	311	VAL
1	C	373	SER
1	C	380	ASN
1	C	388	LEU
1	C	420	GLU
1	C	422	PRO
1	C	453	GLU
1	C	530	THR
1	D	10	SER
1	D	77	CYS
1	D	86	PHE
1	D	171	ARG
1	D	201	ARG
1	D	228	LYS
1	D	247	ARG
1	D	311	VAL
1	D	373	SER
1	D	377	PRO
1	D	422	PRO
1	D	442	LYS
1	D	454	ARG
1	D	510	THR
1	D	530	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	131	ASN
1	B	508	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	380	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

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	549/579 (94%)	0.14	6 (1%)	80	82	36, 54, 82, 102	0
1	B	546/579 (94%)	0.22	7 (1%)	77	78	37, 59, 81, 134	0
1	C	551/579 (95%)	0.32	8 (1%)	73	76	42, 64, 91, 122	0
1	D	548/579 (94%)	0.58	42 (7%)	13	11	41, 71, 99, 164	0
All	All	2194/2316 (94%)	0.32	63 (2%)	51	52	36, 62, 92, 164	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	521	VAL	4.1
1	C	234	VAL	3.8
1	D	501	VAL	3.8
1	D	435	THR	3.7
1	A	210	GLY	3.7
1	D	472	ALA	3.5
1	D	174	MET	3.5
1	D	49	GLU	3.3
1	D	430	ASP	3.2
1	C	514	GLY	3.1
1	D	474	LEU	3.1
1	D	520	ASN	3.1
1	D	380	ASN	3.0
1	B	522	THR	3.0
1	D	498	LYS	2.9
1	D	490	TYR	2.9
1	D	512	LEU	2.9
1	D	384	LYS	2.8
1	D	390	PHE	2.8
1	C	355	ILE	2.8
1	B	52	LEU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	52	LEU	2.7
1	D	510	THR	2.7
1	D	522	THR	2.7
1	D	493	ASP	2.7
1	D	417	LEU	2.7
1	A	67	GLU	2.7
1	D	475	GLN	2.6
1	B	541	GLU	2.6
1	D	78	GLY	2.6
1	A	49	GLU	2.5
1	D	50	GLN	2.5
1	A	572	PHE	2.5
1	D	529	SER	2.5
1	D	356	LEU	2.5
1	D	500	LYS	2.4
1	D	276	SER	2.4
1	B	390	PHE	2.4
1	D	292	TYR	2.3
1	A	375	LYS	2.3
1	B	523	SER	2.3
1	D	307	GLU	2.3
1	C	376	LEU	2.3
1	D	465	VAL	2.3
1	D	343	PHE	2.3
1	B	521	VAL	2.2
1	C	49	GLU	2.2
1	D	464	ASP	2.2
1	D	513	ALA	2.2
1	B	375	LYS	2.2
1	C	8	GLY	2.2
1	D	484	LYS	2.2
1	A	380	ASN	2.2
1	D	519	ASN	2.2
1	D	509	CYS	2.1
1	D	499	ILE	2.1
1	D	104	GLU	2.1
1	C	343	PHE	2.1
1	D	105	HIS	2.1
1	D	391	ALA	2.0
1	C	210	GLY	2.0
1	D	446	VAL	2.0
1	D	508	ASN	2.0

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.