



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 10:33 AM BST

PDB ID : 2HOD  
Title : Crystal Structure of Fragment D from Human Fibrinogen Complexed with Gly-hydroxyPro-Arg-Pro-amide  
Authors : Doolittle, R.F.; Kollman, J.M.; Chen, A.; Pandi, L.  
Deposited on : 2006-07-14  
Resolution : 2.90 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

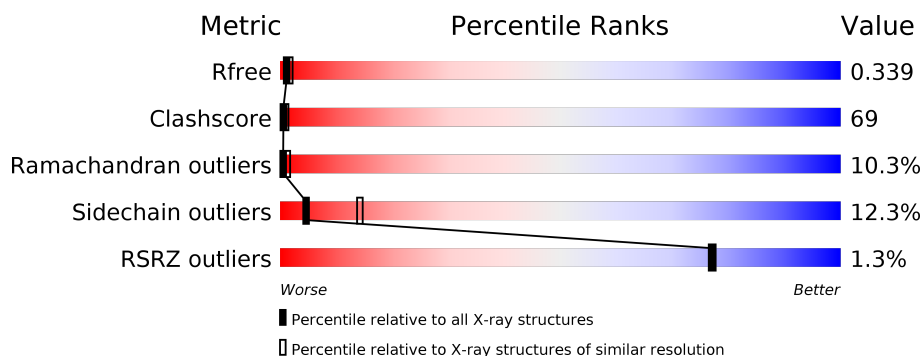
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	A	87	
1	D	87	
1	G	87	
1	J	87	
2	B	328	
2	E	328	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	328	
2	K	328	
3	C	323	
3	F	323	
3	I	323	
3	L	323	
4	M	5	
4	N	5	
4	O	5	
4	P	5	
4	Q	5	
4	R	5	
4	S	5	
4	T	5	

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
4	HYP	S	2	-	-	X	-
5	NAG	E	470	X	-	-	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22278 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 Fibrinogen alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	74	Total	C	N	O	S	0	0	0
			608	377	115	113	3			
1	D	71	Total	C	N	O	S	0	0	0
			584	361	112	108	3			
1	G	74	Total	C	N	O	S	0	0	0
			608	377	115	113	3			
1	J	79	Total	C	N	O	S	0	0	0
			652	402	126	121	3			

- Molecule 2 is a protein called Fibrinogen beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	307	Total	C	N	O	S	0	0	0
			2462	1535	433	472	22			
2	E	304	Total	C	N	O	S	0	0	0
			2434	1514	430	468	22			
2	H	307	Total	C	N	O	S	0	0	0
			2462	1535	433	472	22			
2	K	305	Total	C	N	O	S	0	0	0
			2442	1520	431	469	22			

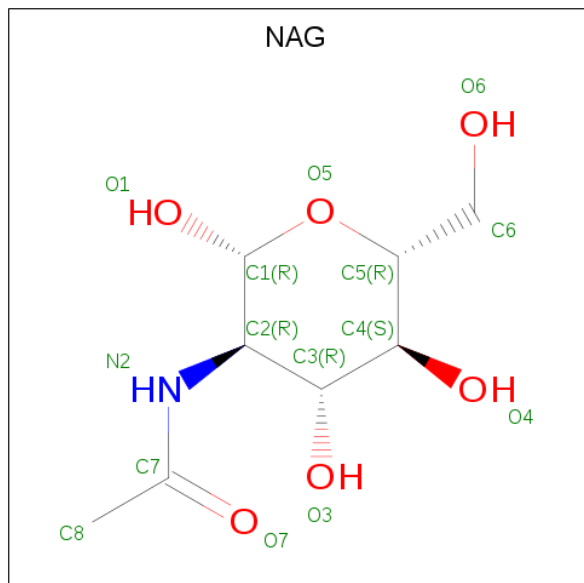
- Molecule 3 is a protein called Fibrinogen, gamma polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	305	Total	C	N	O	S	0	0	0
			2446	1552	410	472	12			
3	F	300	Total	C	N	O	S	0	0	0
			2410	1529	405	464	12			
3	I	305	Total	C	N	O	S	0	0	0
			2446	1552	410	472	12			
3	L	300	Total	C	N	O	S	0	0	0
			2410	1529	405	464	12			

- Molecule 4 is a protein called Gly-hydroxyPro-Arg-Pro-amide peptide ligand.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	M	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	N	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	O	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	P	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	Q	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	R	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	S	5	Total	C	N	O	0	0	1
			31	18	8	5			
4	T	5	Total	C	N	O	0	0	1
			31	18	8	5			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	H	1	Total	C	N	O	0	0
			14	8	1	5		
5	K	1	Total	C	N	O	0	0
			14	8	1	5		

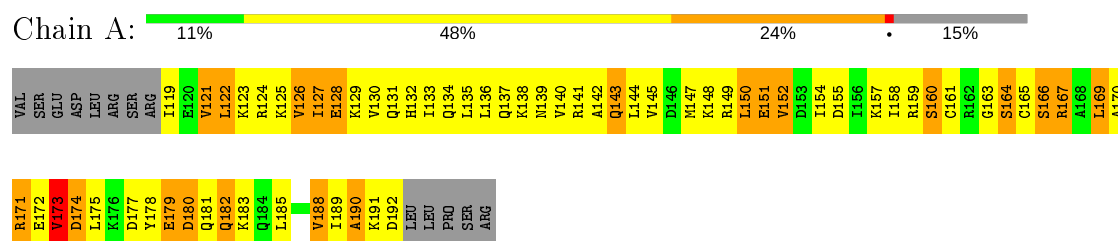
- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	K	2	Total	Ca	0	0
			2	2		
6	E	2	Total	Ca	0	0
			2	2		
6	H	1	Total	Ca	0	0
			1	1		
6	B	1	Total	Ca	0	0
			1	1		
6	I	1	Total	Ca	0	0
			1	1		
6	C	1	Total	Ca	0	0
			1	1		
6	L	1	Total	Ca	0	0
			1	1		
6	F	1	Total	Ca	0	0
			1	1		

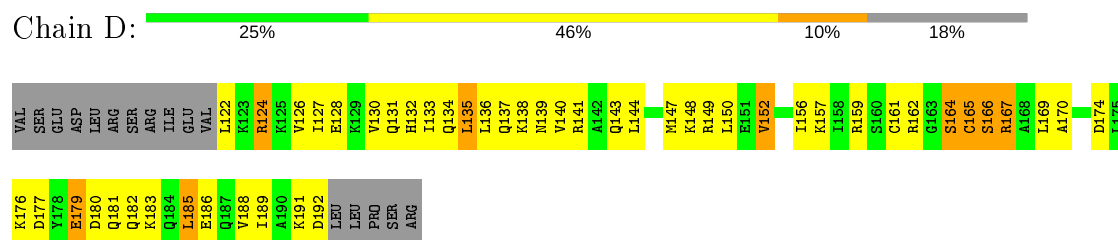
### 3 Residue-property plots

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 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.

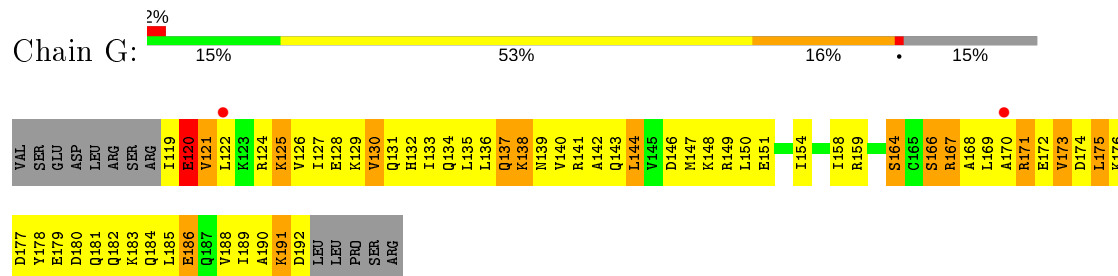
#### • Molecule 1: Fibrinogen alpha chain



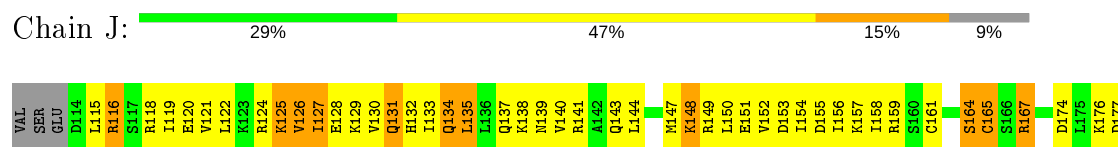
#### • Molecule 1: Fibrinogen alpha chain



#### • Molecule 1: Fibrinogen alpha chain

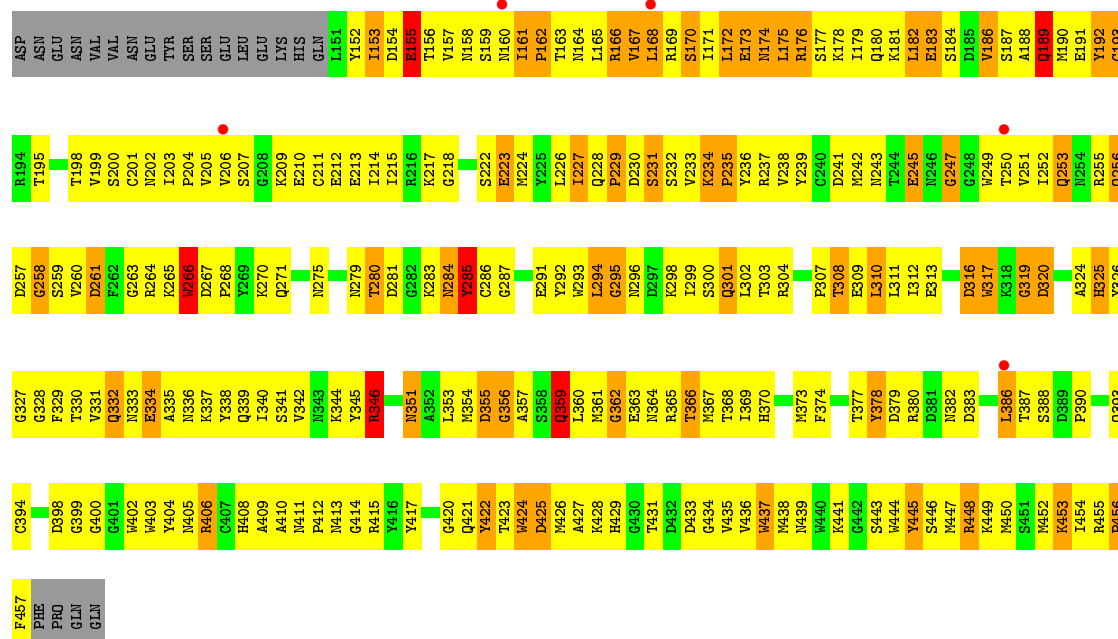


#### • Molecule 1: Fibrinogen alpha chain

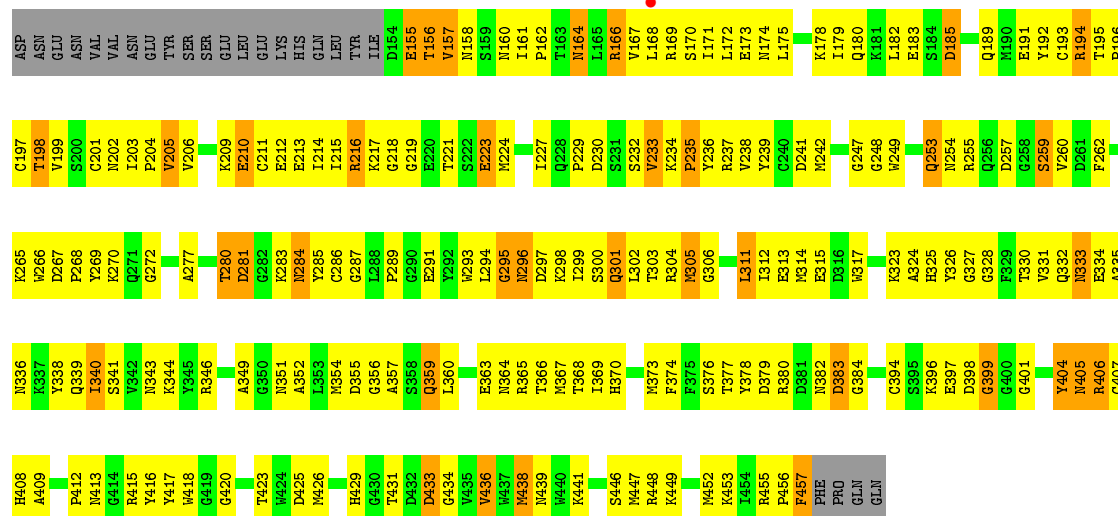




- Molecule 2: Fibrinogen beta chain



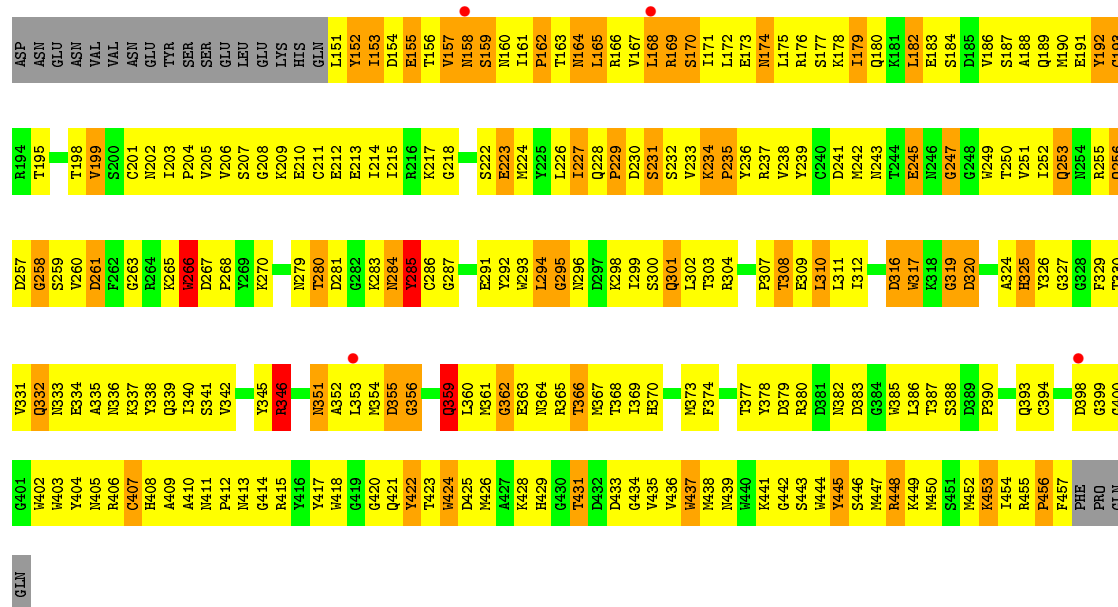
- Molecule 2: Fibrinogen beta chain



- Molecule 2: Fibrinogen beta chain

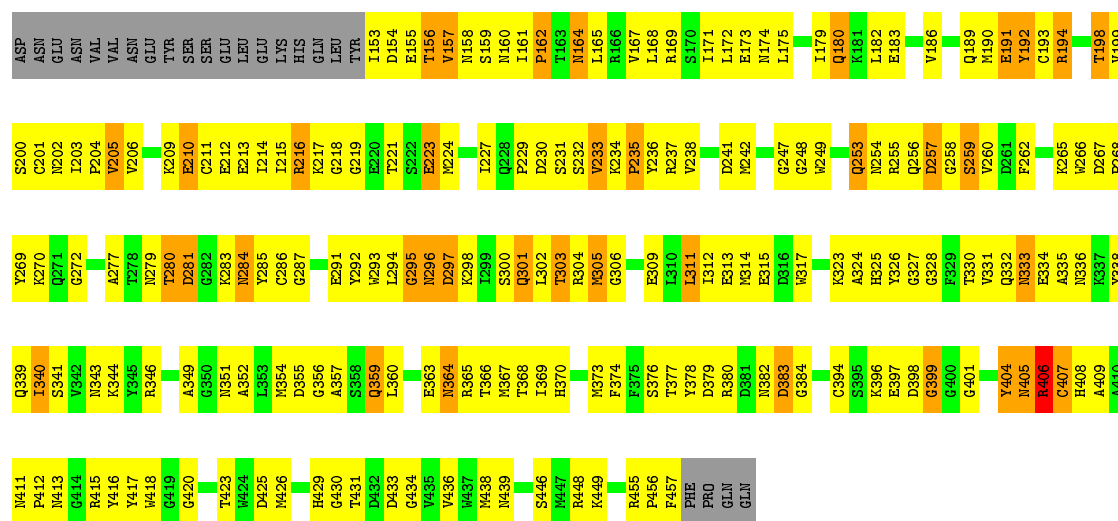






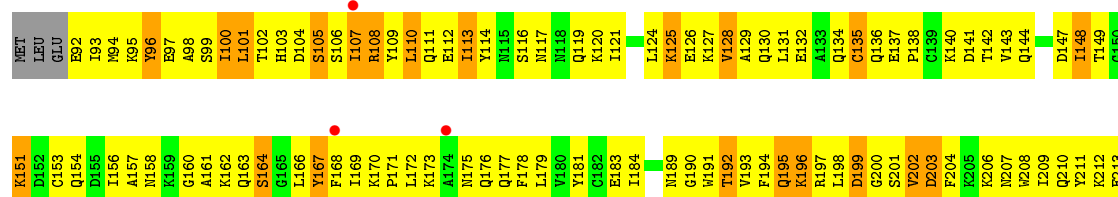
• Molecule 2: Fibrinogen beta chain

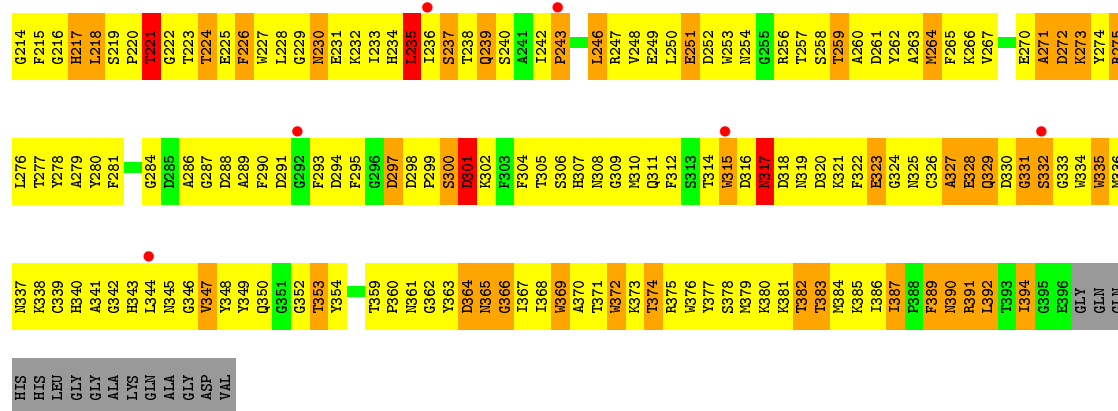
Chain K: 32% 50% 11% 7%



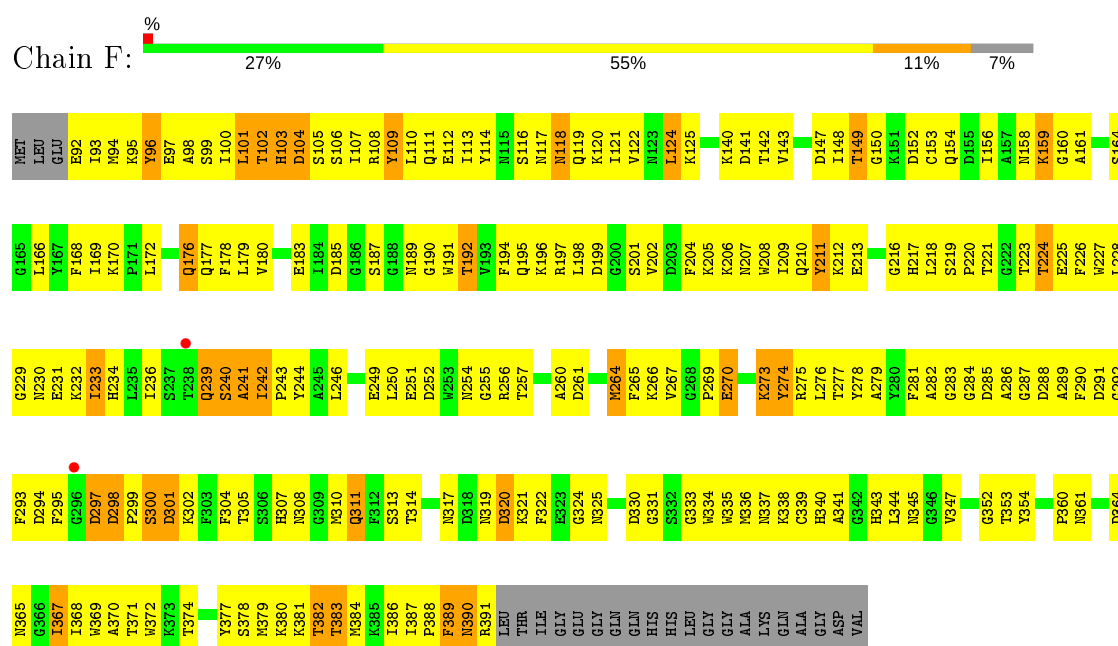
• Molecule 3: Fibrinogen, gamma polypeptide

Chain C: 3% 13% 60% 20% 6%

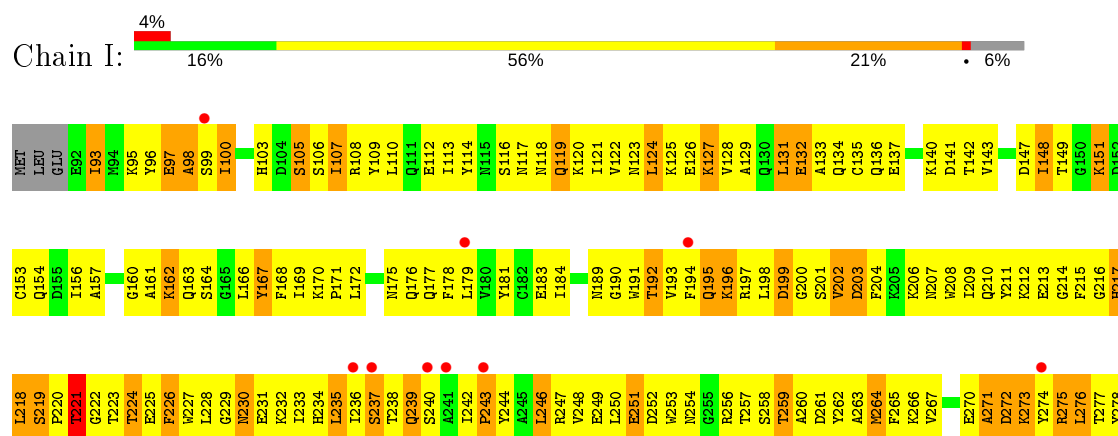


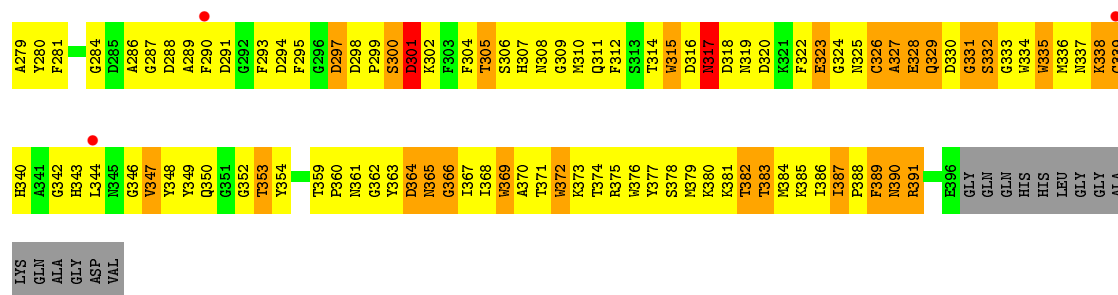


• Molecule 3: Fibrinogen, gamma polypeptide

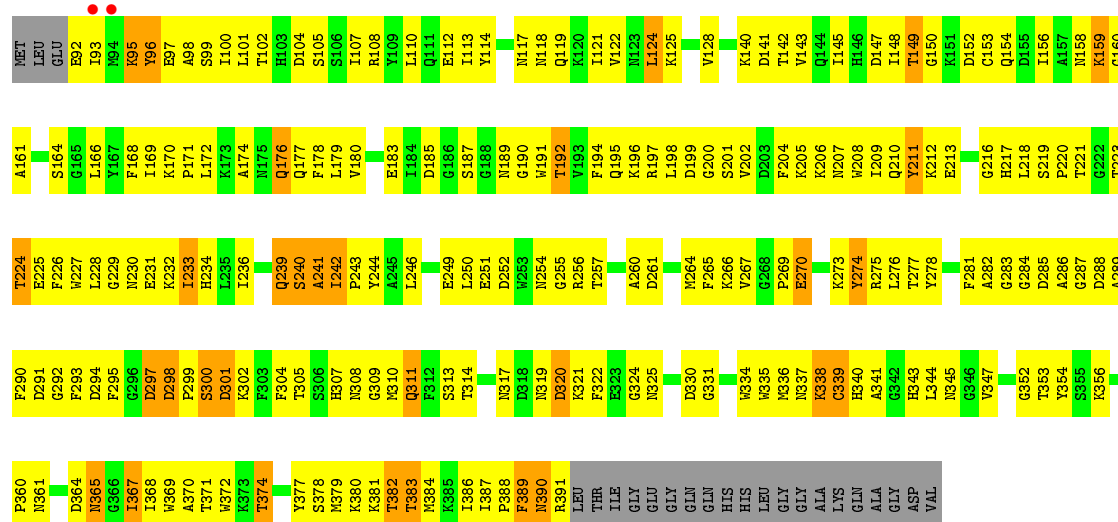


• Molecule 3: Fibrinogen, gamma polypeptide

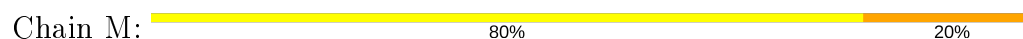




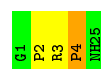
- Molecule 3: Fibrinogen, gamma polypeptide



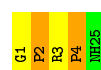
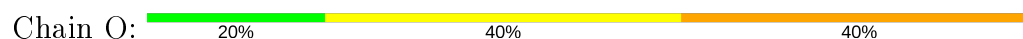
- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand



- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand



- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand

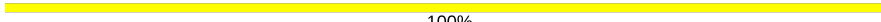


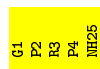
- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand

Chain P:  40% 60%



- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand

Chain Q:  100%



- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand

Chain R:  20% 60% 20%



- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand

Chain S:  40% 40% 20%



- Molecule 4: Gly-hydroxyPro-Arg-Pro-amide peptide ligand

Chain T:  40% 60%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.65Å 47.07Å 431.46Å 90.00° 90.06° 90.00°	Depositor
Resolution (Å)	30.00 – 2.90 49.22 – 2.79	Depositor EDS
% Data completeness (in resolution range)	89.5 (30.00-2.90) 85.0 (49.22-2.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.22 (at 2.77Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.268 , 0.347 0.261 , 0.339	Depositor DCC
$R_{free}$ test set	3441 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	75.8	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.21 , 22.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.477 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	22278	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: HYP, CA, NAG, NH2

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	0.44	0/609	0.76	0/811
1	D	0.47	0/585	0.77	0/778
1	G	0.40	0/609	0.73	0/811
1	J	0.46	0/653	0.76	0/869
2	B	0.44	0/2523	0.77	3/3409 (0.1%)
2	E	0.51	0/2494	0.76	1/3369 (0.0%)
2	H	0.44	0/2523	0.78	3/3409 (0.1%)
2	K	0.52	0/2502	0.79	3/3380 (0.1%)
3	C	0.42	0/2512	0.70	0/3396
3	F	0.45	0/2476	0.72	1/3347 (0.0%)
3	I	0.43	0/2512	0.73	2/3396 (0.1%)
3	L	0.44	0/2476	0.74	3/3347 (0.1%)
4	M	0.66	0/21	0.75	0/25
4	N	0.84	0/21	0.71	0/25
4	O	0.71	0/21	0.56	0/25
4	P	0.51	0/21	0.65	0/25
4	Q	0.58	0/21	0.71	0/25
4	R	0.79	0/21	1.04	0/25
4	S	0.62	0/21	0.53	0/25
4	T	0.62	0/21	0.64	0/25
All	All	0.46	0/22642	0.75	16/30522 (0.1%)

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
2	E	0	1
2	H	0	1
2	K	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	3

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	406	ARG	C-N-CA	6.48	137.89	121.70
2	H	399	GLY	N-CA-C	6.20	128.59	113.10
2	B	399	GLY	N-CA-C	6.07	128.26	113.10
2	H	437	TRP	N-CA-C	-6.05	94.67	111.00
2	B	437	TRP	N-CA-C	-5.72	95.56	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	236	TYR	Sidechain
2	H	152	TYR	Sidechain
2	K	236	TYR	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	608	0	649	117	0
1	D	584	0	623	74	0
1	G	608	0	649	118	0
1	J	652	0	695	94	0
2	B	2462	0	2327	468	0
2	E	2434	0	2295	248	0
2	H	2462	0	2326	410	0
2	K	2442	0	2307	250	0
3	C	2446	0	2294	480	0
3	F	2410	0	2256	270	0
3	I	2446	0	2294	451	0
3	L	2410	0	2256	256	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	M	31	0	32	3	0
4	N	31	0	32	7	0
4	O	31	0	32	10	0
4	P	31	0	32	5	0
4	Q	31	0	32	5	0
4	R	31	0	32	4	0
4	S	31	0	32	12	0
4	T	31	0	32	5	0
5	B	14	0	13	4	0
5	E	14	0	13	5	0
5	H	14	0	13	3	0
5	K	14	0	13	2	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	E	2	0	0	0	0
6	F	1	0	0	0	0
6	H	1	0	0	0	0
6	I	1	0	0	0	0
6	K	2	0	0	0	0
6	L	1	0	0	0	0
All	All	22278	0	21279	3018	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 69.

The worst 5 of 3018 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)
3:F:305:THR:HB	3:F:341:ALA:HB2	1.23	1.17
2:E:355:ASP:HB3	2:E:365:ARG:HH22	1.12	1.15
2:H:245:GLU:HB3	2:H:249:TRP:HE1	1.09	1.15
1:J:127:ILE:HG23	2:K:153:ILE:HG21	1.27	1.15
2:E:191:GLU:HA	2:E:194:ARG:HH21	1.09	1.13

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	72/87 (83%)	43 (60%)	17 (24%)	12 (17%)	0	0
1	D	69/87 (79%)	47 (68%)	19 (28%)	3 (4%)	2	10
1	G	72/87 (83%)	43 (60%)	19 (26%)	10 (14%)	0	0
1	J	77/87 (88%)	58 (75%)	15 (20%)	4 (5%)	2	6
2	B	305/328 (93%)	195 (64%)	70 (23%)	40 (13%)	0	0
2	E	302/328 (92%)	229 (76%)	53 (18%)	20 (7%)	1	3
2	H	305/328 (93%)	191 (63%)	70 (23%)	44 (14%)	0	0
2	K	303/328 (92%)	230 (76%)	47 (16%)	26 (9%)	1	2
3	C	303/323 (94%)	168 (55%)	96 (32%)	39 (13%)	0	0
3	F	298/323 (92%)	210 (70%)	70 (24%)	18 (6%)	1	4
3	I	303/323 (94%)	175 (58%)	85 (28%)	43 (14%)	0	0
3	L	298/323 (92%)	215 (72%)	66 (22%)	17 (6%)	1	5
4	M	2/5 (40%)	1 (50%)	0	1 (50%)	0	0
4	N	2/5 (40%)	1 (50%)	0	1 (50%)	0	0
4	O	2/5 (40%)	0	1 (50%)	1 (50%)	0	0
4	P	2/5 (40%)	1 (50%)	1 (50%)	0	100	100
4	Q	2/5 (40%)	1 (50%)	0	1 (50%)	0	0
4	R	2/5 (40%)	1 (50%)	0	1 (50%)	0	0
4	S	2/5 (40%)	2 (100%)	0	0	100	100
4	T	2/5 (40%)	2 (100%)	0	0	100	100
All	All	2723/2992 (91%)	1813 (67%)	629 (23%)	281 (10%)	0	1

5 of 281 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	121	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	122	LEU
1	A	160	SER
1	A	166	SER
2	B	170	SER

### 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	69/82 (84%)	56 (81%)	13 (19%)	1	4
1	D	66/82 (80%)	57 (86%)	9 (14%)	3	11
1	G	69/82 (84%)	61 (88%)	8 (12%)	5	16
1	J	74/82 (90%)	62 (84%)	12 (16%)	2	7
2	B	265/286 (93%)	231 (87%)	34 (13%)	4	13
2	E	262/286 (92%)	238 (91%)	24 (9%)	9	27
2	H	265/286 (93%)	238 (90%)	27 (10%)	7	22
2	K	263/286 (92%)	238 (90%)	25 (10%)	8	26
3	C	257/269 (96%)	216 (84%)	41 (16%)	2	7
3	F	253/269 (94%)	223 (88%)	30 (12%)	5	15
3	I	257/269 (96%)	216 (84%)	41 (16%)	2	7
3	L	253/269 (94%)	225 (89%)	28 (11%)	6	19
4	M	2/2 (100%)	2 (100%)	0	100	100
4	N	2/2 (100%)	2 (100%)	0	100	100
4	O	2/2 (100%)	2 (100%)	0	100	100
4	P	2/2 (100%)	2 (100%)	0	100	100
4	Q	2/2 (100%)	2 (100%)	0	100	100
4	R	2/2 (100%)	2 (100%)	0	100	100
4	S	2/2 (100%)	2 (100%)	0	100	100
4	T	2/2 (100%)	2 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2369/2564 (92%)	2077 (88%)	292 (12%)	<b>4</b> <b>14</b>

5 of 292 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	180	VAL
2	H	169	ARG
3	L	147	ASP
3	F	211	TYR
3	F	382	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 165 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	239	GLN
2	H	284	ASN
3	L	136	GLN
3	F	308	ASN
2	H	164	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 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
4	HYP	M	2	4	6,8,9	1.10	1 (16%)	5,10,12	1.35	1 (20%)
4	HYP	N	2	4	6,8,9	0.78	0	5,10,12	1.73	1 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	HYP	O	2	4	6,8,9	0.64	0	5,10,12	1.96	2 (40%)
4	HYP	P	2	4	6,8,9	0.92	0	5,10,12	1.10	0
4	HYP	Q	2	4	6,8,9	0.97	0	5,10,12	1.39	1 (20%)
4	HYP	R	2	4	6,8,9	0.83	0	5,10,12	2.14	1 (20%)
4	HYP	S	2	4	6,8,9	0.98	0	5,10,12	1.23	1 (20%)
4	HYP	T	2	4	6,8,9	0.74	0	5,10,12	1.09	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
4	HYP	M	2	4	-	0/0/11/13	0/1/1/1
4	HYP	N	2	4	-	0/0/11/13	0/1/1/1
4	HYP	O	2	4	-	0/0/11/13	0/1/1/1
4	HYP	P	2	4	-	0/0/11/13	0/1/1/1
4	HYP	Q	2	4	-	0/0/11/13	0/1/1/1
4	HYP	R	2	4	-	0/0/11/13	0/1/1/1
4	HYP	S	2	4	-	0/0/11/13	0/1/1/1
4	HYP	T	2	4	-	0/0/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	M	2	HYP	CB-CG	-2.21	1.48	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	R	2	HYP	CG-CB-CA	-3.77	99.20	103.96
4	N	2	HYP	CG-CB-CA	-3.21	99.91	103.96
4	O	2	HYP	CG-CB-CA	-3.06	100.10	103.96
4	M	2	HYP	CG-CB-CA	-2.41	100.91	103.96
4	S	2	HYP	CG-CB-CA	-2.36	100.98	103.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	O	2	HYP	2	0
4	P	2	HYP	2	0
4	S	2	HYP	4	0
4	T	2	HYP	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 10 are monoatomic - leaving 4 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	NAG	K	470	2	14,14,15	0.64	0	17,19,21	0.93	1 (5%)
5	NAG	H	470	2	14,14,15	0.77	0	17,19,21	0.83	1 (5%)
5	NAG	E	470	2	14,14,15	0.75	1 (7%)	17,19,21	0.81	0
5	NAG	B	470	2	14,14,15	0.64	0	17,19,21	0.86	1 (5%)

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	NAG	K	470	2	-	3/6/23/26	0/1/1/1
5	NAG	H	470	2	-	5/6/23/26	0/1/1/1
5	NAG	E	470	2	1/1/5/7	5/6/23/26	0/1/1/1
5	NAG	B	470	2	-	3/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	470	NAG	C1-C2	2.13	1.55	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	K	470	NAG	C2-N2-C7	-2.78	118.94	122.90
5	B	470	NAG	C2-N2-C7	-2.60	119.20	122.90
5	H	470	NAG	C2-N2-C7	-2.05	119.98	122.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	E	470	NAG	C1

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	K	470	NAG	C8-C7-N2-C2
5	K	470	NAG	O7-C7-N2-C2
5	H	470	NAG	C3-C2-N2-C7
5	H	470	NAG	C8-C7-N2-C2
5	H	470	NAG	O7-C7-N2-C2

There are no ring outliers.

4 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	K	470	NAG	2	0
5	H	470	NAG	3	0
5	E	470	NAG	5	0
5	B	470	NAG	4	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	74/87 (85%)	-0.19	0 100 100	61, 81, 96, 116	0
1	D	71/87 (81%)	-0.37	0 100 100	36, 59, 88, 90	0
1	G	74/87 (85%)	-0.18	2 (2%) 54 50	56, 82, 101, 110	0
1	J	79/87 (90%)	-0.40	0 100 100	37, 60, 102, 108	0
2	B	307/328 (93%)	-0.25	5 (1%) 72 71	49, 78, 102, 115	0
2	E	304/328 (92%)	-0.41	1 (0%) 94 94	30, 52, 82, 104	0
2	H	307/328 (93%)	-0.23	4 (1%) 77 77	49, 78, 101, 115	0
2	K	305/328 (92%)	-0.47	0 100 100	29, 52, 83, 105	0
3	C	305/323 (94%)	-0.10	9 (2%) 50 45	49, 85, 110, 121	0
3	F	300/323 (92%)	-0.34	2 (0%) 87 87	35, 70, 93, 123	0
3	I	305/323 (94%)	-0.03	12 (3%) 39 35	47, 85, 108, 123	0
3	L	300/323 (92%)	-0.33	2 (0%) 87 87	38, 70, 92, 113	0
4	M	3/5 (60%)	0.09	0 100 100	70, 70, 73, 91	0
4	N	3/5 (60%)	-0.33	0 100 100	57, 57, 66, 83	0
4	O	3/5 (60%)	0.56	0 100 100	94, 94, 110, 115	0
4	P	3/5 (60%)	-0.27	0 100 100	67, 67, 70, 99	0
4	Q	3/5 (60%)	0.12	0 100 100	66, 66, 67, 87	0
4	R	3/5 (60%)	-0.04	0 100 100	51, 51, 63, 77	0
4	S	3/5 (60%)	-0.26	0 100 100	89, 89, 98, 123	0
4	T	3/5 (60%)	0.05	0 100 100	65, 65, 69, 91	0
All	All	2755/2992 (92%)	-0.27	37 (1%) 77 77	29, 73, 102, 123	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	168	LEU	5.5
3	C	236	ILE	5.4
2	B	206	VAL	5.2
3	C	344	LEU	4.8
3	I	243	PRO	4.1

## 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
4	HYP	O	2	8/9	0.68	0.20	97,105,109,110	0
4	HYP	S	2	8/9	0.82	0.25	90,94,98,103	0
4	HYP	T	2	8/9	0.93	0.16	73,75,76,79	0
4	HYP	P	2	8/9	0.94	0.14	72,76,81,85	0
4	HYP	Q	2	8/9	0.94	0.14	70,72,73,76	0
4	HYP	R	2	8/9	0.95	0.15	43,49,56,59	0
4	HYP	M	2	8/9	0.96	0.14	70,73,74,74	0
4	HYP	N	2	8/9	0.96	0.13	49,56,63,64	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
6	CA	E	3	1/1	0.72	0.12	88,88,88,88	0
6	CA	H	2	1/1	0.78	0.17	87,87,87,87	0
5	NAG	K	470	14/15	0.80	0.22	65,89,100,103	0
5	NAG	B	470	14/15	0.82	0.15	85,88,92,100	0
6	CA	K	3	1/1	0.84	0.12	94,94,94,94	0
5	NAG	H	470	14/15	0.90	0.19	82,95,103,105	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	CA	L	1	1/1	0.92	0.09	62,62,62,62	0
5	NAG	E	470	14/15	0.92	0.14	66,77,88,93	0
6	CA	B	2	1/1	0.94	0.09	86,86,86,86	0
6	CA	F	1	1/1	0.94	0.09	64,64,64,64	0
6	CA	E	2	1/1	0.94	0.26	60,60,60,60	0
6	CA	C	1	1/1	0.95	0.14	70,70,70,70	0
6	CA	K	2	1/1	0.96	0.11	60,60,60,60	0
6	CA	I	1	1/1	0.96	0.05	69,69,69,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.