wwPDB X-ray Structure Validation Summary Report (i)

Sep 3, 2023 – 01:54 AM EDT

PDB ID	:	3RIJ
Title	:	Epitope backbone grafting by computational design for improved presentation of linear epitopes on scaffold proteins
Authors	:	Azoitei, M.L.; Ban, Y.A.; Julien, J.P.; Bryson, S.; Schroeter, A.; Kalyuzhniy, O.; Porter, J.R.; Adachi, Y.; Baker, D.; Szabo, E.; Pai, E.F.; Schief, W.R.
Deposited on	:	2011-04-13
Resolution	:	2.30 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

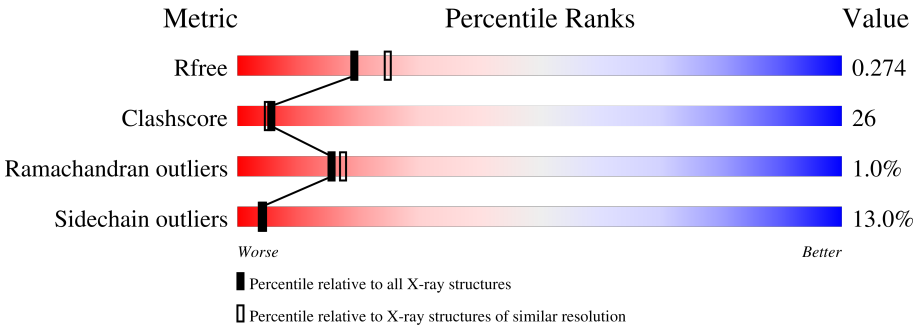
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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%

Mol	Chain	Length	Quality of chain
1	A	165	58% 34% • 5%
1	B	165	56% 35% • 5%
1	C	165	48% 35% 6% • 9%
1	D	165	50% 39% 5% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4991 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 SC_2cx5.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	157	Total	C	N	O	0	0	0
			1149	733	202	214			
1	B	156	Total	C	N	O	0	0	0
			1143	730	201	212			
1	C	150	Total	C	N	O	0	0	0
			1095	702	191	202			
1	D	157	Total	C	N	O	0	0	0
			1149	733	202	214			

There are 48 discrepancies between the modelled and reference sequences:

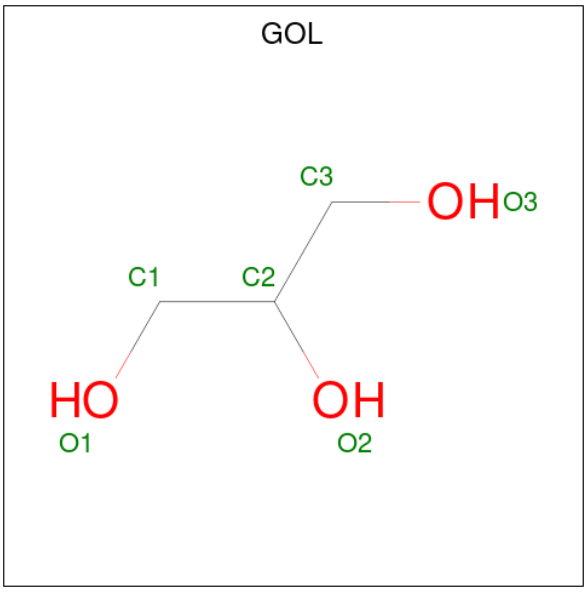
Chain	Residue	Modelled	Actual	Comment	Reference
A	54	TYR	PHE	engineered mutation	UNP Q5SHN1
A	55	GLY	VAL	engineered mutation	UNP Q5SHN1
A	99	ASP	ALA	engineered mutation	UNP Q5SHN1
A	100	ALA	ILE	engineered mutation	UNP Q5SHN1
A	159	LEU	-	insertion	UNP Q5SHN1
A	160	GLU	-	insertion	UNP Q5SHN1
A	161	HIS	-	expression tag	UNP Q5SHN1
A	162	HIS	-	expression tag	UNP Q5SHN1
A	163	HIS	-	expression tag	UNP Q5SHN1
A	164	HIS	-	expression tag	UNP Q5SHN1
A	165	HIS	-	expression tag	UNP Q5SHN1
A	166	HIS	-	expression tag	UNP Q5SHN1
B	54	TYR	PHE	engineered mutation	UNP Q5SHN1
B	55	GLY	VAL	engineered mutation	UNP Q5SHN1
B	99	ASP	ALA	engineered mutation	UNP Q5SHN1
B	100	ALA	ILE	engineered mutation	UNP Q5SHN1
B	159	LEU	-	insertion	UNP Q5SHN1
B	160	GLU	-	insertion	UNP Q5SHN1
B	161	HIS	-	expression tag	UNP Q5SHN1
B	162	HIS	-	expression tag	UNP Q5SHN1
B	163	HIS	-	expression tag	UNP Q5SHN1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	164	HIS	-	expression tag	UNP Q5SHN1
B	165	HIS	-	expression tag	UNP Q5SHN1
B	166	HIS	-	expression tag	UNP Q5SHN1
C	54	TYR	PHE	engineered mutation	UNP Q5SHN1
C	55	GLY	VAL	engineered mutation	UNP Q5SHN1
C	99	ASP	ALA	engineered mutation	UNP Q5SHN1
C	100	ALA	ILE	engineered mutation	UNP Q5SHN1
C	159	LEU	-	insertion	UNP Q5SHN1
C	160	GLU	-	insertion	UNP Q5SHN1
C	161	HIS	-	expression tag	UNP Q5SHN1
C	162	HIS	-	expression tag	UNP Q5SHN1
C	163	HIS	-	expression tag	UNP Q5SHN1
C	164	HIS	-	expression tag	UNP Q5SHN1
C	165	HIS	-	expression tag	UNP Q5SHN1
C	166	HIS	-	expression tag	UNP Q5SHN1
D	54	TYR	PHE	engineered mutation	UNP Q5SHN1
D	55	GLY	VAL	engineered mutation	UNP Q5SHN1
D	99	ASP	ALA	engineered mutation	UNP Q5SHN1
D	100	ALA	ILE	engineered mutation	UNP Q5SHN1
D	159	LEU	-	insertion	UNP Q5SHN1
D	160	GLU	-	insertion	UNP Q5SHN1
D	161	HIS	-	expression tag	UNP Q5SHN1
D	162	HIS	-	expression tag	UNP Q5SHN1
D	163	HIS	-	expression tag	UNP Q5SHN1
D	164	HIS	-	expression tag	UNP Q5SHN1
D	165	HIS	-	expression tag	UNP Q5SHN1
D	166	HIS	-	expression tag	UNP Q5SHN1

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C O 6 3 3	0	0

- Molecule 3 is water.

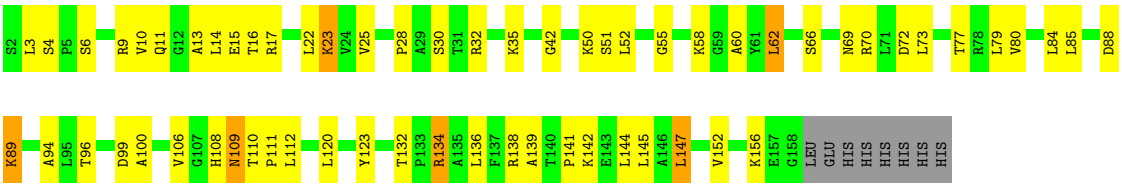
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	120	Total O 120 120	0	0
3	B	115	Total O 115 115	0	0
3	C	75	Total O 75 75	0	0
3	D	139	Total O 139 139	0	0

3 Residue-property plots (i)

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. 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. 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.

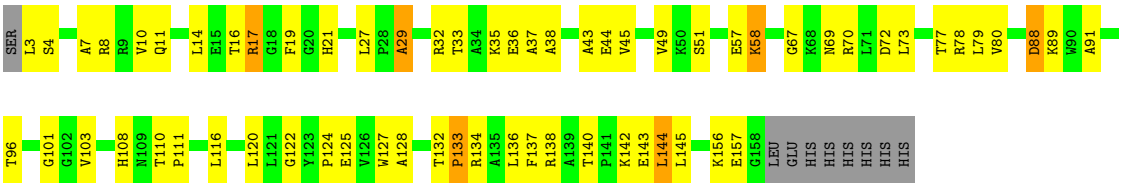
- Molecule 1: SC_2cx5

Chain A:



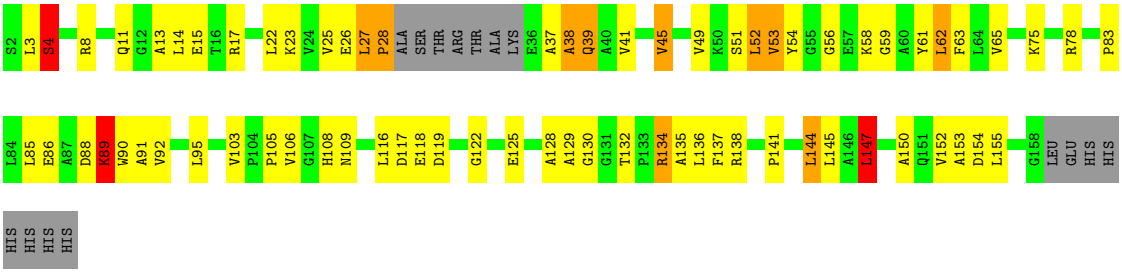
- Molecule 1: SC_2cx5

Chain B:



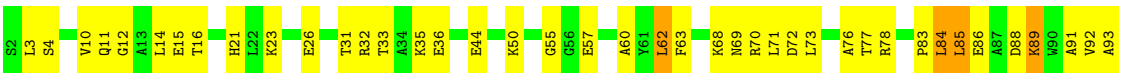
- Molecule 1: SC_2cx5

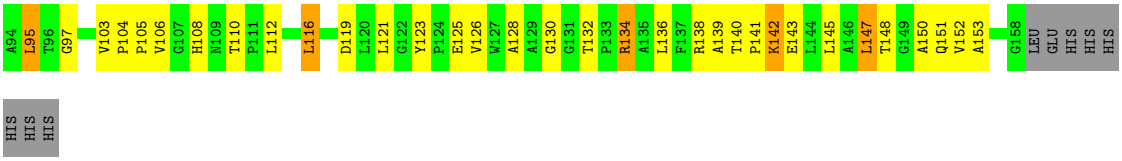
Chain C:



- Molecule 1: SC_2cx5

Chain D:




A94
L95
T96
G97
V103
P104
P105
V106
G107
H108
N109
T110
P111
L112
L116
D119
L120
L121
G122
Y123
P124
E125
V126
W127
A128
A129
G130
G131
T132
P133
R134
A135
L136
F137
R138
A139
T140
P141
K142
E143
L144
L145
A146
L147
T148
G149
A150
Q151
V152
A153
G158
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α, β, γ	70.00Å 70.00Å 131.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.00 – 2.30 16.71 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.3 (17.00-2.30) 98.7 (16.71-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.30Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.252 , 0.294 0.262 , 0.274	Depositor DCC
R_{free} test set	1389 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	32.8	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 36.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.418 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4991	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 36.24 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.1561e-04.*

[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: GOL

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	0.43	0/1172	0.74	1/1593 (0.1%)
1	B	0.39	0/1166	0.71	0/1585
1	C	0.43	0/1117	0.78	2/1518 (0.1%)
1	D	0.45	0/1172	0.73	0/1593
All	All	0.43	0/4627	0.74	3/6289 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	27	LEU	N-CA-C	7.49	131.22	111.00
1	A	136	LEU	CA-CB-CG	5.14	127.12	115.30
1	C	147	LEU	CA-CB-CG	5.08	126.99	115.30

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	1149	0	1174	50	0
1	B	1143	0	1169	52	0
1	C	1095	0	1114	74	0
1	D	1149	0	1174	64	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	6	0	8	0	0
3	A	120	0	0	15	0
3	B	115	0	0	11	0
3	C	75	0	0	9	0
3	D	139	0	0	21	0
All	All	4991	0	4639	236	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 26.

The worst 5 of 236 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:28:PRO:HG2	3:D:409:HOH:O	1.33	1.24
1:D:23:LYS:HD2	3:D:409:HOH:O	1.58	1.02
1:D:106:VAL:HG13	3:D:291:HOH:O	1.65	0.96
1:A:9:ARG:HG3	3:A:300:HOH:O	1.74	0.87
1:C:39:GLN:HG3	1:C:41:VAL:HG22	1.59	0.85

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	155/165 (94%)	145 (94%)	10 (6%)	0	100 100
1	B	154/165 (93%)	142 (92%)	10 (6%)	2 (1%)	12 12
1	C	146/165 (88%)	132 (90%)	10 (7%)	4 (3%)	5 3
1	D	155/165 (94%)	148 (96%)	7 (4%)	0	100 100
All	All	610/660 (92%)	567 (93%)	37 (6%)	6 (1%)	15 17

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	150	ALA
1	C	38	ALA
1	B	29	ALA
1	C	89	LYS
1	C	4	SER

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	113/121 (93%)	96 (85%)	17 (15%)	3	3
1	B	112/121 (93%)	101 (90%)	11 (10%)	8	9
1	C	107/121 (88%)	92 (86%)	15 (14%)	3	3
1	D	113/121 (93%)	98 (87%)	15 (13%)	4	4
All	All	445/484 (92%)	387 (87%)	58 (13%)	4	4

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	4	SER
1	D	142	LYS
1	C	62	LEU
1	D	134	ARG
1	D	95	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	39	GLN
1	C	109	ASN
1	D	109	ASN
1	D	39	GLN
1	B	151	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 monosaccharides in this entry.

5.6 Ligand geometry

1 ligand is 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	GOL	C	3731	-	5,5,5	0.30	0	5,5,5	0.37	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	GOL	C	3731	-	-	0/4/4/4	-

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 (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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.