



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 05:25 PM BST

PDB ID : 1J4Z
Title : Structural and mechanistic basis for allostery in the bacterial chaperonin GroEL; see remark 400
Authors : Wang, J.
Deposited on : 2002-01-02
Resolution : 3.52 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

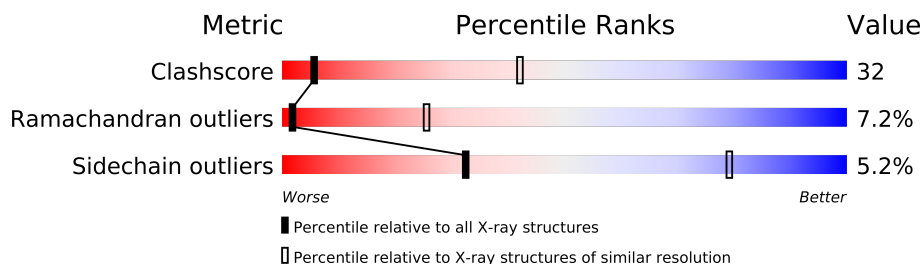
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : **NOT EXECUTED**
Percentile statistics : 21963
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23004

1 Overall quality at a glance

The reported resolution of this entry is 3.52 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	79885	1020 (3.68-3.36)
Ramachandran outliers	78287	1082 (3.70-3.34)
Sidechain outliers	78261	1082 (3.70-3.34)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	547	
1	B	547	
1	C	547	
1	D	547	
1	E	547	
1	F	547	
1	G	547	
1	H	547	
1	I	547	
1	J	547	
1	K	547	
1	L	547	
1	M	547	
1	N	547	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 53998 atoms, of which 0 are hydrogen and 0 are deuterium.

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 GROEL PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	B	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	C	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	D	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	E	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	F	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	G	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	H	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	I	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	J	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	K	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	L	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	M	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			
1	N	525	Total	C	N	O	S	0	0	0
			3857	2400	667	770	20			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
B	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
B	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
C	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
C	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
D	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
D	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
E	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
E	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
F	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
F	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
G	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
G	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
H	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
H	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
I	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
I	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
J	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
J	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
K	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
K	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
L	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
L	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
M	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
M	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
N	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
N	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5

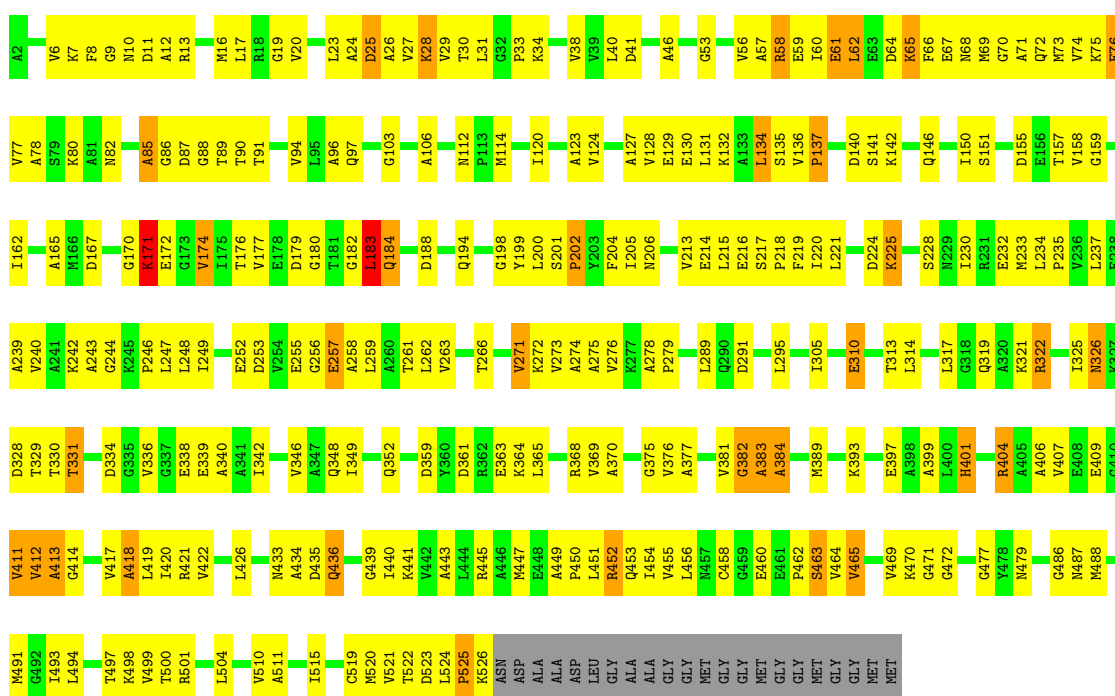
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

Note EDS was not executed.

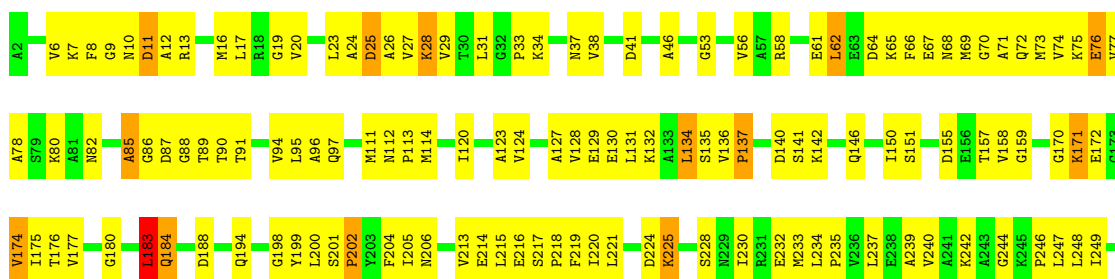
• Molecule 1: GROEL PROTEIN

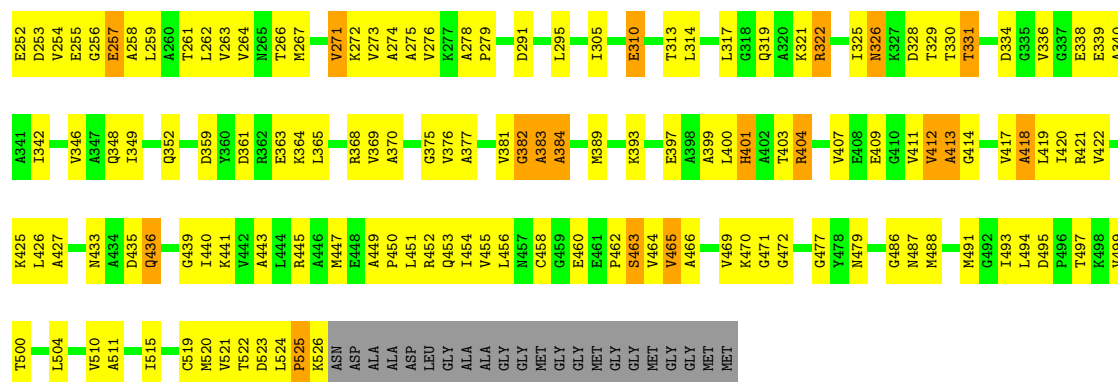
Chain A:



• Molecule 1: GROEL PROTEIN

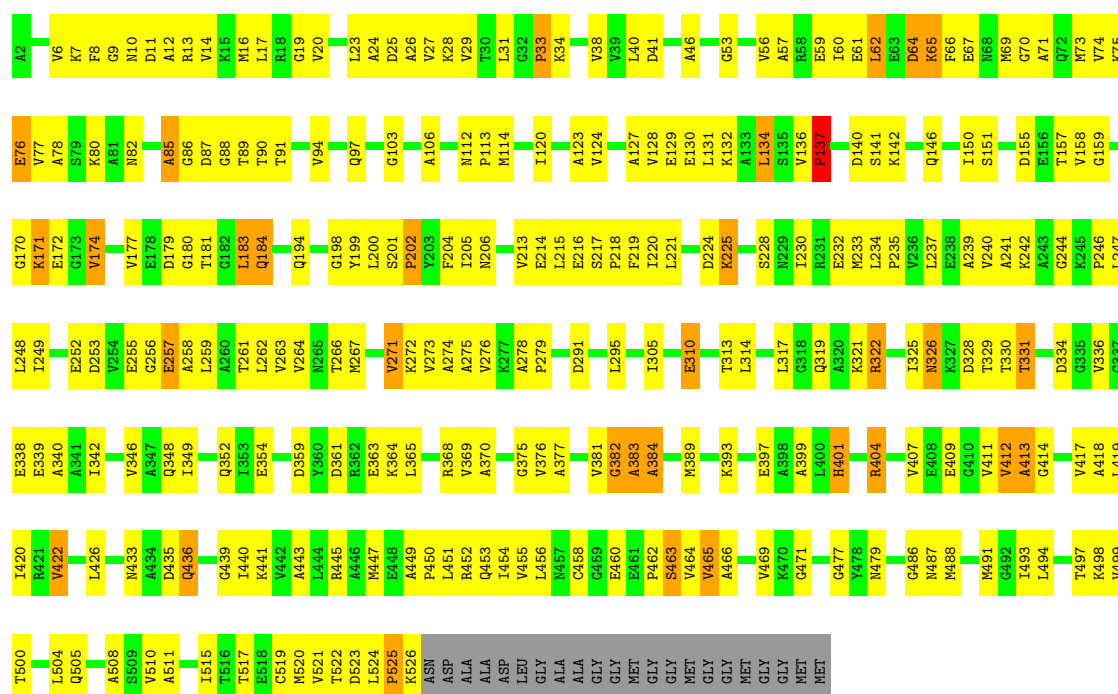
Chain B:





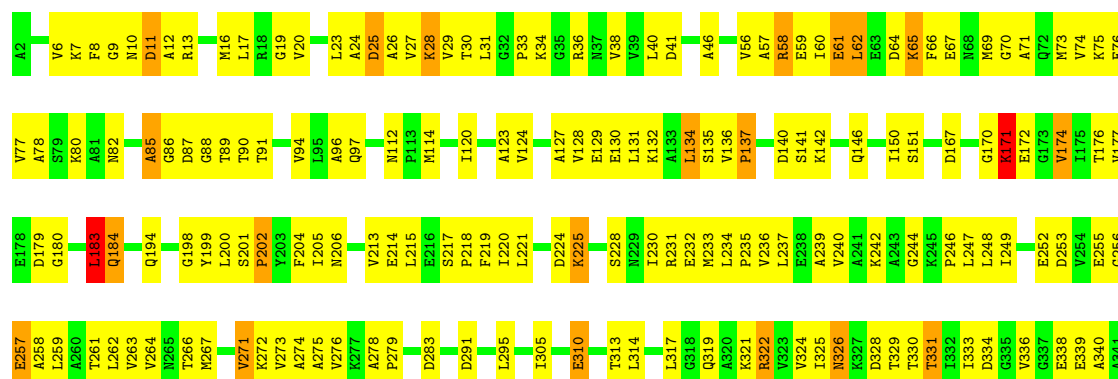
• Molecule 1: GROEL PROTEIN

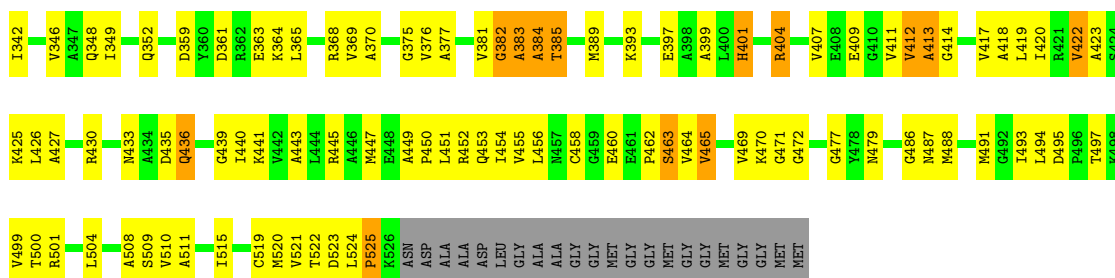
Chain C:



• Molecule 1: GROEL PROTEIN

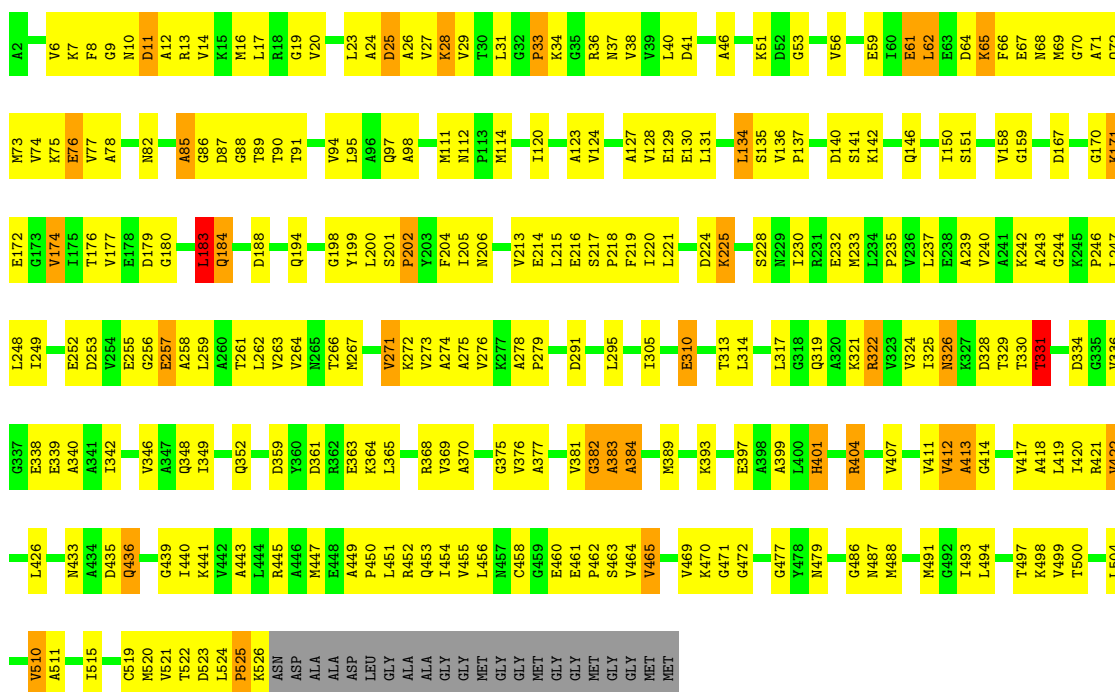
Chain D:





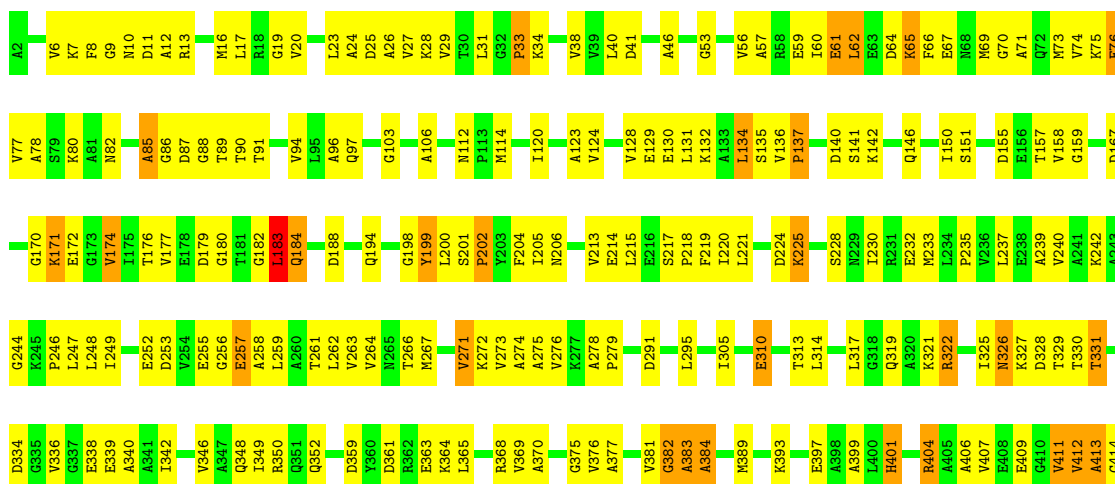
• Molecule 1: GROEL PROTEIN

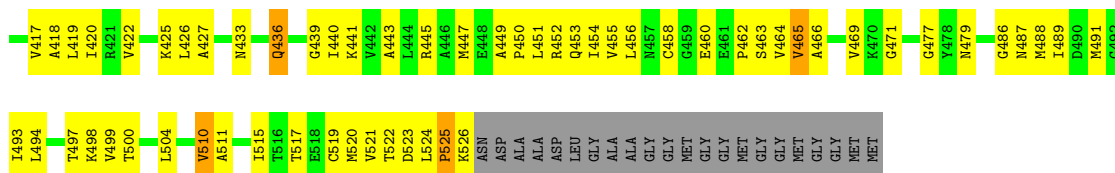
Chain E:



• Molecule 1: GROEL PROTEIN

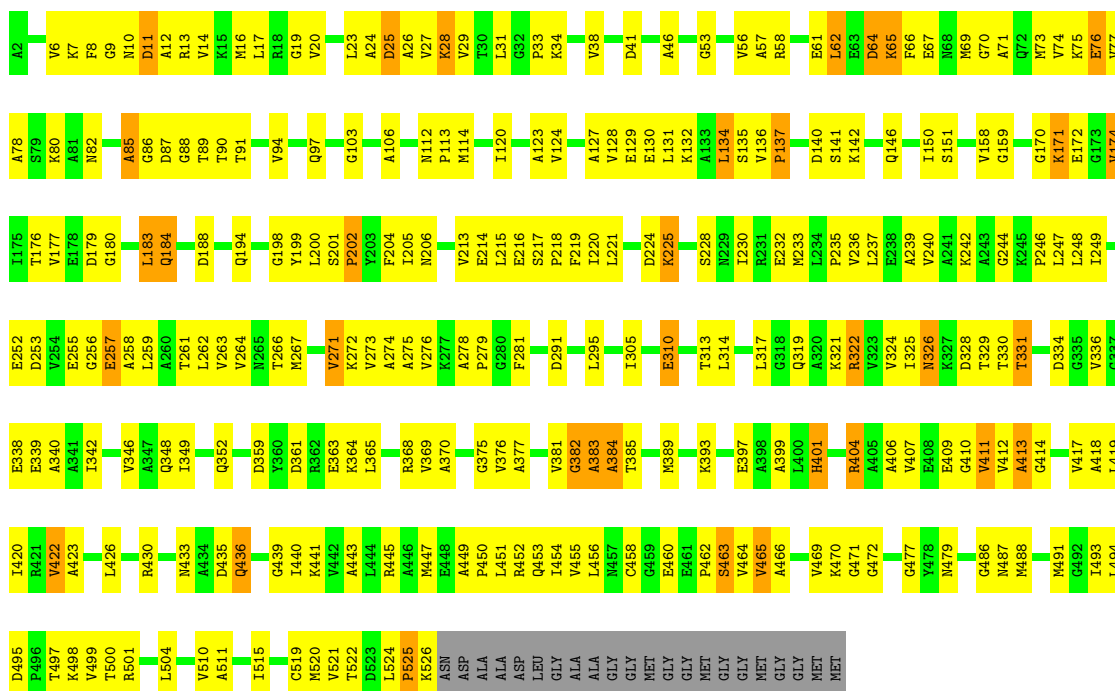
Chain F:





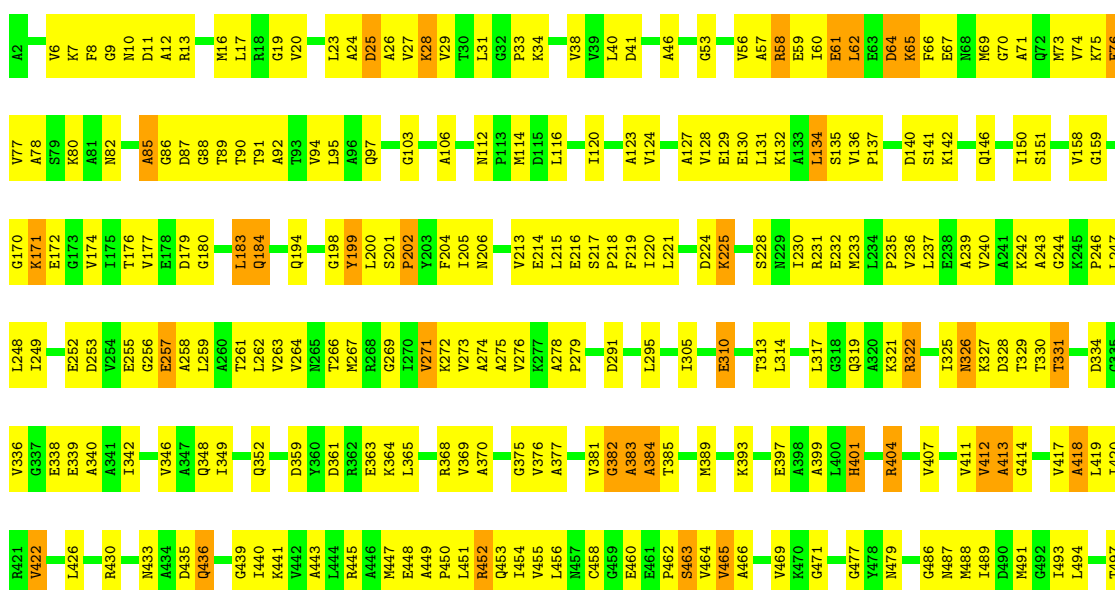
• Molecule 1: GROEL PROTEIN

Chain G:

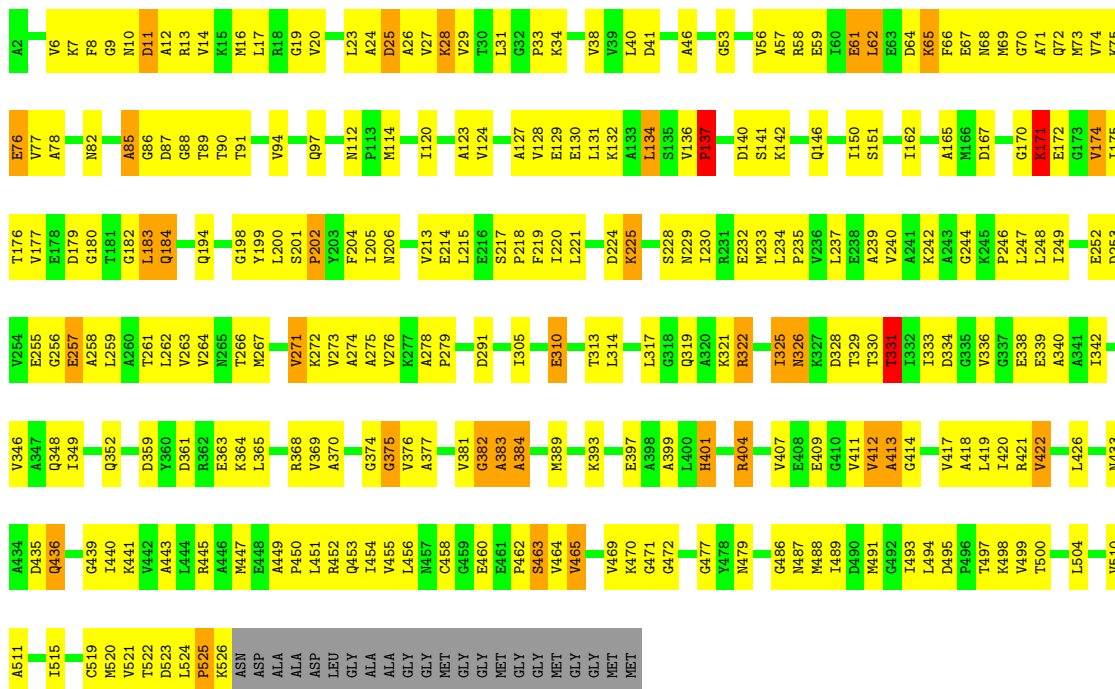


• Molecule 1: GROEL PROTEIN

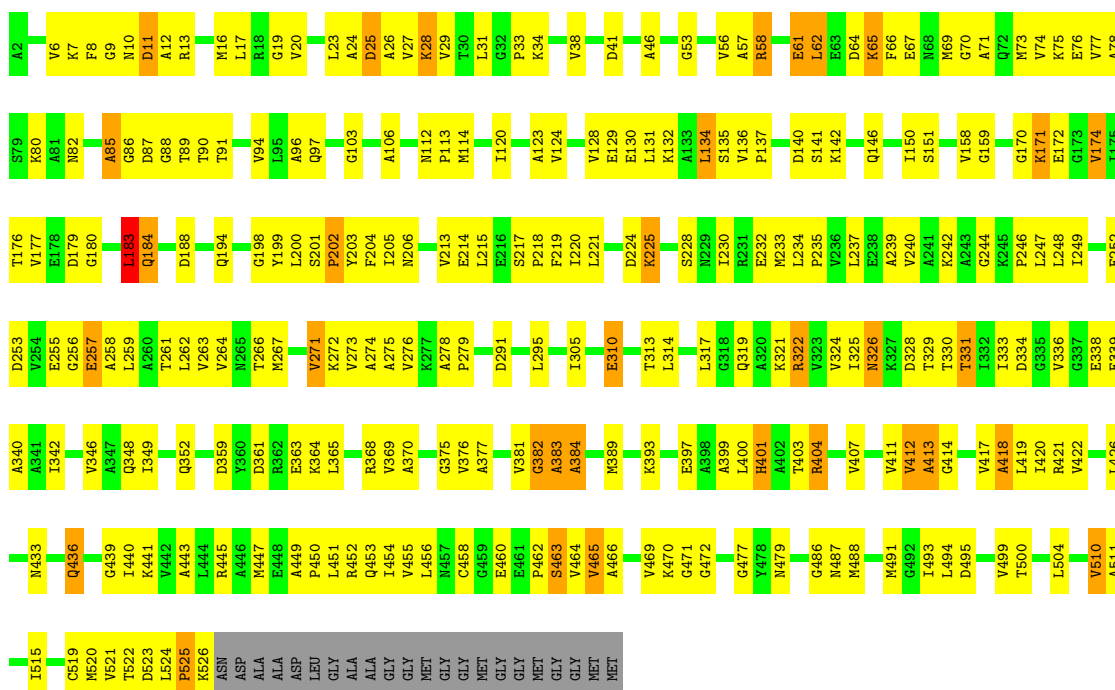
Chain H:



Chain I:

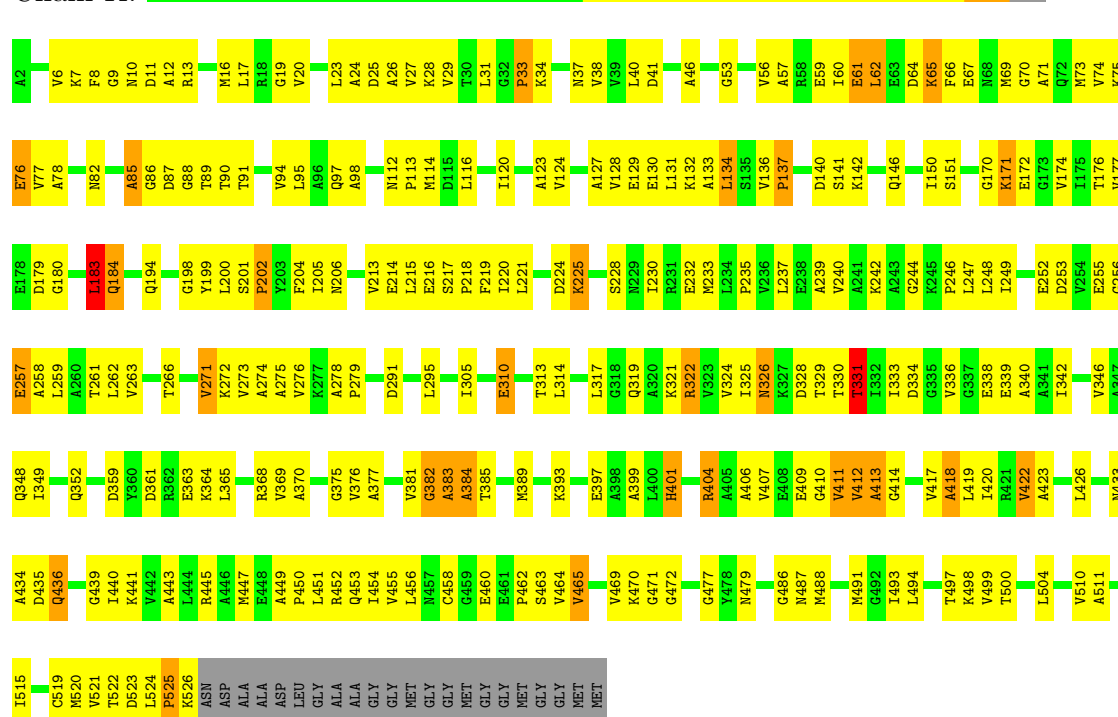


Chain J:



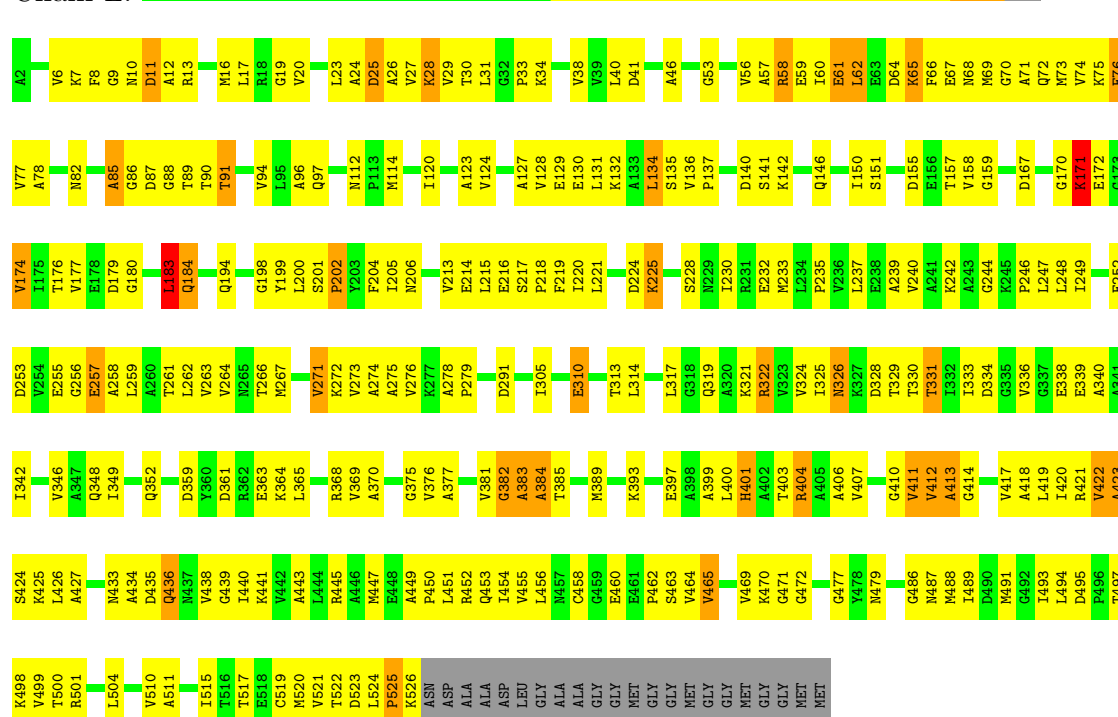
- Molecule 1: GROEL PROTEIN

Chain K:



- Molecule 1: GROEL PROTEIN

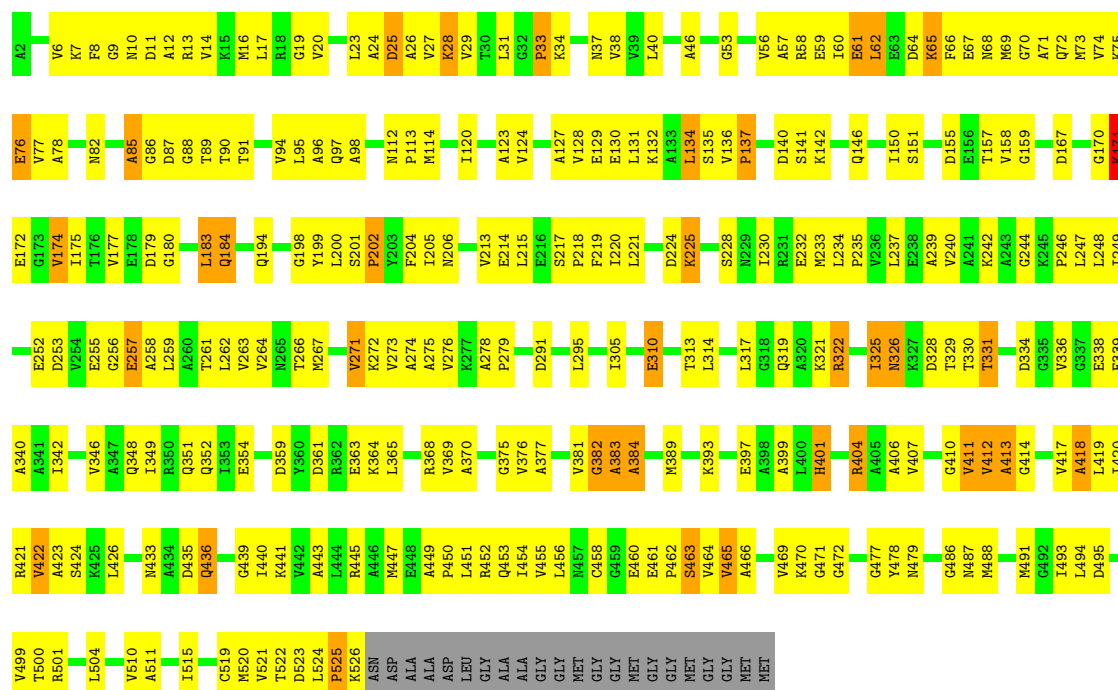
Chain L:



- Molecule 1: GROEL PROTEIN

Chain M:





4 Data and refinement statistics

EDS was not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	137.67Å 264.24Å 294.80Å 90.00° 92.39° 90.00°	Depositor
Resolution (Å)	20.00 – 3.52	Depositor
% Data completeness (in resolution range)	76.5 (20.00-3.52)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.87 (at 3.56Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.291 , 0.298	Depositor
Wilson B-factor (Å ²)	61.4	Xtriage
Anisotropy	0.684	Xtriage
Estimated twinning fraction	0.069 for h,-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	2 of 198920 reflections (0.001%)	Xtriage
Total number of atoms	53998	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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	0.46	0/3885	0.65	0/5245
1	B	0.48	0/3885	0.65	0/5245
1	C	0.48	0/3885	0.66	0/5245
1	D	0.48	0/3885	0.65	0/5245
1	E	0.47	0/3885	0.66	0/5245
1	F	0.47	0/3885	0.66	0/5245
1	G	0.47	0/3885	0.66	0/5245
1	H	0.48	0/3885	0.66	0/5245
1	I	0.51	0/3885	0.68	0/5245
1	J	0.48	0/3885	0.66	0/5245
1	K	0.49	0/3885	0.66	0/5245
1	L	0.49	0/3885	0.66	0/5245
1	M	0.53	0/3885	0.68	0/5245
1	N	0.52	0/3885	0.67	0/5245
All	All	0.49	0/54390	0.66	0/73430

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3857	0	3989	264	0
1	B	3857	0	3989	252	0
1	C	3857	0	3989	251	1
1	D	3857	0	3989	261	0
1	E	3857	0	3989	248	0
1	F	3857	0	3989	244	1
1	G	3857	0	3989	261	0
1	H	3857	0	3989	261	0
1	I	3857	0	3989	246	0
1	J	3857	0	3989	254	0
1	K	3857	0	3989	258	0
1	L	3857	0	3989	262	0
1	M	3857	0	3989	259	0
1	N	3857	0	3989	258	0
All	All	53998	0	55846	3489	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 32.

The worst 5 of 3489 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:230:ILE:HD12	1:E:261:THR:HG21	1.32	1.10
1:D:230:ILE:HD12	1:D:261:THR:HG21	1.33	1.10
1:C:230:ILE:HD12	1:C:261:THR:HG21	1.32	1.10
1:A:183:LEU:H	1:A:383:ALA:HB3	1.16	1.09
1:G:230:ILE:HD12	1:G:261:THR:HG21	1.32	1.09

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:354:GLU:OE1	1:F:350:ARG:NH1[1_455]	2.10	0.10

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	523/547 (96%)	397 (76%)	87 (17%)	39 (8%)	2	24
1	B	523/547 (96%)	398 (76%)	91 (17%)	34 (6%)	2	29
1	C	523/547 (96%)	400 (76%)	87 (17%)	36 (7%)	2	27
1	D	523/547 (96%)	394 (75%)	88 (17%)	41 (8%)	1	22
1	E	523/547 (96%)	393 (75%)	94 (18%)	36 (7%)	2	27
1	F	523/547 (96%)	397 (76%)	91 (17%)	35 (7%)	2	28
1	G	523/547 (96%)	399 (76%)	88 (17%)	36 (7%)	2	27
1	H	523/547 (96%)	393 (75%)	90 (17%)	40 (8%)	2	23
1	I	523/547 (96%)	391 (75%)	95 (18%)	37 (7%)	2	26
1	J	523/547 (96%)	398 (76%)	88 (17%)	37 (7%)	2	26
1	K	523/547 (96%)	396 (76%)	89 (17%)	38 (7%)	2	25
1	L	523/547 (96%)	399 (76%)	84 (16%)	40 (8%)	2	23
1	M	523/547 (96%)	397 (76%)	87 (17%)	39 (8%)	2	24
1	N	523/547 (96%)	399 (76%)	88 (17%)	36 (7%)	2	27
All	All	7322/7658 (96%)	5551 (76%)	1247 (17%)	524 (7%)	2	25

5 of 524 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	PRO
1	A	202	PRO
1	A	256	GLY
1	A	334	ASP
1	A	384	ALA

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	403/412 (98%)	382 (95%)	21 (5%)	32	79
1	B	403/412 (98%)	382 (95%)	21 (5%)	32	79
1	C	403/412 (98%)	383 (95%)	20 (5%)	34	80

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	403/412 (98%)	383 (95%)	20 (5%)	34	80
1	E	403/412 (98%)	380 (94%)	23 (6%)	29	76
1	F	403/412 (98%)	381 (94%)	22 (6%)	30	77
1	G	403/412 (98%)	384 (95%)	19 (5%)	36	82
1	H	403/412 (98%)	385 (96%)	18 (4%)	38	83
1	I	403/412 (98%)	380 (94%)	23 (6%)	29	76
1	J	403/412 (98%)	383 (95%)	20 (5%)	34	80
1	K	403/412 (98%)	382 (95%)	21 (5%)	32	79
1	L	403/412 (98%)	381 (94%)	22 (6%)	30	77
1	M	403/412 (98%)	380 (94%)	23 (6%)	29	76
1	N	403/412 (98%)	381 (94%)	22 (6%)	30	77
All	All	5642/5768 (98%)	5347 (95%)	295 (5%)	32	79

5 of 295 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	174	VAL
1	I	62	LEU
1	M	523	ASP
1	G	401	HIS
1	H	76	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 132 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	146	GLN
1	H	366	GLN
1	M	366	GLN
1	G	265	ASN
1	H	97	GLN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.