



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 08:37 pm BST

PDB ID : 5HCG  
Title : CRYSTAL STRUCTURE OF FREE-CODV.  
Authors : Vallee, F.; Dupuy, A.; Rak, A.  
Deposited on : 2016-01-04  
Resolution : 2.68 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](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) : 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

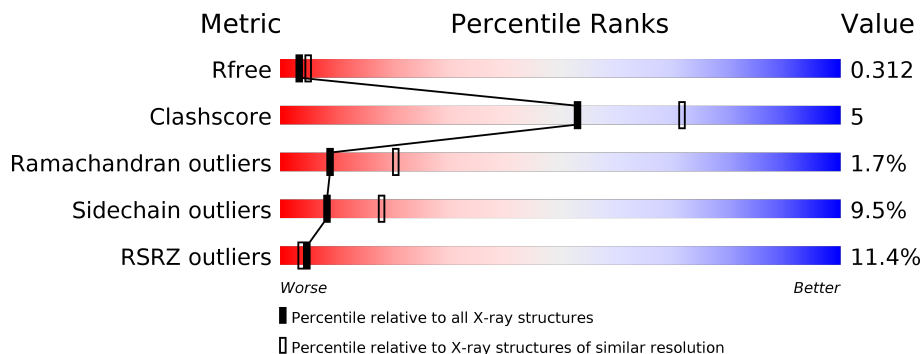
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.68 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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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  $\leq 5\%$ . 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	H	345	
2	L	334	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5102 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 CODV heavy-chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	H	332	2530	1600	419	500	11	0	0	0

- Molecule 2 is a protein called CODV light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	327	2485	1558	417	501	9	0	0	0

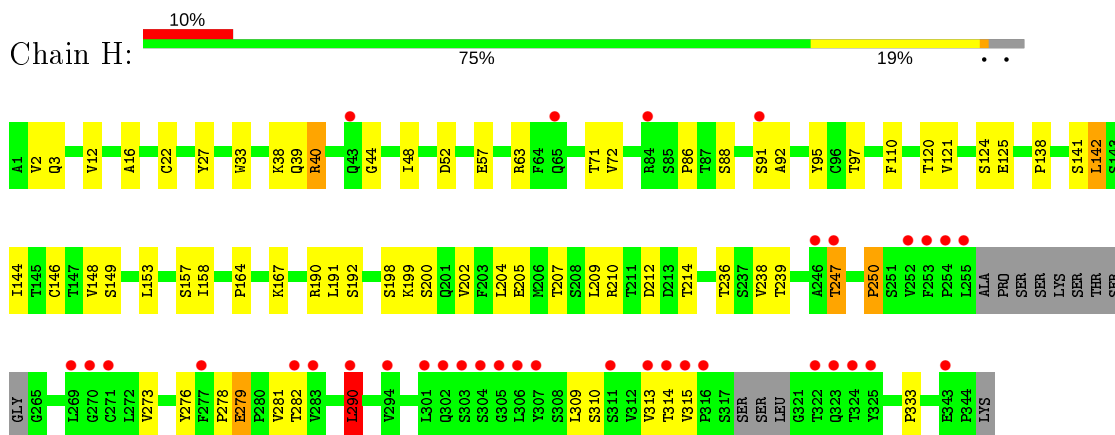
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	H	42	Total	O	0	0
			42	42		
3	L	45	Total	O	0	0
			45	45		

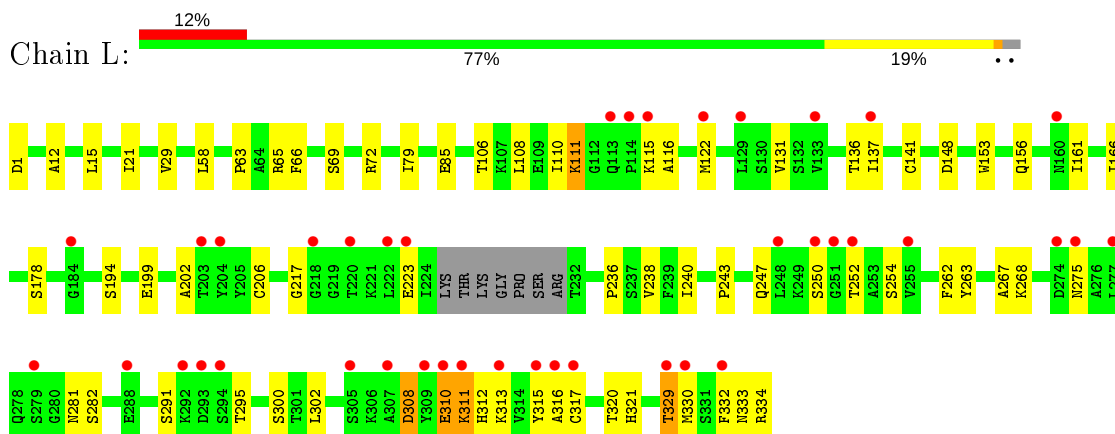
### 3 Residue-property plots [i](#)

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: CODV heavy-chain



- Molecule 2: CODV light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.92Å 70.92Å 272.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.16 – 2.68 47.07 – 2.68	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.16-2.68) 100.0 (47.07-2.68)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.06 (at 2.69Å)	Xtrriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.229 , 0.296 0.238 , 0.312	Depositor DCC
$R_{free}$ test set	1006 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	60.2	Xtrriage
Anisotropy	0.113	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 66.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	5102	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	H	0.52	0/2592	0.78	1/3528 (0.0%)
2	L	0.52	0/2541	0.74	0/3456
All	All	0.52	0/5133	0.76	1/6984 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	H	250	PRO	N-CA-CB	5.81	110.28	103.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	H	2530	0	2439	27	0
2	L	2485	0	2398	27	0
3	H	42	0	0	0	0
3	L	45	0	0	0	0
All	All	5102	0	4837	53	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 5.

All (53) 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)
2:L:141:CYS:HG	2:L:206:CYS:HG	1.27	0.80
2:L:315:TYR:HB2	2:L:332:PHE:CE2	2.31	0.65
1:H:142:LEU:HD12	1:H:238:VAL:HG11	1.82	0.62
2:L:12:ALA:HB1	2:L:111:LYS:HB3	1.82	0.61
2:L:267:ALA:HB2	2:L:321:HIS:HD2	1.67	0.60
1:H:144:ILE:HD12	1:H:204:LEU:HD23	1.86	0.58
1:H:39:GLN:HG3	1:H:44:GLY:O	2.04	0.57
1:H:190:ARG:HD2	1:H:207:THR:O	2.05	0.57
1:H:164:PRO:HB2	1:H:167:LYS:HB2	1.87	0.57
1:H:273:VAL:HG11	1:H:281:VAL:HG11	1.89	0.55
1:H:16:ALA:O	1:H:86:PRO:HD3	2.07	0.54
2:L:268:LYS:HB3	2:L:320:THR:HB	1.90	0.53
2:L:240:ILE:HD12	2:L:317:CYS:HB2	1.89	0.53
1:H:95:TYR:HE1	2:L:161:ILE:HG22	1.74	0.52
2:L:63:PRO:HB2	2:L:65:ARG:HG2	1.92	0.52
1:H:12:VAL:HG13	1:H:16:ALA:HB3	1.92	0.51
2:L:250:SER:HB2	2:L:252:THR:HG22	1.93	0.50
2:L:316:ALA:HB1	2:L:329:THR:HG22	1.93	0.49
1:H:138:PRO:HA	1:H:209:LEU:O	2.13	0.49
1:H:191:LEU:HD22	1:H:204:LEU:HD11	1.94	0.49
2:L:21:ILE:HG12	2:L:106:THR:HG21	1.95	0.48
2:L:282:SER:HB3	2:L:302:LEU:HD12	1.96	0.48
2:L:136:THR:HA	2:L:194:SER:HA	1.95	0.48
2:L:153:TRP:HB2	2:L:166:ILE:HB	1.96	0.47
1:H:2:VAL:HG22	1:H:27:TYR:HB3	1.96	0.47
2:L:308:ASP:HA	2:L:311:LYS:HE3	1.97	0.47
2:L:263:TYR:HB2	2:L:295:THR:HG22	1.98	0.46
1:H:97:THR:OG1	1:H:110:PHE:HB3	2.16	0.46
2:L:58:LEU:HD11	2:L:66:PHE:O	2.16	0.46
1:H:278:PRO:HD2	1:H:333:PRO:HB2	1.99	0.45
2:L:247:GLN:HG2	2:L:252:THR:HG23	1.98	0.45
1:H:39:GLN:O	1:H:40:ARG:HB2	2.17	0.44
1:H:276:TYR:HE2	1:H:279:GLU:HA	1.82	0.44
2:L:122:MET:HB2	2:L:217:GLY:HA2	1.99	0.44
2:L:243:PRO:HB3	2:L:254:SER:H	1.82	0.44
1:H:86:PRO:HB2	1:H:121:VAL:HG21	1.99	0.44
1:H:146:CYS:HB3	1:H:202:VAL:HB	2.00	0.44
2:L:312:HIS:HB2	2:L:315:TYR:CE1	2.54	0.43
2:L:313:LYS:HG3	2:L:334:ARG:HE	1.83	0.43
1:H:290:LEU:HD23	1:H:313:VAL:HG21	2.00	0.43
2:L:236:PRO:HB3	2:L:262:PHE:CD2	2.54	0.42
1:H:38:LYS:HB2	1:H:48:ILE:HD11	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:164:PRO:HB2	1:H:167:LYS:HD2	2.02	0.42
1:H:158:ILE:HG13	1:H:202:VAL:HG21	2.01	0.42
1:H:214:THR:HG23	1:H:239:THR:HG22	2.02	0.41
1:H:214:THR:HG23	1:H:239:THR:HA	2.02	0.41
1:H:192:SER:HB3	1:H:205:GLU:HB2	2.02	0.41
1:H:33:TRP:CH2	1:H:52:ASP:HB2	2.56	0.41
2:L:156:GLN:O	2:L:202:ALA:HB1	2.21	0.41
2:L:311:LYS:HD2	2:L:312:HIS:CE1	2.56	0.40
2:L:317:CYS:O	2:L:329:THR:HA	2.21	0.40
1:H:40:ARG:HG3	1:H:92:ALA:HB2	2.03	0.40
2:L:66:PHE:CD1	2:L:79:ILE:HG12	2.56	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	H	326/345 (94%)	295 (90%)	25 (8%)	6 (2%)	8	19
2	L	323/334 (97%)	295 (91%)	23 (7%)	5 (2%)	10	23
All	All	649/679 (96%)	590 (91%)	48 (7%)	11 (2%)	9	20

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	250	PRO
2	L	115	LYS
2	L	116	ALA
2	L	72	ARG
2	L	148	ASP
1	H	40	ARG
1	H	124	SER

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	H	125	GLU
1	H	247	THR
2	L	310	GLU
1	H	290	LEU

### 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	H	280/294 (95%)	251 (90%)	29 (10%)	7 14
2	L	277/285 (97%)	253 (91%)	24 (9%)	10 21
All	All	557/579 (96%)	504 (90%)	53 (10%)	8 18

All (53) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	3	GLN
1	H	22	CYS
1	H	57	GLU
1	H	63	ARG
1	H	71	THR
1	H	72	VAL
1	H	88	SER
1	H	91	SER
1	H	120	THR
1	H	141	SER
1	H	142	LEU
1	H	148	VAL
1	H	149	SER
1	H	153	LEU
1	H	157	SER
1	H	198	SER
1	H	199	LYS
1	H	200	SER
1	H	210	ARG

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	H	212	ASP
1	H	236	THR
1	H	247	THR
1	H	279	GLU
1	H	282	THR
1	H	290	LEU
1	H	309	LEU
1	H	310	SER
1	H	314	THR
1	H	315	VAL
2	L	1	ASP
2	L	15	LEU
2	L	29	VAL
2	L	69	SER
2	L	85	GLU
2	L	108	LEU
2	L	110	ILE
2	L	111	LYS
2	L	131	VAL
2	L	137	ILE
2	L	178	SER
2	L	199	GLU
2	L	223	GLU
2	L	238	VAL
2	L	275	ASN
2	L	281	ASN
2	L	291	SER
2	L	300	SER
2	L	308	ASP
2	L	310	GLU
2	L	311	LYS
2	L	329	THR
2	L	330	MET
2	L	333	ASN

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

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	H	104	ASN
2	L	113	GLN
2	L	121	GLN

### 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](#)

There are no ligands in this entry.

### 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

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	H	332/345 (96%)	0.55	35 (10%) <b>6</b>   <b>4</b>	29, 66, 113, 129	0
2	L	327/334 (97%)	0.61	40 (12%) <b>4</b>   <b>3</b>	31, 71, 116, 134	0
All	All	659/679 (97%)	0.58	75 (11%) <b>5</b>   <b>4</b>	29, 69, 114, 134	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	269	LEU	8.6
2	L	293	ASP	7.5
1	H	315	VAL	6.0
1	H	306	LEU	5.9
2	L	310	GLU	5.7
2	L	251	GLY	5.6
2	L	311	LYS	5.6
2	L	252	THR	5.4
1	H	253	PHE	5.4
1	H	324	THR	5.2
2	L	113	GLN	5.1
2	L	114	PRO	5.1
2	L	292	LYS	4.9
1	H	294	VAL	4.8
2	L	218	GLY	4.5
1	H	247	THR	4.4
1	H	313	VAL	4.1
2	L	309	TYR	4.1
2	L	129	LEU	4.0
2	L	294	SER	4.0
2	L	315	TYR	3.9
2	L	277	LEU	3.7
2	L	220	THR	3.7
2	L	275	ASN	3.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	L	330	MET	3.5
1	H	270	GLY	3.5
1	H	255	LEU	3.4
1	H	322	THR	3.2
1	H	323	GLN	3.2
1	H	314	THR	3.2
2	L	250	SER	3.1
1	H	325	TYR	3.1
1	H	343	GLU	3.1
2	L	160	ASN	3.0
2	L	307	ALA	3.0
1	H	246	ALA	3.0
1	H	303	SER	2.9
2	L	313	LYS	2.9
1	H	305	GLY	2.9
1	H	301	LEU	2.9
2	L	222	LEU	2.9
1	H	277	PHE	2.9
2	L	332	PHE	2.8
2	L	184	GLY	2.8
1	H	43	GLN	2.8
1	H	316	PRO	2.7
1	H	290	LEU	2.7
1	H	65	GLN	2.6
1	H	282	THR	2.6
1	H	84	ARG	2.6
2	L	248	LEU	2.6
1	H	311	SER	2.4
1	H	304	SER	2.4
2	L	133	VAL	2.4
2	L	305	SER	2.3
2	L	204	TYR	2.3
1	H	254	PRO	2.3
2	L	329	THR	2.3
2	L	203	THR	2.2
2	L	274	ASP	2.2
1	H	283	VAL	2.2
1	H	91	SER	2.2
2	L	279	SER	2.2
2	L	115	LYS	2.2
1	H	307	TYR	2.1
2	L	223	GLU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	L	288	GLU	2.1
2	L	255	VAL	2.1
2	L	122	MET	2.1
2	L	317	CYS	2.1
1	H	252	VAL	2.1
2	L	137	ILE	2.1
1	H	302	GLN	2.1
2	L	316	ALA	2.1
1	H	271	CYS	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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.