18–28 Nucleic Acids Research, 1995, Vol. 23, No. 1

Molecular dissection of the pseudoknot governing the translational regulation of *Escherichia coli* ribosomal protein S15

Claude Philippe, Lionel Bénard[1], C. Portier[1], E. Westhof, B. Ehresmann and C. Ehresmann*

UPR 9002 du CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg cedex, France and [1]Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received October 24, 1994; Revised and Accepted November 21, 1994

ABSTRACT

The ribosomal protein S15 controls its own translation by binding to a mRNA region overlapping the ribosome binding site. That region of the mRNA can fold in two mutually exclusive conformations that are in dynamic equilibrium: a structure with two hairpins and a pseudoknot. A mutational analysis provided evidence for the existence and requirement of the pseudoknot for translational control *in vivo* and S15 recognition *in vitro*. In this study, we used chemical probing to analyze the structural consequences of mutations and their effect on the stem–loop/pseudoknot equilibrium. Interactions between S15 and the pseudoknot structure were further investigated by footprinting experiments. These data, combined with computer modelling and the previously published data on S15 binding and *in vivo* control, provide important clues on pseudoknot formation and S15 recognition. An unexpected result is that the relevant control element, here the pseudoknot form, can exist in a variety of topologically equivalent structures recognizable and shapable by S15. S15 sits on the deep groove of the co-axial stack and makes contacts with both stems, shielding the bridging adenine. The only specific sequence determinants are found in the helix common to the pseudoknot and the hairpin structures.

INTRODUCTION

Several genetic systems from prokaryotic or eukaryotic cells have proved the involvement of a pseudoknot structure in translational control: autoregulation of the α-operon by the r-protein S4 in *E.coli* (1); autoregulation of the bacteriophage T4 gene 32 (2); stimulation of translation of *repZ* (3); ribosomal frameshifting in a vertebrate coronavirus (4). Strong evidence suggested that the autoregulation of *Escherichia coli* ribosomal protein S15 at the translational level depends on the mRNA capacity to adopt a pseudoknot structure (5–7). Indeed, the S15 regulatory region, about seventy nucleotides overlapping the ribosome loading site, can form two mutually exclusive conformations. The first one consists of two stem–loops (hairpins II and III) and the second one is a pseudoknot in which nucleotides from hairpin III are paired with nucleotides of the loop of hairpin II (Fig. 1). It was proposed that these two stuctures are in dynamic equilibrium and that the binding of S15 stabilizes the pseudoknot form. The binding of S15 was also shown to allow ribosome binding and to trap the ribosome on its loading site, preventing the formation of the active ternary 30S/mRNA/initiator tRNA complex (8). Consistently with the previous observation, 30S subunits bind more efficiently to the pseudoknot than to the stem–loop structure (9)

In order to gain more information on the regulatory mechanism, we analyzed the effects of a set of mutations both on the expression of a translational fusion between the S15 gene and the reporter gene *lacZ* as well as on the capacity of the mRNA to bind S15 *in vitro* (10). These results provided convincing genetic evidence for the existence of the pseudoknot *in vivo* and for its requirement for translational control and S15 recognition. One result that emerged from that study was the genetic demonstration of base pairing between A(–47) and U(–38). As a consequence, only one residue [A(–46)] connects stems 1 and 2, and crosses the deep groove in the pseudoknot structure. This study also identified the U(–49)–G(–36) pair as a sequence specific determinant, since its replacement by a canonical U–A pair induces a loss of control. Moreover, the results gave a good correlation between the *in vivo* control and the *in vitro* binding affinity for S15. Furthermore, the mutational analysis suggested that a few specific determinants for S15 recognition are provided by the pseudoknot structure. However, the unambiguous interpretation of genetic data requires the knowledge of the structural consequences of the mutations. Here, this is particularly important due to the complexity of the system which is based on a dynamic equilibrium. Indeed, the mutations are expected to displace the equilibrium, either by stabilizing or destabilizing one of the two conformations in equilibrium (or even by creating another conformation). This is, for example, the case of mutant CFP5517 for which we detect only the pseudoknot form as a

* To whom correspondence should be addressed

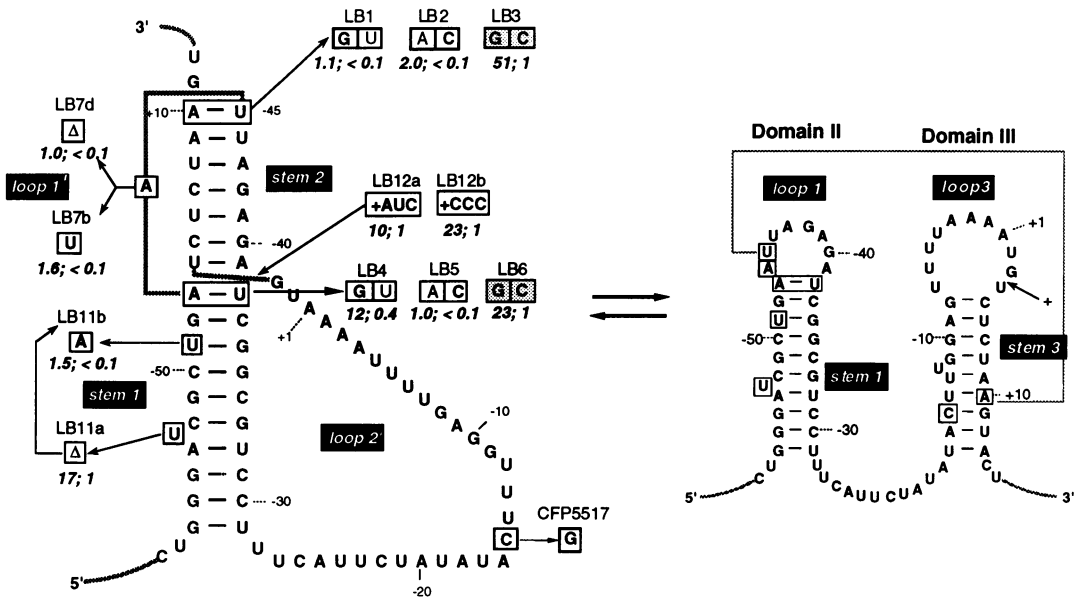

Figure 1. Mutations are indicated on the two conformations which are in equilibrium in the wild-type mRNA. The nomenclature of the different stems and loops is indicated. The studied mutations are shown: boxes indicate nucleotide changes (open boxes for single changes and shaded boxes for double changes. (+) and (Δ) refer to insertion and deletion. The name of the mutant is indicated above the mutation. The effects of the mutations are summerized: the first number refers to the *in vivo* repression rate and the second to the relative binding affinity measured *in vitro*. In the case of the wild-type RNA, these values are 23 and 1, respectively. Data are from Portier *et al.* (1994). In the case of mutant CFP5517, the repression rate was determined in another construction (Portier *et al.* 1990a). It was similar to that of the wild-type RNA measured in the same construction. The binding strength was also found to be similar to that of the wild-type RNA (9).

consequence of a C to G substitution at position −15 because it destabilizes the stem of hairpin III (7,9).

Here, we report a detailed structural analysis of 12 of the mutants mentioned above, together with wild-type RNA and mutant CFP5517 as references. The analysis was based on chemical and enzymatic probing. Further insight into the interactions between S15 and the pseudoknot form was also provided by footprinting experiments. A three dimensional model of the pseudoknot is derived from experimental data. These data are used to dissect the effect of mutations on the translation control and to provide important clues on the structural rules governing the formation of the pseudoknot and S15 recognition.

MATERIALS AND METHODS

Chemicals and enzymes

1-cyclohexyl-3(2-(1methylmorpholino)-ethyl)-carbodiimide (CMCT) and H_2O_2 were from Merck, dimethylsulfate (DMS) from Aldrich Chemicals Company, (β-etoxy-α-ketobutyraldehyde and kethoxal from US Biochemical Co., ethylnitrosourea (ENU) and sodium ascorbate from Sigma. Acrylamide and N, N′ methylene bis-acrylamide were from BDH Chemicals. [γ-^{32}P]ATP was from Amersham. Restriction enzymes were purchased either from Boehringer, Bethesda Research Laboratory or New England Biolabs. T7 RNA polymerase was purified from the overproducing strain BL21/pAR1219, following the purification protocol provided by F.W.Studier. T4 polynucleotide kinase, RNase V1 and RNase T1 were from P.L. Biochemicals. Reverse transcriptase was from Avian myeloblastosis virus (AMV).

Preparation of the biological material

All plasmids were described in (10) and RNAs were prepared by T7 RNA polymerase transcription as described by Philippe *et al.* (1994). The RNA fragments were renatured prior to use by incubation at 42°C for 20 min in the appropriate buffer and were then cooled on ice. Protein S15, kindly provided by C.Cachia, was purified under non denaturing conditions (11).

Probing and footprinting

In a standard assay (containing 12.5 pmoles of mRNA and 5 μg carrier tRNA), RNase digestion or chemical modifications were conducted at 37°C in 50 μl buffer A (20 mM Tris acetate (pH 7.5), 10 mM Mg acetate, 60 mM NH_4Cl, 100 mM KCl, 3 mM (β-mercaptoethanol) unless otherwise stated. Digestion with RNase V1 (0.025 and 0.05 U) was for 5 min. Modification with DMS: 0.5 μl of DMS from 3 to 10 min. Modification with kethoxal: 1, 3 and 5 μl of a 20 mg/ml kethoxal-solution for 10 min at 20°C. Modification with CMCT: 10 μl of CMCT (42 mg/ml) in 50 μl buffer N2 (50 mM Na borate pH 8.0, 20 mM Mg acetate, 300 mM KCl) for 40 and 60 min. Reactions were stopped by ethanol precipitation at −20°C (after addition of borate to 25 mM, for kethoxal). Modification with ENU: 5 μl of a ENU-saturated ethanol solution in 20 μl of buffer N3 (150 mM Na cacodylate pH 7.5, 20 mM Mg acetate, 300 mM KCl) or buffer D3 (150 mM Na cacodylate pH 7.5, 1 mM EDTA, 300 mM KCl) for 30 and 60 min. Cleavage of the resulting phosphotriester bonds was performed in 15 μl of 0.1 M Tris–HCl, pH 9.0. Hydroxyl radical cleavage was initiated by a freshly prepared solution of iron II (0.006 mM) and EDTA (0.012 mM), sodium ascorbate (3 mM) and H_2O_2 (0.09%) as described in (12). Reactions were conducted in a final volume of 70 μl from 45–75 s and stopped

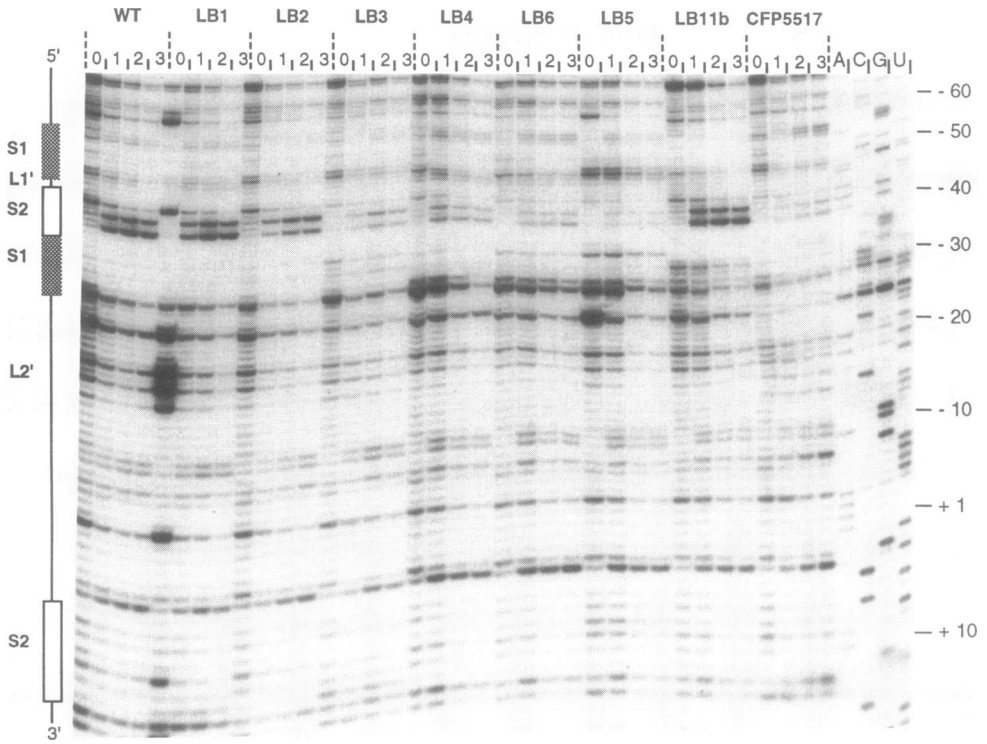

Figure 2. Autoradiogram resulting from kethoxal modification on different mutants. The reaction was performed at 20°C. Lane 0, incubation control; lane 1, 1 µl; lane 2, 3 µl; lane 3, 5 µl of kethoxal. The stem and the loop regions of the pseudoknot are indicated on the left of the autoradiogram. C, U, A and G are sequencing lanes.

by addition of thiourea to a final concentration of 10 mM and by ethanol precipitation. In footprinting experiments, the complex was formed in the presence of 50 pmoles S15 in buffer A (iron–EDTA footprint) or N3 (ENU footprint) containing 0.002% bovine serum albumin. Modified sites or cuts were detected by extension with AMV reverse transcriptase of a primer complementary to nucleotides +38 to +51. Elongation controls were run in parallel in order to detect spontaneous hydrolysis in the RNA template or pauses of reverse transcription.

Computer modelling

The modeled molecule integrating stereochemical constraints and experimental data was constructed with the help of several computer programs (13) and tested by comparing the theoretical accessibility of atoms with the observed experimental reactivity, as described earlier (14).

RESULTS

Enzymatic and chemical probing of the mutants

In order to analyze the stuctural consequence of the mutations shown in Figure 1, the conformation of the various RNA mutants was investigated by probing experiments. The four bases were tested at one of their Watson–Crick positions with DMS, at A(N1) and C(N3), with CMCT, at U(N3), with kethoxal at G(N1, N2) and with RNase V1 which cleaved helical regions. Typical experiments are shown in Figures 2, 3 and 4. Experiments were repeated at least twice and the degree of reactivity was evaluated from 1 to 4 by visual inspection. In the following, the stem and loop elements will be referred to the nomenclature shown in Figure 1.

The reactivity changes induced by mutations are essentially localized in the crucial regions that correspond to nucleotides involved in the pseudoknot formation (in particular nucleotides –45 to –39 in loop 1 which pair with nucleotides +4 to +10 in stem 3 to form the pseudoknot). Therefore, the degree of reactivity of these nucleotides is a measure of the stem–loop/pseudoknot equilibrium with a non-reactivity indicating that the equilibrium is displaced towards the pseudoknot structure (see mutant CFP 5517). Probing data on nucleotides involved in the pseudoknot formation are summarized in Figure 5. The wild-type RNA and mutant CFP 5517 are shown as references for the stem–loop/pseudoknot equilibrium and for the pseudoknot form, respectively. Note that U(–38) is unreactive in the wild-type RNA and in all mutants, while A(–47) is reactive in most cases (with the exception of LB7b and LB7d). Otherwise, mutants can be divided into three classes on the basis of their reactivity profile.

(i) Class 1 mutants. The first class corresponds to mutants that display a pattern of reactivity similar to that of the wild type RNA. They are characterized by an overall accessibility of all nucleotides in loop 1. This is the case of mutants LB1, LB2, LB4, LB11a, LB11b and LB12a. These mutants probably still possess the stem–loop/pseudoknot equilibrium. However, at this stage it cannot be excluded that the equilibrium is displaced towards the stem–loop structure. This is particularly the case for mutants LB1

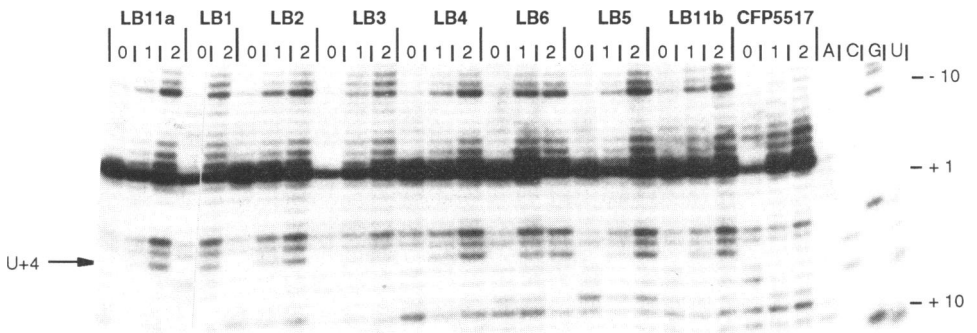

Figure 3. Autoradiogram resulting from CMCT probing experiments on different mutants. The reaction was performed at 37°C. Lane 0, incubation control; lane 1, 40 min; lane 2, 60 min. Portion of the gel showing the reactivity of nucleotides –14 to +10 corresponding to one strand of stem 2 and a portion of loop 2′ in the pseudoknot structure. C, U, A and G are sequencing lanes.

and LB2 in which the disruption of the terminal U(–45)–A(+10) pair is expected to destabilize the pseudoknot.

(ii) Class 2 mutants. Mutants from the second class (LB7b and LB7d) clearly differ from class 1 mutants by the non-reactivity of A(–47) and of the first nucleotides of loop 1, while nucleotides AGA(–41) remain highly reactive (level 3/4) (Fig. 5). This is an indication for the extension of stem 1 by extra base pairs. Indeed, the substitution of A(–46) by U in mutant LB7b favors the existence of the pairs A(–47)–U(–38), U(–46)–A(–39) and possibly U(–45)–G(–40), leading to a tetraloop UAGA (Fig. 6). In mutant LB7d [with A(–46) deleted] the reactivity pattern is in favor of pairs A(–47)–U(–38), U(–45)–A(–39) and possibly U(–44)–G(–40). Therefore, these two mutations, by decreasing the number of nucleotides in loop 1, are detrimental to the formation of the pseudoknot.

(iii) Class 3 mutants. Mutants from the third class [LB3, LB5, LB6 and LB12b] display a strong decrease of reactivity of nucleotides (–45) to (–39) in loop 1, as in mutant CFP5517 (Fig. 5). This general decrease of reactivity can be interpreted as resulting from the involvement of these nucleotides in the pseudoknot structure. Remarkably, A(–47) remains reactive at different degrees in LB3, LB5 and LB12b (as in CFP5517), whereas the mutated G(–47) is unreactive in LB6 (Fig. 2). Otherwise, the reactivity of A(–46) differs also from one mutant to another one.

Mutant CFP5517. In this mutant, A(–47) and A(–46) are both reactive [with A(–47) > A(–46)]. On the other hand, G(+3) is reactive to kethoxal (level 3) and cleaved by RNase T1 (Philippe *et al.*, 1993). These observations favor a pseudoknot with two As crossing the deep groove of stem 2 and G(+3) unpaired (Fig. 7). As suggested by its non-reactivity, U(–38) is most likely stacked between helices 1 and 2.

Mutant LB3. The double substitution C(–45)/G(+10) inserts a C–G pair instead of a U–A pair and increases the stability of stem 2. A(–47) and A(–46) are equally reactive (Fig. 4), but A(–47) is less reactive than in CFP5517. On the other hand, G(+3) is still reactive, but less than in CFP5517 (Fig. 2). These results also favor a pseudoknot with two As crossing the deep groove, as in CFP5517 (Fig. 7). However, the weaker reactivity of A(–47) and G(+3) suggests subtle differences at the junction of the two helices, e. g. possible labile interactions between U(–38) with either A(–47) or G(+3).

Mutant LB6. The two mutated bases, G(–47) and C(–38), are paired, in line with their non-reactivity. However, several features clearly differ in LB3 and CFP5517 (Fig. 7). (i) U(+4) and A(–39)

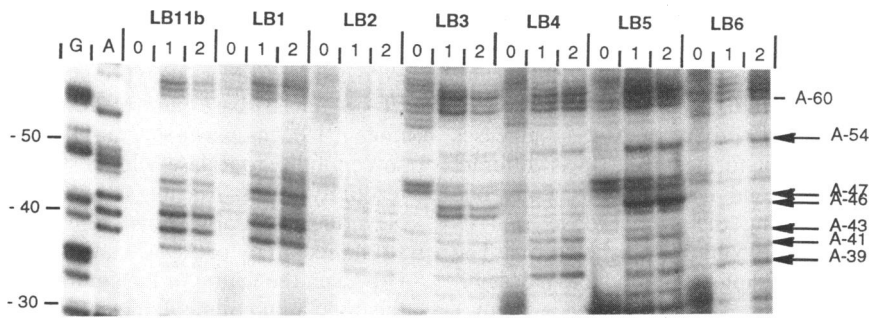

Figure 4. Autoradiogram resulting from DMS probing experiments on different mutants. The reaction was performed at 37°C. Lane 0, incubation control; lane 1, 3 min; lane 2, 6 min. Portion of the gel showing the reactivity of nucleotides –63 to –30. Note that the lanes corresponding to the different mutants do not contain the same amount of radioactive material, so that the degree of reactivity has to be compared to A(–60), which displays an unchanged reactivity. C, U, A and G are sequencing lanes.

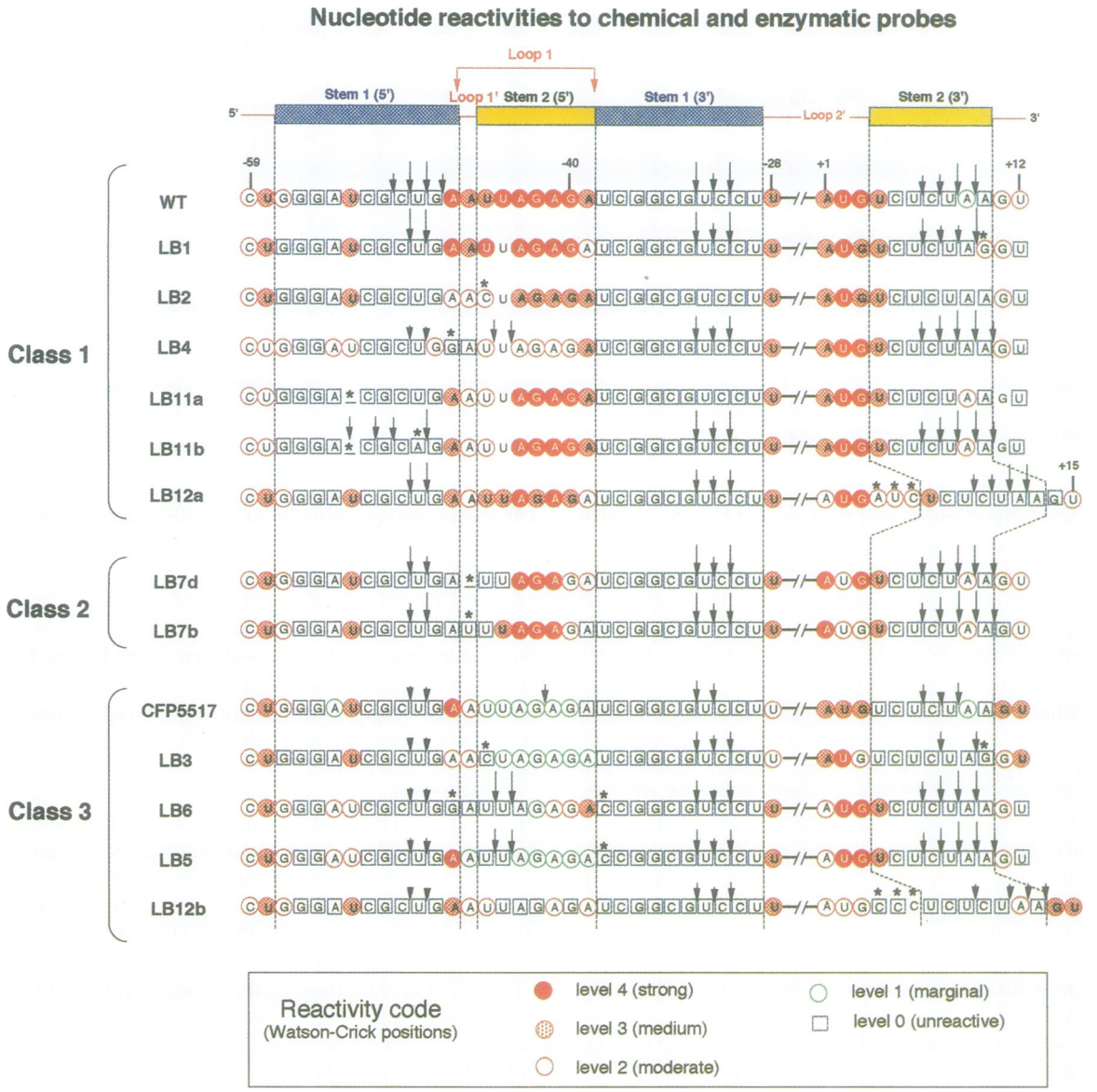

Figure 5. Summary of the reactivity of nucleotides to chemical and enzymatic probes on each mutant. The study is restrained to the sequences involved in the pseudoknot –59 to 28 and +1 to +12 as indicated in Figure 1. The position of the mutations is pointed by an asterisk above the nucleotide. The A of the AUG is numbered +1. Reactivity of Watson–Crick positions are obtained using DMS (A, C), CMCT (U) and kethoxal (G). The reactivity color code is indicated. The RNase V1 cuts are indicated by arrows; the size of the arrow is proportional to the intensity of the cut. The mRNAs are subdivided into three classes, according to their reactivity pattern (see text).

are both reactive (level 3). (ii) Two RNAse V1 cuts appear in stem 2 at positions –43 and –44. (iii) A(–46) is unreactive. (iv) U(–53) becomes less reactive while the reactivity of A(–54) increases (Fig. 5). Therefore the peudoknot formed by LB6 is different from that formed by CFP5517 and LB3 (Fig. 6). Only one adenine [A(–46)] is now crossing the deep groove, and its non-reactivity suggests that it is probably inside the groove. The additional G(–47)–C(–38) pair closing stem 1 induces a destabilisation of the U(+4)–A(–39) pair. As a consequence, stem 1 contains an extra base pair, while stem 2 is shortened by one base pair. The two stems are separated by the unpaired A(–39).

Mutant LB5. The case of this mutant, in which U(–38) is substituted by C, is rather puzzling. The U to C change was expected to favor the pseudoknot form by promoting a C(–38)–G(+3) base pair, by reference to CFP5517. However, mutant LB5 displays several characteristics of LB6: the same susceptibility to RNase V1, a distal effect on A(–54), a marginal reactivity of A(–46), the same reactivity of G(+3) and U(+4). Nevertheless, it also shows some particularities of CFP5517: a strong reactivity of A(–47) and a marginal reactivity of A(–39). However, the nature of the equilibrium and the conformation of this pseudoknot remain unclear.

Mutant LB12b. This mutant, in which a CCC triplet is inserted after the first codon, appears to be similar to CFP5517 and LB3. However, the precise reason for this unexpected displacement of the equilibrium toward the pseudoknot structure is unclear (by

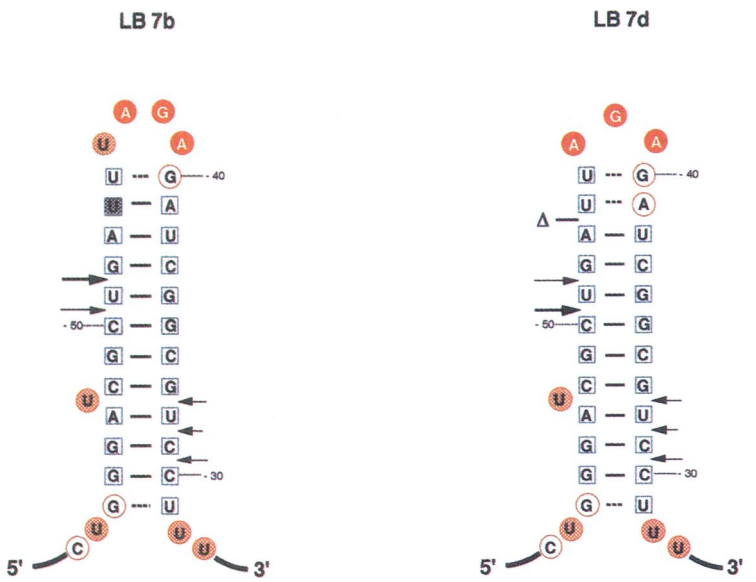

Figure 6. Proposed secondary structure of domain II for mutants LB7b and LB7d. Chemical reactivity at Watson–Crick positions and RNase V1 susceptibility are reported on the secondary structure, with the same color code as in Figure 5.

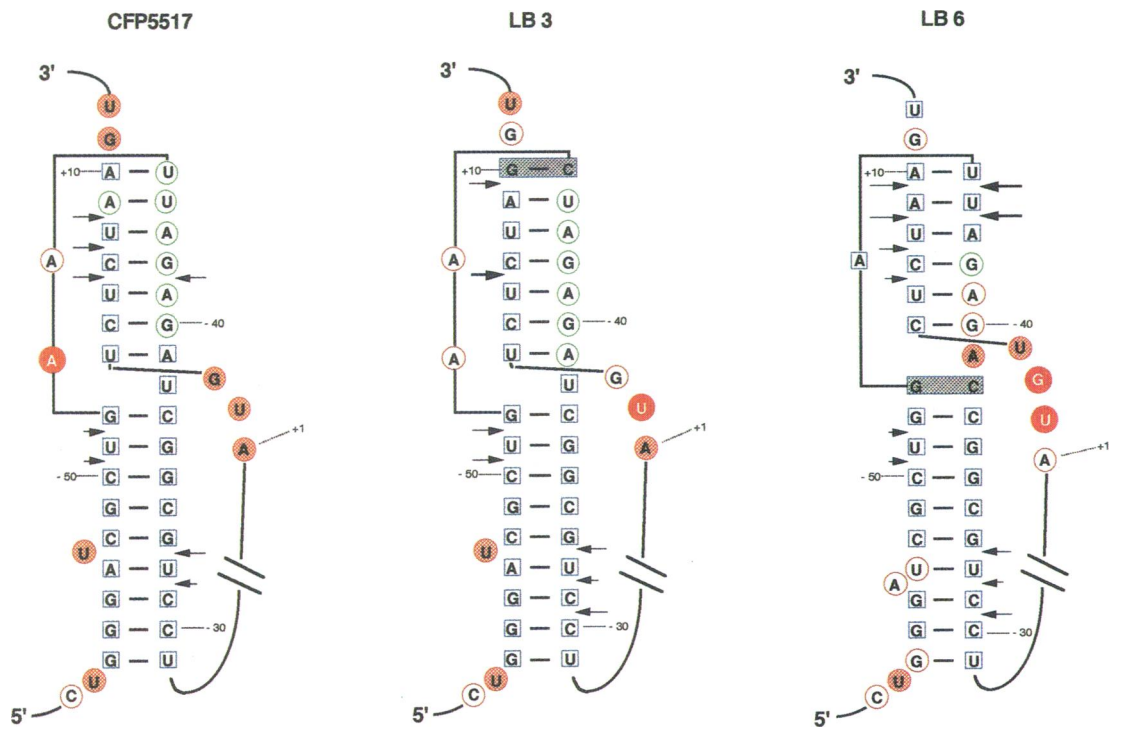

Figure 7. Proposed pseudoknot structure for mutants CFP5517, LB3 and LB6. Chemical reactivity at Watson–Crick positions and RNase V1 susceptibility are reported on the secondary structure, with the same color code as in Figure 5.

destabilizing hairpin III or by stabilizing the pseudoknot?). Since the insertion of the AUC triplet (mutant LB12a) does not induces the same effect, the nature of the first two inserted bases does interfere with this phenomenom. A noticeable particularity of this mutant is the presence of a strong stop of reverse transcriptase at the inserted C residues. Since an end-labelled RNA analysis clearly indicates that there is no cleavage at this position, the pause of reverse transcriptase most likely results from a particularly stable conformational feature. Moreover, the latter analysis indicates that the Cs are unreactive to DMS (not shown). Therefore, these results fit with the existence of interactions involving the additional Cs and residues elsewhere (i.e. in the minor groove of stem 1). These interactions may be responsible for the stabilization of the pseudoknot.

Footprinting experiments

In order to gain further insight into the elements of the pseudoknot that are recognized by S15, we used Fe(II)–EDTA and ENU to identify which portions of the sugar–phosphate backbone are protected by the protein. In contrary to other chemical probes that modify bases, both probes display the advantage of being insensitive to the secondary structure. Fe(II)–EDTA generates free hydroxyl radicals which are able to attack position C4′ of the sugar moiety (15) and ENU is an alkylating reagent of phosphate groups (16). Fe(II)–EDTA footprinting experiments were performed on three RNA mutants which adopt the pseudoknot form and bind S15 with the same affinity as the wild-type mRNA (CFP5517, LB3 and LB6) while ENU footprint was conducted on CFP5517 RNA. An example of gel is shown in Figures 8 and 9 and the results are summarized in Figure 10 on the secondary structure of these mutants. The effect of magnesium on the reactivity of RNA to Fe(II)–EDTA and ENU was also investigated. The presence of magnesium does not induce significant effect on the sugar moiety reactivity (Fig. 8). However, the reactivity of several phosphate groups to ENU alkylation is dependent on the presence of magnesium. Thus, magnesium induces a reduction of reactivity at phosphates –33, –34, –36 and –51 in stem 1, +5, (+9) in stem 2, (Fig. 9). Strikingly, magnesium also increases the reactivity of phosphate –50 in stem 1 and –7 and –6 in loop 2′. Since ENU is not sensitive to the secondary structure, the observed reactivity changes more likely reflect specific interactions with magnesium or the stabilization of tertiary interactions involving phosphate groups.

S15 induces common protection from Fe(II)–EDTA attack on three distinct parts of the pseudoknot: in loop 1′ (position –46), in stem 1 (positions –49 to –47) and stem 2 (positions –45 to –42) in all three RNAs. In addition, specific protections are observed in the various mutants. These variations in the protection pattern probably refect subtle differences in the fine structure of the various pseudoknots. For instance, both LB3 and CFP5517 displays protections in loop 2′ (near position +2), while LB3 and LB6 show additional protections on the opposite strand of stem 2 (near +10) and on stem 1 (near –34). As expected, all positions protected from hydroxyl radical hydrolysis on CFP5517 RNA are also protected from ENU alkylation at their corresponding 5′ phosphate group (Fig. 9). However, additional protections from ENU are observed in stem 1 (positions –36 to –31). The fact that these protections are specifically observed on the phosphate groups but not on the sugar moieties, most likely reflect tertiary interactions (i.e. between nucleotides of loop 2′ and the phosphate backbone of stem 1), rather than S15-induced protections. It should also be noted that the binding of S15 induces an enhanced accessibility of the Shine–Dalgarno sequence to RNase T1 hydrolysis in wild-type RNA as well as in mutants CFP5517, LB3, LB6 and LB12a [(8), and results not shown].

Three-dimensional model of the pseudoknot

Throughout this work, we used computer modelling to build three-dimensional models of the various structures suggested by the experiments. Here we present the two main types of pseudoknots (Fig. 11a and b). The first one contains two residues, A(–47) and A(–46), crossing the deep groove, and corresponds to the free form of mutants CFP5517 and LB3 as deduced from probing experiments (Fig. 11a). Stems 1 and 2 were constructed

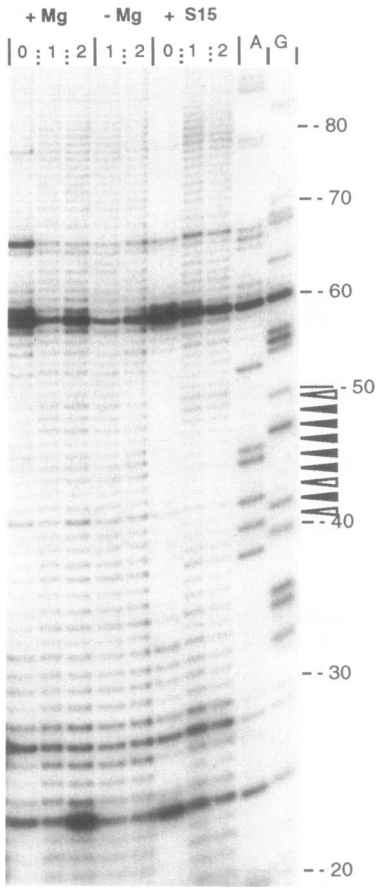

Figure 8. Autoradiogram showing the hydroxyl radical hydrolysis on mutant CFP5517, in the presence or in the absence of magnesium and S15. Incubation with Fe–EDTA: lane 0, incubation control; lane 1, 45 s; lane 2, 60 s. A and G correspond to sequencing lanes. The main S15-induced protections are indicated by arrows: full arrow for strong protection; open arrow for moderate protection.

as regular A-form helices and were assumed to be co-axially stacked. Residues U(–45) to U(–29) form a continuous helix, with U(–38) stacked in the helix. The second model is derived from the first one by pairing A(–47)–U(–38) (Fig. 11b). This pseudoknot form, which contains a single residue, A(–47), crossing the deep groove corresponds to the genetically defined pseudoknot. Stems 1 and 2 do not appear to form a completely contiguous A-like helix and residues A(–39) and U(–38) at the junction between helices 1 and 2 are not perfectly stacked. The absence of strong reactivity of A(–46) in all pseudoknots (note its complete absence of reactivity in LB6) suggests that this residue is probably buried in the deep groove and may form additional stabilizing interactions. Loop 2′, which contains 31 nucleotides, is not constrained in length for crossing the minor groove of stem 1. Although the orientation and the precise structure of loop 2′ is unknown, it can be seen from the model that interactions between the proximal 3′ nucleotides and the minor groove of stem 1 are quite conceivable. Such interactions (i.e. base–base or base–backbone interactions) probably account for the various observations reported above (unexpected behaviour of mutant LB12b and phosphate reactivity to ENU) and for an additional stabilization of the pseudoknot.

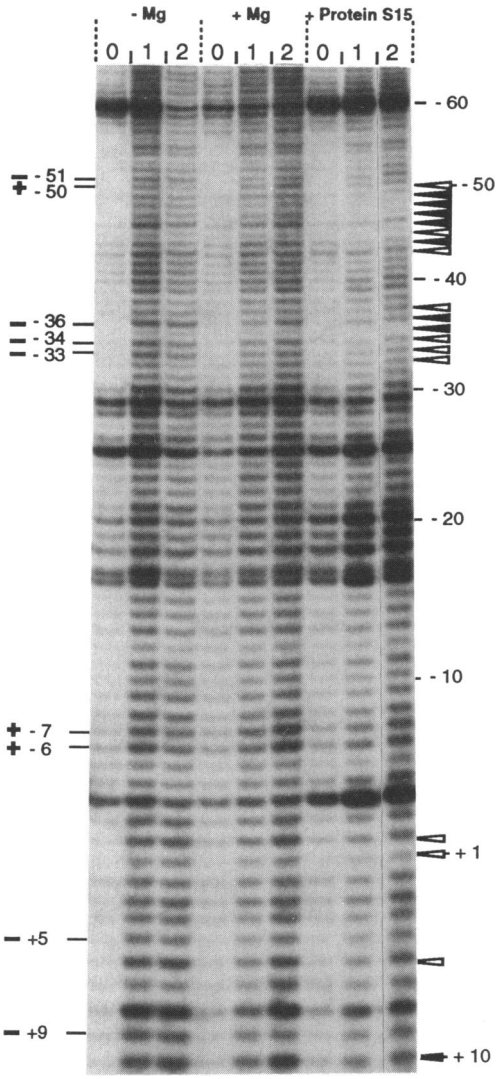

Figure 9. Autoradiogram showing the phosphates protected from ENU alkylation by magnesium and protein on mutant CFP5516. Incubation with ENU: lane 0, incubation control; lane 1, 30 min; lane 2, 60 min. The main protein S15-induced protections are indicated on the right of the figure with the same symbols as in Figure 8. Magnesium effects are indicated on the left of the figure with symbols (+) for an increase of reactivity or (–) for a decrease of reactivity.

The sugar residues protected from the hydroxyl-induced cleavage by S15 in the three pseudoknots recognized by S15 (CFP5517, LB3 and LB6) are indicated by spheres on the three dimensional model (Fig. 11b). Remarkably, on this particular region, which displays a sharp turn of the sugar–phosphate backbone, the protections are clustered on the same side of the pseudoknot.

DISCUSSION

What is governing the pseudoknot/stem–loop equilibrium?

As expected, one effect of the mutations is to alter the pseudoknot/stem–loop equilibrium. According to our results, the mutants can be divided into three classes. Mutants from class 1 probably still form a pseudoknot in equilibrium with a stem–loop structure, with the possible exception of LB1 and LB2. In mutants from class 2, the equilibrium is clearly shifted towards the stem–loop conformation. The pseudoknot does not form, essentially because the mutations induce an extension of stem 1. Thus, the size of loop 1 is reduced to four bases in LB7b and probably to three bases in LB7d. Note that hairpin loops of four nucleotides are known to be thermodynamically more stable than larger loops (17).

In the third class of mutants, the equilibrium is displaced towards the pseudoknot. However, the structures of these pseudoknots show striking differences, especially in the size of loop 1′ and at the junction of helices 1 and 2. Mutant LB3, where stabilization is provided by the replacement of the terminal U(–45)–A(+10) pair at the extremity of stem 2 by a C–G pair, ressembles that formed by mutant CFP5517. In these two mutants, two bases, A(–47) and A(–46), connect stem 1 and stem 2 by crossing the deep groove and U(–38) is unpaired and co-axially stacked between stem 1 and 2 (Fig. 11a). However, differences in the susceptibility of A(–47) to DMS suggest that this residue may form a labile pair with U(–38) in mutant LB3, resulting in a dynamic equilibrium between a pseudoknot with two As crossing the deep groove and a pseudoknot with only one A crossing. Therefore, the gain in free energy provided by the replacement of a A–U pair by a G–C pair at the upper extremity of helix 2 appears to induce a global stabilization of the pseudoknot, as revealed by a possible dynamic pairing between A(–47) and U(–38) at the junction of the two stems. The pseudoknot formed by mutant LB6 clearly differs from CFP5517 by the fact that only one A is crossing the major groove, as the result of the formation of a stable G(–46)–C(–38) pair. Strikingly the stabilization of the (–47)–(–38) pair induces a parallel destabilization of the A(–39)–U(+3) pair. These observations clearly indicate an unexpected flexibility at the junction of helices 1 and 2, at least in the absence of S15. Mutant LB5 appears to form a third kind of pseudoknot with characteristics of both LB3 and LB6. One puzzling result of mutations affecting the base pair closing helix 1 is the distal effect observed at A(–54) which becomes more reactive (with a weaker extent in LB4). We have presently no clear explanation for this behaviour. Unexpectedly, the addition of three Cs after the first codon (mutant LB12b) displaces the equilibrium towards the pseudoknot. This argues in favor of stabilizing contacts between nucleotides of loop 2′ and stem 1 (see below).

The pseudoknot carries determinants for S15 recognition

Our previous mutagenesis analysis showed that there is a very good correlation between the autoregulation efficiency and the affinity of S15 for the various mutants (10). The present structural analysis points out some structural elements required for S15 recognition. The results demontrate that pseudoknot formation, either in equilibrium with two hairpins or not, is a necessary condition for S15 recognition, but that it is not enough to ensure correct binding.

(i) Determinants in stem 2. A clear demonstration of the requirement for the pseudoknot are mutants LB1, LB2 and LB3. Obviously the loss of control and the reduced binding affinity in LB1 and LB2 is due to the destabilization of the pseudoknot. Control and binding are restored in LB3 by a double compensa-

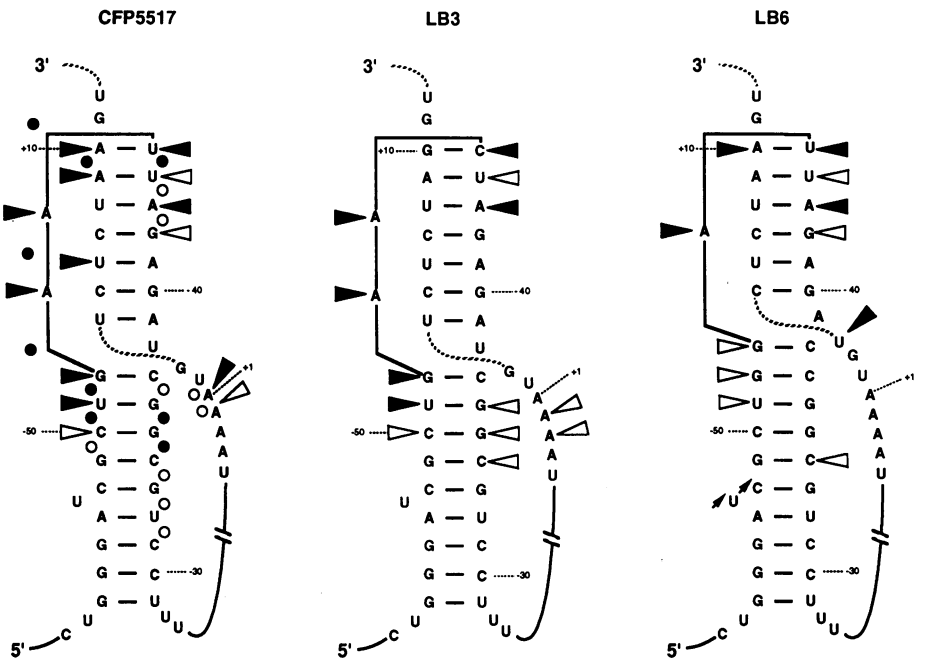

Figure 10. Diagram showing the footprint of S15 on mutants CFP5516, LB3 and LB6. The results are shown on the secondary structure from Figure 7. S15-induced protections are indicated by arrow heads for Fe(II)–EDTA and circles for ENU. The small arrows point to enhanced reactivity to Fe(II)–EDTA.

tory mutation. Moreover, the fact that LB3 is even more efficiently regulated indicates that there is a correlation between the stability of the pseudoknot and the control. Footprinting experiments provide further evidence for a close contact in the external part of stem 1 (Fig. 10). These contacts are clearly not sequence specific since the A–U pair can be replaced by a G–C pair. More likely, S15 interacts with the sugar–phosphate backbone of stem 2. This is further confirmed by the observation that neither the permutations of nucleotides AGA(–41) and UCU(+8) nor the replacement of pair A(–39)–U(+4) by a C–G pair alter the control efficiency (10).

(ii) Determinants in loop 1'. The present results show that the loss of control induced by the deletion of A(–46) or its substitution by U is due to the stabilization of stem–loop 2 and the subsequent incapacity to form the pseudoknot. Thus, this result does not allow to conclude that A(–46) is a S15 determinant. Nevertheless, footprinting experiments show a strong protection of residue A(–46) on both ribose and phosphate groups in all types of pseudoknot recognized by S15 (Fig. 9) and at position N1 in the wild-type RNA (7). On the other hand, the fact that A can be substituted by G but not by C (10) may favor the view that a purine at position –46 is required for specific contact with S15. Alternatively, S15 may recognize the bridging residue through interactions with the sugar–phosphate moiety without discriminating A from G, but a C at this position might impede S15 binding, either by inducing bad contacts, or by preventing favorable ones within the pseudoknot.

(iii) Determinants in stem 1. With mutants LB4, LB5 and LB6, genetic evidence was provided for a requirement of pairing between residues (–47) and (–38) (10). The present results confirm the existence of the G(–47)–C(–38) pair in LB6, thus indicating that this pair is required but without base specificity. However, there is no evidence for a stable A(–47)–U(–38) pair in mutants CFP5517 or LB3 in the free RNA. Furthermore, we could not evidence the simultaneous presence of base pairs A(–39)–U(+4) and A(–47)–U(–38) in any of the pseudoknot mutants. This apparent contradiction can be resolved if we postulate that these two base pairs require the presence of S15 to be stabilized. Therefore, the free RNA appears to be versatile and dynamic at the junction of stems 1 and 2, with mobility disappearing in the presence of S15. Most likely, in the complex, the pseudoknot adopts a single conformation with only A(–46) crossing the deep groove, the conformational transition energy being overcome by S15 binding. The differences observed in the footprints of CFP5517, LB3 and LB6 probably arise from subtle differences of the geometry of helices 1 and 2 due to the substitution of A–U pairs by G–C pairs. Consistently, footprinting experiments indicate that the upper part of stem 1 (5′-strand) is protected from hydroxyl-induced cleavage in the three types of pseudoknot (Fig. 10). Otherwise, the fact that mutant LB5, in which U(–38) is substituted by C, is able to adopt a pseudoknot conformation but is not recognized by S15 indicate that the formation of the pseudoknot is a necessary but not sufficient condition for S15 recognition.

The bulged U(–53) was initially proposed as a specific contact of S15, since the reactivity of N3 was reduced in the presence of S15 (7). However, its deletion (mutant LB11a) does not drastically alter the regulation (10), while the equilibrium does not appear to be affected. Moreover, the present results fail to show any protection at the level of the sugar–phosphate backbone in any of the tested pseudoknot mutants. Therefore, we conclude that U(–53) is not a major determinant for S15. However, the

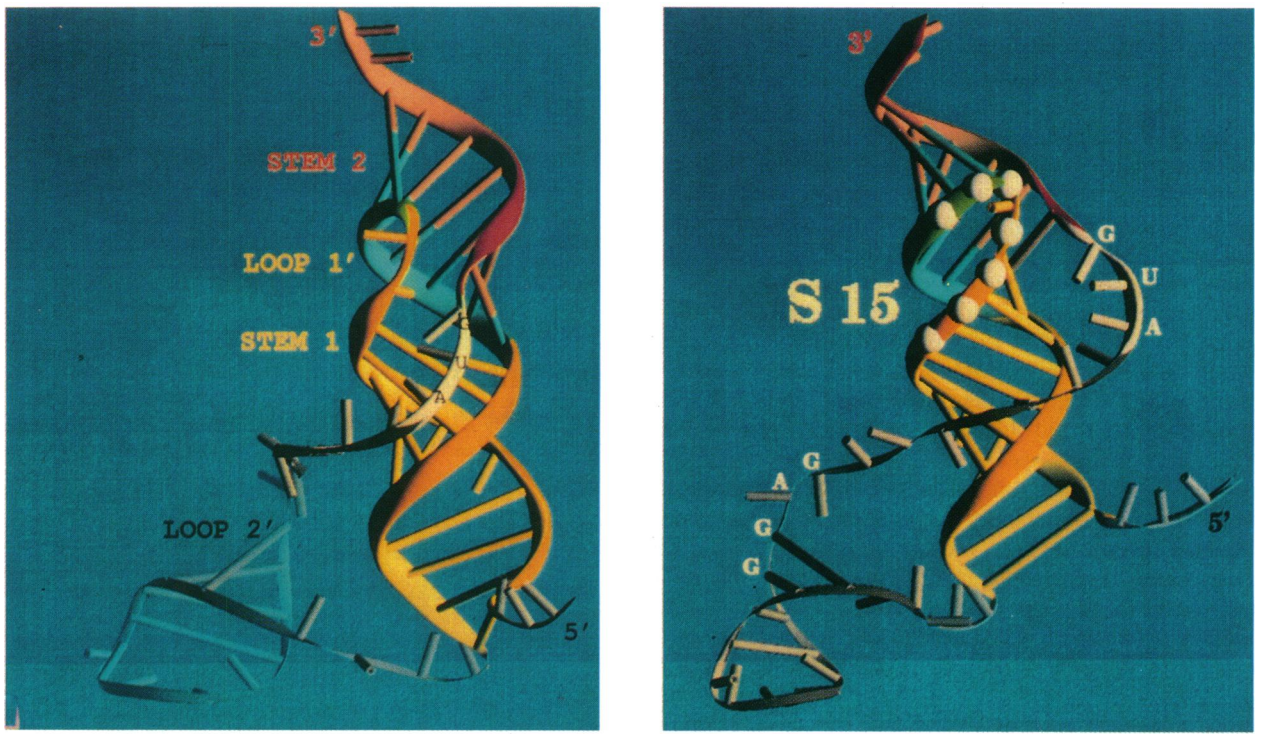

Figure 11. Three-dimensional model of the pseudoknot. Only the backbone is represented, with the base pairs shown as bars. Stem 1 and loop 1′ are shown in yellow, the 5′ strand of stem 2 in red, the 3′ strand in green and loop 2′ in light grey. The AUG initiation codon and the GAGG Shine–Dalgarno sequence are indicated. (**a**) Free RNA with two adenines [A(–47) and A(–46)] crossing the deep groove. This type of pseudoknot corresponds to CFP5517 and LB3. (**b**) The pseudoknot form stabilized by S15, with one single adenine [A(–47)] crossing the deep groove. The spheres correspond to positions that are protected by S15 from hydroxyl radical hydrolysis in the three mutants CFP5516, LB3 and LB6. They are positioned on the phosphate 5′ to the corresponding base. The orientation of loop 2′, although not precisely known, is chosen so that its 3′ part covers the shallow groove of stem 1. Although it might be different in the free and bound state, the same orientation is shown in both views, since at that stage we have not enough information to justify a more detailed model. The drawings were made using the software DRAWNA (20) and SHOWCASE on a EXTREME 2 of Silicon Graphics.

double mutation Δ (–53)/A(–49) (mutant LB11b) dramatically affects both control and S15 recognition. Since the conformational equilibrium does not seem to be affected in this mutant, we can deduce that the wobble U(–49)–G(–36) pair contains S15 determinant(s). In agreement with this assumption is the finding that converting the U(–49)–G(–36) pair into a canonical U–A pair abolishes the control (10). Consistently, S15 does protect the ribose moiety of U(–49) from hydroxyl attack in all three pseudoknot mutants and strong protections against ENU alkylation are observed in CFP5517 (Fig. 10).

(iv) Determinants in loop 2′. The presence of S15 determinants in loop 2′ is hinted at by mutants LB12a and LB12b. In these two mutants, the size of loop 2′ is increased by three additional nucleotides inserted after the AUG codon. The effect of such mutations depends on the sequence inserted. The most unexpected result is that the CCC insertion (LB12b) appears to favor a particular pseudoknot form, which is probably stabilized by interactions between nucleotides of loop 2′ (near their junction with stem 2) and the minor groove of stem 1. On the other hand, there is evidence that similar interactions also exist in the other pseudoknot mutants and that they are stabilized by S15. Indeed, A(+1) is unreactive at position N7 in both CFP5517 and LB3 and G(+3) is poorly accessible to RNase T1 in CFP5517, LB3 and LB6 (results not shown). Moreover, earlier results showed that A(+1) and U(+2) are less reactive to chemicals in the presence of S15 (7), and the present footprinting experiments reveal protections against both Fe(II)–EDTA and ENU at positions (+1) and (–1) in CFP5517, and at (–1) and (–2) against radical hydrolysis in LB3. The extensive protections of phosphates –31 to –36 on the 3 ′strand of stem 1 induced by S15 might be related to the decrease of reactivity of phosphates –33 and –34 induced by magnesium.

The three-dimensional conformation recognized by protein S15

The present data reveal an unexpected number of pseudoknot forms with versatility at the junction between stems 1 and 2, still recognized by S15. Consistently with previous data (10) and footprinting analyses of the different mutants, S15 stabilizes the pseudoknot in a form in which one single nucleotide [A(–46)] crosses the deep groove (Fig. 11b). Such types of peudoknots have been suggested in bacteriophage T4 gene 32 (2) and in plant viral RNAs (18). The stem 2 of these RNAs contains 6 or 7 base pairs and one can expect that a simple nucleotide is sufficient for

crossing the major groove in such a pseudoknot (18–19). The three-dimensional model shows that one nucleotide is indeed enough to bridge the distance across the major groove of a 7 base-pair helix. However, the experimental data show that, in the absence of S15, two mutants (CPF5517 and LB3) favor rather a pseudoknot with a 7 base pair-stem 2 and two bases in loop 1′, while another one (LB6) adopts a pseudoknot with a 6 base pair-stem 2 and one base in loop 1′. Thus, it appears that the pseudoknot with a 7 base pair-stem with one bridging nucleotide is not always thermodynamically the most favorable.

Footprinting experiments conducted on the three pseudoknots recognized by S15 (CFP5517, LB3 and LB6) show that the protein sits on the deep groove of the co-axial stack, especially on the region which displays a sharp turn of the sugar–phosphate backbone (Fig. 11b), and shields the bridging A residue. Most likely S15 recognizes a specific and unique three-dimensional conformation of the sugar–phosphate backbone that is provided by the pseudoknot. Noteworthy, most of the additional protections that are specifically observed with mutants CFP5517, LB3 and LB6 are also located on the same side of the pseudoknot. The reduction of the reactivity of the bulged U(–53) at N3 induced by S15 binding (7) can hardly be interpreted by a direct protection in this model. This also fits with the finding that this residue is not a major determinant (see above). Most likely, a local conformational rearrangement of stem 1 is induced by the interaction of S15. Some structural interdependence within this particular region of stem 1 is further suggested by the effects of mutations in the A(–47)–U(–38) pair that closes the top of stem 1. Therefore, our results reveal an unexpected complexity and subtlety of the pseudoknot and of its recognition by S15.

The autoregulation of protein S15 at the translational level through a dynamic equilibrium between two local structural states of the 5′ region of the S15 mRNA is another example of the biological role of alternative pairings in structure–function relationships of RNA molecules. Alternative pairing depends first of the number and on the types of base pairs formed in each state and on the conformational energetics at the junctions between the helices and with the single-stranded regions. The existence of a dynamic equilibrium at room temperature *in vitro* implies a free energy difference between the two states close to zero. Although there is no evidence for a role of the dynamic equilibrium *in vivo*, biological resilience requires that mutations do not disrupt the pseudoknot form too easily. This implies that the biologically relevant control element, here the pseudoknot, can exist in a variety of topologically equivalent structures recognizable and shapable by the S15 protein. The present results illustrates how the pseudoknot structure is suitable for such requirements.

ACKNOWLEDGEMENTS

We are indebted to F. Eyerman and L. Dondon for skilful technical assistance and to C. Cachia for providing protein S15. This work is supported by grants from the Centre National de la Recherche Scientifique (UPR 9002 and URA1139) and from the Fondation de la Recherche Médicale (UPR 9002). E. W. acknowledges a grant BIO2CT-93-035 from the EEC.

REFERENCES

1 Tang, C. K. and Draper, D. (1989) *Cell* **57**, 531–536.
2 McPheeters, D. S., Stormo, G. D. and Gold, L. (1988) *J. Mol. Biol.* **201**, 517–535.
3 Asano, K., Kato, A., Moriwaki, H., Hama, C., Shiba, K. & Mizobuchi, K. (1991) *J. Biol. Chem.*, **266**, 3774–3781.
4 Brierley, I., Rolley, N. J., Jenner, A. J. and Inglis, S. C. (1991) *J. Mol. Biol.* **220**, 889–902.
5 Portier, C., Dondon, J. and Grunberg-Manago, M. (1990a) *J. Mol. Biol.* **211**, 407–414.
6 Portier, C., Philippe, C., Dondon, J., Grunberg-Manago, M., Ebel, J.P., Ehresmann, B. and Ehresmann, C. (1990b) *Biochim. Biophys. Acta* **1050**, 328–336.
7 Philippe, C., Portier, C., Grunberg-Manago, M., Ebel, J.P., Ehresmann, B. and Ehresmann, C. (1990) *J. Mol. Biol.* **211**, 415–426.
8 Philippe, C., Eyermann, F., Bénard, L., Portier, C., Ehresmann, B. and Ehresmann, C. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4394–4398.
9 Philippe, C., Bénard, L., Eyermann, F., Cachia, C., Kirillov, S.V., Portier, C., Ehresmann, B. and Ehresmann, C. (1994) *Nucleic Acids Res.* **22**, 2538–2546.
10 Bénard, L., Philippe, C., Dondon, L., Grunberg-Manago, M., Ehresmann, B., Ehresmann, C and Portier, C. (1994) *Mol. Microbiol.* **14**, 31–40.
11 Cachia, C., Flamion, P.J. and Schreiber, J.P. (1991) *Biochimie* **73**, 607–610.
12 Tullius, T. D. and Dombroski, B. A. (1987) *Methods Enzymol.* **155**, 537–558.
13 Westhof, E. (1993) *J. Mol. Struct. (Theochem)* **286**, 203–210.
14 Westhof, E., Romby, P., Romaniuk, P.J., Ebel, J.P., Ehresmann, C. and Ehresmann, B. (1989) *J. Mol. Biol.* **207**, 417–431.
15 Celander, D. W. and Cech, T. R. (1990) *Biochemistry* **29**, 1355–1361.
16 Vlassov, V. V., Giegé, R. and Ebel, J.P.(1981) *Eur. J. Biochem.* **119**, 51–59.
17 Groebe, D. R. and Uhlenbeck, O. C. (1988) *Nucleic Acids Res.* **16**, 11725–11735.
18 Pleij, C. W. A. , Rietveld, K. and Bosch, L. (1985) *Nucleic Acids. Res.* **13**, 1717–1731.
19 Westhof, E. and Jaeger, L. (1992) *Current opinion in Structural Biology* **2**, 327–333.
20 Massire, C., Gaspin, C. and Westhof, E. (1994) J. *Mol. Graph.* **12**, 201–206.