



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 17, 2024 – 12:53 PM EST

PDB ID : 3S51  
Title : Structure of FANCI  
Authors : Pavletich, N.P.  
Deposited on : 2011-05-20  
Resolution : 3.30 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 : 2.36  
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) : 2.36

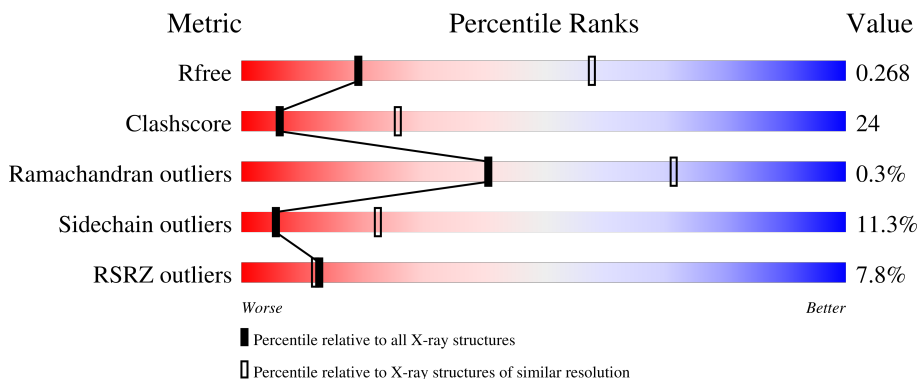
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	A	1308	
1	B	1308	
1	C	1308	
1	D	1308	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 34594 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 Fanconi anemia group I protein homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1134	8960	5762	1489	1656	53	0	0	0
1	B	1071	8487	5468	1409	1559	51	0	0	0
1	C	1134	8960	5762	1489	1656	53	0	0	0
1	D	1034	8187	5277	1358	1504	48	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1303	HIS	-	expression tag	UNP Q8K368
A	1304	HIS	-	expression tag	UNP Q8K368
A	1305	HIS	-	expression tag	UNP Q8K368
A	1306	HIS	-	expression tag	UNP Q8K368
A	1307	HIS	-	expression tag	UNP Q8K368
A	1308	HIS	-	expression tag	UNP Q8K368
B	1303	HIS	-	expression tag	UNP Q8K368
B	1304	HIS	-	expression tag	UNP Q8K368
B	1305	HIS	-	expression tag	UNP Q8K368
B	1306	HIS	-	expression tag	UNP Q8K368
B	1307	HIS	-	expression tag	UNP Q8K368
B	1308	HIS	-	expression tag	UNP Q8K368
C	1303	HIS	-	expression tag	UNP Q8K368
C	1304	HIS	-	expression tag	UNP Q8K368
C	1305	HIS	-	expression tag	UNP Q8K368
C	1306	HIS	-	expression tag	UNP Q8K368
C	1307	HIS	-	expression tag	UNP Q8K368
C	1308	HIS	-	expression tag	UNP Q8K368
D	1303	HIS	-	expression tag	UNP Q8K368
D	1304	HIS	-	expression tag	UNP Q8K368
D	1305	HIS	-	expression tag	UNP Q8K368

*Continued on next page...*

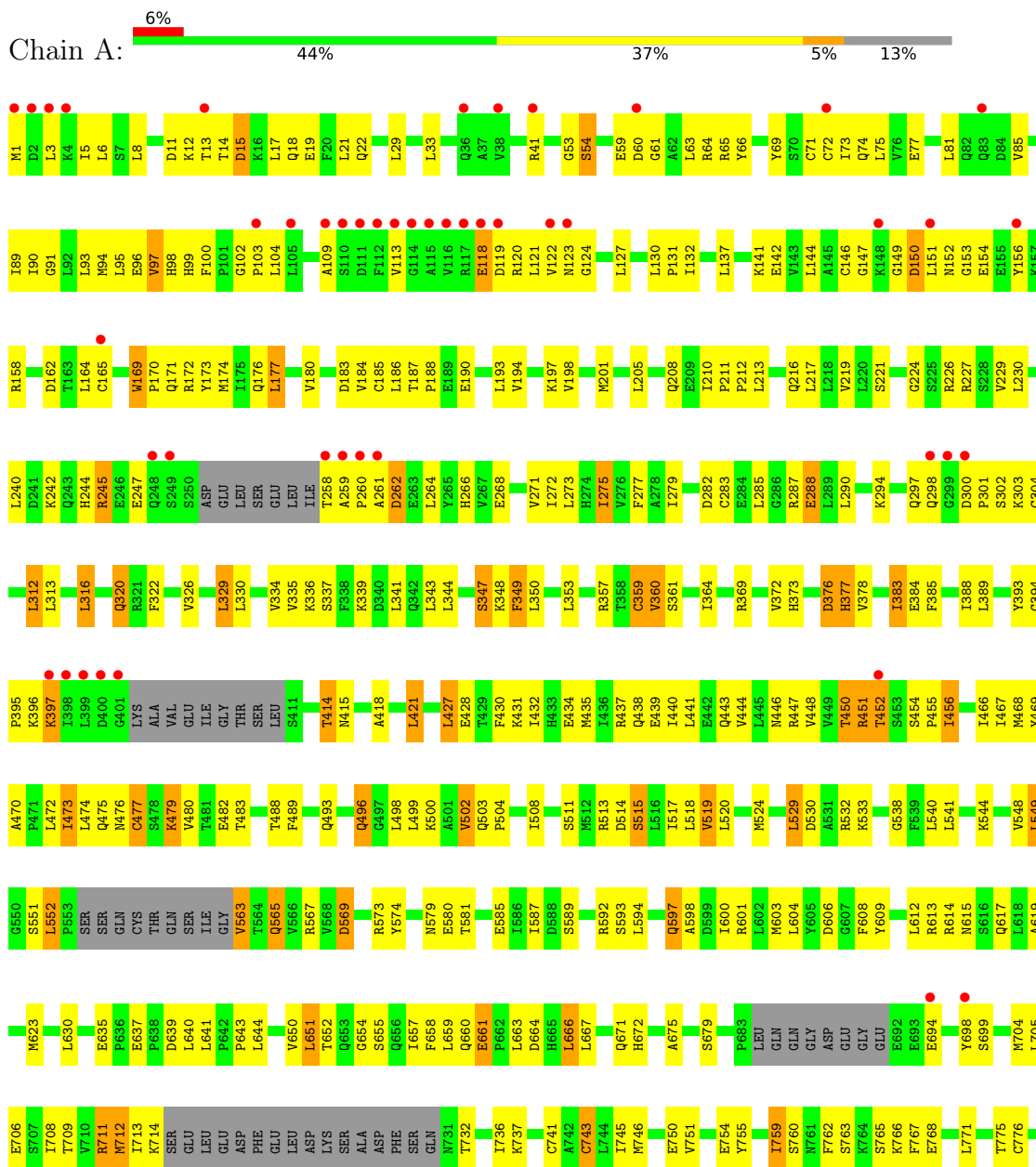
*Continued from previous page...*

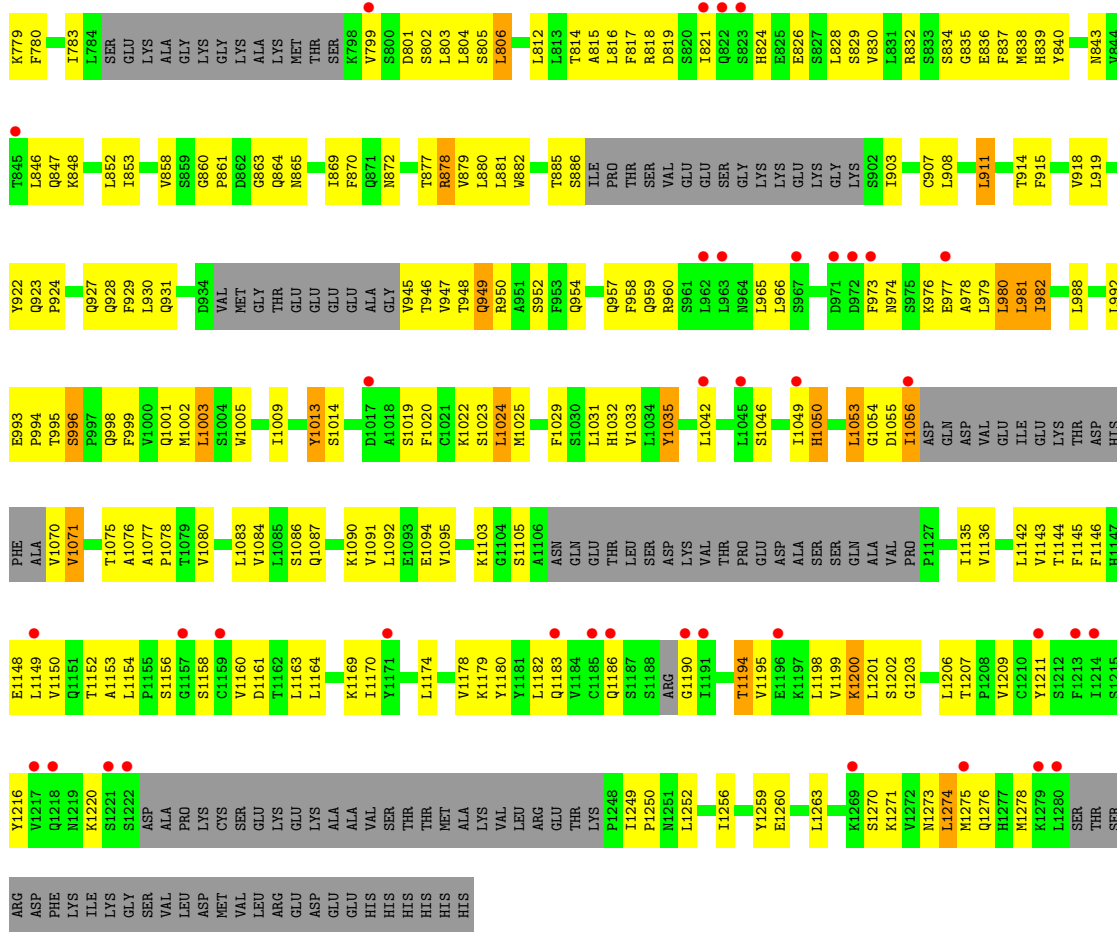
<b>Chain</b>	<b>Residue</b>	<b>Modelled</b>	<b>Actual</b>	<b>Comment</b>	<b>Reference</b>
D	1306	HIS	-	expression tag	UNP Q8K368
D	1307	HIS	-	expression tag	UNP Q8K368
D	1308	HIS	-	expression tag	UNP Q8K368

### 3 Residue-property plots

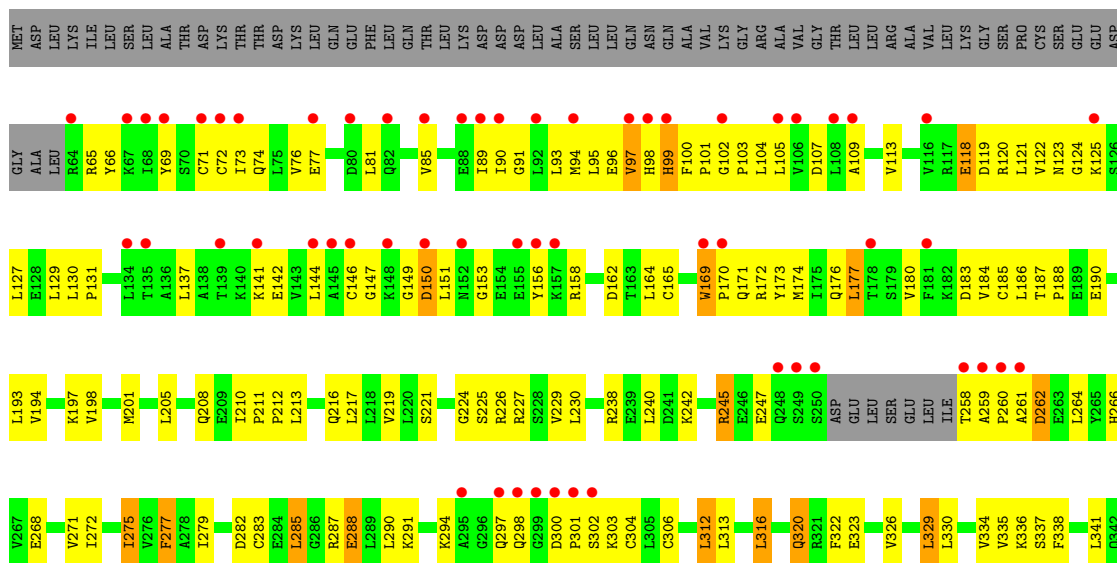
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

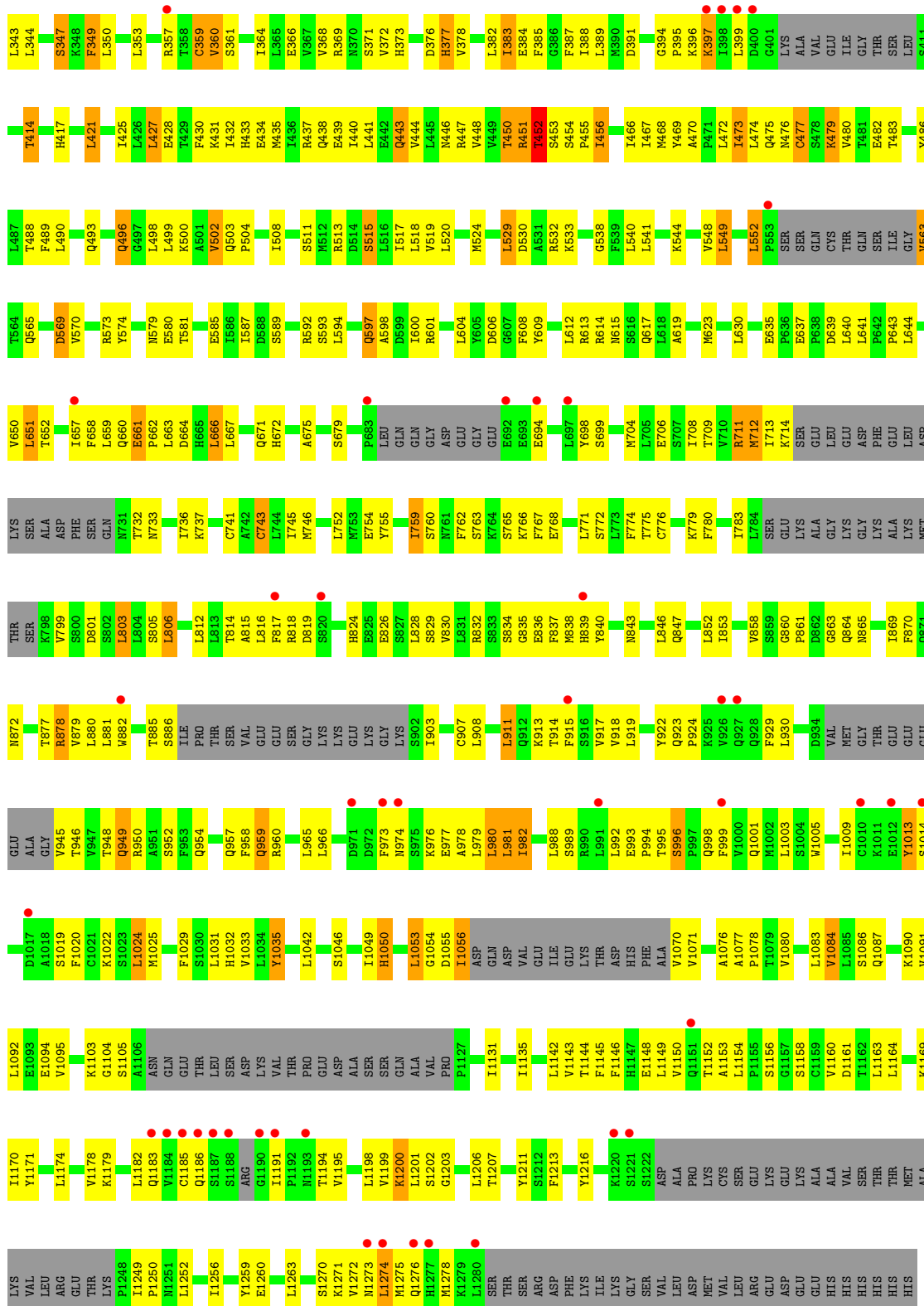
- Molecule 1: Fanconi anemia group I protein homolog



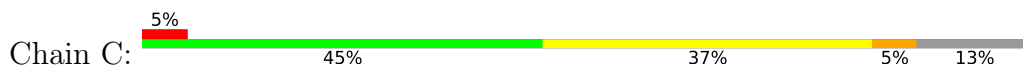


● Molecule 1: Fanconi anemia group I protein homolog



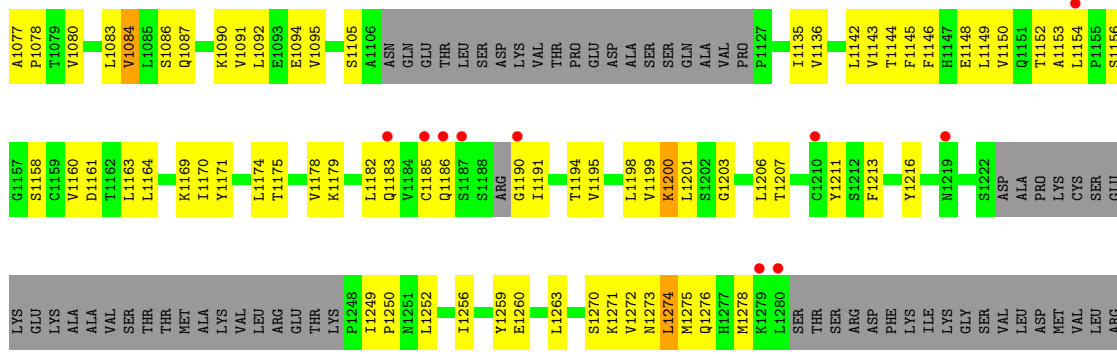


• Molecule 1: Fanconi anemia group I protein homolog

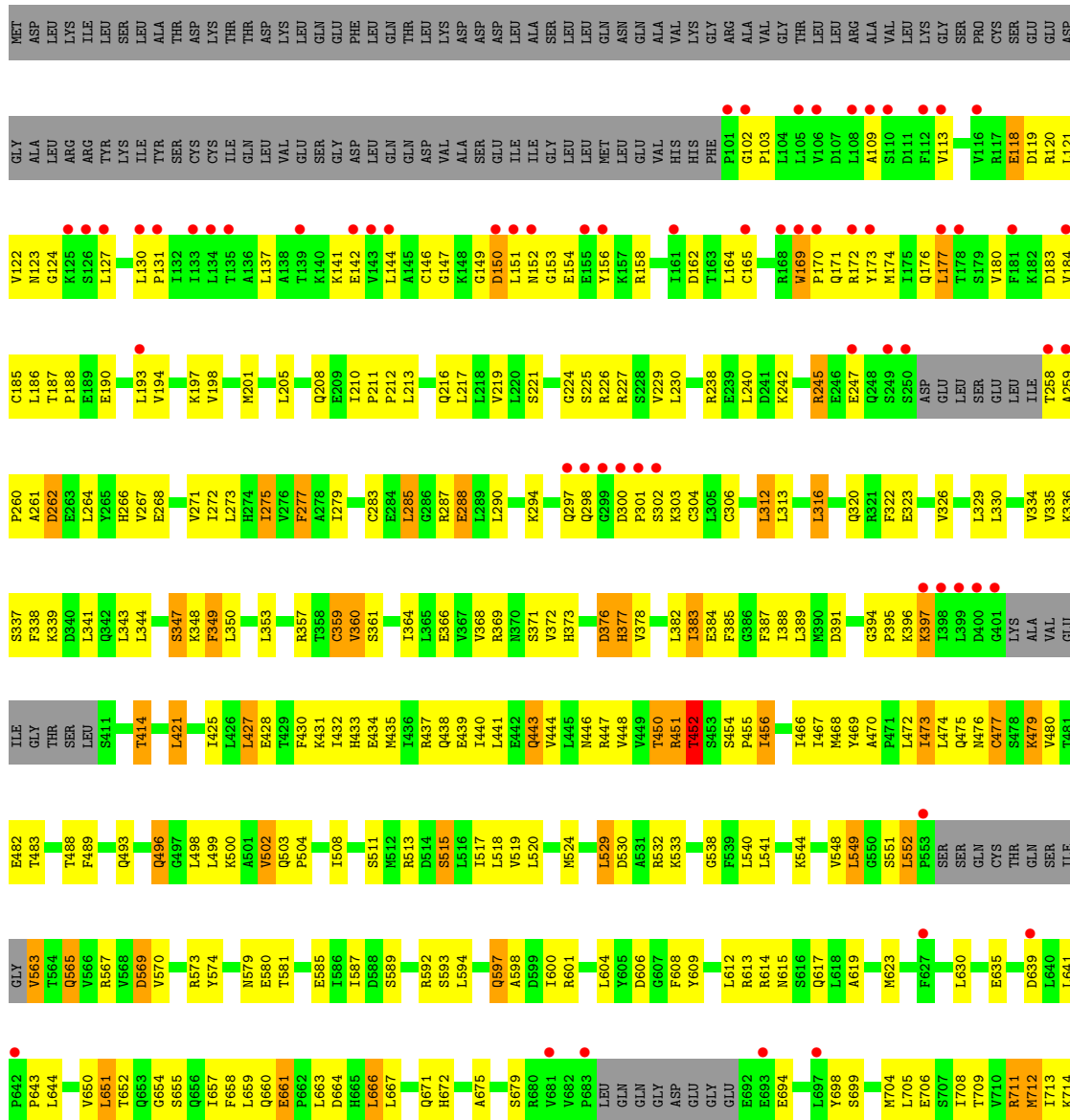








• Molecule 1: Fanconi anemia group I protein homolog



SER	GLU	GLU	LEU	GLY	GLY	ASP	PHE	GLU	LEU	ASP	LYS	SER	ALA	ASP	PHE	SER	GLN	N731	T732	N733	I736	K737	L741	A742	A815	C743	L744	I745	M746	V751	L752	M753	E754	Y755	I759	S760	M761	F762	S763	K764	S765	K766	F767	E768	L771	S772	T775	C776	K779	F780	I783	L784	SER	GLU					
LYS	ALA	GLY	GLY	LYS	LYS	LYS	ALA	LEU	LYS	MET	THR	SER	ALA	K798	V799	S800	S801	S802	L803	L804	S805	L806	L812	L813	T814	A815	L816	F817	R818	D819	H824	E825	E826	S827	L828	S829	L830	L831	R832	S833	S834	G835	E836	F837	M838	H839	Y840	N843	L846	Q847	L852	I853	V858	P861	D862				
G863	Q864	M865	I869	F870	G871	M872	T877	R878	V879	R950	L880	L881	M882	T885	S886	IIF	PRO	THR	SER	VAL	GLU	GLY	SER	GLY	LYS	LYS	GLU	LYS	LYS	LYS	LYS	LYS	I903	I903	C907	L908	L911	T914	F915	V916	L919	Y922	Q923	P924	F929	L930	Q931	I934	VAL	MET	GLY	THR							
GLU	GLU	GLU	GLU	ALA	GLY	V945	T946	T947	Q949	A951	S952	Q954	Q957	F958	Q959	R960	L965	L966	D971	D972	F973	N974	S975	E977	A978	L979	L980	L981	I982	L988	R989	R990	L991	L992	E993	P994	T995	S996	P997	Q998	F999	V1000	M1002	L1003	S1004	H1005	T1009	C1010	K1011										
E1012	Y1013	S1014	Q1015	E1016	S1019	K1022	S1023	L1024	M1025	F1029	S1030	L1031	V1032	L1034	Y1035	K1036	S1037	S1046	H1050	L1053	G1054	D1055	I1056	ASP	GLN	ASP	VAL	GLU	I1E	GLU	LYS	THR	ASP	HIS	PHE	ALA	V1070	V1071	A1076	A1077	P1078	V1079	V1080	L1083	V1084	L1085	S1086	Q1087	K1090										
V1091	L1092	E1093	E1094	V1095	K1103	G1104	S1105	A1106	ASN	GLN	GLU	THR	LEU	SER	ASP	LYS	VAL	THR	PRO	GLU	ASP	ALA	SER	SER	GLN	ALA	VAL	PRO	P1127	T1128	L1129	L1130	I1131	I1135	L1142	V1143	T1144	F1145	F1146	H1147	E1148	L1149	V1150	Q1151	T1152	A1153	L1154	P1155	S1156	G1157	S1158	C1159	V1160	D1161	T1162	L1163			
L1164	K1169	I1170	Y1171	L1174	T1175	V1178	K1179	Y1180	Y1181	L1182	Q1183	V1184	Q1185	Q1186	S1187	S1188	ARG	G1190	I1191	T1194	V1195	L1198	V1199	K1200	L1201	S1202	G1203	L1206	T1207	Y1211	S1212	F1213	Y1216	V1217	S1222	ASP	ALA	PRO	LYS	CYS	SER	GLU	LYS	ARG	GLY	VAL	LEU	ASP	MET	VAL	LEU	ARG	GLU	ASP	GLY	ALA	ALA	VAL	THR
THR	MET	ALA	LYS	VAL	LEU	ARG	GLU	THR	LYS	P1248	I1249	P1250	M1251	L1252	I1256	Y1259	E1260	L1263	I1264	S1270	K1271	V1272	M1273	L1274	M1275	Q1276	H1277	M1278	K1279	L1280	SER	THR	SER	ARG	PHE	LYS	ILE	LYS	GLY	SER	VAL	LEU	ASP	MET	VAL	LEU	ARG	GLU	ASP	GLY	ALA	ALA	VAL	HIS	HIS	HIS			

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.70Å 136.50Å 149.70Å 115.90° 106.00° 95.00°	Depositor
Resolution (Å)	39.82 – 3.30 39.82 – 3.28	Depositor EDS
% Data completeness (in resolution range)	82.8 (39.82-3.30) 83.1 (39.82-3.28)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.52 (at 3.25Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.259 , 0.278 0.245 , 0.268	Depositor DCC
$R_{free}$ test set	2199 reflections (2.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	93.7	Xtrriage
Anisotropy	0.486	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 117.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	34594	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	192.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% 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](#)

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	0.26	2/9099 (0.0%)	0.46	0/12286
1	B	0.29	4/8624 (0.0%)	0.46	0/11646
1	C	0.26	2/9099 (0.0%)	0.46	0/12286
1	D	0.26	2/8319 (0.0%)	0.46	0/11234
All	All	0.27	10/35141 (0.0%)	0.46	0/47452

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	A	0	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	959	GLN	CD-NE2	-9.98	1.07	1.32
1	B	959	GLN	CD-OE1	-7.80	1.06	1.24
1	C	320	GLN	CD-NE2	-6.76	1.16	1.32
1	D	320	GLN	CD-NE2	-6.59	1.16	1.32
1	A	320	GLN	CD-NE2	-6.15	1.17	1.32

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1053	LEU	Peptide
1	B	1053	LEU	Peptide
1	C	1053	LEU	Peptide
1	D	1053	LEU	Peptide

## 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	A	8960	0	9275	454	0
1	B	8487	0	8776	449	0
1	C	8960	0	9275	435	0
1	D	8187	0	8476	411	0
All	All	34594	0	35802	1705	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 24.

The worst 5 of 1705 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:C:799:VAL:O	1:C:847:GLN:NE2	1.87	1.07
1:A:799:VAL:O	1:A:847:GLN:NE2	1.88	1.07
1:B:799:VAL:O	1:B:847:GLN:NE2	1.88	1.06
1:D:799:VAL:O	1:D:847:GLN:NE2	1.88	1.05
1:A:489:PHE:HB3	1:C:450:THR:HG21	1.35	1.04

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	1108/1308 (85%)	1023 (92%)	82 (7%)	3 (0%)	41	71
1	B	1045/1308 (80%)	969 (93%)	72 (7%)	4 (0%)	34	66
1	C	1108/1308 (85%)	1022 (92%)	83 (8%)	3 (0%)	41	71
1	D	1008/1308 (77%)	933 (93%)	71 (7%)	4 (0%)	34	66
All	All	4269/5232 (82%)	3947 (92%)	308 (7%)	14 (0%)	41	71

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	VAL
1	A	150	ASP
1	B	122	VAL
1	B	150	ASP
1	C	122	VAL

### 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	1032/1188 (87%)	917 (89%)	115 (11%)	6	23
1	B	979/1188 (82%)	867 (89%)	112 (11%)	5	22
1	C	1032/1188 (87%)	917 (89%)	115 (11%)	6	23
1	D	945/1188 (80%)	838 (89%)	107 (11%)	6	22
All	All	3988/4752 (84%)	3539 (89%)	449 (11%)	6	22

5 of 449 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	81	LEU
1	D	1142	LEU
1	C	549	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	1050	HIS
1	D	587	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	672	HIS
1	C	377	HIS
1	D	847	GLN
1	D	433	HIS
1	D	672	HIS

### 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	A	1134/1308 (86%)	0.25	85 (7%) 14 13	85, 192, 306, 410	0
1	B	1071/1308 (81%)	0.27	101 (9%) 8 9	87, 188, 306, 415	0
1	C	1134/1308 (86%)	0.25	70 (6%) 20 20	81, 191, 306, 410	0
1	D	1034/1308 (79%)	0.28	84 (8%) 12 11	84, 185, 302, 416	0
All	All	4373/5232 (83%)	0.26	340 (7%) 13 12	81, 189, 304, 416	0

The worst 5 of 340 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	109	ALA	12.3
1	D	250	SER	11.1
1	D	135	THR	10.9
1	D	150	ASP	8.9
1	D	299	GLY	8.7

### 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.