



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 09:50 pm BST

PDB ID : 2GSI  
Title : Crystal Structure of a Murine Fab in Complex with an 11 Residue Peptide  
Derived from Staphylococcal Nuclease  
Authors : Armstrong, A.A.; Amzel, L.M.  
Deposited on : 2006-04-26  
Resolution : 2.81 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](mailto: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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

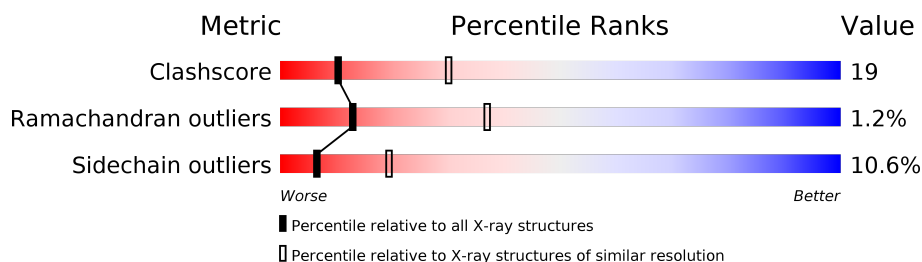
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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))
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	215	67% 30% .
1	C	215	62% 33% 5%
1	E	215	56% 40% 5%
1	G	215	61% 36% .
2	B	221	67% 24% . 5%
2	D	221	71% 24% . .
2	F	221	61% 33% . .
2	H	221	58% 34% . .

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	W	11	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>27%55%9%9%</div>
3	X	11	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>73%27%</div>
3	Y	11	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>27%64%9%</div>
3	Z	11	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>55%18%27%</div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13426 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 Immunoglobulin (kappa) light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	215	Total	C	N	O	S	0	1	0
			1639	1013	282	338	6			
1	C	215	Total	C	N	O	S	0	1	0
			1639	1013	283	337	6			
1	E	215	Total	C	N	O	S	0	3	0
			1651	1020	286	339	6			
1	G	215	Total	C	N	O	S	0	2	0
			1646	1017	286	337	6			

- Molecule 2 is a protein called Immunoglobulin (gamma) heavy chain (VH + CH1 fragment).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	210	Total	C	N	O	S	0	0	0
			1588	1008	252	319	9			
2	D	215	Total	C	N	O	S	0	0	0
			1618	1025	258	326	9			
2	F	214	Total	C	N	O	S	0	0	0
			1613	1022	257	325	9			
2	H	213	Total	C	N	O	S	0	2	0
			1624	1031	257	326	10			

- Molecule 3 is a protein called Thermonuclease.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	X	11	Total	C	N	O	0	0	0
			93	60	17	16			
3	Z	8	Total	C	N	O	0	0	0
			67	44	11	12			
3	W	10	Total	C	N	O	0	0	0
			86	56	16	14			
3	Y	11	Total	C	N	O	0	0	0
			93	60	17	16			

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	2	Total 2	Na 2	0	0
4	D	1	Total 1	Na 1	0	0
4	E	1	Total 1	Na 1	0	0
4	B	1	Total 1	Na 1	0	0
4	W	1	Total 1	Na 1	0	0
4	F	1	Total 1	Na 1	0	0

- Molecule 5 is water.

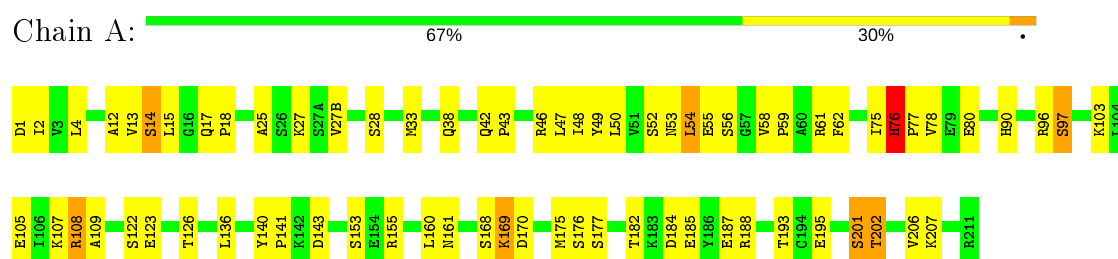
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	7	Total 7	O 7	0	1
5	B	3	Total 3	O 3	0	0
5	C	11	Total 11	O 11	0	0
5	D	17	Total 17	O 17	0	0
5	E	4	Total 4	O 4	0	0
5	F	8	Total 8	O 8	0	0
5	G	4	Total 4	O 4	0	0
5	H	7	Total 7	O 7	0	0
5	W	1	Total 1	O 1	0	0

### 3 Residue-property plots

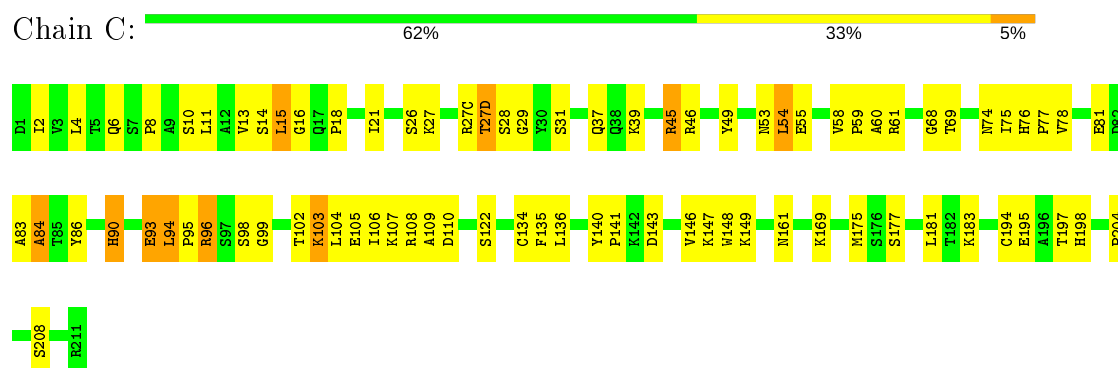
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

Note EDS was not executed.

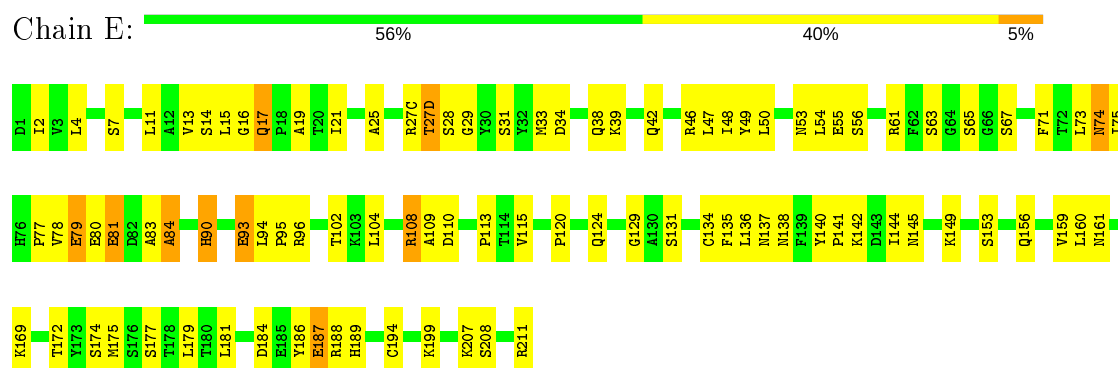
#### • Molecule 1: Immunoglobulin (kappa) light chain



#### • Molecule 1: Immunoglobulin (kappa) light chain

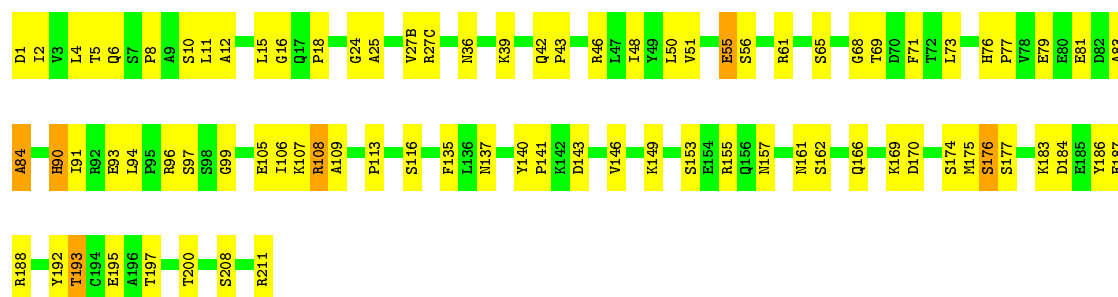


#### • Molecule 1: Immunoglobulin (kappa) light chain



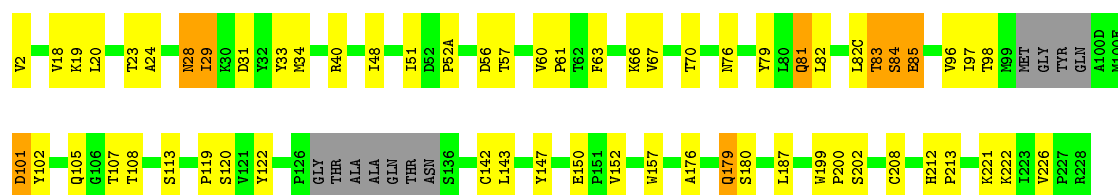
#### • Molecule 1: Immunoglobulin (kappa) light chain

Chain G: 



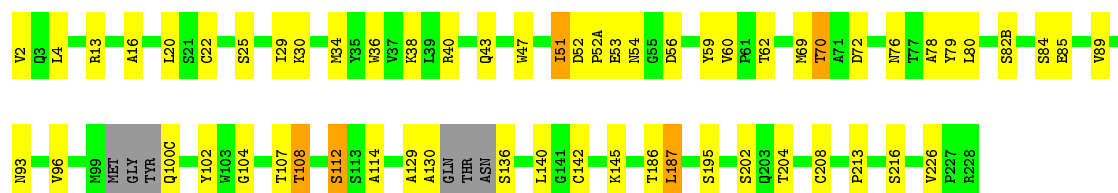
- Molecule 2: Immunoglobulin (gamma) heavy chain (VH + CH1 fragment)

Chain B: 



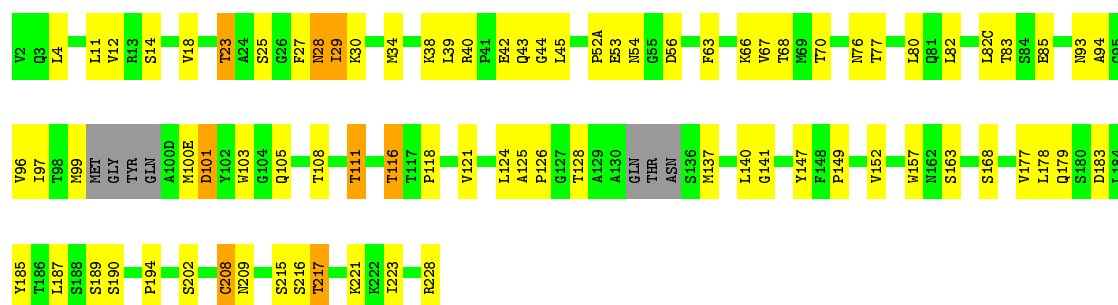
- Molecule 2: Immunoglobulin (gamma) heavy chain (VH + CH1 fragment)

Chain D: 



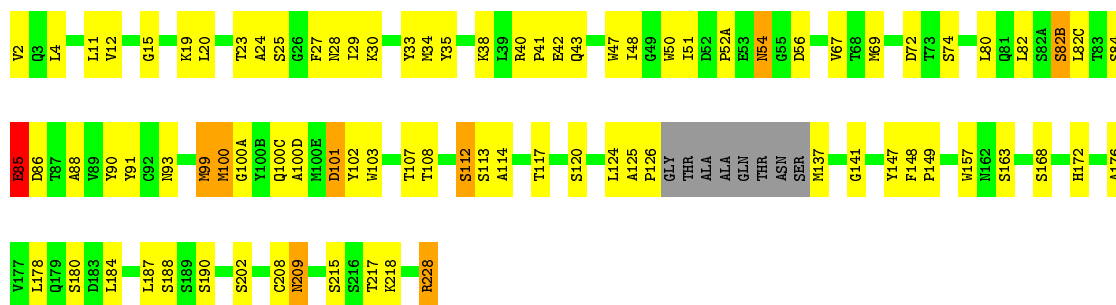
- Molecule 2: Immunoglobulin (gamma) heavy chain (VH + CH1 fragment)

Chain F: 



- Molecule 2: Immunoglobulin (gamma) heavy chain (VH + CH1 fragment)

Chain H: 



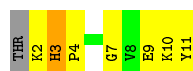
• Molecule 3: Thermonuclease



• Molecule 3: Thermonuclease



• Molecule 3: Thermonuclease



• Molecule 3: Thermonuclease





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.74Å 131.04Å 291.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.74 – 2.81	Depositor
% Data completeness (in resolution range)	84.2 (48.74-2.81)	Depositor
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.221 , 0.293	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	13426	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

## 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:  
NA

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	0.76	0/1679	0.69	0/2281
1	C	0.75	0/1679	0.67	0/2282
1	E	0.74	0/1699	0.68	0/2308
1	G	0.68	0/1690	0.62	0/2296
2	B	0.74	0/1628	0.66	0/2230
2	D	0.78	0/1658	0.68	0/2271
2	F	0.72	0/1653	0.67	0/2264
2	H	0.80	0/1674	0.73	0/2293
3	W	0.61	0/88	0.64	0/113
3	X	0.82	0/95	0.51	0/123
3	Y	0.67	0/95	0.60	0/123
3	Z	0.70	0/68	0.58	0/86
All	All	0.75	0/13706	0.67	0/18670

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	C	0	1
2	H	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	75	ILE	Peptide
2	H	100(A)	GLY	Peptide

## 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	1639	0	1565	68	0
1	C	1639	0	1567	73	0
1	E	1651	0	1578	66	0
1	G	1646	0	1575	66	0
2	B	1588	0	1539	64	0
2	D	1618	0	1566	36	0
2	F	1613	0	1565	66	0
2	H	1624	0	1568	70	0
3	W	86	0	92	10	0
3	X	93	0	102	3	0
3	Y	93	0	102	9	0
3	Z	67	0	73	1	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	2	0	0	0	0
4	W	1	0	0	0	0
5	A	7	0	0	0	0
5	B	3	0	0	0	0
5	C	11	0	0	2	0
5	D	17	0	0	0	0
5	E	4	0	0	3	0
5	F	8	0	0	0	0
5	G	4	0	0	0	0
5	H	7	0	0	0	0
5	W	1	0	0	0	0
All	All	13426	0	12892	491	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 19.

The worst 5 of 491 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:E:46[B]:ARG:HD3	1:E:55[B]:GLU:OE2	1.47	1.14
2:B:2:VAL:HG13	2:B:102:TYR:HE2	1.20	1.05
1:E:46[B]:ARG:CD	1:E:55[B]:GLU:OE2	2.05	1.03
2:B:29:ILE:HB	2:B:34:MET:HE3	1.41	1.03
1:G:2:ILE:HD12	1:G:90:HIS:CE1	1.94	1.02

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	214/215 (100%)	198 (92%)	14 (6%)	2 (1%)	17	44
1	C	214/215 (100%)	190 (89%)	19 (9%)	5 (2%)	6	20
1	E	216/215 (100%)	199 (92%)	14 (6%)	3 (1%)	11	32
1	G	215/215 (100%)	199 (93%)	15 (7%)	1 (0%)	29	59
2	B	204/221 (92%)	187 (92%)	16 (8%)	1 (0%)	29	59
2	D	209/221 (95%)	199 (95%)	10 (5%)	0	100	100
2	F	208/221 (94%)	191 (92%)	15 (7%)	2 (1%)	15	42
2	H	211/221 (96%)	191 (90%)	14 (7%)	6 (3%)	5	15
3	W	8/11 (73%)	6 (75%)	1 (12%)	1 (12%)	0	1
3	X	9/11 (82%)	5 (56%)	4 (44%)	0	100	100
3	Y	9/11 (82%)	8 (89%)	1 (11%)	0	100	100
3	Z	6/11 (54%)	5 (83%)	1 (17%)	0	100	100
All	All	1723/1788 (96%)	1578 (92%)	124 (7%)	21 (1%)	13	37

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	202	THR
1	E	28	SER
1	E	84	ALA
1	G	84	ALA
2	H	42	GLU

### 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	186/188 (99%)	170 (91%)	16 (9%)	10	29
1	C	186/188 (99%)	164 (88%)	22 (12%)	5	15
1	E	188/188 (100%)	164 (87%)	24 (13%)	4	13
1	G	187/188 (100%)	170 (91%)	17 (9%)	9	26
2	B	182/190 (96%)	167 (92%)	15 (8%)	11	31
2	D	184/190 (97%)	163 (89%)	21 (11%)	5	17
2	F	184/190 (97%)	159 (86%)	25 (14%)	3	11
2	H	186/190 (98%)	167 (90%)	19 (10%)	7	21
3	W	9/10 (90%)	9 (100%)	0	100	100
3	X	10/10 (100%)	10 (100%)	0	100	100
3	Y	10/10 (100%)	8 (80%)	2 (20%)	1	3
3	Z	7/10 (70%)	6 (86%)	1 (14%)	3	9
All	All	1519/1552 (98%)	1357 (89%)	162 (11%)	6	19

5 of 162 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	226	VAL
1	E	142	LYS
2	H	101[B]	ASP
1	E	17[A]	GLN
1	E	65	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	5	GLN
1	E	76	HIS
2	H	28	ASN
1	E	74	ASN
1	E	90	HIS

### 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](#)

Of 7 ligands modelled in this entry, 7 are 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 [i](#)

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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.