wwPDB EM Validation Summary Report (i)

Nov 19, 2022 – 10:48 pm GMT

PDB ID	:	6F9E
EMDB ID	:	EMD-4200
Title	:	Model of the Rift Valley fever virus glycoprotein hexamer type 3
Authors	:	Halldorsson, S.; Bowden, T.A.; Huiskonen, J.T.
Deposited on	:	2017-12-14
Resolution	:	13.30 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

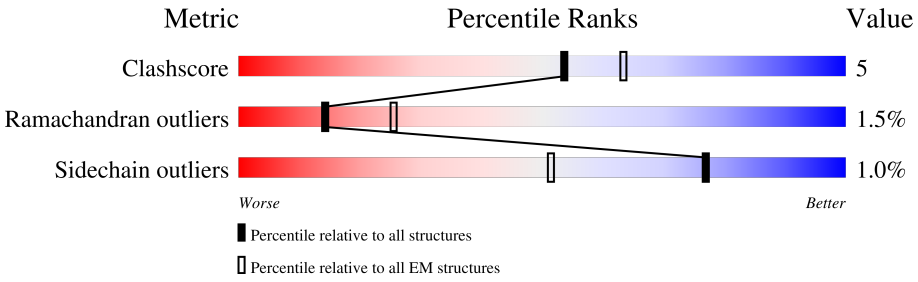
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 13.30 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.



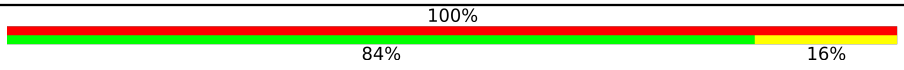
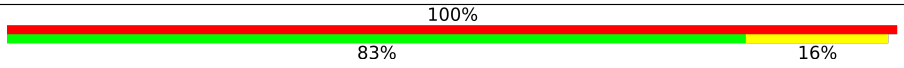
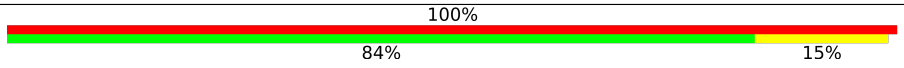
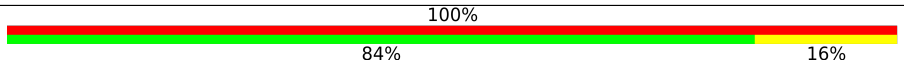
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	316	43% 81% 14% 5%
1	C	316	46% 80% 15% 5%
1	E	316	46% 80% 15% • 5%
1	G	316	43% 78% 17% 5%
1	I	316	41% 82% 13% 5%
1	K	316	41% 80% 15% 5%
2	B	431	100% 85% 15%
2	D	431	100% 85% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	431	
2	H	431	
2	J	431	
2	L	431	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 33048 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	C	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	E	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	G	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	I	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	K	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		

- Molecule 2 is a protein called Glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	D	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	F	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	H	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	J	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	L	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	688	ASP	-	expression tag	UNP A2T072
B	689	PRO	-	expression tag	UNP A2T072

Continued on next page...

Continued from previous page...

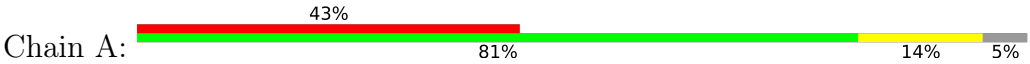
Chain	Residue	Modelled	Actual	Comment	Reference
B	690	GLY	-	expression tag	UNP A2T072
D	688	ASP	-	expression tag	UNP A2T072
D	689	PRO	-	expression tag	UNP A2T072
D	690	GLY	-	expression tag	UNP A2T072
F	688	ASP	-	expression tag	UNP A2T072
F	689	PRO	-	expression tag	UNP A2T072
F	690	GLY	-	expression tag	UNP A2T072
H	688	ASP	-	expression tag	UNP A2T072
H	689	PRO	-	expression tag	UNP A2T072
H	690	GLY	-	expression tag	UNP A2T072
J	688	ASP	-	expression tag	UNP A2T072
J	689	PRO	-	expression tag	UNP A2T072
J	690	GLY	-	expression tag	UNP A2T072
L	688	ASP	-	expression tag	UNP A2T072
L	689	PRO	-	expression tag	UNP A2T072
L	690	GLY	-	expression tag	UNP A2T072

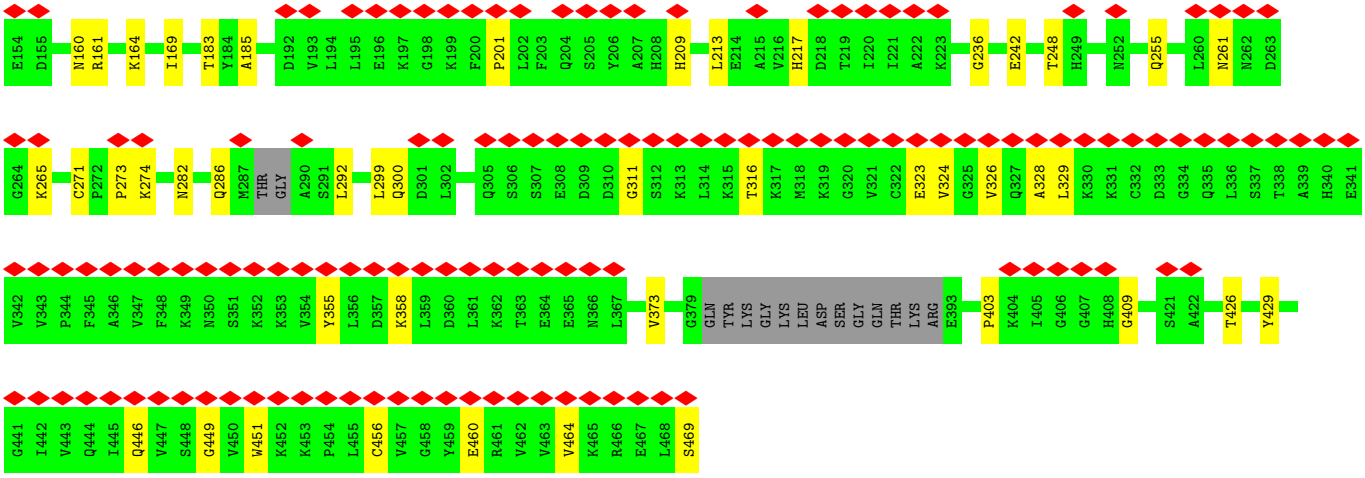
3 Residue-property plots (i)

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 atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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: Glycoprotein

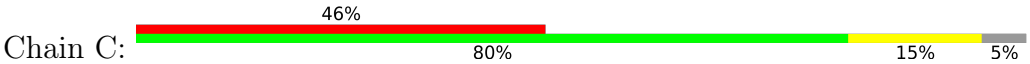
Chain A:

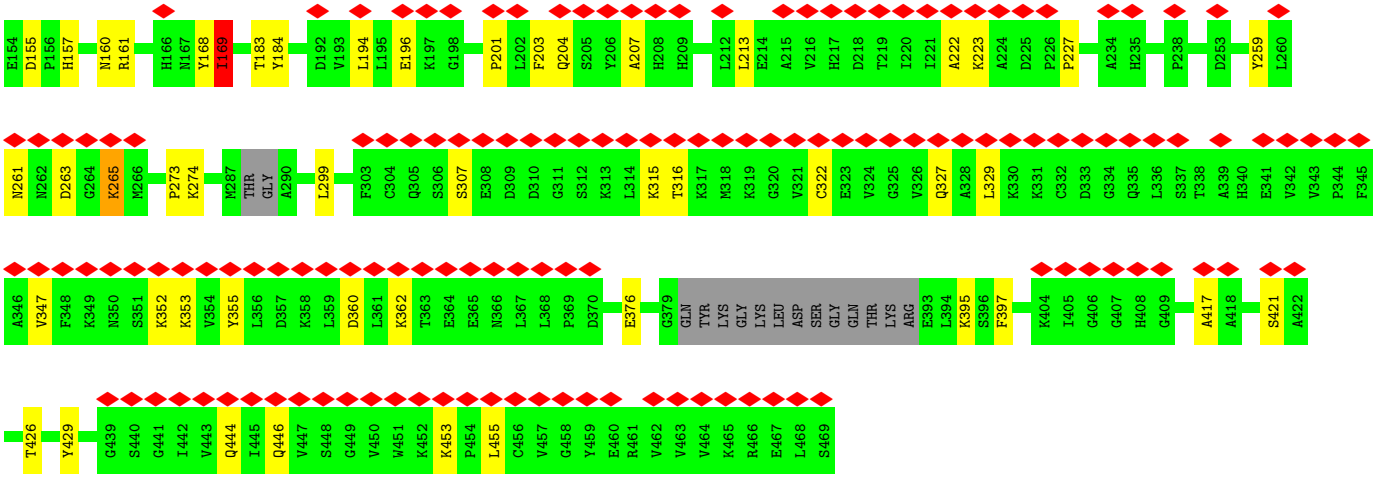




• Molecule 1: Glycoprotein

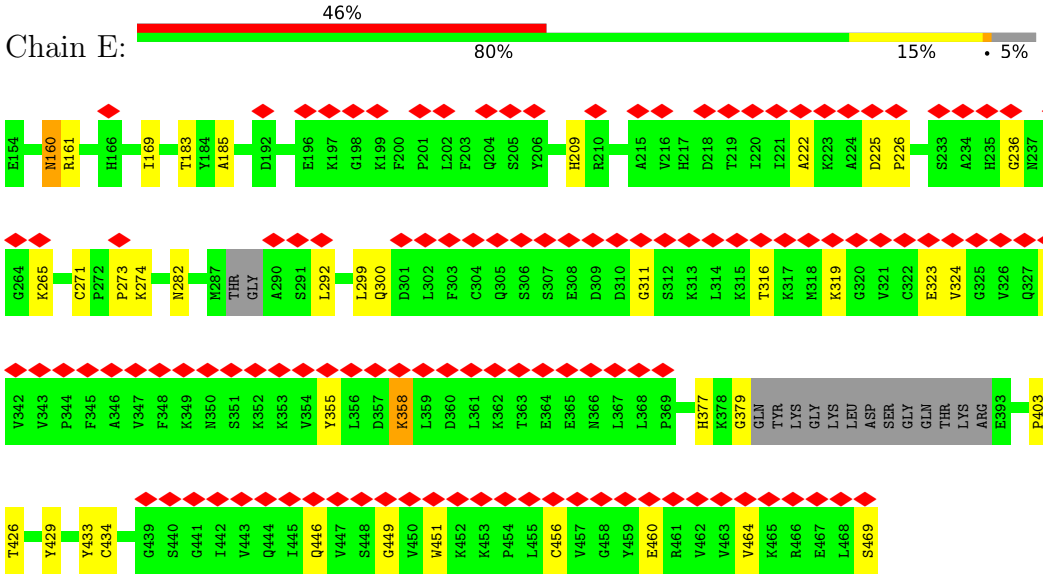
Chain C:





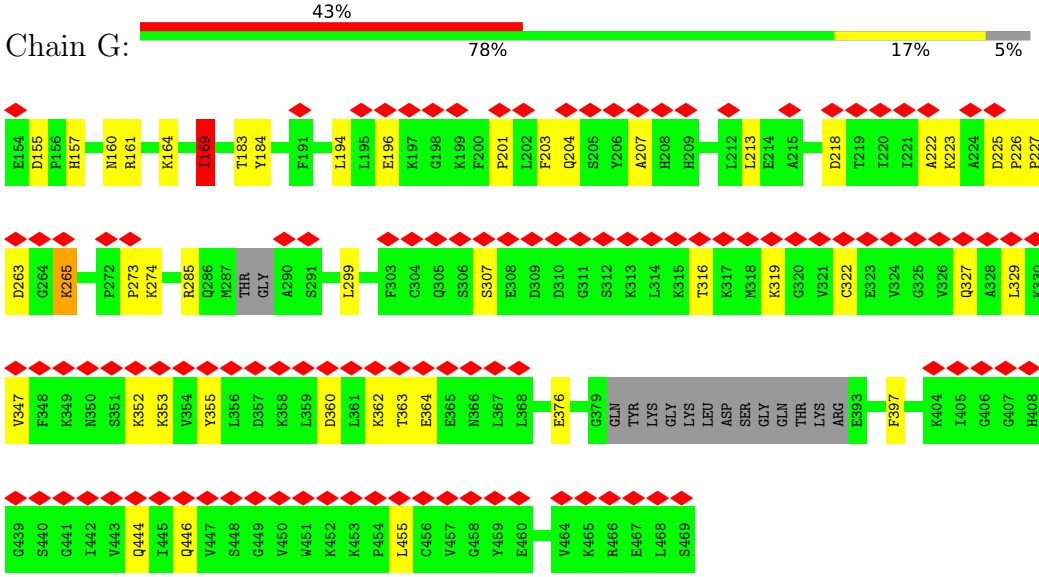
• Molecule 1: Glycoprotein

Chain E:



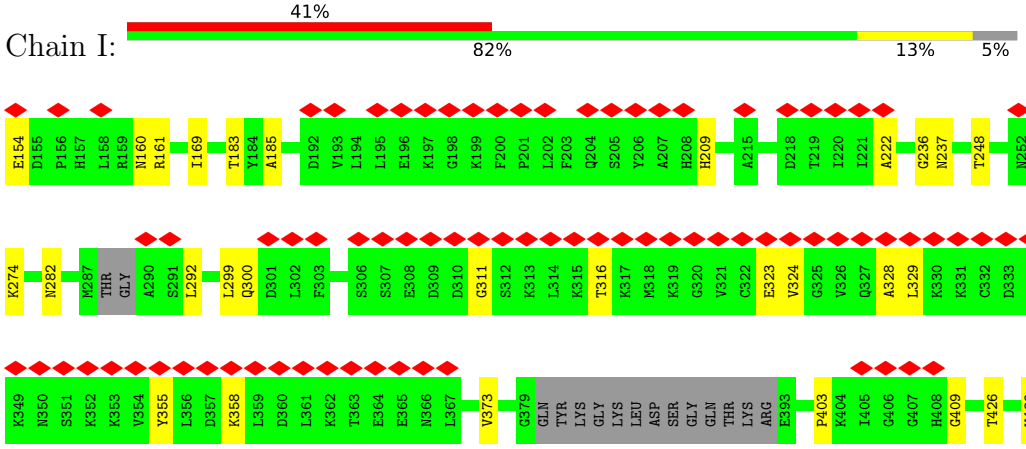
• Molecule 1: Glycoprotein

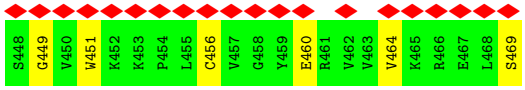
Chain G:



• Molecule 1: Glycoprotein

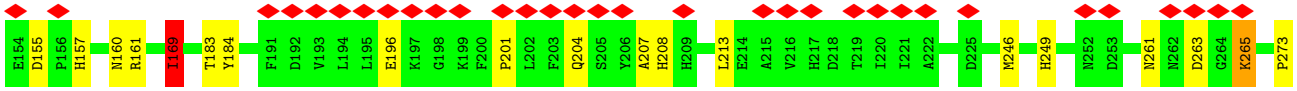
Chain I:

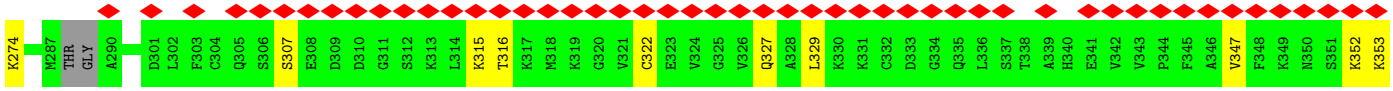


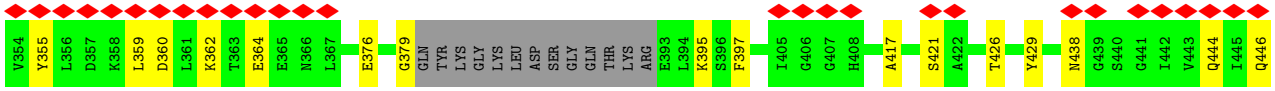


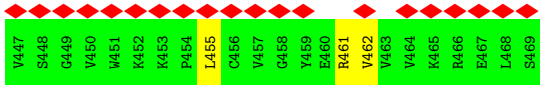
• Molecule 1: Glycoprotein

Chain K:



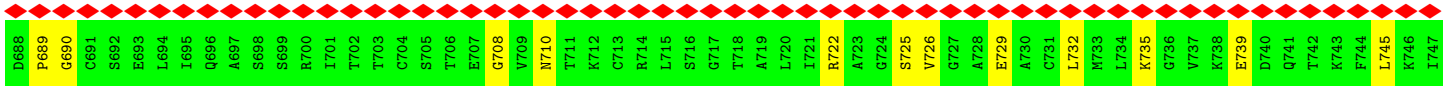


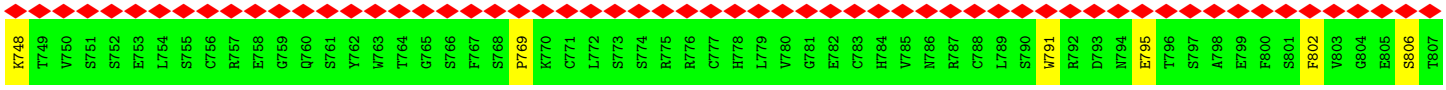


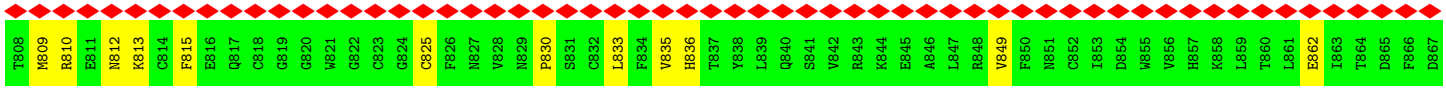


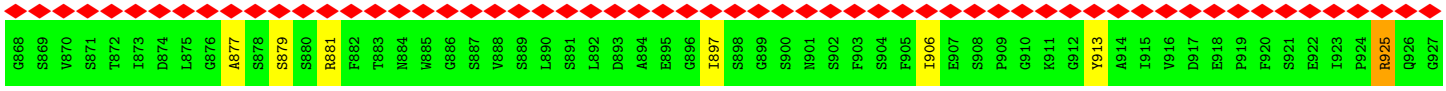
• Molecule 2: Glycoprotein

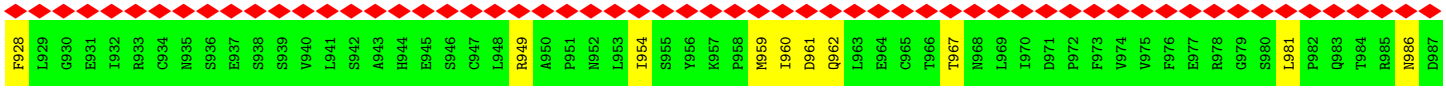
Chain B:

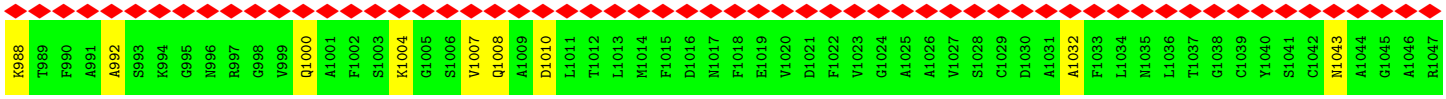












V1048 C1049 L1050 S1051 I1052 T1053 S1054 T1055 G1056 T1057 G1058 S1059 L1060 S1061 A1062 H1063 N1064 K1065 D1066 G1067 S1068 L1069 H1070 I1071 V1072 L1073 P1074 S1075 E1076 N1077 G1078 T1079 K1080 D1081 Q1082 C1083 Q1084 I1085 L1086 H1087 F1088 T1089 V1090 P1091 E1092 V1093 E1094 E1095 E1096 F1097 M1098 Y1099 S1100 C1101 D1102 G1103 D1104 E1105 R1106 P1107

L1108 L1109 V1110 K1111 G1112 T1113 L1114 I1115 A1116 I1117 D1118

• Molecule 2: Glycoprotein

Chain D:

D688 P689 G690 C691 S692 E693 L694 I695 Q696 A697 S698 S699 R700 I701 T702 T703 C704 S705 T706 E707 G708 V709 N710 T711 K712 C713 R714 L715 S716 G717 T718 A719 L720 I721 R722 A723 G724 S725 V726 G727 A728 E729 A730 C731 L732 M733 L734 K735 G736 V737 K738 E739 D740 Q741 T742 K743 F744 L745 K746 I747

K748 T749 V750 S751 S752 E753 L754 S755 C756 R757 E758 G759 Q760 S761 Y762 W763 T764 G765 S766 F767 S768 P769 K770 C771 L772 S773 S774 R775 R776 C777 H778 L779 V780 G781 E782 C783 H784 V785 N786 R787 C788 L789 S790 W791 R792 D793 N794 E795 T796 S797 A798 E799 F800 S801 F802 V803 G804 E805 S806 T807

T808 M809 R810 E811 N812 K813 C814 F815 E816 Q817 C818 G819 G820 W821 G822 C823 G824 C825 F826 N827 V828 N829 P830 S831 C832 L833 F834 V835 H836 T837 Y838 L839 Q840 S841 V842 R843 K844 E845 A846 L847 R848 V849 F850 N851 C852 I853 D854 W855 V856 H857 K858 L859 T860 L861 E862 I863 T864 D865 F866 D867

G868 S869 V870 S871 T872 I873 D874 L875 G876 A877 S878 S879 S880 R881 F882 T883 N884 W885 G886 S887 V888 S889 L890 S891 L892 D893 A894 E895 G896 I897 S898 G899 S900 N901 S902 F903 S904 F905 I906 E907 S908 P909 G910 K911 G912 Y913 A914 I915 V916 D917 E918 P919 F920 S921 E922 I923 P924 R925 Q926 G927

F928 L929 G930 E931 I932 R933 C934 N935 S936 E937 S938 S939 V940 L941 S942 A943 H944 E945 S946 C947 L948 R949 A950 P951 N952 L953 I954 S955 Y956 K957 P958 N959 I960 D961 Q962 L963 E964 C965 T966 T967 N968 L969 I970 D971 P972 F973 V974 V975 F976 E977 R978 G979 S980 L981 P982 Q983 T984 R985 N986 D987

K988 T989 F990 A991 A992 S993 K994 G995 N996 R997 G998 V999 Q1000 A1001 F1002 S1003 K1004 C1005 S1006 V1007 Q1008 A1009 D1010 L1011 T1012 L1013 M1014 F1015 D1016 N1017 F1018 E1019 V1020 D1021 F1022 V1023 G1024 A1025 A1026 V1027 S1028 C1029 D1030 A1031 A1032 F1033 L1034 N1035 L1036 T1037 G1038 C1039 Y1040 S1041 C1042 N1043 A1044 G1045 A1046 R1047

V1048 C1049 L1050 S1051 I1052 T1053 S1054 T1055 G1056 T1057 G1058 S1059 L1060 S1061 A1062 H1063 N1064 K1065 D1066 G1067 S1068 L1069 H1070 I1071 V1072 L1073 P1074 S1075 E1076 N1077 G1078 T1079 K1080 D1081 Q1082 C1083 Q1084 I1085 L1086 H1087 F1088 T1089 V1090 P1091 E1092 V1093 E1094 E1095 E1096 F1097 M1098 Y1099 S1100 C1101 D1102 G1103 D1104 E1105 R1106 P1107

L1108 L1109 V1110 K1111 G1112 T1113 L1114 I1115 A1116 I1117 D1118

• Molecule 2: Glycoprotein

Chain F:

D688 P689 G690 C691 S692 E693 L694 I695 Q696 A697 S698 S699 R700 I701 T702 T703 C704 S705 T706 E707 G708 V709 N710 T711 K712 C713 R714 L715 S716 G717 T718 A719 L720 I721 R722 A723 G724 S725 V726 G727 A728 E729 A730 C731 L732 M733 L734 K735 G736 V737 K738 E739 D740 Q741 T742 K743 F744 L745 K746 I747

K748 T749 V750 S751 S752 E753 L754 S755 C756 R757 E758 G759 Q760 S761 Y762 W763 T764 G765 S766 F767 S768 P769 K770 C771 L772 S773 S774 R775 R776 C777 H778 L779 V780 G781 E782 C783 H784 V785 N786 R787 C788 L789 S790 W791 R792 D793 N794 E795 T796 S797 A798 E799 F800 S801 F802 V803 G804 E805 S806 T807

T808 M809 R810 E811 N812 K813 C814 F815 E816 Q817 C818 G819 G820 W821 G822 C823 G824 C825 F826 N827 V828 N829 P830 S831 C832 L833 F834 V835 H836 T837 Y838 L839 Q840 S841 V842 R843 K844 E845 A846 L847 R848 V849 F850 N851 C852 I853 D854 W855 V856 H857 K858 L859 T860 L861 E862 I863 T864 D865 F866 D867

G868 S869 V870 S871 T872 I873 D874 L875 G876 A877 S878 S879 S880 R881 F882 T883 N884 W885 G886 S887 V888 S889 L890 S891 L892 D893 A894 E895 G896 I897 S898 G899 S900 N901 S902 F903 S904 F905 I906 E907 S908 P909 G910 K911 G912 Y913 A914 I915 V916 D917 E918 P919 F920 S921 E922 I923 P924 R925 Q926 G927

F928 L929 G930 E931 I932 R933 C934 N935 S936 E937 S938 S939 V940 L941 S942 A943 H944 E945 S946 C947 L948 R949 A950 P951 N952 L953 I954 S955 Y956 K957 P958 M959 I960 D961 Q962 L963 E964 C965 T966 T967 N968 L969 I970 D971 P972 F973 V974 V975 F976 E977 R978 G979 S980 L981 P982 Q983 T984 R985 N986 D987

K988 T989 F990 A991 A992 S993 K994 G995 N996 R997 G998 V999 Q1000 A1001 F1002 S1003 K1004 G1005 S1006 V1007 Q1008 A1009 D1010 L1011 T1012 L1013 M1014 F1015 D1016 N1017 F1018 E1019 V1020 D1021 F1022 V1023 G1024 A1025 A1026 V1027 S1028 C1029 D1030 A1031 A1032 F1033 L1034 N1035 L1036 T1037 G1038 C1039 Y1040 S1041 C1042 N1043 A1044 G1045 A1046 R1047

V1048 C1049 L1050 S1051 I1052 T1053 S1054 T1055 G1056 T1057 G1058 S1059 L1060 S1061 A1062 H1063 N1064 K1065 D1066 G1067 S1068 L1069 H1070 I1071 V1072 L1073 P1074 S1075 E1076 N1077 G1078 T1079 K1080 D1081 Q1082 C1083 Q1084 I1085 L1086 H1087 F1088 T1089 V1090 P1091 E1092 V1093 E1094 E1095 E1096 F1097 M1098 Y1099 S1100 C1101 D1102 G1103 D1104 E1105 R1106 P1107

L1108 L1109 V1110 K1111 G1112 T1113 L1114 I1115 A1116 I1117 D1118

• Molecule 2: Glycoprotein

Chain H:

D688 P689 G690 C691 S692 E693 L694 I695 Q696 A697 S698 S699 R700 I701 T702 T703 C704 S705 T706 E707 G708 V709 N710 T711 K712 C713 R714 L715 S716 G717 T718 A719 L720 I721 R722 A723 G724 S725 V726 G727 A728 E729 A730 C731 L732 M733 L734 K735 G736 V737 K738 E739 D740 Q741 T742 K743 F744 L745 K746 L747

K748 T749 V750 S751 S752 E753 L754 S755 C756 R757 E758 G759 Q760 S761 Y762 W763 T764 G765 S766 F767 S768 P769 K770 C771 L772 S773 S774 R775 R776 C777 H778 L779 V780 G781 E782 C783 H784 V785 N786 R787 C788 L789 S790 W791 R792 D793 N794 E795 T796 S797 A798 E799 F800 S801 F802 V803 G804 E805 S806 T807

T808 M809 R810 E811 N812 K813 C814 F815 E816 Q817 C818 G819 G820 W821 G822 C823 G824 C825 F826 N827 V828 N829 P830 S831 C832 L833 F834 V835 H836 T837 Y838 L839 Q840 S841 V842 R843 K844 E845 A846 L847 R848 V849 F850 N851 C852 I853 D854 W855 V856 H857 K858 L859 T860 L861 E862 I863 T864 D865 F866 D867

G868 S869 V870 S871 T872 I873 D874 L875 G876 A877 S878 S879 S880 R881 F882 T883 N884 W885 G886 S887 V888 S889 L890 S891 L892 D893 A894 E895 G896 I897 S898 G899 S900 N901 S902 F903 S904 F905 I906 E907 S908 P909 G910 K911 G912 Y913 A914 I915 V916 D917 E918 P919 F920 S921 E922 I923 P924 R925 Q926 G927

F928 L929 G930 E931 I932 R933 C934 N935 S936 E937 S938 S939 V940 L941 S942 A943 H944 E945 S946 C947 L948 R949 A950 P951 N952 L953 I954 S955 Y956 K957 P958 M959 I960 D961 Q962 L963 E964 C965 T966 T967 N968 L969 I970 D971 P972 F973 V974 V975 F976 E977 R978 G979 S980 L981 P982 Q983 T984 R985 N986 D987

K988 T989 F990 A991 A992 S993 K994 G995 N996 R997 G998 V999 Q1000 A1001 F1002 S1003 K1004 G1005 S1006 V1007 Q1008 A1009 D1010 L1011 T1012 L1013 M1014 F1015 D1016 N1017 F1018 E1019 V1020 D1021 F1022 V1023 G1024 A1025 A1026 V1027 S1028 C1029 D1030 A1031 A1032 F1033 L1034 N1035 L1036 T1037 G1038 C1039 Y1040 S1041 C1042 N1043 A1044 G1045 A1046 R1047

V1048 C1049 L1050 S1051 I1052 T1053 S1054 T1055 G1056 T1057 G1058 S1059 L1060 S1061 A1062 H1063 N1064 K1065 D1066 G1067 S1068 L1069 H1070 I1071 V1072 L1073 P1074 S1075 E1076 N1077 G1078 T1079 K1080 D1081 Q1082 C1083 Q1084 I1085 L1086 H1087 F1088 T1089 V1090 P1091 E1092 V1093 E1094 E1095 E1096 F1097 M1098 Y1099 S1100 C1101 D1102 G1103 D1104 E1105 R1106 P1107

L1108 L1109 V1110 K1111 G1112 T1113 L1114 I1115 A1116 I1117 D1118

- Molecule 2: Glycoprotein

Chain J:

D688 P689 G690 C691 S692 E693 L694 I695 Q696 A697 S698 S699 R700 I701 T702 T703 C704 S705 T706 E707 G708 V709 N710 T711 K712 C713 R714 L715 S716 G717 T718 A719 L720 T721 R722 A723 G724 S725 V726 G727 A728 E729 A730 C731 L732 M733 L734 K735 G736 V737 K738 E739 D740 Q741 T742 K743 F744 L745 K746 I747

K748 T749 V750 S751 S752 E753 L754 S755 C756 R757 E758 G759 Q760 S761 Y762 W763 T764 G765 S766 F767 S768 P769 K770 C771 L772 S773 S774 R775 R776 C777 H778 L779 V780 G781 E782 C783 H784 V785 N786 R787 C788 L789 S790 W791 R792 D793 N794 E795 T796 S797 A798 E799 F800 S801 F802 V803 G804 E805 S806 T807

T808 M809 R810 E811 N812 K813 C814 F815 E816 Q817 C818 G819 G820 W821 G822 C823 G824 C825 F826 N827 V828 N829 P830 S831 C832 L833 F834 V835 H836 T837 Y838 L839 Q840 S841 V842 R843 K844 E845 A846 L847 R848 V849 F850 N851 C852 I853 D854 W855 V856 H857 K858 L859 T860 L861 E862 I863 T864 D865 F866 D867

G868 S869 V870 S871 T872 I873 D874 L875 G876 A877 S878 S879 S880 R881 F882 T883 N884 W885 G886 S887 V888 S889 L890 S891 L892 D893 A894 E895 G896 I897 S898 G899 S900 N901 S902 F903 S904 F905 I906 E907 S908 P909 G910 K911 G912 Y913 A914 I915 V916 D917 E918 P919 F920 S921 E922 I923 P924 R925 Q926 G927

F928 L929 G930 E931 I932 R933 C934 N935 S936 E937 S938 S939 V940 L941 S942 A943 H944 E945 S946 C947 L948 R949 A950 P951 N952 L953 I954 S955 Y956 K957 P958 M959 I960 D961 Q962 L963 E964 C965 T966 T967 N968 L969 I970 D971 P972 F973 V974 V975 F976 E977 R978 G979 S980 L981 P982 Q983 T984 R985 N986 D987

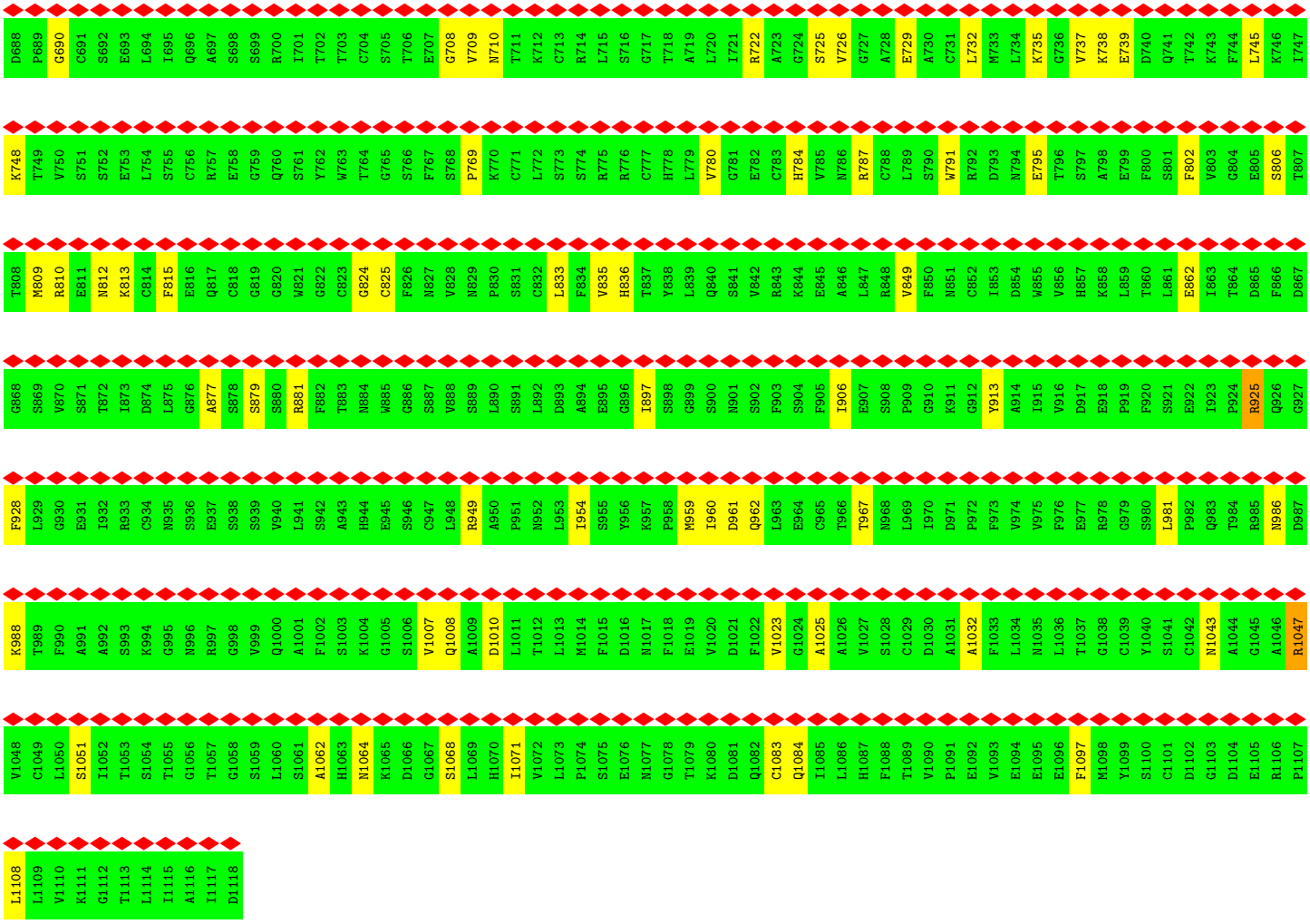
K988 T989 F990 A991 A992 S993 K994 G995 N996 R997 G998 V999 Q1000 A1001 F1002 S1003 K1004 G1005 S1006 V1007 Q1008 A1009 D1010 L1011 T1012 L1013 M1014 F1015 D1016 N1017 F1018 E1019 V1020 D1021 F1022 V1023 G1024 A1025 A1026 V1027 S1028 C1029 D1030 A1031 A1032 F1033 L1034 N1035 L1036 T1037 G1038 C1039 Y1040 S1041 C1042 N1043 A1044 G1045 A1046 R1047

V1048 C1049 L1050 S1051 I1052 T1053 S1054 T1055 G1056 T1057 G1058 S1059 L1060 S1061 A1062 H1063 N1064 K1065 D1066 G1067 S1068 L1069 H1070 I1071 V1072 L1073 P1074 S1075 E1076 N1077 G1078 T1079 K1080 D1081 Q1082 C1083 Q1084 I1085 L1086 H1087 F1088 T1089 V1090 P1091 E1092 V1093 E1094 E1095 E1096 F1097 M1098 Y1099 S1100 C1101 D1102 G1103 D1104 E1105 R1106 P1107

L1108 L1109 V1110 K1111 G1112 T1113 L1114 I1115 A1116 I1117 D1118

- Molecule 2: Glycoprotein

Chain L:


D688 P689 G690 C691 S692 E693 L694 I695 Q696 A697 S698 S699 R700 I701 T702 T703 C704 S705 T706 E707 G708 V709 N710 T711 K712 C713 R714 L715 S716 G717 T718 A719 L720 I721 R722 A723 G724 S725 V726 G727 A728 E729 A730 C731 L732 M733 L734 K735 G736 V737 K738 E739 D740 Q741 T742 K743 F744 L745 K746 I747
K748 T749 V750 S751 S752 E753 L754 S755 C756 R757 E758 G759 Q760 S761 Y762 W763 T764 G765 S766 F767 S768 P769 K770 C771 L772 S773 S774 R775 R776 C777 H778 L779 V780 G781 E782 C783 H784 V785 N786 R787 C788 L789 S790 W791 R792 D793 N794 E795 T796 S797 A798 E799 F800 S801 F802 V803 G804 E805 S806 T807
T808 M809 R810 E811 N812 K813 C814 F815 E816 Q817 C818 G819 G820 W821 G822 C823 G824 C825 F826 N827 V828 N829 P830 S831 C832 L833 F834 V835 H836 T837 Y838 L839 Q840 S841 V842 R843 K844 E845 A846 L847 R848 V849 F850 N851 C852 I853 D854 W855 V856 H857 K858 L859 T860 L861 E862 I863 T864 D865 F866 D867
G868 S869 V870 S871 T872 I873 D874 L875 G876 A877 S878 S879 S880 R881 F882 T883 N884 W885 G886 S887 V888 S889 L890 S891 L892 D893 A894 E895 G896 I897 S898 G899 S900 N901 S902 F903 S904 F905 I906 E907 S908 P909 G910 K911 G912 Y913 A914 I915 V916 D917 E918 P919 F920 S921 E922 I923 P924 R925 Q926 G927
F928 L929 G930 E931 I932 R933 C934 N935 S936 E937 S938 S939 V940 L941 S942 A943 H944 E945 S946 C947 L948 R949 A950 P951 N952 L953 I954 S955 Y956 K957 P958 M959 I960 D961 Q962 L963 E964 C965 T966 T967 N968 L969 I970 D971 P972 F973 V974 V975 F976 E977 R978 G979 S980 L981 P982 Q983 T984 R985 N986 D987
K988 T989 F990 A991 A992 S993 K994 G995 N996 R997 G998 V999 Q1000 A1001 F1002 S1003 K1004 G1005 S1006 V1007 Q1008 A1009 D1010 L1011 T1012 L1013 M1014 F1015 D1016 N1017 F1018 E1019 V1020 D1021 F1022 V1023 G1024 A1025 A1026 V1027 S1028 C1029 D1030 A1031 A1032 F1033 L1034 N1035 L1036 T1037 G1038 C1039 Y1040 S1041 C1042 N1043 A1044 G1045 A1046 R1047
V1048 C1049 L1050 S1051 I1052 T1053 S1054 T1055 G1056 T1057 G1058 S1059 L1060 S1061 A1062 H1063 N1064 K1065 D1066 G1067 S1068 L1069 H1070 I1071 V1072 L1073 P1074 S1075 E1076 N1077 G1078 T1079 K1080 D1081 Q1082 C1083 Q1084 I1085 L1086 H1087 F1088 T1089 V1090 P1091 E1092 V1093 E1094 E1095 E1096 F1097 M1098 Y1099 S1100 C1101 D1102 G1103 D1104 E1105 R1106 P1107
L1108 L1109 V1110 K1111 G1112 T1113 L1114 I1115 A1116 I1117 D1118

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	2995	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	22	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.121	Depositor
Minimum map value	-0.060	Depositor
Average map value	-0.003	Depositor
Map value standard deviation	0.019	Depositor
Recommended contour level	0.03	Depositor
Map size (Å)	345.6, 345.6, 345.6	wwPDB
Map dimensions	128, 128, 128	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.7, 2.7, 2.7	Depositor

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	0.24	0/2333	0.40	0/3136
1	C	0.24	0/2333	0.41	0/3136
1	E	0.24	0/2333	0.40	0/3136
1	G	0.24	0/2333	0.41	0/3136
1	I	0.24	0/2333	0.40	0/3136
1	K	0.24	0/2333	0.41	0/3136
2	B	0.24	0/3284	0.42	0/4431
2	D	0.24	0/3284	0.42	0/4431
2	F	0.24	0/3284	0.43	1/4431 (0.0%)
2	H	0.24	0/3284	0.42	0/4431
2	J	0.24	0/3284	0.42	0/4431
2	L	0.24	0/3284	0.42	0/4431
All	All	0.24	0/33702	0.42	1/45402 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	689	PRO	C-N-CA	5.95	134.78	122.30

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	2284	0	2205	21	0
1	C	2284	0	2205	27	0
1	E	2284	0	2205	27	0
1	G	2284	0	2205	32	0
1	I	2284	0	2205	21	0
1	K	2284	0	2205	27	0
2	B	3224	0	3071	34	0
2	D	3224	0	3071	35	0
2	F	3224	0	3071	36	0
2	H	3224	0	3071	42	0
2	J	3224	0	3071	35	0
2	L	3224	0	3071	38	0
All	All	33048	0	31656	353	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.

The worst 5 of 353 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:H:926:GLN:NE2	2:H:931:GLU:OE2	2.27	0.67
2:D:959:MET:HB2	2:D:962:GLN:HB2	1.79	0.65
2:F:959:MET:HB2	2:F:962:GLN:HB2	1.79	0.64
2:D:1062:ALA:HB3	2:D:1071:ILE:HB	1.80	0.64
2:B:959:MET:HB2	2:B:962:GLN:HB2	1.79	0.64

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 PDB entries followed by that with respect to all EM entries.

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	295/316 (93%)	251 (85%)	39 (13%)	5 (2%)	9 42

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	295/316 (93%)	253 (86%)	38 (13%)	4 (1%)	11 46
1	E	295/316 (93%)	254 (86%)	37 (12%)	4 (1%)	11 46
1	G	295/316 (93%)	250 (85%)	39 (13%)	6 (2%)	7 38
1	I	295/316 (93%)	254 (86%)	37 (12%)	4 (1%)	11 46
1	K	295/316 (93%)	252 (85%)	39 (13%)	4 (1%)	11 46
2	B	429/431 (100%)	372 (87%)	51 (12%)	6 (1%)	11 46
2	D	429/431 (100%)	370 (86%)	53 (12%)	6 (1%)	11 46
2	F	429/431 (100%)	372 (87%)	50 (12%)	7 (2%)	9 44
2	H	429/431 (100%)	374 (87%)	49 (11%)	6 (1%)	11 46
2	J	429/431 (100%)	372 (87%)	50 (12%)	7 (2%)	9 44
2	L	429/431 (100%)	372 (87%)	50 (12%)	7 (2%)	9 44
All	All	4344/4482 (97%)	3746 (86%)	532 (12%)	66 (2%)	14 46

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	456	CYS
2	B	825	CYS
2	D	825	CYS
1	E	456	CYS
2	F	709	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 PDB entries followed by that with respect to all EM entries.

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	241/271 (89%)	238 (99%)	3 (1%)	71 83
1	C	241/271 (89%)	239 (99%)	2 (1%)	81 89
1	E	241/271 (89%)	237 (98%)	4 (2%)	60 78
1	G	241/271 (89%)	239 (99%)	2 (1%)	81 89
1	I	241/271 (89%)	238 (99%)	3 (1%)	71 83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	K	241/271 (89%)	238 (99%)	3 (1%)	71 83
2	B	347/371 (94%)	344 (99%)	3 (1%)	78 87
2	D	347/371 (94%)	345 (99%)	2 (1%)	86 92
2	F	347/371 (94%)	344 (99%)	3 (1%)	78 87
2	H	347/371 (94%)	344 (99%)	3 (1%)	78 87
2	J	347/371 (94%)	343 (99%)	4 (1%)	71 83
2	L	347/371 (94%)	344 (99%)	3 (1%)	78 87
All	All	3528/3852 (92%)	3493 (99%)	35 (1%)	77 86

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	1047	ARG
1	K	160	ASN
2	L	925	ARG
1	E	358	LYS
1	E	282	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	438	ASN
2	L	1000	GLN
2	H	1000	GLN
2	L	1070	HIS
1	K	160	ASN

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.

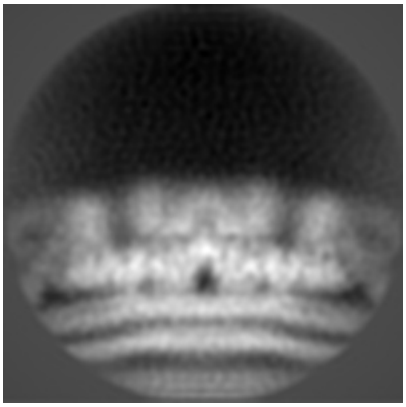
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-4200. These allow visual inspection of the internal detail of the map and identification of artifacts.

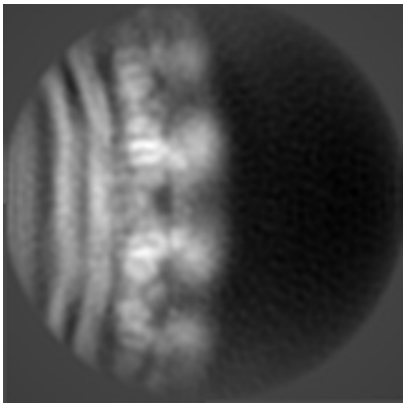
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

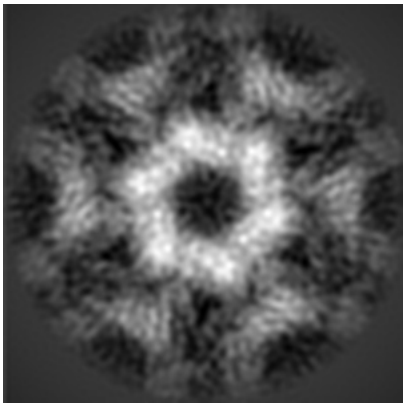
6.1.1 Primary map



X



Y

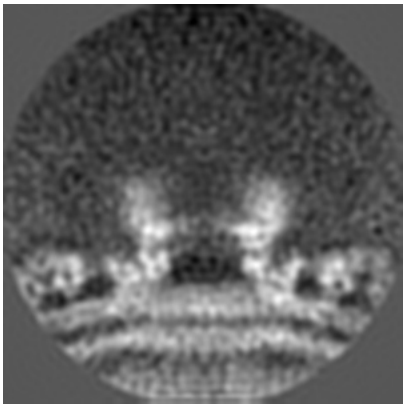


Z

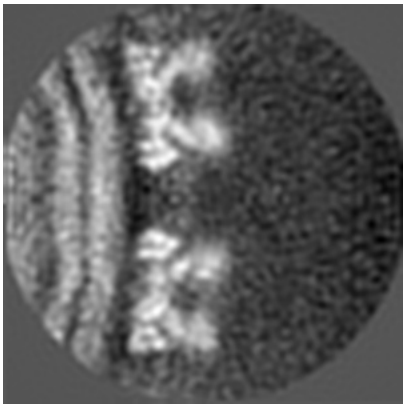
The images above show the map projected in three orthogonal directions.

6.2 Central slices

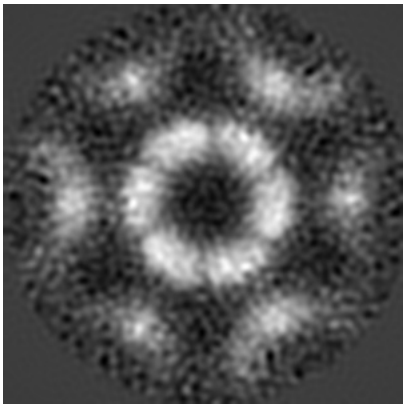
6.2.1 Primary map



X Index: 64



Y Index: 64

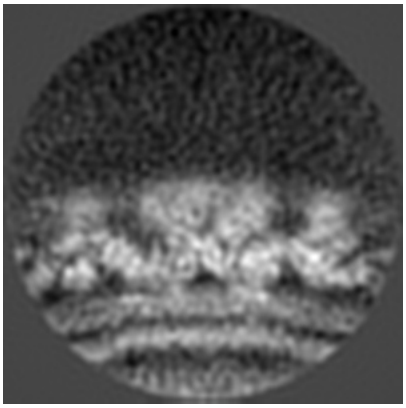


Z Index: 64

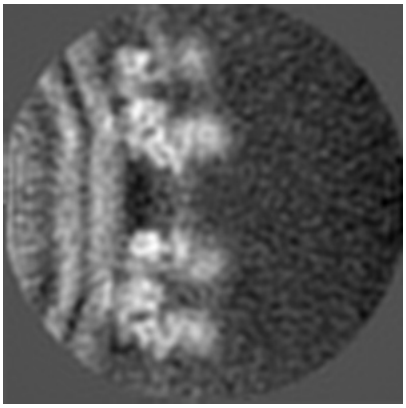
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

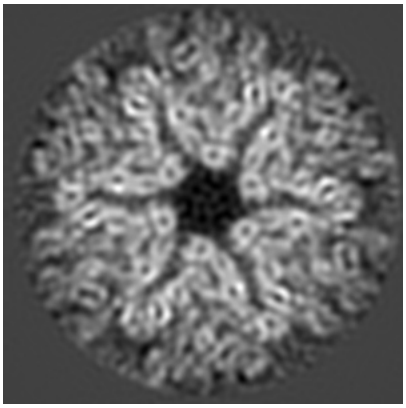
6.3.1 Primary map



X Index: 82



Y Index: 60



Z Index: 43

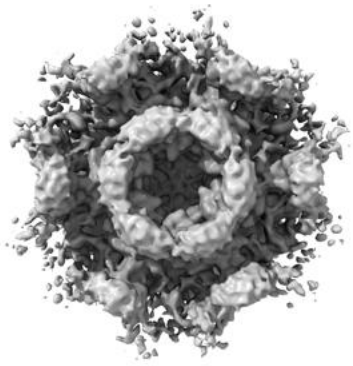
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

6.4.1 Primary map

X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.03. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

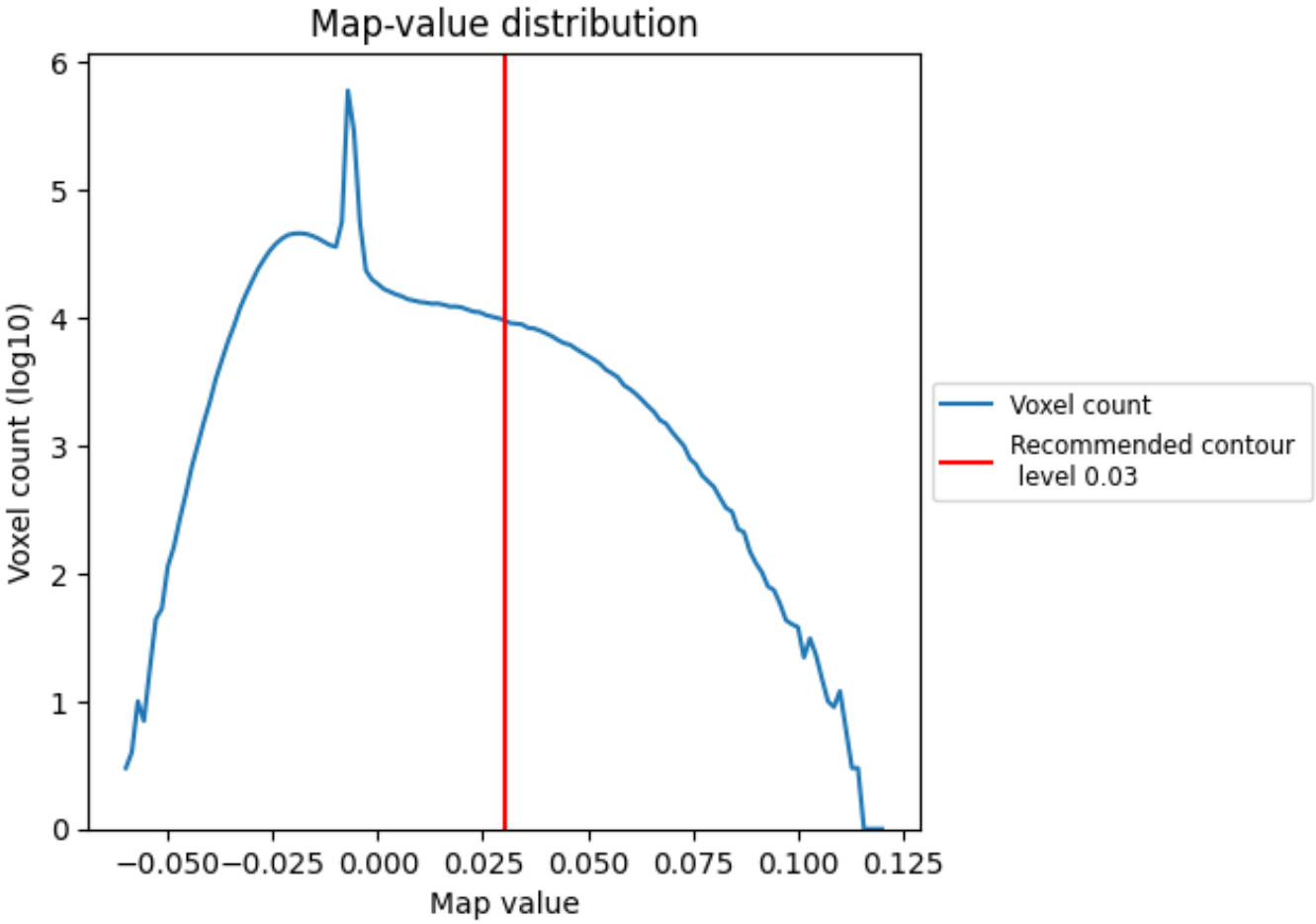
6.5 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

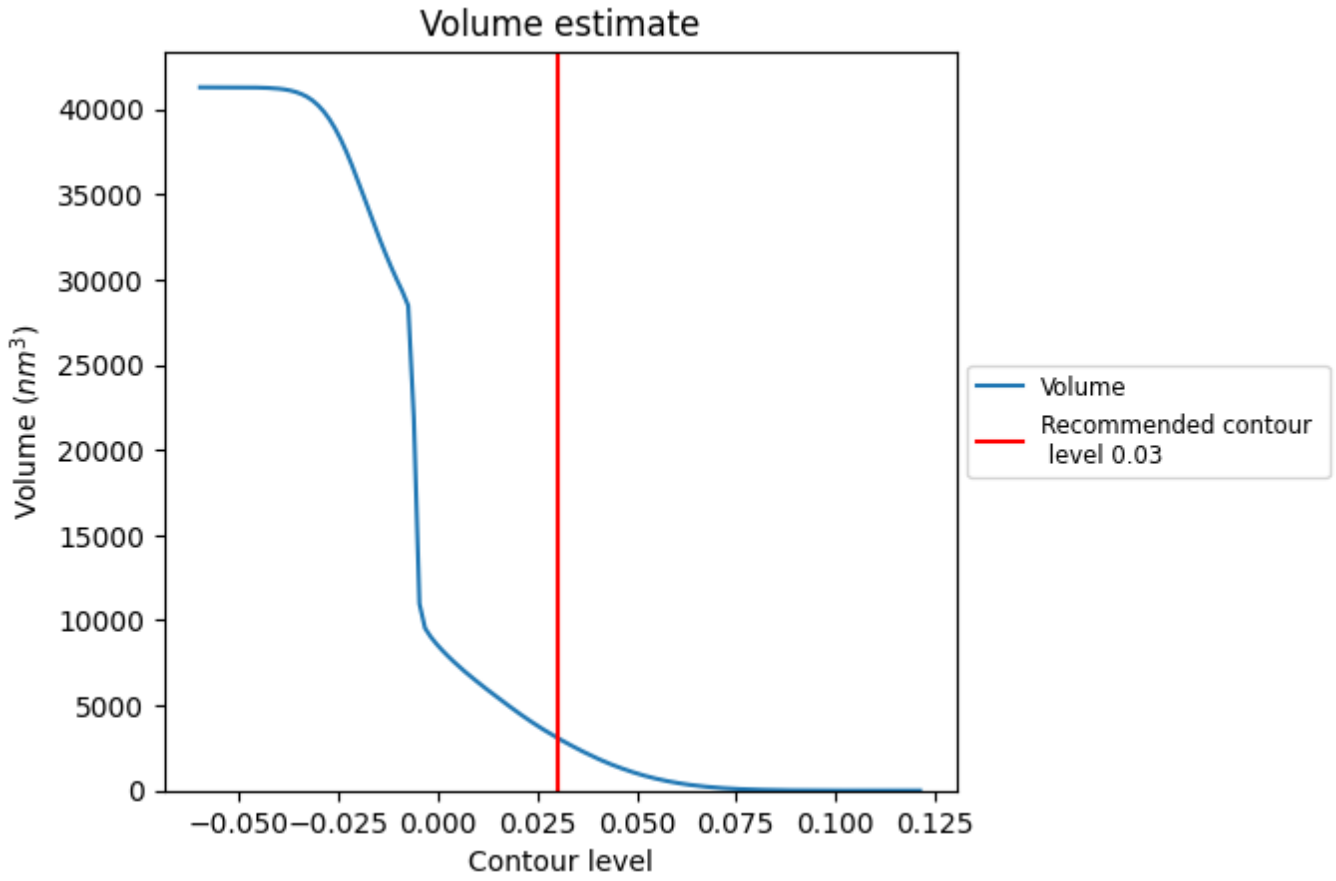
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

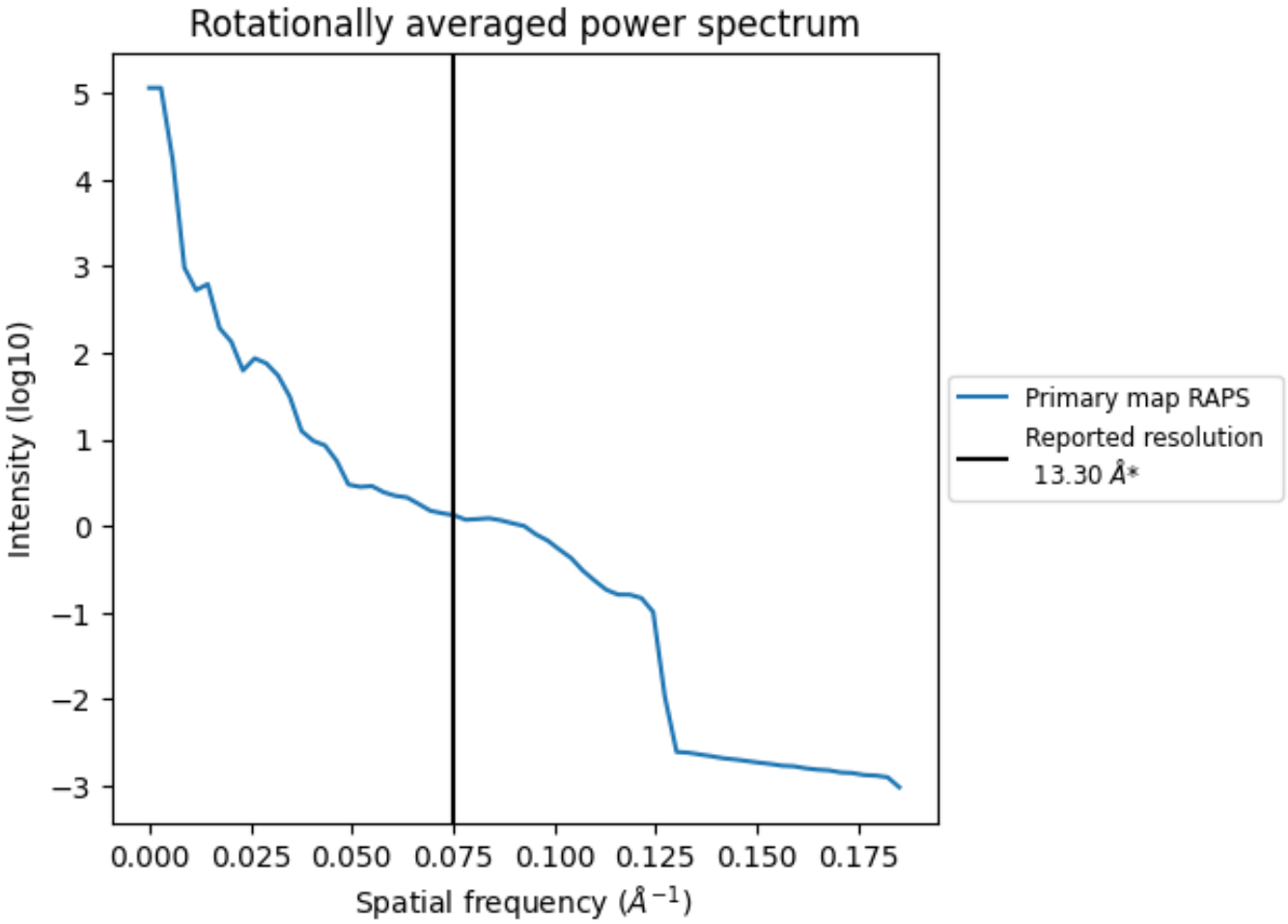
7.2 Volume estimate (i)



The volume at the recommended contour level is 3102 nm^3; this corresponds to an approximate mass of 2802 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum (i)

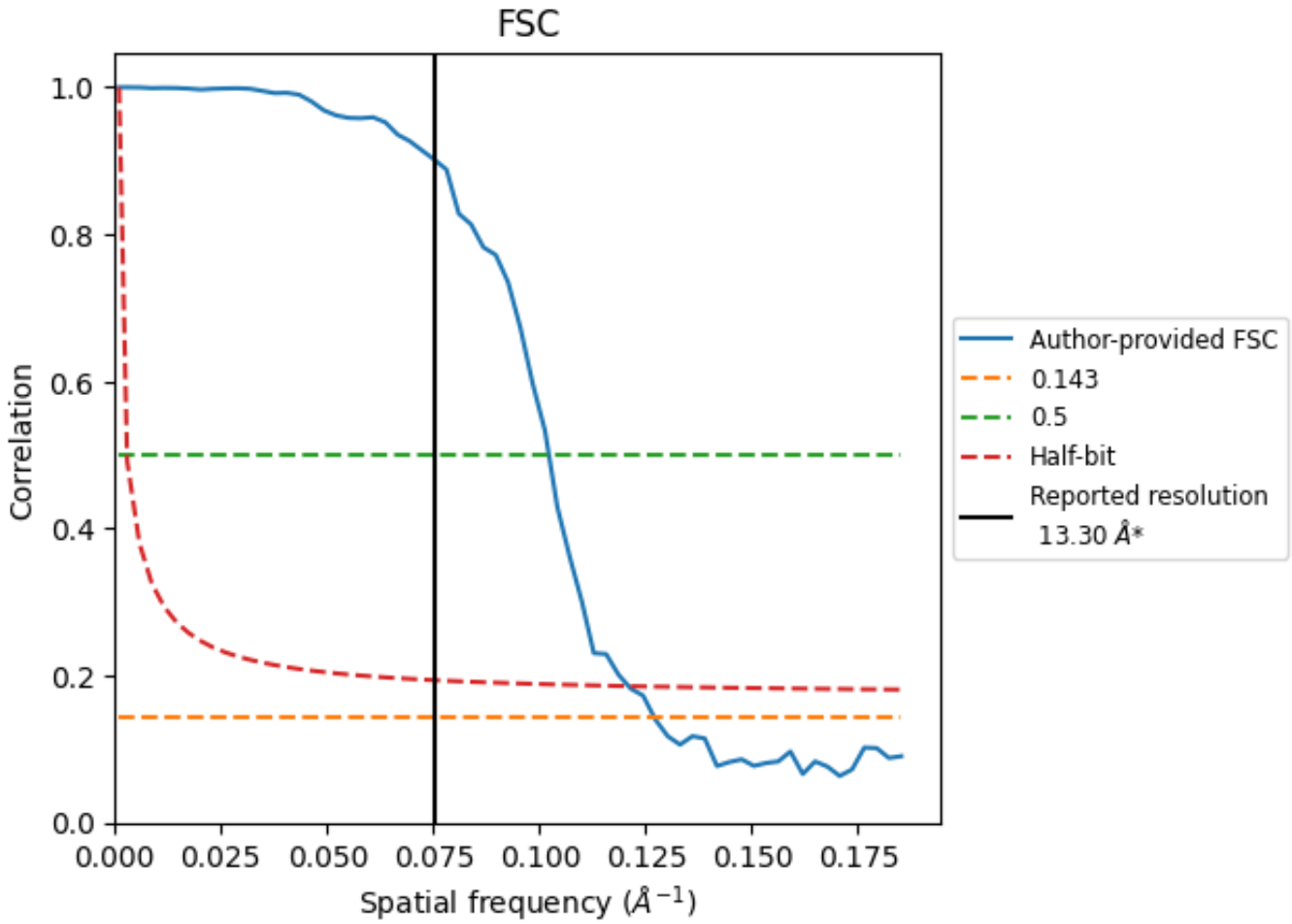


*Reported resolution corresponds to spatial frequency of 0.075 Å^{-1}

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.075 Å^{-1}

8.2 Resolution estimates (i)

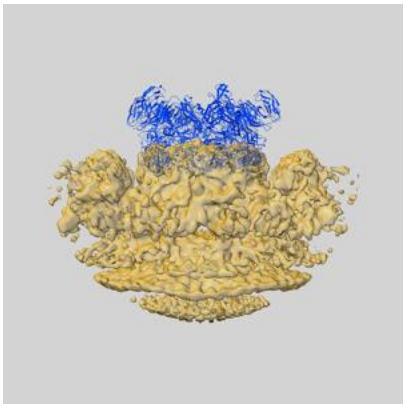
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	13.30	-	-
Author-provided FSC curve	7.87	9.78	8.26
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from author-provided FSC intersecting FSC 0.143 CUT-OFF 7.87 differs from the reported value 13.3 by more than 10 %

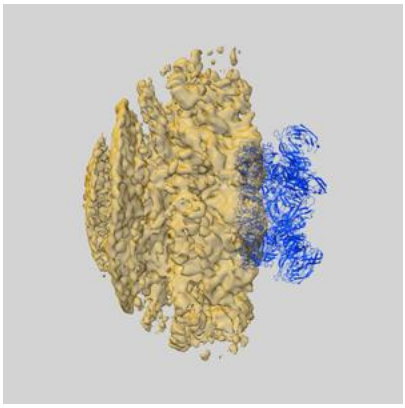
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-4200 and PDB model 6F9E. Per-residue inclusion information can be found in section 3 on page 6.

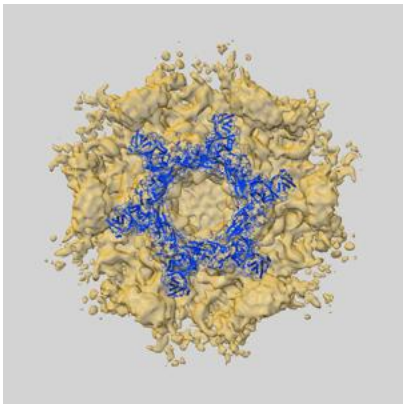
9.1 Map-model overlay



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.03 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)

X Y Z

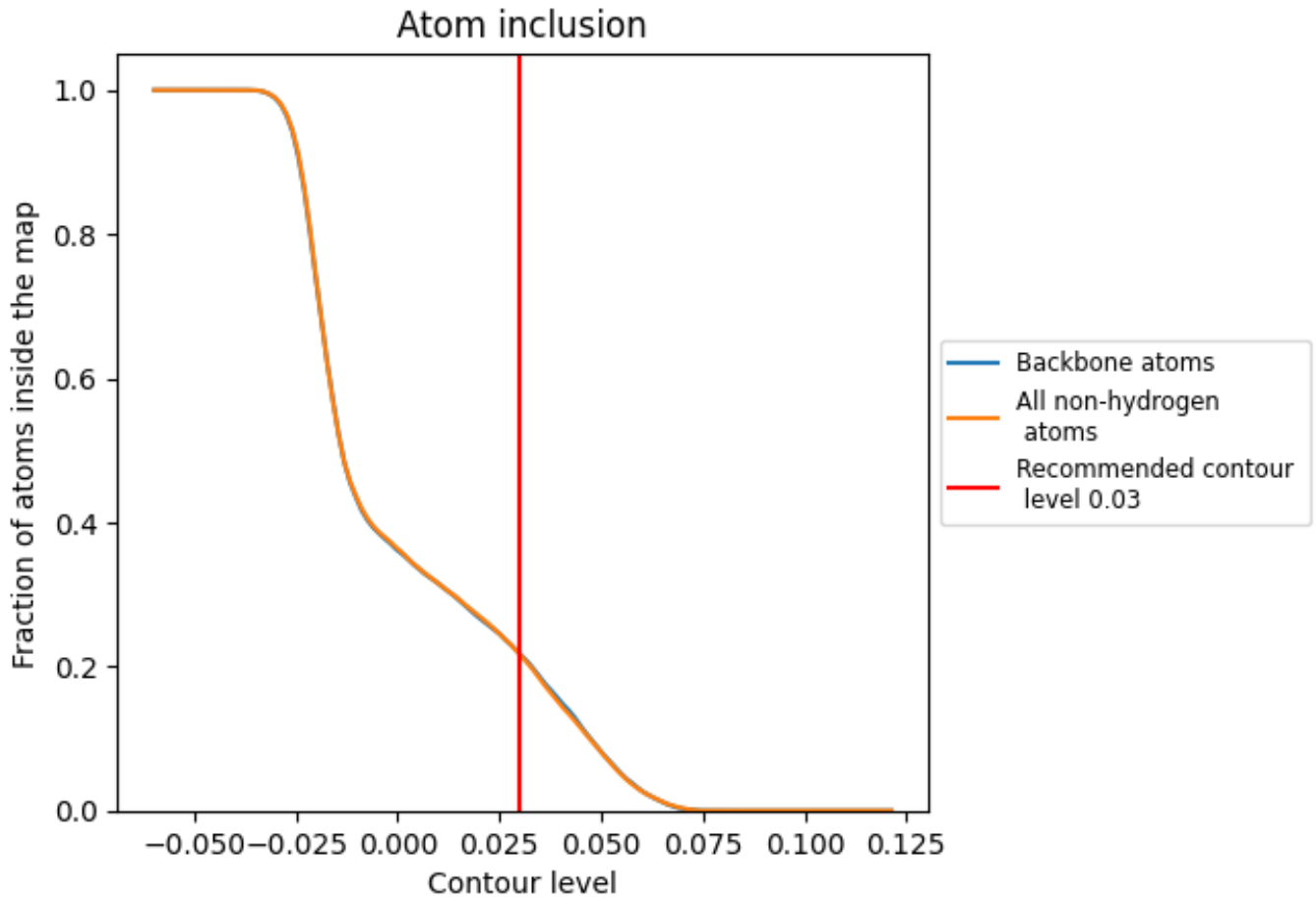
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model (i)

X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.03).

9.4 Atom inclusion (i)



At the recommended contour level, 22% of all backbone atoms, 22% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.03) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.2167	0.0090
A	0.5179	0.0220
B	0.0013	0.0060
C	0.4958	0.0190
D	0.0000	0.0010
E	0.5064	0.0260
F	0.0000	-0.0000
G	0.5210	0.0210
H	0.0000	-0.0040
I	0.5383	0.0250
J	0.0013	-0.0020
K	0.5427	0.0180
L	0.0003	-0.0010