BioTechnologia
vol. 94(1) • pp. 59-65 • 2013 LECTURES

Introns of plant pri-miRNAs are required for proper biogenesis and function of miRNAs

DAWID BIELEWICZ [1], MALGORZATA KALAK [1], MARIA KALYNA [2], DAVID WINDELS [3], ANDREA BARTA [2], FRANCK VAZQUEZ [3], ZOFIA SZWEYKOWSKA-KULINSKA [1], ARTUR JARMOLOWSKI [1]

[1] Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

[2] Max F. Perutz Laboratories, Medical University of Vienna, Vienna, Austria

[3] Botanical Institute of the University of Basel, Zürich-Basel Plant Science Center, Basel, Switzerland

Plant *MIR* genes are independent transcription units that encode long primary miRNA precursors which usually contain introns. In the majority of cases a miRNA/miRNA* hairpin is found in the first exon of pri-miRNAs, but there are also miRNA genes where a miRNA is encoded in the second or third exon. Biogenesis of plant miRNAs requires complex and multistep processing. The miRNA maturation steps include: splicing of pri-miRNAs, miRNA-containing hairpin excision, miRNA/miRNA* duplex formation, and miRNA incorporation into the RISC complex. To answer the question whether splicing plays an essential role in the efficiency of mature miRNA production, we analyzed processing of five different intron-containing *A. thaliana* pri-miRNAs in several Arabidopsis SR protein mutants. The data show that in some of the SR mutants tested the level of miRNAs which are processed from the intron-containing genes is significantly decreased. Next, we asked the question whether the observed by us changes in the level of mature miRNAs are caused by direct or rather indirect effects. To test it we introduced three variants of the *MIR163* gene: a native form containing one intron, a gene containing mutated 5′ and 3′ splice sites, and an intronless *MIR163* variant, into the *A. thaliana mir163-2* mutant (SALK_034556), in which the inserted T-DNA sequence had disrupted the endogenous *MIR163* gene. At least two independent lines for each transgenic construct were analyzed for the pri-miRNA163 and mature miR163 levels. Introduction of the wild type form of *MIR163* showed the same level of pri-miRNA 163 and its mature form as it was observed in wild type plants. In the case of the intronless *MIR163* construct we observed the accumulation of pri-miRNA 163, while the level of mature miRNA 163 was decreased about three times. In the case of transgenic plants containing the *MIR163* gene with mutated both splice sites, the level of mature miRNA 163 was again lower when compared to wild type plants. Altogether, our results show that introns stimulate biogenesis of plant miRNAs derived from intron-containing genes. We demonstrate also that the stimulating effects of the intron mostly reside in the 5′ss rather than on a genuine splicing event. Our findings are biologically significant since the presence of functional splice sites in the *MIR163* gene appears mandatory for pathogen-triggered accumulation of miR163 and proper regulation of at least one of its targets.

Common sense applied to some PDB structures

ZBIGNIEW DAUTER

Synchrotron Radiation Research Section, MCL, National Cancer Institute,
Argonne National Laboratory, Argonne, USA

dauter#@anl.gov

The Protein Data Bank (PDB) serves as an archive of macromolecular structures acquired by various biophysical methods. Currently it contains about 90 000 atomic models, majority of which were obtained by X-ray crystallography, but ca. 10 000 by NMR and a few hundred by other methods. Before submission to the PDB each of these structures is validated by one of the dedicated programs, automatically checking for correctness of various statistical, geometrical, symmetry and other aspects of the investigated models. The programs used for validation (PROCHECK, WHAT_IF, MOLPROBITY and many others) are highly sophisticated and very useful, also as tools during the process of model building and refinement. However, macromolecular crystallography is a very interdisciplinary branch of science, integrating such fields as physics (diffraction experiments), chemistry (investigated compounds), biology (source of samples), medicine (ultimate goal of the research), with advanced computing (need for fast computers and elaborate graphical tools) and mathematics (symmetry theory and other sophisticated algorithms). The task of validation programs is therefore enormous, if one takes into account how widely different all these structures are in terms of diffraction data resolution, crystal symmetry, chemical properties and biological function. It is not surprising that among many PDB structures there are some examples containing certain erroneously interpreted aspects in terms of physics, chemistry, biology, or symmetry. Such errors may result from the wrong interpretation of results provided by various computer programs or from the lack of experience in one of the mentioned above branches of science. Usually such misinterpretations do not invalidate the biological conclusions, but sometimes may have serious negative consequences. The presence of small amount of errors among so many crystal structures in the PDB is a "statistically inevitable" consequence of the enormous success of macromolecular crystallography, where the availability of very powerful and highly automatic programs makes it possible for investigators less experienced in the underlining methodology to relatively easily solve structures of the molecules of interest. Many examples of misinterpretations could be avoided by critical analysis of all available information by the knowledgeable specialist. In that respect a common sense provided by the human brain can be treated as a valuable validation tool.

Downregulation of *Cap Binding Protein 80* gene expression as a strategy to engineer a drought-tolerant potato

Zofia Szweykowska-Kulinska [1], Marcin Pieczynski [1], Waldemar Marczewski [2], Jacek Hennig [3], Jakub Dolata [1], Dawid Bielewicz [1], Paulina Piontek [1], Anna Wyrzykowska [1], Dominika Krusiewicz [2], Danuta Strzelczyk-Zyta [2], Dorota Konopka-Postupolska [3], Magdalena Krzeslowska [4], Artur Jarmolowski [1]

[1] Department of Gene Expression, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

[2] Plant Breeding and Acclimatization Institute, National Research Institute, Młochow, Poland

[3] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa, Poland

[4] Department of General Botany, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

Developing new strategies for crop plants to respond to drought is crucial for their innovative breeding. Recent studies have revealed that the proteins involved in RNA processing affect the ABA (abscisic acid) signal transduction occurring in drought-stressed plants (Fedoroff, 2002; Hugouvieux et al., 2001). Among these proteins, the cap-binding protein 80 (CBP80, also known as Abscisic Acid Hypersensitive 1, ABH1) gene in *Arabidopsis thaliana* has been shown to be an important player in the regulation of the ABA transduction pathway and in drought tolerance. Interestingly, its inactivation in *A. thaliana* leads to an ABA-hypersensitive stomatal closing and reduced wilting during drought (Hugouvieux et al., 2001, 2002; Kmieciak et al., 2002). In addition, Papp et al. (2004) have shown that the loss of function of the Arabidopsis cap-binding protein 20 (CBP20) also confers a hypersensitivity to ABA during germination and increased water-deficit tolerance during drought stress. The CBP80 protein forms a dimer with the CBP20 protein (Kierzkowski et al., 2009), producing the CBC (Cap-Binding Complex), a complex that recognizes and binds to the cap structure of RNA Pol II transcripts in the nucleus.

The *CBP80* gene in the potato cultivar Desiree was silenced using artificial microRNAs. Transgenic plants (amiR80.2-14) displayed a higher tolerance to drought, ABA-hypersensitive stomatal closing, an increase in leaf stomata and trichome density, and compact cuticle structures with a lower number of microchannels. These findings were correlated with a higher tolerance to water stress. The level of miR159 was decreased, and the levels of its target mRNAs MYB33 and MYB101 increased in the transgenic plants subjected to drought. Similar trends were observed in an Arabidopsis cbp80 mutant. Tunnel (field) experiment allowed us to characterize tuber progeny. Tuber yield, structure, starch and reducing sugars content analyses showed no differences between Desiree and amiR80.2-14 transgenic potato plants. The evolutionary conservation of CBP80, a gene that plays a role in the response to drought, suggests that it is a candidate for genetic manipulations that aim to obtain improved water-deficit tolerance of crop plants.

References

Fedoroff N.V. (2002) Curr. Opin. Plant Biol. 5: 452-459.

Hugouvieux V., Kwak J.M., Schroeder J.I. (2001) Cell 106: 477-487.

Hugouvieux V., Murata Y., Young J.J., Kwak J.M., Mackesy D.Z., Schroeder J.I. (2002) Plant Physiol. 130: 1276-1287.

Kmieciak M., Simpson C.G., Lewandowska D., Brown J.W.S., Jarmolowski A. (2002) Gene 283: 171-183.

Papp I., Mur L.A., Dalmadi A., Dulai S., Koncz C. (2004) Plant Mol. Biol. 55: 679-686.

Kierzkowski D., Kmieciak M., Piontek P., Wojtaszek P., Szweykowska-Kulinska Z., Jarmolowski A. (2009) Plant J. 59: 813-823.

Pieczynski M., Marczewski W., Hennig J., Dolata J., Bielewicz D., Piontek P., Wyrzykowska A., Krusiewicz D., Strzelczyk-Zyta D., Konopka-Postupolska D., Krzeslowska M., Jarmolowski A., Szweykowska-Kulinska Z. (2012) Plant Biotech. J. Plant Biotechnol. J. 2012 Dec 11. doi: 10.1111/pbi.12032.

CpG underrepresentation leads to methyltransferase M.MpeI

MAREK WOJCIECHOWSKI [1], HONORATA CZAPINSKA [1], MATTHIAS BOCHTLER [1, 2, 3]

[1] International Institute of Molecular and Cell Biology, Warszawa, Poland

[2] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa, Poland

[3] Schools of Chemistry and Biosciences, Cardiff University, Cardiff, UK

mbochtler@iimcb.gov.pl

In many eukaryotic species DNA cytosine methylation is predominantly found in the context of the CpG dinucleotide (Jeltsch, 2010). It is thought to be essential in mammals and many other model animals, with the notable exception of the fruit fly *Drosophila melanogaster* (Jeltsch, 2010). However, cytosine methylation comes at a price. Genome-wide studies have consistently shown that CpG methylation in eukaryotes is associated with CpG depletion (Cooper et al., 1989). As a result, the CpG dinucleotide is found several-fold less frequently in the nuclear DNA of higher mammals, including humans, than one might expect based on the GC content of the DNA (Fig. 1). The reasons for the link between CpG methylation and depletion are both chemical and biological. At the chemical level, cytosine methylation promotes deamination and leads to thymines (Shen et al., 1992, Shen et al., 1994). More importantly, at the biological level, methylcytosines converted to thymines are difficult to identify as a damage product and difficult to repair to cytosines through DNA repair pathways. Both the chemical and biological arguments for the link between CpG methylation and depletion are fundamental and should apply to all kingdoms of life.

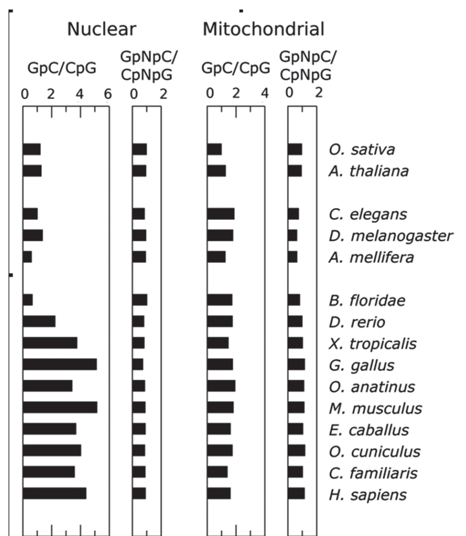

Fig. 1. CpG/GpC ratios of nuclear and mitochondrial DNA in different organisms

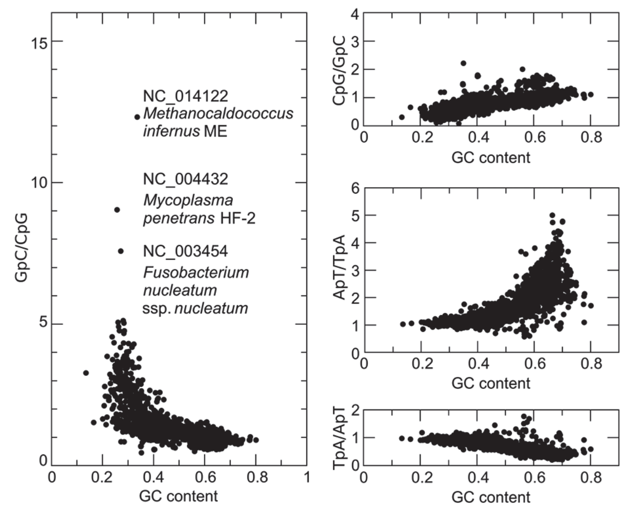

Fig. 2. Screen for CpG depletion in bacterial genomes with controls

Hence, one can ask the question, "Is it possible to discover novel prokaryotic CpG methyltransferases (such as the previously found CpG specific M.SssI (Renbaum et al., 1990)) by searching bacterial genomes for CpG depletion?" The suggestion has previously been made, but was not followed up by experiments (Xia, 2003). We have carried out a more extended search for CpG depletion and scanned all fully sequenced bacterial genomes in the NCBI sequence collection for CpG underrepresentation (Fig. 2). We found several drastically (i.e., approximately 10-fold) CpG-depleted bacterial species. Although there were differences in the statistical signatures of CpG depletion when compared with

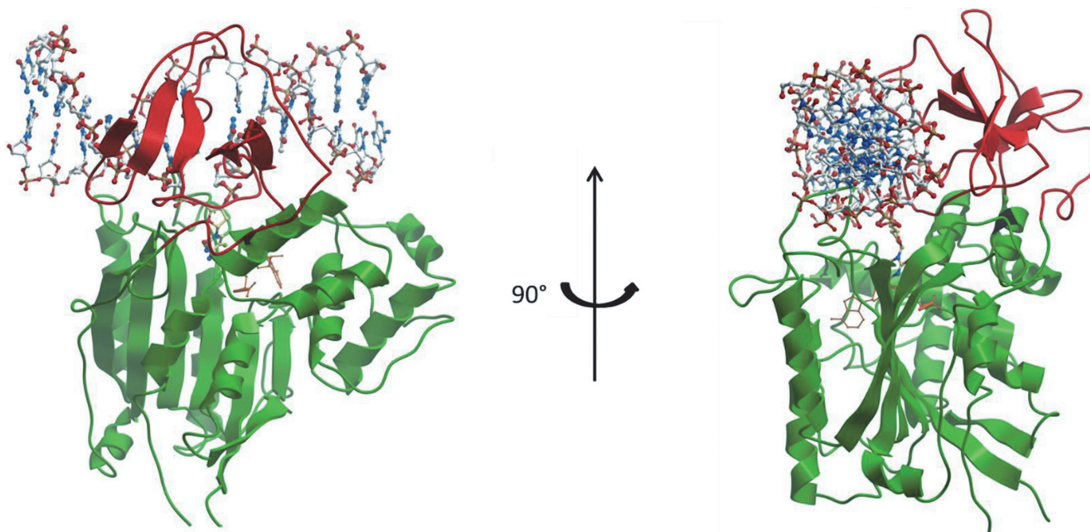

Fig. 3. Structure of the M.MpeI DNA complex. Target recognition domain in red, S-adenosylmethionine binding domain in green

eukaryotes (e.g., in the comparison of coding and non-coding regions), bacteria with drastic CpG depletion were next analyzed for CpG methylation. In the case of *Mycoplasma penetrans*, a genome-wide study of cytosine methylation using bisulfite sequencing identified global CpG methylation and several other universally methylated sequences. To identify the *M. penetrans* CpG methyltransferase, we picked a candidate protein on the basis of remote amino acid sequence similarity to M.SssI. Using bisulfite sequencing and other CpG methylation assays (i.e., HpaII/MspI digestion), we demonstrated *in vitro* that our candidate enzyme was indeed a CpG-specific DNA methyltransferase; and hence, named it M.MpeI, in accordance with nomenclature guidelines (Roberts et al., 2003).

How does M.MpeI recognize its CpG target sequence with extraordinary specificity? To answer this question, we crystallized M.MpeI with target DNA and solved the structure at 2.15 Å resolution (Fig. 3). Unsurprisingly, we found typical features of CpG methyltransferase DNA complexes, such as a flipping of the substrate cytosine and the proximity of the co-factor to the substrate base for direct transfer of a methyl group. Very interestingly, the DNA structure was perturbed not only in the substrate strand but also in the complementary strand. In this latter strand, we detected intercalation of a phenylalanine residue between the C and G nucleotides of the CpG site (Fig. 4). The 5′-pyrimidine purine-3′ steps

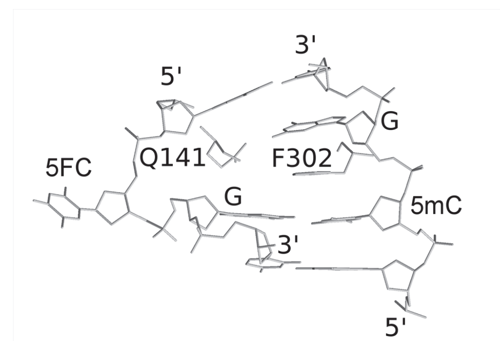

Fig. 4. Detail of the M.MpeI-DNA complex. 5-fluorocytosine (5FC) was used instead of cytosine in the position of the substrate base

are thought to be easier to unstack than other dinucleotide steps (Olson et al., 1998). Hence, intercalation might contribute to CpG readout. This concept is supported by the recent structure of the eukaryotic DNA maintenance methyltransferase Dnmt1 in complex with target DNA, which also shows unstacking of the CpG step.

If CpG methylation damages genomes, then what is the benefit for bacteria to retain a CpG-specific DNA methyltransferase? We are presently unable to answer this question, but several possibilities exist. The methyltransferase might be part of a CpG-specific restriction modification system. Alternatively and somewhat improbably in light of our genome-wide methylation data, it might play a role as an epigenetic regulator. Finally, bacterial CpG methylation might involve host pathogen interactions. Although the claim is still debated, most authors now agree

that CpG-unmethylated DNA is far more immunogenic than CpG-methylated DNA. Hence, CpG methylation might help bacteria dodge the host immune system. If so, then our findings could also have medical applications because at least some of the CpG-specific DNA methyltransferases are found in human pathogens.

Our work on the link between CpG depletion and methylation and the CpG specific methyltransferase M.MpeI has recently been published (Wojciechowski et al., 2013).

References

Jeltsch A. (2010) Science 328: 837-838.

Cooper D.N., Krawczak M. (1989) Hum. Genet. 83: 181-188.

Shen J.C., Rideout W.M. 3rd, Jones P.A. (1992) Cell 71: 1073-1080.

Shen J.C., Rideout W.M. 3rd, Jones P.A. (1994) Nucl. Acids Res. 22: 972-976.

Renbaum P., Abrahamove D., Fainsod A. et al. (1990) Nucl. Acids Res. 18: 1145-1152.

Xia X. (2003) J. Mol. Evol. 57, Suppl. 1: S21-S28.

Roberts R.J., Belfort M., Bestor T. et al. (2003) Nucl. Acids Res. 31: 1805-1812.

Olson W.K., Gorin A.A., Lu X.J. et al. (1998) Proc. Natl. Acad. Sci. USA 95: 11163-11168.

Wojciechowski M., Czapinska H., Bochtler M. (2013) Proc. Natl. Acad. Sci. USA 110: 105-110.

Plant proteins with anti-cancer properties: structural studies of two members of β-trefoil family that combine serine protease inhibition with activities as lectins

Dongwen Zhou [1], Rodrigo S. Ferreira [2], Yara A. Lobo [2],
Alla Gustchina [1], Maria L.V. Oliva [2], Alexander Wlodawer [1]

[1] Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, USA
[2] Departamento de Bioquímica, Universidade Federal de São Paulo, São Paulo, Brazil

Two related plant proteins that belong to the β-trefoil family, which includes Kunitz-type serine protease inhibitors, have been investigated by X-ray crystallography as well as by biochemiques. A potent inhibitor of trypsin-related enzymes isolated from the seeds of *Enterolobium contortisiliquum* (EcTI) inhibits the invasion of gastric cancer cells through alterations in integrin-dependent cell-signaling pathway, whereas a much weaker inhibitor isolated from the bark of *Crataeva tapia* (CrataBL) also functions as a lectin. We determined high-resolution crystal structures of free EcTI and in complex with bovine trypsin, in the process re-determining the amino acid sequence. The structure of the complex confirmed the standard inhibitory mechanism in which the reactive loop of the inhibitor is docked into trypsin active site with the side chains of Arg64 and Ile65 occupying the S1 and S1' pockets, respectively. The overall conformation of the reactive loop undergoes only minor adjustments upon binding to trypsin. A comparison of the EcTI-trypsin complex with the complexes of related Kunitz inhibitors has shown that rigid body rotation of the inhibitors by as much as 15° is required for accurate juxtaposition of the reactive loop with the active site while preserving its conformation. Modeling of the putative complexes of EcTI with several serine proteases and a comparison with equivalent models for other Kunitz inhibitors elucidated the structural basis for the fine differences in their specificity, providing tools that might allow modification of their potency towards the individual enzymes. We have also determined the high-resolution crystal structures of two different crystal forms of glycosylated CrataBL and identified dimers of the protein forming the crystals. CrataBL shows relatively weak inhibitory activity against trypsin and is more potent against Factor Xa, but not active against a number of other serine proteases. We have shown that, as a lectin, CrataBL binds only sulfated oligosaccharides, most likely heparin and its derivatives. We have observed that addition of CrataBL to DU145 and PC3 cell lines leads to their apoptosis, with release of mitochondrial cytochrome *c* through activation of caspase-3. Further studies aimed at more complete elucidation of the structural basis of the anti-cancer activity of these proteins are in progress.