



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:59 AM GMT

PDB ID : 3CUS  
Title : Structure of a double ILE/PHE mutant of NI-FE hydrogenase refined at 2.2 angstrom resolution  
Authors : Volbeda, A.  
Deposited on : 2008-04-17  
Resolution : 2.20 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  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

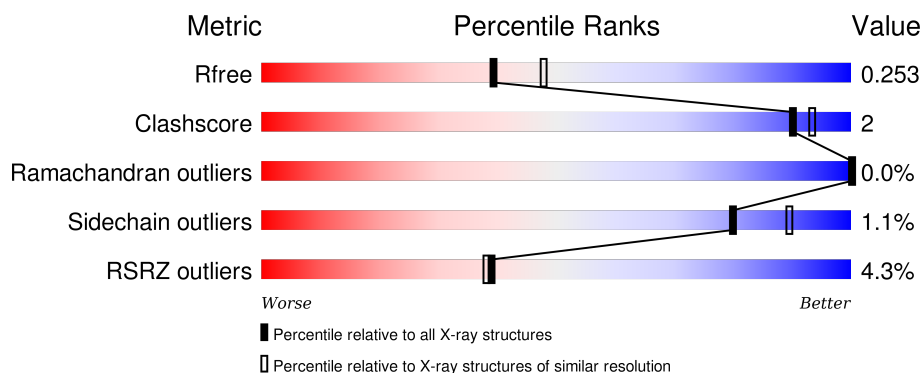
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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.



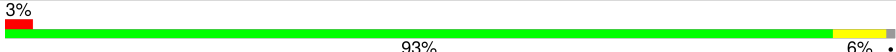
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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. 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	A	264	<div> <div>2%</div> <div>93%</div> <div>6%</div> </div>
1	B	264	<div> <div>19%</div> <div>92%</div> <div>7%</div> </div>
1	C	264	<div> <div>%</div> <div>95%</div> <div>• •</div> </div>
2	Q	549	<div> <div>%</div> <div>93%</div> <div>6%</div> </div>
2	R	549	<div> <div>5%</div> <div>95%</div> <div>• •</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	S	549	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
8	GOL	Q	562	-	-	-	X

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 19590 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 Periplasmic [NiFe] hydrogenase small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	262	Total	C	N	O	S	0	0	0
			1973	1256	330	372	15			
1	B	262	Total	C	N	O	S	0	2	0
			1980	1260	330	375	15			
1	C	260	Total	C	N	O	S	0	1	0
			1965	1251	329	370	15			

- Molecule 2 is a protein called Periplasmic [NiFe] hydrogenase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Q	544	Total	C	N	O	S	0	4	0
			4185	2668	724	770	23			
2	R	545	Total	C	N	O	S	0	1	0
			4178	2661	725	770	22			
2	S	544	Total	C	N	O	S	0	1	0
			4170	2657	724	767	22			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	74	ILE	VAL	ENGINEERED MUTATION	UNP P18188
Q	122	PHE	LEU	ENGINEERED MUTATION	UNP P18188
R	74	ILE	VAL	ENGINEERED MUTATION	UNP P18188
R	122	PHE	LEU	ENGINEERED MUTATION	UNP P18188
S	74	ILE	VAL	ENGINEERED MUTATION	UNP P18188
S	122	PHE	LEU	ENGINEERED MUTATION	UNP P18188

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	R	1	Total	Ni	0	0
			1	1		

*Continued on next page...*

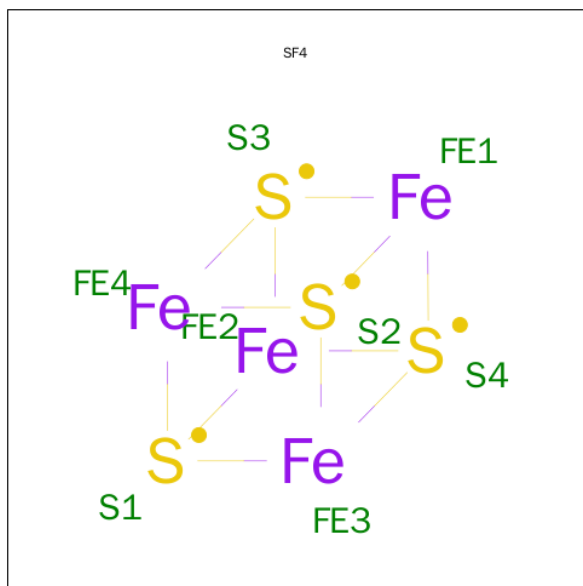
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	Q	1	Total	Ni	0	0
			1	1		
3	S	1	Total	Ni	0	0
			1	1		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	R	1	Total	Mg	0	0
			1	1		
4	Q	1	Total	Mg	0	0
			1	1		
4	S	1	Total	Mg	0	0
			1	1		

- Molecule 5 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



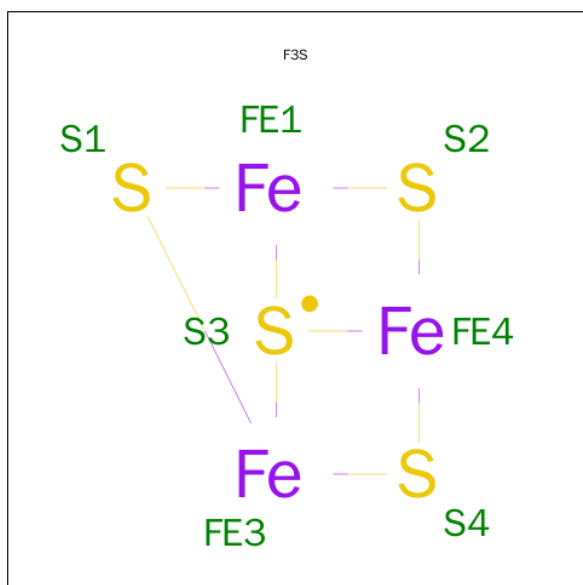
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	Fe	S	0	0
			8	4	4		
5	A	1	Total	Fe	S	0	0
			8	4	4		
5	B	1	Total	Fe	S	0	0
			8	4	4		
5	B	1	Total	Fe	S	0	0
			8	4	4		

*Continued on next page...*

Continued from previous page...

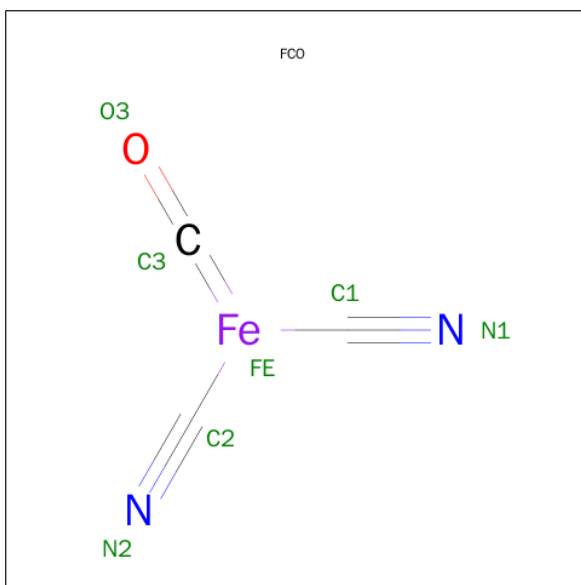
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	Fe	S	0	0
			8	4	4		
5	C	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 6 is FE3-S4 CLUSTER (three-letter code: F3S) (formula:  $\text{Fe}_3\text{S}_4$ ).



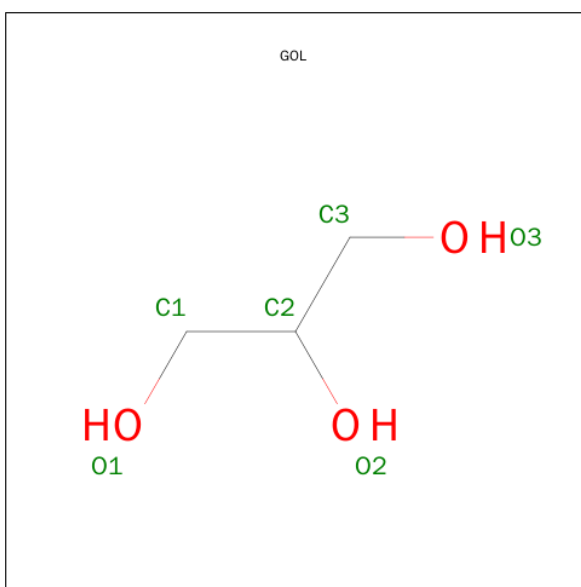
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	Fe	S	0	0
			7	3	4		
6	B	1	Total	Fe	S	0	0
			7	3	4		
6	C	1	Total	Fe	S	0	0
			7	3	4		

- Molecule 7 is CARBONMONOXIDE-(DICYANO) IRON (three-letter code: FCO) (formula:  $\text{C}_3\text{FeN}_2\text{O}$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	Q	1	Total	C	Fe	N	O	0	0
			7	3	1	2	1		
7	R	1	Total	C	Fe	N	O	0	0
			7	3	1	2	1		
7	S	1	Total	C	Fe	N	O	0	0
			7	3	1	2	1		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	Q	1	Total 6	C 3	O 3	0	0
8	Q	1	Total 6	C 3	O 3	0	0
8	R	1	Total 6	C 3	O 3	0	0
8	S	1	Total 6	C 3	O 3	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	134	Total 134	O 134	0	0
9	Q	245	Total 245	O 245	0	0
9	B	91	Total 91	O 91	0	0
9	R	215	Total 215	O 215	0	0
9	C	123	Total 123	O 123	0	0
9	S	205	Total 205	O 205	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 errors 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.

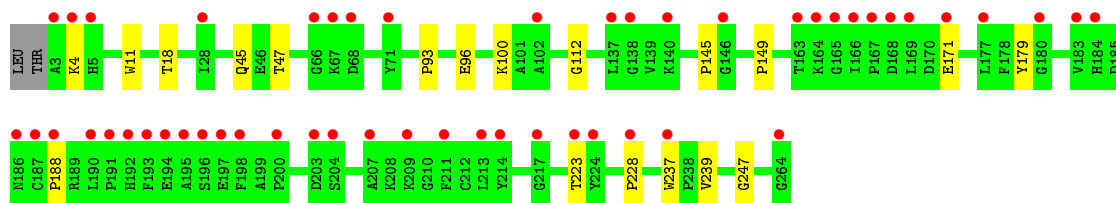
- Molecule 1: Periplasmic [NiFe] hydrogenase small subunit

Chain A: 



- Molecule 1: Periplasmic [NiFe] hydrogenase small subunit

Chain B: 

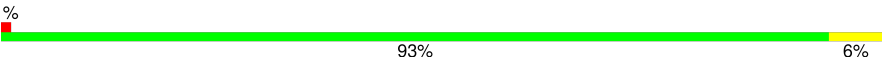


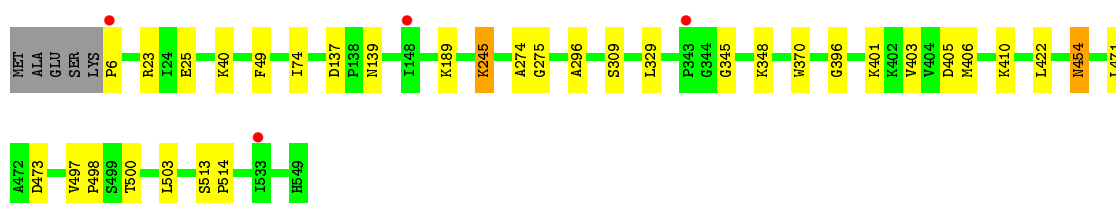
- Molecule 1: Periplasmic [NiFe] hydrogenase small subunit

Chain C: 

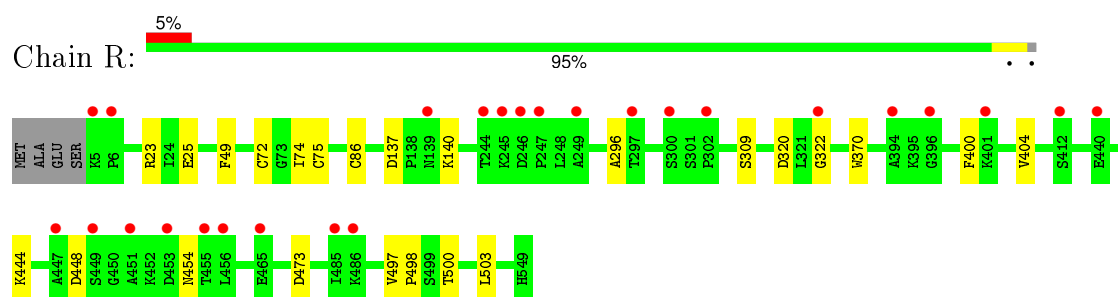


- Molecule 2: Periplasmic [NiFe] hydrogenase large subunit

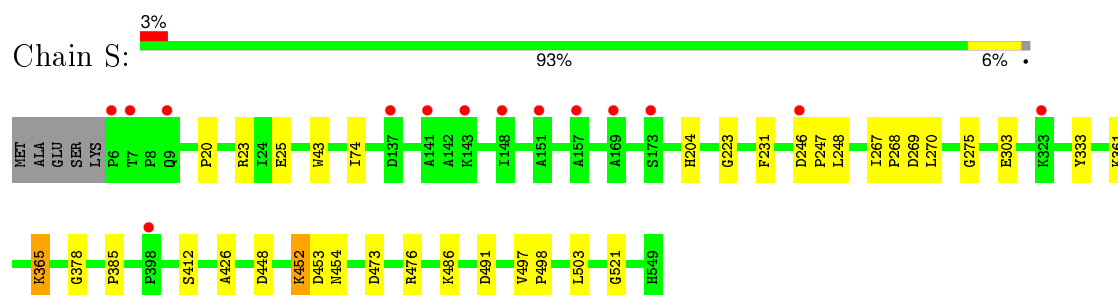
Chain Q: 



- Molecule 2: Periplasmic [NiFe] hydrogenase large subunit



- Molecule 2: Periplasmic [NiFe] hydrogenase large subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.40 Å 99.96 Å 182.69 Å 90.00° 92.23° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20 19.99 – 2.20	Depositor EDS
% Data completeness (in resolution range)	94.2 (20.00-2.20) 94.2 (19.99-2.20)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.25 (at 2.19 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.185 , 0.222 0.219 , 0.253	Depositor DCC
$R_{free}$ test set	5404 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.7	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 39.2	EDS
Estimated twinning fraction	0.048 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 107158 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	19590	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% 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, 0.375 respectively for untwinned datasets, and 0.333, 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, CSO, NI, SF4, MG, F3S, FCO

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	0.62	0/2027	0.60	0/2759
1	B	0.50	0/2043	0.55	0/2781
1	C	0.54	0/2025	0.59	0/2758
2	Q	0.61	0/4302	0.65	0/5835
2	R	0.56	0/4281	0.60	0/5809
2	S	0.55	1/4273 (0.0%)	0.62	0/5798
All	All	0.57	1/18951 (0.0%)	0.61	0/25740

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	S	303	GLU	CB-CG	-5.21	1.42	1.52

There are no bond angle outliers.

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	1973	0	1911	11	0
1	B	1980	0	1918	11	0
1	C	1965	0	1900	5	0
2	Q	4185	0	4161	23	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	R	4178	0	4142	15	0
2	S	4170	0	4138	22	0
3	Q	1	0	0	0	0
3	R	1	0	0	0	0
3	S	1	0	0	0	0
4	Q	1	0	0	0	0
4	R	1	0	0	0	0
4	S	1	0	0	0	0
5	A	16	0	0	0	0
5	B	16	0	0	0	0
5	C	16	0	0	0	0
6	A	7	0	0	0	0
6	B	7	0	0	0	0
6	C	7	0	0	0	0
7	Q	7	0	0	0	0
7	R	7	0	0	1	0
7	S	7	0	0	1	0
8	A	6	0	8	0	0
8	Q	12	0	16	0	0
8	R	6	0	8	0	0
8	S	6	0	8	0	0
9	A	134	0	0	3	0
9	B	91	0	0	1	0
9	C	123	0	0	5	0
9	Q	245	0	0	4	0
9	R	215	0	0	4	0
9	S	205	0	0	4	0
All	All	19590	0	18210	83	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 2.

All (83) 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:R:322:GLY:HA3	9:R:721:HOH:O	1.71	0.89
9:C:341:HOH:O	2:S:361:LYS:HE3	1.81	0.81
1:C:57:GLU:HB2	9:C:390:HOH:O	1.84	0.78
2:S:385:PRO:HB2	9:S:614:HOH:O	1.86	0.75
1:A:215:GLU:HG3	9:A:373:HOH:O	1.88	0.72
2:Q:245:LYS:HD3	9:Q:785:HOH:O	1.91	0.69

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:S:412:SER:HB2	9:S:715:HOH:O	1.92	0.68
2:R:137:ASP:OD2	2:R:140:LYS:HG3	1.97	0.64
1:A:98:THR:HG22	1:A:137:LEU:HD11	1.79	0.63
2:R:497:VAL:CG1	2:R:498:PRO:HD2	2.31	0.60
1:B:188:PRO:HA	9:B:331:HOH:O	2.02	0.60
1:B:47:THR:O	2:R:23:ARG:HA	2.01	0.59
2:S:25:GLU:HB2	2:S:74:ILE:HD12	1.86	0.56
2:Q:497:VAL:CG1	2:Q:498:PRO:HD2	2.38	0.53
2:Q:6:PRO:HG3	9:R:721:HOH:O	2.09	0.53
1:A:215:GLU:CG	9:A:373:HOH:O	2.52	0.53
2:Q:396:GLY:HA2	2:Q:401:LYS:HE3	1.91	0.52
2:Q:275:GLY:HA3	9:Q:653:HOH:O	2.11	0.50
2:Q:6:PRO:HG3	2:R:320:ASP:OD1	2.11	0.50
2:R:497:VAL:HG13	2:R:498:PRO:HD2	1.93	0.50
2:S:497:VAL:CG1	2:S:498:PRO:HD2	2.42	0.50
2:S:246:ASP:HB2	2:S:247:PRO:HD3	1.93	0.49
2:S:267:ILE:HB	2:S:268:PRO:HD3	1.94	0.49
2:R:49:PHE:HB2	2:R:370:TRP:CD2	2.48	0.49
2:S:270:LEU:HD21	2:S:426:ALA:HA	1.95	0.49
2:Q:25:GLU:HB2	2:Q:74:ILE:HD12	1.93	0.49
2:Q:403:VAL:HA	9:Q:736:HOH:O	2.12	0.49
2:Q:329:LEU:HD11	2:Q:471:LEU:HD11	1.95	0.49
9:C:390:HOH:O	2:S:20:PRO:CD	2.61	0.49
9:C:390:HOH:O	2:S:20:PRO:HD2	2.14	0.47
2:R:25:GLU:HB2	2:R:74:ILE:HD12	1.96	0.47
2:Q:245:LYS:O	2:Q:245:LYS:HE2	2.15	0.47
2:S:204:HIS:ND1	2:S:269:ASP:OD2	2.48	0.47
2:R:444:LYS:HG3	9:R:737:HOH:O	2.14	0.47
2:Q:454:ASN:H	2:Q:454:ASN:HD22	1.62	0.46
1:C:18:THR:HG22	1:C:18:THR:O	2.15	0.46
1:C:237:TRP:CH2	1:C:239:VAL:HB	2.51	0.46
2:Q:405:ASP:HB2	9:Q:735:HOH:O	2.15	0.46
1:A:145:PRO:HD2	1:A:179:TYR:CZ	2.51	0.46
2:Q:274:ALA:HB1	2:Q:422:LEU:HD11	1.98	0.46
1:A:96:GLU:HB2	9:A:400:HOH:O	2.16	0.45
2:S:43:TRP:CE2	2:S:365:LYS:HE2	2.51	0.45
2:Q:49:PHE:HB2	2:Q:370:TRP:CD2	2.52	0.45
1:B:237:TRP:CH2	1:B:239:VAL:HB	2.52	0.45
2:S:275:GLY:HA3	9:S:646:HOH:O	2.17	0.45
2:R:448:ASP:HB3	9:R:659:HOH:O	2.17	0.45
1:C:47:THR:O	2:S:23:ARG:HA	2.17	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:184:HIS:HB2	1:A:220:GLY:C	2.36	0.44
2:Q:497:VAL:HG12	2:Q:498:PRO:HD2	1.98	0.44
2:S:448:ASP:O	2:S:452:LYS:HE2	2.17	0.44
1:A:33:ASP:O	1:A:37:LEU:HG	2.18	0.44
2:Q:137:ASP:OD2	2:Q:139:ASN:HB2	2.17	0.44
2:Q:274:ALA:CB	2:Q:422:LEU:HD11	2.47	0.44
1:B:223:THR:OG1	1:B:247:GLY:HA2	2.18	0.44
2:R:72:CYS:HB3	2:R:75:CYS:SG	2.57	0.44
2:S:476:ARG:HD2	7:S:550:FCO:C2	2.48	0.44
1:C:145:PRO:HD2	1:C:179:TYR:CZ	2.53	0.43
1:B:112:GLY:HA2	1:B:149:PRO:HD3	2.01	0.43
1:A:18:THR:HG22	1:A:18:THR:O	2.19	0.43
2:R:296:ALA:HA	2:R:309:SER:HA	2.02	0.42
2:Q:189:LYS:HE2	9:C:299:HOH:O	2.19	0.42
1:A:47:THR:O	2:Q:23:ARG:HA	2.20	0.42
1:B:237:TRP:CZ2	1:B:239:VAL:HB	2.55	0.42
2:Q:296:ALA:HA	2:Q:309:SER:HA	2.00	0.42
2:S:521:GLY:O	9:S:762:HOH:O	2.21	0.42
1:B:18:THR:HG22	1:B:18:THR:O	2.20	0.41
1:A:112:GLY:HA2	1:A:149:PRO:HD3	2.01	0.41
2:R:497:VAL:HG12	2:R:498:PRO:HD2	2.01	0.41
1:B:228:PRO:HB3	1:B:237:TRP:CZ2	2.56	0.41
2:S:248:LEU:HD23	2:S:248:LEU:HA	1.79	0.41
1:B:93:PRO:HB2	1:B:96[B]:GLU:HG3	2.03	0.41
2:S:223:GLY:HA2	2:S:231:PHE:CD1	2.55	0.41
2:S:486:LYS:HB3	2:S:491:ASP:HB2	2.03	0.41
2:S:453:ASP:OD1	2:S:453:ASP:C	2.59	0.41
2:R:400:PHE:O	2:R:404:VAL:HG23	2.20	0.41
2:Q:513:SER:HB2	2:Q:514:PRO:HD2	2.02	0.41
2:Q:406[A]:MET:HG2	2:Q:410:LYS:HD2	2.03	0.41
1:B:145:PRO:HD2	1:B:179:TYR:CZ	2.56	0.40
2:Q:345:GLY:O	2:Q:348:LYS:HG2	2.21	0.40
1:B:11:TRP:HB3	1:B:45:GLN:HG2	2.03	0.40
2:S:333:TYR:OH	2:S:378:GLY:HA2	2.21	0.40
2:R:75:CYS:CB	7:R:550:FCO:C2	2.99	0.40
1:A:228:PRO:HB3	1:A:237:TRP:CZ2	2.56	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	260/264 (98%)	254 (98%)	6 (2%)	0	100	100
1	B	262/264 (99%)	252 (96%)	10 (4%)	0	100	100
1	C	259/264 (98%)	252 (97%)	6 (2%)	1 (0%)	39	42
2	Q	545/549 (99%)	532 (98%)	13 (2%)	0	100	100
2	R	543/549 (99%)	529 (97%)	14 (3%)	0	100	100
2	S	542/549 (99%)	527 (97%)	15 (3%)	0	100	100
All	All	2411/2439 (99%)	2346 (97%)	64 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	231	LEU

### 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	208/210 (99%)	207 (100%)	1 (0%)	92	96
1	B	210/210 (100%)	207 (99%)	3 (1%)	74	85
1	C	208/210 (99%)	207 (100%)	1 (0%)	92	96
2	Q	438/438 (100%)	432 (99%)	6 (1%)	74	85
2	R	435/438 (99%)	430 (99%)	5 (1%)	80	89
2	S	434/438 (99%)	429 (99%)	5 (1%)	78	88

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	1933/1944 (99%)	1912 (99%)	21 (1%)	80	89

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	67	LYS
2	Q	40	LYS
2	Q	245	LYS
2	Q	454	ASN
2	Q	473	ASP
2	Q	500	THR
2	Q	503	LEU
1	B	4	LYS
1	B	100	LYS
1	B	171	GLU
2	R	86	CYS
2	R	454	ASN
2	R	473	ASP
2	R	500	THR
2	R	503	LEU
1	C	95	ILE
2	S	365	LYS
2	S	452	LYS
2	S	454	ASN
2	S	473	ASP
2	S	503	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	ASN
1	A	61	HIS
2	Q	454	ASN
1	B	14	ASN
2	R	454	ASN
1	C	14	ASN
2	S	454	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues are 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	CSO	Q	543	3,2	3,6,7	0.70	0	1,6,8	1.91	0
2	CSO	R	543	3,2	3,6,7	0.74	0	1,6,8	1.22	0
2	CSO	S	543	3,2	3,6,7	0.85	0	1,6,8	1.66	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	CSO	Q	543	3,2	-	0/1/5/7	0/0/0/0
2	CSO	R	543	3,2	-	0/1/5/7	0/0/0/0
2	CSO	S	543	3,2	-	0/1/5/7	0/0/0/0

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.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 23 ligands modelled in this entry, 6 are monoatomic - leaving 17 for Mogul analysis.

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	$\# Z  > 2$
5	SF4	A	265	1	0,12,12	0.00	-	0,24,24	0.00	-
6	F3S	A	266	1	0,9,9	0.00	-	0,15,15	0.00	-
5	SF4	A	267	1	0,12,12	0.00	-	0,24,24	0.00	-
8	GOL	A	271	-	5,5,5	0.34	0	5,5,5	0.18	0
5	SF4	B	265	1	0,12,12	0.00	-	0,24,24	0.00	-
6	F3S	B	266	1	0,9,9	0.00	-	0,15,15	0.00	-
5	SF4	B	267	1	0,12,12	0.00	-	0,24,24	0.00	-
5	SF4	C	265	1	0,12,12	0.00	-	0,24,24	0.00	-
6	F3S	C	266	1	0,9,9	0.00	-	0,15,15	0.00	-
5	SF4	C	267	1	0,12,12	0.00	-	0,24,24	0.00	-
7	FCO	Q	550	2	0,6,6	0.00	-	0,6,6	0.00	-
8	GOL	Q	561	-	5,5,5	0.48	0	5,5,5	0.22	0
8	GOL	Q	562	-	5,5,5	0.43	0	5,5,5	0.53	0
7	FCO	R	550	2	0,6,6	0.00	-	0,6,6	0.00	-
8	GOL	R	563	-	5,5,5	0.38	0	5,5,5	0.37	0
7	FCO	S	550	2	0,6,6	0.00	-	0,6,6	0.00	-
8	GOL	S	561	-	5,5,5	0.31	0	5,5,5	0.44	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
5	SF4	A	265	1	-	0/0/48/48	0/6/5/5
6	F3S	A	266	1	-	0/0/24/24	0/0/3/3
5	SF4	A	267	1	-	0/0/48/48	0/6/5/5
8	GOL	A	271	-	-	0/4/4/4	0/0/0/0
5	SF4	B	265	1	-	0/0/48/48	0/6/5/5
6	F3S	B	266	1	-	0/0/24/24	0/0/3/3
5	SF4	B	267	1	-	0/0/48/48	0/6/5/5
5	SF4	C	265	1	-	0/0/48/48	0/6/5/5

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	F3S	C	266	1	-	0/0/24/24	0/0/3/3
5	SF4	C	267	1	-	0/0/48/48	0/6/5/5
7	FCO	Q	550	2	-	0/0/6/6	0/0/0/0
8	GOL	Q	561	-	-	0/4/4/4	0/0/0/0
8	GOL	Q	562	-	-	0/4/4/4	0/0/0/0
7	FCO	R	550	2	-	0/0/6/6	0/0/0/0
8	GOL	R	563	-	-	0/4/4/4	0/0/0/0
7	FCO	S	550	2	-	0/0/6/6	0/0/0/0
8	GOL	S	561	-	-	0/4/4/4	0/0/0/0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	R	550	FCO	1	0
7	S	550	FCO	1	0

## 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	A	262/264 (99%)	0.15	6 (2%) 64 63	13, 15, 17, 26	8 (3%)
1	B	262/264 (99%)	0.97	51 (19%) 1 1	13, 15, 17, 23	9 (3%)
1	C	260/264 (98%)	0.01	3 (1%) 81 80	13, 15, 17, 21	5 (1%)
2	Q	543/549 (98%)	0.07	4 (0%) 89 88	12, 14, 17, 23	12 (2%)
2	R	544/549 (99%)	0.35	26 (4%) 34 34	12, 15, 18, 22	13 (2%)
2	S	543/549 (98%)	0.17	14 (2%) 59 58	12, 15, 17, 22	15 (2%)
All	All	2414/2439 (98%)	0.25	104 (4%) 39 38	12, 15, 17, 26	62 (2%)

All (104) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	211	PHE	4.8
1	B	228	PRO	4.6
1	B	163	THR	4.5
1	B	168	ASP	4.3
2	R	453	ASP	4.3
1	B	3	ALA	4.2
2	R	451	ALA	4.2
2	S	157	ALA	4.0
2	R	249	ALA	3.9
1	B	213	LEU	3.9
1	B	191	PRO	3.9
1	B	140	LYS	3.8
1	C	5	HIS	3.8
1	B	183	VAL	3.8
1	B	264	GLY	3.8
1	B	188	PRO	3.8
1	B	190	LEU	3.8
2	R	447	ALA	3.7
1	B	180	GLY	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	S	6	PRO	3.6
2	R	246	ASP	3.5
1	B	164	LYS	3.5
2	R	247	PRO	3.5
2	R	322	GLY	3.5
1	B	169	LEU	3.4
1	B	197	GLU	3.4
1	B	165	GLY	3.4
1	B	184	HIS	3.4
1	B	195	ALA	3.4
2	R	449	SER	3.3
1	B	196	SER	3.3
1	B	67	LYS	3.3
1	A	3	ALA	3.2
2	S	398	PRO	3.2
2	R	6	PRO	3.1
2	S	173	SER	3.1
1	B	171	GLU	3.1
2	S	151	ALA	3.1
2	Q	6	PRO	3.0
2	R	455	THR	3.0
2	R	244	THR	2.9
1	B	68	ASP	2.9
1	B	167	PRO	2.9
1	A	4	LYS	2.9
2	R	245	LYS	2.9
1	B	198	PHE	2.8
2	R	5	LYS	2.8
1	B	237	TRP	2.8
1	B	5	HIS	2.7
2	R	297	THR	2.7
2	S	141	ALA	2.7
1	B	4	LYS	2.7
1	B	186	ASN	2.6
1	B	223	THR	2.6
1	B	203	ASP	2.6
1	B	71	TYR	2.6
2	S	9	GLN	2.6
1	B	193	PHE	2.6
1	B	137	LEU	2.5
2	S	7	THR	2.5
2	R	485	ILE	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	207	ALA	2.5
1	B	177	LEU	2.5
1	B	217	GLY	2.5
1	B	28	ILE	2.5
1	B	102	ALA	2.4
1	B	224	TYR	2.4
2	R	302	PRO	2.4
1	A	165	GLY	2.4
2	Q	148	ILE	2.4
1	B	209	LYS	2.4
2	S	246	ASP	2.3
1	B	214	TYR	2.3
2	S	323	LYS	2.3
2	S	137	ASP	2.3
1	B	200	PRO	2.3
1	A	67	LYS	2.3
1	B	187	CYS	2.3
2	R	412	SER	2.3
1	B	66	GLY	2.3
2	R	456	LEU	2.2
1	B	192	HIS	2.2
2	S	148	ILE	2.2
1	B	138	GLY	2.2
1	C	166	ILE	2.2
2	R	396	GLY	2.2
2	R	394	ALA	2.1
2	S	143	LYS	2.1
2	Q	343	PRO	2.1
1	B	204	SER	2.1
2	R	300	SER	2.1
1	A	164	LYS	2.1
2	R	401	LYS	2.1
1	B	166	ILE	2.1
1	C	165	GLY	2.1
2	R	440	GLU	2.1
2	R	486	LYS	2.1
1	B	146	GLY	2.0
2	S	169	ALA	2.0
1	B	194	GLU	2.0
1	A	218	CYS	2.0
2	Q	533	ILE	2.0
2	R	139	ASN	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	R	465	GLU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	CSO	R	543	7/8	0.96	0.11	-	14,14,14,16	0
2	CSO	Q	543	7/8	0.96	0.13	-	14,14,16,16	0
2	CSO	S	543	7/8	0.96	0.11	-	14,14,15,16	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	GOL	Q	562	6/6	0.89	0.18	2.67	13,16,16,16	0
8	GOL	A	271	6/6	0.87	0.19	1.88	18,20,20,21	0
8	GOL	R	563	6/6	0.86	0.19	0.25	16,18,19,19	0
8	GOL	S	561	6/6	0.91	0.14	0.22	13,14,14,15	0
7	FCO	S	550	7/7	0.99	0.12	-0.45	14,14,14,15	0
8	GOL	Q	561	6/6	0.95	0.11	-0.50	10,11,12,12	0
5	SF4	A	267	8/8	0.99	0.14	-0.78	12,13,14,15	0
5	SF4	A	265	8/8	0.97	0.12	-0.90	12,13,13,14	0
4	MG	Q	553	1/1	0.97	0.12	-1.00	13,13,13,13	0
4	MG	S	553	1/1	0.84	0.12	-1.09	13,13,13,13	0
5	SF4	C	265	8/8	0.98	0.10	-1.11	12,13,14,14	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	FCO	R	550	7/7	0.98	0.10	-1.16	13,13,14,14	0
7	FCO	Q	550	7/7	0.99	0.13	-1.41	13,14,15,15	0
5	SF4	C	267	8/8	0.98	0.11	-1.46	13,14,14,15	0
4	MG	R	553	1/1	0.79	0.09	-1.49	13,13,13,13	0
6	F3S	C	266	7/7	0.98	0.11	-1.50	13,14,14,15	0
6	F3S	B	266	7/7	0.90	0.11	-1.54	13,14,15,15	0
6	F3S	A	266	7/7	0.98	0.13	-1.62	13,14,14,14	0
5	SF4	B	265	8/8	0.88	0.11	-1.99	14,14,15,16	0
5	SF4	B	267	8/8	0.96	0.07	-2.57	13,14,15,15	0
3	NI	S	551	1/1	1.00	0.09	-	14,14,14,14	0
3	NI	Q	551	1/1	0.99	0.10	-	15,15,15,15	0
3	NI	R	551	1/1	0.99	0.04	-	15,15,15,15	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.