Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 6, No. 10, pp. 1157-1162, 1996

0960-894X/96 $15.00 + 0.00

PII: S0960-894X(96)00189-8

STRUCTURAL CHARACTERIZATION OF A FLAVIN-SPECIFIC RNA APTAMER BY CHEMICAL PROBING

Petra Burgstaller and Michael Famulok*

Institut für Biochemie - Genzentrum der LMU München, Würmtalstraße 221, 81375 München, Germany. Fax: +49-89-7401-7448; e-mail: Famulok@lmb.uni-muenchen.de

Abstract: In a previous study, an RNA aptamer for the specific recognition of flavin cofactors was isolated from a combinatorial RNA-pool. We have now carried out a structural analysis of this RNA by chemical probing and identified residues protected from modification in the presence of the cofactor and modifications that interfere with its complexation. An induced fit of the RNA motif occurs upon cofactor binding. Our analysis led to a refined secondary structure model for the FMN/RNA complex. Copyright © 1996 Elsevier Science Ltd

Through screening of combinatorial nucleic acid libraries, a considerable number of RNA and DNA sequences ("Aptamers") have been isolated which bind with high specificity and affinity to different target molecules[1]. Aptamers can complex both high- and low molecular weight molecules; in the latter case the nucleic acid serves as the ligand which forms a specific binding pocket complementary to the complexed molecule - analogous to the "lock and key" principle[2]. In some cases, secondary structure elements such as pseudoknots[3], internal loops[4], hairpin loops[5], and G-quartets[6] were proposed as functionally important motifs on the basis of covariations among the selected sequences. These motifs are usually identified by defining helical regions in which base pairing occurs in the Watson-Crick sense and which differ from residues in which Watson-Crick pairing is not possible. Such "unpaired" residues are mostly found to be highly conserved and to contain functionally important structural elements. Secondary structures based on the artificial phylogeny of covarying sequences must be considered a relatively crude model as details about structural features can only rarely be extracted from sequence comparisons. Secondary structure refinements can, however, be obtained by chemical modification analyses[7]; this technique allows to probe individual atoms in a nucleotide residue and to test whether they are directly or indirectly involved in the complexation. In a previous study, we reported the isolation of an RNA motif for the specific complexation of isoalloxazine containing biological cofactors such as flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD)[8]. The RNA motif contained an internal loop which consisted of six conserved bases in the upper and five in the lower strand, flanked by Watson-Crick helices. This RNA recognizes exclusively the isoalloxazine moiety of the cofactor with values of K_d between 300 and 500 nM (Figure 1).

Here we used chemical probing to examine the secondary structure of the FMN binding site of this RNA aptamer with the specific structure probes dimethylsulfate (DMS), kethoxal (KE), and 1-cyclohexyl-3-[2-(morpholino ethyl)-carbodiimide] metho-p-toluene sulfonate (CMCT). These probes specifically modify bases at atom positions involved in Watson-Crick pairing, but can access their site of modification only in the unpaired state. DMS reacts with positions N1 in adenosine and N3 in cytosine, KE with N1 and the exocyclic NH_2-group in guanosine, and CMCT with N3 in uridine. To identify those base positions which are directly or indirectly involved into the binding of the cofactor we carried out a modification/selection analysis[9]. The RNA was modified so that, on average, each molecule contained one modification. The modified pool was then separated by affinity chromatography with FMN-agarose into binding and non-binding molecules. Reverse transcription

with a radiolabelled 3'-primer and separation of the resulting cDNA by polyacrylamide gel electrophoresis allowed the identification of those base positions at which modification results in a loss or retainment of binding

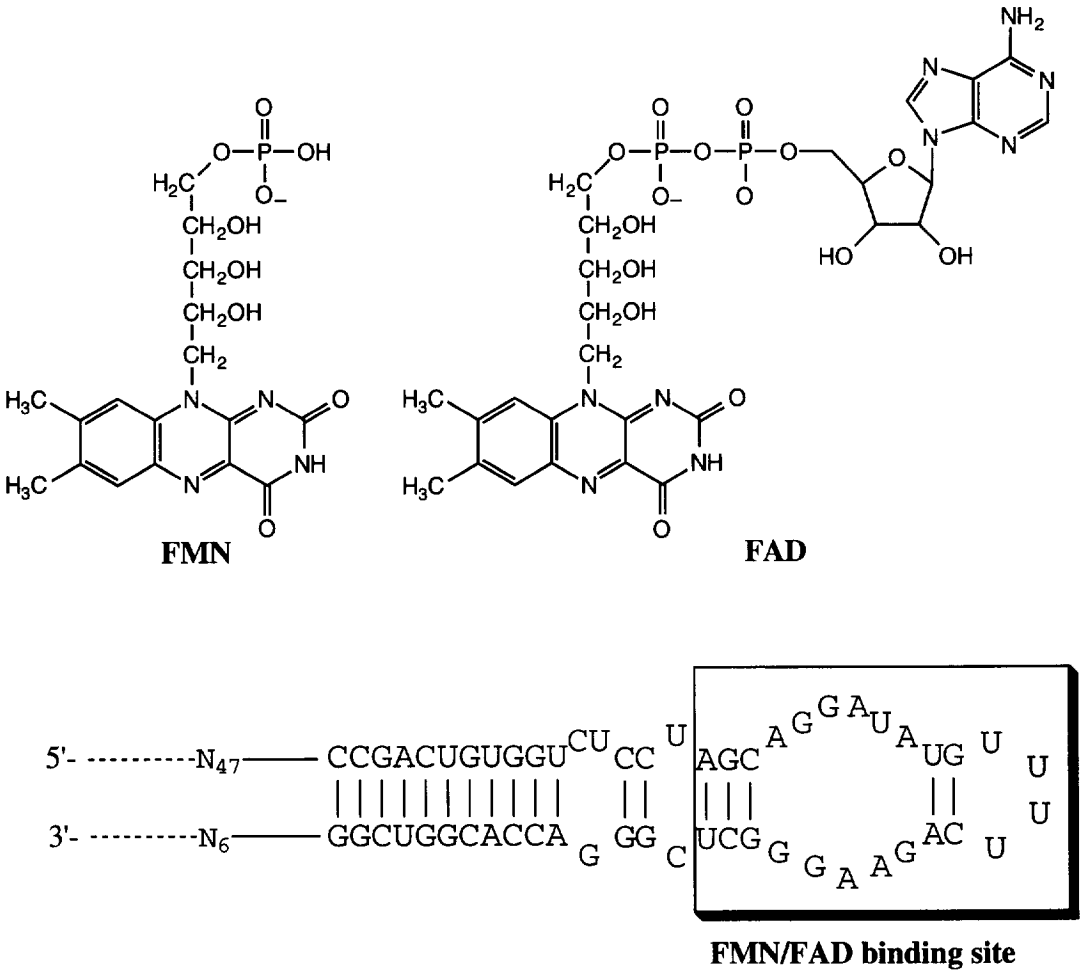

Figure 1: Chemical structure of the FMN and FAD cofactors which are complexed by the RNA shown below. The minimal cofactor binding site of this RNA is shown in the boxed region. The total length of the RNA aptamer was 109 nucleotides. Only the isoalloxazine moiety of the cofactors are required for the specific binding.

activity. As a control, a modification experiment in the presence of the cofactor was carried out. The same technique has recently been used to identify bases within the SunY self splicing group I intron which are essential for splicing[10]. The result of this analysis is shown in Figure 2[11].

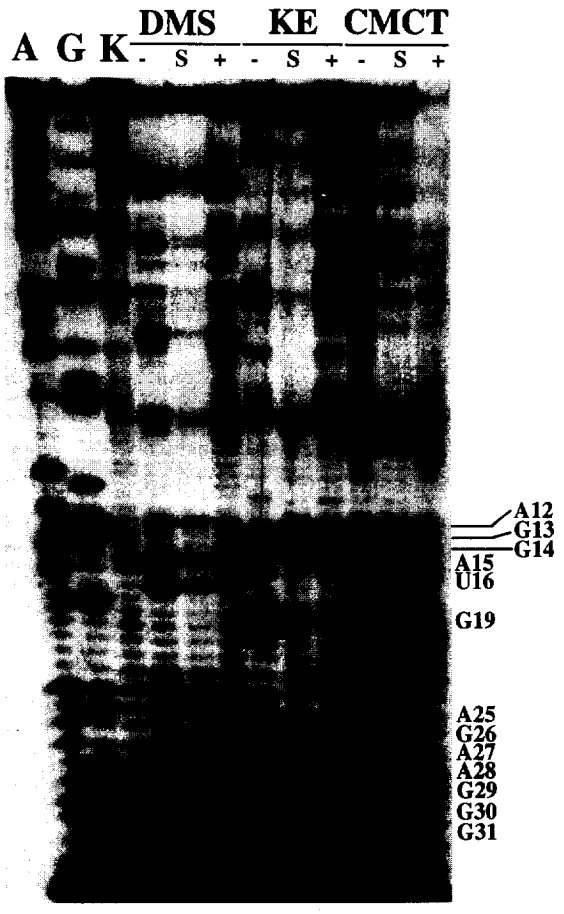

Figure 2: Chemical Modification, damage selection, and footprinting data of the RNA clone FMN-2 with DMS, kethoxal, and CMCT. A and G indicate the dideoxy sequencing lanes, K the primer extension of unmodified RNA. (-) indicates modification of the free RNA, (s) damage selected RNA (RNA that bound to the FMN column and eluted with FMN after modification with the probe shown above), (+) indicates modification of RNA in the presence of 100 μM FMN in solution.

In the free RNA every base within the two conserved consensus sequences A12-G13-G14-A15-U16-A17-U18 and A25-G26-A27-A28-G29-G30 is accessible for chemical modification. In accordance with the secondary structure model shown in Figure 3, the bases upstream of the binding site are not modified, due to Watson-Crick pairing and helix formation. Based on the covariations of the selected sequences we previously suggested that U18/A25 and the non-conserved nucleotides G19/C24 form a short helix. However, this is not observed in the free RNA. In contrast, modification in the presence of 100 μM FMN clearly results in a reduced intensity of the corresponding bands, indicating that the formation of the short helix U18-G19/A25-C24 is induced and stabilized by the bound cofactor[14]. This result indicates a conformational change ("induced fit") of the RNA aptamer upon ligand binding. Additional support for this proposal is obtained from the modification at A15. This nucleotide is the only position within the conserved loop region at which an increased intensity of the band is observed after FMN binding. Hence, N1 in A15 is more accessible to DMS-modification in the presence of FMN than in its absence. Interestingly, A15 is the only position within the invariant consensus sequence which is not conserved among different isolates and can be substituted by the three other bases without altering the strength of binding[3]. Taken together, these observations suggest that the contribution of residue 15 to the recognition of the flavine through the formation of a direct contact or by tertiary interactions must be considered highly unlikely. The

increased intensity of the modification signal upon complexation to the cofactor indicates that residue 15 is bulged out of the overall structure forming the recognition motif and is directed towards the solvent.

At residues G13, G14, U16, G26, G29, and G30 within the conserved loop region a clear reduction of modification in the presence of the cofactor can be detected. The intensity of modification at A17, A27, and A28 remains unchanged. This result is in accordance with the modification/selection analysis: modification at the former residues results in loss of binding activity whereas binding of the RNA to FMN is retained despite modification at the latter bases. The only exception is residue A12; the N1-atom of A12 is protected from modification when the probing reaction is carried out in the presence of 100 μM FMN but modification at N1 performed with the free RNA does not significantly interfere with FMN binding in the damage selection. Taken together, these results confirm the formation of the internal loop consisting of the conserved A12-G13-G14, U16-A17, and G26-A27-A28-G29-G30. Furthermore, they indicate that extensive tertiary interactions among the residues and possibly direct interactions with the bound cofactor take place within the conserved internal bulge region. The damage selection data are consistent with the notion that these interactions do not involve the Watson-Crick side of A17, A27, and A28, but do involve the Watson-Crick side of G13, G14, U16, G26, G29, G30 and possibly that of A12. The summary of the chemical probing data of the free RNA clone FMN-2 is shown in Figure 3; the damage selection and footprinting data are summarized in Table 1.

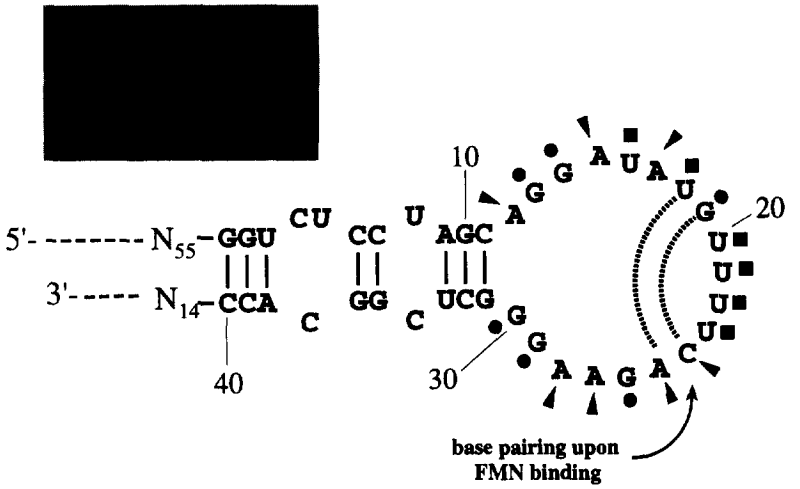

Figure 3: Summary of the reactivity of nucleotides in the free RNA to the chemical probes DMS (triangle), kethoxal (circle), and CMCT (square). Only the base positions corresponding to the FMN binding site are shown. The bases were numbered accordingly. The pairing of U18·A25 and G19·C24 occuring upon cofactor binding is indicated by the dashed lines.

In conclusion, we showed that binding of the cofactor to the RNA not only induces conformational changes as corroborated by the hypersensitivity of residue 15 in the complex, but might also stabilize extensive tertiary interactions and non canonical base pairing within the unpaired loop regions. To get further insight into the structural details of the complex detailed NMR structural analyses will have to be carried out. In the context of an

NMR study, the structural refinement described here will prove to be helpful to obtain a higher resolved picture of the solution structure of the RNA/FMN complex.

Table 1: Summary of the chemical modification and damage selection data

Base position	Tolerated in damage selection	Modified in presence of cofactor
A12	+	-
G13	-	-
G14	-	-
A15	+	+
U16	-	-
A17	+	+
U20	+	+
U21	+	+
U22	+	+
U23	+	+
G26	-	-
A27	+	+
A28	+	+
G29	-	-
G30	-	-

(+) indicates that modification is tolerated in the damage selection or that the residue is modified in the presence of 100 μM cofactor, (-) indicates that modification is not tolerated in the damage selection or that the residue is protected from modification in the presence of 100 μM cofactor.

Acknowledgement: This work was supported by grants from the Deutsche Forschungsgemeinschaft and the European Union (Project no: Biot-CT93-0345). We thank E.-L. Winnacker for his generous support and A. Hüttenhofer and T. Tiedchen for helpful discussions.

References and Notes:

1. (a) Gold, L. *J. Biol. Chem.* **1995**, *270*, 13581. (b) Burke, J. M.; Berzal-Herranz, A. *FASEB J.* **1993**, *7*, 106. (c) Klug, S. J.; Famulok, M. *Mol. Biol. Rep.* **1994**, *20*, 97. (d) Gold, L.; Polisky, B.; Uhlenbeck, O.; Yarus, M. *Annu. Rev. Biochem.* **1995**, *64*, 763. (e) Famulok, M.; Szostak, J. W. *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 979. (f) Famulok, M.; Szostak, J. W. In *Nucleic Acids and Molecular Biology*; Eckstein, F.; Lilley, D.M.J., Eds. Springer Verlag: New York, 1993; Vol. 7, pp. 271-284. (g) Szostak, J. W. *Trends Biochem. Sci.* **1992**, *17*, 89. (h) McGown, L.B.; Joseph, M.J.; Pitner, J. B.; Vonk, G. P. Linn, C. P. *Anal. Chem.* **1995**, 663A.
2. (a) Joyce, G. F. *Curr. Opin. Struct. Biol.* **1994**, *4*, 331. (b) Ellington, A. D. *Curr. Biol.* **1994**, *4*, 427. (c) Hirao, I.; Ellington, A.D. *Curr. Biol.* **1995**, *5*, 1017.
3. (a) Tuerk, C.; MacDougal, S.; Gold, L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6988. (b) Lorsch, J. R.; Szostak, J. W. *Biochemistry* **1994**, *33*, 973.

4. (a) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550. (b) Famulok, M. *J. Am. Chem. Soc.* **1994**, *116*, 1698. (c) Bartel, D.P.; Zapp, M. L.; Green, M.; Szostak, J. W. *Cell* **1991**, *67*, 529. (d) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425.
5. (a) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505. (b) Ellington, A. D.; Szostak, J. W. *Nature* **1992**, *355*, 850. (c) Lato, S. M.; Boles, A. M.; Ellington, A. D. *Chem. Biol.* **1995**, *2*, 291. (d) Wang, Y.; Rando, R. R. *Chem. Biol.* **1995**, *2*, 281. (e) Wallis, M. G.; von Ahsen, U.; Schroeder, R.; Famulok, M.*Chem. Biol.* **1995**, *2*, 543.
6. (a) Huizenga, D. E.; Szostak, J. W. *Biochemistry* **1995**, *34*, 656. (b) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, *355*, 564. (c) Wang, K. Y.; McCurdy, S.; Shea, R. G.; Swaminathan, S.; Bolton, P. H. *Biochemistry* **1993**, *32*, 1899. (d) Macaya, R. F.; Schulze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3745. (e) Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235*, 1532. (f) Lauhon, C. T.; Szostak, J. W. *J. Am. Chem. Soc.* **1995**, *117*, 1246.
7. Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J.-P.; Ehresmann, B. *Nucleic Acids Res.* **1987**, *15*, 9109.
8. Burgstaller, P.; Famulok, M. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1084.
9. Burgstaller, P.; Kochoyan, M.; Famulok, M. *Nucleic Acids Res.* **1995**, *23*, 4769.
10. von Ahsen, U.; Noller, H. F. *EMBO J.* **1993**, *12*, 4747.
11. Experimental conditions: 1-5 pmol RNA FMN-2 were denatured for 3 min at 95° C and renatured in the absence or presence of 100 µM of FMN in 250 mM NaCl, 50 mM sodium cacodylate pH 7.4 and 5 mM $MgCl_2$ for 10 min at room temperature. For probing with CMCT 50 mM potassium borate pH 8.0 instead of sodium cacodylate was used. Chemical modification was performed by addition of 1 µl DMS (1:5 dilution in 96 % Ethanol), 1 µl kethoxal (1:5 dilution in H_2O, stock solution: 37 mg/ml) or 12.5 µl CMCT (32 mg/ml in H_2O) to a final reaction volume of 50 µl, followed by incubation at 25 °C for 20 min. After precipitation in the presence of 10 µg tRNA, samples were dissolved in H_2O (DMS and CMCT modified samples) or 25 mM potassium borate pH 7.0 (kethoxal modified samples). For damage selection experiments, modified RNA was renatured in selection buffer (250 mM NaCl, 50 mM Tris pH 7.6, 5 mM $MgCl_2$; for kethoxal modified RNAs 25 mM K-borate was added) and applied on FMN-derivatized agarose (0.5 mM FMN). After washing with five column volumes of buffer, the remaining bound molecules were eluted with a 1.0 mM solution of the cofactor in binding buffer. The eluted RNA was precipitated and dissolved in water. Detection of modified positions by primer extension and polyacrylamide gel electrophoresis was performed as described previously (9, 12, 13) using 5'-^{32}P end-labeled primer M20.106 (5'-GTGGATCCGACCGTGGTGCC-3').
12. Stern, S.; Moazed, D.; Noller, H. F. *Methods Enzymol.* **1988**, *164*, 481.
13. Moazed, D.; Noller, H. F. *Cell* **1989**, *57*, 585.
14. When the free RNA was probed in the absence of the bivalent cation Mg^{2+} the modification at positions U18, G19, C24, and A25 was even stronger than in the presence of 5 mM Mg^{2+} (data not shown). In the absence of Mg^{2+} FMN binding is significantly reduced. These data show that the short helix is required for complexation of the cofactor but requires Mg^{2+} for stabilization in addition to FMN.

(*Received in Belgium* 19 *February* 1996; *accepted* 19 *April* 1996)