wwPDB X-ray Structure Validation Summary Report (i)

Aug 9, 2020 – 06:18 AM BST

PDB ID	:	1M1J
Title	:	Crystal structure of native chicken fibrinogen with two different bound ligands
Authors	:	Yang, Z.; Kollman, J.M.; Pandi, L.; Doolittle, R.F.
Deposited on	:	2002-06-19
Resolution	:	2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

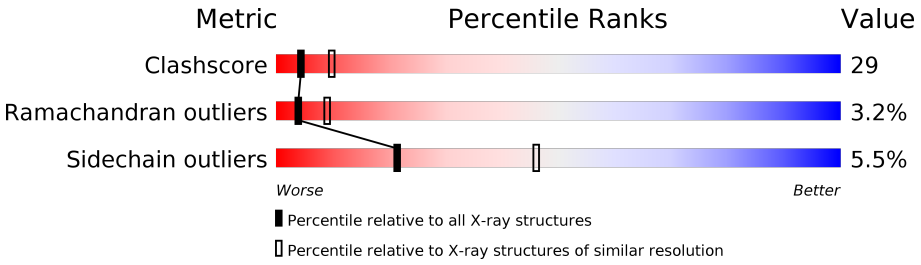
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)

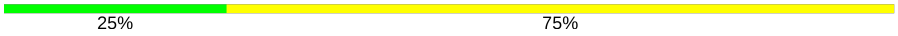
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	491	20% 16% • 61%
1	D	491	18% 19% • 60%
2	B	464	53% 31% • 13%
2	E	464	50% 33% • 14%
3	C	409	56% 32% 6% 5%
3	F	409	54% 36% 5% 5%
4	G	4	75% 25%
4	H	4	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	I	4	
5	J	4	

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
6	NDG	B	470	-	-	X	-
6	NDG	C	420	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 16117 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 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total	C	N	O	S	0	0	0
			1544	947	282	305	10			
1	D	194	Total	C	N	O	S	0	0	0
			1565	962	286	307	10			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	GLY	CYS	SEE REMARK 999	UNP P14448
D	49	GLY	CYS	SEE REMARK 999	UNP P14448

- Molecule 2 is a protein called Fibrinogen beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	402	Total	C	N	O	S	0	0	0
			3225	2023	554	623	25			
2	E	401	Total	C	N	O	S	0	0	0
			3216	2019	553	619	25			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLN	-	SEE REMARK 999	UNP Q02020
E	1	GLN	-	SEE REMARK 999	UNP Q02020

- Molecule 3 is a protein called Fibrinogen gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	390	Total	C	N	O	S	0	0	0
			3162	1987	539	620	16			
3	F	389	Total	C	N	O	S	0	0	0
			3155	1983	538	618	16			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	286	ALA	ARG	SEE REMARK 999	UNP O93568
F	286	ALA	ARG	SEE REMARK 999	UNP O93568

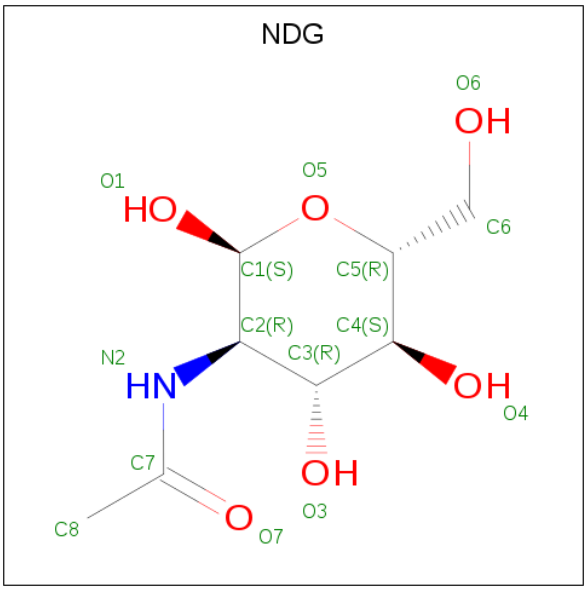
- Molecule 4 is a protein called GLY-PRO-ARG-PRO peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	G	4	Total	C	N	O	0	0	0
			30	18	7	5			
4	H	4	Total	C	N	O	0	0	0
			30	18	7	5			

- Molecule 5 is a protein called GLY-HIS-ARG-PRO peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	4	Total	C	N	O	0	0	0
			33	19	9	5			
5	J	4	Total	C	N	O	0	0	0
			33	19	9	5			

- Molecule 6 is 2-acetamido-2-deoxy-alpha-D-glucopyranose (three-letter code: NDG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	O	0	0
			15	8	1	6		

Continued on next page...

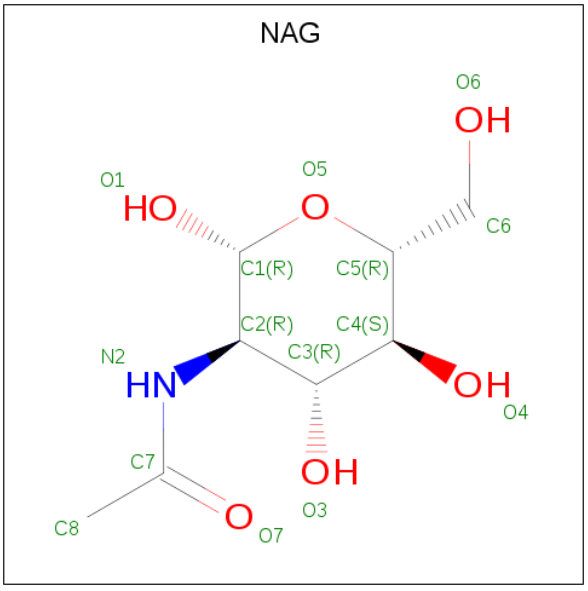
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total C N O 15 8 1 6	0	0
6	E	1	Total C N O 15 8 1 6	0	0
6	F	1	Total C N O 15 8 1 6	0	0
6	J	1	Total C N O 15 8 1 6	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total Ca 1 1	0	0
7	C	1	Total Ca 1 1	0	0
7	F	1	Total Ca 1 1	0	0
7	E	1	Total Ca 1 1	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	1	Total C N O 15 8 1 6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	F	1	Total	C	N	O	0	0
			15	8	1	6		
8	I	1	Total	C	N	O	0	0
			15	8	1	6		

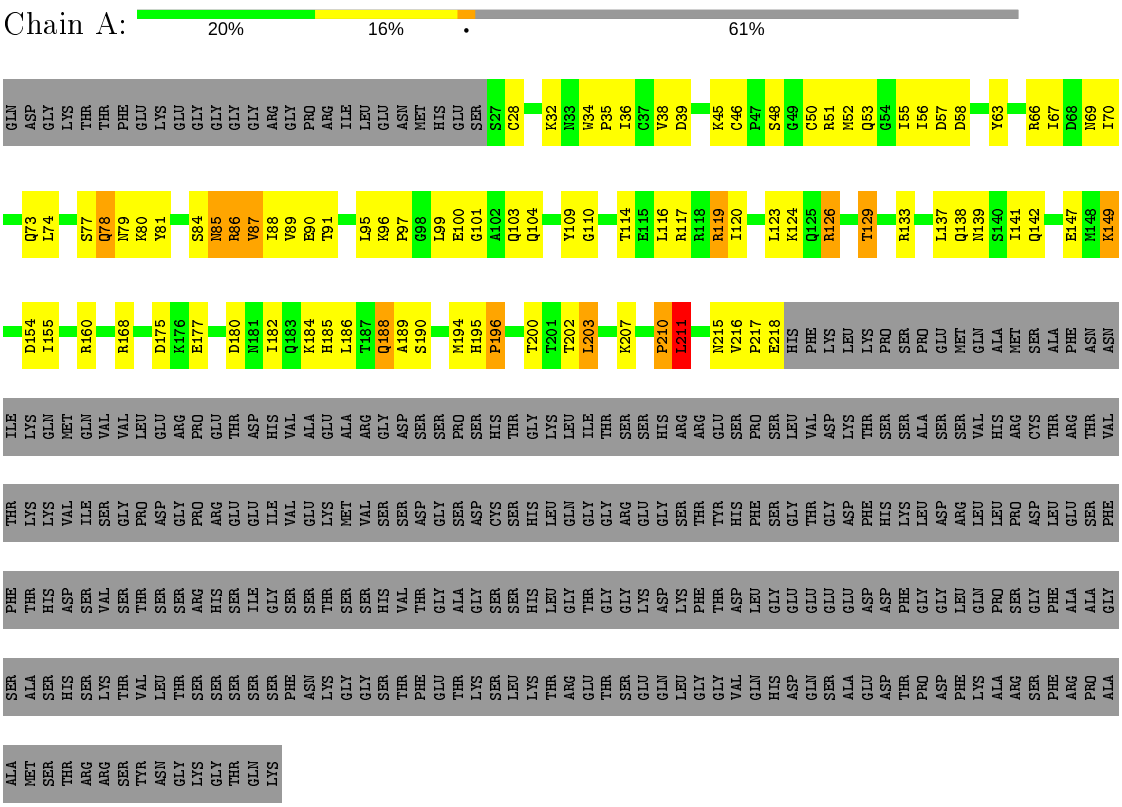
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

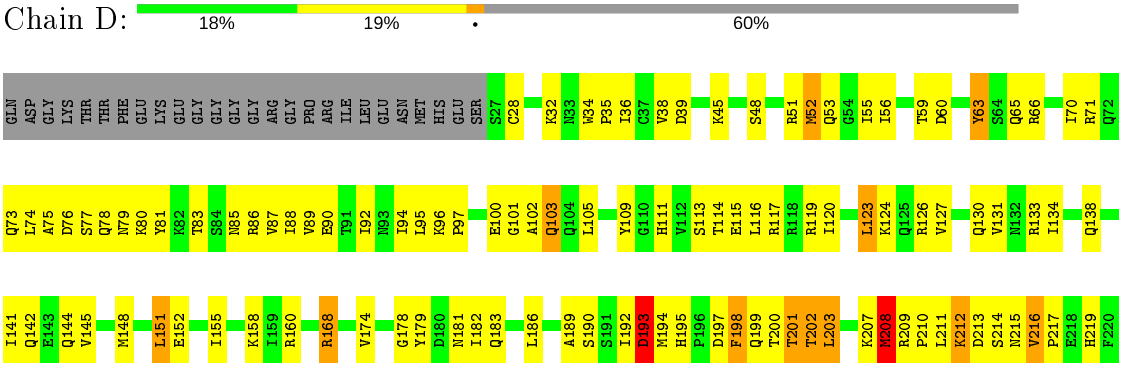
- Molecule 1: Fibrinogen alpha subunit

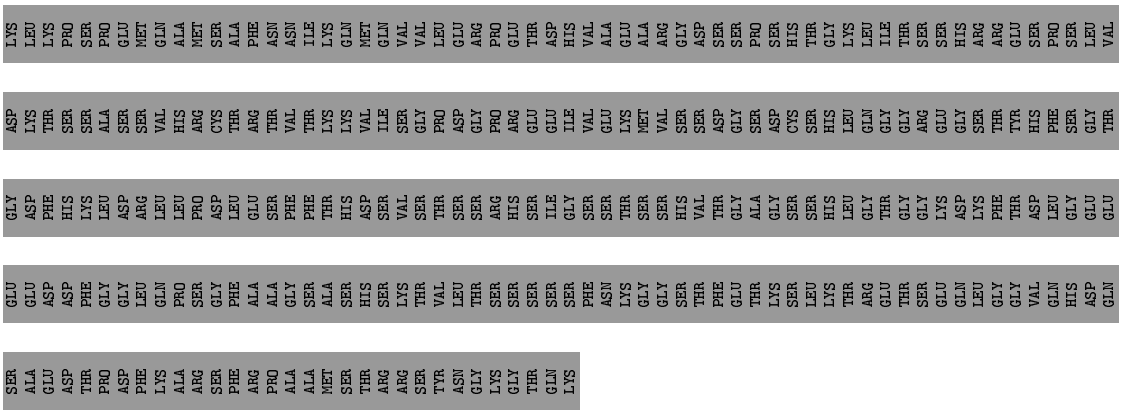
Chain A:



- Molecule 1: Fibrinogen alpha subunit

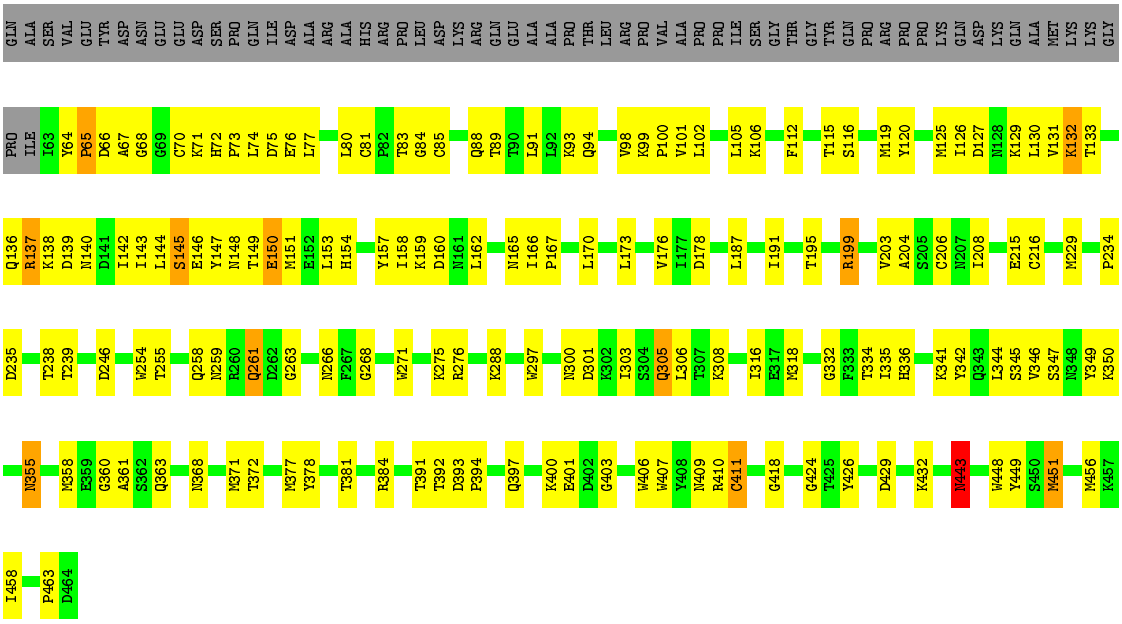
Chain D:





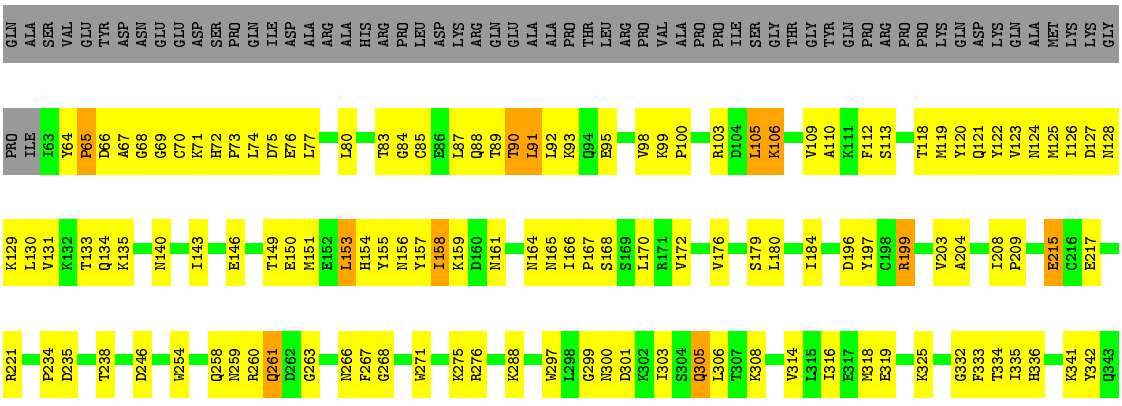
- Molecule 2: Fibrinogen beta chain

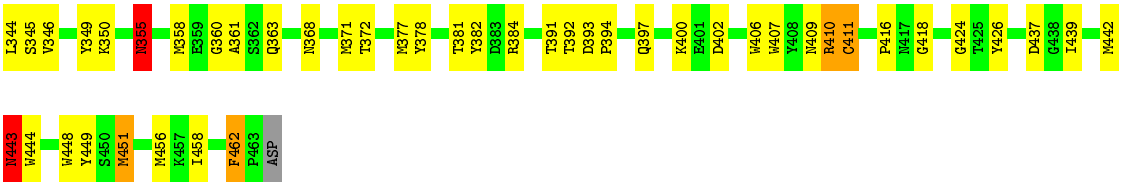
Chain B:



- Molecule 2: Fibrinogen beta chain

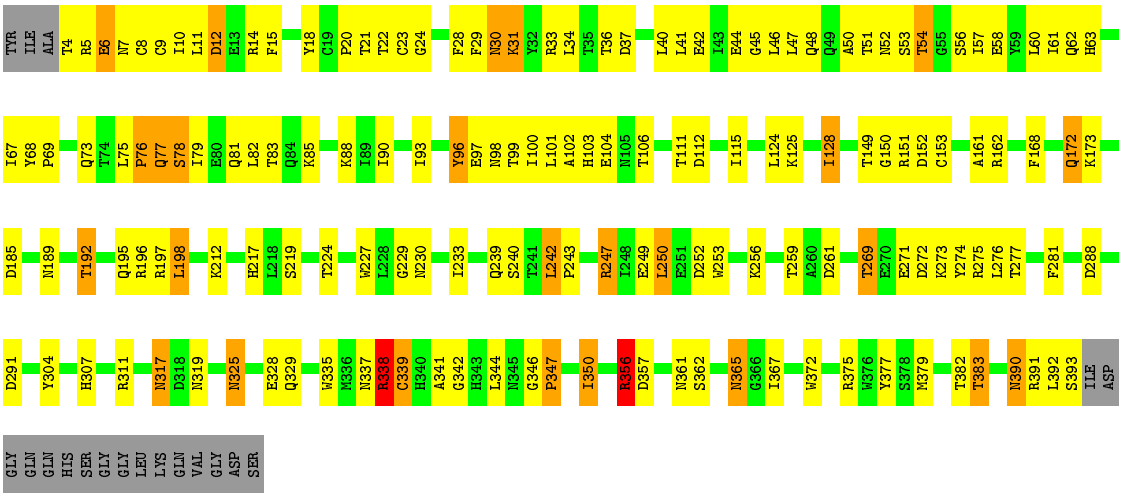
Chain E:





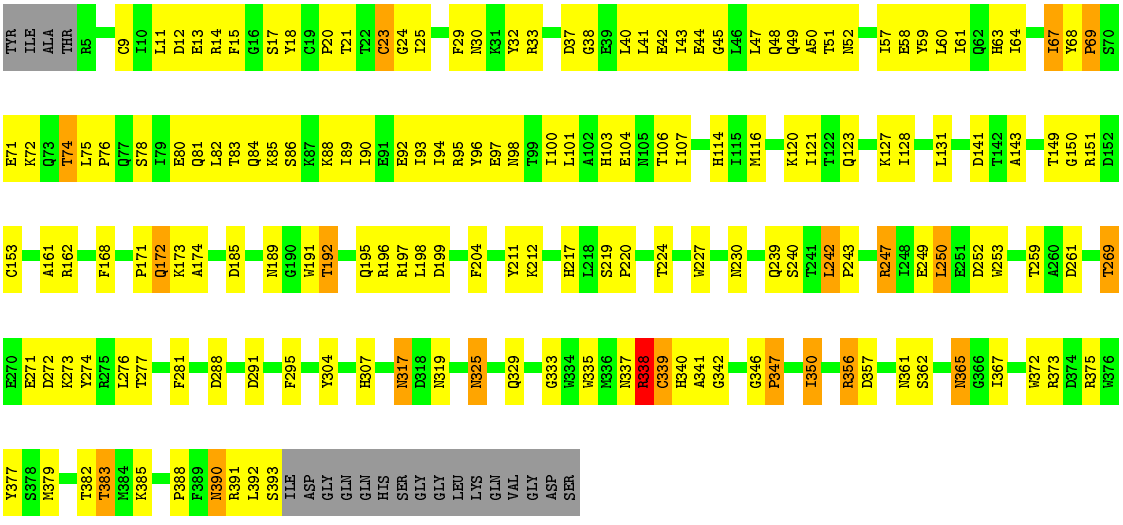
- Molecule 3: Fibrinogen gamma chain

Chain C:



- Molecule 3: Fibrinogen gamma chain

Chain F:



- Molecule 4: GLY-PRO-ARG-PRO peptide

Chain G:

- Molecule 4: GLY-PRO-ARG-PRO peptide

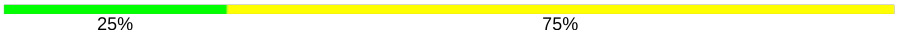
Chain H:

- Molecule 5: GLY-HIS-ARG-PRO peptide

Chain I:

- Molecule 5: GLY-HIS-ARG-PRO peptide

Chain J:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	114.09Å 100.02Å 200.09Å 90.00° 105.79° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70	Depositor
% Data completeness (in resolution range)	93.1 (20.00-2.70)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.227 , 0.256	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16117	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	0.33	0/1564	0.57	0/2108
1	D	0.30	0/1587	0.58	0/2139
2	B	0.39	0/3304	0.63	1/4467 (0.0%)
2	E	0.34	0/3295	0.61	1/4456 (0.0%)
3	C	0.42	0/3236	0.65	1/4374 (0.0%)
3	F	0.39	0/3229	0.65	1/4364 (0.0%)
4	G	0.62	0/31	0.80	0/40
4	H	0.60	0/31	0.69	0/40
5	I	0.55	0/34	0.69	0/43
5	J	0.46	0/34	0.52	0/43
All	All	0.37	0/16345	0.63	4/22074 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	338	ARG	N-CA-C	-6.79	92.66	111.00
3	C	338	ARG	N-CA-C	-6.79	92.67	111.00
2	B	403	GLY	N-CA-C	5.09	125.82	113.10
2	E	410	ARG	N-CA-C	-5.09	97.26	111.00

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	1544	0	1532	133	0
1	D	1565	0	1548	161	0
2	B	3225	0	3081	179	0
2	E	3216	0	3077	233	0
3	C	3162	0	2992	189	0
3	F	3155	0	2985	190	0
4	G	30	0	32	1	0
4	H	30	0	32	2	0
5	I	33	0	32	1	0
5	J	33	0	32	4	0
6	B	15	0	12	8	0
6	C	15	0	12	10	0
6	E	15	0	12	4	0
6	F	15	0	12	3	0
6	J	15	0	12	0	0
7	B	1	0	0	0	0
7	C	1	0	0	0	0
7	E	1	0	0	0	0
7	F	1	0	0	0	0
8	C	15	0	15	1	0
8	F	15	0	15	0	0
8	I	15	0	15	3	0
All	All	16117	0	15448	913	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 29.

The worst 5 of 913 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:216:VAL:HG22	2:B:129:LYS:HE3	1.24	1.15
3:F:52:ASN:HD21	6:F:520:NDG:H8C3	1.01	1.13
2:E:443:ASN:H	2:E:443:ASN:HD22	1.05	1.01
2:E:371:MET:HB2	2:E:410:ARG:CB	1.92	0.99
3:F:356:ARG:HB3	3:F:356:ARG:NH1	1.79	0.97

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	190/491 (39%)	159 (84%)	24 (13%)	7 (4%)	3 7
1	D	192/491 (39%)	147 (77%)	32 (17%)	13 (7%)	1 1
2	B	400/464 (86%)	353 (88%)	39 (10%)	8 (2%)	7 19
2	E	399/464 (86%)	335 (84%)	52 (13%)	12 (3%)	4 10
3	C	388/409 (95%)	350 (90%)	27 (7%)	11 (3%)	5 11
3	F	387/409 (95%)	342 (88%)	34 (9%)	11 (3%)	5 11
4	G	2/4 (50%)	2 (100%)	0	0	100 100
4	H	2/4 (50%)	1 (50%)	1 (50%)	0	100 100
5	I	2/4 (50%)	2 (100%)	0	0	100 100
5	J	2/4 (50%)	2 (100%)	0	0	100 100
All	All	1964/2744 (72%)	1693 (86%)	209 (11%)	62 (3%)	4 9

5 of 62 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	203	LEU
1	A	210	PRO
1	A	211	LEU
2	B	65	PRO
2	B	411	CYS

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	177/430 (41%)	164 (93%)	13 (7%)	14	33
1	D	179/430 (42%)	167 (93%)	12 (7%)	16	37
2	B	350/402 (87%)	335 (96%)	15 (4%)	29	57
2	E	349/402 (87%)	335 (96%)	14 (4%)	31	60
3	C	341/355 (96%)	317 (93%)	24 (7%)	15	35
3	F	340/355 (96%)	322 (95%)	18 (5%)	22	48
4	G	3/3 (100%)	3 (100%)	0	100	100
4	H	3/3 (100%)	3 (100%)	0	100	100
5	I	3/3 (100%)	3 (100%)	0	100	100
5	J	3/3 (100%)	3 (100%)	0	100	100
All	All	1748/2386 (73%)	1652 (94%)	96 (6%)	21	46

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	317	ASN
1	D	63	TYR
3	F	325	ASN
3	C	325	ASN
3	C	356	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 93 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	325	ASN
1	D	130	GLN
3	F	239	GLN
3	C	361	ASN
1	D	69	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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	$\#\|Z\| > 2$
6	NDG	E	570	-	15,15,15	0.45	0	21,21,21	0.57	0
6	NDG	J	571	-	15,15,15	0.55	0	21,21,21	0.55	0
6	NDG	C	420	-	15,15,15	0.42	0	21,21,21	0.54	0
8	NAG	C	421	-	15,15,15	0.48	0	21,21,21	0.60	0
8	NAG	I	471	-	15,15,15	0.49	0	21,21,21	0.54	0
8	NAG	F	521	-	15,15,15	0.43	0	21,21,21	0.54	0
6	NDG	B	470	-	15,15,15	0.47	0	21,21,21	0.78	0
6	NDG	F	520	-	15,15,15	0.45	0	21,21,21	0.58	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
6	NDG	E	570	-	-	2/6/26/26	0/1/1/1
6	NDG	J	571	-	-	3/6/26/26	0/1/1/1
6	NDG	C	420	-	-	2/6/26/26	0/1/1/1
8	NAG	C	421	-	-	2/6/26/26	0/1/1/1
8	NAG	I	471	-	-	2/6/26/26	0/1/1/1
8	NAG	F	521	-	-	2/6/26/26	0/1/1/1
6	NDG	B	470	-	-	4/6/26/26	0/1/1/1
6	NDG	F	520	-	-	2/6/26/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	J	571	NDG	C1-C2-N2-C7
6	J	571	NDG	C8-C7-N2-C2
6	J	571	NDG	O7-C7-N2-C2
6	C	420	NDG	C8-C7-N2-C2
6	C	420	NDG	O7-C7-N2-C2

There are no ring outliers.

6 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	570	NDG	4	0
6	C	420	NDG	10	0
8	C	421	NAG	1	0
8	I	471	NAG	3	0
6	B	470	NDG	8	0
6	F	520	NDG	3	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

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.