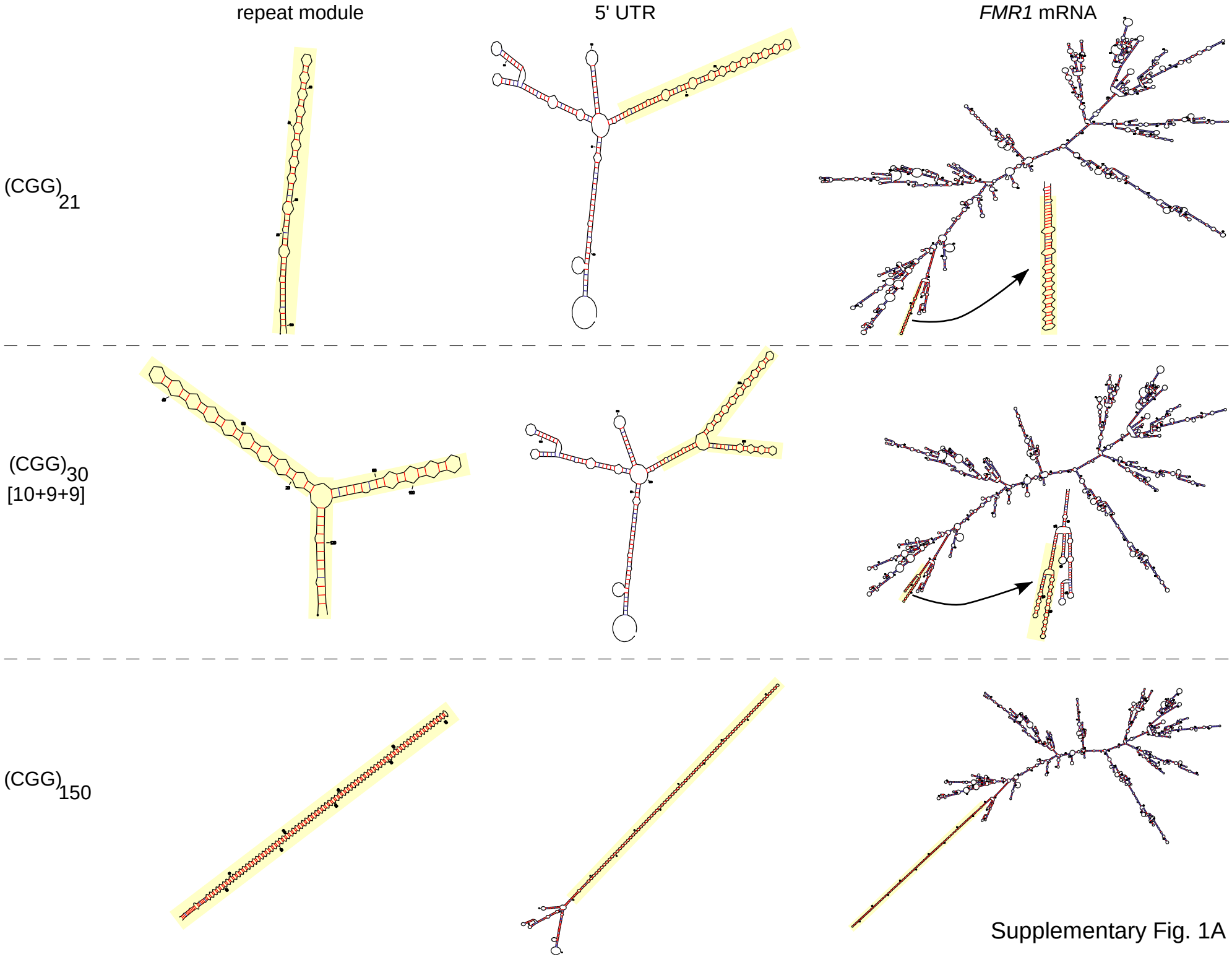
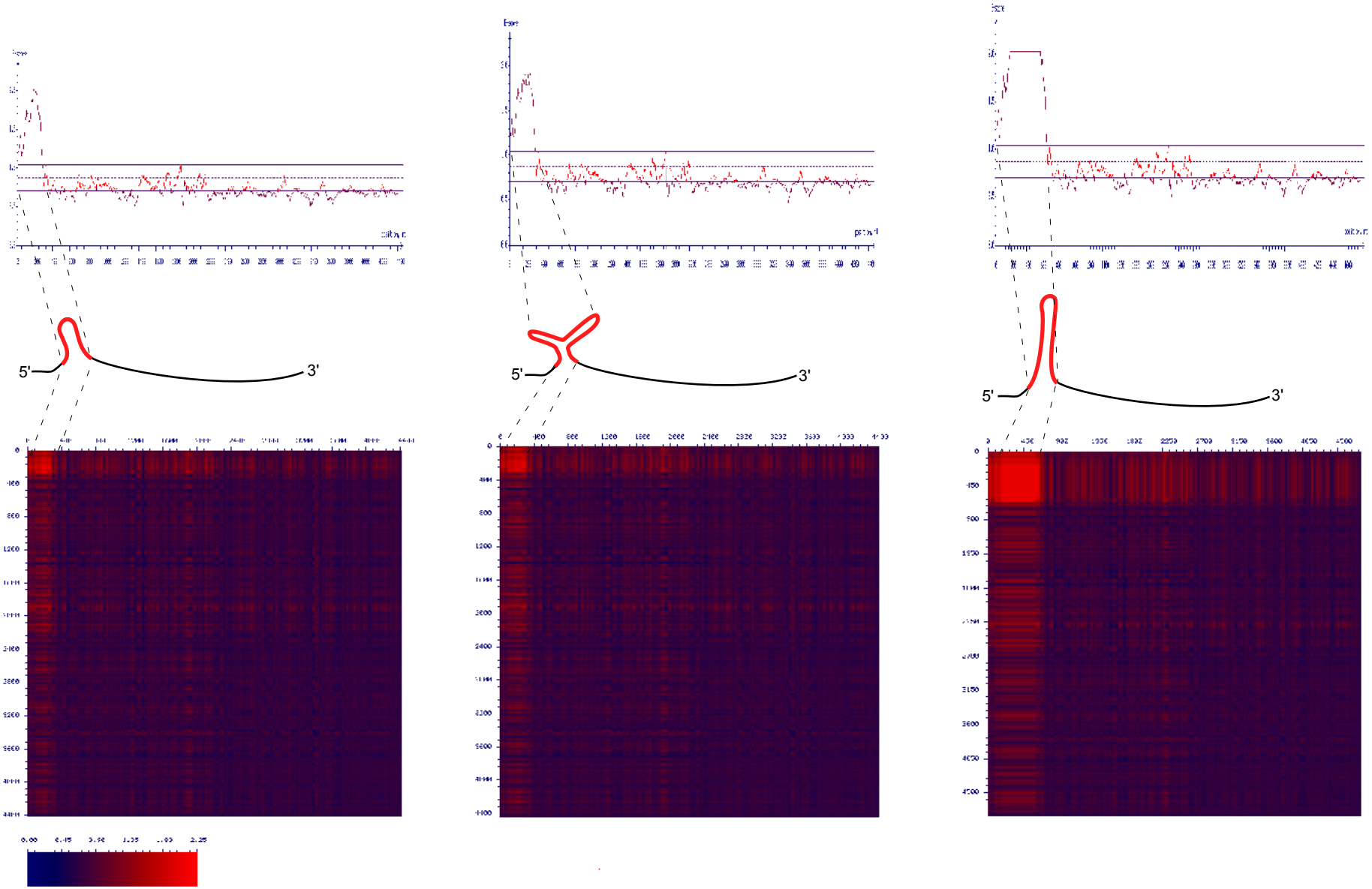


Supplementary Fig. 1A

$(CGG)_{19}$ *FMR1* mRNA

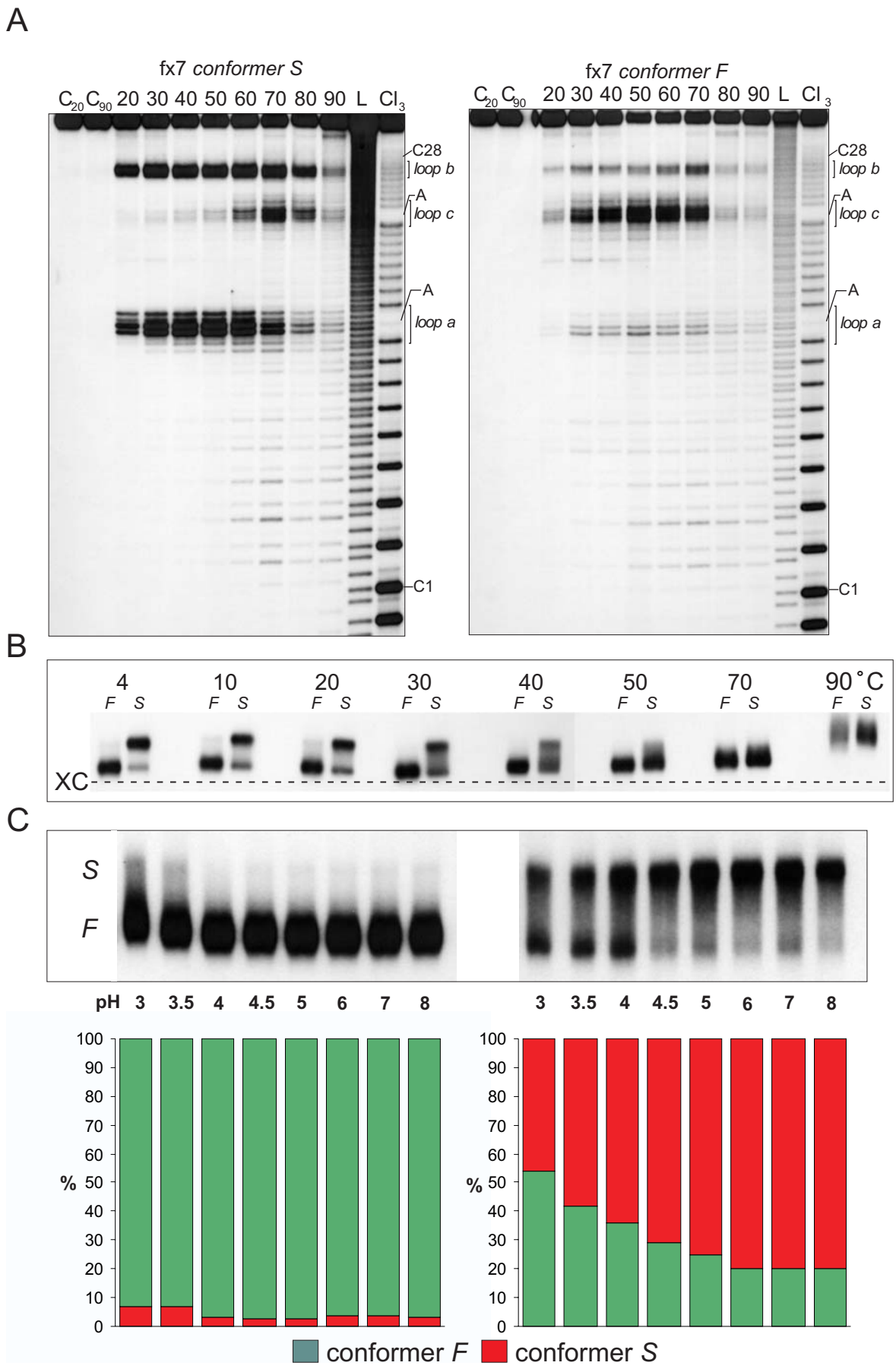
$(CGG)_{30}$ *FMR1* mRNA

$(CGG)_{150}$ *FMR1* mRNA



Supplementary Fig. 1B

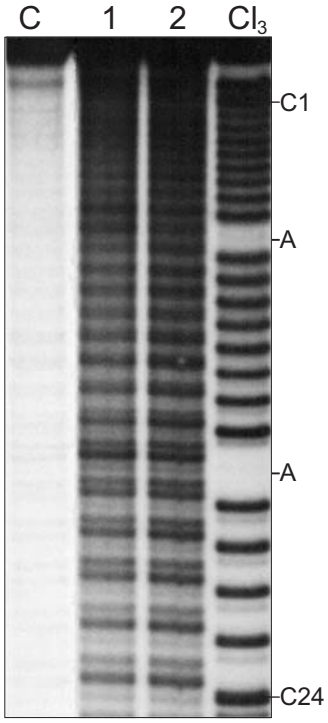
Supplementary Fig. 1. The computer simulation of the *FMR1* mRNA folding. (A) Results of the free energy minimization of the RNA molecules containing (i) 21 CGG repeats, (ii) 30 repeats with 2 AGG interruptions and 150 pure CGG repeats (iii). Analyses were performed in three different sequence contexts: RNAs corresponding to the transcript studied using enzymatic and chemical probes (RNA module), entire 5'UTR of the *FMR1* mRNA and complete mRNA molecule (NM_002024 modified using 5'UTR sequence data from (37)). Analyses were performed with the Mfold v. 3.1 software (http://www.bioinfo.rpi.edu/) at the default parameters. The optimal structures with the lowest calculated ΔG are shown. Yellow shading indicates the structural module subjected to the probing experiments. Notice, that irrespectively of the length of the repeat region and length of the RNA (including entire *FMR1* mRNA) the structural integrity of these modules is preserved. This strongly suggests that transcripts fx1 – fx9 are folded in the same manner regardless of the broader sequence context and therefore can be considered as the independent structural modules. (B) Analysis of the *FMR1* mRNAs using MatrixSS software (GArna MODULE package). The MatrixSS is used to search for the RNA regions harboring potentially significant secondary structure. High "E-score" indicated as a pick on the "E-score profile" (top) or red area on the "E-score matrix" (bottom). The dotted lines on the "E-score profile" indicate typical "E-score" for the structural RNAs (tRNA, rRNA). The "E-score" calculated for *FMR1* mRNA is much higher suggesting the formation of very stable secondary structures by fragment of *FMR1* 5'UTR. The region of the highest values corresponds to the CGG repeat tract. The results of the analyses conducted for the entire *FMR1* mRNAs harboring 21, 30 (with 2 AGG interruptions) and 150 repeats are presented. The mfold and MatrixSS predictions were performed on several RNAs containing different CGG length and AGG interruption status (including transcripts fx1 – fx9).



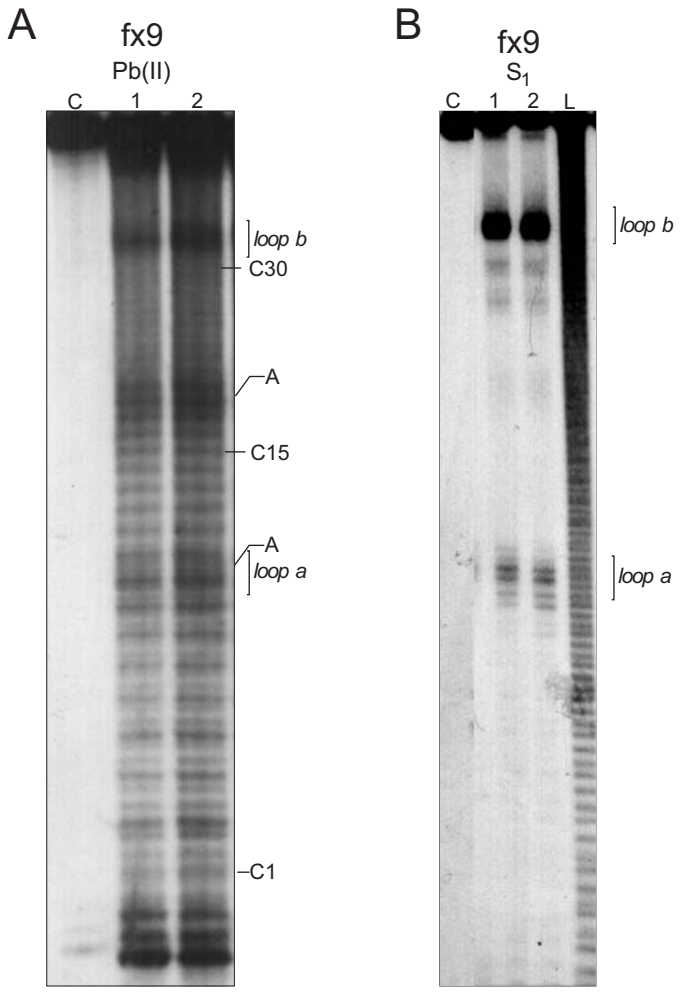
Supplementary Fig. 2

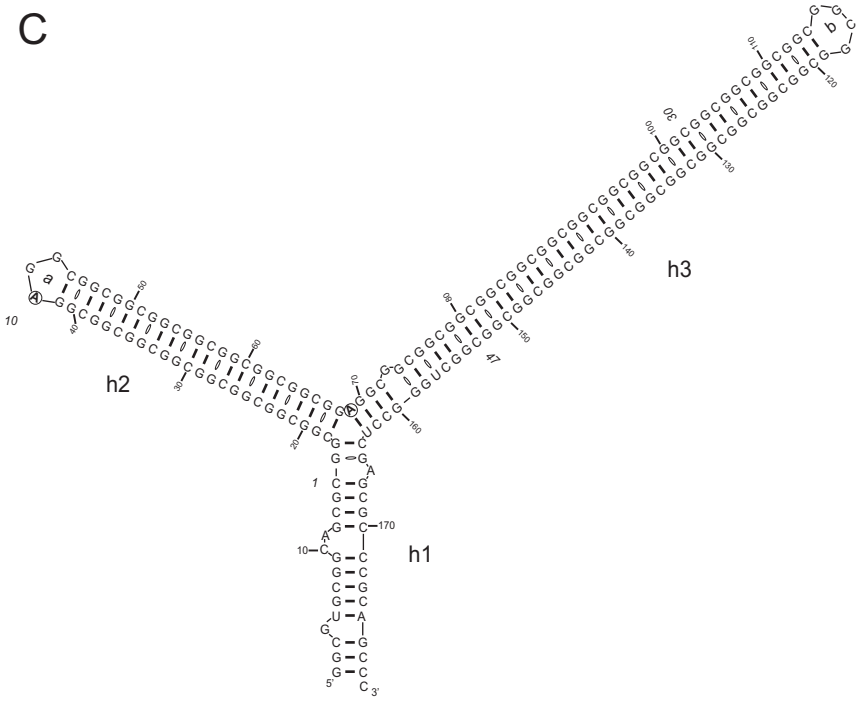
Supplementary Fig. 2. Effect of temperature and pH on transition between conformers F and S of the fx7 transcript. A, melting of the conformer S and conformer F structures analyzed by S_1 nuclease cleavages at different temperatures (shown above each lane). 1.35 units/μl of S_1 nuclease was used in each reaction. The incubation periods were 40 min., 30 min., 10 min., 3 min., 1 min., 30 s., 20 s., and 20 s., for reactions carried out between 20 – 90 °C, respectively. The reaction times were adjusted for each temperature to give the same ratio between the unreacted substrate and cleavage products. Lanes C_{20} and C_{90} – incubation controls (without a probe) performed at 20 and 90 °C, respectively. The first and the last cytosines of the CGG tract as well as both the adenine residues of the AGG interruptions are indicated. The regions corresponding to the loops *a*, *b* and *c* are marked. B, effect of the temperature on the transition between the conformers of the fx7 transcript. Results of 8 independent gel electrophoresis conducted in 10% native polyacrylamide gels at various temperatures are shown. The migration of the xylene cyanol (XC) is indicated. Prior to the electrophoresis the gels were equilibrated to the appropriate temperature for 2 h. RNA samples were incubated for 3 min. at the temperature corresponding to the temperature of the gel and immediately loaded into the wells. C, Influence of the pH on the transition between the conformers of the fx7 transcript. Isolated conformers (S and F) were incubated for 30 min. in 10 mM Tris-HCl buffer (for pH 6.5 - 8) and 10 mM sodium acetate (for pH 3 – 6) and immediately loaded on a 10% native polyacrylamide gel followed by electrophoresis at 20 °C. After autoradiography, bands corresponding to the S and F conformers were excised from the gel and quantitated using a scintillation counter.

At the range of 20 – 70 °C both S_1 cleavage pattern and gel migration of the fx7 F were very similar, suggesting no significant structural changes of this molecule (panel A and B). On the contrary, conformer S had a slight spontaneous rate of conversion into the F form even at 10 °C at pH 7.0. At 40 °C, along with strong cleavages of the loops *a* and *b*, weak cuts at the region of the second AGG interruption (3-way junction→loop *c*) were observed. At 70 °C the pattern of S_1 cleavages was almost identical for fx7 S and fx7 F with the central loop *c* region cut much stronger than loops *a* and *b* (at this temperature a 100% of the fx7 S was converted into the faster migrating conformer – panel B). Both forms were completely denatured at 90 °C (panel B). The pH had no influence on the stability of the F conformer, but low pH (<4.5) readily induced transformation from fx7 S to fx7 F (panel C) demonstrating that a high energy input was not required to achieve the S→F conformational change. These results strongly suggest that structural polymorphism of the fx7 transcript arose from the changes in tertiary interactions.



Supplementary Fig. 3. DMS modification of the fx7 F conformer. Modification patterns obtained for 3'-end labeled transcript; lanes 1 and 2 – reactions carried out at 20 °C and 90 °C, respectively, lane C – control reaction performed at 90 °C without DMS; Cl_3 – cytosine specific ladder. Positions of the AGG interruptions are indicated. Electrophoresis conducted in a 10% denaturing polyacrylamide gel.





Supplementary Fig. 4

Supplementary Fig. 4. Structure analysis of 5' UTR fragment of the *FMR 1* mRNA containing 47 CGG repeats with two AGG interruptions. Due to very small amount of transcript available for analyses we were unable separate individual conformers and therefore the mixture of conformers was subjected to the probing reactions. A, cleavage patterns obtained for 5' end labeled fx9 transcript treated with: Pb(II), lanes 1, 2 correspond to 0.5 and 1.0 mM lead acetate, respectively; lane C – incubation control. The 1^{st}, 15^{th} and 30^{th} cytosine of the CGG tract are indicated, as well as the adenines of the AGG interruptions. The regions of the RNA molecules predicted to form single-stranded loops *a* and *b* are shown. B, S_1 nuclease cleavage patterns, lanes 1 and 2 correspond to 0.675, and 2.7 units/μl of S_1, respectively; lane C – incubation control; L – formamide ladder. Electrophoresis was conducted in a 6% denaturing polyacrylamide gel. C, secondary structure model of the predominant conformer of fx9 transcript. Terminal loops of the helical regions h2 and h3 and are designated as *a* and *b*, respectively. Numbers indicating the trinucleotide repeats are italicized.