



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 02:12 am BST

PDB ID : 3F3G
Title : Crystal structure of the nucleoporin pair Nup85-Seh1, space group P212121
Authors : Debler, E.W.; Hseo, H.; Ma, Y.; Blobel, G.; Hoelz, A.
Deposited on : 2008-10-30
Resolution : 3.75 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 ⓘ symbol.

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	:	2.11
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)	:	2.11

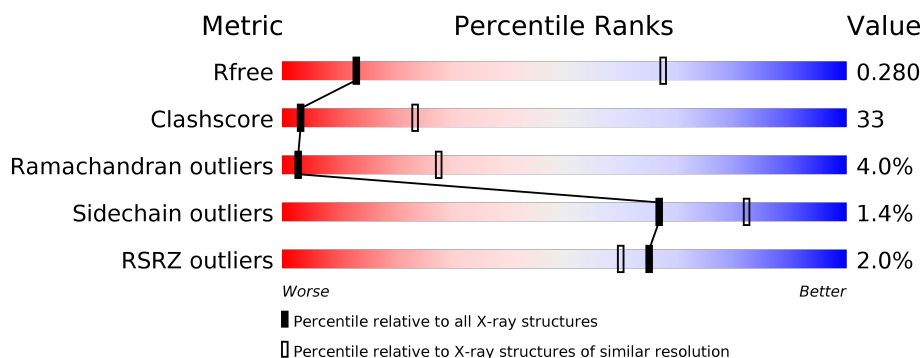
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.75 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	1039 (3.94-3.58)
Clashscore	141614	1051 (3.92-3.60)
Ramachandran outliers	138981	1015 (3.92-3.60)
Sidechain outliers	138945	1011 (3.92-3.60)
RSRZ outliers	127900	1050 (3.96-3.56)

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 $\leq 5\%$. 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	351	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 41%, yellow 43%, orange 15%, grey 15%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> % 41% 43% • 15% </div> </div>
1	B	351	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, green 38%, yellow 45%, orange 15%, grey 15%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> 2% 38% 45% • 15% </div> </div>
1	E	351	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, green 38%, yellow 45%, orange 15%, grey 15%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> 3% 38% 45% • 15% </div> </div>
1	F	351	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, green 40%, yellow 43%, orange 15%, grey 15%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> 3% 40% 43% • 15% </div> </div>
2	C	570	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, green 39%, yellow 43%, orange 14%, grey 14%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> 2% 39% 43% • 14% </div> </div>
2	D	570	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 39%, yellow 42%, orange 16%, grey 16%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> % 39% 42% • 16% </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	570	<div><div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>40%41%•15%</div></div></div>
2	H	570	<div><div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>39%43%•14%</div></div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 25114 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 Nucleoporin SEH1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	300	Total	C	N	O	S	0	0	0
			2379	1506	410	452	11			
1	B	300	Total	C	N	O	S	0	0	0
			2379	1506	410	452	11			
1	E	300	Total	C	N	O	S	0	0	0
			2379	1506	410	452	11			
1	F	300	Total	C	N	O	S	0	0	0
			2379	1506	410	452	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	PRO	-	EXPRESSION TAG	UNP P53011
A	0	HIS	-	EXPRESSION TAG	UNP P53011
B	-1	PRO	-	EXPRESSION TAG	UNP P53011
B	0	HIS	-	EXPRESSION TAG	UNP P53011
E	-1	PRO	-	EXPRESSION TAG	UNP P53011
E	0	HIS	-	EXPRESSION TAG	UNP P53011
F	-1	PRO	-	EXPRESSION TAG	UNP P53011
F	0	HIS	-	EXPRESSION TAG	UNP P53011

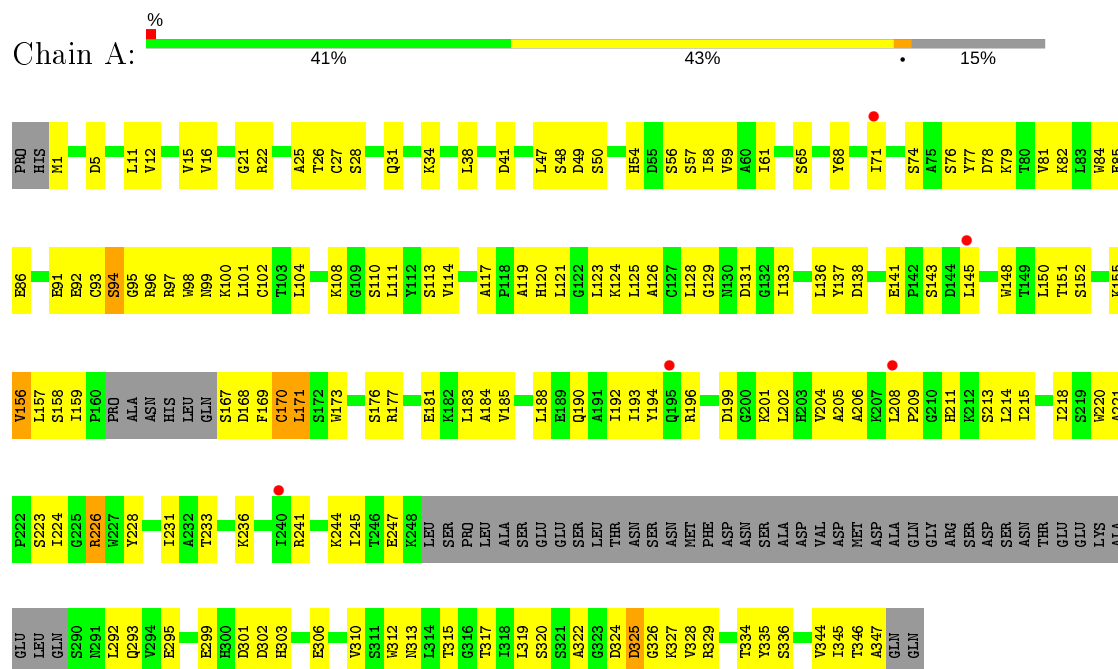
- Molecule 2 is a protein called Nucleoporin NUP85.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	492	Total	C	N	O	S	0	0	0
			3938	2526	634	756	22			
2	D	480	Total	C	N	O	S	0	0	0
			3850	2473	618	737	22			
2	G	482	Total	C	N	O	S	0	0	0
			3863	2482	620	739	22			
2	H	493	Total	C	N	O	S	0	0	0
			3947	2532	636	757	22			

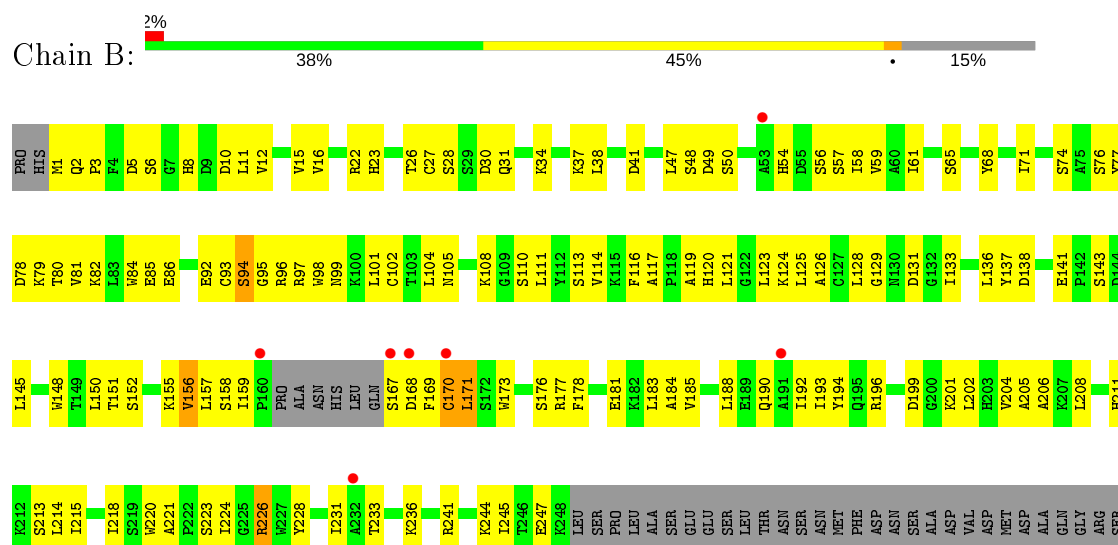
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Nucleoporin SEH1

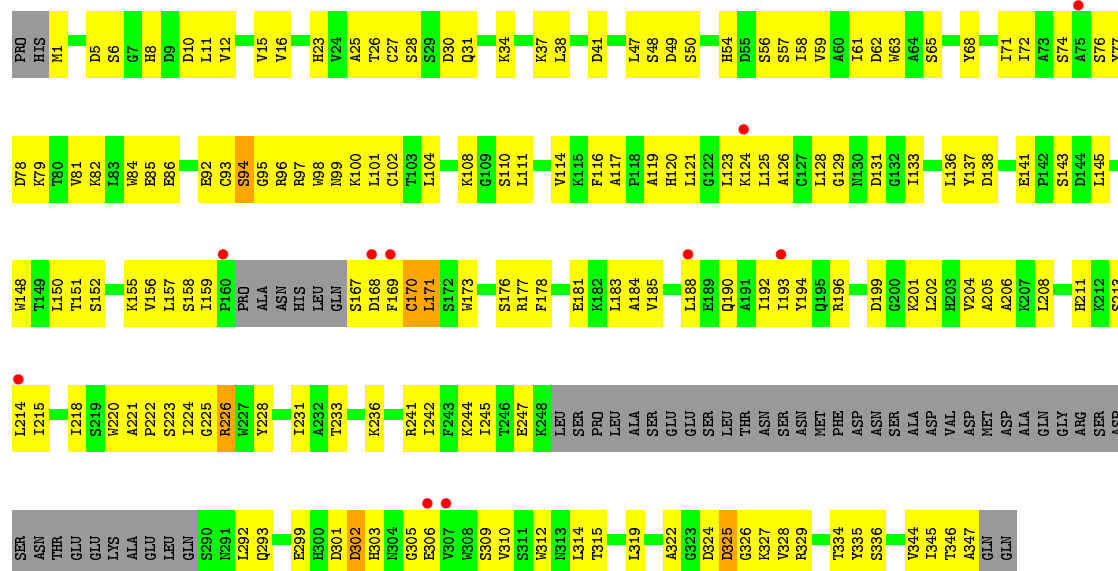


• Molecule 1: Nucleoporin SEH1

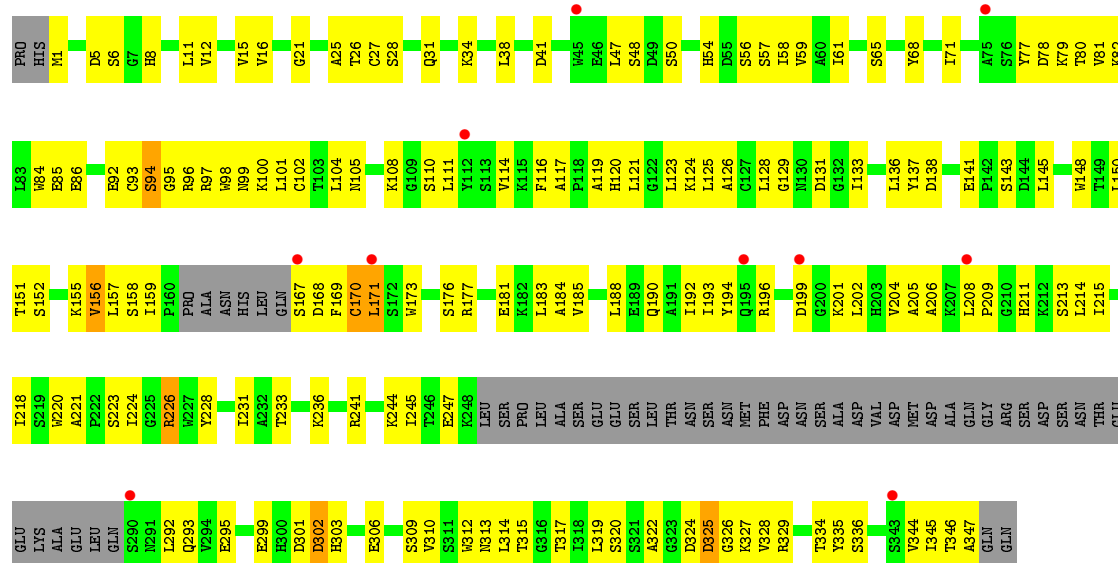




• Molecule 1: Nucleoporin SEH1

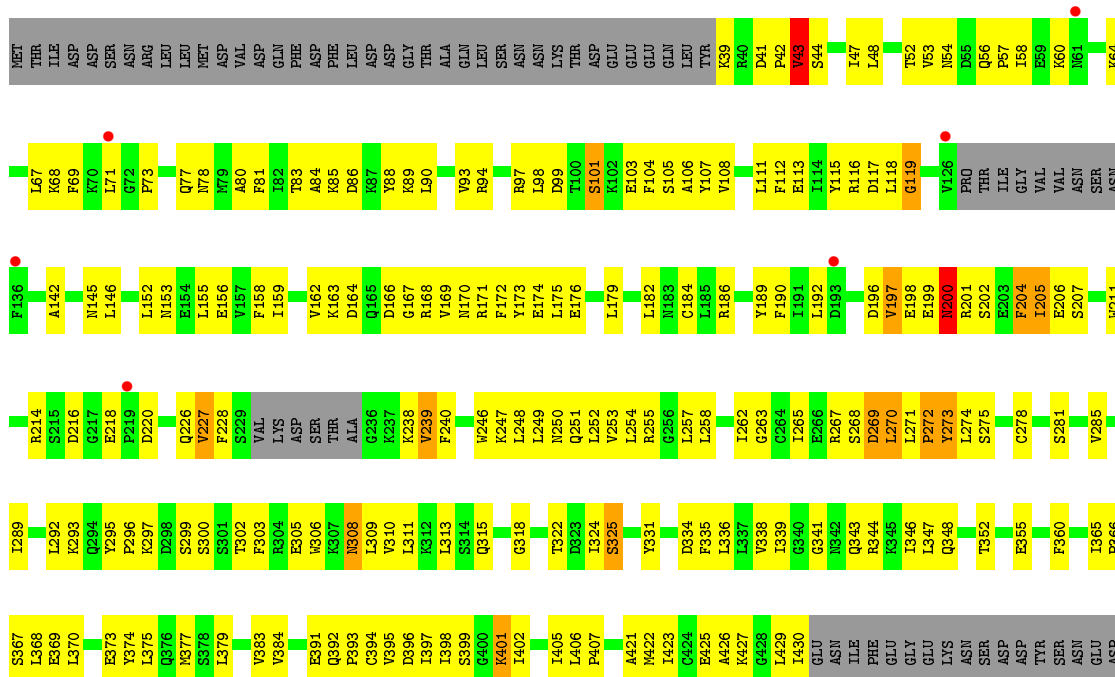


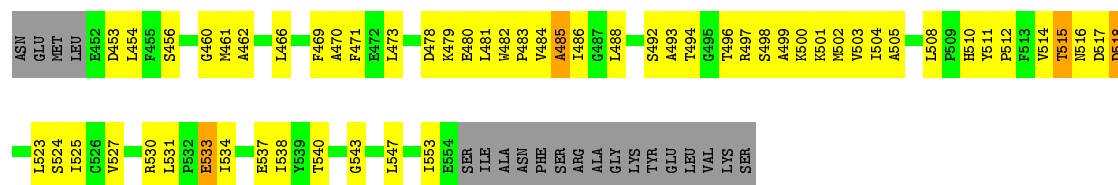
• Molecule 1: Nucleoporin SEH1



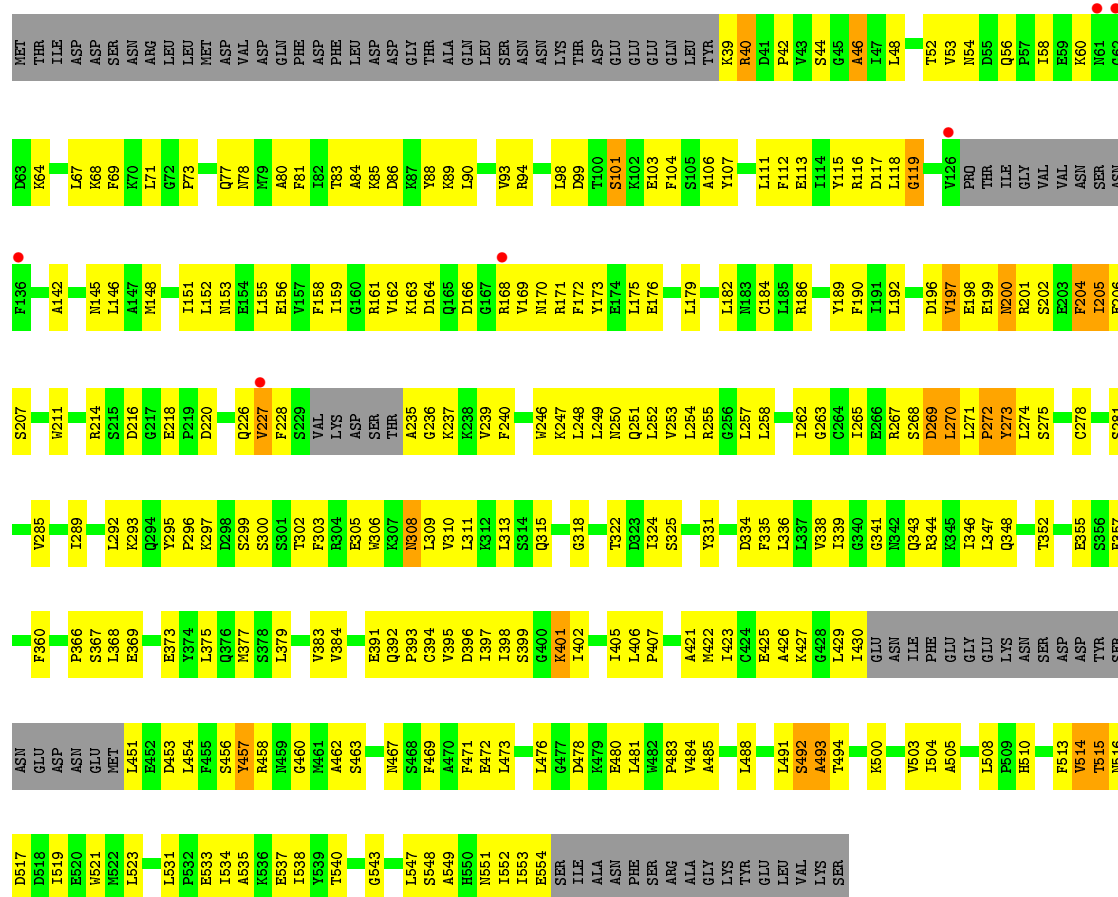
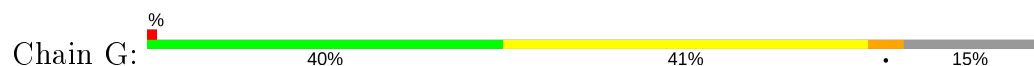
• Molecule 2: Nucleoporin NUP85



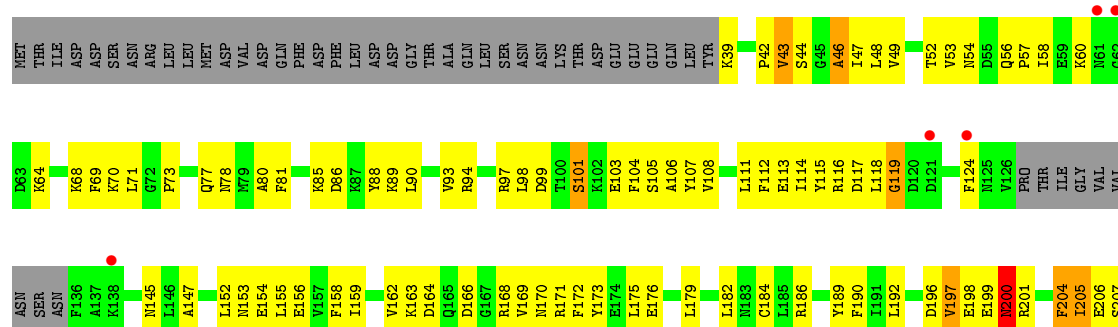




• Molecule 2: Nucleoporin NUP85



• Molecule 2: Nucleoporin NUP85



M522	L523	S524	C526	W527	R530	L531	P532	E533	I534	A535	K536	E537	I538	T541	Q545	M546	L547	S548	A549	M550	M551	I552	I553	E554	S555	I556	A557	M558	R561	ALA	GLY	LYS	TYR	GLU	LEU	VAL	LYS	SER																	
F455	S456	Y457	R458	M459	G460	M461	M465	L466	M467	S468	F469	A470	F471	E472	L473	G474	S475	L476	G477	D478	K479	E480	I481	K482	P483	V484	A485	I486	G487	L488	I489	A490	L491	S492	A493	T494	G495	T496	R497	S498	A499	K500													
M377	S378	L379	V383	G384	V384	E391	Q392	P393	C394	V395	D396	I397	I398	S399	G400	K401	I402	I405	L406	P407	C415	A421	M422	I423	G424	E425	A426	K427	G428	L429	I430	GLU	ASN	ILE	PHE	GLU	GLY	GLU	LYS	ASN	SER	ASP	ASP	TYR	SER	ASN	GLU	ASP	ASN	GLU	MET	LEU	E452	D453	L454
L289	L209	M210	W211	R214	S215	D216	G217	E218	P219	D220	Q226	V227	F228	K231	D232	K238	V239	F240	F245	W246	R247	L248	L249	N250	Q251	L252	V253	L254	R255	G256	L257	L258	I262	G263	C264	I265	E266	R267	S268	D269	L270	L271	P272	Y273	L274	S275	C278	S281	V285						
L289	L292	K293	Q294	Y295	P296	S299	T302	E305	W306	K307	N308	L309	V310	L313	T322	D323	I324	S325	Y331	D334	F335	L336	I337	V338	I339	G340	G341	N342	Q343	R344	K345	I346	T352	E355	F360	L361	L362	P366	S367	L368	E369	L370	E373	Y374	Q376										

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.38Å 106.49Å 358.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.75 47.52 – 3.53	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.75) 86.4 (47.52-3.53)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.18	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 3.57Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.243 , 0.272 0.254 , 0.280	Depositor DCC
R_{free} test set	2062 reflections (4.47%)	wwPDB-VP
Wilson B-factor (Å ²)	107.4	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 78.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.337 for k,h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	25114	wwPDB-VP
Average B, all atoms (Å ²)	127.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	A	0.40	0/2437	0.64	0/3301
1	B	0.39	0/2437	0.63	0/3301
1	E	0.39	0/2437	0.64	0/3301
1	F	0.40	0/2437	0.64	0/3301
2	C	0.58	2/4018 (0.0%)	0.67	1/5440 (0.0%)
2	D	0.53	1/3928 (0.0%)	0.67	2/5316 (0.0%)
2	G	0.50	1/3941 (0.0%)	0.66	1/5334 (0.0%)
2	H	0.59	3/4027 (0.1%)	0.67	1/5451 (0.0%)
All	All	0.50	7/25662 (0.0%)	0.66	5/34745 (0.0%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	39	LYS	CD-CE	5.48	1.65	1.51
2	G	39	LYS	CD-CE	5.47	1.65	1.51
2	H	43	VAL	CB-CG2	5.46	1.64	1.52
2	D	39	LYS	CD-CE	5.43	1.64	1.51
2	C	552	ILE	CA-CB	5.22	1.66	1.54

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	41	ASP	CB-CG-OD1	5.93	123.64	118.30
2	C	43	VAL	CB-CA-C	-5.84	100.30	111.40
2	D	43	VAL	N-CA-C	5.59	126.11	111.00
2	G	40	ARG	CB-CA-C	-5.31	99.77	110.40
2	H	39	LYS	CD-CE-NZ	5.21	123.68	111.70

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	2379	0	2319	165	0
1	B	2379	0	2319	165	0
1	E	2379	0	2319	170	0
1	F	2379	0	2319	167	0
2	C	3938	0	3892	274	0
2	D	3850	0	3807	257	0
2	G	3863	0	3823	251	0
2	H	3947	0	3905	279	0
All	All	25114	0	24703	1642	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 33.

The worst 5 of 1642 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:43:VAL:HG13	2:C:48:LEU:HD22	1.27	1.10
2:D:42:PRO:HB3	2:D:48:LEU:HD22	1.37	1.04
2:H:42:PRO:HB3	2:H:48:LEU:HD22	1.42	0.99
2:H:42:PRO:HD3	2:H:48:LEU:HB2	1.40	0.99
2:D:42:PRO:CG	2:D:48:LEU:HB2	1.94	0.98

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	294/351 (84%)	240 (82%)	48 (16%)	6 (2%)	7	41
1	B	294/351 (84%)	243 (83%)	45 (15%)	6 (2%)	7	41
1	E	294/351 (84%)	242 (82%)	47 (16%)	5 (2%)	9	43
1	F	294/351 (84%)	242 (82%)	46 (16%)	6 (2%)	7	41
2	C	486/570 (85%)	378 (78%)	82 (17%)	26 (5%)	2	22
2	D	472/570 (83%)	371 (79%)	76 (16%)	25 (5%)	2	22
2	G	474/570 (83%)	370 (78%)	79 (17%)	25 (5%)	2	22
2	H	487/570 (85%)	384 (79%)	78 (16%)	25 (5%)	2	23
All	All	3095/3684 (84%)	2470 (80%)	501 (16%)	124 (4%)	3	28

5 of 124 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	302	ASP
1	B	302	ASP
2	C	197	VAL
2	C	200	ASN
2	C	270	LEU

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	262/307 (85%)	259 (99%)	3 (1%)	73	85
1	B	262/307 (85%)	259 (99%)	3 (1%)	73	85
1	E	262/307 (85%)	259 (99%)	3 (1%)	73	85
1	F	262/307 (85%)	259 (99%)	3 (1%)	73	85
2	C	438/510 (86%)	431 (98%)	7 (2%)	62	80
2	D	428/510 (84%)	421 (98%)	7 (2%)	62	80
2	G	429/510 (84%)	424 (99%)	5 (1%)	71	84
2	H	439/510 (86%)	430 (98%)	9 (2%)	53	74

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2782/3268 (85%)	2742 (99%)	40 (1%)	67 82

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	530	ARG
1	E	226	ARG
2	H	497	ARG
1	E	5	ASP
1	F	5	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	308	ASN
2	C	545	GLN
2	D	308	ASN
2	G	551	ASN
2	H	308	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 [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, 95th 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 ²)	Q<0.9
1	A	300/351 (85%)	0.31	5 (1%) 70 65	97, 126, 151, 158	0
1	B	300/351 (85%)	0.30	8 (2%) 54 47	96, 125, 152, 160	0
1	E	300/351 (85%)	0.25	10 (3%) 46 40	97, 126, 152, 160	0
1	F	300/351 (85%)	0.33	10 (3%) 46 40	98, 126, 151, 158	0
2	C	492/570 (86%)	0.08	13 (2%) 56 49	90, 124, 170, 179	0
2	D	480/570 (84%)	0.02	6 (1%) 77 73	92, 122, 167, 180	0
2	G	482/570 (84%)	0.06	6 (1%) 79 75	92, 123, 167, 179	0
2	H	493/570 (86%)	0.16	6 (1%) 79 75	91, 125, 170, 183	0
All	All	3147/3684 (85%)	0.16	64 (2%) 65 60	90, 124, 162, 183	0

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	126	VAL	6.5
2	G	61	ASN	4.9
2	C	124	PHE	4.8
1	F	195	GLN	3.9
2	H	121	ASP	3.9

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.