Full wwPDB X-ray Structure Validation Report ⓘ

Jan 2, 2024 – 11:33 pm GMT

PDB ID	:	5FU7
Title	:	drosophila nanos NBR peptide bound to the NOT module of the human CCR4-NOT complex
Authors	:	Raisch, T.; Bhandari, D.; Sabath, K.; Helms, S.; Valkov, E.; Weichenrieder, O.; Izaurralde, E.
Deposited on	:	2016-01-21
Resolution	:	3.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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:

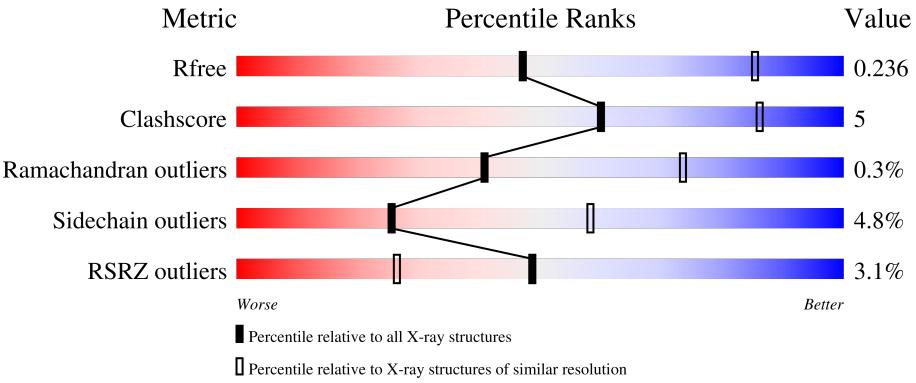
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.10 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.



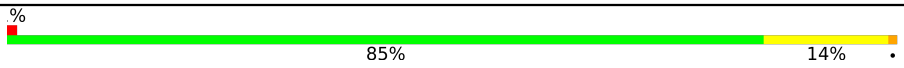
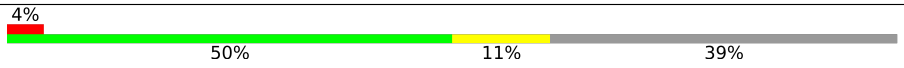
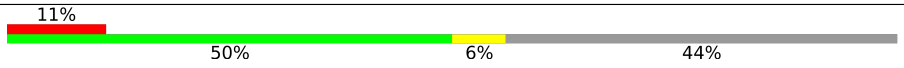
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 >=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% 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	535	% 80% 16%
1	E	535	2% 80% 14%
2	B	197	10% 82% 15%
2	F	197	3% 85% 12%
3	C	148	% 82% 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	148	
4	D	54	
4	H	54	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 14548 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 CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	519	Total	C	N	O	S	0	0	0
			4213	2724	722	744	23			
1	E	514	Total	C	N	O	S	0	0	0
			4178	2703	717	735	23			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1827	GLY	-	expression tag	UNP A5YKK6
A	1828	PRO	-	expression tag	UNP A5YKK6
A	1829	HIS	-	expression tag	UNP A5YKK6
A	1830	MET	-	expression tag	UNP A5YKK6
A	1831	LEU	-	expression tag	UNP A5YKK6
A	1832	GLU	-	expression tag	UNP A5YKK6
A	2344	GLU	HIS	engineered mutation	UNP A5YKK6
A	2345	GLU	CYS	engineered mutation	UNP A5YKK6
A	2346	GLU	ALA	engineered mutation	UNP A5YKK6
E	1827	GLY	-	expression tag	UNP A5YKK6
E	1828	PRO	-	expression tag	UNP A5YKK6
E	1829	HIS	-	expression tag	UNP A5YKK6
E	1830	MET	-	expression tag	UNP A5YKK6
E	1831	LEU	-	expression tag	UNP A5YKK6
E	1832	GLU	-	expression tag	UNP A5YKK6
E	2344	GLU	HIS	engineered mutation	UNP A5YKK6
E	2345	GLU	CYS	engineered mutation	UNP A5YKK6
E	2346	GLU	ALA	engineered mutation	UNP A5YKK6

- Molecule 2 is a protein called CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	191	Total	C	N	O	S	0	0	0
			1570	1012	263	287	8			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	191	Total	C	N	O	S	0	0	0
			1570	1012	263	287	8			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	344	GLY	-	expression tag	UNP Q9NZN8
B	345	PRO	-	expression tag	UNP Q9NZN8
B	346	HIS	-	expression tag	UNP Q9NZN8
B	347	MET	-	expression tag	UNP Q9NZN8
B	348	LEU	-	expression tag	UNP Q9NZN8
B	349	GLU	-	expression tag	UNP Q9NZN8
F	344	GLY	-	expression tag	UNP Q9NZN8
F	345	PRO	-	expression tag	UNP Q9NZN8
F	346	HIS	-	expression tag	UNP Q9NZN8
F	347	MET	-	expression tag	UNP Q9NZN8
F	348	LEU	-	expression tag	UNP Q9NZN8
F	349	GLU	-	expression tag	UNP Q9NZN8

- Molecule 3 is a protein called CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	148	Total	C	N	O	S	0	0	0
			1292	840	216	229	7			
3	G	148	Total	C	N	O	S	0	0	0
			1293	840	216	230	7			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	601	GLY	-	expression tag	UNP O75175
C	602	PRO	-	expression tag	UNP O75175
C	603	HIS	-	expression tag	UNP O75175
C	604	MET	-	expression tag	UNP O75175
C	605	LEU	-	expression tag	UNP O75175
C	606	GLU	-	expression tag	UNP O75175
G	601	GLY	-	expression tag	UNP O75175
G	602	PRO	-	expression tag	UNP O75175
G	603	HIS	-	expression tag	UNP O75175
G	604	MET	-	expression tag	UNP O75175
G	605	LEU	-	expression tag	UNP O75175
G	606	GLU	-	expression tag	UNP O75175

- Molecule 4 is a protein called NANOS, ISOFORM B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	33	Total	C	N	O	S	0	0	0
			238	148	35	53	2			
4	H	30	Total	C	N	O	S	0	0	0
			194	120	31	42	1			

There are 12 discrepancies between the modelled and reference sequences:

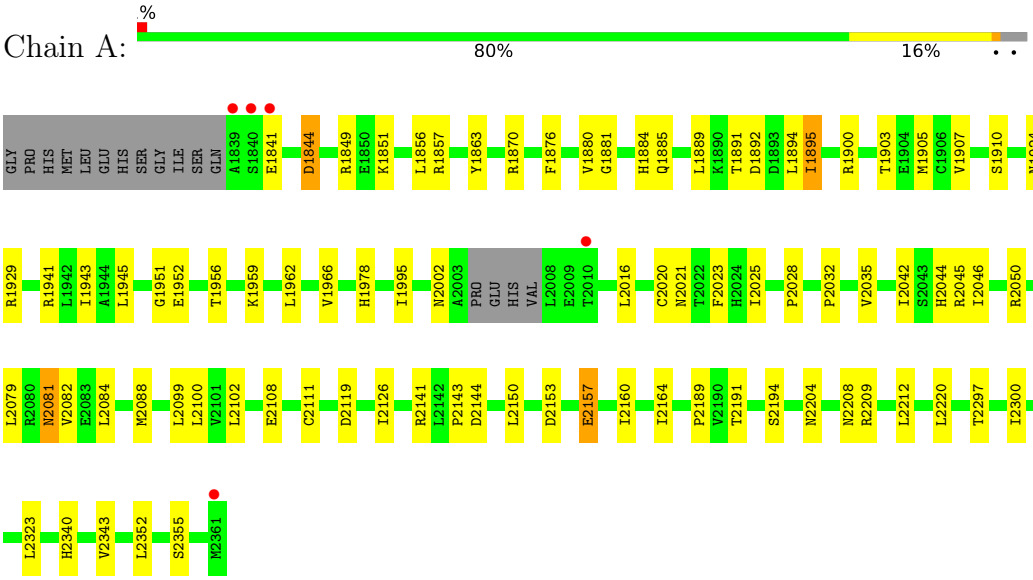
Chain	Residue	Modelled	Actual	Comment	Reference
D	110	GLY	-	expression tag	UNP A0A0B4KGY
D	111	PRO	-	expression tag	UNP A0A0B4KGY
D	112	HIS	-	expression tag	UNP A0A0B4KGY
D	113	MET	-	expression tag	UNP A0A0B4KGY
D	114	LEU	-	expression tag	UNP A0A0B4KGY
D	115	GLU	-	expression tag	UNP A0A0B4KGY
H	110	GLY	-	expression tag	UNP A0A0B4KGY
H	111	PRO	-	expression tag	UNP A0A0B4KGY
H	112	HIS	-	expression tag	UNP A0A0B4KGY
H	113	MET	-	expression tag	UNP A0A0B4KGY
H	114	LEU	-	expression tag	UNP A0A0B4KGY
H	115	GLU	-	expression tag	UNP A0A0B4KGY

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.

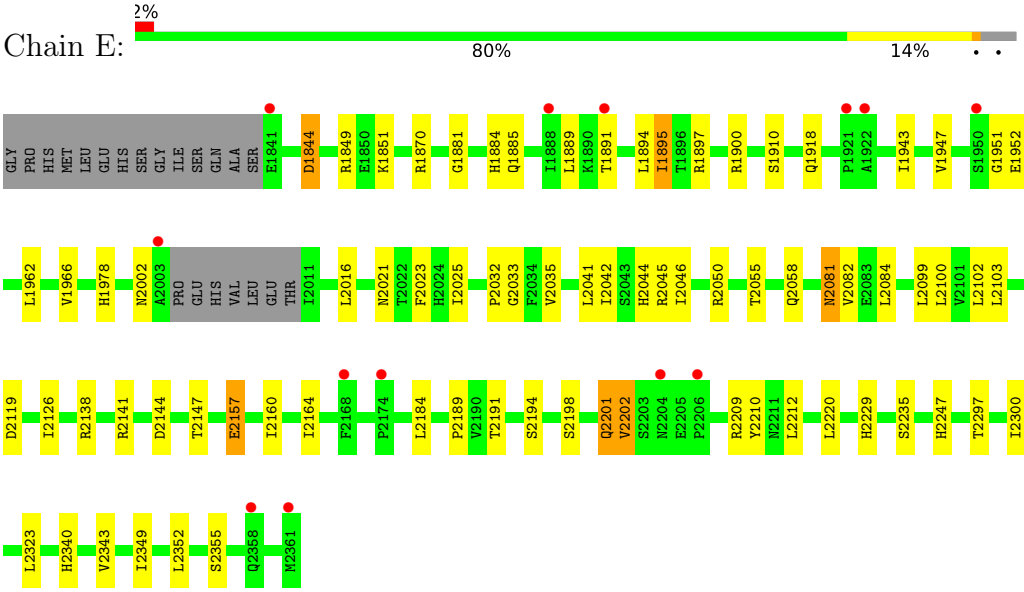
• Molecule 1: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1

Chain A:



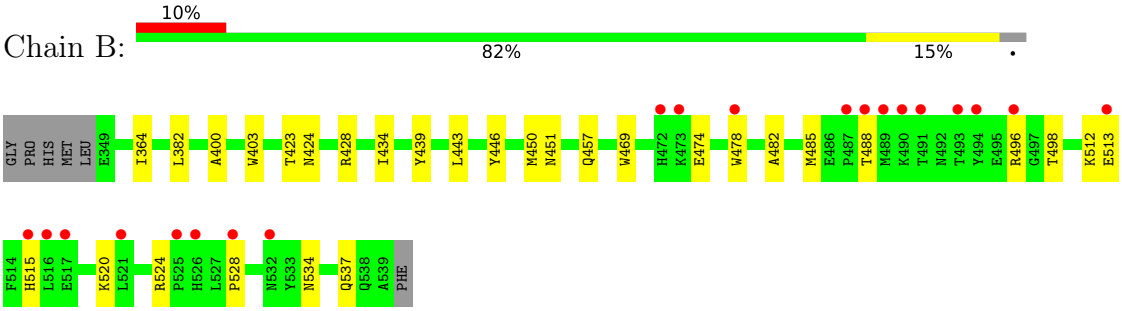
• Molecule 1: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1

Chain E:



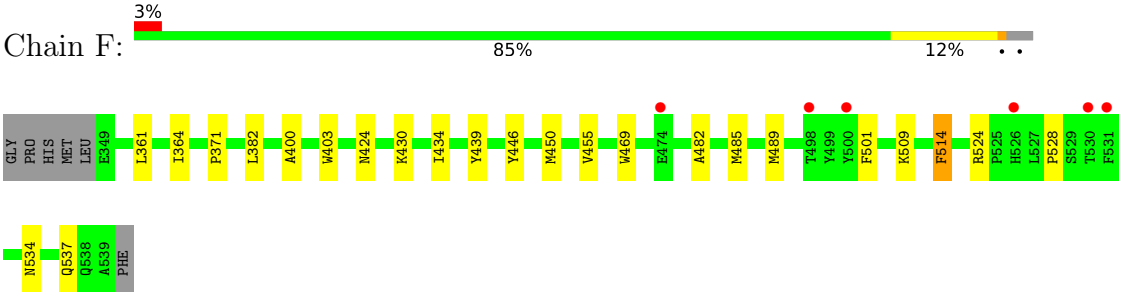
• Molecule 2: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2

Chain B:



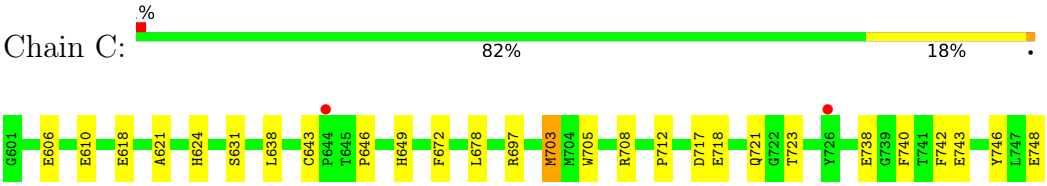
• Molecule 2: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2

Chain F:



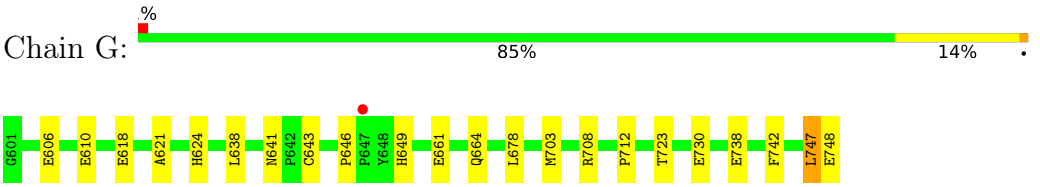
• Molecule 3: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3

Chain C:



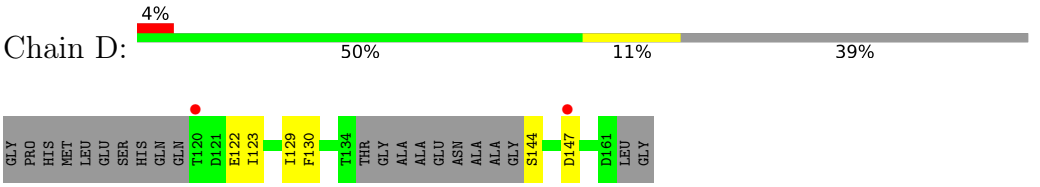
• Molecule 3: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3

Chain G:



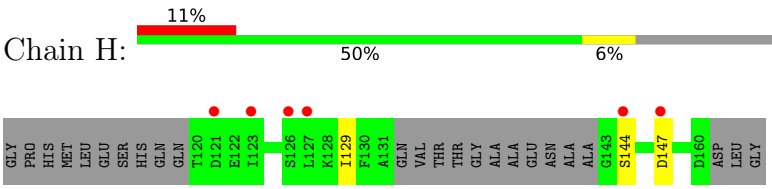
• Molecule 4: NANOS, ISOFORM B

Chain D:



• Molecule 4: NANOS, ISOFORM B

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	75.95Å 135.67Å 104.97Å 90.00° 107.98° 90.00°	Depositor
Resolution (Å)	48.58 – 3.10 48.58 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.58-3.10) 99.7 (48.58-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 3.12Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.165 , 0.227 0.180 , 0.236	Depositor DCC
R_{free} test set	1892 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	82.0	Xtriage
Anisotropy	0.595	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 73.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14548	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.31% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.54	0/4320	0.70	0/5874
1	E	0.53	0/4285	0.69	0/5826
2	B	0.49	0/1617	0.70	0/2198
2	F	0.49	0/1617	0.70	0/2198
3	C	0.49	0/1343	0.64	0/1819
3	G	0.50	0/1344	0.66	0/1819
4	D	0.51	0/239	0.71	0/322
4	H	0.49	0/194	0.67	0/262
All	All	0.52	0/14959	0.69	0/20318

There are no bond length outliers.

There are no bond angle outliers.

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	4213	0	4224	49	0
1	E	4178	0	4190	44	0
2	B	1570	0	1512	13	0
2	F	1570	0	1512	12	0
3	C	1292	0	1207	15	0
3	G	1293	0	1207	12	0
4	D	238	0	208	5	0
4	H	194	0	158	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	14548	0	14218	133	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 5.

All (133) 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:2184:LEU:HD22	1:E:2247:HIS:HD2	1.35	0.92
1:E:1978:HIS:HE1	1:E:2033:GLY:H	1.31	0.79
1:A:1952:GLU:HB2	4:H:129:ILE:HG12	1.72	0.71
2:B:498:THR:HG22	2:B:513:GLU:HB3	1.72	0.71
1:E:2021:ASN:O	1:E:2025:ILE:HG12	1.91	0.71
3:G:638:LEU:HD12	3:G:678:LEU:HD21	1.74	0.70
1:A:2021:ASN:O	1:A:2025:ILE:HG12	1.89	0.70
1:E:2184:LEU:HD22	1:E:2247:HIS:CD2	2.23	0.70
3:C:638:LEU:HD12	3:C:678:LEU:HD21	1.72	0.70
1:A:1884:HIS:HB2	1:A:1889:LEU:HD12	1.76	0.68
1:E:1884:HIS:HB2	1:E:1889:LEU:HD12	1.76	0.67
1:A:1903:THR:O	1:A:1907:VAL:HG23	1.97	0.65
2:F:528:PRO:HG2	3:G:643:CYS:HB2	1.79	0.63
1:A:2081:ASN:HD21	3:C:606:GLU:HB2	1.66	0.60
4:D:129:ILE:HG12	1:E:1952:GLU:HB2	1.84	0.58
2:B:451:ASN:O	2:B:457:GLN:HB2	2.03	0.58
1:E:2023:PHE:HB3	1:E:2042:ILE:HD11	1.85	0.58
1:E:1881:GLY:O	1:E:1885:GLN:HG2	2.03	0.58
2:B:423:THR:HG22	3:C:703:MET:HG2	1.85	0.57
1:A:1943:ILE:HG21	1:A:1966:VAL:HG11	1.86	0.57
1:A:1952:GLU:CB	4:H:129:ILE:HG12	2.35	0.57
1:A:1943:ILE:CG2	1:A:1966:VAL:HG11	2.34	0.56
1:A:2189:PRO:HB2	1:A:2191:THR:HG22	1.86	0.56
1:A:1926:THR:HG22	1:A:1929:ARG:HH22	1.70	0.56
1:E:2081:ASN:HD21	3:G:606:GLU:HB2	1.71	0.56
1:E:1943:ILE:HG21	1:E:1966:VAL:HG11	1.86	0.56
1:A:2309:LEU:HD21	1:A:2352:LEU:HB3	1.88	0.56
3:C:708:ARG:HD3	3:C:712:PRO:HD3	1.87	0.56
1:A:2157:GLU:HA	1:A:2160:ILE:HD12	1.87	0.56
2:F:455:VAL:HG12	3:G:641:ASN:HB3	1.88	0.56
1:A:1881:GLY:O	1:A:1885:GLN:HG2	2.05	0.56
1:A:1863:TYR:HE1	1:A:1941:ARG:HD2	1.71	0.55
1:E:2144:ASP:HB3	1:E:2147:THR:HG23	1.87	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:2189:PRO:HB2	1:E:2191:THR:HG22	1.87	0.55
3:G:708:ARG:HD3	3:G:712:PRO:HD3	1.88	0.54
1:E:1943:ILE:CG2	1:E:1966:VAL:HG11	2.37	0.54
1:E:2002:ASN:HD21	1:E:2050:ARG:HH11	1.55	0.54
2:F:534:ASN:HB3	2:F:537:GLN:HG3	1.89	0.54
1:A:2002:ASN:HD21	1:A:2050:ARG:HH11	1.56	0.53
1:A:1880:VAL:HG11	4:D:130:PHE:HD2	1.73	0.53
1:E:2309:LEU:HD21	1:E:2352:LEU:HB3	1.89	0.53
2:B:534:ASN:HB3	2:B:537:GLN:HG3	1.90	0.53
1:E:2157:GLU:HA	1:E:2160:ILE:HD12	1.90	0.53
3:C:723:THR:HG22	3:C:738:GLU:HA	1.90	0.53
1:A:2044:HIS:HD2	1:A:2046:ILE:HB	1.75	0.51
1:A:2100:LEU:HD21	2:B:382:LEU:HD13	1.91	0.51
1:E:1951:GLY:HA3	1:E:1959:LYS:HE3	1.91	0.51
1:A:2023:PHE:HB3	1:A:2042:ILE:HD11	1.92	0.51
1:E:2100:LEU:HD21	2:F:382:LEU:HD13	1.92	0.51
3:G:723:THR:HG22	3:G:738:GLU:HA	1.93	0.50
1:A:2297:THR:HB	1:A:2300:ILE:HG13	1.94	0.50
1:A:1962:LEU:HD13	4:D:123:ILE:HG21	1.94	0.49
1:E:2099:LEU:HD23	1:E:2102:LEU:HD12	1.94	0.49
2:B:528:PRO:HG2	3:C:643:CYS:HB2	1.93	0.49
3:C:646:PRO:HG2	3:C:649:HIS:HD2	1.77	0.49
1:A:2099:LEU:HD23	1:A:2102:LEU:HD12	1.94	0.49
1:A:2313:ILE:HD11	1:A:2355:SER:HB2	1.95	0.49
2:B:482:ALA:HB3	2:B:485:MET:HB2	1.94	0.49
3:G:646:PRO:HG2	3:G:649:HIS:HD2	1.78	0.48
1:E:2044:HIS:CD2	1:E:2046:ILE:H	2.31	0.48
1:E:2055:THR:HG23	1:E:2058:GLN:HA	1.95	0.48
1:A:2164:ILE:HD13	1:A:2220:LEU:HD21	1.95	0.48
1:E:2313:ILE:HD11	1:E:2355:SER:HB2	1.95	0.48
2:F:469:TRP:CZ2	2:F:509:LYS:HE3	2.49	0.48
1:A:1978:HIS:CE1	1:A:2032:PRO:HD2	2.49	0.47
1:E:2164:ILE:HD13	1:E:2220:LEU:HD21	1.96	0.47
2:F:430:LYS:O	3:G:730:GLU:HB3	2.15	0.46
1:A:1889:LEU:HA	1:A:1895:ILE:HD11	1.97	0.46
1:A:1951:GLY:HA3	1:A:1959:LYS:NZ	2.30	0.46
1:A:1995:ILE:HG21	3:C:631:SER:HA	1.96	0.46
2:F:482:ALA:HB3	2:F:485:MET:HB2	1.96	0.46
3:C:672:PHE:CE1	3:C:705:TRP:HB2	2.50	0.46
1:A:1884:HIS:NE2	1:E:1953:ALA:HB1	2.31	0.46
1:A:2016:LEU:HD11	1:A:2050:ARG:HB3	1.98	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1889:LEU:HA	1:E:1895:ILE:HD11	1.96	0.46
1:E:2044:HIS:HD2	1:E:2046:ILE:HB	1.79	0.46
1:E:2103:LEU:HA	1:E:2110:LEU:HD12	1.97	0.46
1:A:1844:ASP:CG	1:A:1900:ARG:HH22	2.19	0.46
1:E:2108:GLU:OE2	1:E:2138:ARG:HG2	2.15	0.46
1:E:2016:LEU:HD11	1:E:2050:ARG:HB3	1.96	0.45
3:G:747:LEU:O	3:G:748:GLU:HG2	2.16	0.45
1:E:2229:HIS:CD2	1:E:2247:HIS:HE1	2.33	0.45
1:A:2079:LEU:HD23	1:A:2084:LEU:HD11	1.97	0.45
2:F:446:TYR:CE2	2:F:450:MET:HG3	2.52	0.45
3:G:621:ALA:HA	3:G:624:HIS:CE1	2.51	0.45
3:C:718:GLU:HG2	4:H:147:ASP:HA	1.99	0.45
2:B:496:ARG:HG2	2:B:515:HIS:HB2	1.99	0.45
1:E:2349:ILE:HD13	2:F:361:LEU:HD13	1.98	0.45
1:A:1891:THR:HG23	1:A:1894:LEU:HB2	2.00	0.44
2:B:446:TYR:CE2	2:B:450:MET:HG3	2.52	0.44
1:E:2201:GLN:HA	1:E:2210:TYR:HA	2.00	0.44
1:A:1892:ASP:OD2	4:D:123:ILE:HD12	2.18	0.44
1:E:2297:THR:HB	1:E:2300:ILE:HG13	1.99	0.44
1:A:2143:PRO:HB2	1:A:2150:LEU:HD11	2.00	0.44
2:F:469:TRP:HZ2	2:F:509:LYS:HE3	1.83	0.44
2:F:501:PHE:HB3	2:F:514:PHE:HZ	1.82	0.44
4:D:144:SER:HB2	4:D:147:ASP:OD2	2.18	0.44
1:E:1947:VAL:HA	1:E:1962:LEU:HD23	2.00	0.44
1:E:1891:THR:HG23	1:E:1894:LEU:HB2	2.01	0.43
3:C:621:ALA:HA	3:C:624:HIS:CE1	2.52	0.43
1:E:2108:GLU:CD	1:E:2108:GLU:H	2.22	0.43
1:E:1951:GLY:HA3	1:E:1959:LYS:CE	2.49	0.43
1:E:1952:GLU:HG3	1:E:1955:ASN:HD22	1.83	0.43
1:A:1863:TYR:CE1	1:A:1941:ARG:HD2	2.54	0.42
2:B:474:GLU:HB2	2:B:520:LYS:HD2	2.01	0.42
4:H:144:SER:HB3	4:H:147:ASP:OD2	2.18	0.42
1:A:1876:PHE:HB2	1:A:1945:LEU:HD13	2.01	0.42
1:E:2126:ILE:HG21	1:E:2314:VAL:HG21	2.01	0.42
1:A:2208:ASN:O	1:A:2209:ARG:HG2	2.19	0.42
3:C:743:GLU:HB2	3:C:746:TYR:CD2	2.55	0.42
3:C:646:PRO:HG2	3:C:649:HIS:CD2	2.54	0.41
2:B:439:TYR:HB3	2:B:443:LEU:HD23	2.02	0.41
1:E:2023:PHE:CE1	1:E:2041:LEU:HD13	2.55	0.41
1:A:2088:MET:HE2	1:A:2088:MET:HB3	1.87	0.41
2:B:434:ILE:HA	2:B:439:TYR:OH	2.19	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:G:661:GLU:HA	3:G:664:GLN:HE21	1.84	0.41
1:A:2126:ILE:HG21	1:A:2314:VAL:HG21	2.02	0.41
1:A:1857:ARG:HG3	1:A:1905:MET:CE	2.51	0.41
1:A:1951:GLY:HA3	1:A:1959:LYS:HD3	2.03	0.41
2:F:434:ILE:HA	2:F:439:TYR:OH	2.20	0.41
1:A:2044:HIS:CD2	1:A:2046:ILE:HB	2.54	0.41
1:A:1856:LEU:HD23	1:A:1905:MET:SD	2.61	0.41
1:A:1891:THR:HG23	1:A:1894:LEU:H	1.86	0.41
1:A:1924:ASN:O	1:A:1928:ILE:HG12	2.21	0.41
1:E:1844:ASP:CG	1:E:1900:ARG:HH22	2.25	0.41
1:E:1978:HIS:CE1	1:E:2032:PRO:HD2	2.56	0.41
1:E:2035:VAL:HG11	3:G:618:GLU:HG3	2.02	0.41
1:A:2035:VAL:HG11	3:C:618:GLU:HG3	2.01	0.41
3:C:721:GLN:HA	3:C:740:PHE:O	2.21	0.41
1:E:1897:ARG:HG2	1:E:1900:ARG:HH21	1.86	0.41
1:A:2002:ASN:HD21	1:A:2050:ARG:HD2	1.86	0.40
1:A:2028:PRO:HB2	1:A:2074:TYR:CD1	2.56	0.40
2:B:469:TRP:HB3	2:B:478:TRP:CE3	2.56	0.40

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	515/535 (96%)	499 (97%)	15 (3%)	1 (0%)	47 79
1	E	510/535 (95%)	488 (96%)	20 (4%)	2 (0%)	34 69
2	B	189/197 (96%)	174 (92%)	13 (7%)	2 (1%)	14 46
2	F	189/197 (96%)	175 (93%)	13 (7%)	1 (0%)	29 64
3	C	146/148 (99%)	138 (94%)	8 (6%)	0	100 100
3	G	146/148 (99%)	139 (95%)	7 (5%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
4	D	29/54 (54%)	29 (100%)	0	0	100 100
4	H	26/54 (48%)	25 (96%)	1 (4%)	0	100 100
All	All	1750/1868 (94%)	1667 (95%)	77 (4%)	6 (0%)	41 73

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	428	ARG
2	F	400	ALA
2	B	400	ALA
1	E	2202	VAL
1	A	2082	VAL
1	E	2082	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	464/479 (97%)	439 (95%)	25 (5%)	22 53
1	E	460/479 (96%)	434 (94%)	26 (6%)	20 52
2	B	167/172 (97%)	161 (96%)	6 (4%)	35 67
2	F	167/172 (97%)	160 (96%)	7 (4%)	30 62
3	C	136/136 (100%)	130 (96%)	6 (4%)	28 61
3	G	136/136 (100%)	132 (97%)	4 (3%)	42 72
4	D	24/41 (58%)	23 (96%)	1 (4%)	30 62
4	H	16/41 (39%)	16 (100%)	0	100 100
All	All	1570/1656 (95%)	1495 (95%)	75 (5%)	25 58

All (75) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1841	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1844	ASP
1	A	1849	ARG
1	A	1851	LYS
1	A	1870	ARG
1	A	1895	ILE
1	A	1910	SER
1	A	1956	THR
1	A	2020	CYS
1	A	2045	ARG
1	A	2055	THR
1	A	2081	ASN
1	A	2108	GLU
1	A	2111	CYS
1	A	2119	ASP
1	A	2141	ARG
1	A	2144	ASP
1	A	2153	ASP
1	A	2157	GLU
1	A	2194	SER
1	A	2204	ASN
1	A	2212	LEU
1	A	2323	LEU
1	A	2340	HIS
1	A	2343	VAL
2	B	364	ILE
2	B	403	TRP
2	B	424	ASN
2	B	488	THR
2	B	512	LYS
2	B	524	ARG
3	C	610	GLU
3	C	697	ARG
3	C	703	MET
3	C	717	ASP
3	C	742	PHE
3	C	748	GLU
4	D	122	GLU
1	E	1844	ASP
1	E	1849	ARG
1	E	1851	LYS
1	E	1870	ARG
1	E	1895	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	1910	SER
1	E	1918	GLN
1	E	1959	LYS
1	E	2045	ARG
1	E	2081	ASN
1	E	2084	LEU
1	E	2108	GLU
1	E	2111	CYS
1	E	2119	ASP
1	E	2141	ARG
1	E	2157	GLU
1	E	2194	SER
1	E	2198	SER
1	E	2201	GLN
1	E	2202	VAL
1	E	2209	ARG
1	E	2212	LEU
1	E	2235	SER
1	E	2323	LEU
1	E	2340	HIS
1	E	2343	VAL
2	F	364	ILE
2	F	371	PRO
2	F	403	TRP
2	F	424	ASN
2	F	489	MET
2	F	514	PHE
2	F	524	ARG
3	G	610	GLU
3	G	703	MET
3	G	742	PHE
3	G	747	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2002	ASN
1	A	2044	HIS
1	A	2081	ASN
1	A	2232	ASN
2	B	515	HIS
2	B	526	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	694	GLN
3	C	699	HIS
1	E	1917	GLN
1	E	1961	ASN
1	E	1964	ASN
1	E	1978	HIS
1	E	2002	ASN
1	E	2044	HIS
1	E	2081	ASN
1	E	2201	GLN
1	E	2229	HIS
1	E	2232	ASN
1	E	2247	HIS
3	G	664	GLN
3	G	694	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

6.1 Protein, DNA and RNA chains

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^{th} 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	519/535 (97%)	-0.25	5 (0%) 82 67	49, 81, 137, 163	0
1	E	514/535 (96%)	-0.13	13 (2%) 57 34	55, 92, 142, 185	0
2	B	191/197 (96%)	0.20	20 (10%) 6 2	67, 109, 205, 217	0
2	F	191/197 (96%)	-0.04	6 (3%) 49 26	73, 102, 174, 190	0
3	C	148/148 (100%)	-0.18	2 (1%) 75 56	50, 101, 133, 157	0
3	G	148/148 (100%)	-0.10	1 (0%) 87 75	51, 89, 123, 136	0
4	D	33/54 (61%)	0.55	2 (6%) 21 9	109, 137, 158, 164	0
4	H	30/54 (55%)	0.68	6 (20%) 1 0	120, 136, 165, 181	0
All	All	1774/1868 (94%)	-0.10	55 (3%) 49 26	49, 95, 157, 217	0

All (55) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	488	THR	4.2
2	B	496	ARG	4.1
1	E	2361	MET	3.9
1	E	1841	GLU	3.8
2	B	516	LEU	3.5
2	B	473	LYS	3.4
2	B	521	LEU	3.4
1	E	1891	THR	3.2
4	H	127	LEU	3.2
2	B	526	HIS	3.2
4	H	126	SER	3.1
1	E	2206	PRO	3.1
2	F	531	PHE	3.1
2	B	532	ASN	3.1
1	E	1950	SER	3.1
2	F	498	THR	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	528	PRO	3.0
1	E	2204	ASN	3.0
1	E	1922	ALA	3.0
2	B	517	GLU	3.0
2	B	515	HIS	2.9
2	B	472	HIS	2.9
2	F	530	THR	2.8
2	B	489	MET	2.8
2	B	513	GLU	2.7
1	E	1921	PRO	2.7
1	A	2010	THR	2.7
1	A	2361	MET	2.7
2	B	493	THR	2.6
1	E	1888	ILE	2.6
2	F	526	HIS	2.6
4	H	144	SER	2.6
2	F	474	GLU	2.5
2	B	491	THR	2.5
4	H	147	ASP	2.4
1	E	2168	PHE	2.4
2	B	487	PRO	2.4
4	D	120	THR	2.3
4	D	147	ASP	2.3
1	E	2358	GLN	2.3
1	A	1839	ALA	2.3
3	C	644	PRO	2.3
4	H	121	ASP	2.2
1	E	2174	PRO	2.2
2	B	490	LYS	2.2
4	H	123	ILE	2.1
2	B	525	PRO	2.1
1	A	1840	SER	2.1
2	B	478	TRP	2.1
1	A	1841	GLU	2.1
2	B	494	TYR	2.1
3	C	726	TYR	2.1
3	G	647	PRO	2.1
1	E	2003	ALA	2.0
2	F	500	TYR	2.0

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.