



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 14, 2023 – 09:59 AM EDT

PDB ID : 4V43  
Title : Structural and mechanistic basis for allostery in the bacterial chaperonin GroEL  
Authors : Wang, J.  
Deposited on : 2002-01-02  
Resolution : 3.52 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

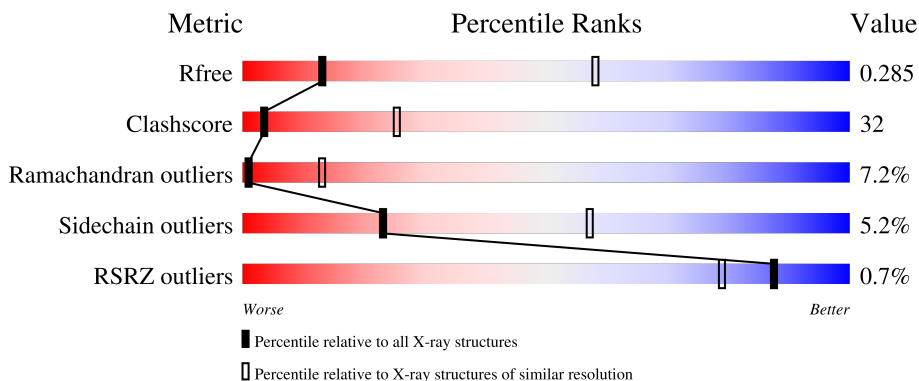
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	130704	1161 (3.60-3.44)
Clashscore	141614	1244 (3.60-3.44)
Ramachandran outliers	138981	1206 (3.60-3.44)
Sidechain outliers	138945	1207 (3.60-3.44)
RSRZ outliers	127900	1080 (3.60-3.44)

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 of 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	1	547	
1	2	547	
1	A	547	
1	B	547	
1	C	547	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	D	547	47% 43% 6% .
1	E	547	% 48% 42% 6% .
1	F	547	% 47% 43% 6% .
1	G	547	% 47% 43% 6% .
1	H	547	% 47% 43% 7% .
1	I	547	48% 42% 6% . .
1	J	547	48% 41% 6% .
1	K	547	48% 42% 5% .
1	L	547	% 45% 44% 6% .
1	M	547	46% 43% 7% .
1	N	547	47% 43% 6% .
1	O	547	% 46% 43% 6% .
1	P	547	47% 43% 6% .
1	Q	547	% 46% 43% 6% .
1	R	547	% 48% 42% 6% .
1	S	547	% 48% 42% 5% .
1	T	547	% 47% 43% 6% .
1	U	547	% 46% 43% 6% .
1	V	547	47% 43% 6% .
1	W	547	% 45% 44% 7% .
1	X	547	46% 44% 5% .
1	Y	547	47% 43% 6% .
1	Z	547	2% 46% 44% 6% .

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 107996 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 GROEL PROTEIN.

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

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	Q	525	3857	2400	667	770	20	0	0	0
1	R	525	3857	2400	667	770	20	0	0	0
1	S	525	3857	2400	667	770	20	0	0	0
1	T	525	3857	2400	667	770	20	0	0	0
1	U	525	3857	2400	667	770	20	0	0	0
1	V	525	3857	2400	667	770	20	0	0	0
1	W	525	3857	2400	667	770	20	0	0	0
1	X	525	3857	2400	667	770	20	0	0	0
1	Y	525	3857	2400	667	770	20	0	0	0
1	Z	525	3857	2400	667	770	20	0	0	0
1	1	525	3857	2400	667	770	20	0	0	0
1	2	525	3857	2400	667	770	20	0	0	0

There are 56 discrepancies between the modelled and reference sequences:

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

Continued on next page...

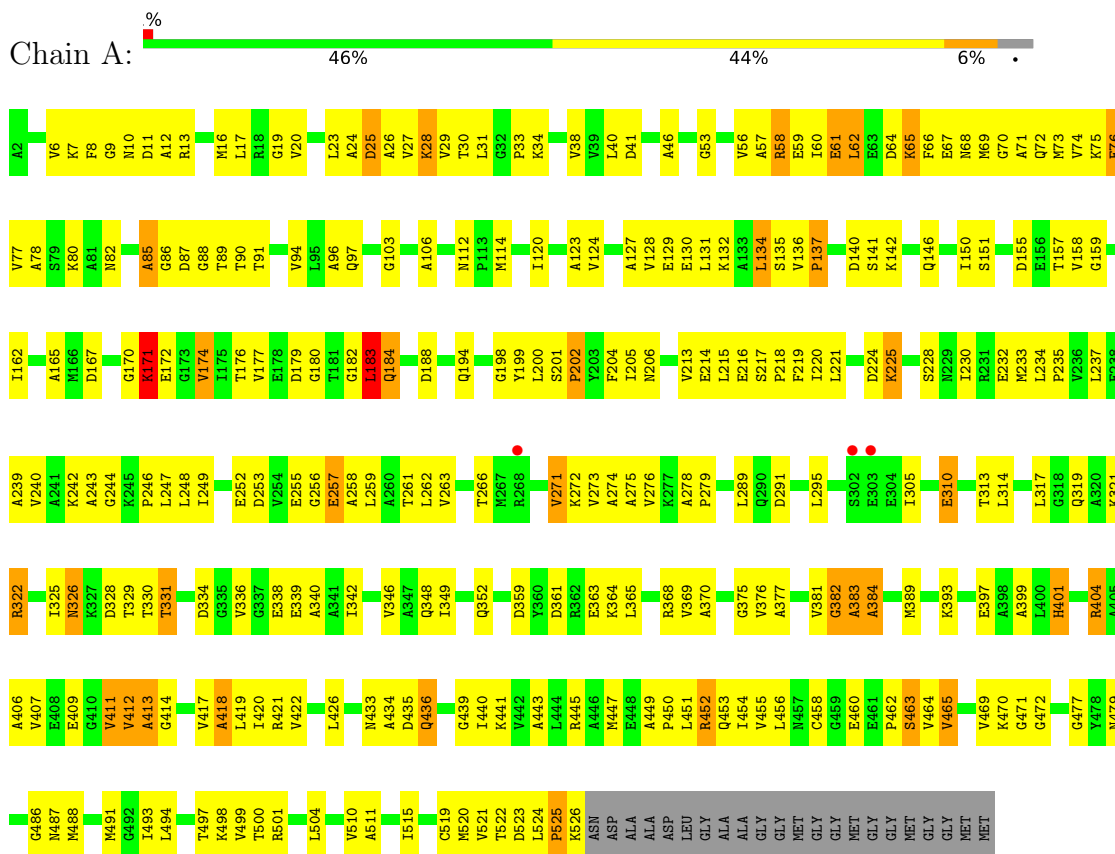
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
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
O	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
O	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
P	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
P	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
Q	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
Q	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
R	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
R	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
S	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
S	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
T	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
T	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
U	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
U	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
V	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
V	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
W	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
W	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
X	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
X	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
Y	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
Y	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
Z	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
Z	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
1	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
1	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5
2	398	ALA	ASP	SEE REMARK 999	UNP P0A6F5
2	434	ALA	GLU	SEE REMARK 999	UNP P0A6F5

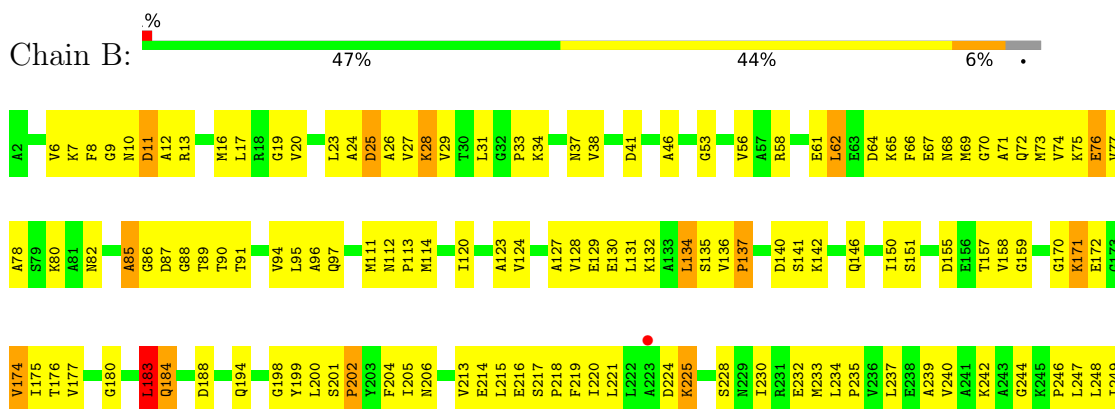
### 3 Residue-property plots i

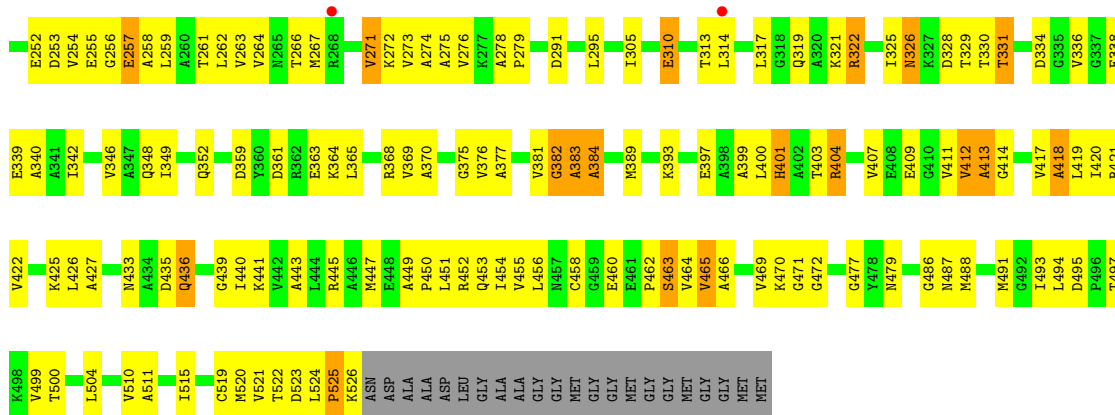
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: GROEL PROTEIN

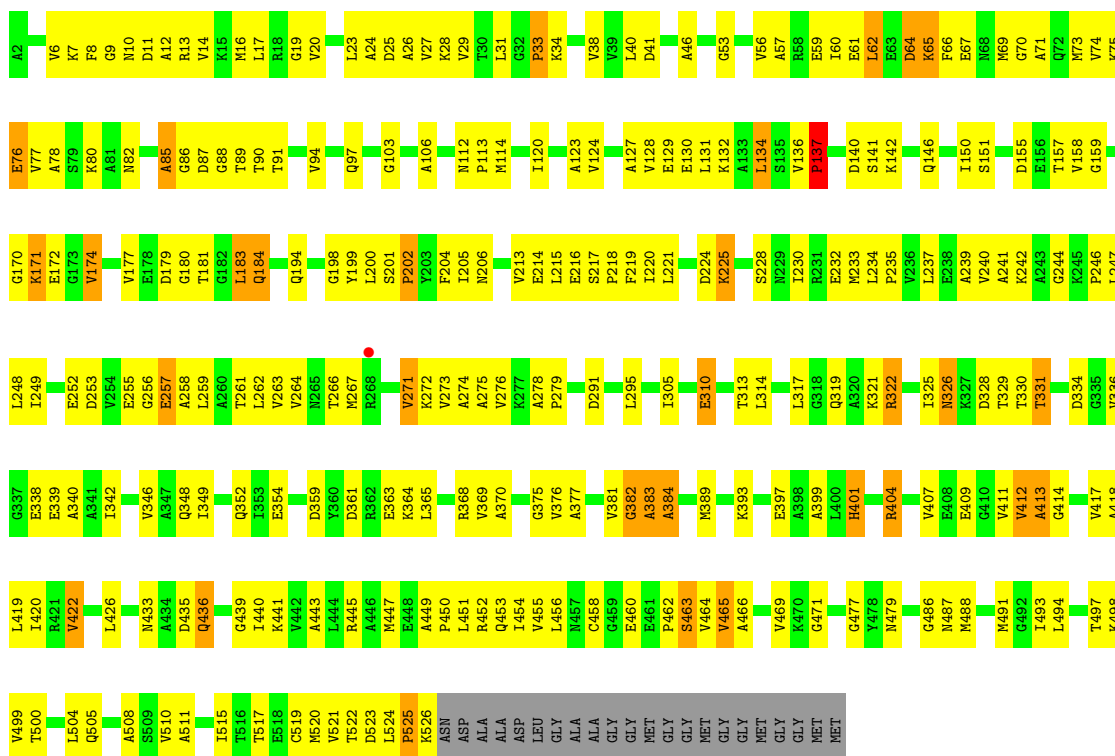


#### • Molecule 1: GROEL PROTEIN

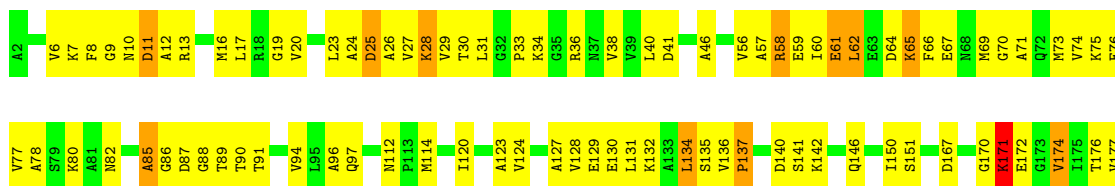




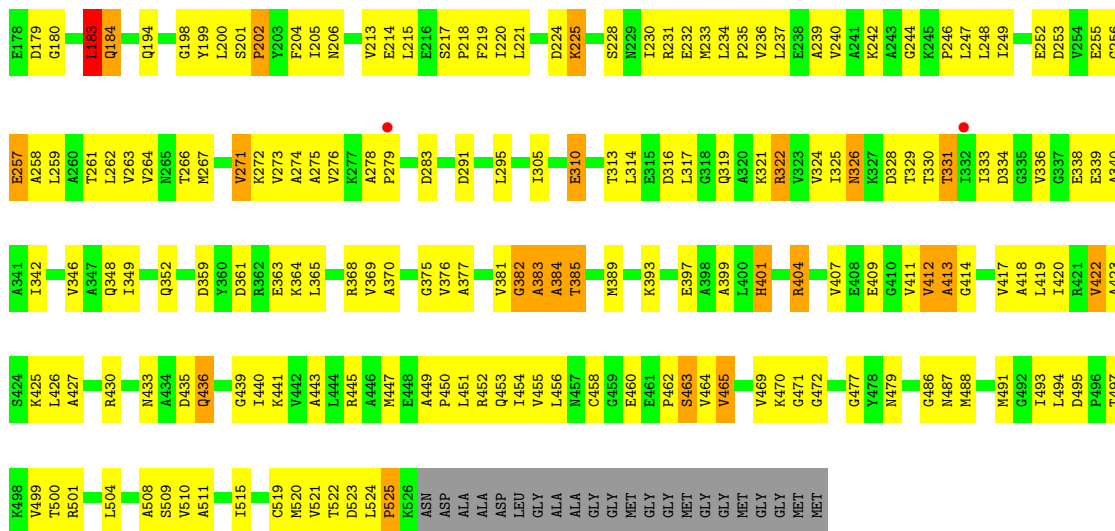
• Molecule 1: GROEL PROTEIN



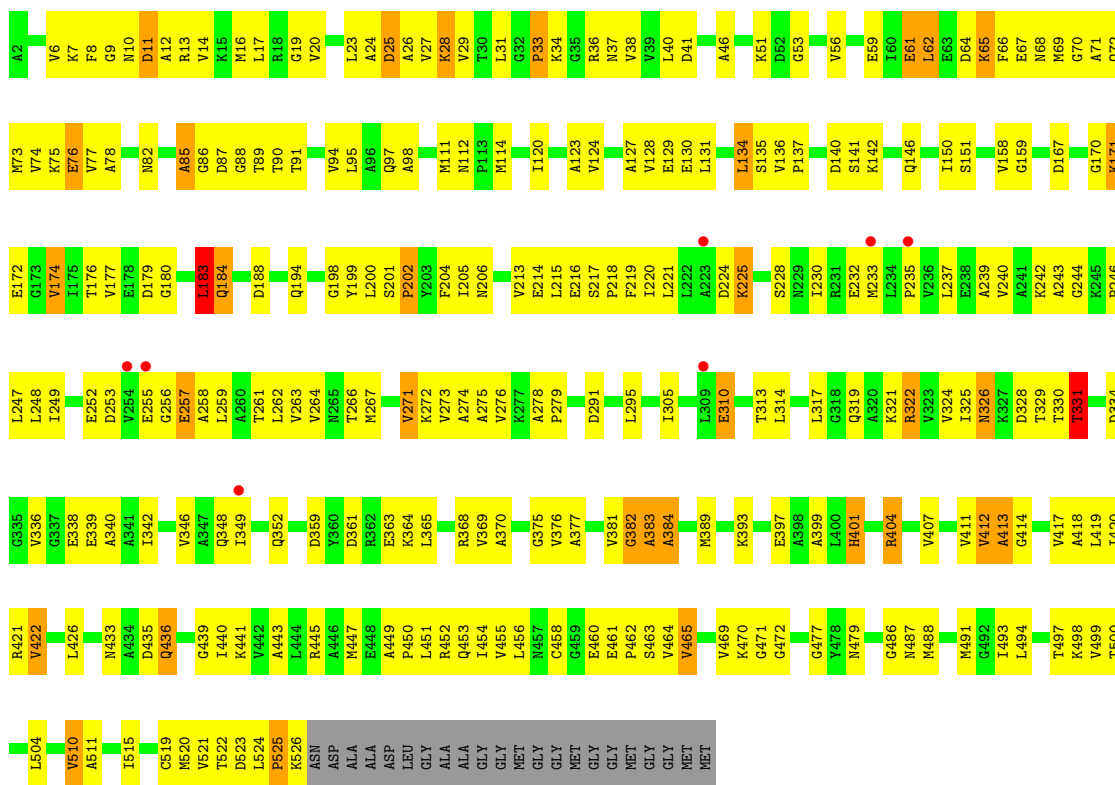
• Molecule 1: GROEL PROTEIN



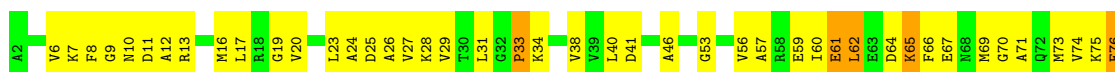


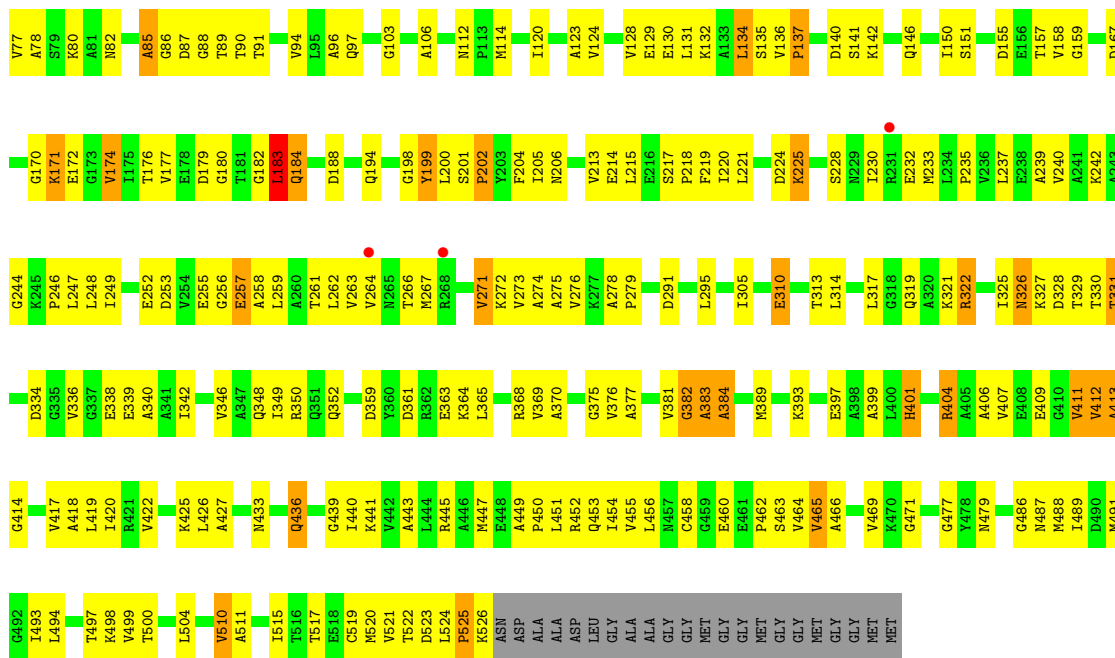


• Molecule 1: GROEL PROTEIN

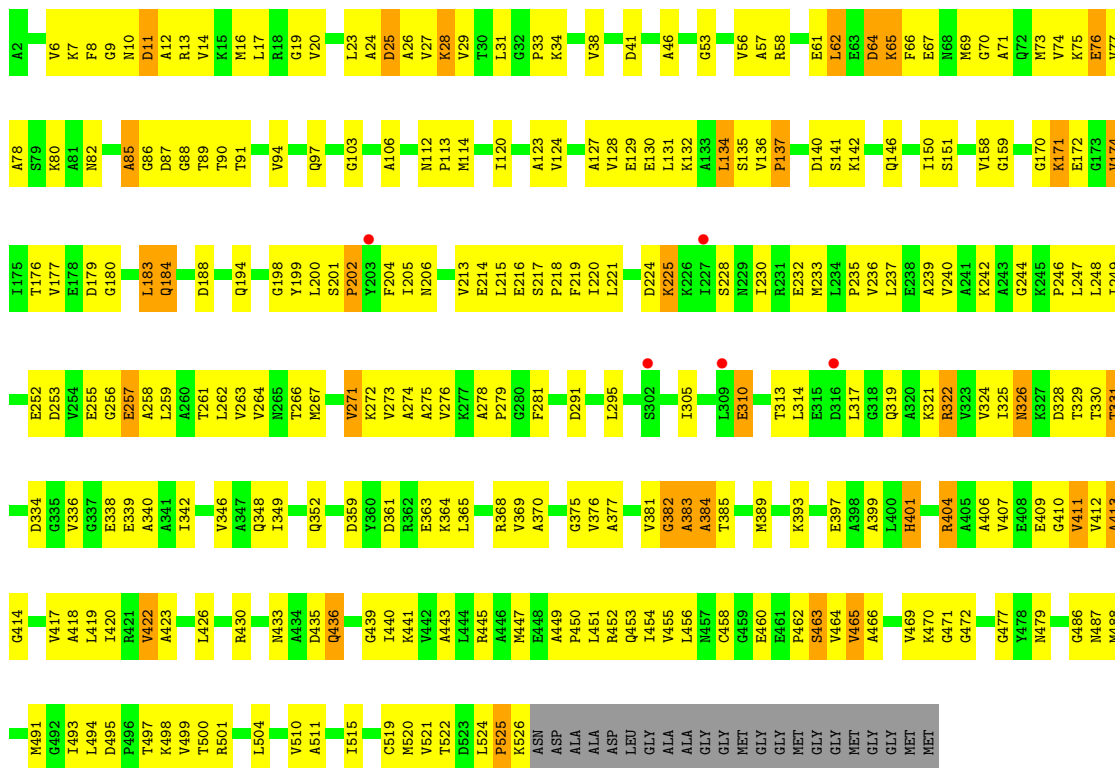


• Molecule 1: GROEL PROTEIN



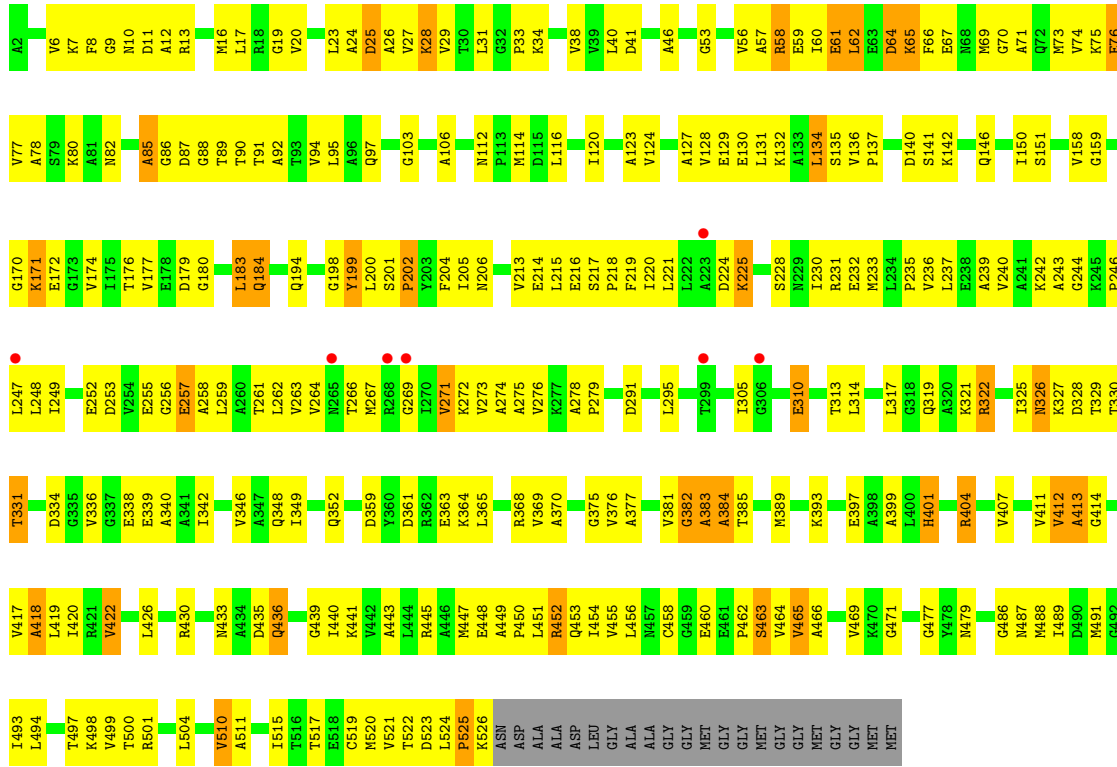


• Molecule 1: GROEL PROTEIN

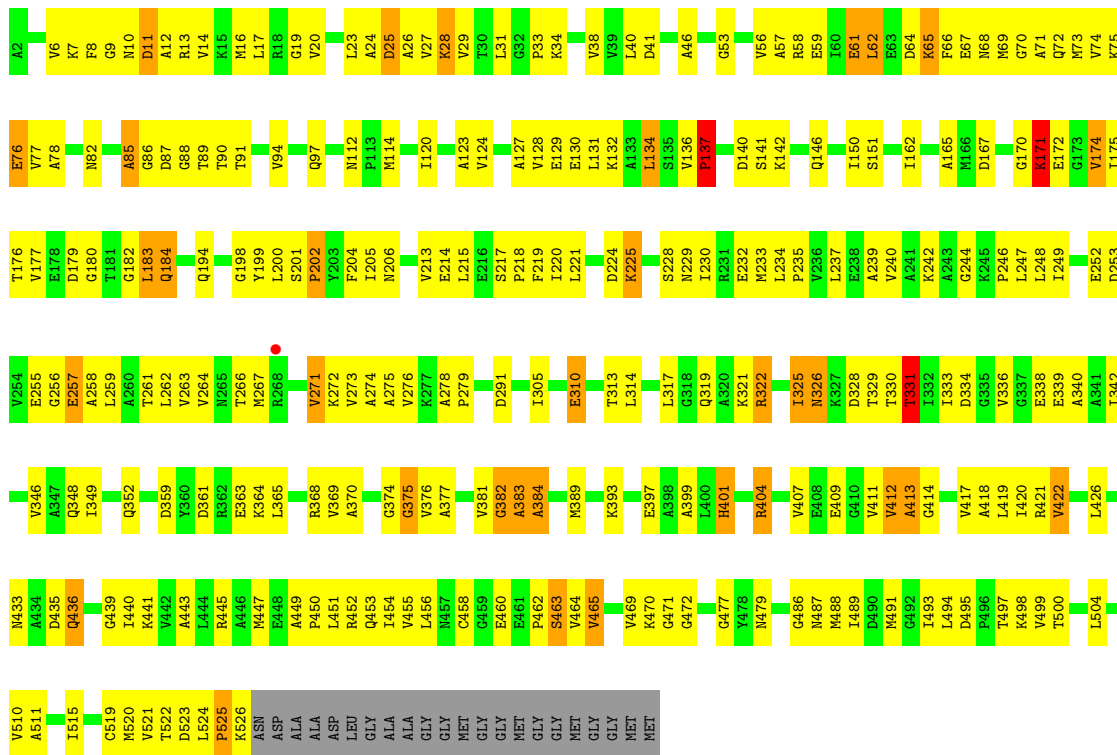


• Molecule 1: GROEL PROTEIN





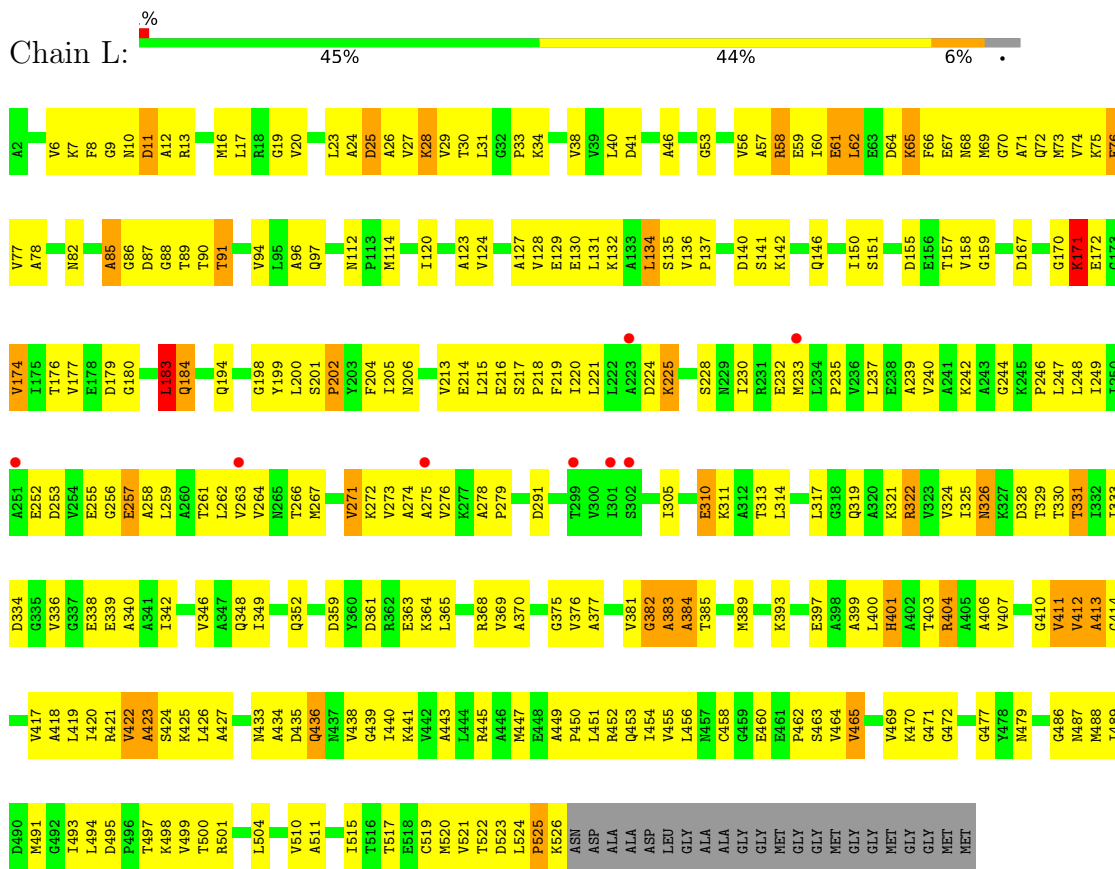
• Molecule 1: GROEL PROTEIN



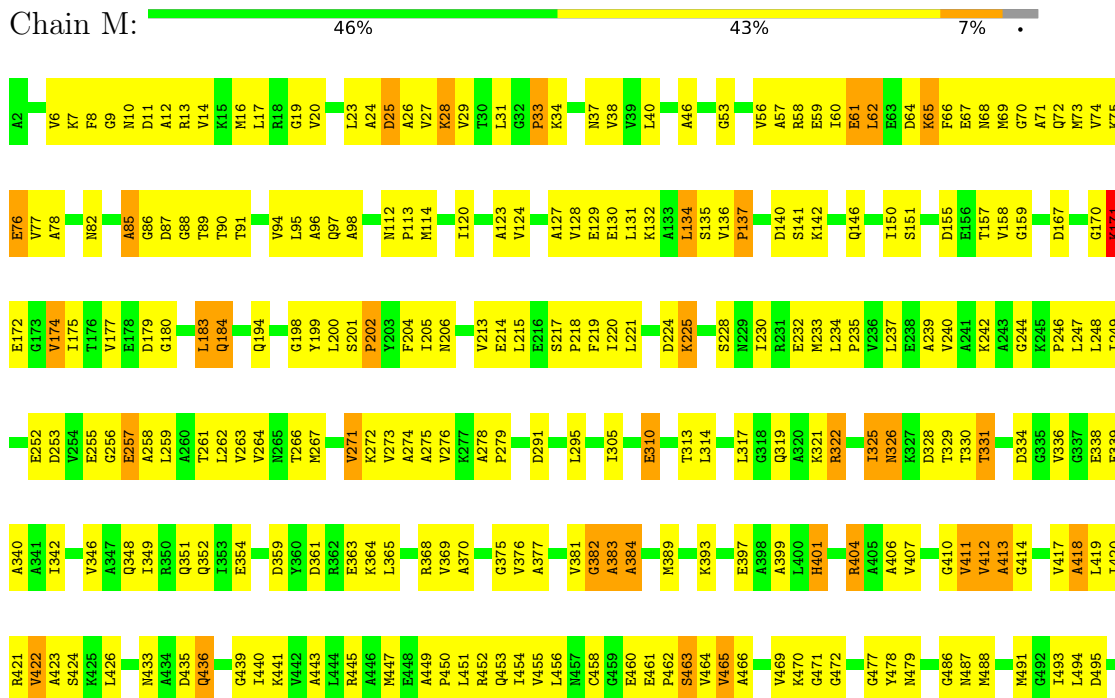
• Molecule 1: GROEL PROTEIN



• Molecule 1: GROEL PROTEIN

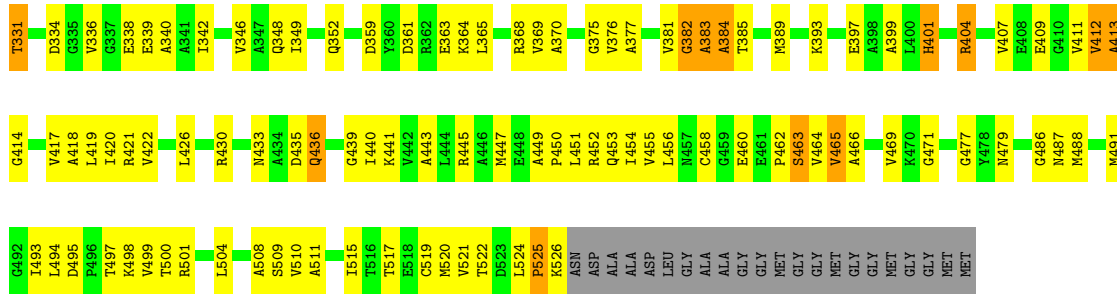


• Molecule 1: GROEL PROTEIN

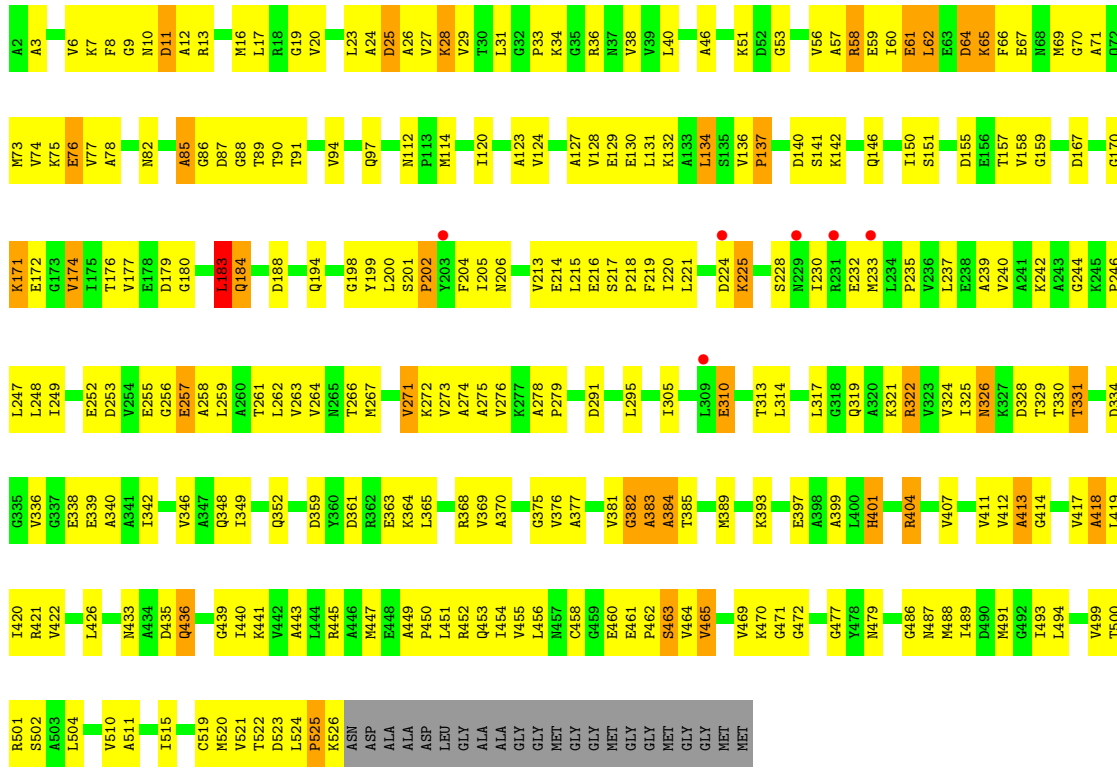




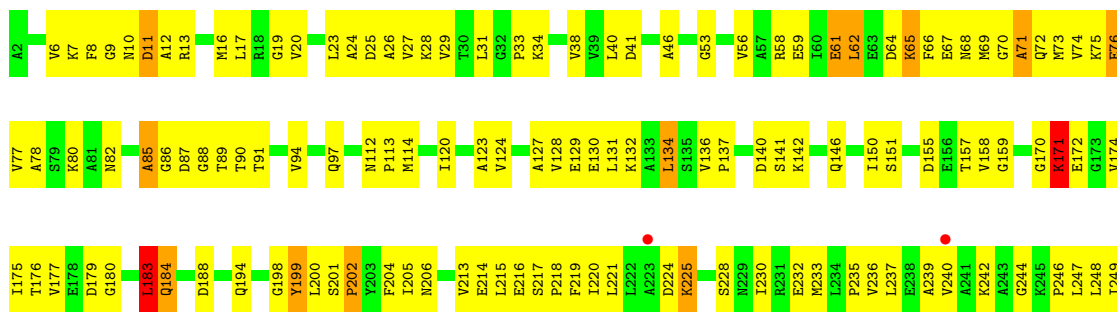




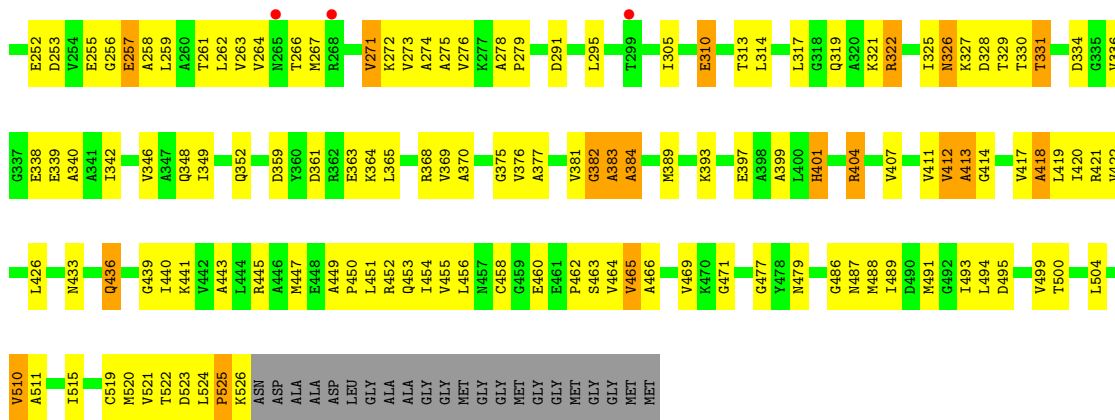
• Molecule 1: GROEL PROTEIN



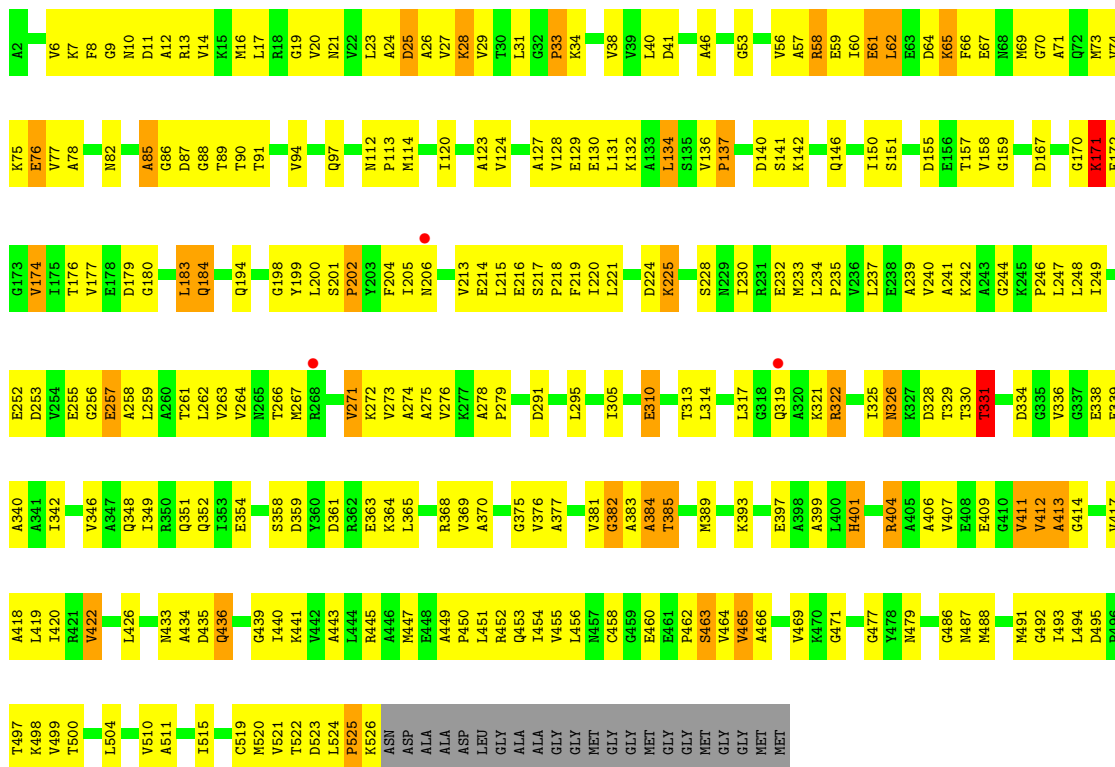
• Molecule 1: GROEL PROTEIN



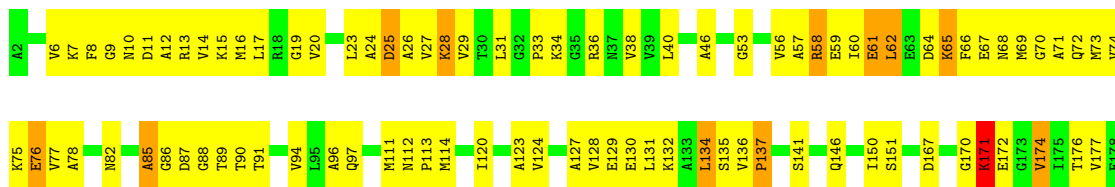




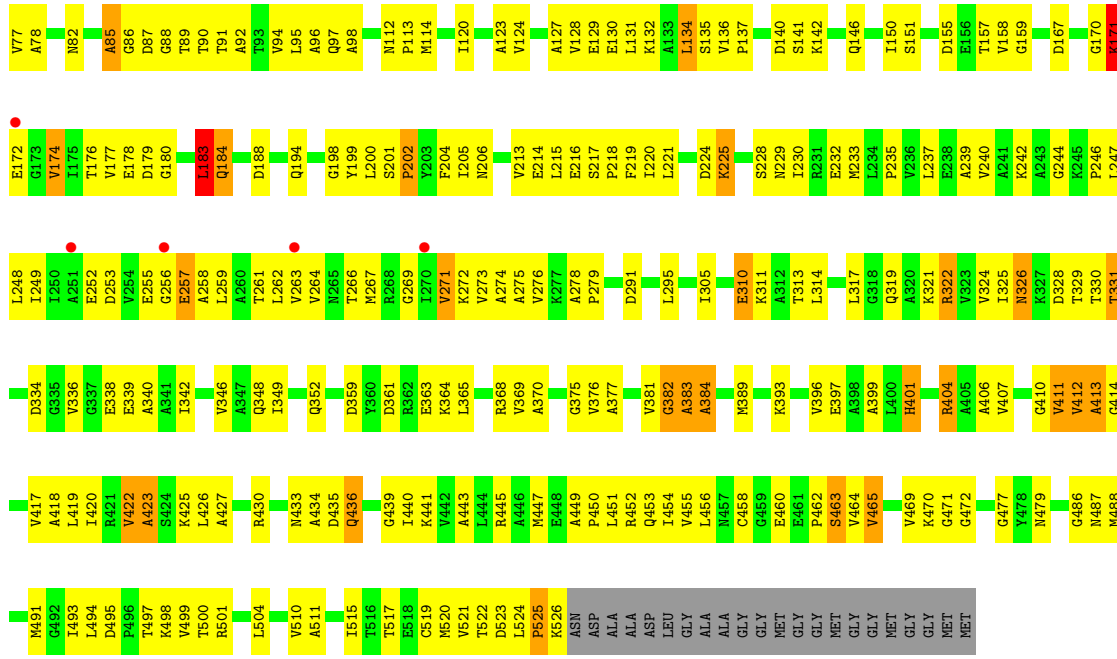
• Molecule 1: GROEL PROTEIN



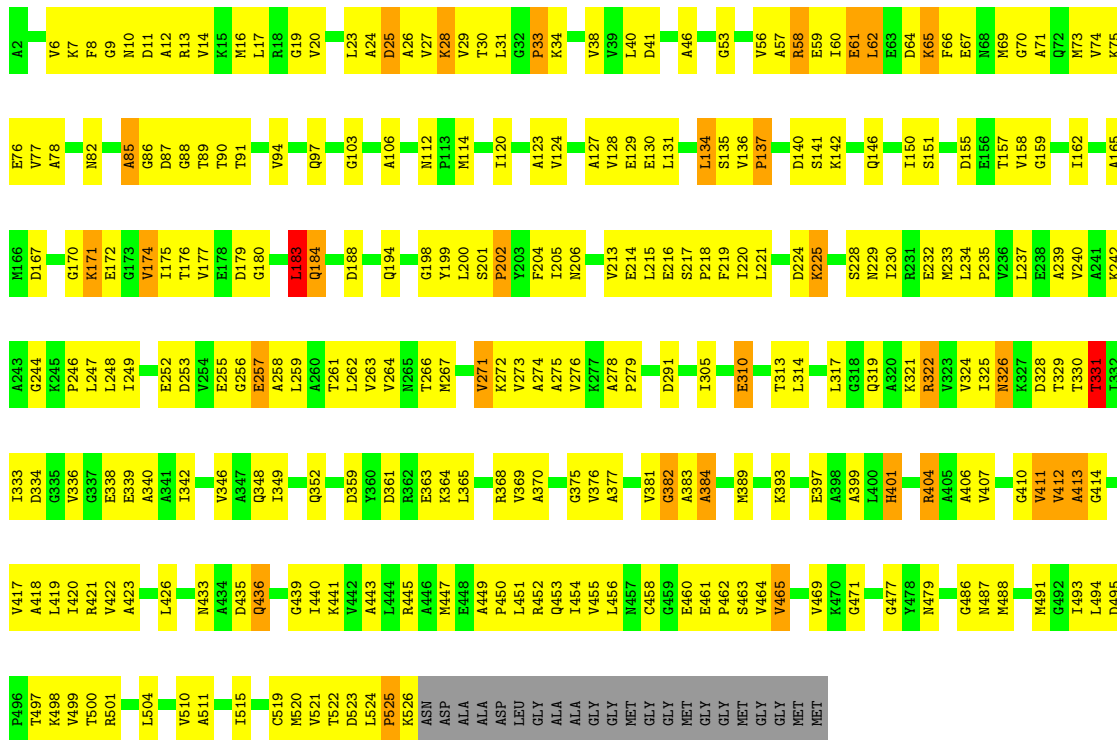
• Molecule 1: GROEL PROTEIN







• Molecule 1: GROEL PROTEIN



• Molecule 1: GROEL PROTEIN







## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	137.67Å 264.24Å 294.80Å 90.00° 92.39° 90.00°	Depositor
Resolution (Å)	20.00 – 3.52 30.25 – 3.52	Depositor EDS
% Data completeness (in resolution range)	76.5 (20.00-3.52) 76.5 (30.25-3.52)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.87 (at 3.56Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.291 , 0.298 0.280 , 0.285	Depositor DCC
$R_{free}$ test set	9897 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.4	Xtrriage
Anisotropy	0.684	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 8.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.069 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	107996	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtrriage'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.*

<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

### 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	1	0.51	0/3885	0.67	0/5245
1	2	0.49	0/3885	0.66	0/5245
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
1	O	0.47	0/3885	0.66	0/5245
1	P	0.47	0/3885	0.66	0/5245
1	Q	0.48	0/3885	0.66	0/5245
1	R	0.45	0/3885	0.65	0/5245
1	S	0.47	0/3885	0.65	0/5245
1	T	0.47	0/3885	0.65	0/5245
1	U	0.47	0/3885	0.65	0/5245
1	V	0.50	0/3885	0.67	0/5245
1	W	0.48	0/3885	0.66	0/5245
1	X	0.54	0/3885	0.67	0/5245
1	Y	0.51	0/3885	0.67	0/5245
1	Z	0.47	0/3885	0.66	0/5245
All	All	0.49	0/108780	0.66	0/146860

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
1	Y	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	Y	476	TYR	Sidechain

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

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	1	3857	0	3989	260	0
1	2	3857	0	3989	255	0
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	2
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	265	0
1	M	3857	0	3989	259	0
1	N	3857	0	3989	258	0
1	O	3857	0	3989	262	0
1	P	3857	0	3989	254	2
1	Q	3857	0	3989	260	0
1	R	3857	0	3989	246	0
1	S	3857	0	3989	244	0
1	T	3857	0	3989	254	2
1	U	3857	0	3989	264	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	V	3857	0	3989	258	0
1	W	3857	0	3989	265	2
1	X	3857	0	3989	259	0
1	Y	3857	0	3989	260	0
1	Z	3857	0	3989	261	0
All	All	107996	0	111692	7010	5

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 32.

The worst 5 of 7010 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:E:230:ILE:HD12	1:E:261:THR:HG21	1.32	1.10
1:W:230:ILE:HD12	1:W:261:THR:HG21	1.33	1.10
1:D:230:ILE:HD12	1:D:261:THR:HG21	1.33	1.10
1:P:230:ILE:HD12	1:P:261:THR:HG21	1.33	1.10
1:C:230:ILE:HD12	1:C:261:THR:HG21	1.32	1.10

All (5) 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	Interatomic distance (Å)	Clash overlap (Å)
1:D:313:THR:N	1:W:311:LYS:NZ[1_556]	1.94	0.26
1:C:354:GLU:OE1	1:F:350:ARG:NH1[1_455]	2.10	0.10
1:P:167:ASP:OD2	1:T:358:SER:OG[1_455]	2.15	0.05
1:P:350:ARG:NH1	1:T:354:GLU:OE1[1_455]	2.17	0.03
1:D:316:ASP:OD2	1:W:311:LYS:CE[1_556]	2.18	0.02

## 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	1	523/547 (96%)	393 (75%)	93 (18%)	37 (7%)	1	13
1	2	523/547 (96%)	397 (76%)	90 (17%)	36 (7%)	1	14
1	A	523/547 (96%)	397 (76%)	87 (17%)	39 (8%)	1	12
1	B	523/547 (96%)	398 (76%)	91 (17%)	34 (6%)	1	15
1	C	523/547 (96%)	400 (76%)	87 (17%)	36 (7%)	1	14
1	D	523/547 (96%)	394 (75%)	88 (17%)	41 (8%)	1	11
1	E	523/547 (96%)	393 (75%)	94 (18%)	36 (7%)	1	14
1	F	523/547 (96%)	397 (76%)	91 (17%)	35 (7%)	1	15
1	G	523/547 (96%)	399 (76%)	88 (17%)	36 (7%)	1	14
1	H	523/547 (96%)	393 (75%)	90 (17%)	40 (8%)	1	11
1	I	523/547 (96%)	391 (75%)	95 (18%)	37 (7%)	1	13
1	J	523/547 (96%)	398 (76%)	88 (17%)	37 (7%)	1	13
1	K	523/547 (96%)	396 (76%)	89 (17%)	38 (7%)	1	13
1	L	523/547 (96%)	399 (76%)	84 (16%)	40 (8%)	1	11
1	M	523/547 (96%)	397 (76%)	87 (17%)	39 (8%)	1	12
1	N	523/547 (96%)	399 (76%)	88 (17%)	36 (7%)	1	14
1	O	523/547 (96%)	394 (75%)	92 (18%)	37 (7%)	1	13
1	P	523/547 (96%)	392 (75%)	94 (18%)	37 (7%)	1	13
1	Q	523/547 (96%)	395 (76%)	91 (17%)	37 (7%)	1	13
1	R	523/547 (96%)	395 (76%)	88 (17%)	40 (8%)	1	11
1	S	523/547 (96%)	400 (76%)	84 (16%)	39 (8%)	1	12
1	T	523/547 (96%)	401 (77%)	84 (16%)	38 (7%)	1	13
1	U	523/547 (96%)	394 (75%)	91 (17%)	38 (7%)	1	13
1	V	523/547 (96%)	392 (75%)	94 (18%)	37 (7%)	1	13
1	W	523/547 (96%)	396 (76%)	87 (17%)	40 (8%)	1	11
1	X	523/547 (96%)	400 (76%)	87 (17%)	36 (7%)	1	14
1	Y	523/547 (96%)	400 (76%)	87 (17%)	36 (7%)	1	14
1	Z	523/547 (96%)	391 (75%)	94 (18%)	38 (7%)	1	13
All	All	14644/15316 (96%)	11091 (76%)	2503 (17%)	1050 (7%)	1	13

5 of 1050 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 [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	1	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	2	403/412 (98%)	379 (94%)	24 (6%)	19	53
1	A	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	B	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	C	403/412 (98%)	383 (95%)	20 (5%)	24	58
1	D	403/412 (98%)	383 (95%)	20 (5%)	24	58
1	E	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	F	403/412 (98%)	381 (94%)	22 (6%)	21	55
1	G	403/412 (98%)	384 (95%)	19 (5%)	26	60
1	H	403/412 (98%)	385 (96%)	18 (4%)	27	62
1	I	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	J	403/412 (98%)	383 (95%)	20 (5%)	24	58
1	K	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	L	403/412 (98%)	381 (94%)	22 (6%)	21	55
1	M	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	N	403/412 (98%)	381 (94%)	22 (6%)	21	55
1	O	403/412 (98%)	385 (96%)	18 (4%)	27	62
1	P	403/412 (98%)	384 (95%)	19 (5%)	26	60
1	Q	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	R	403/412 (98%)	382 (95%)	21 (5%)	23	57

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	S	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	T	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	U	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	V	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	W	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	X	403/412 (98%)	380 (94%)	23 (6%)	20	54
1	Y	403/412 (98%)	382 (95%)	21 (5%)	23	57
1	Z	403/412 (98%)	386 (96%)	17 (4%)	30	63
All	All	11284/11536 (98%)	10693 (95%)	591 (5%)	23	57

5 of 591 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	W	74	VAL
1	2	174	VAL
1	W	412	VAL
1	W	62	LEU
1	Y	411	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 263 such sidechains are listed below:

Mol	Chain	Res	Type
1	Y	326	ASN
1	Z	194	GLN
1	2	351	GLN
1	K	97	GLN
1	J	366	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	1	525/547 (95%)	-0.28	1 (0%) 95 91	14, 53, 105, 117	0
1	2	525/547 (95%)	-0.21	2 (0%) 92 87	17, 68, 135, 147	0
1	A	525/547 (95%)	-0.17	3 (0%) 89 81	17, 76, 156, 168	0
1	B	525/547 (95%)	-0.16	3 (0%) 89 81	21, 69, 146, 158	0
1	C	525/547 (95%)	-0.24	1 (0%) 95 91	17, 67, 130, 142	0
1	D	525/547 (95%)	-0.19	2 (0%) 92 87	21, 80, 143, 155	0
1	E	525/547 (95%)	-0.13	7 (1%) 77 65	20, 79, 169, 181	0
1	F	525/547 (95%)	-0.19	3 (0%) 89 81	18, 71, 148, 160	0
1	G	525/547 (95%)	-0.09	5 (0%) 82 71	17, 77, 168, 180	0
1	H	525/547 (95%)	-0.13	7 (1%) 77 65	15, 71, 157, 169	0
1	I	525/547 (95%)	-0.28	1 (0%) 95 91	13, 52, 107, 119	0
1	J	525/547 (95%)	-0.23	2 (0%) 92 87	15, 72, 143, 155	0
1	K	525/547 (95%)	-0.19	1 (0%) 95 91	12, 58, 146, 158	0
1	L	525/547 (95%)	-0.11	8 (1%) 73 61	16, 74, 160, 171	0
1	M	525/547 (95%)	-0.26	0 100 100	12, 50, 109, 121	0
1	N	525/547 (95%)	-0.21	1 (0%) 95 91	14, 58, 133, 145	0
1	O	525/547 (95%)	-0.11	5 (0%) 82 71	20, 80, 164, 176	0
1	P	525/547 (95%)	-0.17	2 (0%) 92 87	19, 68, 149, 161	0
1	Q	525/547 (95%)	-0.15	7 (1%) 77 65	16, 82, 162, 174	0
1	R	525/547 (95%)	-0.10	6 (1%) 80 69	20, 73, 174, 186	0
1	S	525/547 (95%)	-0.13	5 (0%) 82 71	21, 75, 155, 167	0
1	T	525/547 (95%)	-0.20	3 (0%) 89 81	18, 71, 141, 153	0
1	U	525/547 (95%)	-0.14	5 (0%) 82 71	21, 78, 159, 171	0
1	V	525/547 (95%)	-0.20	2 (0%) 92 87	15, 68, 148, 160	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	W	525/547 (95%)	-0.17	5 (0%) 82 71	16, 73, 153, 165	0
1	X	525/547 (95%)	-0.27	0 100 100	11, 50, 112, 124	0
1	Y	525/547 (95%)	-0.25	1 (0%) 95 91	11, 60, 142, 154	0
1	Z	525/547 (95%)	-0.12	9 (1%) 70 57	18, 67, 161, 173	0
All	All	14700/15316 (95%)	-0.18	97 (0%) 87 79	11, 67, 155, 186	0

The worst 5 of 97 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	223	ALA	5.1
1	R	203	TYR	4.5
1	F	268	ARG	4.3
1	B	223	ALA	4.2
1	K	268	ARG	4.2

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.