1 **A conserved homo-dimerization interface in human IFIT1 provides**

2 **insights into IFIT interactome assembly**

3 **Yazan M. Abbas[1], Saúl Martínez-Montero[2], Regina Cencic[3], Jerry Pelletier[3], Peter D.**

4 **Pawelek[4,5], Masad J. Damha[2], & Bhushan Nagar[1,5]**

5 [1]Department of Biochemistry, McGill University, 3649 Promenade Sir William Osler, Montreal,

6 Quebec H3G 0B1, Canada

7 [2]Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec

8 H3A 0B8, Canada

9 [3]Departments of Biochemistry and Oncology and The Rosalind and Morris Goodman Cancer

10 Research Center, McGill University, Montréal, Québec H3G 1Y6, Canada

11 [4]Department of Chemistry, Concordia University, 7141 Sherbrooke Street West, Montréal,

12 Québec H4B 1R6, Canada

13 [5]Groupe de Recherche Axé sur la Structure des Protéines (GRASP)

14 Running title: Crystal structures of RNA-bound human IFIT1 dimers

15 To whom correspondence should be addressed: Dr. Bhushan Nagar, Room 464, 3649 promenade

16 Sir-William-Osler, Montréal, Québec, H3G 0B1, Canada. Telephone: 514-398-7272. Fax: 514-

17 398-2983. E- mail: bhushan.nagar@mcgill.ca

20

Abstract

22 The Interferon-Induced Proteins with Tetratricopeptide Repeats (IFITs) are a group of potently

23 expressed Interferon Stimulated Genes that mediate antiviral innate immunity. Previous studies

24 have revealed that most IFITs partake in higher order structures, potentially as part of an 'IFIT

25 interactome' that results in viral inhibition. Recent crystal structures of a mutated, monomeric form

26 of IFIT1 revealed the molecular basis of how it recognizes non-self, capped viral mRNAs to

27 selectively inhibit their translation. However, wild-type IFIT1 forms dimers in solution and the

28 role of dimerization was not examined in detail. Here we present a structural and biochemical

29 analysis of wild-type IFIT1 in complex with capped and uncapped RNA. Wild-type IFIT1 forms

30 an antiparallel, elongated dimer that is in stark contrast to the domain-swapped, parallel dimer

31 found in IFIT2. Dimerization takes place through a small, C-terminal interface that is

32 evolutionarily conserved in IFIT1 and IFIT1B proteins. The interface is modular and can be grafted

33 onto IFIT5, which is natively monomeric, to induce dimerization. Mutational analysis of this

34 interface showed that homo-dimerization is not required for full RNA binding or translational

35 inhibition by IFIT1. Sedimentation velocity analytical ultracentrifugation measurements

36 demonstrated a reversible monomer-dimer equilibrium, suggesting that dimerization is of low

37 affinity and could play a role under physiological concentrations, possibly in regulating IFIT

38 interactome assembly. Finally, conformational changes in IFIT1 that occur upon RNA binding

39 provide insight into how RNA enters its binding site in solution.

Introduction

41 Interferon (IFN) stimulated genes (ISGs) are a broad class of genes whose expression is triggered

42 downstream of virally activated signaling pathways (1). They number in the hundreds and

43 altogether mediate diverse antiviral effects to promote an antiviral state within infected and

44 uninfected cells. The IFN-induced proteins with tetratricopeptide repeats (IFITs) are among the

45   most potently induced ISGs (2). They are conserved throughout vertebrate evolution, and in

46   humans and most mammals, consist of 5 paralogues: IFIT1, IFIT1B, IFIT2, IFIT3, and IFIT5 (3).

47   IFITs are cytosolic proteins composed of multiple tandem copies of the tetratricopeptide repeat

48   (TPR), a helix-turn-helix motif with the propensity to assemble into super-helical arrays (4).

49   Crystal structures of several IFITs revealed that their TPRs coalesce into three small, super-helical

50   subdomains that come together to form clamp-shaped structures (5-9). TPRs generally mediate

51   protein-protein interactions (10), through which IFITs were previously implicated in modulating

52   several biological processes including cell proliferation (11, 12), migration (13, 14), apoptosis

53   (15), and cytokine signaling (16-18), although the molecular details of these are poorly defined.

54   However, more recently it was discovered that the primary mode of IFIT action is to deter

55   viral replication through direct interaction with the 5′ ends of viral RNAs (19). This role is best

56   characterized in the ability of IFIT1 and IFIT1B to compete with the eukaryotic initiation factor

57   4F (eIF4F) for binding to the *N7*-methylguanosine triphosphate (m7Gppp-) cap and cap-proximal

58   nucleotides at the 5′ end of viral mRNAs (6, 20-23). In doing so, IFIT1 and IFIT1B inhibit viral

59   mRNA translation, limiting viral protein production. Endogenous mRNA is protected from IFIT1

60   and IFIT1B translational inhibition through 'self' markers such as ribose 2′-O methylation on the

61   first and sometimes second cap proximal nucleotides (N1 and N2, where N is any nucleotide),

62   which interfere with IFIT1 and IFIT1B binding (6, 20-22, 24, 25). Thus, IFITs play a critical role

63   in antiviral innate immunity by discerning self from non-self. As such, in an attempt to evade IFIT1

64   and IFIT1B activity, many viruses have acquired the means to produce mRNA displaying 2′-O

65   methylation at N1 (26).

66   IFIT proteins also target other RNAs as part of their antiviral program. IFIT1 and IFIT5

67   can recognize uncapped 5′ triphosphate RNA (PPP-RNA) to interfere with the replication of some

68   negative-sense single-stranded RNA viruses (5, 19), and IFIT2 has been shown to interact with

69   AU-rich RNA (8). Conversely, IFIT3 is incapable of binding RNA on its own (5, 19). Thus, by

3

70   encoding different RNA binding properties and targeting different virus-derived RNAs, IFIT

71   proteins are thought to mediate non-redundant antiviral activities, which is supported by the

72   distinct and context-dependent induction pattern for each IFIT gene (2, 27). However, the

73   expression of multiple IFITs simultaneously in many cell types, and the recent discovery of an

74   IFN-dependent IFIT 'interactome' made up of IFIT1, IFIT2, IFIT3, and other host RNA-binding

75   factors suggested cooperation amongst the different IFITs resulting in synergistic or

76   complementary activities (19). At the core of this interactome, IFIT proteins interact directly with

77   each other to form homo- and hetero-oligomers (19), but the nature of these interactions and how

78   they regulate interactome assembly is not clear. Furthermore, how these interactions impact IFIT-

79   RNA binding is not known.

80   Crystal structures of a mutated form of full-length IFIT1 with capped and uncapped RNA

81   showed that it engages the 5′ end of only single-stranded RNA through a narrow, positively-

82   charged RNA-binding tunnel at the center of the protein (6). The crystal structure of IFIT5 in

83   complex with PPP-RNA showed a similar RNA-binding tunnel (5), which recognizes only

84   uncapped RNA due to protein residues at one end of the tunnel blocking any progression beyond

85   the PPP moiety (5, 6). Notably, biochemical and structural analyses of IFIT5 revealed that it exists

86   exclusively in a monomeric state in solution that undergoes a conformational change upon RNA

87   binding (5). Prior to RNA binding, IFIT5 appears to exist in a more open, discrete conformation

88   in solution that compacts and wraps itself around the incoming RNA. Two extended, non-TPR α-

89   helices, α15 and α16, form a pivot region in IFIT5 that regulate the closure of the protein around

90   the RNA. On the other hand, the crystal structure of RNA-free IFIT2 showed that it formed parallel

91   domain-swapped dimers mediated by the exchange of three α-helices (α7-α9) from its second

92   subdomain, resulting in two adjacent RNA-binding channels with relatively flexible C-terminal

93   regions (8). The crystal structures of an N-terminal fragment of IFIT1 and RNA-bound IFIT1

94   showed no evidence for domain swapping and therefore how IFIT1 forms higher order structures

95   and the nature of RNA-dependent conformational changes in IFIT1 are as yet unclear (5, 6).

96   We report here crystal structures of wild-type, human IFIT1 bound to capped and uncapped

97   RNAs. The structures show that IFIT1 forms an antiparallel, end-to-end, S-shaped dimer using its

98   extreme C-terminus to form a pseudo-continuous TPR super-helix, in stark contrast to the globular,

99   domain-swapped, parallel dimer of IFIT2. Biophysical and mutational analysis confirm the

100   presence of IFIT1 dimers in solution, and combined with functional assays, suggest that homo-

101   dimerization is not required for full RNA binding and translational inhibition. Finally, comparison

102   of the different RNA-bound forms of wild-type IFIT1, complemented with limited proteolysis,

103   suggests a role for the C-terminal subdomain and an mRNA 'cap-binding loop' in mediating

104   conformational changes that have implications for its antiviral activity. Our work uncovers a novel

105   dimerization mechanism within the IFIT family and suggests a model for mRNA binding in

106   solution.

Results

IFIT1 forms dimers in solution

109   IFIT1 has been reported to oligomerize and bind RNA with positive cooperativity (22), but the

110   stoichiometry and mechanism of oligomerization are unknown. Therefore, we first wanted to

111   validate the IFIT1 molecular weight (MW) in solution by subjecting it to multi-angle light

112   scattering with inline size-exclusion chromatography (SEC-MALS, **Fig. 1A**). IFIT1 was injected

113   into SEC-MALS at 180 µM (10 mg/ml), and the protein migrated predominantly as a single peak

114   with an estimated MW of 101.4 kDa, corresponding to an IFIT1 dimer (expected dimer MW 111

115   kDa). Interestingly, the elution profiles consistently displayed a non-symmetric Gaussian shape

116   with a tail on the low MW side, suggesting that the dimeric state may be in a dynamic equilibrium.

117   To better understand the potential dynamic nature of dimerization in solution, we turned to

118   sedimentation velocity analytical ultracentrifugation (AUC, **Fig. 1B** and **Table S1**), performed at

119   a range of temperatures (12-24 $^{\circ}$C) and sample concentrations (3.3, 8.5 and 18.0 µM). SEDFIT

120   $c(S)$ analysis of sedimentation velocity data collected at 16 $^{\circ}$C revealed the presence of a major

121   sedimenting species in the IFIT1 preparation that exhibited a concentration-dependent increase in

122   sedimentation coefficient (**Fig. 1B**). Such a concentration-dependent change in sedimentation

123   coefficient is consistent with a reversible monomer-dimer interaction occurring during the run

124   (28). Furthermore, the observed major $c(S)$ peak was observed to narrow as a function of

125   concentration, being broadest at the lowest concentration measured, consistent with an overall

126   change in shape of the sedimenting species. Similar concentration-dependent increases in $S_{20,W}$

127   were observed between 12 $^{\circ}$C and 24 $^{\circ}$C (**Table S1**, column 3). The apparent MW of the major

128   sedimenting species calculated by SEDFIT varied between 85-90 kDa, between the monomer MW

129   (55.5 kDa) and dimer MW (111 kDa), as expected for a rapidly reversible monomer-dimer

130   equilibrium occurring during the sedimentation velocity run. That the apparent MW of the major

131   sedimenting species is closer to the predicted dimer MW, suggests that the equilibrium position is

132   towards dimer formation at all concentrations examined. At the lowest concentration, the major

133   sedimenting species accounted for roughly ~75 % of the total optical density at 280 nm, which

134   increased to ~95 % at the highest concentration (**Table S1**, column 4).

135   **Crystallization and structure determination of human IFIT1**

136   To gain insight into the mechanism of IFIT1 dimerization, we crystallized full-length human IFIT1

137   (residues 1-478) in complex with PPP- and m7Gppp- containing oligoadenylates (oligoA, four

138   nucleotides in length). RNA-bound IFIT1 crystallized as a dimer with two molecules in the

139   asymmetric unit (ASU), and the crystals diffracted X-rays to resolutions of 2.65 Å (PPP-oligoA),

140   and 2.8 Å (m7Gppp-oligoA). The initial structures were solved by molecular replacement using

141   RNA-bound IFIT5 as a search model, and subsequent structures determined by rigid body

142   refinement (see **Table S2** for data collection and refinement statistics and **Fig. S1 and Fig. S2** for

143   representative protein and RNA electron density, respectively). Superposition of all four protein

144 chains from the two co-crystal structures yields backbone root mean square deviations (r.m.s.d.)

145 of 0.2-0.5 Å for all C_α atoms, indicating minimal differences between the two molecules of the

146 ASU or between the different RNA-bound forms (**Fig. S3A**).

147   We also determined a high resolution, 1.58 Å structure of an RNA-bound, monomeric

148 variant of IFIT1, generated with the aid of the IFIT1 dimer structure determined here (discussed

149 below). The structure of each molecule of IFIT1 from the wild-type dimer form is similar to the

150 monomeric variant, except for small changes at flexible regions and near the protein C-terminus

151 that are likely due to differences in crystal packing (**Fig. S3, B and C**). The details of IFIT1-RNA

152 binding have already been described in an earlier study (6), and here we focus mainly on the

153 mechanism of dimerization and conformational changes associated with RNA binding.

154 **Overall structure of the RNA-bound human IFIT1 dimer**

155 The two molecules of the ASU interact in a tail-to-tail manner exclusively through their C-terminal

156 regions to form a near two-fold symmetric, S-shaped dimer with ~173° rotation between them and

157 approximate dimensions of 140 Å x 80 Å x 50 Å (**Fig. 2A**). Each monomer of IFIT1 is composed

158 of 23 α-helices, 18 originating from its 9 TPR motifs, organized into 3 distinct subdomains (**Fig.**

159 **2A**). As with IFIT5, the subdomains come together to form a positively-charged RNA binding

160 tunnel at the core of the protein, which in IFIT1 encircles the cap and four RNA nucleotides (**Fig.**

161 **2, A, B and C**) giving rise to the preference for binding only ssRNA with unstructured 5′ ends or

162 dsRNA with 5′ overhangs (5, 6, 23). The m7G moiety of the cap sits at one end of the tunnel inside

163 a relatively hydrophobic cap-binding pocket between subdomains I and II (**Fig. 2A, B and D**). The

164 m7G electron density in the 2.8 Å model is not well defined (**Fig. S2**), consistent with the presence

165 of two base conformations (*syn* and *anti*) observed in the 1.58 Å structure of monomeric IFIT1

166 with m7Gppp-RNA (6). As before, one side of the m7G base interacts with Trp 147 (through π-π

167 stacking) and Ile 183, while the other contacts Leu 46 and Thr 48 from a 'cap-binding loop' (**Fig.**

168 **2D**).

169   The RNA-binding tunnel continues along the central super-helix of the protein, spanning

170   subdomain II (TPR3-TPR6), the pivot region (α15+α16), and the N-terminal portion of subdomain

171   III (TPR7-TPR8) (**Fig. 2A**). The 3′ end of the RNA (N4) emerges from the tunnel and points

172   towards a positively-charged, solvent exposed groove which may contribute to non-specific RNA

173   binding (**Fig. 2C**) (6). As with IFIT5, α21 intervenes between TPR8 and TPR9 to cap the central

174   super-helix and alter the right-handed super-helical trajectory at the end of each molecule (**Fig.**

175   **2A**). TPR9, in turn, mediates IFIT1 dimerization. The overall arrangement of IFIT1 dimers thus

176   positions the two RNA-binding tunnels in a roughly antiparallel fashion with respect to one

177   another, and results in two non-contiguous RNA-binding surfaces (**Fig. 2, A and C**).

178   The elongated IFIT1 dimer is distinct from that of IFIT2 (44 % sequence identity), which

179   forms a parallel side-to-side dimer through N-terminal domain swapping (**Fig. 2E**) (8). However,

180   the structure of individual IFIT1 monomers is very similar to IFIT2 and IFIT5 with backbone

181   r.m.s.d's of 2-4 Å (**Fig. S4A**), which improves to 0.44-1.1 Å when only subdomains I, II, or III are

182   used in the superposition (**Fig. S4B**).

183   **Mechanism of IFIT1 dimerization**

184   TPR motifs are generally composed of a degenerate, 34 amino-acid sequence that folds into a pair

185   of antiparallel helices (helices A and B) (4). Tandemly arranged TPR motifs will then assemble

186   into super-helical arrays through conserved hydrophobic interactions between helices A and B

187   within one motif, and between one TPR and helix A′ of the following motif. In this manner, the

188   minimal building block for a TPR array is 1.5 TPR motifs (or AB-A′), and the repeating pattern

189   (i.e. AB-A′B′-A′′B′′) continues until the array is capped at its end by an A-like helix. When

190   present at the C-terminus of a protein, this capping helix also functions as a 'solubility helix'. For

191   example, in IFIT5, the very C-terminal helix α24 follows TPR9 to bury its hydrophobic surface

192   and prevent further extension of the super-helix (**Fig. 3A**). In comparison, IFIT1 is prematurely

193   truncated because it lacks a corresponding helix α24 (**Fig. 3A and 4A**). The resulting exposed

194   hydrophobic surface on TPR9 is instead capped by an opposing TPR9 from another molecule, thus

195   forming a tail-to-tail elongated dimer.

196   The dimer interface buries ~ 530 Å^2 of surface area (determined by PISA, (29)) and is

197   composed of several small and large hydrophobic residues, including Leu 457, Tyr 460, and Leu

198   464 from each TPR9 motif (**Fig. 3, B, C and D**). The interface is further stabilized by reciprocal

199   hydrogen bonds between Tyr 460 and Glu 461, and salt bridges between Lys 448 and Glu 461

200   (**Fig. 3B**). The residues involved in dimerization are conserved across mammalian IFIT1 genes,

201   indicating an evolutionarily conserved mechanism for dimerization (**Fig. 4, A and B**). The

202   interface is also conserved in IFIT1B proteins, suggesting that they use the same mechanism of

203   homo-dimerization.

Mutational analysis of dimerization

205   The above described dimer is but one of several distinct dimerization interfaces in the crystal

206   lattice. To confirm that the IFIT1 dimerization observed in our solution experiments is the one

207   mediated by its C-terminus and not one of the other crystal packing interfaces, we performed a

208   mutational analysis of TPR9 which we assayed by analytical gel filtration (**Fig. 5A**). Wild-type

209   and mutant IFIT1 proteins were injected at 2 mg/ml (~ 36 µM) and, as controls, we compared their

210   gel-filtration profiles to that of IFIT2 (109 kDa dimer), IFIT5 (56 kDa monomer), and BSA. The

211   single point mutants L457E, Y460E, and L464E, as well as the double mutants L457A/L464A

212   (IFIT1^{DM-A}) and L457E/L464E (IFIT1^{DM-E}) all disrupted dimerization. K448E had no impact,

213   indicating that the K448-E461 salt bridge is dispensable, and E461A and E461K mutants could

214   not be expressed. SEC-MALS was repeated to confirm the monomer MW of IFIT1^{DM-E} (**Fig. S5A**),

215   and sedimentation velocity AUC of IFIT1^{DM-E} revealed that at 20 °C the SEDFIT-calculated

216   apparent MW of the major sedimenting species (52.7, 45.1, and 49.1 kDa at 3.3, 8.5 and 18.0 µM,

217   respectively) corresponded to the predicted monomer MW, with no monomer-dimer equilibrium

218   apparent. The mutations do not disrupt the protein fold, since circular dichroism (CD) spectroscopy

219   of IFIT1, IFIT1^{DM-A}, and IFIT1^{DM-E} revealed no discernable differences (**Fig. S5B**). Additionally,

220   the crystal structure of RNA-bound IFIT1^{DM-E} described in our previous analysis of IFIT1-RNA

221   interactions revealed no major differences between the wild-type and mutant protein (**Fig. S3B**).

222   To determine if TPR9 alone is responsible for dimerization, we created a chimera of human

223   IFIT5 α1-α21 (residues 1-435) and human IFIT1 α22-α23 (TPR9, residues 437-478). Analytical

224   gel filtration showed that the chimera can form homo-dimers in solution (**Fig. 5B**), and electro-

225   mobility gel shift assays (EMSAs) between the chimera and a 42-nucleotide RNA (GGG42) shows

226   that it retains binding to PPP-RNA (**Fig. 5C**). However, PPP-RNA binding by the chimera is

227   weaker compared to IFIT5, possibly because dimerization interferes with the dynamics of IFIT5

228   closure around the RNA. Taken together, the mutational analysis confirms the role of TPR9 and

229   its hydrophobic surface in mediating IFIT1 dimerization, and shows that this motif can behave as

230   an independent dimerization module.

231   **IFIT1 dimers form a pseudo-continuous TPR super-helix**

232   The central portion of IFIT1 dimers (TPR9s and several preceding helices) where they associate

233   bears a striking resemblance to small TPR domains such as the synthetic TPR protein, cTPR3 (4),

234   made up of three idealized motifs (**Fig. 6A**). In fact, the residues involved in IFIT1 dimerization

235   from the second helix of TPR9 are highly conserved when compared to consensus TPR motif

236   sequences derived from a large database of TPR containing proteins (**Fig. 6C**) (4). These residues

237   are usually critical for mediating the intra-molecular TPR stacking interactions that give rise to

238   continuous TPR arrays (**Fig. 6B**). Interestingly, however, there is one key difference between the

239   inter-molecular stack at the IFIT1 C-terminus and canonical intra-molecular TPR stacking.

240   Whereas canonical TPR stacking is between one TPR and helix A of the next, in IFIT1 the inter-

241   molecular stacking is between TPR9 and helix B of the adjoining TPR9 instead of another A-like

242   helix (**Fig. 6, A and B**). Thus, the two TPR9 motifs of each IFIT1 molecule interact in a manner

10

243   that is analogous to intra-molecular TPR stacking, and form a pseudo-continuous TPR super-helix

244   which is responsible for IFIT1 dimerization.

245   IFIT1 dimerization is dispensable for full RNA binding and inhibition of translation

246   To investigate the role of dimerization in RNA binding by IFIT1, we performed an EMSA utilizing

247   a 43-nucleotide m7Gppp-RNA with a five-residue overhang at the 5′ end, which is optimal for

248   IFIT1 binding (6). The capped RNA was fluorescently labelled at the 3′ end with pCp-Cy5, and

249   binding to wild-type IFIT1 and the monomeric mutant IFIT1^{DM-E} was tested (**Fig. 7A**). Both

250   proteins bound the capped RNA with the same apparent affinity of ~ 50-100 nM, suggesting that

251   dimerization is not required for optimal RNA binding. Moreover, SYBR-Gold stained EMSAs

252   with other unlabeled m7Gppp-RNAs, including two virus derived sequences, showed similar

253   results (**Fig. S6**).

254   We next looked at the impact of dimerization on IFIT1 translational inhibition. To address

255   this, we utilized an *in vitro* translation system consisting of Krebs extracts programmed with a

256   bicistronic mRNA reporter (**Fig. 7B**) (30). The first open reading frame (ORF) produces Firefly

257   luciferase (FF) in a cap-dependent manner, and should be susceptible to IFIT1 inhibition, while

258   the second ORF expresses *Renilla* luciferase (Ren) under the control of an internal ribosome entry

259   site (IRES) from hepatitis C virus (HCV), and would reveal the levels non-specific IFIT1 activity.

260   Titrating IFIT1 or IFIT1^{DM-E} into the extracts at concentrations of 30 nM to 5 μM resulted in

261   similar reductions of FF levels (compared to buffer control) (**Fig. 7C**). In both cases, Ren levels

262   only dropped by 5-10 % indicating minimal non-specific activity. Thus, IFIT1 and IFIT1^{DM-E} exert

263   similar translation inhibition, confirming that homo-dimerization is not required for IFIT1

264   inhibitory activity in an *in vitro* setting.

265   RNA binding is associated with large-scale conformational changes

266   The means by which RNA gains entry to its binding site in IFIT1 is not immediately apparent,

267   particularly because of the narrow shape and depth of the tunnel. To gain insight into the RNA-

11

268   dependent conformational changes, we performed limited protease digestion of IFIT1 with trypsin

269   and endoproteinase Glu-C, which revealed that both PPP-oligoA- and capped-oligoA-bound IFIT1

270   exhibit a greater degree of stabilization than RNA-free IFIT1 (**Fig. 8A**). The capped-RNA bound

271   complex was more resistant to proteolysis than PPP-RNA-bound IFIT1, which reflects either

272   specific conformational changes associated with capped-RNA binding or, more likely, the stronger

273   affinity of IFIT1 towards capped-RNA (6, 22).

274   The two major stable fragments resulting from endoproteinase Glu-C treatment of IFIT1

275   without RNA have apparent MWs of ~ 35 and ~ 32 kDa (**Fig. 8A, bottom**). We further examined

276   these using in-gel tryptic digestion followed by LC/MS/MS, which showed that both originate

277   from the protein N-terminus, namely subdomains I and II (**Fig. S7, A and B**). Intact protein MS

278   of the second fragment indicates that it is the result of cleavage at Glu 272 (**Fig. S7, C and D**), and

279   corresponds to the elastase-resistant fragment of human IFIT1 previously crystallized (**Fig. 8B**).

280   When Glu-C cleavage was repeated in 0.1 M ammonium bicarbonate buffer, we observed slower

281   kinetics of degradation, which we took advantage of to follow the IFIT1/Glu-C degradation profile

282   over a period of 17 hours (**Fig. 8C**). The extended time course shows that 32 kDa fragment is

283   derived from the ~ 35 kDa fragment, and also reveals the presence of several intermediate ~ 40-

284   48 kDa fragments within the first hour, which are likely due to cleavage at one of several positions

285   on subdomain III. Thus, it would appear that the IFIT1 C-terminus exists in one or more open

286   conformations that are susceptible to limited protease digestion, facilitating RNA entry into the

287   narrow tunnel.

288   **A flexible cap-binding loop facilitates m7G recognition**

289   Given that mRNA likely enters the tunnel from the C-terminal side of the protein, it is unclear how

290   the m7G moiety is accommodated inside the cap-binding pocket, which forms a narrow, circular

291   extension at the N-terminal side of the tunnel (**Fig. 2B**). To address this, we compared the

292   structures of PPP-RNA-bound and m7Gppp-RNA-bound IFIT1, which revealed small-scale

12

293   conformational within the cap-binding loop (residues 45-50). This loop forms one wall of the cap-

294   binding pocket and is stabilized by binding of the m7G moiety, as it is somewhat disordered in the

295   PPP-RNA structure (**Fig. 8, D and E**). However, based on crystal temperature factors it seems that

296   the loop retains its dynamic nature even in the presence of the cap (**Fig. S8A**). In the crystal

297   structure of RNA-free, N-terminal IFIT1 (composed of SD I+II (5)), the cap-binding loop also

298   appears to be mobile as it has high temperature factors (**Fig. S8B**). Interestingly, these small-scale

299   conformational changes were not evident from the high-resolution structures of the monomeric

300   IFIT1^{DM-E} bound to PPP- and m7Gppp-RNA, since the cap-binding pocket of its PPP-RNA bound

301   form was occupied by PEG molecules from the crystallization solution, artificially stabilizing the

302   cap-binding loop (6). Nevertheless, the plasticity of the cap-binding loop is likely to be an

303   important element in m7G recognition, as it would facilitate mRNA binding by leaving the cap-

304   binding pocket in a more open and accessible state (**Fig. 8E**). On the other hand, the critical cap-

305   binding residue Trp 147 is coordinated by Glu 176 in all structures, and appears to form a pre-

306   organized 'landing pad' for m7G binding (**Fig. 2D**). Supporting the importance of the W147-E176

307   interaction for IFIT1 function, disrupting this coordinating pair through mutation (W147F or

308   E176A) reduced binding to capped-RNA and translational inhibitory activity (6).

Discussion

310   The Interferon-Induced Proteins with Tetratricopeptide Repeats family has been reported to

311   mediate their antiviral effects through various mechanisms (31). Central to this ability is the TPR

312   motif, a versatile protein-protein interaction module which in IFITs has also adapted to interact

313   with virus-derived RNA (5, 6, 32). One facet of IFIT antiviral activity is the formation of an IFN-

314   dependent, multi-protein complex with IFIT1, IFIT2, and IFIT3 at its core (19). The mechanisms

315   regulating homo- and hetero-oligomerization of the IFIT core of this complex remain poorly

316   characterized. The crystal structure of IFIT2 showed that it homo-dimerizes *via* domain swapping

317   in which the N-terminal portion of its subdomain II (helices α7-α9) is exchanged between the

13

318   interacting IFIT2 molecules (8). We report here the structural basis for IFIT1 dimerization, which

319   utilizes a different mechanism for self-association. IFIT1 dimerizes through the formation of a C-

320   terminal, pseudo-continuous TPR super-helix, which forms in a manner that mimics naturally

321   occurring, intra-molecular TPR-stacking interactions that normally give rise to continuous TPR

322   arrays.

323   TPR containing proteins have been shown to assemble into homo-dimers and higher order

324   complexes using diverse mechanisms, ranging from those which resemble TPR-peptide

325   interactions (e.g. the self-assembling, magnetosome-forming MamA (33)), to interactions which

326   involve entire TPR domains encircling each other and burying up to 6000 Å^2 (e.g. subunits of the

327   anaphase-promoting complex (34)). The crystal structure of an engineered TPR domain made up

328   of consensus motifs (CTPR3Y3) has been shown to assemble into trimers with one interface

329   forming a C-terminal pseudo-continuous super-helix (35), similar to IFIT1. As with IFIT1, the

330   pseudo-continuous super-helix of CTPR3Y3 involves displacement of a C-terminal capping helix,

331   except in IFIT1 this capping helix appears to be lost due to a C-terminal truncation (when

332   compared to IFIT5). Nevertheless, our structure confirms the notion that displacement or loss of a

333   capping-helix giving rise to pseudo-continuous TPR arrays may be a general mechanism for TPR

334   protein oligomerization (35). We propose that TPR9 can be used as a TPR dimerization module to

335   facilitate the design of self-assembling TPR molecules. We demonstrated the feasibility of this by

336   grafting the IFIT1 TPR9 sequence onto human IFIT5, thereby converting it from a monomer to a

337   dimer.

338   TPR proteins have also been reported to form crystallographic dimers that wouldn't

339   otherwise form in solution (36), but we thoroughly validated the presence of IFIT1 dimerization

340   in solution using gel-filtration, SEC-MALS, and AUC. These data are consistent with the

341   formation of IFIT1 oligomers during RNA-binding gel shift assays and during protein blue-native

342   PAGE electrophoresis (22). Additionally, human and rabbit IFIT1 were shown to bind capped-

343   RNA with low-level positive cooperativity (Hill coefficient, n = 1.6-1.7) (22). However, co-

344   immunoprecipitation assays failed to detect an IFIT1 self-interaction in human cells (19), possibly

345   because the interaction is relatively weak. This is supported by our AUC analysis where we

346   observed that wild-type IFIT1 had established a rapidly reversible monomer-dimer equilibrium

347   during centrifugation. Furthermore, the relatively small dimerization interface burying only ~ 530

348    $Å^2$ is smaller than expected for a stable self-interaction (37).

349   Regardless, during AUC IFIT1 is still predominantly dimeric at 3.3 μM, and given that

350   IFIT1 is highly expressed upon induction where its levels can approach 1-2 million copies per cell

351   (19), it is likely that physiological concentrations of IFIT1 are within a range that promotes dimer

352   formation. Moreover, IFIT1 has been shown to re-localize to discrete spots following transfection

353   of PPP-RNA (19), where even higher local concentrations would favor dimerization. However,

354   although IFIT1 dimer formation *in vivo* is a likely scenario, the possibility remains that it is non-

355   physiological. Since human IFIT1 is at the center of a large multi-protein complex (19), isolating

356   it *in vitro* could give rise to artifacts. In this case, IFIT1 homo-dimerization may be due to residual

357   binding that is a consequence of IFIT1 hetero-oligomerization with other members of the

358   interactome, such as IFIT2 or IFIT3.

359   Most mammals encode an IFIT1 protein and an IFIT1B protein (3). However, in mice,

360   IFIT1 appears to be deleted and instead they have three copies of IFIT1B (3), referred to as mouse

361   Ifit1, mouse Ifit1b, and mouse Ifit1c (with 58-77 % sequence identity between them). On the other

362   hand, in humans, IFIT1B expression has not been shown to change with IFN treatment (27), and

363   heterologous overexpression in yeast suggests that it lacks a function in RNA binding and

364   translational inhibition (3). Nevertheless, the conserved nature of the IFIT1 dimerization interface

365   suggests that mammalian IFIT1B proteins homo-dimerize in the same manner, and may even form

366   hetero-dimers with IFIT1 or other IFIT1B-like proteins through the same interface. Supporting

367   this notion, mouse Ifit1c was found to co-purify with mouse Ifit1 in capped-RNA pulldowns (21).

368   Although it remains to be shown, hetero-dimerization through this interface could regulate RNA

369   binding or serve to expand the protein-protein interaction network of IFIT1, therefore diversifying

370   IFIT1-mediated antiviral responses.

371   Although foot-printing primer-extension assays performed under equilibrium binding

372   conditions revealed that human IFIT1 can bind capped-RNA with positive cooperativity (22), our

373   gel-shift RNA binding assays did not reveal a difference in apparent affinity between monomeric

374   and dimeric IFIT1. One possible explanation for this discrepancy is that EMSAs are not a true

375   equilibrium method, and complex formation is susceptible to variations in the assay conditions

376   (38, 39), therefore the apparent RNA-binding affinities of IFIT1 and the monomeric mutant could

377   be affected. Future assays performed under equilibrium conditions (e.g. primer-extension foot-

378   printing, fluorescence polarization, or isothermal titration calorimetry) could be performed to

379   clarify the role of dimerization in RNA binding. Of note, our attempts at isothermal titration

380   calorimetry with human IFIT1 were unsuccessful due protein precipitation during the experiment.

381   Regardless, the translation assays in Krebs extracts indicate that homo-dimerization is dispensable

382   for translational inhibition by IFIT1, at least under this *in vitro* setting.

383   Finally, comparison of the different RNA bound forms of IFIT1 complemented with

384   limited protease digestion revealed large-scale and small-scale conformational changes associated

385   with RNA binding. Based on the structural analysis, limited proteolysis, and previous work with

386   human IFIT5 (5), we propose the following model for capped mRNA binding by IFIT1 in solution.

387   IFIT1 exists in one or more open conformations that are amenable for RNA entry into the tunnel.

388   Upon binding the 5′ end of mRNA (nucleotides 1-4), IFIT1 undergoes conformational changes at

389   its C-terminus that are facilitated by the pivot region, as was seen in IFIT5 (5). At the same time,

390   or in a subsequent step, the m7G moiety is directed towards the cap-binding pocket, which would

391   initially be in a more open and accessible state as indicated by the dynamic nature of the cap-

392   binding loop. Recognition by Trp 147 and stabilization of the cap-binding loop would then be the

393   last in a series of coordinated steps required for mRNA binding. The RNA-dependent

394   conformational changes would help IFIT1 clamp down on viral mRNA, thus decreasing its rate of

395   dissociation, and prevent translation initiation factors from outcompeting IFIT1 once it is bound.

396   Both the large scale and small scale conformational changes described here are also likely

397   impacted by other members of the IFIT interactome, further enhancing or limiting IFIT1 antiviral

398   activity *in vivo*.

399   In summary, we have determined the crystal structure of human IFIT1, a central component

400   of the IFIT interactome, in its native, dimeric form. The structure shows that IFIT1 uses a small

401   C-terminal interface to form extended, antiparallel dimers that are distinct from the parallel,

402   domain-swapped dimers of IFIT2. This arrangement of IFIT1 dimers could combine with the

403   parallel interface found in IFIT2 to initiate interactome assembly. Alternatively, IFIT2 or IFIT3

404   could compete with the IFIT1 homo-dimerization interface to build up the interactome. As these

405   interactions could be important for regulating IFIT1 conformational changes and inhibitory

406   activity, our results will contribute to piecing together the network of interactions that regulate

407   IFIT proteins' role in the antiviral innate immune response.

408   **Experimental Procedures**

409   *Plasmids*– IFIT1, IFIT2, and IFIT5 were sub-cloned into pSMT3 between BamHI and NotI sites,

410   resulting in a fusion protein with an N-terminal, Ulp1-cleavable 6xHis-Sumo tag (40). IFIT1 point

411   mutants were generated by site-directed mutagenesis using iProof High-Fidelity DNA polymerase

412   (Bio-Rad). The chimera made up of IFIT5 (residues 1-435) and IFIT1 (residues 437-478) was

413   generated using overlap-extension PCR, and sub-cloned into pSMT3 between BamHI and NotI.

414   *Expression and purification of IFIT proteins*– IFIT1, IFIT1 mutants and IFIT5 expression and

415   purification are described elsewhere (6). The IFIT5/IFIT1 chimera was expressed in Rosetta 2

416   (DE3) pLysS BL21 cells (Novagen) by first growing transformed cells in LB at 37 $^\circ$C until an

417   OD_{600} of 0.6-0.8, followed by induction with 1 mM IPTG and growing the cells for 4 hours at 30

418 $^{\circ}$C. The protein was purified by Ni-affinity chromatography followed by cleavage of the tag. The

419 cleaved protein was further purified by passing over a 5 ml HiTrap SP HP column (GE Healthcare)

420 equilibrated with 50 mM HEPES pH 7.5 and 1 mM DTT, followed by washing with 50 mM

421 HEPES pH 7.5, 100 mM NaCl, and 1 mM DTT, and eluted with 50 mM HEPES pH 7.5, 400 mM

422 NaCl, and 1 mM DTT. IFIT2 was expressed in BL21 (DE3) and purified using 2-step Ni-affinity

423 chromatography (with cleavage of the tag) using standard protocols. All proteins were further

424 purified by gel filtration using superdex 200 10/300 or 16/60 columns (GE Healthcare) in 20 mM

425 Tris pH 7.6, 150 mM NaCl and 1 mM DTT.

426 *Preparation of PPP and m7Gppp containing oligoadenylate RNA–* PPP-AAAA and Gppp-AAAA

427 were prepared using phosphoramidite solid phase synthesis followed by installation of the

428 triphosphate or cap moieties as previously described (6); the oligos were purified by HPLC and

429 characterized by LC-MS as before (6). m7Gppp-AAAA was prepared by enzymatically modifying

430 Gppp-AAAA with purified mRNA cap guanine-N7 methyltransferase as before (6).

431 *Crystallization and data collection–* Wild-type human IFIT1 was mixed with molar excess

432 chemically synthesized oligos either directly before crystallization or prior to gel filtration. The

433 protein buffer was 20 mM Tris pH 7.6, 150 mM NaCl and 1 mM TCEP. The co-crystals were

434 obtained with 2-4 mg/ml protein drops mixed 1:1 with reservoir solution containing 27-32 % PEG

435 200 (Sigma), 0.1 M Tris pH 8.1, and 200 mM CaCl2 using the hanging drop vapour diffusion

436 method at 4 $^{\circ}$C. Crystals were flash frozen in liquid nitrogen without additional cryo-protection.

437 All data were collected at 100K with 0.979 Å X-rays on beamline 08ID-1 at the CLS, which was

438 equipped with a Mar300CCD detector (41), and the images integrated, scaled, and merged using

439 the HKL2000 suite (43). Data were truncated with Ctruncate (ccp4 suite (42)) and 5 % of

440 reflections were set aside for the R_{free} set. Only the highest resolution dataset was used as a master

441 file to generate all R_{free} sets. Data collection statistics from HKL2000 are in **Table S2**.

18

442   *Structure determination, model building, and refinement*– The initial crystal structure was

443   determined by molecular replacement using the PHASER program in Phenix (43). The structure

444   of RNA-bound human IFIT5 (residues 1-434) was used as a search model. The remaining 3 helices

445   of IFIT1 (residues 437-469) were built manually after one macro-cycle of refinement in Phenix

446   (45). Subsequent structures were determined with rigid body refinement with the protein only.

447   Restraints for m7GpppA were calculated using Phenix eLBOW (44). The structures were refined

448   iteratively using Phenix with manual model building in Coot (46). The refinement strategy

449   included all-isotropic B-factor refinement and TLS. The final models contained protein residues

450   9-84, 91-305, and 310-469 for m7Gppp-RNA bound IFIT1 chain A; 9-84, 90-304, and 315-469

451   for m7Gppp-RNA bound IFIT1 chain C; 9-46, 51-81, 92-247, 251-301, and 316-469 for PPP-RNA

452   bound IFIT1 chain A; and 9-46, 51-81, 92-247, 251-305, and 314-469 for PPP-RNA bound IFIT1

453   chain C. Structure validation was performed with MOLPROBITY in Phenix (47). Ramachandran

454   statistics are as follows: IFIT1 with m7Gppp-RNA, 98.5 % favored, 0 % outliers, 0.25 % rotamer

455   outliers; and IFIT1 with PPP-RNA: 98.2 % favored, 0 % outliers, 0.4 % rotamer outliers. The

456   MOLPROBITY overall scores for the m7Gppp-RNA and PPP-RNA-bound structures were 0.58

457   and 0.63, respectively. Refinement statistics are in **Table S2**.

458   *Sequence and structure analysis*– APBS was used to calculate the surface electrostatic potential

459   (48), and PyMol to general all molecular figures (https://www.pymol.org/). For surface

460   electrostatic analysis, all surfaces are colored by electrostatic potential from negative (-10 kTe-1;

461   red) to positive (+10 kTe-1; blue). ESPript was used to generate the sequence alignment in Fig.

462   4A (49). Sequence conservation analysis by WebLogo (50) in Fig. 4B was performed as described

463   (6).

464   *Analytical gel filtration*– 500 μL (Fig. 5A) or 250 μL (Fig. 5B) of purified protein at 2 mg/ml were

465   injected onto a superdex 200 10/300 increase column (GE Healthcare) equilibrated in 20 mM Tris

19

466   pH 7.6 and 150 mM NaCl. 1 µg (Fig. 5A) or 2 µg (Fig. 5B) of the input proteins were run on 12

467   % SDS-PAGE and visualized by coomassie staining.

468   *Analytical ultracentrifugation*– Sedimentation velocity experiments were performed in a Beckman

469   XL-I analytical ultracentrifuge using an An-60Ti rotor. Solutions containing 20 mM Tris pH 7.6,

470   150 mM NaCl, 0.5 mM TCEP and varying concentrations of IFIT1 (3.3 µM, 8.5 µM , and 18.0

471   µM) were centrifuged at 50,000 rpm over approximately 8h. A total of 200 scans were collected

472   per run, with absorbances measured at 280 nm. Sedimentation velocity data were fit to the $c(S)$

473   model using the computer program SEDFIT (28). Sedimentation coefficients were determined by

474   integration of c(S) distributions in SEDFIT and then transformed into $S_{20,W}$ values. The program

475   SEDNTERP (v1.09, http://www.jphilo.mailway.com/) was used for calculation of solution

476   viscosity, solution density, and IFIT1 partial specific volume (from primary amino acid sequence)

477   at each temperature examined.

478   *Multi-angle light scattering with inline size-exclusion chromatography (SEC-MALS)*– 50 µL of

479   purified IFIT1 (10 mg/ml), IFIT1[DM-E] (12 mg/ml), or BSA (10.5 mg/ml) were injected onto

480   superdex 200 10/300 increase (GE Healthcare) equilibrated with 20 mM Tris pH 7.6 and 150 mM

481   NaCl. Measurements were performed using a Waters e2796 Separations module (Waters

482   Corporation) equipped with an in-line Waters 2489 UV/Visible detector set to 280nm (Waters

483   corporation), a miniDAWN TREOS multi-angle light scattering detector with a 656nm laser

484   (Wyatt Technology), and an Optilab rEX differential refractive index detector (Wyatt

485   Technology). Data was processed and analyzed using the ASTRA software package (version

486   5.3.4.18 from Wyatt technology).

487   *Limited Protease Digestion*–PPP-oligoA or Gppp-oligoA were re-suspended in water at a final

488   concentration of 10 nmol/µL. 1.26 nmol IFIT1 (70 µg) in gel filtration buffer were mixed with

489   either 15 µL water, 15 µL PPP-oligoA (1.5 nmol), or 15 µL Gppp-oligoA (1.5 nmol) at a final

490   volume of 70 µL, resulting in 10 mg/ml (10X) protein stocks with or without RNA. 10X protease

20

491   stocks were prepared in 20 mM HEPES pH 7.5, 50 mM KCl, and 20 mM Mg_2SO_4 (Cleavage

492   buffer), resulting in a 1 mg/ml endoproteinase Glu-C stock and 0.04 mg/ml trypsin stock. For each

493   reaction (corresponding to each lane in Fig. 8A), 1.6 µL of 10X protein and 1.6 µL of 10X protease

494   were mixed at a final volume of 16 µL in Cleavage buffer and the reaction carried out at room

495   temperature. At the indicated time points, the reactions were stopped with 4 µL of 5X SDS loading

496   buffer, boiled for 10 minutes, and stored at -20 °C until the samples were ready to be run on 12 %

497   SDS-PAGE. For the extended time-course (also performed at room temperature), 150 µg IFIT1

498   was mixed with 15 µg Glu-C in 150 µL containing 0.1 ammonium bicarbonate buffer pH 7.8. At

499   the indicated time points, 16 µL of the reaction were removed and mixed with 4 µL of 5X SDS

500   loading buffer, then boiled for 10 minutes, and stored at -20 °C until the samples were ready to be

501   run on 12 % SDS-PAGE.

502   *RNA electro-mobility gel shift assays (EMSA)*– Capped RNAs for EMSAs were prepared by *in*

503   *vitro* transcription with T7 RNA polymerase followed by PAGE purification and post-

504   transcriptional capping with the vaccinia capping system (NEB) as previously described (6).

505   Preparation of pCp-labelled RNA was performed with T4 RNA ligase (NEB) as previously

506   described (6). Gel shifts were performed exactly as described before (6).

507   *Circular dichroism (CD) spectroscopy*– IFIT1, IFIT1[DM-A], or IFIT1[DM-E] were buffer exchanged

508   into 50 mM Na_2HPO_4/NaH_2PO_4 pH 8 and diluted to a final concentration of 0.5

509   mg/ml. Measurements were performed on a Chirascan CD spectrometer using a 0.2 mm quartz

510   cell. CD wavelength scans were obtained at 21 °C, with 1 nm step size and 0.5 s time intervals.

511   The signals obtained were in CD millidegrees and are represented as an average of three spectra

512   obtained for each sample. The averaged spectra were smoothened using the Savitsky-Golay

513   method and a window size of 3.

514   *In vitro translation assay with Krebs extracts*– Preparation of reporter mRNA and *in vitro*

515   translation assay were performed as before (6). Briefly, pSP-(CAGless)/FF/HCV/Ren plasmid was

21

516   linearized with BamHI and transcribed with SP6 RNA polymerase in the presence of

517   m7G(5′)ppp(5′)G cap analog (NEB). *In vitro* translations were set up at a final volume of 10 µl

518   with 4 ng/μL reporter mRNA (~ 4 nM final) and 1 µL wild-type IFIT1 or IFIT1^{DM-E} in untreated

519   Krebs-2 extracts. The reactions were set up on ice and translation was allowed to proceed for 1 hr

520   at 30 °C. The reactions were stopped on ice, and FF and Ren luciferase activities (RLU) were

521   measured on a Berthold Lumat LB 9507 luminometer. Values obtained were normalized against

522   buffer control, which was set at 1. The following final protein concentrations were used: 0.031,

523   0.062, 0.125, 0.250, 0.5, 1, and 5 µM.

524   **Accession codes**

525   Coordinates and structure factor data for the structures reported in this study were deposited in the

526   Protein Data Bank under the accession codes 5W5H (IFIT1 + m7Gppp-AAAA) and 5W5I (IFIT1

527   + PPP-AAAA)

528   **Acknowledgements**

541   Council of Canada (M.J.D. and P.D.P.); and the McGill CIHR Drug Development Training

542   Program (S.M-M.).

543   **Conflict of interest**

544   The authors declare that they have no conflicts of interest with the contents of this article.

545   **Author Contributions**

546   B.N. and Y.M.A. designed the study and prepared the manuscript with input from all authors;

547   B.N., J.P., and M.J.D. supervised the work; Y.M.A. prepared proteins and RNA, and performed

548   structural work, SEC-MALS, limited proteolysis, CD spectroscopy, analytical gel filtration, and

549   EMSAs; S.M-M. performed chemical synthesis and purification of PPP- and Gppp-oligoA; R.C.

550   prepared translation extracts and reporter mRNA, and performed translation assays; P.D.P.

551   performed AUC data collection and analysis.

552   **References**

553   1.   Fensterl, V., Chattopadhyay, S., and Sen, G. C. (2015) No Love Lost Between Viruses
554   and Interferons. *Annual Review of Virology*. **2**, 549–572
555   2.   Fensterl, V., and Sen, G. C. (2015) Interferon-induced Ifit proteins: their role in viral
556   pathogenesis. *J Virol*. **89**, 2462–2468
557   3.   Daugherty, M. D., Schaller, A. M., Geballe, A. P., and Malik, H. S. (2016) Evolution-
558   guided functional analyses reveal diverse antiviral specificities encoded by IFIT1 genes in
559   mammals. *elife*. **5**, 60
560   4.   Main, E. R. G., Xiong, Y., Cocco, M. J., D'Andrea, L., and Regan, L. (2003) Design of
561   stable alpha-helical arrays from an idealized TPR motif. *Structure*. **11**, 497–508
562   5.   Abbas, Y. M., Pichlmair, A., Górna, M. W., Superti-Furga, G., and Nagar, B. (2013)
563   Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature*. **494**,
564   60–64
565   6.   Abbas, Y. M., Laudenbach, B. T., Martínez-Montero, S., Cencic, R., Habjan, M.,
566   Pichlmair, A., Damha, M. J., Pelletier, J., and Nagar, B. (2017) Structure of human IFIT1
567   with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and
568   N2 ribose 2'-O methylations. *Proceedings of the National Academy of Sciences*.
569   10.1073/pnas.1612444114
570   7.   Katibah, G. E., Lee, H. J., Huizar, J. P., Vogan, J. M., Alber, T., and Collins, K. (2013)
571   tRNA binding, structure, and localization of the human interferon-induced protein IFIT5.
572   *Mol Cell*. **49**, 743–750
573   8.   Yang, Z., Liang, H., Zhou, Q., Li, Y., Chen, H., Ye, W., Chen, D., Fleming, J., Shu, H.,
574   and Liu, Y. (2012) Crystal structure of ISG54 reveals a novel RNA binding structure and
575   potential functional mechanisms. *Cell Res*. **22**, 1328–1338

576 9. Feng, F., Yuan, L., Wang, Y. E., Crowley, C., Lv, Z., Li, J., Liu, Y., Cheng, G., Zeng, S.,
577   and Liang, H. (2013) Crystal structure and nucleotide selectivity of human IFIT5/ISG58.
578   *Cell Res*. **23**, 1055–1058

579 10. D'Andrea, L. D., and Regan, L. (2003) TPR proteins: the versatile helix. *Trends Biochem*
580   *Sci*. **28**, 655–662

581 11. Feng, X., Wang, Y., Ma, Z., Yang, R., Liang, S., Zhang, M., Song, S., Li, S., Liu, G., Fan,
582   D., and Gao, S. (2014) MicroRNA-645, up-regulated in human adencarcinoma of gastric
583   esophageal junction, inhibits apoptosis by targeting tumor suppressor IFIT2. *BMC*
584   *Cancer*. **14**, 633

585 12. Xiao, S., Li, D., Zhu, H.-Q., Song, M.-G., Pan, X.-R., Jia, P.-M., Peng, L.-L., Dou, A.-X.,
586   Chen, G.-Q., Chen, S.-J., Chen, Z., and Tong, J.-H. (2006) RIG-G as a key mediator of the
587   antiproliferative activity of interferon-related pathways through enhancing p21 and p27
588   proteins. *Proc Natl Acad Sci USA*. **103**, 16448–16453

589 13. Lai, K. C., Liu, C. J., Chang, K. W., and Lee, T. C. (2013) Depleting IFIT2 mediates
590   atypical PKC signaling to enhance the migration and metastatic activity of oral squamous
591   cell carcinoma cells. *Oncogene*. **32**, 3686–3697

592 14. Lai, K.-C., Chang, K.-W., Liu, C.-J., Kao, S.-Y., and Lee, T.-C. (2008) IFN-induced
593   protein with tetratricopeptide repeats 2 inhibits migration activity and increases survival of
594   oral squamous cell carcinoma. *Mol. Cancer Res.* **6**, 1431–1439

595 15. Stawowczyk, M., Van Scoy, S., Kumar, K. P., and Reich, N. C. (2011) The interferon
596   stimulated gene 54 promotes apoptosis. *Journal of Biological Chemistry*. **286**, 7257–7266

597 16. Berchtold, S., Manncke, B., Klenk, J., Geisel, J., Autenrieth, I. B., and Bohn, E. (2008)
598   Forced IFIT-2 expression represses LPS induced TNF-alpha expression at
599   posttranscriptional levels. *BMC Immunol.* **9**, 75

600 17. Li, Y., Li, C., Xue, P., Zhong, B., Mao, A.-P., Ran, Y., Chen, H., Wang, Y.-Y., Yang, F.,
601   and Shu, H.-B. (2009) ISG56 is a negative-feedback regulator of virus-triggered signaling
602   and cellular antiviral response. *Proceedings of the National Academy of Sciences*. **106**,
603   7945–7950

604 18. Liu, X. Y., Chen, W., Wei, B., Shan, Y. F., and Wang, C. (2011) IFN-Induced TPR
605   Protein IFIT3 Potentiates Antiviral Signaling by Bridging MAVS and TBK1. *The Journal*
606   *of Immunology*. **187**, 2559–2568

607 19. Pichlmair, A., Lassnig, C., Eberle, C.-A., Górna, M. W., Baumann, C. L., Burkard, T. R.,
608   Bürckstümmer, T., Stefanovic, A., Krieger, S., Bennett, K. L., Rülicke, T., Weber, F.,
609   Colinge, J., Müller, M., and Superti-Furga, G. (2011) IFIT1 is an antiviral protein that
610   recognizes 5′-triphosphate RNA. *Nat Immunol*. **12**, 624–630

611 20. Kimura, T., Katoh, H., Kayama, H., Saiga, H., Okuyama, M., Okamoto, T., Umemoto, E.,
612   Matsuura, Y., Yamamoto, M., and Takeda, K. (2013) Ifit1 inhibits JEV replication
613   through binding to 5′ capped 2′-O unmethylated RNA. *J Virol*. **87**, 9997–10003

614 21. Habjan, M., Hubel, P., Lacerda, L., Benda, C., Holze, C., Eberl, C. H., Mann, A., Kindler,
615   E., Gil-Cruz, C., Ziebuhr, J., Thiel, V., and Pichlmair, A. (2013) Sequestration by IFIT1
616   Impairs Translation of 2′O-unmethylated Capped RNA. *PLoS Pathog*. **9**, e1003663

617 22. Kumar, P., Sweeney, T. R., Skabkin, M. A., Skabkina, O. V., Hellen, C. U. T., and
618   Pestova, T. V. (2014) Inhibition of translation by IFIT family members is determined by
619   their ability to interact selectively with the 5"-terminal regions of cap0-, cap1- and 5"ppp-
620   mRNAs. *Nucleic Acids Res*. **42**, 3228–3245

621 23. Hyde, J. L., Gardner, C. L., Kimura, T., White, J. P., Liu, G., Trobaugh, D. W., Huang, C.,
622   Tonelli, M., Paessler, S., Takeda, K., Klimstra, W. B., Amarasinghe, G. K., and Diamond,
623   M. S. (2014) A viral RNA structural element alters host recognition of nonself RNA.
624   *Science*. **343**, 783–787
625 24. Daffis, S., Szretter, K. J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.-Y., Schneller, S.,
626   Zust, R., Dong, H., Thiel, V., Sen, G. C., Fensterl, V., Klimstra, W. B., Pierson, T. C.,
627   Buller, R. M., Gale, M., Jr, Shi, P.-Y., and Diamond, M. S. (2010) 2′-O methylation of the
628   viral mRNA cap evades host restriction by IFIT family members. *Nature*. **468**, 452–456
629 25. Zust, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B. W., Ziebuhr, J.,
630   Szretter, K. J., Baker, S. C., Barchet, W., Diamond, M. S., Siddell, S. G., Ludewig, B., and
631   Thiel, V. (2011) Ribose 2′-O-methylation provides a molecular signature for the
632   distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol*.
633   **12**, 137–143
634 26. Diamond, M. S. (2014) IFIT1: A dual sensor and effector molecule that detects non-2′-O
635   methylated viral RNA and inhibits its translation. *Cytokine Growth Factor Rev*.
636   10.1016/j.cytogfr.2014.05.002
637 27. Fensterl, V., and Sen, G. C. (2011) The ISG56/IFIT1 gene family. *Journal of Interferon &*
638   *Cytokine Research*. **31**, 71–78
639 28. Schuck, P., and Balbo, A. (2005) Analytical Ultracentrifugation in the Study of Protein
640   Self- association and Heterogeneous Protein-Protein Interactions. *Cold Spring Harbor*
641   *Laboratory Press*
642 29. Krissinel, E., and Henrick, K. (2007) Inference of Macromolecular Assemblies from
643   Crystalline State. *J Mol Biol*. **372**, 774–797
644 30. Novac, O., Guenier, A. S., and Pelletier, J. (2004) Inhibitors of protein synthesis identified
645   by a high throughput multiplexed translation screen. *Nucleic Acids Res*. **32**, 902–915
646 31. Diamond, M. S., and Farzan, M. (2013) The broad-spectrum antiviral functions of IFIT
647   and IFITM proteins. *Nat Rev Immunol*. **13**, 46–57
648 32. Pichlmair, A., Schulz, O., Tan, C. P., Rehwinkel, J., Kato, H., Takeuchi, O., Akira, S.,
649   Way, M., Schiavo, G., and Reis e Sousa, C. (2009) Activation of MDA5 Requires Higher-
650   Order RNA Structures Generated during Virus Infection. *J Virol*. **83**, 10761–10769
651 33. Zeytuni, N., Ozyamak, E., Ben-Harush, K., Davidov, G., Levin, M., Gat, Y., Moyal, T.,
652   Brik, A., Komeili, A., and Zarivach, R. (2011) Self-recognition mechanism of MamA, a
653   magnetosome-associated TPR-containing protein, promotes complex assembly.
654   *Proceedings of the National Academy of Sciences*. **108**, E480–7
655 34. Zhang, Z., Chang, L., Yang, J., Conin, N., Kulkarni, K., and Barford, D. (2013) The Four
656   Canonical TPR Subunits of Human APC/C Form Related Homo-Dimeric Structures and
657   Stack in Parallel to Form a TPR Suprahelix. *J Mol Biol*. **425**, 4236–4248
658 35. Krachler, A. M., Sharma, A., and Kleanthous, C. (2010) Self-association of TPR domains:
659   Lessons learned from a designed, consensus-based TPR oligomer. *Proteins*.
660   10.1002/prot.22726
661 36. Kajander, T., Cortajarena, A. L., Mochrie, S., and Regan, L. (2007) Structure and stability
662   of designed TPR protein superhelices: unusual crystal packing and implications for natural
663   TPR proteins. *Acta Crystallogr D Biol Crystallogr*. **63**, 800–811
664 37. Conte, Lo, L., Chothia, C., and Janin, J. (1999) The atomic structure of protein-protein
665   recognition sites. *J Mol Biol*. **285**, 2177–2198

25

666   38.   Gaudreault, M., Gingras, M.-E., Lessard, M., Leclerc, S., and Guérin, S. L. (2009)
667   Electrophoretic Mobility Shift Assays for the Analysis of DNA-Protein Interactions. in
668   *Nucleic Acid Crystallography* (Ennifar, E. ed), pp. 15–35, Methods in Molecular
669   Biology™, Humana Press, Totowa, NJ, **543**, 15–35
670   39.   Fried, M. G., and Liu, G. (1994) Molecular sequestration stabilizes CAP–DNA complexes
671   during polyacrylamide gel electrophoresis. *Nucleic Acids Res*. **22**, 5054–5059
672   40.   Mossessova, E., and Lima, C. D. (2000) Ulp1-SUMO crystal structure and genetic
673   analysis reveal conserved interactions and a regulatory element essential for cell growth in
674   yeast. *Mol Cell*. **5**, 865–876
675   41.   Grochulski, P., Fodje, M. N., Gorin, J., Labiuk, S. L., Berg, R., IUCr (2011) Beamline
676   08ID-1, the prime beamline of the Canadian Macromolecular Crystallography Facility. *J
677   Synchrotron Radiat*. **18**, 681–684
678   42.   Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R.,
679   Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J.,
680   Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A.,
681   Wilson, K. S., IUCr (2011) Overview of the CCP4 suite and current developments. *Acta
682   Crystallogr D Biol Crystallogr*. **67**, 235–242
683   43.   McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., Read,
684   R. J., IUCr (2007) Phaser crystallographic software. *J Appl Crystallogr*. **40**, 658–674
685   44.   Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009) electronic Ligand
686   Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint
687   generation. *Acta Crystallogr D Biol Crystallogr*. **65**, 1074–1080
688   45.   Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd,
689   J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W.,
690   Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and
691   Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular
692   structure solution. *Acta Crystallogr D Biol Crystallogr*. **66**, 213–221
693   46.   Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development
694   of Coot. *Acta Crystallogr D Biol Crystallogr*. **66**, 486–501
695   47.   Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G.
696   J., Murray, L. W., Richardson, J. S., Richardson, D. C., IUCr (2010) MolProbity: all-atom
697   structure validation for macromolecular crystallography. *Acta Crystallogr D Biol
698   Crystallogr*. **66**, 12–21
699   48.   Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001)
700   Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings
701   of the National Academy of Sciences*. **98**, 10037–10041
702   49.   Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the
703   new ENDscript server. *Nucleic Acids Res*. **42**, W320–4
704   50.   Schneider, T. D., and Stephens, R. M. (2004) Sequence logos: a new way to display
705   consensus sequences. *Nucleic Acids Res*. **18**, 1–4
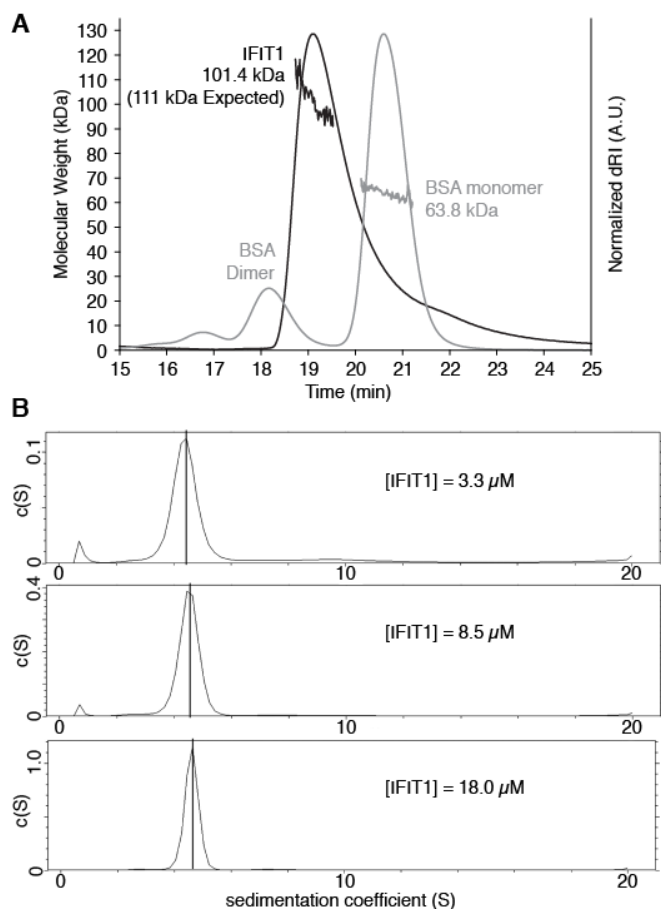
Figure 1. Human IFIT1 forms dimers in solution.
(A) SEC-MALS of human IFIT1 injected at 10 mg/ml on a Superdex 200 Increase 10/300 column. Shown are the estimated molecular weight and normalized differential refractive index (dRI) as a function of time. BSA (10.5 mg/ml) was performed as a control. (B) Sedimentation velocity analytical ultracentrifugation of IFIT1. Sedimentation data were collected at 16 °C and analyzed using SEDFIT. Upper panel: c(S) distribution of IFIT1 at 3.3 μM; middle panel: c(S) distribution of IFIT1 at 8.5 μM; lower panel: c(S) distribution of IFIT1 at 18.0 μM. Vertical lines in peaks correspond to untransformed sedimentation coefficients (4.40 S at 3.3 μM, 4.52 S at 8.5 μM, and 4.61 S at 18.0 μM) calculated by peak integration of c(S) distribution using SEDFIT. Sedimentation coefficients transformed to $S_{20,w}$ values can be found in **Table S1**.
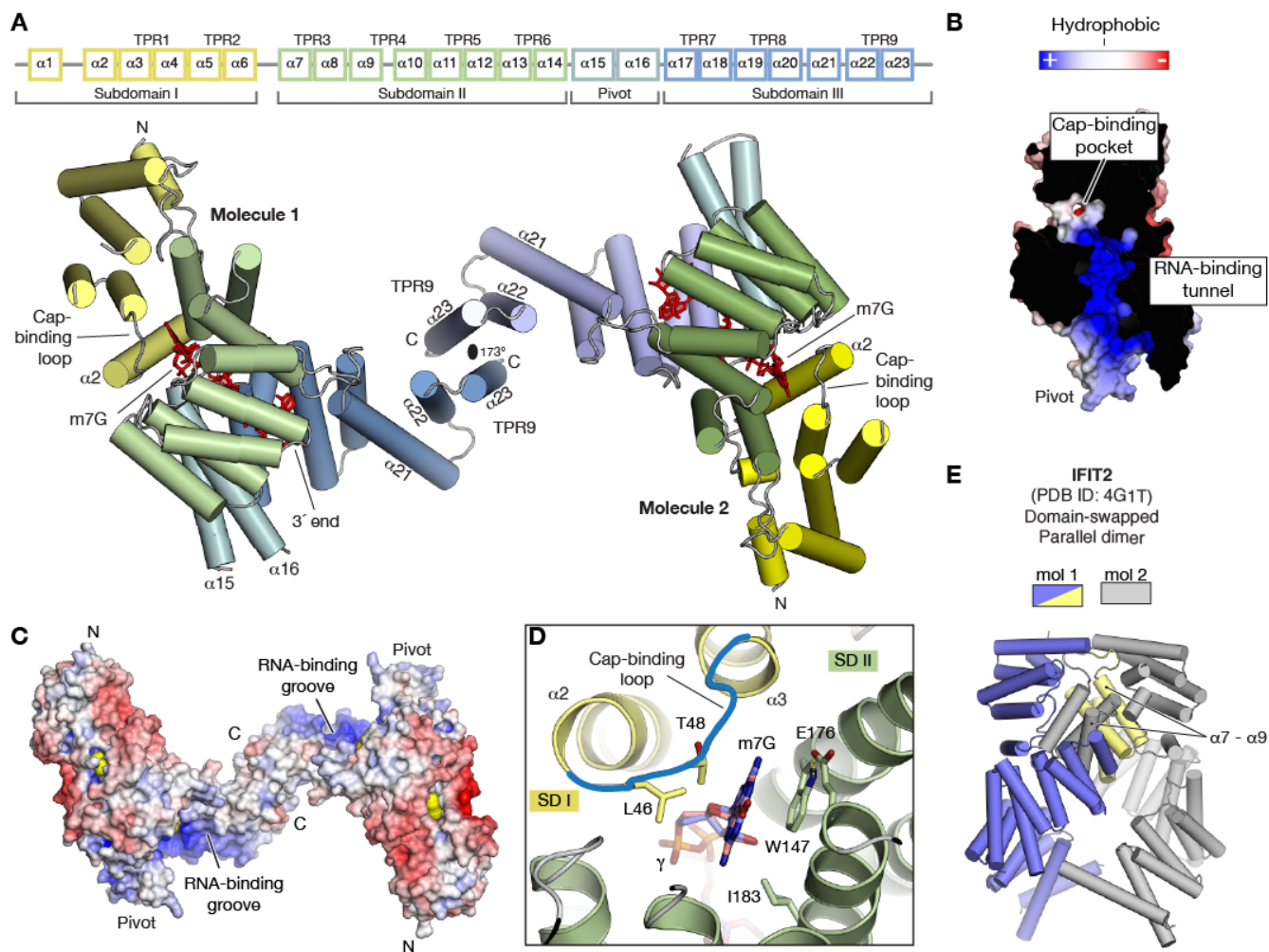
Figure 2. Overall structure of RNA-bound human IFIT1 dimer.
(A) Schematic of human IFIT1 subdomains and cartoon representation of the overall structure colored by subdomain. **(B)** Cross section of the IFIT1 RNA-binding tunnel and **(C)** IFIT1 dimers colored by surface electrostatic potential from negative (-10 kTe^{-1}; red) to positive (+10 kTe^{-1}; blue). The 3′ end of the RNA exits from the RNA-binding tunnel and points towards an RNA-binding groove at the C-terminus of each molecule. **(D)** Close-up of the cap-binding pocket with m7G interacting residues represented as sticks and colored according to subdomain. **(E)** In IFIT2, dimerization is through domain-swapping by exchanging helices α7-α9 (pale yellow and grey) from subdomain II between the two molecules.

28
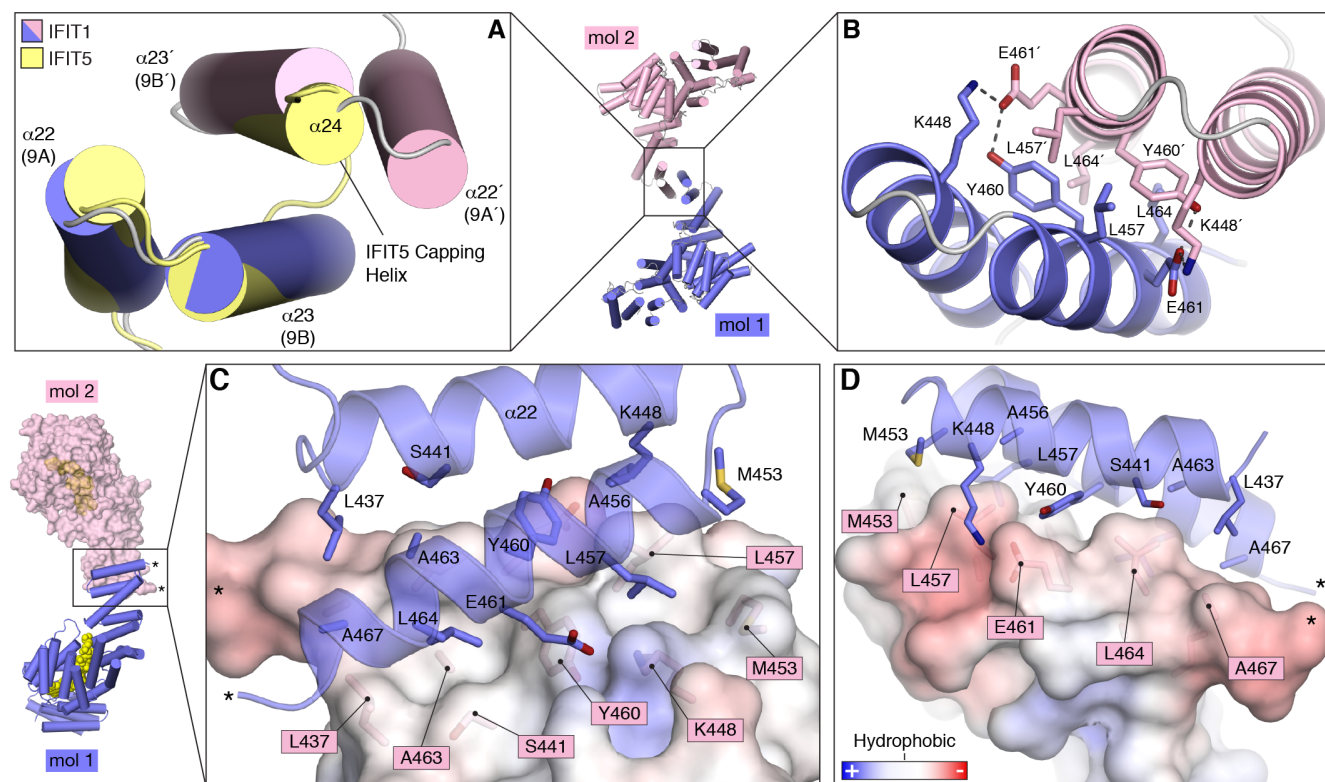
Figure 3. Structural basis for IFIT1 dimerization.

(A) Close-up of TPR9 (helices 9A and 9B) from human IFIT1 with IFIT5 superposed. Compared to IFIT5, IFIT1 proteins are truncated and lack helix α24, which in IFIT5 functions as a C-terminal capping helix. The two protomers are referred to as IFIT1 (blue) and IFIT1′ (light pink). **(B)** Cartoon/stick representation of the critical residues at the IFIT1 dimerization interface in the same orientation as (A). These residues make up most of the buried surface area between the dimers, and were targeted in the mutational analysis. **(C and D)** Two orthogonal views showing cartoon representation of IFIT1 molecule 1 (blue) and surface representation of IFIT1 molecule 2 colored according to electrostatic potential from -10 kTe^{-1} (red) to + 10 kTe^{-1} (blue). Residues at the interface are shown as sticks. The C-terminus of each molecule is indicated by asterisks for reference.
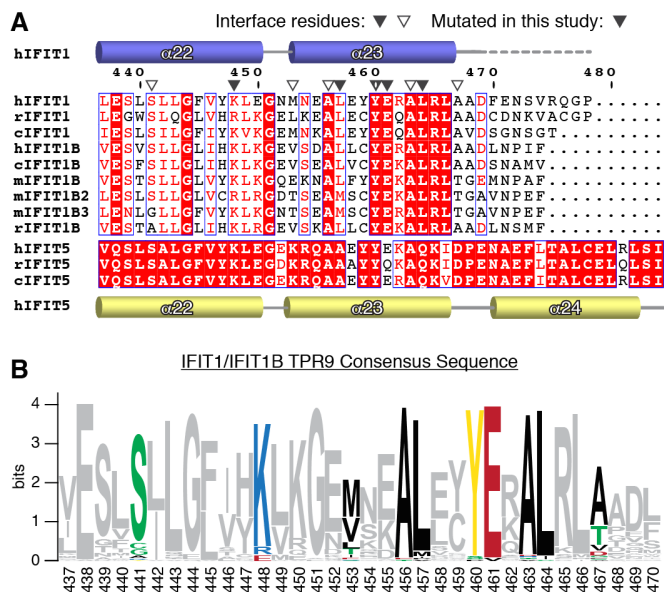
Figure 4. The IFIT1 dimerization interface is conserved in IFIT1 and IFIT1B proteins.
(A) TPR9 Sequence alignment of select IFIT1, IFIT1B and IFIT5 residues from human, mouse, cat, and rabbit. Secondary structure of IFIT1 and IFIT5 are depicted above and below. The ruler corresponds to human IFIT1 residue numbers. Compared to IFIT1 proteins, IFIT5 proteins are extended and have an additional α-helix (α24). **(B)** WebLogo sequence consensus of 86 IFIT1 and IFITB genes showing only TPR9. Interface residues are colored according to property, and other residues are in gray. The dimerization interface residues are highly conserved.
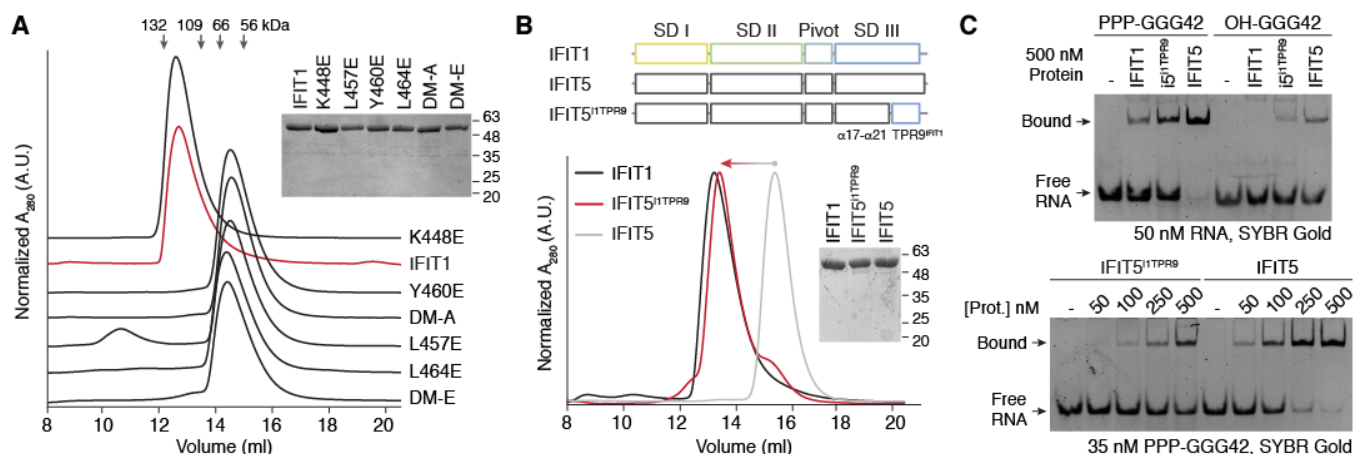
Figure 5. Mutational analysis of IFIT1 dimerization.
(A) Gel filtration (Superdex 200 Increase 10/300) analysis of IFIT1 mutants (at 2 mg/ml each) with SDS-PAGE of the input proteins (1 µg per lane). Each UV trace was offset along the y-axis for clarity. The peak maximum of IFIT1 wt was arbitrarily set to 1, and each trace shows the peak height relative to IFIT1. The bars above correspond to the migration of BSA dimer (132 kDa), IFIT2 dimer (109 kDa), BSA monomer (66 kDa), and IFIT5 monomer (56 kDa). **(B)** Schematic of the IFIT5-IFIT1 chimera (IFIT5^i1TPR9), and gel filtration (Superdex 200 Increase 10/300) analysis of the indicated proteins (at 2 mg/ml each) with SDS-PAGE of the input proteins (2 µg per lane). The chimera behaves as a dimer on gel filtration similar to IFIT1. **(C)** SYBR Gold stained electromobility gel-shift assays between a 42-nucleotide PPP-RNA (GGG42) and the indicated proteins. OH-RNA was utilized as a negative control for non-specific RNA binding in the upper panel.
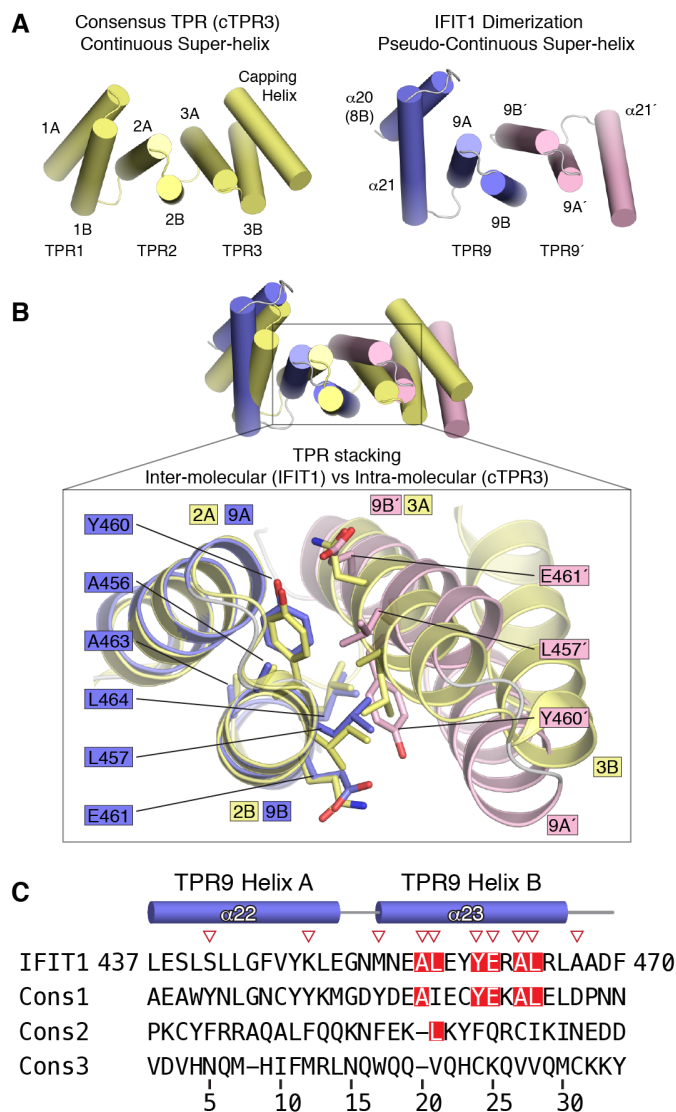
Figure 6. IFIT1 forms a pseudo-continuous TPR super-helix that mediates dimerization.
(A) Left, structure of cTPR3 (PDB ID: 1NA0) made up of 3 consensus repeats, and Right, the central region from the IFIT1 dimer colored according to the molecule each helix originates from. Together, the C-terminal helices from each IFIT1 molecule associate in a manner that resembles a continuous TPR array. **(B)** Superposition of cTPR3 over the IFIT1 dimerization helices and close-up of residues involved in IFIT1 inter-molecular TPR stacking vs cTPR3 intra-molecular TPR stacking. **(C)** Comparison of IFIT1 TPR9 to consensus TPR sequences, which represent the top 3 residue preferences for each position of the 34 amino acid TPR motif. The consensus sequences were adapted from ref. (4).
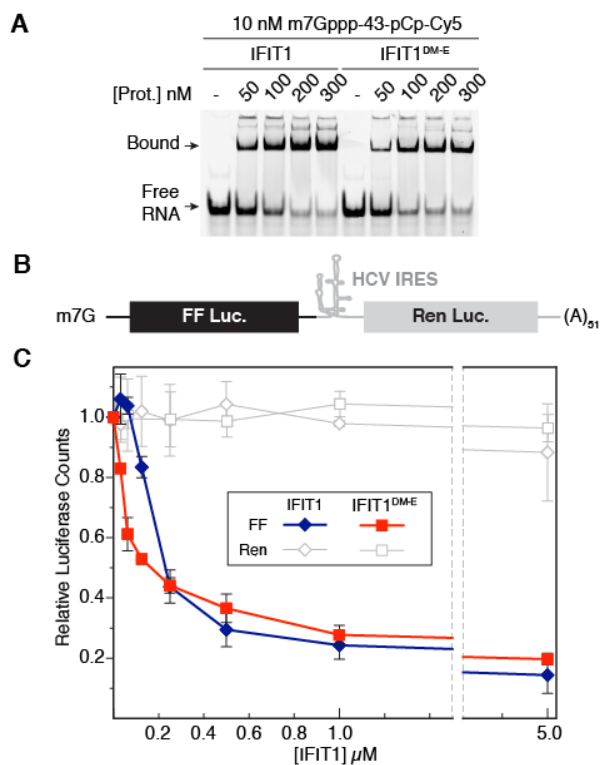
Figure 7. Functional analysis of IFIT1 dimerization.
(A) Gel shift between wt IFIT1 or monomeric IFIT1^{DM-E} and a 43-nucleotide, capped-RNA which is fluorescently labelled with pCp-Cy5 at the 3′ end. **(B)** Schematic of bicistronic mRNA reporter utilized in the translation assay. **(C)** Comparison of translational inhibition by IFIT1 or IFIT1^{DM-E}. Both proteins reduce cap-dependent FF translation equally without affecting cap-independent Ren translation. Data are represented as the mean of 2 independent measurements performed in duplicate ± standard deviation.
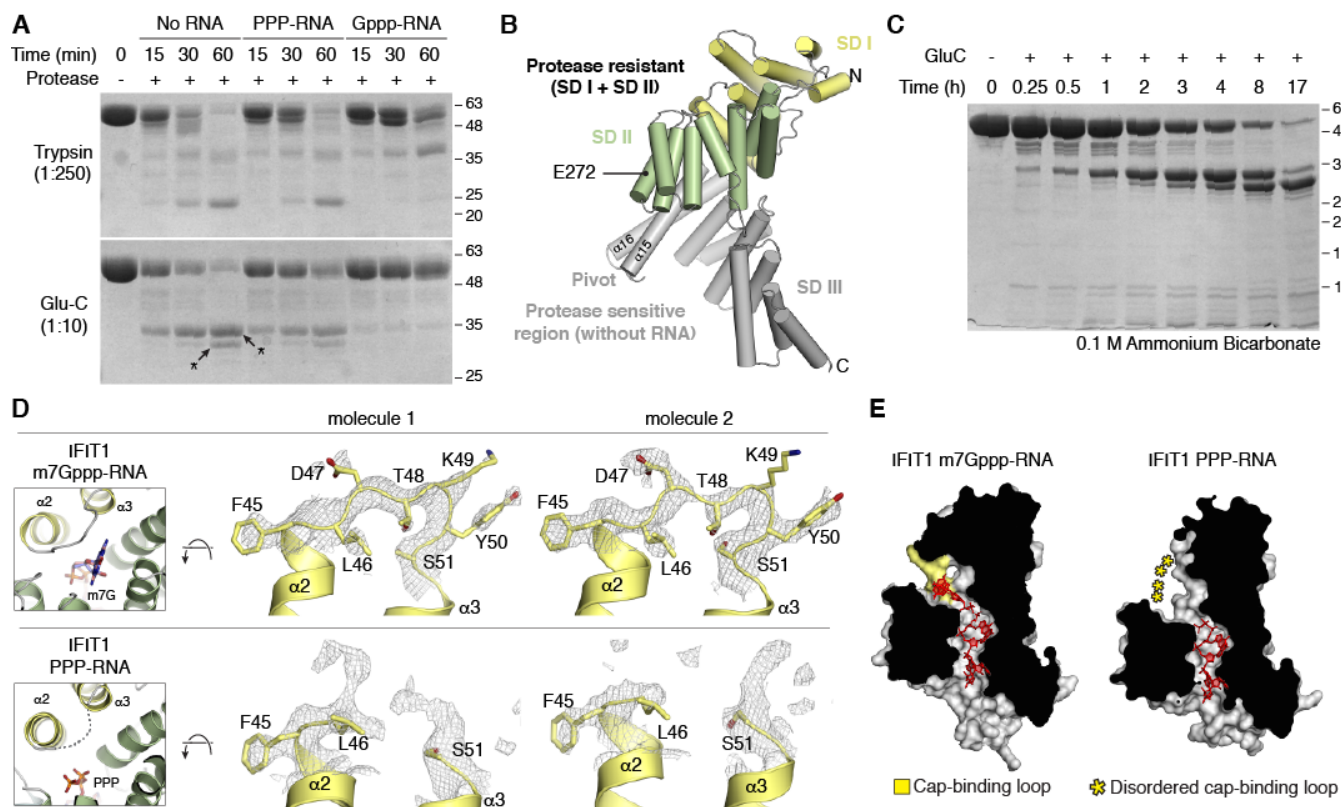
Figure 8. Limited protease digestion of IFIT1 and conformational changes associated with RNA binding.
(A) IFIT1, IFIT1 with PPP-oligoA, and IFIT1 with Gppp-oligoA were subjected to limited protease digestion for the indicated times. Top, 1:250 (w/w) IFIT1:trypisin and Bottom, 1:10 (w/w) IFIT1:Endoproteinase Glu-C. In this assay, Gppp-RNA was used instead of m7Gppp-RNA since it has been shown to bind similarly (6). The two indicated bands were excised and analyzed by in-gel tryptic digestion and LC/MS/MS to map the in-solution Glu-C cleavage sites. **(B)** Location of Glu 272 on subdomain II, where Glu-C cleaves IFIT1 to produce the 32 kDa bottom fragment. **(C)** Extended time course of Glu-C proteolysis in 0.1M Ammonium Bicarbonate pH 7.8 with 1:10 protease:protein. **(D)** Simulated annealing $2F_o$-F_c omit map of IFIT1 residues 45-51. The cap-binding loop is more ordered in the m7Gppp-RNA bound structure compared to the PPP-RNA bound structure. **(E)** Cross-section of the IFIT1 RNA-binding tunnel (grey protein surface with RNA depicted as red sticks). The cap-binding pocket is an open state in the absence of an m7G moiety.